Looking for the Tangent Portfolio: Risk Optimization Techniques on Equity Style Buckets

Boris Fays\textsuperscript{1}  Marie Lambert\textsuperscript{1}  Nicolas Papageorgiou\textsuperscript{2}

\textsuperscript{1}HEC Liege, ULiège, Belgium
\textsuperscript{2}HEC Montréal

Research Seminar, HEC Montréal
February 21\textsuperscript{st}, 2019
AGENDA

THE PAPER IN A NUTSHELL

OUR CONTRIBUTION

DATA

EMPIRICAL RESULTS

CONCLUSION
RESEARCH OBJECTIVE

Looking for the tangent portfolio using risk-optimization techniques.

Our objective is to propose a simple intermediary method to proxy for •
**Methodology and Results**

- Stratification of the US equity universe (NYSE, AMEX, Nasdaq) into size and book-to-markets equity style buckets
  - Extension to momentum

- Risk-based investment strategies (MV, MD, RP) are shown to provide
  - Better pricing of characteristic-sorted portfolios than existing multifactor models
  - Higher Sharpe ratio than a portfolio made of:
    - market portfolio (Mkt)
    - 30-year US treasury bond (B30)
    - size (SMB) and value (HML) factors
**Motivation**

We ground our research into the following papers:

- Daniel et al. (2017, JF)
- Grinblatt and Saxena (2018, JFQA Forthcoming)
- Ao, Li, and Zheng (2018, RFS)

We rely on the following facts and evidence:

1. Caveats over the cap-weighted market benchmarks
2. Sharp rise in multi-factor models and in the number of index-funds and ETFs
3. Inefficiencies of long-short factors
4. Finding MSR is a noisy exercise
Motivation - Caveats over the Cap-Weighted Market Benchmarks

- “Market indices [...] are if anything inside that [mean-variance] frontier” (Cochrane 2001, Asset Pricing)
- “Cap-weighted stock portfolios are inefficient investments. [...] Even the most comprehensive cap-weighted portfolios occupy positions inside the efficient set” (Haugen and Baker 1991, JPM, p.35)

Based on data for the period 1979-1998. The efficient frontier assumes a perfect forecast of the future covariance matrix and of the future mean return. Figure taken from Schwartz(2000, Figure 3, p. 19).
Motivation - Sharp Rise in the number of multi-factor models

- From 50 significant characteristics (Subrahmanyam, 2010 EFM)
- To over 300!
  - 316 anomaly-based firm characteristics, see Harvey and Liu (2016, RFS)
  - 330 characteristics, see Green, Hand, and Zhang (2013, RAS)
  - +430 characteristics, see Hou, Xue, and Zhang (2018, WP)

Source: Harvey and Liu (2016, RFS)
Motivation - Sharp Rise in the number of ETFs versus listed stocks

The number of stocks reached 7,487 in 1995 but has fallen 42 percent.

From 2010 to 2012, the number of indexes quadrupled to 1,000.

Source: Bloomberg.com
MOTIVATION – INEFFICIENCIES OF LONG/SHORT FACTORS

Daniel et al. (2017, WP, p. 3):
“This set of portfolios will explain the returns of portfolios sorted on the same characteristics, but are unlikely to span the MVE portfolio of all assets, because they do not take into account the asset covariance structure.”

Grinblatt and Saxena (2018, JFQA Forthcoming, p. 5):
“The optimal combination of the factor mimicking portfolios has a significantly lower Sharpe ratio than the optimal combination of the basis portfolios they are created from.”
Motivation – Sample errors with MSR estimate

- Sample and specification errors
- Low-risk portfolios: giving up on estimating expected returns
- Robust variance-covariance matrix
FINDING A CANDIDATE FOR THE MVE PORTFOLIO

1. The opportunity sets: the DNS versus the original Fama-French sorting procedure
   ▶ Independent versus dependent (D) sorting
   ▶ NYSE breakpoints vs all names (N) breakpoints
   ▶ Double and triple sort (size, value and momentum): 2x3, 3x3 and 3x3x3 (Asymmetric versus Symmetric sort)
1. The opportunity sets

2. MSR weights replaced by smart beta (risk-based) optimization
   ▶ Minimum Variance (MV) (Clarke, Silva, and Thorley 2013, JPM)
   ▶ Maximum Diversification (MD) (Choueifaty and Coignard 2008, JPM)
   ▶ Risk parity (RP) (Maillard, Roncalli, and Teiletche 2010, JPM)

<table>
<thead>
<tr>
<th>Strategy</th>
<th>Objective Function</th>
<th>Constraints</th>
</tr>
</thead>
<tbody>
<tr>
<td>Minimum Variance (MV)</td>
<td>( \min f(w) = \sum_i^N \sum_j^N w_i \sigma_{ij} w_j )</td>
<td>( w_i \in [0, 1] ) and ( \sum_{i=1}^N w_i = 1 )</td>
</tr>
<tr>
<td>Maximum Diversification (MD)</td>
<td>( \max f(w) = \frac{\sum_i^N w_i \sigma_i}{\sqrt{\sum_i^N \sum_j^N w_i \sigma_{ij} w_j}} )</td>
<td></td>
</tr>
<tr>
<td>Risk parity (RP)</td>
<td>( \min f(w) = \sum_i^N \sum_j^N (w_i \times (\Sigma w)_i - w_j \times (\Sigma w)_j)^2 )</td>
<td></td>
</tr>
</tbody>
</table>
FINDING A CANDIDATE FOR THE MVE PORTFOLIO - THE OPPORTUNITY SETS

The DNS sorting procedure allows for:

▶ A better stratification of the US equity universe

![Growth MFs](chart1.png)
![Blend MFs](chart2.png)
![Value MFs](chart3.png)

- Portfolio-weighted average B/M score for each MF/month:
  - Growth MFs: 9.17%, 21.32%, 70.91%, 8.35%, 0.20%, 1.11%
  - Blend MFs: 0.52%, 55.01%, 49.42%, 49.42%, 0.31%, 0.32%
  - Value MFs: 0.10%, 16.64%, 78.45%, 4.81%, 0.18%, 10.13%
The DNS sorting procedure allows for:

- A better stratification of the US equity universe

![Graphs showing equity style box (Long) vs. Large Growth, Large Blend, Large Value](image-url)

Percent distribution and portfolio-weighted average B/M score for each MF/month.
Finding a candidate for the MVE portfolio - The opportunity sets

The DNS sorting procedure allows for:

- A better stratification of the US equity universe

![Equity Style Box (Long) - Mid Growth](chart1)

![Equity Style Box (Long) - Mid Blend](chart2)

![Equity Style Box (Long) - Mid Value](chart3)
The DNS sorting procedure allows for:

- A better stratification of the US equity universe
The DNS sorting procedure allows for:

- A better stratification of the US equity universe
- Better diversification

<table>
<thead>
<tr>
<th># Portfolios</th>
<th>Independent Sort (1)</th>
<th>Dependent Sort (2)</th>
<th>Difference (1)-(2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2x3</td>
<td>84.99</td>
<td>78.00</td>
<td>6.99</td>
</tr>
<tr>
<td>3x3</td>
<td>84.99</td>
<td>75.81</td>
<td>9.18</td>
</tr>
<tr>
<td>3x3x3</td>
<td>78.38</td>
<td>66.8</td>
<td>11.58</td>
</tr>
</tbody>
</table>
Finding a candidate for the MVE portfolio - The opportunity sets

The DNS sorting procedure allows for:

- A better stratification of the US equity universe
- Better diversification
- Similar to other portfolio sorts, a reduction of the complexity of the universe (consistent with the categorization process of Barberis and Shleifer (2003))
FINDING A CANDIDATE FOR THE MVE PORTFOLIO - MSR WEIGHTS REPLACED BY SMART BETA (RISK-BASED) OPTIMIZATION

- Long-only investment scheme
- Avoid the empirical challenge of estimating expected returns
US EQUITIES

We employ:

- Dataset from the merge of CRSP and Compustat.
- All stocks listed on NYSE, NASDAQ, and AMEX stocks and share code of 10 or 11.
- Sample period ranges from July 1963 to December 2015.
US EQUITIES

We employ:

- Dataset from the merge of CRSP and Compustat.
- All stocks listed on NYSE, NASDAQ, and AMEX stocks and share code of 10 or 11.
- Sample period ranges from July 1963 to December 2015.

Filtering criteria following Fama and French (1993, JF):

- Shares (SHROUT) and price (PRC)
- Stock return (RET) data for month t
- 2 years of listing on COMPUSTAT (survival bias)
US EQUITIES

We employ:

▶ Dataset from the merge of CRSP and Compustat.
▶ All stocks listed on NYSE, NASDAQ, and AMEX stocks and share code of 10 or 11.
▶ Sample period ranges from July 1963 to December 2015.

Filtering criteria following Fama and French (1993, JF):

▶ Shares (SHROUT) and price (PRC)
▶ Stock return (RET) data for month \( t \)
▶ 2 years of listing on COMPUSTAT (survival bias)

Characteristics:

▶ Market equity (firm size) as \( \text{SHARE} \times \text{PRICE} \)
▶ Book-to-market equity as \( \text{BE}/\text{ME} \)
▶ Momentum is the \( t-2 \) to \( t-12 \) cumulative return of stock
CHARACTERISTIC-SORTED PORTFOLIOS

Each year in June, we sort US stocks on the following traditional characteristics.

