Appendix
A Large-Scale Approach for Evaluating Asset Pricing Models

Laurent Barras*

This version: March 24, 2018

*Desautels Faculty of Management, McGill University, Montreal. E-mail: laurent.barras@mcgill.ca.

I The Cross-Section of Micro Portfolios

A Portfolio Formation Procedure

We describe the procedure for forming the set of micro portfolios in each size group (tiny-, small-, and big-cap). First, we sort the N_{t_a} stocks in each formation year t_a ($t_a = 1, ..., T_a$) according to their estimated average returns. To compute this variable denoted by $\hat{\mu}^s_{i,t_a}$ ($i = 1, ..., N_{t_a}$), we use a linear combination of firm characteristics.[1] For each month t prior to the formation date, we run a cross-sectional regression of the monthly stock excess returns on the most recently observed characteristics: $r_{i,t} = \gamma_t' c_{i,t} + e_{i,t}$, where $c_{i,t}$ denotes the vector of characteristics including a constant. Then, we estimate the characteristic-based average return as

$$\hat{\mu}_{i,t_a} = \hat{\gamma}_{t_a}' c_{i,t_a}, \tag{A1}$$

where c_{i,t_a} is the vector of characteristics observed in the formation year t_a, and $\hat{\gamma}_{t_a}$ is the time-series average of the monthly vector of coefficients. To facilitate the chaining of portfolio returns over consecutive years, we work with the standardized average return computed as

$$\hat{\mu}^s_{i,t_a} = \frac{\hat{\mu}_{i,t_a} - \frac{1}{N_{t_a}} \sum_{i=1}^{N_{t_a}} \hat{\mu}_{i,t_a}}{\left(\frac{1}{N_{t_a}} \sum_{i=1}^{N_{t_a}} \hat{\mu}^2_{i,t_a} - \left(\frac{1}{N_{t_a}} \sum_{i=1}^{N_{t_a}} \hat{\mu}_{i,t_a} \right)^2 \right)^{\frac{1}{2}}}. \tag{A2}$$

Second, we construct, for each stock i, a micro portfolio by equally weighting the stock itself and $n-1$ additional stocks with the nearest values to $\hat{\mu}^s_{i,t_a}$. This technique is called local averaging and borrows from Efron (2010, ch. 9). Third, we chain the portfolio returns over time to obtain stable average returns. For each pair (i, j) of micro portfolios in years t_a and t_a+1, we compute the distance between them as $\left| \hat{\mu}^s_{i,t_a} - \hat{\mu}^s_{j,t_a+1} \right|$.[2] Then, we match the portfolios with the lowest distance (each year-t_a portfolio can only be paired with one year-t_a+1 portfolio). To minimize changes in portfolio composition, we match the pair (i, i) first if $\left| \hat{\mu}^s_{i,t_a} - \hat{\mu}^s_{i,t_a+1} \right|$ is in the bottom 1% of all measured distances.

[1] The characteristic-average return relation used here should be distinguished from previous studies that impose a linear relation between characteristics and pricing errors (Avramov and Chordia (2006), Brennan, Chordia, and Subrahmanyam (1998)). In our case, two stocks can have similar characteristics and yet different pricing errors because they are not exposed to the same risk factors.

[2] Alternatively, we can compute the estimated average return of the newly-created portfolio as $\frac{1}{n}\left(\hat{\mu}^s_{i,t_a} + \sum_m \hat{\mu}^s_{i_m,t_a}\right)$, where i_m ($m = 1, ..., n-1$) denotes the identity of the additional stocks included in the portfolio. Using this approach leaves the results unchanged.

In Figure A1, we illustrate the portfolio formation procedure in a population of 50 stocks ($N_{t_a} = N = 50$) over a 2-year sample period ($T_a = 2$). Each dot denotes the ordered value $\hat{\mu}^s_{i,t_a}$ at the start of each year. We see that the portfolio composition changes each year to account for the time-variation in characteristics. For instance, the portfolio associated with the median average return $\hat{\mu}^s_{25,t_a}$ includes stocks S_{10}, $S_{42}, ...,$ S_3 in year 1, and stocks S_{18}, $S_6, ...,$ S_{46} in year 2. The formation procedure yields a total number of micro portfolios, M, equal to the number of stocks ($M = N$).

In practice, the formation procedure is more complicated because the number of stocks changes over time. Suppose that the number of stocks in year 2 is equal to 60 (instead of 50). Applying the matching procedure described above, we can pair 50 year-1 and 50 year-2 portfolios, which leaves 10 year-2 portfolios unmatched. In this example, the cross-section includes 60 micro portfolios ($M = 60$) with unequal time-series lengths: (i) 50 portfolios created in year 1 with complete return history (24 monthly returns), (ii) 10 unmatched portfolios created in year 2 with 12 monthly return observations. Conversely, suppose that we have 60 portfolios in year 1 (instead of 50). In this case, we can only pair 50 year-1 portfolios, which leaves 10 year-1 portfolios unmatched ($M = 60$). In general, the total number of micro portfolio is therefore equal to $M = \max_{t_a}(N_{t_a})$.

Please insert Figure A1 here

B Definition of the Characteristics

To compute the firm's average return, we follow Fama and French (2008) and use a linear combination of three characteristics—book-to-market, profitability, and investment. We use the definitions of Fama and French (2008, 2015) to measure these characteristics at the end of June of each year t_a and winsorize the data at the 1% and 99% levels to remove outliers. The book-to-market ratio is equal to the ratio of the book value of equity to the market value of equity. The book value for year t_a is defined as total assets minus liabilities, plus balance sheet deferred taxes and investment tax credit (if available), minus preferred shares stock liquidating values (if available), or carrying value (if available) in the fiscal year ending in the calendar year $t_a - 1$. The market value for year t_a equals the price times shares outstanding at the end of December of year $t_a - 1$. Investment for year t_a is computed as the relative change in total assets between the fiscal years ending in calendar years $t_a - 2$ and $t_a - 1$. Finally, profitability for year t_a is defined as revenues minus cost of goods sold, minus selling, general, and administrative expenses, minus interest expense all divided by the book value of equity. Each of these variables is computed using data in the fiscal year ending in the calendar year $t_a - 1$.

We also use the definitions in Hou, Xue, and Zhang (2015) to construct alternative proxies for the three characteristics. First, we replace the book-to-market ratio with the earnings-to-price and cash flow-to-price ratios. Earnings for year t_a are defined as income before extraordinary items. Cash flows for year t_a are computed for year as income before extraordinary items, plus equity's share of depreciation, plus deferred taxes (if available). Each of these variables is computed using data in the fiscal year ending in the calendar year $t_a - 1$. Price for year t_a equals the market value measured at the end of December of year $t_a - 1$.

Second, we measure investment using infrastructure growth and inventory growth. Infrastructure growth for year t_a is defined as the change in gross property, plant, and equipment, plus changes in inventory in the fiscal year ending in the calendar year $t_a - 1$ scaled by total assets in the fiscal year ending in the calendar year $t_a - 2$. Inventory growth is defined as the relative change in capital expenditure between the fiscal years ending in calendar years $t_a - 2$ and $t_a - 1$.

Third, we measure profitability using the Return On Equity (ROE) and the Return On Assets (ROA). ROE for year t_a is defined as income before extraordinary items in the fiscal year ending in the calendar year $t_a - 1$ divided by book equity in the fiscal year ending in the calendar year $t_a - 2$. ROA for year t_a is defined as income before extraordinary items in the fiscal year ending in the calendar year $t_a - 1$ divided by total assets in the fiscal year ending in the calendar year $t_a - 2$.

II Estimation Procedure

A Extended Two-Pass Regression

We provide a general description of the econometric framework for estimating the pricing errors of the micro portfolios. Under model k ($k = 1, ..., K$), the excess return of each portfolio j ($j = 1, ..., M$) can be written as

$$r_{j,t} = a_j^k + b_{jm} r_{m,t} + b_j^{k\prime} f_t^k + c_j^{\prime} z_t + e_{j,t}, \tag{A3}$$

where $r_{m,t}$ is the market excess return, f_t^k is the J-vector of risk factors specific to model k, z_t the P-vector of factors included in the other models and orthogonal to $r_{m,t}$ and f_t^k, and $e_{j,t}$ denotes the residual term. The intercept is equal to

$$a_j^k = \alpha_j^k - b_j^{k\prime} p_c^k, \tag{A4}$$

where α_j^k is the portfolio pricing error and p_c^k is the J-vector of forward prices of the risk factors.[3] Using Equation (A4), we can write the pricing error as

$$\alpha_j^k = \omega^{k\prime} \beta_j^k, \tag{A5}$$

where the $(J+1)$-vectors ω^k and β_j^k are defined as $\omega^k = [1, p_c^{k\prime}]'$ and $\beta_j^k = [a_j^k, b_j^{k\prime}]'$. To estimate α_j^k, we build on recent work by Gagliardini, Ossola, and Scaillet (2016; GOS hereafter) who extend the traditional two-pass regression to a large and unbalanced panel of test assets—two important features exhibited by micro portfolios.

In the first step, we run a time-series regression of $r_{j,t}$ on the $(J+P+2)$-vector $x_t = [1, r_{m,t}, f_t^{k\prime}, z_t']'$ for each portfolio j. The OLS estimator of the $(J+P+2)$-vector of coefficients $b_j = [a_j, b_{jm}, b_j^{k\prime}, c_j']'$ is given by

$$\hat{b}_j = \left(\sum_{t=1}^{T} \mathit{1}_{j,t} x_t x_t' \right)^{-1} \sum_{t=1}^{T} \mathit{1}_{j,t} x_t r_{j,t}, \tag{A6}$$

where T is the total number of observations, and $\mathit{1}_{j,t}$ equals one if $r_{j,t}$ is non-missing. The matrix inversion in Equation (A6) is numerically unstable if only few return observations are available. To address this issue, GOS introduce the following trimming device:

$$\mathit{1}_j^{\chi} = \mathit{1} \left\{ \tau_{j,T} \leq \chi_{1,T},\ CN(\hat{Q}_{x,j}) \leq \chi_{2,T} \right\}, \tag{A7}$$

where $\tau_{j,T} = \frac{T}{T_j}$, $T_j = \sum_{t=1}^{T} \mathit{1}_{j,t}$, $CN(\hat{Q}_{x,j}) = \left(eig_{\max}(\hat{Q}_{x,j}) / eig_{\min}\left(\hat{Q}_{x,j}\right) \right)^{\frac{1}{2}}$ denotes the condition number of the matrix $\hat{Q}_{x,j} = \frac{1}{T} \sum_{t=1}^{T} \mathit{1}_{j,t} x_t x_t'$. Following GOS, we set $\chi_{1,T} = \frac{606}{60}$ (a minimum of 60 monthly observations) and $\chi_{2,T} = 15$.

In the second step, we estimate the J-vector of forward prices p_c^k using a cross-sectional regression of the estimated intercept $\hat{a}_j^k$ on the J-vector of estimated betas $\hat{b}_j^k$ keeping the non-trimmed portfolios only:

$$\hat{p}_{c(1)}^k = -\left(\sum_{j=1}^{M} \mathit{1}_j^{\chi} \hat{b}_j^{k\prime} \hat{b}_j^k \right)^{-1} \sum_{j=1}^{M} \mathit{1}_j^{\chi} \hat{b}_j^{k\prime} \hat{a}_j^k. \tag{A8}$$

We adjust $\hat{p}_{c(1)}^k$ for the bias component $\Psi_{p_c^k} = Q_{b_k}^{-1} \left(\frac{1}{M} \sum_{j=1}^{M} \tau_{j,T} E_1' V_j \omega^k \right)$, where $Q_{b_k} = E[b_j^{k\prime} b_j^k]$, $E_1 = [\mathbf{0}_{J \times 1}, I_J]'$, I_J is the $J \times J$ identity matrix, $V_j = E_2' Q_x^{-1} S_{jj} Q_x^{-1} E_2$,

[3] The forward price of the market factor does not appear in equation (A4) because $r_{m,t}$ is an excess return which, by definition, has a forward price equal to zero.