- size and value (2x3)
- size and value (3x3)
- size and value and momentum (3x3x3)

Average distribution of stock in portfolios
EMPIRICAL TESTS AND RESULTS IN A NUTSHELL

1. Smart investment strategies on DSN portfolios achieve better diversification return than other smart investment strategies and equally weighted scheme
   ▶ Diversification return framework of Booth and Fama (1992, FAJ) and Willenbrock (2011, FAJ)

2. Strategic beta portfolios constructed on dependent equity style buckets outperform a single-index model (using CW factor), a multi-factor model (FF-3 Factors) and other strategic beta portfolios
   ▶ Mean-variance spanning of Kan and Zhou (2012, AEF)
   ▶ Bootstrap procedure similar to Fama and French (2010, JF) and Harvey and Liu (2016, WP)
   ▶ Factor selection technique from Harvey and Liu (2016, WP)
Diversification return

Following Booth and Fama (1992, FAJ) and Willenbrock (2011, FAJ), the diversification return is given by,

\[
\text{DR}^{FW} = \mu_p - \sum_{i} w_i \mu_i + \frac{1}{2} \left( \sum_{i} w_i \sigma_i^2 - \sigma_p^2 \right)
\]

where

\[\text{DR}_1^{FW} = 0 \text{ if weights are constant} \quad \text{and} \quad \text{DR}_2^{FW} = \text{variance reduction benefit}\]

The relationship assumes that,

- weights \( w_i \) are held constant over the estimation period,
- \( i \) stands for the \( i^{th} \) security in the portfolio \( p \),
- \( FW \) denotes Fixed-Weight.
DIVERSIFICATION RETURN

Issues:

▶ Weights of the low risk strategies are not constant over time. For rebalancing strategies (non fixed weight), Erb and Harvey (2006, FAJ) use of the average of the weights over the sample period ($\bar{w}_i = \frac{1}{T} \sum_{1}^{T} w_i$).
**Diversification Return**

**Issues:**

- Weights of the low risk strategies are not constant over time. For rebalancing strategies (non fixed weight), Erb and Harvey (2006, FAJ) use of the average of the weights over the sample period ($\bar{w}_i = \frac{1}{T} \sum_{t=1}^{T} w_t$).

- The endogenous fixed weights benchmark used in the FW configuration differ sharply across the strategies.
DIVERSIFICATION RETURN

Issues:

- Weights of the low risk strategies are not constant over time. For rebalancing strategies (non fixed weight), Erb and Harvey (2006, FAJ) use of the average of the weights over the sample period ($\bar{w}_i = \frac{1}{T} \sum^T_1 w^t_i$).

- The endogenous fixed weights benchmark used in the FW configuration differ sharply across the strategies.

Proposition: diversification return with regard to an EW benchmark.

\[ DR^{EW} = \mu_p - \frac{1}{N} \sum^N_i \mu_i + \frac{1}{2} \left( \frac{1}{N} \sum^N_i \sigma^2_i - \sigma^2_p \right) \]

(2)
## Diversification return

We test the difference in the diversification components \( (DR_1, DR_2, \text{ and } DR) \) using the bootstrap method of Ledoit and Wolf (2008, JEF).

<table>
<thead>
<tr>
<th></th>
<th>( DR_1^{FW} )</th>
<th></th>
<th>( DR_2^{FW} )</th>
<th></th>
<th>( DR^{FW} )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Ind</td>
<td>Dep</td>
<td>( \Delta ) Dep-Ind</td>
<td>p-val\textsuperscript{b}</td>
<td>Ind</td>
</tr>
<tr>
<td>MD\textsubscript{2x3}</td>
<td>-0.005</td>
<td>-0.003</td>
<td>0.002</td>
<td>0.936</td>
<td>0.025</td>
</tr>
<tr>
<td>MD\textsubscript{3x3}</td>
<td>0.004</td>
<td>-0.032</td>
<td>-0.035</td>
<td>0.257</td>
<td>0.031</td>
</tr>
<tr>
<td>MD\textsubscript{3x3x3}</td>
<td>-0.041</td>
<td>-0.075</td>
<td>-0.034</td>
<td>0.428</td>
<td>0.046</td>
</tr>
<tr>
<td>MV\textsubscript{2x3}</td>
<td>0.013</td>
<td>-0.025</td>
<td>-0.038</td>
<td>0.554</td>
<td>0.024</td>
</tr>
<tr>
<td>MV\textsubscript{3x3}</td>
<td>-0.001</td>
<td>-0.036</td>
<td>-0.035</td>
<td>0.581</td>
<td>0.024</td>
</tr>
<tr>
<td>MV\textsubscript{3x3x3}</td>
<td>-0.018</td>
<td>-0.115</td>
<td>-0.097</td>
<td>0.093</td>
<td>0.047</td>
</tr>
<tr>
<td>RP\textsubscript{2x3}</td>
<td>0.005</td>
<td>0.005</td>
<td>0.000</td>
<td>0.992</td>
<td>0.024</td>
</tr>
<tr>
<td>RP\textsubscript{3x3}</td>
<td>0.005</td>
<td>0.000</td>
<td>-0.005</td>
<td>0.679</td>
<td>0.027</td>
</tr>
<tr>
<td>RP\textsubscript{3x3x3}</td>
<td>-0.001</td>
<td>-0.010</td>
<td>-0.009</td>
<td>0.522</td>
<td>0.043</td>
</tr>
</tbody>
</table>

*number of bootstraps=4999  
** figures are from gross return on a monthly basis (in %)  
*** Block size for bootstrap = 10
In this framework, results suggest that dependent-sorted portfolios provide significantly

- greater variance reduction benefits
- greater diversification return

### Equal-Weight (EW) Benchmark

<table>
<thead>
<tr>
<th></th>
<th>DR\textsubscript{1EW}</th>
<th></th>
<th></th>
<th></th>
<th>DR\textsubscript{2EW}</th>
<th></th>
<th></th>
<th></th>
<th>DR\textsubscript{EW}</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Ind</td>
<td>Dep</td>
<td>Δ</td>
<td>p-val</td>
<td>Ind</td>
<td>Dep</td>
<td>Δ</td>
<td>p-val</td>
<td>Ind</td>
<td>Dep</td>
<td>Δ</td>
</tr>
<tr>
<td>MD\textsubscript{2x3}</td>
<td>0.005</td>
<td>0.088</td>
<td>0.083</td>
<td>0.012</td>
<td>0.027</td>
<td>0.038</td>
<td>0.011</td>
<td>0.003</td>
<td>0.032</td>
<td>0.126</td>
<td>0.094</td>
</tr>
<tr>
<td>MD\textsubscript{3x3}</td>
<td>0.010</td>
<td>0.070</td>
<td>0.060</td>
<td>0.125</td>
<td>0.033</td>
<td>0.048</td>
<td>0.015</td>
<td>0.000</td>
<td>0.043</td>
<td>0.118</td>
<td>0.075</td>
</tr>
<tr>
<td>MD\textsubscript{3x3x3}</td>
<td>-0.034</td>
<td>0.063</td>
<td>0.097</td>
<td>0.107</td>
<td>0.046</td>
<td>0.061</td>
<td>0.014</td>
<td>0.010</td>
<td>0.012</td>
<td>0.124</td>
<td>0.112</td>
</tr>
<tr>
<td>MV\textsubscript{2x3}</td>
<td>0.082</td>
<td>0.218</td>
<td>0.136</td>
<td>0.071</td>
<td>0.033</td>
<td>0.028</td>
<td>-0.004</td>
<td>0.407</td>
<td>0.115</td>
<td>0.246</td>
<td>0.131</td>
</tr>
<tr>
<td>MV\textsubscript{3x3}</td>
<td>0.050</td>
<td>0.143</td>
<td>0.093</td>
<td>0.163</td>
<td>0.034</td>
<td>0.046</td>
<td>0.012</td>
<td>0.066</td>
<td>0.084</td>
<td>0.189</td>
<td>0.105</td>
</tr>
<tr>
<td>MV\textsubscript{3x3x3}</td>
<td>0.022</td>
<td>0.019</td>
<td>-0.003</td>
<td>0.955</td>
<td>0.061</td>
<td>0.080</td>
<td>0.018</td>
<td>0.002</td>
<td>0.084</td>
<td>0.099</td>
<td>0.015</td>
</tr>
<tr>
<td>RP\textsubscript{2x3}</td>
<td>0.020</td>
<td>0.056</td>
<td>0.036</td>
<td>0.013</td>
<td>0.026</td>
<td>0.036</td>
<td>0.010</td>
<td>0.001</td>
<td>0.046</td>
<td>0.092</td>
<td>0.047</td>
</tr>
<tr>
<td>RP\textsubscript{3x3}</td>
<td>0.019</td>
<td>0.048</td>
<td>0.029</td>
<td>0.063</td>
<td>0.029</td>
<td>0.042</td>
<td>0.013</td>
<td>0.000</td>
<td>0.048</td>
<td>0.090</td>
<td>0.041</td>
</tr>
<tr>
<td>RP\textsubscript{3x3x3}</td>
<td>0.011</td>
<td>0.035</td>
<td>0.024</td>
<td>0.102</td>
<td>0.046</td>
<td>0.068</td>
<td>0.022</td>
<td>0.000</td>
<td>0.057</td>
<td>0.102</td>
<td>0.046</td>
</tr>
</tbody>
</table>

*number of bootstraps=4999
** figures are from gross return on a monthly basis (in %)
*** Block size for bootstrap = 10

Bootstrap
TEST OF MEAN-VARIANCE SPANNING

Illustration of Kan and Zhou (2012, AEF) mean-variance spanning test:

$H_0^1$: Tangency Portfolio

Benchmark Assets ($R_1$):

- US Bond
- Portfolio A
**Test of Mean-Variance Spanning**

Illustration of Kan and Zhou (2012, AEF) test of Mean-Variance spanning:

- **$H_0^1$: Tangency Portfolio**
- **Benchmark Assets ($R_1$):**
  - US Bond
  - Portfolio A
- **Test Asset ($R_2$):**
  - Portfolio B
**Test of Mean-Variance Spanning**

Illustration of Kan and Zhou (2012, AEF) test of Mean-Variance spanning:

\[ H_0^2: \text{GMV Portfolio} \]

**Benchmark Assets** \((R_1):\)

- US Bond
- Portfolio A
TEST OF MEAN-VARIANCE SPANNING

Illustration of Kan and Zhou (2012, AEF) test of mean-variance spanning:

$H_0^2$: GMV Portfolio

Benchmark Assets ($R_1$):
- US Bond
- Portfolio A

Test Asset ($R_2$):
- Portfolio B
**Test of Mean-Variance Spanning**

Huberman and Kandel (1987, JF) define the following regression test:

\[
R_t^2 = \alpha + \beta R_t^1 + e^t
\]  

(3)

The null hypothesis \( H_0 \) sets \( \alpha = 0 \) and \( \delta = 1 - \beta = 0 \).

Considering an efficient frontier comprising \( K + N \) assets, the weights of the \( N \) assets into the tangent (\( Qw_1 \)) and GMV (\( Qw_2 \)) portfolios are defined as:

\[
Qw_1 = \frac{QV^{-1}\mu^{1}}{1'_{N+K}V^{-1}\mu^{1}} = \frac{\Sigma^{-1}\alpha}{1'_{N+K}V^{-1}\mu^{1}}
\]

(4)

\[
Qw_2 = \frac{QV^{-1}1_{N+K}}{1'_{N+K}V^{-1}1_{N+K}} = \frac{\Sigma^{-1}\delta}{1'_{N+K}V^{-1}1_{N+K}}
\]

where \( Q = [0_{N \times K}, I_N] \), \( I_N \) is an \( N \times N \) identity matrix, \( \Sigma = V_{22} - V_{21}V_{11}^{-1}V_{12} \), and \( V \) is the variance-covariance matrix of the \( K \) benchmark assets (\( R_1 \)) plus the \( N \) test assets (\( R_2 \)) such that,

\[
V = \text{Var}[R_1, R_2] = \begin{bmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{bmatrix}
\]
**Test of Mean-Variance Spanning**

Step-down procedure to test the spanning hypothesis (Kan and Zhou 2012, AEF):

\[ H_0^1 = \alpha = 0_N, \text{ such that } Qw_1 = 0. \]
**TEST OF MEAN-VARIANCE SPANNING**

Step-down procedure to test the spanning hypothesis (Kan and Zhou 2012, AEF):

\[ H_0^1 = \alpha = 0_N, \text{ such that } Qw_1 = 0. \]

The \( F \)-test (\( H_0^1 \)):

\[ F_1 = \frac{T - K - N}{N} \left( \frac{\hat{a} - \hat{a}_1}{1 + \hat{a}_1} \right) \]

- \( T \) is the number of observations
- \( K \) is the number of benchmark assets
- \( N \) is the number of test assets
- \( \hat{a}_1 = \hat{\mu}' \hat{V}^{-1} \hat{\mu}_1 \)
- \( \hat{V}_{11} \): the variance of the benchmark assets
- \( \hat{\mu}_1 \): the vector of mean return of the benchmark assets
- \( \hat{a} \) but refers to the benchmark assets \( (R_1) \) plus the new test asset \( (R_2) \)
TEST OF MEAN-VARIANCE SPANNING

Step-down procedure to test the spanning hypothesis (Kan and Zhou 2012, AEF):

\[ H^2_0 : \delta = 1_N - \beta_1 K = 0_N | \alpha = 0_N, \text{ such that } Qw_2 = 0 \text{ conditional on } Qw_1 = 0. \]

The F-test \( H^2_0 \):

\[
F_2 = \frac{T - K - N + 1}{N} \left[ \frac{\hat{c} + \hat{d}}{\hat{c}_1 + \hat{a}_1} \frac{1 + \hat{a}_1}{1 + \hat{a}} - 1 \right]
\]

\[ \hat{a}_1 = \hat{\mu}_1 \hat{V}_{11}^{-1} \hat{\mu}_1 \]
\[ \hat{b}_1 = \hat{\mu}_1 \hat{V}_{11}^{-1} 1_K \]
\[ \hat{c}_1 = 1_K \hat{V}_{11}^{-1} 1_K \]
\[ \hat{d}_1 = \hat{a}_1 \hat{c}_1 - \hat{b}_1^2 \]
\[ \hat{V}_{11} : \text{ the variance of the benchmark assets} \]
\[ \hat{\mu}_1 : \text{ the vector of mean return of the benchmark assets} \]
\[ \hat{a}, \hat{b}, \hat{c} \text{ and } \hat{d} \text{ refers to the benchmark assets (} R_1 \text{) plus the new test asset (} R_2 \text{)} \]
BOOTSTRAP APPROACH ON TEST OF MEAN-VARIANCE SPANNING

The null hypothesis should be true in-sample (Harvey and Liu (2016, WP) and White (2000, ECO)):

$$ \rightarrow Qw_1 = 0 \text{ and } Qw_2 = 0. $$
The null hypothesis should be true in-sample (Harvey and Liu (2016, WP) and White (2000, ECO)):

\[ Qw_1 = 0 \quad \text{and} \quad Qw_2 = 0. \]

**Step 1: Orthogonalization under the null**

\[ R_t^2 = \alpha + \beta R_t^1 + R_t^{2,e} \]  \hspace{1cm} (5)

- \( R^{orth}_2 = R^{t,e}_2 + R^{t}_{1,MVE} \) by construction \( \alpha = 0 \) and \( \beta_{MVE} = 1 \rightarrow Qw_1 = 0 \) and \( Qw_2 = 0 \)
- \( R_{1,MVE} \) is the proxy for the market portfolio present in \( R_1 \)
Bootstrap Approach on Test of Mean-Variance Spanning

The null hypothesis should be true in-sample (Harvey and Liu (2016, WP) and White (2000, ECO)):

\[
\rightarrow Qw_1 = 0 \text{ and } Qw_2 = 0.
\]

**Step 1: Orthogonalization under the null**

\[
R^t_2 = \alpha + \beta R^t_1 + R^{t,e}_2
\]

- \(R^{orth}_2 = R^{t,e}_2 + R^{t,MVE}_1\) by construction \(\alpha = 0\) and \(\beta_{MVE} = 1\) → \(Qw_1 = 0\) and \(Qw_2 = 0\)
- \(R^{t,MVE}_1\) is the proxy for the market portfolio present in \(R_1\)

**Step 2: Bootstrap (Harvey and Liu 2016, WP)**

- preserves the cross-sectional correlations among the benchmark \((R_1)\) and test \((R^{orth}_2)\) assets
- preserves the uncertainty of the time-series: bootstrap sampling length=original time-series length
**Bootstrap Approach on Test of Mean-Variance Spanning**

**Step 3: Test of Mean-Variance Spanning**

- Apply the test of mean-variance spanning from Kan and Zhou (2012, AEF)

- Outputs:
  - Tangency portfolio: range of $F_{1,\text{ind}}^b$ and $F_{1,\text{dep}}^b$ with $\{b = 1, 2, \ldots, B\}$
  - GMV portfolio: range of $F_{2,\text{ind}}^b$ and $F_{2,\text{dep}}^b$ with $\{b = 1, 2, \ldots, B\}$
  - Where,

\[
F_1 = \frac{T - K - N}{N} \frac{\hat{a} - \hat{a}_1}{1 + \hat{a}_1}
\]

\[
F_2 = \frac{T - K - N + 1}{N} \left[ \frac{\hat{c} + \hat{d}}{\hat{c}_1 + \hat{d}_1} \frac{1 + \hat{a}_1}{1 + \hat{a}} - 1 \right]
\]
Bootstrap Approach on Test of Mean-Variance Spanning

Step 3: Mean-Variance Spanning Test (cont’d)

- Harvey and Liu (2016, WP)’s bootstrap approach robust for multiple testing
  - Conservative reference point:
    - $F^b_1 = \max(F^b_{1,ind}, F^b_{1,dep})$ with $\{b = 1, 2, \ldots, B\}$
    - $F^b_2 = \max(F^b_{2,ind}, F^b_{2,dep})$ with $\{b = 1, 2, \ldots, B\}$

- Bootstrap p-value for the F-tests

\[
\begin{align*}
  F^0_1 \text{ p-value:} & \\
  p-val_{ind}^b &= \frac{\#\{F^b_1 > F^0_{1,ind}\}}{B} \\
  p-val_{dep}^b &= \frac{\#\{F^b_1 > F^0_{1,dep}\}}{B} \\
  F^0_2 \text{ p-value:} & \\
  p-val_{ind}^b &= \frac{\#\{F^b_2 > F^0_{1,ind}\}}{B} \\
  p-val_{dep}^b &= \frac{\#\{F^b_2 > F^0_{1,dep}\}}{B}
\end{align*}
\]
SPANNING OF SINGLE INDEX AND MULTIFACTOR MODELS

Sample period: July 1963 - December 2015

- Benchmark assets \( (R_1) = \text{CW-Market Portfolio (Mkt)} + \text{30-Year US Treasury Bond (B30)} \)
- Test asset \( (R_2) = \text{Smart Beta (SB)} \)

\[
R_t^2 = \alpha + \beta_1 B30_t + \beta_2 Mkt_t + R_t^{*,e}
\]

\[
R_2^{orth} = R_2^{*,e} + Mkt_t
\]