$Q_x = [x_t x_t']$, $S_{jj} = E[e_{j,t}^2 x_t x_t']$, and E_2 is a $(J+P+2) \times (J+1)$ matrix whose sth row e_s $(s = 1, 3, 4, ..., J+1)$ has one for the sth element and zeros everywhere else. The final estimate of p_c^k is equal to

$$\hat{p}_c^k = \hat{p}_{c(1)}^k + \frac{1}{T}\hat{\Psi}_{p_c^k}, \tag{A9}$$

where $\hat{\Psi}_{p_c^k}$ is computed as $\hat{Q}_{b_k}^{-1}\left(\frac{1}{M_\chi}\sum_{j=1}^{M} 1_j^\chi \tau_{j,T} E_1' \hat{V}_j \hat{\omega}_{(1)}^k\right)$, $M_\chi = \sum_{j=1}^{M} 1_j^\chi$, $\hat{Q}_{b_k} = \frac{1}{M_\chi}\sum_{j=1}^{M} 1_j^\chi \hat{b}_j^{k\prime}\hat{b}_j^k$, $\hat{V}_j = E_2'\hat{Q}_{x,j}^{-1}\hat{S}_{jj}\hat{Q}_{x,j}^{-1}E_2$, and $\hat{\omega}_{(1)}^k = [1, \hat{p}_{c(1)}^{k\prime}]'$. Following GOS, we estimate S_{jj} using the White estimator (1980): $\hat{S}_{jj} = \frac{1}{T}\sum_{t=1}^{T} 1_{j,t}\hat{e}_{j,t}^2 x_t x_t'$, where $\hat{e}_{j,t} = r_{j,t} - \hat{b}_j' x_t$. Plugging the estimated quantities in Equation (A5), we obtain

$$\hat{\alpha}_j^k = \hat{\omega}^{k\prime}\hat{\beta}_j^k. \tag{A10}$$

B Estimation of the Portfolio t-Statistics

We now prove Proposition 1 which provides an analytical expression for the t-statistic of the portfolio pricing error and its asymptotic distribution.

Proof of Proposition 1. We consider the misspecified model k and suppose that the residual terms $e_{j,t}$ $(j = 1, ..., M)$ are weakly correlated. When the number of portfolios and return observations grow large $(T, M \to \infty)$, Proposition 7 of GOS shows that $\hat{p}_c^k$ converges towards p_c^k at a rate equal to $\sqrt{M}$. In addition, standard results in regression analysis reveal that the vector of estimated coefficients $\hat{\beta}_j^k$ is asymptotically distributed as

$$\sqrt{T}(\hat{\beta}_j^k - \beta_j^k) \xrightarrow{d} N\left(0, V_j\right), \tag{A11}$$

where $V_j = E_2' Q_x^{-1} S_{jj} Q_x^{-1} E_2$. With M in the thousands and T in the hundreds, the asymptotic sampling variation in $\hat{\alpha}_j^k$ is therefore only driven by that of $\hat{\beta}_j^k$, i.e.,

$$\sqrt{T}(\hat{\alpha}_j^k - \alpha_j^k) \xrightarrow{d} N\left(0, \omega^{k\prime} V_j \omega^k\right). \tag{A12}$$

Using this result, we compute the portfolio t-statistic t_j as $\frac{\hat{\alpha}_j^k}{\hat{\sigma}_{\alpha_j^k}}$, where

$$\hat{\sigma}_{\alpha_j^k}^2 = \frac{1}{T_j}\hat{\omega}^{k\prime}\hat{V}_j\hat{\omega}^k = \hat{\omega}^{k\prime}\hat{V}_{\beta_j^k}\hat{\omega}^k. \tag{A13}$$

The variance term $\hat{V}_{\beta_j^k}$ is equal to $\frac{1}{T_j}\hat{V}_j$, where $\hat{V}_j$ is a consistent estimator of the covariance matrix V_j. In addition, Equation (A12) implies that the t-statistic follows a normal distribution, $t_j \xrightarrow{d} N\left(\frac{\alpha_j^k}{\sigma_{\alpha_j^k}}, 1\right)$, where $\alpha_j^k = \omega^{k\prime}\beta_j^k$ and $\sigma_{\alpha_j^k}^2 = \omega^{k\prime} V_{\beta_j^k}\omega^k$.

III Statistical Inference

A Proportion of Mispriced Portfolios

We compute the proportion of portfolios that are mispriced by model k as

$$\hat{\pi}_k = 1 - \frac{\hat{P}_k(I)}{P_0(I)} = 1 - \frac{\frac{1}{M}\sum_{j=1}^{M} \mathit{1}(t_j^k)}{\Phi_0(I)}, \tag{A14}$$

where $\hat{P}_k(I)$ is the estimated probability that the t-statistic falls in the interval $I = [-a, a]$, $\mathit{1}(t_j^k)$ is an indicator function equal to one if t_j^k falls in I and $P_0(I)$ is computed from the standard normal cdf $\Phi_0(I) = \Phi_0(a) - \Phi_0(-a) = 2\Phi_0(a) - 1$.

Proof of Proposition 2. We consider the misspecified model k and suppose that the residual terms $e_{j,t}$ $(j = 1, ..., M)$ are weakly correlated. We further assume that the t-statistics are spatially ordered such that nearby t-statistics exhibit higher correlation. When the number of portfolios grows large $(M \to \infty)$, Lemma 2 of Farcomeni (2006) shows that

$$\sqrt{M}(\hat{P}_k(I) - P_k(I)) \xrightarrow{d} N\left(0, \sigma_k^2\right), \tag{A15}$$

where $\sigma_k^2 = var(\mathit{1}(t_1^k)) + 2\sum_{j=2}^{\infty} cov(\mathit{1}(t_1^k), \mathit{1}(t_j^k))$ and t_j $(j = 1, ..., \infty)$ are the ordered t-statistics. Because the variance of the estimated proportion $\hat{\pi}_k$ only depends on that of $\hat{P}_k(I)$ (Equation (A14)), the asymptotic distribution of the vector of estimated proportions for two misspecified models k and l is given by

$$\sqrt{M} \begin{bmatrix} \hat{\pi}_k - \pi_k^* \\ \hat{\pi}_l - \pi_l^* \end{bmatrix} \xrightarrow{d} N \left(\begin{matrix} 0 \\ 0 \end{matrix}, \begin{matrix} \frac{\sigma_k^2}{\Phi_0(I)^2} & \frac{\sigma_{k,l}}{\Phi_0(I)^2} \\ \frac{\sigma_{k,l}}{\Phi_0(I)^2} & \frac{\sigma_l^2}{\Phi_0(I)^2} \end{matrix} \right). \tag{A16}$$

where $\pi_k^* = E(\hat{\pi}_k)$ and $\pi_l^* = E(\hat{\pi}_l)$. The variance terms are given by

$$\begin{aligned} \sigma_k^2 &= var(\mathit{1}(t_1^k)) + 2\sum_{j=2}^{\infty} cov(\mathit{1}(t_1^k), \mathit{1}(t_j^k)), \\ \sigma_l^2 &= var(\mathit{1}(t_1^l)) + 2\sum_{j=2}^{\infty} cov(\mathit{1}(t_1^l), \mathit{1}(t_j^l)), \\ \sigma_{k,l} &= cov(\mathit{1}(t_1^k), \mathit{1}(t_1^l)) + \sum_{j=2}^{\infty} cov(\mathit{1}(t_1^k), \mathit{1}(t_j^l)) + cov(\mathit{1}(t_1^l), \mathit{1}(t_j^k)), \end{aligned} \tag{A17}$$

where t_j^k, t_j^l $(j = 1, ..., \infty)$ are the ordered t-statistics for models k and l.

Using Proposition 2, we can test the null hypothesis that model k is correctly specified. Under the null hypothesis $H_0 : \pi_k^* = 0$, Genovese and Wasserman (2004) show that the estimated mispricing proportion $\hat{\pi}_k$ is asymptotically distributed as

$$\sqrt{M}\hat{\pi}_k \xrightarrow{d} \frac{1}{2}\delta_0 + \frac{1}{2}N^+\left(0, \frac{\sigma_k^2}{\Phi_0(I)^2}\right), \tag{A18}$$

where δ_0 is a point-mass at zero and N^+ is a positive-truncated normal distribution. To test this hypothesis at the size level ϕ, we determine whether $\hat{\pi}_k$ is sufficiently far away from zero using the following threshold:

$$\hat{\pi}_k > \chi_\phi \frac{1}{\sqrt{M}} \frac{\hat{\sigma}_k}{\Phi_0(I)}, \tag{A19}$$

where $\hat{\sigma}_k$ is the consistent estimator of σ_k and χ_ϕ is the quantile of the standard normal distribution at (1-ϕ). To compute $\hat{\sigma}_k$, we use the following estimator proposed by Newey-West (1987):

$$\hat{\sigma}_k^2 = \left[\frac{1}{M}\sum_{j=1}^{M} \mathit{1}(t_j^k)\right] - \hat{P}_k(I)^2 + 2\sum_{l_1=1}^{L}\left[\frac{1}{M-l}\sum_{j=1}^{M-l_1} \mathit{1}(t_j^k)\mathit{1}(t_{l_1+j}^k)\right] - \hat{P}_k(I)^2, \tag{A20}$$

where L is the number of cross-sectional lags.[4]

Proposition 2 also allows us to test the null hypothesis of equal performance between two misspecified, possibly non-nested models. Under the null hypothesis $H_0 : \Delta\pi^* = \pi_k^* - \pi_l^* = 0$, the estimated difference $\Delta\hat{\pi} = \hat{\pi}_k - \hat{\pi}_l$ is asymptotically distributed as

$$\sqrt{M}(\Delta\hat{\pi} - \Delta\pi^*) \xrightarrow{d} N\left(0, \frac{\sigma_k^2 + \sigma_l^2 - 2\sigma_{k,l}}{\Phi_0(I)^2}\right). \tag{A21}$$

To implement this testing procedure, we compute the covariance term $\sigma_{k,l}$ using the following consistent estimator:

$$\begin{aligned}\hat{\sigma}_{k,l} &= \left[\frac{1}{M}\sum_{j=1}^{M} \mathit{1}(t_1^k)\mathit{1}(t_1^l)\right] - \hat{P}_k(I)\hat{P}_l(I) \\ &\quad + \sum_{l_1=1}^{L}\left[\frac{1}{M-l}\sum_{j=1}^{M-l_1} \mathit{1}(t_1^k)\mathit{1}(t_{l_1+j}^l) + \mathit{1}(t_1^l)\mathit{1}(t_{l_1+j}^k)\right] - \hat{P}_k(I)\hat{P}_l(I). \end{aligned} \tag{A22}$$

[4] In the baseline specification, we set L equal to 1% of the total number of portfolios ($L = 40$ in the entire population) to account for potential weak dependencies between portfolios. Setting $L = n$ yields similar volatility estimates.

B Sign of the Pricing Errors

We can extend the large-scale approach to conduct inference on the estimated proportions of portfolios with negative and positive pricing errors. To compute both proportions denoted by $\hat{\pi}_k^-$ and $\hat{\pi}_k^+$, we use the procedure of Barras, Scaillet, and Wermers (2010). First, we determine the proportions of portfolios with low or high estimated pricing errors $\hat{P}_k(I^-)$ and $\hat{P}_k(I^+)$, where $I^- = [-\infty, -a]$, $I^+ = [a, +\infty]$, and $-a$ and a denote the lower and upper bounds of the interval I. Second, we deduct the proportions of "false discoveries", $(1-\hat{\pi}_k)\Phi_0(I^-)$ and $(1-\hat{\pi}_k)\Phi_0(I^+)$—both expressions measure the proportions of correctly-priced portfolios which, by chance, have t-statistics falling in the intervals I^- and I^+. This two-step approach yields the following expressions for $\hat{\pi}_\kappa^-$ and $\hat{\pi}_\kappa^+$ and their variances:

$$\begin{aligned} \hat{\pi}_k^- &= \hat{P}_k(I^-) - (1-\hat{\pi}_k)\Phi_0(I^-), \\ \hat{\pi}_k^+ &= \hat{P}_k(I^+) - (1-\hat{\pi}_k)\Phi_0(I^+), \end{aligned} \tag{A23}$$

$$\begin{aligned} \sigma^2_{\pi_k^-} &= \sigma^2_{P_k(I^-)} + \Phi_0(I^-)^2\sigma^2_{\pi_\kappa} - 2\Phi_0(I^-)\sigma_{P_k(I^-),\pi_k}, \\ \sigma^2_{\pi_k^+} &= \sigma^2_{P_k(I^+)} + \Phi_0(I^+)^2\sigma^2_{\pi_k} - 2\Phi_0(I^+)\sigma_{P_k(I^+),\pi_k}, \end{aligned} \tag{A24}$$

where $\sigma^2_{P_k(I^-)} = \frac{1}{M}\sigma^2_{k^-}$, $\sigma^2_{P_k(I^+)} = \frac{1}{M}\sigma^2_{k^+}$, $\sigma^2_{\pi_k} = \frac{1}{M}\frac{\sigma_k^2}{\Phi_0(I)^2}$, $\sigma_{P_k(I^-),\pi_k} = \frac{1}{M\Phi_0(I)}\sigma_{k^-,k}$, and $\sigma_{P_k(I^+),\pi_k} = \frac{1}{M\Phi_0(I)}\sigma^2_{k^+,k}$. The different components are given by

$$\begin{aligned} \sigma^2_{k^-} &= var(1^-(t_1^k)) + 2\sum_{j=2}^{\infty} cov(1^-(t_1^k), 1^-(t_j^k)), \\ \sigma^2_{k^+} &= var(1^+(t_1^k)) + 2\sum_{j=2}^{\infty} cov(1^+(t_1^k), 1^+(t_j^k)), \\ \sigma_{k^-,k} &= cov(1^-(t_1^k), 1(t_1^k)) + \sum_{j=2}^{\infty} cov(1^-(t_1^k), 1(t_j^k)) + cov(1(t_1^k), 1^-(t_j^k)), \\ \sigma_{k^+,k} &= cov(1^+(t_1^k), 1(t_1^k)) + \sum_{j=2}^{\infty} cov(1^+(t_1^k), 1(t_j^k)) + cov(1(t_1^k), 1^+(t_j^k)), \end{aligned} \tag{A25}$$

where $1^-(t_j^k)$ $(1^+(t_j^k))$ is an indicator function equal to one if t_j^k falls in I^- (I^+). After replacing the above expressions with the consistent estimators proposed by Newey-West (1987) we can conduct inference on the two proportions $\hat{\pi}_k^-$ and $\hat{\pi}_k^+$.

C Testing for Useless Factors

We now explain how to test whether the sth factor $f_{s,t}^k$ included in model k is useless. For each portfolio j ($j = 1, ..., M$), the beta on factor $f_{s,t}^k$ obtained from the first-pass regression in Equation (A6) is asymptotically distributed as

$$\sqrt{T}(\hat{b}_{js}^k - b_{js}) \xrightarrow{d} N\left(0, e_{s+1}' V_j e_{s+1}\right), \tag{A26}$$

where e_{s+1} is a $(J+1)$-vector whose $(s+1)$th element is one and the others are zero. Using this result, we compute the associated t-statistic $t_{b_{js}}^k$ as $\frac{\hat{b}_{js}^k}{\hat{\sigma}_{b_{js}^k}}$, where the estimated variance of $\hat{b}_{js}^k$ is given by

$$\hat{\sigma}_{b_{js}^k} = \frac{1}{T_j} e_{s+1}' \hat{V}_j e_{s+1}. \tag{A27}$$

Then, we compute the proportion of portfolios with non-zero betas on factor $f_{s,t}^k$ using the same expression as in Equation (A14):

$$\hat{\pi}_k(b_s) = 1 - \frac{\hat{P}_{b_s^k}(I)}{P_{b_s^k,0}(I)} = 1 - \frac{\frac{1}{M}\sum_{j=1}^{M} 1(t_{b_{js}}^k)}{\Phi_0(I)}, \tag{A28}$$

where $\hat{P}_{b_s^k}(I)$ is the estimated probability that the beta t-statistic falls in the interval I and $1(t_{b_{js}}^k)$ is an indicator function equal to one if $t_{b_{js}}^k$ falls in I.

If factor $f_{s,t}^k$ is useless, the *true* betas are all equal to zero and we obtain $H_0 : E(\hat{\pi}_k(b_s)) = \pi_k^*(b_s) = 0$. Similar to $\hat{\pi}_k$, we can write the distribution of the estimated proportion $\hat{\pi}_k(b_s)$ under H_0 as

$$\sqrt{M}\hat{\pi}_k(b_s) \xrightarrow{d} \frac{1}{2}\delta_0 + \frac{1}{2}N^+\left(0, \frac{\sigma_k^2(b_s)}{\Phi_0(I)^2}\right), \tag{A29}$$

where δ_0 is a point-mass at zero, N^+ is a positive-truncated normal distribution, and $\sigma_k^2(b_s)$ follows the same expression as in Equation (A17) except that we use the ordered beta t-statistics $t_{b_{js}}^k$. To test this hypothesis at the size level ϕ, we determine whether $\hat{\pi}(b_s^k)$ is sufficiently far away from zero using the following threshold:

$$\hat{\pi}(b_s^k) > \chi_\phi \frac{1}{\sqrt{M}} \frac{\hat{\sigma}_k(b_s)}{\Phi_0(I)}, \tag{A30}$$

where $\hat{\sigma}_k(b_s^k)$ is the consistent estimator of $\sigma_k(b_s^k)$, and χ_ϕ is the quantile of the standard normal distribution at (1-ϕ).

IV Monte Carlo Analysis

A Setting

We conduct a Monte Carlo analysis to evaluate the finite-sample properties of the proportion estimators for two misspecified models a and b. We extend the illustrative example presented in the paper on several important dimensions to closely replicate the salient features of the data. First, we match the total number of micro portfolios across the three size groups (before imposing any filters on the data). Specifically, we construct a set of 2,349 tiny-cap portfolios, 938 small-cap portfolios, and 1,302 big-cap portfolios based on the empirical characteristics of the individual stocks in each size group.

Second, we account for the unbalanced nature of the panel of portfolio returns. To guarantee the same unbalanced structure as in the data, we apply the empirical $T \times M$ matrix of indicators $1_{j,t}$ ($j = 1, ..., M$ and $t = 1, ..., T$) to each simulated panel of portfolio returns, where T denotes the total sample size equal to 606 monthly observations and M is equal to 4,589 micro portfolios.

Third, we jointly match the average proportion of mispriced portfolios across the proposed models examined in the empirical section by adding a size premium λ^s to the average excess return of each individual stock i ($i = 1, ..., N$ with $N = M$):

$$\mu_i = b_{is}\lambda_s + b_{im}\lambda_m + b_{ia}\lambda_a + b_{ib}\lambda_b, \tag{A31}$$

where λ_m is the premium of the market return $r_{m,t}$, and λ_a, λ_b denote the premia of the two additional risk factors $f_{a,t}$ and $f_{b,t}$.

Model a includes the market and factor a, which implies that the vector of explanatory variables is defined as $x_{a,t} = [1, r_{m,t}, f_{a,t}, z_{b,t}]'$. The term $z_{b,t}$ is the estimated component of the omitted factor $f_{b,t}$ that is orthogonal to $x^0_{a,t} = [1, r_{m,t}, f_{a,t}]'$, i.e., $z_{b,t} = f_{b,t} - \hat{B}'_{z_b} x^0_{a,t}$, where $\hat{B}_{z_b}$ is the vector of estimated coefficients from a time-series regression of $f_{b,t}$ on $x^0_{a,t}$ over the entire sample period. Model b includes the market and factor b which implies that $x_{b,t} = [1, r_{m,t}, f_{b,t}, z_{a,t}]'$, where $z_{a,t} = f_{a,t} - \hat{B}'_{z_a} x^0_{b,t}$, $x^0_{b,t} = [1, r_{m,t}, f_{b,t}]'$ and $\hat{B}_{z_a}$ is the vector of estimated coefficients from a time-series regression of the omitted factor $f_{a,t}$ on $x^0_{b,t}$ over the entire sample period. We assume that $r_{m,t}$, $f_{a,t}$, $f_{b,t}$, and the residual term $e_{i,t}$ are all independent and normally distributed as $N(\lambda_m, \sigma^2_m)$, $N(\lambda_a, \sigma^2_a)$, $N(\lambda_b, \sigma^2_b)$, and $N(0, \sigma^2_e)$, respectively. We further assume that b_{is}, b_{im}, b_{ia}, b_{ib}, are randomly drawn from the normal distribution $N(E(b_s), var(b_s))$, $N(E(b_m), var(b_m))$, $N(E(b_a), var(b_a))$, $N(E(b_b), var(b_b))$.

To calibrate the model, we use monthly data on individual stocks and the Fama-

French three factors (market, size, value) over the entire sample period. The calibration of the distribution parameters for the betas and the residual term is done separately for each size group. To attribute each individual stock to a specific size group, we form, each year, the three size groups by taking as breakpoints the 20th and 50th percentiles of the market capitalization for the NYSE stocks (similar to Fama and French (2008)). We then classify each stock based on the frequencies at which it falls in the three groups. For each size group, we set $E(b_s)$ and $var(b_s)$ equal to the median and variance of the estimated size betas, and $E(b_m)$ and $var(b_m)$ equal to the median and variance of the estimated market betas. We further set $E(b_a)$, $E(b_b)$, and $var(b_a)$, $var(b_b)$ equal to the median and variance of the estimated value betas. Finally, σ_e is set equal to the cross-sectional average of the estimated residual volatility.

We set λ_s equal to 0.5% per month so as to approximate the median value for the proportions of mispriced portfolios (around 45%). We set λ_m and σ_m equal to the average return and volatility of the CRSP value-weighted index (0.5% and 4.4% per month). For the volatilities σ_a and σ_b of the additional risk factors a and b, we split the volatility of the value factor in two (2.8% per month). To determine the values for λ_a and λ_b, we choose two scenarios to capture the minimum and maximum proportion differences observed in the data. Under the first scenario, λ_a and λ_b are set equal to 1.0% and 0.0% per month so as to produce a large proportion difference between the two models. Under the second scenario, λ_a and λ_b are both equal to 0.5% per month, which implies that both models yield the same moderate performance.

B Simulation Procedure

For each scenario, we compute the estimated proportions of mispriced portfolios over 1,000 iterations and five sets of values for the stock betas ($S = 5,000$). For each iteration s ($s = 1, ..., S$), we first construct a T-vector of monthly return observations for each stock i ($i = 1, ..., N$ with $N = M$) :

$$r_{i,t}(s) = b_{is}\lambda_s + b_{im}r_{m,t}(s) + b_{ia}f_{a,t}(s) + b_{ib}f_{b,t}(s) + e_{i,t}(s), \tag{A32}$$

where $r_{m,t}(s)$, $f_{a,t}(s)$, $f_{b,t}(s)$, and $e_{i,t}(s)$ are drawn from their respective distributions. Second, we form the cross-section of micro portfolios using the average stock return μ_i as the sorting variable and apply the portfolio formation described above.[5] The resulting cross-section consists of M micro portfolios, each containing 10 stocks ($n = 10$)—stock

[5]We assume that the book equity of each firm is proportional to its future expected cash flows. In this case, the average return can be directly inferred from the observable book-to-market (bm) of each firm, i.e., μ_i is proportional to bm_i (see Berk (2000)).

i and nine additional stocks with the nearest average return to stock i. We keep track of the identity of the stocks included in each micro portfolio via a $M \times M$ matrix ID whose jth row id_j has zeros everywhere except for the stocks included in the portfolio. Third, we construct the monthly return of each micro portfolio j from the M-vector of stock returns $r_t(s) = [r_{1,t}, ..., r_{M,t}]'$:

$$r_{j,t}(s) = \mathit{1}_{j,t} \frac{1}{n} \left(id_j r_t(s) \right), \tag{A33}$$

where $\mathit{1}_{j,t}$ takes the value of one if the return is observed in the data (and zero otherwise). Fourth, we compute the vector of t-statistics for all portfolios using the extended two-pass regression described above and estimate the proportions of mispriced portfolios for the two models and its difference,

$$\begin{aligned} \hat{\pi}_a(s) &= 1 - \frac{\hat{P}_a(I)(s)}{\Phi_0(I)}, \\ \hat{\pi}_b(s) &= 1 - \frac{\hat{P}_b(I)(s)}{\Phi_0(I)}, \end{aligned}$$

$$\Delta\hat{\pi}(s) = \hat{\pi}_a(s) - \hat{\pi}_b(s), \tag{A34}$$

as well as the estimated variances of these estimators using Equations (A20) and (A22),

$$\begin{aligned} \hat{\sigma}^2_{\pi_a}(s) &= \frac{\hat{\sigma}^2_a(s)}{\Phi_0(I)^2}, \\ \hat{\sigma}^2_{\pi_b}(s) &= \frac{\hat{\sigma}^2_b(s)}{\Phi_0(I)^2}, \\ \hat{\sigma}^2_{\Delta\pi}(s) &= \frac{\hat{\sigma}^2_a(s) + \hat{\sigma}^2_b(s) - 2\hat{\sigma}_{a,b}(s)}{\Phi_0(I)^2}. \end{aligned} \tag{A35}$$

Repeating these three steps S times, we can then compute the average values of the estimated proportions and their difference as

$$\begin{aligned} E(\hat{\pi}_a) &= \pi^*_a = \frac{1}{S}\sum_{s=1}^{S} \hat{\pi}_a(s), \\ E(\hat{\pi}_b) &= \pi^*_b = \frac{1}{S}\sum_{s=1}^{S} \hat{\pi}_b(s), \end{aligned}$$

$$E(\Delta\hat{\pi}) = \Delta\pi^* = \pi^*_a - \pi^*_b. \tag{A36}$$

We also compare the *true* variance of the estimators with the average estimated values:

$$\sigma^2_{\pi_a} = \frac{1}{S}\sum_{s=1}^{S}\hat{\pi}^2_a(s) - (\pi^*_a)^2 \quad \text{versus} \quad E(\hat{\sigma}^2_{\pi_a}) = \frac{1}{S}\sum_{s=1}^{S}\hat{\sigma}^2_{\pi_a}(s),$$

$$\sigma^2_{\pi_b} = \frac{1}{S}\sum_{s=1}^{S}\hat{\pi}^2_b(s) - (\pi^*_b)^2 \quad \text{versus} \quad E(\hat{\sigma}^2_{\pi_b}) = \frac{1}{S}\sum_{s=1}^{S}\hat{\sigma}^2_{\pi_b}(s),$$

$$\sigma^2_{\Delta\pi} = \frac{1}{S}\sum_{s=1}^{S}\Delta\hat{\pi}^2(s) - (\Delta\pi^*)^2 \quad \text{versus} \quad E(\hat{\sigma}^2_{\Delta\pi}) = \frac{1}{S}\sum_{s=1}^{S}\hat{\sigma}^2_{\Delta\pi}(s). \qquad \text{(A37)}$$