<table>
<thead>
<tr>
<th>MVE candidates</th>
<th>( \alpha_{ind} )</th>
<th>( F_{1,ind} )</th>
<th>p-val(^b)</th>
<th>( F_{2,ind} )</th>
<th>p-val(^b)</th>
<th>( \alpha_{dep} )</th>
<th>( F_{1,dep} )</th>
<th>p-val(^b)</th>
<th>( F_{2,dep} )</th>
<th>p-val(^b)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MD(_{2x3})</td>
<td>0.0020</td>
<td>11.69</td>
<td>0.00</td>
<td>2.37</td>
<td>0.31</td>
<td>0.0031</td>
<td>13.00</td>
<td>0.00</td>
<td>13.52</td>
<td>0.01</td>
</tr>
<tr>
<td>MD(_{3x3})</td>
<td>0.0022</td>
<td>11.89</td>
<td>0.00</td>
<td>2.96</td>
<td>0.25</td>
<td>0.0034</td>
<td>10.66</td>
<td>0.00</td>
<td>21.30</td>
<td>0.00</td>
</tr>
<tr>
<td>MD(_{3x3x3})</td>
<td>0.0019</td>
<td>7.21</td>
<td>0.01</td>
<td>1.42</td>
<td>0.58</td>
<td>0.0038</td>
<td>9.45</td>
<td>0.00</td>
<td>19.39</td>
<td>0.00</td>
</tr>
<tr>
<td>MV(_{2x3})</td>
<td>0.0032</td>
<td>13.63</td>
<td>0.00</td>
<td>11.83</td>
<td>0.01</td>
<td>0.0049</td>
<td>15.55</td>
<td>0.00</td>
<td>23.12</td>
<td>0.00</td>
</tr>
<tr>
<td>MV(_{3x3})</td>
<td>0.0029</td>
<td>10.60</td>
<td>0.00</td>
<td>6.68</td>
<td>0.07</td>
<td>0.0044</td>
<td>13.39</td>
<td>0.00</td>
<td>28.21</td>
<td>0.00</td>
</tr>
<tr>
<td>MV(_{3x3x3})</td>
<td>0.0029</td>
<td>13.60</td>
<td>0.00</td>
<td>10.24</td>
<td>0.02</td>
<td>0.0034</td>
<td>10.02</td>
<td>0.00</td>
<td>19.87</td>
<td>0.00</td>
</tr>
<tr>
<td>RP(_{2x3})</td>
<td>0.0021</td>
<td>13.03</td>
<td>0.00</td>
<td>1.62</td>
<td>0.42</td>
<td>0.0027</td>
<td>10.60</td>
<td>0.00</td>
<td>7.48</td>
<td>0.04</td>
</tr>
<tr>
<td>RP(_{3x3})</td>
<td>0.0022</td>
<td>11.57</td>
<td>0.00</td>
<td>0.49</td>
<td>0.72</td>
<td>0.0029</td>
<td>8.70</td>
<td>0.00</td>
<td>8.49</td>
<td>0.02</td>
</tr>
<tr>
<td>RP(_{3x3x3})</td>
<td>0.0023</td>
<td>12.09</td>
<td>0.00</td>
<td>0.25</td>
<td>0.84</td>
<td>0.0031</td>
<td>10.28</td>
<td>0.00</td>
<td>7.34</td>
<td>0.04</td>
</tr>
</tbody>
</table>

* Mkt, B30 and SB are taken in excess of the risk-free rate (one-month T-bill from Ibbotson)
SPANNING OF SINGLE INDEX AND MULTIFACTOR MODELS

Sample period: July 1963 - December 2015

- Benchmark assets ($R_1$) = CW-Market Portfolio (Mkt) + 30-Year US Treasury Bond (B30) + SMB + HML
- Test asset ($R_2$) = Smart Beta (SB)

$$R_t^2 = \alpha + \beta_1 B30_t + \beta_2 Mkt_t + \beta_3 SMB_t + \beta_4 HML_t + R_t^{e}$$

- $R_2^{orth} = R_2^{e} + Mkt_t$

<table>
<thead>
<tr>
<th>MVE candidates</th>
<th>$\alpha_{ind}$</th>
<th>$F_{1,ind}$</th>
<th>p-val$^b$</th>
<th>$F_{2,ind}$</th>
<th>p-val$^b$</th>
<th>$\alpha_{dep}$</th>
<th>$F_{1,dep}$</th>
<th>p-val$^b$</th>
<th>$F_{2,dep}$</th>
<th>p-val$^b$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Panel B: $R_1 = MKT + B30 + SMB + HML$, $R_2 = SB, R_2^{orth} = SB^e + Mkt$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MD$_{2\times3}$</td>
<td>0.0001</td>
<td>0.21</td>
<td>0.85</td>
<td>1419.96</td>
<td>0.00</td>
<td>0.00011</td>
<td>4.89</td>
<td>0.05</td>
<td>578.95</td>
<td>0.00</td>
</tr>
<tr>
<td>MD$_{3\times3}$</td>
<td>0.0002</td>
<td>0.66</td>
<td>0.63</td>
<td>1392.77</td>
<td>0.00</td>
<td>0.00011</td>
<td>3.13</td>
<td>0.14</td>
<td>365.03</td>
<td>0.00</td>
</tr>
<tr>
<td>MD$_{3\times3\times3}$</td>
<td>-0.0003</td>
<td>0.65</td>
<td>0.63</td>
<td>824.06</td>
<td>0.00</td>
<td>0.00012</td>
<td>2.22</td>
<td>0.22</td>
<td>285.69</td>
<td>0.00</td>
</tr>
<tr>
<td>MV$_{2\times3}$</td>
<td>0.0009</td>
<td>2.22</td>
<td>0.22</td>
<td>317.78</td>
<td>0.00</td>
<td>0.00022</td>
<td>6.21</td>
<td>0.02</td>
<td>249.22</td>
<td>0.00</td>
</tr>
<tr>
<td>MV$_{3\times3}$</td>
<td>0.0006</td>
<td>1.02</td>
<td>0.47</td>
<td>438.93</td>
<td>0.00</td>
<td>0.00018</td>
<td>4.59</td>
<td>0.06</td>
<td>254.99</td>
<td>0.00</td>
</tr>
<tr>
<td>MV$_{3\times3\times3}$</td>
<td>0.0007</td>
<td>2.22</td>
<td>0.23</td>
<td>564.35</td>
<td>0.00</td>
<td>0.00010</td>
<td>1.92</td>
<td>0.28</td>
<td>310.91</td>
<td>0.00</td>
</tr>
<tr>
<td>RP$_{2\times3}$</td>
<td>0.0002</td>
<td>0.92</td>
<td>0.47</td>
<td>1786.24</td>
<td>0.00</td>
<td>0.00007</td>
<td>3.21</td>
<td>0.11</td>
<td>882.21</td>
<td>0.00</td>
</tr>
<tr>
<td>RP$_{3\times3}$</td>
<td>0.0002</td>
<td>0.88</td>
<td>0.52</td>
<td>2015.93</td>
<td>0.00</td>
<td>0.00007</td>
<td>2.36</td>
<td>0.20</td>
<td>801.67</td>
<td>0.00</td>
</tr>
<tr>
<td>RP$_{3\times3\times3}$</td>
<td>0.0003</td>
<td>1.16</td>
<td>0.44</td>
<td>1895.22</td>
<td>0.00</td>
<td>0.00010</td>
<td>4.27</td>
<td>0.06</td>
<td>811.01</td>
<td>0.00</td>
</tr>
</tbody>
</table>

* Mkt, B30 and SB are taken in excess of the risk-free rate (one-month T-bill from Ibbotson)
SPANNING OF SINGLE INDEX AND MULTIFACTOR MODELS

Sample period: July 1963 - December 2015

- Benchmark assets ($R_1$) = Smart Beta (SB) + 30-Year US Treasury Bond (B30)
- Test asset ($R_2$) = CW-Market Portfolio (Mkt)

\[ R_t^* = \alpha + \beta_1 B30_t^r + \beta_2 SB_t^r + R_t^{r,e} \]

- $R_2^{orth} = R_2^{r,e} + SB_t$

<table>
<thead>
<tr>
<th>MVE candidates</th>
<th>$\alpha_{ind}$</th>
<th>$F_{1,ind}$</th>
<th>p-val$^b$</th>
<th>$F_{2,ind}$</th>
<th>p-val$^b$</th>
<th>$\alpha_{dep}$</th>
<th>$F_{1,dep}$</th>
<th>p-val$^b$</th>
<th>$F_{2,dep}$</th>
<th>p-val$^b$</th>
</tr>
</thead>
<tbody>
<tr>
<td>MD$_{2\times3}$</td>
<td>-0.0013</td>
<td>5.92</td>
<td>0.01</td>
<td>7.93</td>
<td>0.03</td>
<td>-0.0016</td>
<td>4.20</td>
<td>0.05</td>
<td>8.11</td>
<td>0.03</td>
</tr>
<tr>
<td>MD$_{3\times3}$</td>
<td>-0.0014</td>
<td>5.51</td>
<td>0.02</td>
<td>9.40</td>
<td>0.02</td>
<td>-0.0013</td>
<td>1.99</td>
<td>0.22</td>
<td>10.40</td>
<td>0.01</td>
</tr>
<tr>
<td>MD$_{3\times3x3}$</td>
<td>-0.0010</td>
<td>2.22</td>
<td>0.18</td>
<td>17.25</td>
<td>0.01</td>
<td>-0.0010</td>
<td>0.92</td>
<td>0.49</td>
<td>19.97</td>
<td>0.01</td>
</tr>
<tr>
<td>MV$_{2\times3}$</td>
<td>-0.0017</td>
<td>3.97</td>
<td>0.05</td>
<td>12.51</td>
<td>0.01</td>
<td>-0.0016</td>
<td>2.31</td>
<td>0.17</td>
<td>21.01</td>
<td>0.00</td>
</tr>
<tr>
<td>MV$_{3\times3}$</td>
<td>-0.0014</td>
<td>2.63</td>
<td>0.13</td>
<td>18.31</td>
<td>0.00</td>
<td>-0.0014</td>
<td>1.87</td>
<td>0.22</td>
<td>15.30</td>
<td>0.00</td>
</tr>
<tr>
<td>MV$_{3\times3x3}$</td>
<td>-0.0017</td>
<td>4.60</td>
<td>0.05</td>
<td>10.31</td>
<td>0.01</td>
<td>-0.0012</td>
<td>1.45</td>
<td>0.30</td>
<td>14.54</td>
<td>0.00</td>
</tr>
<tr>
<td>RP$_{2\times3}$</td>
<td>-0.0014</td>
<td>6.82</td>
<td>0.01</td>
<td>9.52</td>
<td>0.02</td>
<td>-0.0014</td>
<td>3.44</td>
<td>0.07</td>
<td>10.80</td>
<td>0.01</td>
</tr>
<tr>
<td>RP$_{3\times3}$</td>
<td>-0.0014</td>
<td>5.35</td>
<td>0.03</td>
<td>16.33</td>
<td>0.00</td>
<td>-0.0011</td>
<td>1.81</td>
<td>0.21</td>
<td>16.07</td>
<td>0.00</td>
</tr>
<tr>
<td>RP$_{3\times3x3}$</td>
<td>-0.0014</td>
<td>5.50</td>
<td>0.02</td>
<td>19.14</td>
<td>0.00</td>
<td>-0.0013</td>
<td>2.46</td>
<td>0.14</td>
<td>17.64</td>
<td>0.00</td>
</tr>
</tbody>
</table>