To further measure the accuracy of the variance estimators, we compute the coverage ratio of the confidence intervals at ϕ equal to 90% and 95% as

$$CR(\hat{\pi}_a) = \frac{1}{S}\sum_{s=1}^{S} \mathit{1}\{abs(\hat{\pi}_a(s) - \pi^*_a) < \chi_\phi \hat{\sigma}_{\pi_a}(s)\},$$

$$CR(\hat{\pi}_b) = \frac{1}{S}\sum_{s=1}^{S} \mathit{1}\{abs(\hat{\pi}_b(s) - \pi^*_b) < \chi_\phi \hat{\sigma}_{\pi_b}(s)\},$$

$$CR(\hat{\pi}_{\Delta\pi}) = \frac{1}{S}\sum_{s=1}^{S} \mathit{1}\{abs(\Delta\hat{\pi}(s) - \Delta\pi^*) < \chi_\phi \hat{\sigma}^2_{\Delta\pi}(s)\}, \qquad \text{(A38)}$$

where $\mathit{1}\{.\}$ equals one if the condition inside the parenthesis is satisfied (and zero otherwise), and χ_ϕ equals the quantile of the standard normal distribution at (1-$\frac{\phi}{2}$).

C Main Results

In Panel A of Table AI, we examine the properties of the different estimators under the first scenario where the two models a and b achieve a large difference in performance (34.5% in the entire population). The *true* volatilities of the different estimators range between 3.6% and 9.0% and are typically higher for the two largest size groups which contain fewer portfolios. Turning to the properties of the variance estimators, we find that the average value for each model in the entire population is slightly below average (0.5% for model a and 0.3% for model b). In contrast, the volatility estimator for the difference yields an average value that closely matches the *true* volatility (5.1% versus 5.0%). This last property is maintained across all three size groups. Finally, the coverage ratios of the two confidence intervals at 90% and 95% are, in most cases, remarkably accurate. For instance, the coverage ratios for the proportion difference in the entire population are equal to 90.1% and 95.3%, respectively.

In Panel B, we repeat the analysis for the second scenario where the two models yield the same moderate performance. Similar to the previous scenario, the volatility estimators precisely capture the variability of the estimated mispricing proportions for the entire population (they are identical to the *true* values for both models). We also find that the coverage ratios stay close to their theoretical values (88.0% and 93.4% for the intervals at 90% and 95%, respectively). While the results are similar for the big-cap group, they are less accurate in the two smallest size groups (micro- and small-cap). In both groups, the volatility estimators underestimate the *true* volatilities by 12% on average (in relative terms) which implies that the coverage ratios of the confidence intervals are slightly lower than their theoretical values.

Please insert Table AI here

V Overlapping versus Non-overlapping Portfolios

In this section, we show that the mispricing proportion $\hat{\pi}_k$ is estimated more precisely with overlapping portfolios. For simplicity, we consider a population of N stocks whose residual terms $e_{i,t}$ $(i = 1, ..., N)$ are homoscedastic and uncorrelated both across stocks and over time. We also assume that both the number of portfolios and return observations grow large. The number of overlapping portfolios M_o is equal to N and the number of non-overlapping portfolios M_{no} satisfies $\frac{N}{n} \leq M_{no} \leq \frac{N}{n} + 1$, where n is the number of stocks included in each portfolio. We denote by $\sigma^2_{\pi_k}(o)$ the variance of $\hat{\pi}_k$ obtained with overlapping portfolios and by $\sigma^2_{\pi_k}(no)$ the variance of $\hat{\pi}_k$ obtained with non-overlapping portfolios. The two asymptotic variances can be written as

$$\begin{aligned} \sqrt{N}\sigma^2_{\pi_k}(o) &= \frac{1}{\Phi_0(I)^2}\left(1 + 2\sum_{d=1}^{n-1}\rho_d\right)\sigma^2, \\ \sqrt{N}\sigma^2_{\pi_k}(no) &= \frac{1}{\Phi_0(I)^2} n\sigma^2, \end{aligned} \tag{A39}$$

where σ^2 denotes the variance of the indicator function $\mathit{1}(t^k_j)$, and ρ_d is the correlation between the indicator functions associated with the ordered t-statistics $\mathit{1}(t^k_j)$ and $\mathit{1}(t^k_{j+d})$. From Equation (A39), we infer that the overlapping scheme provides efficiency gains if

$$\sigma^2_{\pi_k}(o) < \sigma^2_{\pi_k}(no) \Longleftrightarrow \left(1 + 2\sum_{d=1}^{n-1}\rho_d\right) < n. \tag{A40}$$

To show that the above inequality holds, we proceed in three steps. First, we compare the variances of the averages of the t-statistics for both overlapping and non-overlapping portfolios. Second, we infer from this comparison that a sufficient condition for Equation (A40) to hold is that the correlation between the pair (t_j^k, t_{j+d}^k) is higher than the correlation between the pair $(\mathit{1}(t_j^k), \mathit{1}(t_{j+d}^k))$. Third, we verify that this is the case.

We write the t-statistic averages for overlapping and non-overlapping portfolios as

$$\begin{aligned} \hat{m}_k(o) &= \frac{1}{M_o}\sum_{j=1}^{M_o} t_j^k, \\ \hat{m}_k(no) &= \frac{1}{M_{no}}\sum_{j=1}^{M_{no}} t_j^k, \end{aligned} \tag{A41}$$

and their asympotic variances as

$$\begin{aligned} \sqrt{N}\sigma_{m_k}^2(o) &= 1 + 2\sum_{d=1}^{n-1} \rho_{ts,d}, \\ \sqrt{N}\sigma_{m_k}^2(no) &= n, \end{aligned} \tag{A42}$$

where $\rho_{ts,d}$ is the correlation between the pair (t_j^k, t_{j+d}^k). To determine $\rho_{ts,d}$, we note that the correlation between the portfolio residuals $e_{j,t}$ and $e_{j+d,t}$ is equal to

$$\rho_{e,d} = \max(1 - \frac{d}{n}, 0), \tag{A43}$$

which means that the correlation progressively drops from one to zero as the portfolio distance d approaches $n-1$. Then, we repeat the analysis for the estimated pricing errors. Building on Equation (A10) and Proposition 1, we can write $\sqrt{T}(\hat{\alpha}_j^k - \alpha_j^k) = (\hat{\omega}^k - \omega^k)'\beta_j^k + \frac{1}{\sqrt{T}}\sum_{t=1}^T \hat{\omega}^{k\prime} E_2' Q_x^{-1} x_t e_{j,t} = \frac{1}{\sqrt{T}}\sum_{t=1}^T q e_{j,t}$, where q is a scalar equal to $\omega^{k\prime} E_2' Q_x^{-1} x_t$. Therefore, the asymptotic correlation between $\sqrt{T}\hat{\alpha}_j^k$ and $\sqrt{T}\hat{\alpha}_{j+d}^k$ is

$$\rho_{\alpha,d} = \frac{cov(\sqrt{T}\hat{\alpha}_j^k, \sqrt{T}\hat{\alpha}_{j+d}^k)}{\left(var(\sqrt{T}\hat{\alpha}_j^k) var(\sqrt{T}\hat{\alpha}_{j+d}^k)\right)^{\frac{1}{2}}} = \frac{q^2 \rho_{e,d}\sigma_{e_j}\sigma_{e_j+d}}{q^2\sigma_{e_j}\sigma_{e_j+d}} = \rho_{e,d} = \max(1 - \frac{d}{n}, 0). \tag{A44}$$

Finally, we use Theorem 8.5 in Efron (2010, ch. 8) to show that asymptotically

$$\rho_{ts,d} = \rho_{\alpha,d} = \max(1 - \frac{d}{n}, 0). \tag{A45}$$

Plugging the above expression in Equation (A42), we find $\sqrt{N}\sigma^2_{m_k}(o)$ and $\sqrt{N}\sigma^2_{m_k}(mo)$ are identical because

$$1 + 2\sum_{d=1}^{n-1} \rho_{ts,d} = 1 + 2\left(\frac{n-1}{2}\right) = n. \tag{A46}$$

Therefore, a sufficient condition for the inequality in Equation (A40) to hold is that

$$\forall d \in [1, ..., n-1],\ \rho_d \leq \rho_{ts,d} \text{ and } \exists d \in [1, ..., n-1] \text{ s.t. } \rho_d < \rho_{ts,d}. \tag{A47}$$

To examine the relationship between these two correlations, we denote the bivariate normal distribution for the t-statistics (t^k_j, t^k_{j+d}) by $\phi\left(t^k_j, t^k_{j+d}; \mu_{ts}, \rho_{ts,d}\right)$ to obtain

$$\rho_d = \frac{\int_I \int_I \phi\left(x, y; \mu_{ts}, \rho_{ts,d}\right) dx\ dy}{\sigma^2}. \tag{A48}$$

where μ_{ts} is the t-statistic mean.[6] The double integral in the numerator of Equation (A48) does not have a closed-form expression but can be easily solved numerically. The results show that if $\rho_{ts,d} \in (0,1)$, we have $\rho_d < \rho_{ts,d}$ for all the intervals and t-statistic means that belong to the sets S and $S_{\mu_{ts}}$ defined as $S = \{I = [-a, a];\ a \in [0.15, 0.65]\})$ and $S_{\mu_{ts}} = \{\mu_{ts}; P_A(I) = prob(t^k_j \in I) > 0)$. This implies that the condition in Equation (A47) holds and that $\sigma^2_{\pi_k}(o) < \sigma^2_{\pi_k}(no)$. To illustrate, Figure A2 shows the function $\rho_d = f(\rho_{ts,d})$ for different values for μ_{ts} ($\mu_{ts} = 0,\ 0.5,\ 1,\ 1.5$) and $I = [-0.4, 0.4]$. In all cases, we see that (i) $\rho_{ts,d} = \rho_d$ when $\rho_{ts,d}$ is equal to zero or one; (ii) the function $f(\rho_{ts,d})$ is convex. Therefore, ρ_d is strictly lower than $\rho_{ts,d}$ when $\rho_{ts,d} \in (0,1)$.

Please insert Figure A2 here

VI Additional Results

A Changes in the Estimation Procedure

A.1 Different Values for the Interval I

In the baseline specification, we set the interval I for estimating the mispricing proportion equal to [-0.4,0.4]. To examine if our results are sensitive to this choice, we re-compute the mispricing proportions for each interval I in the set $S = \{I = [-a, a];$ $a \in [0.15, 0.2, ...0.65]\}$. Table AII shows the results for the entire population (Panel A)

[6] For simplicity, we set the mean of t^k_j and t^k_{j+d} equal to the same value. This assumption is motivated by the fact that the t-statistics are spatially ordered and thus likely to have similar means. Allowing for different means does not change the results.

and the three size groups (Panels B to D). We find that the estimated proportions remain largely unchanged—for instance, the averages in the entire population range between 53.4% and 56.7%. This stability is consistent with the observations made by Barras, Scaillet, and Wermers (2010) and Storey (2002).

Please insert Table AII here

A.2 Bootstrap Analysis

In the baseline specification, we rely on asymptotic theory to estimate the mispricing proportion $\hat{\pi}_k$ in Equation (A14)—that is, we assume that the t-statistics of correctly-priced portfolios follow a standard normal distribution $N(0,1)$ in order to replace $P_0(I)$ with $\Phi_0(I)$. We now relax this assumption using the bootstrap approach of Efron (2010, ch. 2) in which the t-statistic of each portfolio is transformed into a statistic called the z-value. This transformation guarantees that the z-value of a correctly-priced portfolio is distributed as a normal $N(0,1)$. Therefore, we can still use Equation (A14) to compute $\hat{\pi}_k$ provided that we use z-values instead of t-statistics.