* Mkt, B30 and SB are taken in excess of the risk-free rate (one-month T-bill from Ibbotson)
**SPANNING OF SINGLE INDEX AND MULTIFACTOR MODELS**

**Sample period:** July 1993 - December 2015

- Benchmark assets \( (R_1) = \text{CW-Market Portfolio (Mkt)} + 30\text{-Year US Treasury Bond (B30)} \)
- Test asset \( (R_2) = \text{Smart Beta (SB)} \)

\[
R_t^2 = \alpha + \beta_1 B30^t + \beta_2 Mkt^t + R^t,\, e
\]

- \( R^\text{orth}_2 = R^t,\, e + Mkt^t \)

<table>
<thead>
<tr>
<th>MVE candidates</th>
<th>( \alpha_{\text{ind}} )</th>
<th>( F_{1,\text{ind}} )</th>
<th>p-val( ^b )</th>
<th>( F_{2,\text{ind}} )</th>
<th>p-val( ^b )</th>
<th>( \alpha_{\text{dep}} )</th>
<th>( F_{1,\text{dep}} )</th>
<th>p-val( ^b )</th>
<th>( F_{2,\text{dep}} )</th>
<th>p-val( ^b )</th>
</tr>
</thead>
<tbody>
<tr>
<td>MD(_{2\times3})</td>
<td>0.0019</td>
<td>3.77</td>
<td>0.11</td>
<td>6.48</td>
<td>0.09</td>
<td>0.0048</td>
<td>12.70</td>
<td>0.00</td>
<td>16.12</td>
<td>0.00</td>
</tr>
<tr>
<td>MD(_{3\times3})</td>
<td>0.0024</td>
<td>5.08</td>
<td><strong>0.06</strong></td>
<td>7.26</td>
<td>0.05</td>
<td>0.0049</td>
<td>9.14</td>
<td><strong>0.01</strong></td>
<td>16.55</td>
<td>0.00</td>
</tr>
<tr>
<td>MD(_{3\times3}\times3)</td>
<td>0.0016</td>
<td>1.86</td>
<td>0.33</td>
<td>5.87</td>
<td>0.17</td>
<td>0.0053</td>
<td>9.38</td>
<td><strong>0.01</strong></td>
<td>16.66</td>
<td>0.02</td>
</tr>
<tr>
<td>MV(_{2\times3})</td>
<td>0.0039</td>
<td>9.03</td>
<td><strong>0.00</strong></td>
<td>19.93</td>
<td>0.00</td>
<td>0.0071</td>
<td>14.13</td>
<td><strong>0.00</strong></td>
<td>17.80</td>
<td>0.00</td>
</tr>
<tr>
<td>MV(_{3\times3})</td>
<td>0.0038</td>
<td>7.55</td>
<td><strong>0.02</strong></td>
<td>11.39</td>
<td>0.01</td>
<td>0.0066</td>
<td>13.10</td>
<td><strong>0.00</strong></td>
<td>21.37</td>
<td>0.00</td>
</tr>
<tr>
<td>MV(_{3\times3}\times3)</td>
<td>0.0042</td>
<td>11.36</td>
<td><strong>0.00</strong></td>
<td>18.15</td>
<td>0.00</td>
<td>0.0050</td>
<td>9.92</td>
<td><strong>0.00</strong></td>
<td>19.77</td>
<td>0.00</td>
</tr>
<tr>
<td>RP(_{2\times3})</td>
<td>0.0021</td>
<td>4.73</td>
<td><strong>0.07</strong></td>
<td>4.94</td>
<td>0.12</td>
<td>0.0039</td>
<td>9.39</td>
<td><strong>0.01</strong></td>
<td>9.07</td>
<td>0.03</td>
</tr>
<tr>
<td>RP(_{3\times3})</td>
<td>0.0024</td>
<td>5.02</td>
<td><strong>0.05</strong></td>
<td>3.44</td>
<td>0.21</td>
<td>0.0039</td>
<td>6.64</td>
<td><strong>0.02</strong></td>
<td>7.98</td>
<td>0.05</td>
</tr>
<tr>
<td>RP(_{3\times3}\times3)</td>
<td>0.0027</td>
<td>6.19</td>
<td><strong>0.03</strong></td>
<td>4.37</td>
<td>0.13</td>
<td>0.0043</td>
<td>8.69</td>
<td><strong>0.01</strong></td>
<td>9.42</td>
<td>0.02</td>
</tr>
</tbody>
</table>

* Mkt, B30 and SB are taken in excess of the risk-free rate (one-month T-bill from Ibbotson)
**SPANNING OF SINGLE INDEX AND MULTIFACTOR MODELS**

**Sample period:** July 1993 - December 2015

- Benchmark assets ($R_1$) = CW-Market Portfolio (Mkt) + 30-Year US Treasury Bond (B30) + SMB + HML
- MVE market portfolio proxy is Mkt ($R_{1,MVE}$)
- Test asset ($R_2$) = Smart Beta (SB)

\[
R_t^2 = \alpha + \beta_1 B30_t + \beta_2 Mkt_t + \beta_3 SMB_t + \beta_4 HML_t + R_{t,e}^t
\]

- $R_{2,orth} = R_{2,e}^t + Mkt^t$

<table>
<thead>
<tr>
<th>MVE candidates</th>
<th>$\alpha_{ind}$</th>
<th>$F_{1,ind}$</th>
<th>p-val$^b$</th>
<th>$F_{2,ind}$</th>
<th>p-val$^b$</th>
<th>$\alpha_{dep}$</th>
<th>$F_{1,dep}$</th>
<th>p-val$^b$</th>
<th>$F_{2,dep}$</th>
<th>p-val$^b$</th>
</tr>
</thead>
<tbody>
<tr>
<td>MD$_{2x3}$</td>
<td>0.0006</td>
<td>1.52</td>
<td>0.40</td>
<td>384.37</td>
<td>0.00</td>
<td>0.0036</td>
<td>16.78</td>
<td>0.00</td>
<td>118.87</td>
<td>0.00</td>
</tr>
<tr>
<td>MD$_{3x3}$</td>
<td>0.0010</td>
<td>3.71</td>
<td>0.11</td>
<td>382.82</td>
<td>0.00</td>
<td>0.0036</td>
<td>9.31</td>
<td>0.00</td>
<td>63.08</td>
<td>0.00</td>
</tr>
<tr>
<td>MD$_{3x3x3}$</td>
<td>0.0001</td>
<td>0.02</td>
<td>0.99</td>
<td>249.23</td>
<td>0.00</td>
<td>0.0041</td>
<td>8.88</td>
<td>0.01</td>
<td>39.76</td>
<td>0.00</td>
</tr>
<tr>
<td>MV$_{2x3}$</td>
<td>0.0027</td>
<td>6.52</td>
<td>0.03</td>
<td>42.73</td>
<td>0.00</td>
<td>0.0056</td>
<td>14.39</td>
<td>0.00</td>
<td>44.38</td>
<td>0.00</td>
</tr>
<tr>
<td>MV$_{3x3}$</td>
<td>0.0025</td>
<td>5.89</td>
<td>0.02</td>
<td>92.31</td>
<td>0.00</td>
<td>0.0052</td>
<td>13.19</td>
<td>0.00</td>
<td>43.67</td>
<td>0.00</td>
</tr>
<tr>
<td>MV$_{3x3x3}$</td>
<td>0.0028</td>
<td>11.43</td>
<td>0.00</td>
<td>124.93</td>
<td>0.00</td>
<td>0.0038</td>
<td>9.86</td>
<td>0.00</td>
<td>51.57</td>
<td>0.00</td>
</tr>
<tr>
<td>RP$_{2x3}$</td>
<td>0.0008</td>
<td>3.50</td>
<td>0.11</td>
<td>563.63</td>
<td>0.00</td>
<td>0.0026</td>
<td>14.80</td>
<td>0.00</td>
<td>217.80</td>
<td>0.00</td>
</tr>
<tr>
<td>RP$_{3x3}$</td>
<td>0.0010</td>
<td>5.07</td>
<td>0.05</td>
<td>652.92</td>
<td>0.00</td>
<td>0.0026</td>
<td>9.23</td>
<td>0.00</td>
<td>174.22</td>
<td>0.00</td>
</tr>
<tr>
<td>RP$_{3x3x3}$</td>
<td>0.0013</td>
<td>7.72</td>
<td>0.02</td>
<td>618.11</td>
<td>0.00</td>
<td>0.0031</td>
<td>13.91</td>
<td>0.00</td>
<td>158.94</td>
<td>0.00</td>
</tr>
</tbody>
</table>

* Mkt, B30 and SB are taken in excess of the risk-free rate (one-month T-bill from Ibbotson)
HORSE RACE BETWEEN SMART INVESTMENT PORTFOLIOS

- Panel A: MVE market portfolio proxy is $SB_{dep}^{net} (R_1,MVE)$
- Panel B: MVE market portfolio proxy is $SB_{ind}^{net} (R_1,MVE)$

<table>
<thead>
<tr>
<th></th>
<th>Panel A:</th>
<th></th>
<th>Panel B:</th>
<th></th>
<th>MVE Candidate</th>
<th>GMV Candidate</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$R_1 = B30 + SB_{dep}^{net}$</td>
<td>$R_2 = SB_{ind}^{net}$</td>
<td>$R_1 = B30 + SB_{ind}^{net}$</td>
<td>$R_2 = SB_{dep}^{net}$</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$F_{1,ind}^{o}$</td>
<td>$p-val^b$</td>
<td>$F_{2,ind}^{o}$</td>
<td>$p-val^b$</td>
<td>$F_{1,dep}^{o}$</td>
<td>$p-val^b$</td>
</tr>
<tr>
<td>MD_{2x3}</td>
<td>0.801</td>
<td>0.448</td>
<td>0.064</td>
<td>0.944</td>
<td>5.823</td>
<td>0.012</td>
</tr>
<tr>
<td>MD_{3x3}</td>
<td>0.016</td>
<td>0.973</td>
<td>0.121</td>
<td>0.914</td>
<td>4.160</td>
<td>0.049</td>
</tr>
<tr>
<td>MD_{3x3x3}</td>
<td>0.001</td>
<td>1.000</td>
<td>2.470</td>
<td>0.361</td>
<td>5.959</td>
<td>0.026</td>
</tr>
<tr>
<td>MV_{2x3}</td>
<td>0.118</td>
<td>0.871</td>
<td>6.515</td>
<td>0.054</td>
<td>5.254</td>
<td>0.015</td>
</tr>
<tr>
<td>MV_{3x3}</td>
<td>0.035</td>
<td>0.956</td>
<td>0.090</td>
<td>0.954</td>
<td>6.095</td>
<td>0.011</td>
</tr>
<tr>
<td>MV_{3x3x3}</td>
<td>1.490</td>
<td>0.285</td>
<td>3.210</td>
<td>0.178</td>
<td>0.788</td>
<td>0.471</td>
</tr>
<tr>
<td>RP_{2x3}</td>
<td>0.006</td>
<td>0.984</td>
<td>0.299</td>
<td>0.773</td>
<td>1.743</td>
<td>0.173</td>
</tr>
<tr>
<td>RP_{3x3}</td>
<td>0.035</td>
<td>0.931</td>
<td>0.116</td>
<td>0.885</td>
<td>1.497</td>
<td>0.204</td>
</tr>
<tr>
<td>RP_{3x3x3}</td>
<td>0.009</td>
<td>0.984</td>
<td>0.087</td>
<td>0.922</td>
<td>2.315</td>
<td>0.109</td>
</tr>
</tbody>
</table>

* SB strategies are net of transactions costs estimated according to Hasbrouck (2009, JF)’s model as in Novy-Marx and Velikov (2016, RFS).
CROSS-SECTIONAL ASSET PRICING TEST

Harvey and Liu (2016, WP) define a scaled intercept (SI) to

- measure the incremental contribution of an augmented model w.r.t. a baseline model to explain the cross-sectional variations of the \( J \) test assets returns
- overcome the over-rejection issues of the GRS test

\[
SI_{med}^{lew} = \frac{\text{median}(\{|a_i^g|/s_i^b\}_{i=1}^J) - \text{median}(\{|a_i^b|/s_i^b\}_{i=1}^J)}{\text{median}(\{|a_i^b|/s_i^b\}_{i=1}^J)}
\]  

(6)
Cross-sectional asset pricing test

Harvey and Liu (2016, WP) define a scaled intercept (SI) to

- measure the incremental contribution of an augmented model w.r.t. a baseline model to explain the cross-sectional variations of the $J$ test assets returns
- overcome the over-rejection issues of the GRS test

$$SI_{med}^{ew} = \frac{\text{median} \left( \frac{\left| a^g_i \right|}{s^g_i} \right)_{i=1}^J - \text{median} \left( \frac{\left| a^b_i \right|}{s^b_i} \right)_{i=1}^J}{\text{median} \left( \frac{\left| a^b_i \right|}{s^b_i} \right)_{i=1}^J}$$

where,

- $\text{median}(.)$ is the median value of the ratio $\left| a^g_i \right|/s^g_i$ or $\left| a^b_i \right|/s^b_i$
- $s$ denotes the standard errors for the regression intercept $a$
- superscript $b$ is for the baseline model
- superscript $g$ is for the augmented model
- subscript $i$ refers to the i-th portfolio among the $J$ test assets
CROSS-SECTIONAL ASSET PRICING TEST

Harvey and Liu (2016, WP) define a scaled intercept (SI) to

- measure the incremental contribution of an augmented model w.r.t. a baseline model to explain the cross-sectional variations of the $J$ test assets returns
- overcome the over-rejection issues of the GRS test

$$SI_{med}^{med} = \frac{\text{median}\left(\left\{|a_i^g|/s_i^b\right\}_{i=1}^J\right) - \text{median}\left(\left\{|a_i^b|/s_i^b\right\}_{i=1}^J\right)}{\text{median}\left(\left\{|a_i^b|/s_i^b\right\}_{i=1}^J\right)}$$ (6)

Outputs:

- if $SI<0$ then augmented $>$ baseline model
CROSS-SECTIONAL ASSET PRICING TEST

To test the significance of the model improvement, Harvey and Liu (2016, WP) define the following bootstrap procedure,

**Step 1: Orthogonalization of the list of $K$ candidates**

- Baseline assets ($R_1$) = 30-Year US Treasury Bond (B30) + SMB + HML
- Test asset ($R_2^i$) = i-th candidate among the list of $K$ candidates
**CROSS-SECTIONAL ASSET PRICING TEST**

To test the significance of the model improvement, Harvey and Liu (2016, WP) define the following bootstrap procedure,

**Step 1: Orthogonalization of the list of $K$ candidates**

- Baseline assets ($R_1$) = 30-Year US Treasury Bond (B30) + SMB + HML
- Test asset ($R_i^2$) = i-th candidate among the list of $K$ candidates

\[
R_i^2 = \alpha_i + \beta_i R_1 + e_i
\]

\[
R_{i,2}^{\alpha,i} = R_i^2 - \alpha_i = \beta_i R_1 + e_i
\] (7)

Such that, $R_{i,2}^{\alpha,i}$ does not bring any additional information to the baseline model.
CROSS-SECTIONAL ASSET PRICING TEST

**Step 2: Bootstrap** (Similar to the method presented earlier)

In each sample of the $B$ bootstrap:

- a score for the scaled intercept $SI_{ew}^{med}$ can be obtained for the $K$ number of orthogonalized candidates (i.e., $R_2^\alpha,i$ with the $i = \{1, 2, ..., K\}$ candidate)
- take the minimum value among the $b$-th bootstrap to control for multiple testing

\[
SI^{b,*} = \min_{i \in \{1,2,\ldots,K\}} \{SI^{b,i}\}
\]
CROSS-SECTIONAL ASSET PRICING TEST

**Step 2:** Bootstrap (Similar to the method presented earlier)

In each sample of the $B$ bootstrap:
- a score for the scaled intercept $SI_{med}$ can be obtained for the $K$ number of orthogonalized candidates (i.e, $R^\alpha_{2,i}$ with the $i = \{1, 2, ..., K\}$ candidate)
- take the minimum value among the $b$-th bootstrap to control for multiple testing

$$SI^{b,*} = \min_{i \in \{1,2,...,K\}} \{SI^{b,i}\} \quad (8)$$

**Step 3:** Single test p-value

Select the candidate with the lowest $SI^o$ value and significant p-val

$$p-val = \frac{\#\{SI^o > SI^b\}}{B} \quad (9)$$
CROSS-SECTIONAL ASSET PRICING TEST

**Step 2:** Bootstrap (Similar to the method presented earlier)

In each sample of the $B$ bootstrap:

- a score for the scaled intercept $SI_{ew}^{med}$ can be obtained for the $K$ number of orthogonalized candidates (i.e, $R_{2,i}$ with the $i = \{1, 2, ..., K\}$ candidate)
- take the minimum value among the $b$-th bootstrap to control for multiple testing

$$SI_{b,*} = \min_{i \in \{1,2,\ldots,K\}} \{SI_{b,i}\} \quad (8)$$

**Step 3:** Single test p-value

Select the candidate with the lowest $SI^{o}$ value and significant p-val

$$p-val = \frac{\#\{SI^{o} > SI_{b}\}}{B} \quad (9)$$

**Step 4:** Multiple test p-value

$$p-val = \frac{\#\{SI^{o} > SI_{b,*}\}}{B} \quad (10)$$
## Cross-sectional Asset Pricing Test

<table>
<thead>
<tr>
<th>MVE Candidates →</th>
<th>Mkt</th>
<th>MV\textsubscript{dep}</th>
<th>MD\textsubscript{dep}</th>
<th>RP\textsubscript{dep}</th>
<th>MV\textsubscript{ind}</th>
<th>MD\textsubscript{ind}</th>
<th>RP\textsubscript{ind}</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>p-value</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>Scaled intercept (SI)</td>
<td>0.042</td>
<td>0.070</td>
<td>0.041</td>
<td>0.036</td>
<td>0.066</td>
<td>-0.012</td>
</tr>
<tr>
<td></td>
<td>Single test p-value</td>
<td>0.701</td>
<td>0.984</td>
<td>0.929</td>
<td>0.894</td>
<td>0.893</td>
<td>0.325</td>
</tr>
<tr>
<td></td>
<td>SI sequence</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Selected candidate(s)</td>
<td>RP\textsubscript{ind}</td>
<td>[0.000]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Multiple test p-value</td>
<td>[0.000]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Baseline** = US30 + SMB + HML