To compute the z-value of each portfolio j ($j = 1, ..., M$), we use the following procedure. First, we draw, for each bootstrap iteration s ($s = 1, ..., 1,000$), random observations from the original sample of risk factors and residuals to reconstruct the portfolio returns:

$$r_{j,t}(s) = a_{j,0}^k + \hat{b}_{jm} r_{m,t}(s) + \hat{b}_j^k f_t^k(s) + \hat{c}_j z_t(s) + \hat{e}_{j,t}(s), \tag{A49}$$

where we impose that the portfolio is correctly priced ($\alpha_j = 0$) by setting

$$a_{j,0}^k = -\hat{b}_j^k \hat{p}_c^k. \tag{A50}$$

Second, we re-estimate the portfolio t-statistic by regressing the bootstrapped returns on the bootstrapped factors, i.e.,

$$t_j^k(s) = \frac{\hat{\omega}^{k\prime} \hat{\beta}_j^k(s)}{\left(\hat{\omega}^{k\prime} \hat{V}_{\beta_j^k}(s) \hat{\omega}^k\right)^{\frac{1}{2}}}, \tag{A51}$$

where $\hat{\beta}_j(s) = (\hat{a}_j(s), \hat{b}_{j\kappa}(s))'$, $\hat{V}_{\beta_j}(s)$ denote the bootstrapped coefficient vector and its covariance matrix. Third, we repeat the first two steps 1,000 times and compute the bootstrapped cumulative distribution function (cdf) associated with the original

t-statistic as

$$F_{0,j}(t_j^k) = \frac{1}{1,000} \sum_{s=1}^{1,000} 1\{t_j^k(s) \leqslant t_j^k\}. \tag{A52}$$

Finally, we obtain the z-value by inverting the quantile $F_{0,j}(t_j^k)$ using the standard normal cdf Φ_0^{-1}, i.e.,

$$z_j^k = \Phi_0^{-1}(F_{0,j}(t_j^k)). \tag{A53}$$

The empirical results obtained with the bootstrap procedure are reported in Table AIII. The estimated proportions of mispriced portfolios remain largely unchanged. This result implies that the sample size is sufficiently large for the normal distribution to be a good approximation of the *true* t-statistic distribution.

Please insert Table AIII here

B Changes in the Portfolio Formation Procedure

B.1 Different Portfolio Sizes

In this section, we examine the sensitivity of the results to changes in the portfolio formation procedure. To begin, we decrease the number of stocks n in each micro portfolio from 10 to 5 stocks ($n = 5$). The results in Panel A of Table AIV are qualitatively similar except that the mispricing proportions are generally lower. With only 5 stocks in each portfolio, the benefits of diversification are not fully exploited and the detection of the mispriced portfolios becomes more difficult.

Next, Panel B reports the mispricing proportions for micro portfolios formed with 15 stocks ($n = 15$). Overall, the results remain similar to those documented in Table III. We also observe that the volatilities of the estimators are slightly higher because micro portfolios have a higher degree of overlap.

Please insert Table AIV here

B.2 Identical Stock Representation

Next, we tackle the issue of stock representation. While the vast majority of stocks are selected n times in each formation year, some of them are included more or less often. Therefore, the baseline portfolio formation procedure could potentially overweight the importance of some stocks and underweight the importance of others.

To address this issue, we modify the formation procedure to guarantee that each stock is selected exactly n times. For each formation year t_a ($t_a = 1, ..., T_a$), we create the set of micro portfolios following the procedure described above and count the number of times n_i a given stock i ($i = 1, ..., N_{t_a}$) is included in different portfolios. If $n_i < n$, we include stock i in $n - n_i$ additional portfolios with the nearest values to the average return $\hat{\mu}^s_{i,t_a}$. If $n_i > n$, we exclude stock i from $n_i - n$ randomly selected portfolios.

Table AV shows that the estimated mispricing proportions under this alternative portfolio formation remain largely unchanged. The performance differences documented in the paper are therefore not driven by variations in representation across stocks.

Please insert Table AV here

B.3 Alternative Set of Characteristics

We re-build the cross-section of micro portfolios using nine different sets of characteristics for estimating average returns in Equation (A1). The first three specifications simply use each characteristic in isolation (book-to-market, investment, profitability). Specifications 4 and 5 keep our initial measures of investment and profitability but replace the book-to-market ratio with the earnings-to-price and cash flow-to-price ratios. Specifications 6 and 7 keep our initial definitions of book-to-market and profitability but measure investment using infrastructure growth and inventory growth (instead of growth in total assets). Specifications 8 and 9 keep our initial definitions of book-to-market and investment but measure profitability using Return On Equity (ROE) and the Return On Assets (ROA) (instead of operating profitability).

To begin, we examine whether these alternative micro portfolios still produce gains in power and a reduction in beta correlation. In Panel A of Table AVI, we confirm that the interquartile spread in average returns, the median return volatility, and the median number of observations are similar to those reported in Table I. In Panel B, we also measure the independent variation in betas. For each factor included in the CAPM-based models, we compute the beta residuals (the components orthogonal to the other betas) and report the length of the 90%-interval spanned by these residuals. The results show that the dispersion in betas is similar to that of Table II. The overall evidence suggests that the alternative cross-sections of micro portfolios contain sufficient pricing information to discriminate between models.

Next, we estimate the mispricing proportions for the different sets of characteristics. The results in Table AVII reveal strong similarities with those reported in Table III. First, the average mispricing proportions (across the nine specifications) obtained with the standard CAPM remains high, i.e., they reach 72.2%, 57.5%, and 42.7% in the three size groups (74.1%, 60.5%, and 46.4% in the baseline case). Second, the human capital

CAPM maintains its solid performance in all three size groups, e.g., in the tiny- and big-cap groups, the average mispricing proportions are equal to 31.6% and 12.5% (37.7% and 14.9% in the baseline case). Third, the conditional CAPM still performs well in the two largest size groups—the average mispricing proportions are equal to 30.3% and 23.4% (27.3% and 17.9% in the baseline case). Fourth, the liquidity CAPM continues to price tiny-cap portfolios well as the average mispricing proportion equals 33.7% (35.3% in the baseline case). Finally, the three characteristic-based models generally dominate the CAPM-based models and perform equally well except in the tiny-cap group.

Please insert Tables AVI and AVII here

C Traditional Performance Measure

C.1 Testing Procedure

In the paper, we advocate for the use of the mispricing proportion to evaluate models. An alternative approach is to use a version of the traditional performance measures proposed by Gagliardini, Ossola, and Scaillet (2016; GOS hereafter) which can be applied in large cross-sections. This measure is defined as the sum of squared pricing errors, $SS_k = \frac{1}{M}\sum_{i=1}^{M}(\alpha_j^k)^2$, where α_j^k is the *true* pricing error of portfolio j.

The asset pricing test is based on the statistic $\widehat{SS}_k = \frac{\hat{\xi}_k}{\sigma_{\xi_k}}$, where $\hat{\xi}_k = T\sqrt{M}(\hat{Q}_k - \frac{1}{T})$, $\hat{Q}_k = \frac{1}{M}\sum_{i=1}^{M}(\hat{\alpha}_j^k)^2$, $\hat{\alpha}_j^k$ is the estimated portfolio alpha, and $\sigma_{\xi_k}^2$ is the variance of $\hat{\xi}_k$ defined as

$$\sigma_{\xi_k}^2 = 2\lim_{M\to\infty} E\left[\frac{1}{M}\sum_{i=1}^{M}\sum_{j=1}^{M}\frac{\tau_{i,T}^2\tau_{j,T}^2}{\tau_{i,j,T}^2}\left(\omega^{k\prime}E_2'Q_x^{-1}S_{ij}Q_x^{-1}E_2\omega^k\right)^2\right], \tag{A54}$$

where $\tau_{i,T} = \frac{1}{T}\sum_{t=1}^{T} 1_{i,t}$, $\tau_{j,T} = \frac{1}{T}\sum_{t=1}^{T} 1_{j,t}$, $\tau_{i,j,T} = \frac{1}{T}\sum_{t=1}^{T} 1_{i,t}1_{j,t}$, $\omega^k = [1, p_c^{k\prime}]'$, $Q_x = [x_t x_t']$, $S_{ij} = E[e_{i,t}e_{j,t}x_t x_t']$, and E_2 is a $(J+P+2)\times(J+1)$ matrix whose sth row e_s $(s = 1, 3, 4, ..., J+1)$ has one for the sth element and zeros everywhere else.

When the number of portfolios and return observations grow large $(T, M \to \infty)$, Proposition 6 of GOS shows that if model k is correctly specified, we have

$$\widehat{SS}_k = \frac{\hat{\xi}_k}{\hat{\sigma}_{\xi_k}} \xrightarrow{d} N(0,1). \tag{A55}$$

To estimate the variance term $\sigma^2_{\xi_k}$, we use the following consistent estimator

$$\hat{\sigma}^2_{\xi_k} = 2\frac{1}{M}\sum_{i=1}^{M}\sum_{j=1}^{M}\frac{\tau^2_{i,T}\tau^2_{j,T}}{\tau^2_{i,j,T}}\left(\hat{\omega}^{k\prime}E_2'\hat{Q}_{x,i}^{-1}\tilde{S}_{ij}^{-1}\hat{Q}_{x,j}^{-1}E_2^k\hat{\omega}^k\right)^2, \tag{A56}$$

where $\hat{Q}_{x,i} = \frac{1}{T}\sum_{t=1}^{T} 1_{i,t}x_tx_t'$, $\hat{Q}_{x,j} = \frac{1}{T}\sum_{t=1}^{T} 1_{j,t}x_tx_t'$, $\hat{\omega}^k = [1, \hat{p}_c^k]'$, and $\hat{p}_c^k$ is given by Equation (A9). We also need to impose a sparsity condition on the terms S_{ij} $(i, j = 1, ..., M)$ in the double sum (see assumption A.4 in GOS). To this end, we use the estimator $\tilde{S}_{ij} = \hat{S}_{ij}1(\left\|\hat{S}_{ij}\right\| \geq \kappa_{ij})$, where $\hat{S}_{ij} = \frac{1}{T}\sum_{t=1}^{T} 1_{i,t}1_{j,t}\hat{e}_{i,t}\hat{e}_{j,t}x_tx_t'$ and κ_{ij} is the threshold parameter set equal to $0.067\cdot(\log(M))^{\frac{1}{2}}/\tau_{i,j,T}$.

C.2 Empirical Results

Table AVIII reports the statistic $\widehat{SS}_k$ for the entire portfolio population and the three size groups. The results closely mirror those reported in Table III. First, the null hypothesis of correct specification is rejected in all but one case—when the five-factor model is tested on big-cap portfolios (the p-value is equal to 0.22). Second, the ranking of the CAPM-based models is the same in each size group. Third, the three characteristics-based models generally produce lower pricing errors than the CAPM-based models. Because $\widehat{SS}_k$ is consistently lower for the five-factor model, it is tempting to say that it dominates the three- and q-factor models. However, we cannot make such claims without comparison tests which have not been developed for large cross-sections yet.

Please insert Table AVIII here

D Comparison with Individual Stocks

D.1 Construction of the Sample

In this section, we evaluate the different models using individual stocks as test assets. Similar to micro portfolios, we classify stocks in three size groups (tiny-, small-, and big-cap). At the end of June each year, we partition all existing stocks using as breakpoints the 20th and 50th percentiles of the market capitalization for NYSE stocks. Then, we classify each stock in one of the three size groups based on the highest frequency of observations. We also require that each individual stock has a minimum of 60 monthly return observations to compute its t-statistic. The resulting cross-section includes a total of 6,651 individual stocks (3,548 tiny-cap, 1,379 small-cap, 1,724 big-cap).

D.2 Mispricing Proportion

We begin our analysis by examining the proportion of mispriced stocks. Panel A of Table AIX shows that the estimated proportions are significantly lower than those obtained with micro portfolios, e.g., the averages in the tiny- and small-cap groups are equal to 7.5% and 8.9% (versus 55.3% and 49.4% for micro portfolios). Coupled with high estimation uncertainty, these low estimated proportions lead us to conclude that none of the models are misspecified, i.e., we cannot reject the null hypothesis of correct specification. Because the estimated pricing errors of individual stocks are too volatile, we are unable to detect mispricing in the data.

D.3 Traditional Performance Measure

Alternatively, we can use the SS measure proposed by GOS. Because this measure aggregates pricing errors, it allows us to sidestep the challenge of detecting mispricing at the individual stock level. As shown in Panel B, the tests based on the SS measure indicates that the models are all mispecified in the entire stock population (the p-values are all equal to zero). This performance analysis is consistent with (i) the previous results of GOS which show that commonly-used models are strongly rejected at the individual stock level; (ii) the performance evaluation obtained with micro portfolios (see Tables III and AIV).