### Panel A: 2x3 cap-weighted independent portfolios as test assets

- GRS: 4.836, 4.189, 4.391, 4.538, 4.155, 4.445, 4.341
- p-value: 0.000, 0.000, 0.000, 0.000, 0.000, 0.000, 0.000
- Scaled intercept (SI): 0.042, 0.070, 0.041, 0.036, 0.066, -0.012, -0.842
- Single test p-value: 0.701, 0.984, 0.929, 0.894, 0.893, 0.325, 0.000
- SI sequence: 2, 2, 2, 2, 2, 2, 1
- Selected candidate(s): RP\textsubscript{ind}
- Multiple test p-value: [0.000]

### Panel B: 2x3 cap-weighted dependent portfolios as test assets

- p-value: 0.000, 0.000, 0.000, 0.000, 0.000, 0.000, 0.000
- Scaled intercept (SI): 0.049, 0.061, -0.838, 0.023, 0.009, -0.001, -0.008
- Single test p-value: 0.880, 0.865, 0.000, 0.782, 0.569, 0.547, 0.468
- SI sequence: 2, 2, 1, 2, 2, 2, 2
- Selected candidate(s): MD\textsubscript{dep}
- Multiple test p-value: [0.000]
# Cross-sectional Asset Pricing Test

<table>
<thead>
<tr>
<th>MVE Candidates →</th>
<th>Mkt</th>
<th>MV&lt;sub&gt;dep&lt;/sub&gt;</th>
<th>MD&lt;sub&gt;dep&lt;/sub&gt;</th>
<th>RP&lt;sub&gt;dep&lt;/sub&gt;</th>
<th>MV&lt;sub&gt;ind&lt;/sub&gt;</th>
<th>MD&lt;sub&gt;ind&lt;/sub&gt;</th>
<th>RP&lt;sub&gt;ind&lt;/sub&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Baseline</strong> = US30 + SMB + HML</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Panel C</strong>: 3x3 cap-weighted independent portfolios as test assets</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GRS</td>
<td>3.403</td>
<td>2.792</td>
<td>2.892</td>
<td>2.975</td>
<td>2.909</td>
<td>2.918</td>
<td>2.929</td>
</tr>
<tr>
<td>p-value</td>
<td>0.000</td>
<td>0.003</td>
<td>0.002</td>
<td>0.002</td>
<td>0.002</td>
<td>0.002</td>
<td>0.002</td>
</tr>
<tr>
<td>Scaled intercept (SI)</td>
<td>-0.006</td>
<td>0.041</td>
<td>0.255</td>
<td>0.373</td>
<td>-0.876</td>
<td>0.100</td>
<td>0.169</td>
</tr>
<tr>
<td>Single test p-value</td>
<td>0.542</td>
<td>0.755</td>
<td>0.883</td>
<td>0.895</td>
<td>0.000</td>
<td>0.697</td>
<td>0.755</td>
</tr>
<tr>
<td>SI sequence</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Selected candidate(s)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>MV&lt;sub&gt;ind&lt;/sub&gt;</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Multiple test p-value</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>[0.000]</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Panel D</strong>: 3x3 cap-weighted dependent portfolios as test assets</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>p-value</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Scaled intercept (SI)</td>
<td>0.210</td>
<td>0.249</td>
<td>0.221</td>
<td>-0.806</td>
<td>0.001</td>
<td>0.150</td>
<td>0.107</td>
</tr>
<tr>
<td>Single test p-value</td>
<td>0.937</td>
<td>0.979</td>
<td>0.916</td>
<td>0.000</td>
<td>0.528</td>
<td>0.861</td>
<td>0.849</td>
</tr>
<tr>
<td>SI sequence</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Selected candidate(s)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>RP&lt;sub&gt;dep&lt;/sub&gt;</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Multiple test p-value</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>[0.000]</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
# Cross-sectional Asset Pricing Test

<table>
<thead>
<tr>
<th>MVE Candidates →</th>
<th>Mkt</th>
<th>MV&lt;sub&gt;dep&lt;/sub&gt;</th>
<th>MD&lt;sub&gt;dep&lt;/sub&gt;</th>
<th>RP&lt;sub&gt;dep&lt;/sub&gt;</th>
<th>MV&lt;sub&gt;ind&lt;/sub&gt;</th>
<th>MD&lt;sub&gt;ind&lt;/sub&gt;</th>
<th>RP&lt;sub&gt;ind&lt;/sub&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline = US30 + SMB + HML</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Panel E: 3x3x3 cap-weighted independent portfolios as test assets</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GRS</td>
<td>2.262</td>
<td>2.135</td>
<td>2.164</td>
<td>2.213</td>
<td>2.085</td>
<td>2.332</td>
<td>2.110</td>
</tr>
<tr>
<td>p-value</td>
<td>0.000</td>
<td>0.001</td>
<td>0.001</td>
<td>0.000</td>
<td>0.001</td>
<td>0.000</td>
<td>0.001</td>
</tr>
<tr>
<td>Scaled intercept (SI)</td>
<td>0.245</td>
<td>0.030</td>
<td>-0.790</td>
<td>0.301</td>
<td>0.283</td>
<td>0.111</td>
<td>0.225</td>
</tr>
<tr>
<td>Single test p-value</td>
<td>0.943</td>
<td>0.661</td>
<td>0.000</td>
<td>0.970</td>
<td>0.962</td>
<td>0.816</td>
<td>0.952</td>
</tr>
<tr>
<td>SI sequence</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Selected candidate(s)</td>
<td></td>
<td></td>
<td></td>
<td>MD&lt;sub&gt;dep&lt;/sub&gt;</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Multiple test p-value</td>
<td></td>
<td></td>
<td></td>
<td>0.000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Panel F: 3x3x3 cap-weighted dependent portfolios as test assets</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>p-value</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Scaled intercept (SI)</td>
<td>-0.015</td>
<td>0.007</td>
<td>-0.668</td>
<td>-0.039</td>
<td>0.006</td>
<td>0.013</td>
<td>0.020</td>
</tr>
<tr>
<td>Single test p-value</td>
<td>0.566</td>
<td>0.571</td>
<td>0.000</td>
<td>0.431</td>
<td>0.610</td>
<td>0.686</td>
<td>0.739</td>
</tr>
<tr>
<td>SI sequence</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Selected candidate(s)</td>
<td></td>
<td></td>
<td></td>
<td>MD&lt;sub&gt;dep&lt;/sub&gt;</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Multiple test p-value</td>
<td></td>
<td></td>
<td></td>
<td>0.000</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
CONCLUSION AND TAKEAWAYS

Testing the MVE of Smart Beta strategies on characteristic-sorted portfolios

▶ Context
  ▶ Multidimensional market risks, especially after 1993
  ▶ Inefficiencies of long-short factors
  ▶ Sample errors for estimating MVE
  ▶ Need for long-only solutions

▶ Contribution
  ▶ Risk-based optimization on DNS opportunity sets span a single-index model, other MVE candidates (market-cap and other risk-based strategies) and improves a 3-factor model
  ▶ Risk-based strategies on DNS opportunity sets have incremental significance for pricing characteristics-sorted portfolios
  ▶ Dependent-sorted portfolios provide a better investment opportunity set to investors compared to independent-sorted portfolio

▶ Robustness
  ▶ Out-of-sample, multiple testing
END...

Thank you for your attention!
Contact: Marie.Lambert@uliege.be


Diversification Return: Bootstrap

Method:
- block-bootstrap method from Politis and Romano (1992)
- studentized test statistic following Ledoit and Wolf (2008)

Bootstrap:
1. block length = 10 observations (robust to other length \{2, 4, 6, 8, 10\})
2. match the length of the original time-series (630 observations)
3. randomly resample with replacement the original time-series
4. keep the same sequence for all assets in each sample (cross-dependence)
5. 4999 simulations similar to Ledoit and Wolf (2008)
TESTING THE INCREMENTAL DIVERSIFICATION RETURN

Hypothesis testing the spread in Sharpe ratio between the strategy \(i\) and \(j\),

\[
\hat{\Delta} = \frac{\hat{\mu}_i}{\hat{\sigma}_i} - \frac{\hat{\mu}_j}{\hat{\sigma}_j}
\]  

(11)

**Assumption:**
Difference between the first and second moments of the distributions between the two series converge towards zero

\[
\sqrt{T}(\hat{u} - u) \stackrel{d}{\rightarrow} N(0, \Omega)
\]

(12)

- \(\hat{u} = (\hat{\mu}_i, \hat{\mu}_j, \hat{\sigma}_i^2, \hat{\sigma}_j^2)\) are the sample estimates of \(u = (\mu_i, \mu_j, \sigma_i^2, \sigma_j^2)\)
- \(\stackrel{d}{\rightarrow}\) refers to the convergence in distribution of the parameters
- \(\Omega\) not valid when returns exhibit non-normal distribution or serial autocorrelation

**Solution:**

\[
\sqrt{T}(\hat{v} - v) \stackrel{d}{\rightarrow} N(0, \Psi)
\]

(13)

where \(\hat{v} = (\hat{\mu}_i, \hat{\mu}_j, \hat{\gamma}_i, \hat{\gamma}_j)\) is the sample estimates of \(v = (\mu_i, \mu_j, \gamma_i, \gamma_j)\), \(\hat{\gamma}_i = E(r_i^2)\) and \(\hat{\gamma}_j = E(r_j^2)\) and a HAC kernel estimate of \(\Psi\).
**Diversification Return: Hypothesis Testing**

**Spread in Sharpe ratio:**

\[
 f(a, b, c, d) = \frac{a}{\sqrt{c - a^2}} - \frac{b}{\sqrt{d - b^2}}
\]  
(14)

With \( a = \hat{\mu}_i \), \( b = \hat{\mu}_j \), \( c = \hat{\gamma}_i \), and \( d = \hat{\gamma}_j \).