There is no available testing procedure for comparing models based on the SS measure in large cross-sections. However, a casual observation of the estimated values reveals no striking performance differences across the models. A likely culprit for this result is the time-variation in individual stock betas as firms go through different business cycles and stages of development—the empirical evidence shows that this variation is significantly stronger for stocks than for portfolios (e.g., Andersen et al. (2006), Fama and French (1997)). The time-variation in betas introduces an additional source of misspecification (e.g., Jagannathan and Wang (1996)) which affect all time-invariant models. It can therefore smooth out the performance differences observed with micro portfolios.[7]

Please insert Table AIX here

[7] We could explicitly specify the dynamics of the individual stock betas. However, time-varying models are difficult to estimate because of the large number of parameters. In addition, Ghysels (1998) shows that a wrong specification of time-varying betas may result in large pricing errors, possibly greater than those produced by a constant-beta model. The evidence in GOS reveals that the time-varying specifications of the tested models are also strongly rejected in the data.

References

[1] Andersen T. G., T. Bollerslev, F. X. Diebold, and G. Wu, 2006, Realized Beta: Persistence and Predictability, Advances in Econometrics: Econometric Analysis of Economic and Financial Time Series, Elsevier.

[2] Avramov D., and T. Chordia, 2006, Asset Pricing Models and Financial Market Anomalies, Review of Financial Studies 19, 1001-1038.

[3] Barras L., O. Scaillet, and R. Wermers, 2010, False Discoveries in Mutual Fund Performance: Measuring Luck in Estimated Alphas, Journal of Finance 65, 179-216.

[4] Berk J., 2000, Sorting Out Sorts, Journal of Finance 55, 407-427.

[5] Brennan M. J., T. Chordia, and A. Subrahmanyam, 1998, Alternative Factor Specifications, Security Characteristics and the Cross-Section of Expected Stock Returns, Journal of Financial Economics 49, 345-373.

[6] Efron B., 2010, Large-Scale Inference, Cambridge University Press.

[7] Fama E. F., and K. R. French, 1997, Industry Cost of Equity, Journal of Financial Economics 43, 153-193.

[8] Fama E. F., and K. R. French, 2006, Profitability, Investment, and Average Returns, Journal of Financial Economics 82, 491-518.

[9] Fama E. F., and K. R. French, 2008, Dissecting Anomalies, Journal of Finance 63, 1653-1678.

[10] Fama E. F., and K. R. French, 2015, A Five-Factor Asset Pricing Model, Journal of Financial Economics 116, 1-22.

[11] Farcomeni A., 2006, Some Results on the Control of the False Discovery Rate under Dependence, The Scandinavian Journal of Statistics 34, 275-297.

[12] Gagliardini P., E. Ossola, and O. Scaillet, 2016, Time-varying Risk Premium in Large Cross-sectional Equity Datasets, Econometrica 84, 985-1046.

[13] Genovese C., and L. Wasserman, 2004, A Stochastic Process Approach to False Discovery Control, Annals of Statistics 32, 1035-1061.

[14] Ghysels E., 1998, On Stable Factor Structures in the Pricing of Risk: Do Time-Varying Betas Help or Hurt?, Journal of Finance 53, 549–573.

[15] Hou K., C. Xue, and L. Zhang, 2015, Digesting Anomalies: An Investment Approach, Review of Financial Studies 28, 650–705.

[16] Jagannathan R., and Z. Wang, 1996, The Conditional CAPM and the Cross-Section of Expected Returns, Journal of Finance 51, 3-53.

[17] Newey W. K., and K. D. West, 1987, A Simple, Positive Semi-Definite, Heteroscedasticity and Autocorrelation Consistent Covariance Matrix, Econometrica 55, 703-708.

[18] Storey J. D., 2002, A Direct Approach to False Discovery Rates, Journal of the Royal Statistical Society 64, 479–498.

[19] White H., 1980, A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity, Econometrica 48, 817–838.

Table AI

Monte Carlo Analysis

Panel A reports the properties of the proportion estimators under the first scenario where there is a large performance difference between the two misspecified models a and b. For the entire population and each size group (tiny-, small-, and big-cap), the first column shows the average values of the estimated proportions of mispriced portfolios for both models and their difference. The second and third columns compare the true volatilities of the estimated proportions and their difference with the estimated volatilities. The fourth and fifth columns show the coverage ratios of the confidence intervals at 90% and 95% for the estimated proportions and their difference. In Panel B, we repeat the analysis under the second scenario where the two models produce the same moderate performance. The total number of iterations is equal to 5,000.

Panel A: Large Performance Difference

		Volatility		Confidence Interval	
	Mean	True	Estimated	90%-coverage	95%-coverage
All Portfolios					
Model a	30.6	3.6	3.1	91.3	94.5
Model b	65.1	4.3	4.0	92.5	95.8
Difference	-34.5	5.0	5.1	90.1	95.3
Tiny-Cap Portfolios					
Model a	32.4	4.8	4.6	93.2	96.1
Model b	62.5	6.7	5.7	89.8	93.5
Difference	-30.1	8.0	7.7	87.9	93.7
Small-Cap Portfolios					
Model a	32.4	7.4	6.6	92.1	95.3
Model b	63.1	5.6	5.9	94.0	96.3
Difference	-30.6	9.0	9.1	90.3	94.8
Big-Cap Portfolios					
Model a	15.4	6.2	5.6	92.9	95.8
Model b	68.3	4.6	5.3	95.4	97.7
Difference	-52.9	7.3	7.6	91.0	95.5

Table AI

Monte Carlo Analysis (Continued)

Panel B: No Performance Difference

		Volatility		Confidence Interval	
	Mean	True	Estimated	90%-coverage	95%-coverage
All Portfolios					
Model a	40.0	2.9	2.9	94.1	96.9
Model b	40.0	2.9	2.9	94.1	96.7
Difference	0.0	4.0	3.8	88.0	93.4
Tiny-Cap Portfolios					
Model a	35.9	4.4	3.9	92.0	93.5
Model b	35.7	4.4	3.8	90.1	95.3
Difference	0.2	5.9	5.3	85.2	91.3
Small-Cap Portfolios					
Model a	33.3	6.5	5.8	90.0	93.3
Model b	33.0	6.3	5.8	92.7	95.3
Difference	0.3	9.5	8.2	84.4	90.3
Big-Cap Portfolios					
Model a	34.9	5.3	5.2	93.1	96.1
Model b	35.0	5.1	5.3	94.7	97.2
Difference	-0.1	7.4	7.2	88.7	93.6

Table AII

Performance Analysis with Different Intervals

Panel A reports, for the entire population, the estimated proportions of micro portfolios that are mispriced by the standard CAPM, the CAPM-based models (conditional, human capital, intertemporal, and liquidity CAPMs), and the characteristic-based models (three-factor, q-factor, and five-factor models) across the set of intervals S={$I = [-a, a]$; $a = [0.15, 0.20, ..., 0.65]$}. In Panels B to D, we repeat the analysis for the three size groups (tiny-, small- and big-cap).

Panel A: All Portfolios

	Interval Bound a										
	0.15	0.20	0.25	0.3	0.35	0.40	0.45	0.50	0.55	0.60	0.65
CAPM	64.5	63.8	62.5	63.5	63.6	64.3	63.6	63.3	62.7	62.2	61.5
Conditional	41.2	42.5	42.1	42.5	42.9	43.4	43.7	42.5	41.0	41.0	40.8
Human Capital	36.8	36.1	37.3	36.9	36.2	35.1	35.6	35.8	35.6	35.5	34.4
Intertemporal	61.8	60.6	60.9	60.0	59.8	60.2	59.2	58.7	58.6	58.3	57.7
Liquidity	36.8	37.3	35.0	34.7	34.0	34.4	34.0	33.4	33.8	33.6	32.9
Average	48.2	48.1	47.6	47.5	47.2	47.4	47.2	46.7	46.4	46.1	45.4
3-factor Model	34.8	36.8	38.0	38.6	37.5	36.4	36.2	36.4	35.9	35.3	35.0
q-factor Model	48.9	45.0	45.1	45.8	45.0	45.3	43.3	41.8	41.8	41.8	41.4
5-factor Model	32.9	30.7	29.6	30.8	30.4	30.3	29.1	29.2	29.0	29.7	29.2
Average	38.8	37.5	37.5	38.3	37.6	37.3	36.2	35.8	35.5	35.5	35.2

Panel B: Tiny-Cap Portfolios

	Interval Bound a										
	0.15	0.20	0.25	0.3	0.35	0.40	0.45	0.50	0.55	0.60	0.65
CAPM	76.1	74.9	73.9	74.2	73.6	74.1	74.1	73.7	73.4	73.1	72.1
Conditional	58.7	63.2	63.2	61.3	62.6	61.5	60.9	60.5	60.1	58.9	58.3
Human Capital	37.3	37.8	38.0	38.8	39.2	37.7	39.4	39.0	38.3	38.5	36.4
Intertemporal	72.4	70.5	70.2	70.1	69.0	68.0	67.6	67.7	67.0	65.6	64.9
Liquidity	37.2	37.2	38.2	37.7	36.0	35.3	35.2	36.1	30.0	35.4	35.4
Average	56.3	56.7	56.7	56.4	56.1	55.3	55.5	55.4	55.0	54.3	53.4
3-factor Model	55.8	58.0	58.0	55.8	54.5	53.5	51.7	50.9	49.9	48.5	47.9
q-factor Model	72.8	71.1	69.8	69.4	68.6	68.7	67.4	66.1	65.8	66.0	66.1
5-factor Model	47.1	46.5	45.7	47.0	45.3	43.5	43.1	44.0	43.2	44.0	43.0
Average	58.5	58.5	57.8	57.4	56.1	55.3	54.0	53.6	53.0	52.8	52.3

Table AII

Performance Analysis with Different Intervals (Continued)

Panel C: Small-Cap Portfolios

	Interval Bound a										
	0.15	0.20	0.25	0.3	0.35	0.40	0.45	0.50	0.55	0.60	0.65
CAPM	63.3	61.0	60.5	60.2	59.4	60.5	59.1	58.8	57.1	57.7	58.6
Conditional CAPM	32.0	30.7	31.4	30.8	29.1	27.3	26.0	27.3	27.3	28.7	28.0
Human Capital CAPM	57.0	51.5	53.0	51.1	47.0	45.1	47.1	45.9	44.3	44.5	45.6
Intertemporal CAPM	62.4	67.7	70.3	68.8	68.0	67.0	65.9	64.6	63.2	62.7	63.0
Liquidity CAPM	57.0	54.8	52.9	49.3	46.5	47.1	46.2	44.5	44.0	45.6	43.5
Average	54.3	53.1	53.6	52.0	50.0	49.4	48.8	48.2	47.1	47.8	47.7
3-factor Model	23.0	19.9	18.4	19.9	19.7	18.3	20.1	22.0	21.3	23.2	23.5
q-factor Model	30.2	22.7	23.9	24.1	21.3	22.8	22.0	20.9	20.6	19.2	19.0
5-factor Model	12.3	15.8	14.6	18.5	22.0	23.4	21.6	19.5	19.8	20.6	20.0
Average	24.8	19.5	18.9	20.8	21.0	21.5	21.3	20.8	20.6	20.0	20.8

Panel B: Big-Cap Portfolios

	Interval Bound a										
	0.15	0.20	0.25	0.3	0.35	0.40	0.45	0.50	0.55	0.60	0.65
CAPM	40.7	42.4	39.6	43.5	45.6	46.4	44.8	45.0	44.7	42.7	41.4
Conditional CAPM	12.5	16.6	20.3	18.3	16.6	17.9	18.5	17.6	17.7	18.1	17.5
Human Capital CAPM	2.3	7.2	9.0	11.6	13.2	14.9	11.8	13.0	12.4	12.4	10.8
Intertemporal CAPM	62.5	63.6	60.9	57.0	56.4	57.2	55.5	55.0	53.4	51.5	51.2
Liquidity CAPM	29.7	28.9	27.8	31.2	30.9	30.5	28.9	28.2	27.1	26.8	26.9
Average	29.1	31.7	31.5	32.3	32.5	33.3	31.9	31.8	31.1	30.3	29.6
3-factor Model	8.1	6.0	12.3	17.9	16.6	15.2	16.9	17.8	18.6	17.3	17.3
q-factor Model	5.5	8.4	10.4	13.9	14.9	14.6	10.2	7.6	8.4	9.3	7.9
5-factor Model	14.5	9.5	8.0	6.4	5.8	7.7	5.6	5.9	6.6	7.0	7.3
Average	8.7	8.0	10.2	12.7	12.4	12.5	10.9	10.4	11.2	11.1	10.8

Table AIII

Performance Analysis with the Bootstrap

This table reports the estimated proportions of micro portfolios that are mispriced by the standard CAPM, the CAPM-based models (conditional, human capital, intertemporal, and liquidity CAPMs), and the characteristic-based models (three-factor, q-factor, and five-factor models) for the entire portfolio population (All) and the three size groups (tiny-, small-, and big-cap) using a bootstrap approach. Figures in parentheses denote the estimated volatilities of the proportion estimates.

		Size Groups		
	All	Tiny-cap	Small-cap	Big-cap
CAPM	62.9 (2.3)	72.8 (5.7)	57.7 (5.0)	43.3 (4.4)
Conditional CAPM	41.3 (1.9)	60.0 (4.8)	26.2 (4.8)	14.3 (5.7)
Human Capital CAPM	34.8 (3.2)	36.2 (5.7)	45.1 (5.9)	8.9 (5.6)
Intertemporal CAPM	58.4 (2.0)	67.8 (4.2)	66.3 (5.1)	54.2 (4.2)
Liquidity CAPM	32.2 (2.4)	33.6 (6.0)	45.7 (5.4)	28.4 (5.4)
Average	45.9	54.1	48.2	29.8
3-factor Model	35.4 (2.8)	52.1 (5.7)	16.6 (6.6)	14.0 (5.4)
q-factor Model	42.6 (3.0)	67.4 (6.1)	22.4 (4.5)	11.3 (5.2)
5-factor Model	27.4 (2.4)	43.7 (5.5)	19.3 (5.5)	4.7 (5.3)
Average	35.1	54.4	19.4	10.0

Table AIV

Performance Analysis with Different Portfolio Sizes

Panel A reports the estimated proportions of micro portfolios that are mispriced by the standard CAPM, the CAPM-based models (conditional, human capital, intertemporal, and liquidity CAPMs), and the characteristic-based models (three-factor, q-factor, and five-factor models) for the entire portfolio population (All) and the three size groups (tiny-, small-, and big-cap) using five stocks (n=5). Figures in parentheses denote the estimated volatilities of the proportion estimates. In Panel B, we repeat the analysis using micro portfolios made up of 15 stocks (n=15).

Panel A: Five Stocks

		Size Groups		
	All	Tiny-cap	Small-cap	Big-cap
CAPM	48.3 (2.2)	58.8 (5.6)	41.6 (5.2)	31.7 (4.7)
Conditional CAPM	37.2 (2.1)	51.4 (5.3)	26.2 (5.2)	19.1 (5.6)
Human Capital CAPM	26.0 (3.1)	28.5 (4.6)	31.6 (6.1)	15.2 (5.2)
Intertemporal CAPM	46.6 (2.4)	54.2 (4.4)	50.5 (4.6)	39.4 (4.3)
Liquidity CAPM	36.1 (2.3)	42.8 (4.2)	39.6 (4.9)	19.7 (5.0)
Average	38.9	47.1	37.9	25.0
3-factor Model	28.3 (2.5)	34.9 (4.2)	26.8 (4.8)	15.5 (5.4)
q-factor Model	32.4 (2.4)	51.3 (5.2)	13.2 (4.8)	8.3 (5.0)
5-factor Model	24.3 (2.2)	36.7 (3.8)	24.8 (4.8)	0.0 (4.9)
Average	28.4	41.0	21.6	8.0

Panel B: Fifteen Stocks

		Size Groups		
	All	Tiny-cap	Small-cap	Big-cap
CAPM	71.2 (2.5)	78.3 (5.7)	70.8 (4.6)	56.3 (5.1)
Conditional CAPM	47.2 (2.5)	69.7 (5.2)	34.8 (5.4)	18.9 (5.5)
Human Capital CAPM	37.1 (3.8)	41.8 (5.1)	43.3 (6.1)	23.0 (6.1)
Intertemporal CAPM	67.1 (2.0)	75.9 (3.6)	75.9 (4.6)	59.5 (4.7)
Liquidity CAPM	34.0 (2.7)	32.0 (5.5)	57.7 (5.7)	22.7 (6.2)
Average	51.3	59.6	56.5	36.1
3-factor Model	44.0 (2.9)	63.6 (6.0)	24.8 (7.2)	18.2 (6.0)
q-factor Model	51.3 (3.0)	73.2 (6.1)	23.9 (5.0)	10.7 (4.5)
5-factor Model	34.4 (2.2)	55.5 (6.7)	17.5 (7.1)	3.8 (5.0)
Average	41.8	64.1	22.1	10.9

Table AV

Performance Analysis with Equal Stock Representation

This table reports the estimated proportions of micro portfolios that are mispriced by the standard CAPM, the CAPM-based models (conditional, human capital, intertemporal, and liquidity CAPMs), and the characteristic-based models (three-factor, q-factor, and five-factor models) for the entire portfolio population (All) and the three size groups (tiny-, small-, and big-cap) using a portfolio formation procedure in which all stocks are selected the same number of times in each formation year. Figures in parentheses denote the estimated volatilities of the proportion estimates.

		Size Groups		
	All	Tiny-cap	Small-cap	Big-cap
CAPM	62.6 (2.2)	71.6 (5.5)	60.5 (4.5)	44.8 (5.4)
Conditional CAPM	46.2 (1.7)	65.2 (4.7)	33.1 (4.7)	13.7 (5.9)
Human Capital CAPM	34.7 (3.5)	35.8 (5.8)	49.2 (5.4)	13.1 (5.4)
Intertemporal CAPM	57.9 (2.3)	65.1 (5.0)	62.3 (5.4)	55.1 (4.2)
Liquidity CAPM	35.4 (2.5)	41.5 (5.7)	48.8 (5.3)	29.6 (5.0)
Average	47.5	55.8	50.9	31.2
3-factor Model	35.2 (2.5)	45.7 (6.5)	26.5 (6.6)	20.0 (5.7)
q-factor Model	47.4 (2.7)	68.6 (5.7)	24.5 (4.7)	14.0 (4.8)
5-factor Model	27.6 (2.9)	43.9 (6.9)	14.5 (5.6)	3.8 (4.4)
Average	36.7	52.7	21.8	12.6

Table AVI

Properties of Micro Portfolio with Different Sets of Characteristics

Panel A reports the interquartile spread in average returns (annualized), the median return volatility (annualized), and the median number of return observations across micro portfolios for the following sets of characteristics: (i) three cases where book-to-market (bm), investment, and profitability are used in isolation (bm, inv, prof); (ii) two cases where bm is replaced with the earnings/price and cash flows/price ratios (bm_1, bm_2); (iii) two cases where investment is measured with infrastructure and inventory growth (inv_1, inv_2); (iv) two cases where profitability is measured with the return on equity and the return on assets ($prof_1$, $prof_2$); (v) the baseline case used in the paper (Base). Panel B measures the dispersion in the portfolio betas on each risk factor specific to the CAPM-based models (conditional, human capital, intertemporal, and liquidity CAPMs) for the different sets of characteristics. Each column measures the length of the 90%-interval of the beta residuals (the components orthogonal to the other factor betas).

Panel A: Portfolio Returns

	Set of Characteristics										
	bm	inv	prof	bm_1	bm_2	inv_1	inv_2	$prof_1$	$prof_2$	Avg	Base
Median Obs.	384	390	384	366	378	384	402	372	372	381	390
Return Spread	7.5	7.2	6.3	6.5	6.6	7.4	6.6	7.1	7.4	7.0	7.3
Median Vol.	29.7	29.7	29.2	26.5	27.0	29.6	29.7	27.9	27.9	28.6	30.2

Panel B: Independent Variation in Betas

	Set of Characteristics										
	bm	inv	prof	bm_1	bm_2	inv_1	inv_2	$prof_1$	$prof_2$	Avg	Base
Conditional	1.91	1.90	1.80	1.35	1.40	1.80	1.60	2.28	2.21	1.81	1.75
Human Capital	1.35	1.44	1.49	1.10	1.17	1.47	1.28	1.66	1.67	1.40	1.41
Intertemporal	2.25	2.56	2.29	1.27	1.32	2.31	1.67	3.23	3.15	2.22	2.56
Liquidity	1.25	1.25	1.32	1.09	1.1	1.24	1.14	1.40	1.39	1.24	1.26
Average	1.70	1.79	1.73	1.21	1.25	1.71	1.43	2.1	2.1	1.67	1.75

Table AVII

Performance Analysis with Different Sets of Characteristics

Panel A reports, for the entire population, the estimated proportions of micro portfolios that are mispriced by the standard CAPM, the CAPM-based models (conditional, human capital, intertemporal, and liquidity CAPMs), and the characteristic-based models (three-factor, q-factor, and five-factor models) for different sets of characteristics. The specifications are the following: (i) three cases where book-to-market (bm), investment, and profitability are used in isolation (bm, inv, prof); (ii) two cases where bm is replaced with the earnings/price and cash flows/price ratios (bm_1, bm_2); (iii) two cases where investment is measured with infrastructure and inventory growth (inv_1, inv_2); (iv) two cases where profitability is measured with the return on equity and the return on assets (prof_1, prof_2); (v) the baseline case used in the paper (Base). In Panels B to D, we repeat the analysis for the three size groups (tiny-, small- and big-cap).

Panel A: All Portfolios

	Set of Characteristics										
	bm	inv	prof	bm_1	bm_2	inv_1	inv_2	prof_1	prof_2	Avg	Base
CAPM	52.9	62.3	52.8	63.6	66.6	62.4	59.8	66.5	66.0	61.4	64.3
Conditional	31.9	47.3	31.5	37.7	40.2	50.3	39.1	60.8	60.1	44.3	43.4
Human Capital	29.7	27.0	15.4	38.1	37.5	39.7	28.6	43.7	33.1	31.4	35.1
Intertemporal	38.6	57.9	35.0	46.9	47.7	55.5	53.6	56.8	56.1	49.8	60.2
Liquidity	21.2	36.0	22.6	29.9	18.9	32.4	24.1	48.6	52.4	31.8	34.4
Average	34.8	46.1	31.5	43.2	42.2	48.0	41.1	55.3	53.5	44.0	47.5
3-factor Model	26.2	35.7	15.1	38.2	33.8	36.6	33.9	34.2	398	32.6	36.4
q-factor Model	38.9	43.7	44.9	36.3	34.4	44.3	39.3	40.1	43.5	40.6	45.3
5-factor Model	27.1	33.7	18.8	26.5	26.4	31.7	31.3	27.2	29.8	28.1	30.3
Average	30.8	37.7	26.3	33.7	31.6	37.6	34.9	33.9	37.8	33.8	37.3

Panel B: Tiny-Cap Portfolios

	Set of Characteristics										
	bm	inv	prof	bm_1	bm_2	inv_1	inv_2	prof_1	prof_2	Avg	Base
CAPM	63.9	73.5	69.5	74.7	77.6	73.7	73.7	70.9	72.2	72.2	74.1
Conditional	42.4	70.0	47.1	46.8	39.2	64.5	46.9	71.6	70.8	55.5	61.5
Human Capital	45.2	26.6	14.2	48.2	45.3	36.6	22.4	40.7	28.6	31.6	37.7
Intertemporal	44.0	63.1	33.1	37.4	32.7	54.5	53.4	57.8	58.7	48.3	68.0
Liquidity	24.1	35.8	29.7	30.3	16.9	34.6	19.8	57.0	55.1	33.7	35.3
Average	43.9	53.8	38.7	47.5	42.3	52.8	43.2	59.7	57.1	48.8	55.3
3-factor Model	38.9	52.4	20.5	52.6	50.4	49.1	40.5	48.6	49.2	44.7	53.5
q-factor Model	63.5	67.8	69.6	61.4	66.8	73.5	64.3	67.5	65.8	66.7	68.7
5-factor Model	39.5	49.1	28.3	43.1	41.7	47.1	37.4	43.5	44.8	41.6	43.5
Average	47.3	56.5	39.5	52.4	53.0	56.6	47.4	53.2	53.3	51.0	55.2

Table AVII

Performance Analysis with Different Sets of Characteristics (Continued)

Panel C: Small-Cap Portfolios

	Set of Characteristics										
	bm	inv	prof	bm_1	bm_2	inv_1	inv_2	$prof_1$	$prof_2$	Avg	Base
CAPM	50.9	50.0	40.6	62.2	62.0	64.3	49.2	69.8	68.1	57.5	60.5
Conditional	21.8	23.2	13.6	27.0	32.8	48.4	33.2	40.2	32.8	30.3	27.3
Human Capital	22.1	27.6	40.0	18.1	40.8	59.8	52.1	68.1	54.1	42.5	45.1
Intertemporal	50.6	55.1	50.0	47.4	54.8	61.9	54.2	69.4	66.7	56.6	67.0
Liquidity	32.8	42.7	19.4	49.9	36.4	54.9	37.3	55.9	61.1	43.4	47.1
Average	35.6	39.7	32.7	40.9	45.4	57.8	45.2	60.7	56.5	46.1	49.4
3-factor Model	17.0	18.0	7.8	32.1	14.4	27.3	32.8	25.8	30.8	22.9	18.3
q-factor Model	9.8	22.8	13.6	20.6	16.9	9.3	13.4	17.2	21.1	14.3	22.8
5-factor Model	18.0	22.1	5.3	20.6	12.9	17.2	26.2	14.1	17.5	17.1	23.4
Average	15.0	21.0	8.9	24.5	9.7	17.9	24.1	19.1	23.1	18.1	21.5