Gradient of this function (delta-method) is

\[
 \nabla' f(\hat{\nu}) = \left( \frac{c}{(c - a^2)^{1.5}}, -\frac{d}{(d - b^2)^{1.5}}, -\frac{1}{2} \frac{a}{(c - a^2)^{1.5}}, \frac{1}{2} \frac{b}{(d - b^2)^{1.5}} \right)
\]

The standard error of delta estimate is,

\[
 s(\hat{\Delta}) = \sqrt{\frac{\nabla' f(\hat{\nu}) \hat{\Psi} \nabla f(\hat{\nu})}{T}}
\]  
(15)

The kernel estimator \( \hat{\Psi} \) ensures that the estimation of the standard error is robust to heteroskedasticity and autocorrelation (HAC).
DIVERSIFICATION RETURN: HYPOTHESIS TESTING

Studentized test statistic:

On the original time-series,

\[ d = \frac{|\hat{\Delta}|}{s(\hat{\Delta})} \]  \hspace{1cm} (16)

On the b-th bootstrap sample,

\[ d^b = \frac{|\hat{\Delta}^b - \hat{\Delta}|}{s(\hat{\Delta}^b)} \]  \hspace{1cm} (17)

The bootstrap 1-\(\alpha\) confidence interval is defined as:

\[ \left[ \hat{\Delta} - z_{|.|,1-\alpha/2}s(\hat{\Delta}), \hat{\Delta} + z_{|.|,1-\alpha/2}s(\hat{\Delta}) \right] \]  \hspace{1cm} (18)

with \(z_{|.|,1-\alpha}\) the quantile of the distribution of \(d^b\) denoted \(\mathcal{L}(d^b)\).

p-value:

\[ p\text{-val} = \frac{\#\{d^b \geq d\} + 1}{B + 1} \]  \hspace{1cm} (19)
TRANSACTION COSTS: GIBBS ESTIMATES

Hasbrouck (2009) extend Roll (1984)’s price dynamics model with a market factor

\[ \Delta p_t = c \Delta q_t + \beta_{rm} r m_t + u_t \]  

with

\[ \Delta p_t = p_t - p_{t-1} \]
\[ = m_t + c q_t - m_{t-1} - c q_{t-1} \]  

\[ = c \Delta q_t + u_t \]  

▶ \( m_t \) is the log midpoint of the prior bid-ask price
▶ \( p_t \) is the log trade price
▶ \( q_t \) is the sign of the last trade of the day (+1 for a buy and −1 for a sale)
▶ \( u_t \) is assumed to be unrelated to the sign of the trade \( q_t \)
▶ \( r m_t \) is the market return on day \( t \)
▶ \( \beta_{rm} \) is the slope on the market return
▶ \( c \) is the effective cost
**TRANSACTION COSTS: GIBBS ESTIMATES**

\[ \Delta p_t = c \Delta q_t + \beta_{rm} m_{rt} + u_t \]  (22)

Iterative Bayesian methodology to estimate the effective costs (c):

1. Initialize \( q_1 \) to +1 and \( \sigma_u^2 \) to 0.001.
   - if no trade \( q_t = 0 \) (in CRSP, PRC < 0) else \( q_t = \text{sign}(\Delta p_t) \)
   - minimum of 60 to a max 250 daily observations

2. Initialize the distribution from where the values \( c, \beta_{rm}, \) and \( \sigma_u^2 \) will be drawn:
   - \( c \sim N^+ (\mu = 0.01, \sigma^2 = 0.01^2) \)
   - \( \beta_{rm} \sim N (\mu = 1, \sigma^2 = 1) \)
   - \( \sigma^2_u \sim IG (\alpha = 10^{-12}, \beta = 10^{-12}) \)
TRANSACTION COSTS: GIBBS ESTIMATES

for 1 to 1000 sweeps

1. Perform a Bayesian OLS regression on a 250-day of lagged observations to estimate the new values of $c$ and $\beta_{rm}$, update the posterior distribution of the parameters and make a new draw of the coefficients.

2. Back out $u_t$ according to $c$, $\beta_{rm}$, $\Delta p_t$, $rm_t$, $\Delta q_t$

$$u_t = \Delta p_t - \beta_{rm}rm_t - c\Delta q_t$$  \hspace{1cm} (23)

$\triangleright$ update $\sigma_u^2$

3. Draw new series of $q_t$ according to the posterior $\sigma_u^2$

$$u_t = \Delta p_t - \beta_{rm}rm_t - cq_t + cq_{t-1}$$  \hspace{1cm} (24)

$\triangleright$ estimate $u_t(q_t = +1)$ and $u_t(q_t = -1)$ given $u_t \sim N(0, \sigma_u^2)$

$$\text{Odds} = \frac{f(u_t(q_t = +1))}{f(u_t(q_t = -1))} \begin{cases} q_t = +1 & \text{if Odds} > 1 \\ q_t = -1 & \text{if Odds} < 1 \end{cases}$$  \hspace{1cm} (25)

end

$\rightarrow c$ is the average of the last 800 estimations ("burn in" the 200 first obs.)
PORTFOLIO SORTING AND FACTOR PERFORMANCE

A small exercise:
Sorting on two characteristics with a perfect negative correlation (Lambert, Fays, and Hübner 2018, WP)

- First sort on market capitalization, ME: SMB factor
- Second sort on –ME: HML factor
- The factors should be similar

<table>
<thead>
<tr>
<th>Factor construction</th>
<th># Obs</th>
<th>Mean (in %)</th>
<th>SD (in %)</th>
<th>t-stat</th>
<th>Correlation Matrix</th>
</tr>
</thead>
<tbody>
<tr>
<td>SMB</td>
<td>618</td>
<td>0.136</td>
<td>1.651</td>
<td>2.045</td>
<td>1</td>
</tr>
<tr>
<td>HML</td>
<td>618</td>
<td>0.164</td>
<td>1.962</td>
<td>2.078</td>
<td>0.98041***</td>
</tr>
<tr>
<td>SMB</td>
<td>618</td>
<td>0.151</td>
<td>2.102</td>
<td>1.783</td>
<td>1</td>
</tr>
<tr>
<td>HML</td>
<td>618</td>
<td>0.204</td>
<td>2.411</td>
<td>2.105</td>
<td>0.97861***</td>
</tr>
<tr>
<td>SMB</td>
<td>618</td>
<td>0.4</td>
<td>4.105</td>
<td>2.421</td>
<td>1</td>
</tr>
<tr>
<td>HML</td>
<td>618</td>
<td>-0.244</td>
<td>3.248</td>
<td>-1.869</td>
<td>-0.0184</td>
</tr>
<tr>
<td>SMB</td>
<td>618</td>
<td>0.705</td>
<td>7.601</td>
<td>2.304</td>
<td>1</td>
</tr>
<tr>
<td>HML</td>
<td>618</td>
<td>-0.097</td>
<td>4.112</td>
<td>-0.585</td>
<td>0.0579</td>
</tr>
</tbody>
</table>
PORTFOLIO SORTING AND FACTOR PERFORMANCE

A small exercise:
Sorting on two characteristics with a perfect negative correlation (Lambert, Fays, and Hübner 2018, WP)

<table>
<thead>
<tr>
<th>Panel</th>
<th>Independent - 2x3 - NYSE</th>
<th>Independent - 2x3 - Name</th>
<th>Independent - 3x3 - NYSE</th>
<th>Independent - 3x3 - Name</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>SL</td>
<td>SM</td>
<td>SH</td>
<td>BL</td>
</tr>
<tr>
<td>Average (# stocks)</td>
<td>0</td>
<td>416</td>
<td>2226</td>
<td>394</td>
</tr>
<tr>
<td>Min (# stocks)</td>
<td>0</td>
<td>133</td>
<td>439</td>
<td>182</td>
</tr>
<tr>
<td>Max (# stocks)</td>
<td>0</td>
<td>696</td>
<td>4085</td>
<td>638</td>
</tr>
</tbody>
</table>

|       | SL          | SM  | SH   | BL   | BM  | BH  | ML  | MM  | MH  |
|-------|--------------------------|--------------------------|--------------------------|--------------------------|
| Average (# stocks) | 0 | 1 | 977 | 1021 | 0 | 0 | 1 | 1345 | 0   |
| Min (# stocks)      | 0 | 1 | 255 | 268  | 0 | 0 | 1 | 353  | 0   |
| Max (# stocks)      | 0 | 1 | 1726| 1726 | 0 | 0 | 1 | 2302| 0   |
PORTFOLIO SORTING AND FACTOR PERFORMANCE

A small exercise:
Sorting on two characteristics with a perfect negative correlation (Lambert, Fays, and Hübner 2018, WP)

- First sort on market capitalization, ME: SMB factor
- Second sort on −ME: HML factor
- The factors should be similar

<table>
<thead>
<tr>
<th>Factor construction</th>
<th># Obs</th>
<th>Mean (in %)</th>
<th>SD (in %)</th>
<th>t-stat</th>
<th>Correlation Matrix</th>
</tr>
</thead>
<tbody>
<tr>
<td>Panel E: Dependent - 2x3 - NYSE</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SMB</td>
<td>618</td>
<td>0.138</td>
<td>1.264</td>
<td>2.706</td>
<td>1</td>
</tr>
<tr>
<td>HML</td>
<td>618</td>
<td>0.193</td>
<td>2.351</td>
<td>2.038</td>
<td>0.88938***</td>
</tr>
</tbody>
</table>

| Panel F: Dependent - 2x3 - Name |       |             |           |        |                    |
| SMB                 | 618   | 0.117       | 1.367     | 2.128  | 1                  |
| HML                 | 618   | 0.231       | 2.36      | 2.434  | 0.8237***         |

| Panel G: Dependent - 3x3 - NYSE |       |             |           |        |                    |
| SMB                 | 618   | 0.104       | 1.476     | 1.743  | 1                  |
| HML                 | 618   | 0.104       | 1.478     | 1.748  | 0.9993***         |

| Panel H: Dependent - 3x3 - Name |       |             |           |        |                    |
| SMB                 | 618   | 0.190       | 1.765     | 2.670  | 1                  |
| HML                 | 618   | 0.193       | 1.763     | 2.724  | 0.99953***        |