Panel D: Big-Cap Portfolios

	Set of Characteristics										
	bm	inv	prof	bm_1	bm_2	inv_1	inv_2	$prof_1$	$prof_2$	Avg	Base
CAPM	30.8	49.1	27.2	44.7	50.3	36.3	38.5	55.0	52.1	42.7	46.4
Conditional	10.0	20.4	8.4	29.0	33.4	20.5	21.2	32.0	35.7	23.4	17.9
Human Capital	5.0	10.8	10.0	14.0	10.2	16.9	17.8	18.0	10.3	12.5	14.9
Intertemporal	32.2	61.3	27.3	44.7	53.3	47.5	48.2	57.7	59.4	48.0	57.2
Liquidity	28.3	33.9	11.1	41.6	28.8	23.9	19.0	27.0	45.7	28.9	30.5
Average	21.3	35.1	16.8	34.8	35.2	29.0	29.0	38.0	40.7	31.1	33.3
3-factor Model	15.7	15.3	10.0	17.5	19.5	18.1	20.1	13.0	29.3	17.6	15.3
q-factor Model	10.9	10.2	18.6	4.2	2.6	11.8	7.0	5.6	18.8	9.9	14.6
5-factor Model	7.9	11.1	9.9	13.9	9.5	11.2	22.4	6.2	11.1	10.0	7.7
Average	11.5	12.2	12.8	7.7	10.5	13.7	16.5	8.3	19.8	12.6	12.5

Table AVIII

Performance Analysis with the Traditional Measure

This table reports the traditional performance measure proposed by Gagliardini, Ossola, and Scaillet (2016) under the standard CAPM, the CAPM-based models (conditional, human capital, intertemporal, and liquidity CAPMs), and the characteristic-based models (three-factor, q-factor, and five-factor models) for the entire portfolio population (All) and the three size groups (tiny-, small-, and big-cap). The performance measure is defined as the sum of squared pricing errors and provides a valid inference in large cross-sections. Figures in parentheses denote the p-values under the null hypothesis that the model is correctly specified.

		Size Groups		
	All	Tiny-cap	Small-cap	Big-cap
CAPM	146.0 (0.0)	131.8 (0.0)	62.6 (0.0)	40.3 (0.0)
Conditional CAPM	85.8 (0.0)	84.6 (0.0)	30.7 (0.0)	19.8 (0.0)
Human Capital CAPM	71.1 (0.0)	47.4 (0.0)	54.4 (0.0)	14.7 (0.0)
Intertemporal CAPM	133.5 (0.0)	100.7 (0.0)	68.7 (0.0)	43.3 (0.0)
Liquidity CAPM	76.3 (0.0)	59.8 (0.0)	51.4 (0.0)	25.7 (0.0)
Average	102.6	84.9	53.6	28.8
3-factor Model	56.2 (0.0)	48.1 (0.0)	29.8 (0.0)	14.0 (0.0)
q-factor Model	72.7 (0.0)	88.6 (0.0)	9.2 (0.0)	5.4 (0.0)
5-factor Model	5.5 (0.0)	6.6 (0.0)	3.5 (0.0)	0.7 (0.23)
Average	44.8	47.8	14.1	6.7

Table AIX

Performance Analysis with Individual Stocks

Panel A reports the estimated proportions of individual stocks that are mispriced by the standard CAPM, the CAPM-based models (conditional, human capital, intertemporal, and liquidity CAPMs), and the characteristic-based models (three-factor, q-factor, and five-factor models) for the entire portfolio population (All) and the three size groups (tiny-, small-, and big-cap). Figures in parentheses denote the estimated volatilities of the proportion estimates. Panel B conducts the same analysis using the traditional performance measure proposed by Gagliardini, Ossola, and Scaillet (2016). Figures in parentheses denote the p-values under the null hypothesis that the model is correctly specified.

Panel A: Proportion of Mispriced Stocks

		Size Groups		
	All	Tiny-cap	Small-cap	Big-cap
CAPM	14.7 (15.5)	8.5 (21.7)	8.3 (17.8)	32.8 (14.3)
Conditional CAPM	15.5 (15.4)	9.1 (21.4)	9.0 (17.5)	32.3 (14.3)
Human Capital CAPM	12.8 (15.7)	5.2 (21.9)	10.3 (17.6)	29.6 (14.6)
Intertemporal CAPM	14.1 (15.6)	7.6 (21.8)	7.8 (17.8)	32.0 (14.4)
Liquidity CAPM	14.4 (15.6)	7.2 (21.8)	9.3 (17.8)	34.0 (14.2)
Average	14.3	7.5	9.0	32.2
3-factor Model	5.7 (16.0)	3.2 (21.9)	0.0 (18.4)	20.7 (14.2)
q-factor Model	14.3 (16.0)	3.8 (22.0)	0.0 (18.1)	13.1 (15.4)
5-factor Model	7.2 (16.0)	3.7 (21.9)	2.8 (18.1)	18.0 (15.1)
Average	6.1	3.6	1.0	17.2

Panel B: Traditional Performance Measure

		Size Groups		
	All	Tiny-cap	Small-cap	Big-cap
CAPM	15.1 (0.0)	3.7 (0.0)	4.5 (0.0)	20.3 (0.0)
Conditional CAPM	15.0 (0.0)	3.8 (0.0)	4.3 (0.0)	20.4 (0.0)
Human Capital CAPM	14.4 (0.0)	3.5 (0.0)	4.0 (0.0)	18.0 (0.0)
Intertemporal CAPM	15.4 (0.0)	3.7 (0.0)	4.8 (0.0)	21.1 (0.0)
Liquidity CAPM	15.2 (0.0)	3.6 (0.0)	4.8 (0.0)	21.2 (0.0)
Average	15.0	3.7	4.5	20.2
3-factor Model	12.0 (0.0)	3.9 (0.00)	0.5 (0.30)	17.4 (0.0)
q-factor Model	9.2 (0.0)	3.7 (0.0)	0.4 (0.34)	12.5 (0.0)
5-factor Model	6.0 (0.0)	3.6 (0.00)	1.4 (0.08)	8.2 (0.23)
Average	9.1	3.8	0.8	12.7

Figure A1

Forming the Cross-Section of Micro Portfolios

This figure illustrates the procedure for forming the set of micro portfolios sorted based on average returns using an hypothetical population of 50 individual stocks and a two-year sample period. Each dot represents the estimated average return taken by each stock (S1, S2,...). For each stock, the procedure consists of forming an equally-weighted portfolio that includes the stock itself and 9 additional stocks with the nearest estimated average returns. Then, the portfolio returns are chained across years 1 and 2 to maintain a stable average return over time. This procedure yields a cross-section of 50 micro portfolios ranging from P(Return 1) to P(Return 50).

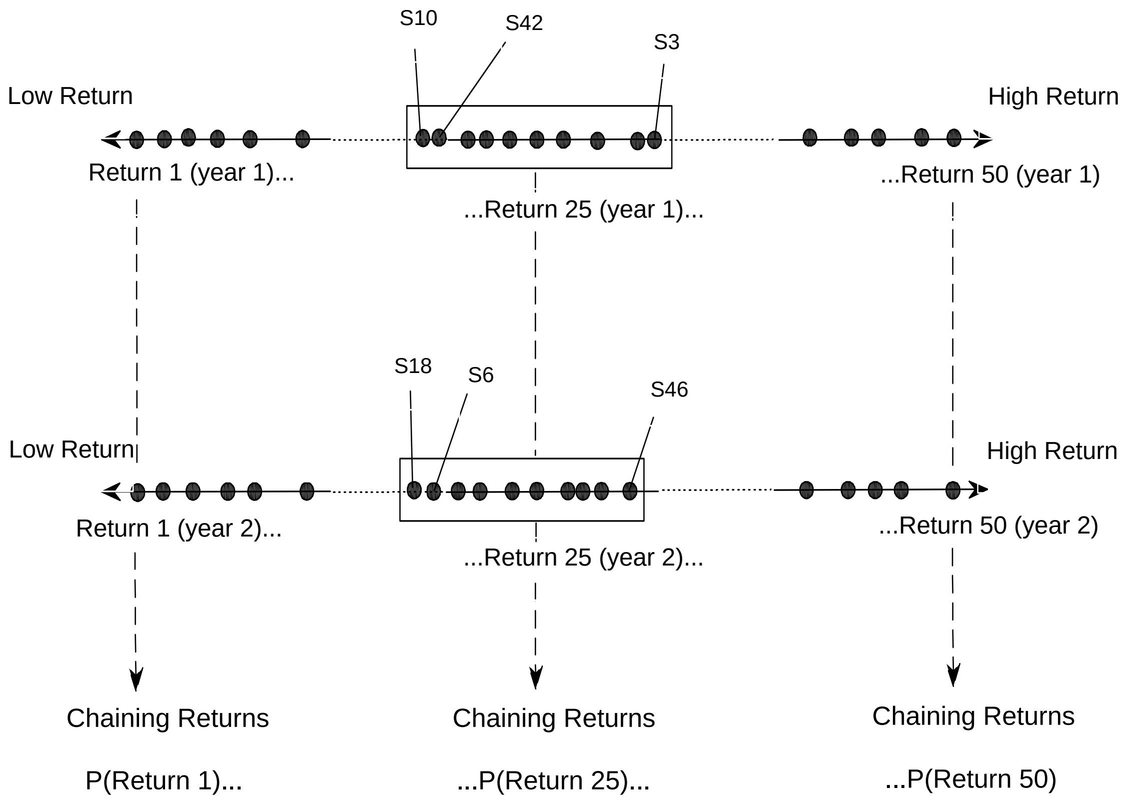
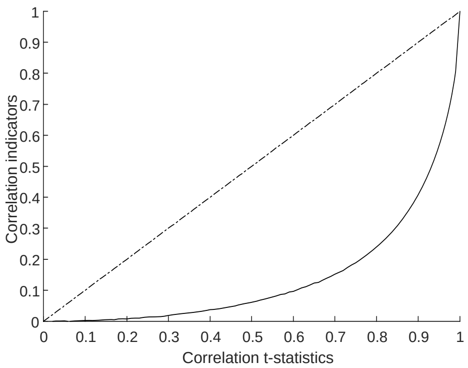


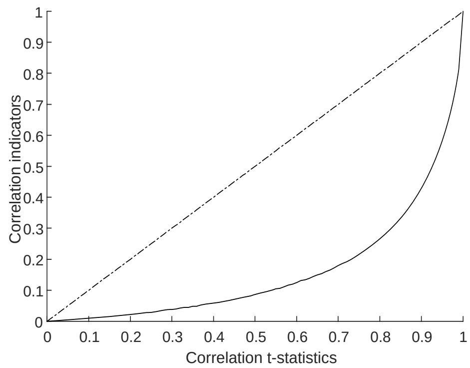
Figure A2

Overlapping versus Non-overlapping Portfolios

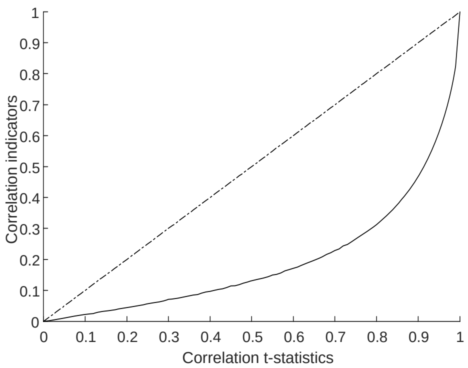
The four panels plot the pairwise correlation between the indicator functions $1(t_j^k)$ and $1(t_{j+d}^k)$ and the pairwise correlation between the t-statistics t_j^k and t_{j+d}^k for different mean values for the average t-statistic (μ_{ts}=0, 0.5, 1, 1.5). The interval I is set equal to [-0.4,0.4]. If the correlation between indicator functions is lower than the correlation between t-statistics, using overlapping instead of non-overlapping porfolios brings efficiency gains (i.e., it reduces the variance of the estimated mispricing proportion).



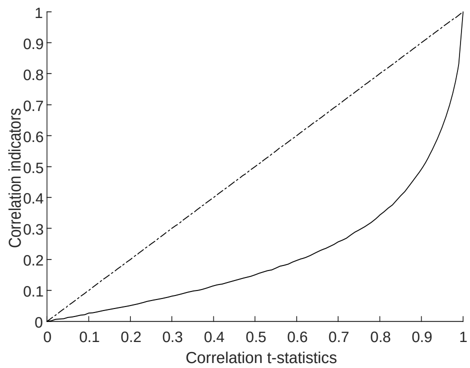
(a) t-statistic mean of 0



(b) t-statistic mean of 0.5



(c) t-statistic mean of 1



(d) t-statistic mean of 1.5