Risk, Return, and the Optimal Exploitation of Stock Characteristics

Christopher G. Lamoureux
&
Huacheng Zhang

May 29, 2014
1. We know that the cross-section of stock returns has a strong factor structure, and we also know that stock characteristics are linked to factor structure.

2. Furthermore characteristics can predict the cross-section of expected returns out-of-sample: Lewellen (2013). Although Nagel (2012) notes that this is a pseudo out-of-sample test.


4. Recent addition to variables that explain cross-sectional predictability of expected returns: lagged same month returns (Heston and Sadka 2008).
This paper’s contributions

1. We use characteristics to form optimal portfolios using the Brandt, Santa-Clara, and Valkanov (2009) algorithm.
   - Maximize expected utility by a fully-invested portfolio with weights linear in measurable characteristics.

2. Question is whether a risk inverse investor can use characteristics to significantly increase expected utility out-of-sample.
   - Addresses to some extent the pseudo out-of-sample critique with a different loss function.
   - Provides a convenient and natural way to combine characteristics, and assess each one’s marginal contribution.
This paper’s contributions

3. By including lagged same-month returns as characteristics we explore this in the context of the other characteristics.

4. Because we have a relatively long sample we isolate factor loadings conditional on month of the year.

5. Period includes 2009, and the momentum crash:
   - We analyze effect of characteristics on portfolio weights on a yearly basis.
   - We consider the (small sample) bias in the effects of characteristics on portfolio weights.

6. We compare a 15-year rolling “training period” with updating (over our 38 year out-of-sample period).
7. Since we have a CRRA (isoelastic) utility function, we examine the effect of changing relative risk aversion on the desired characteristic exposure.

$$E(U) = \frac{1}{T} \sum_{t=0}^{T-1} \frac{(1+r_{p,t+1})^{1-\gamma}}{1-\gamma}.$$ 

8. We bootstrap over the optimization to provide statistical context to all results.
Results

1. **The loss function matters.**
   - Optimal portfolio for $\gamma$ investor with all characteristics, using rolling protocol has 45% Fama-French-Carhart alpha, and CE of -23%.

2. The only characteristics that generate statistically significant improvement in EU are book-to-market and beta.
   - The *dynamic* characteristics do not help—**even before transactions costs**.

3. Portfolios formed based on lagged same-month returns do have higher (lower) exposures to known factors in months when the factors have high (low) expected returns. This is not evident in unconditional betas (Keloharju, Linnainmaa and Nyberg, 2013).

4. There is material bias and imprecision in the sample $\theta$ estimates.
5. There is an in-sample complementarity between multiple factors; ex.: size and beta.

6. While there is weak in-sample substitution between lagged same-month returns and momentum, they complement size and book-to-market.

7. $\theta$ coefficients are lower in later years–notably on momentum after 2009. But all are significant; (McLean and Pontiff, 2012).

8. The updating protocol dominates the rolling protocol out-of-sample.

9. Investors’ desired exposure to characteristics diminishing in $\gamma$, with the exception of beta.
10. **New – not yet in paper:** Note that this optimizing algorithm asks the investor to put a dogmatic prior on a model wherein the relationships between return moments (and co-moments) between lagged characteristics and next month returns are robust and stable over time.

- We find significant improvement in Certainty Equivalent by using the optimal portfolio selected for an investor who is more risk averse.
- Although, there is no explicit model of the return distribution.
- This overturns the results on the *dynamic* characteristics.
- Raises the important question of *why* this does so well.
Portfolio Formation

\[
\max_\theta \sum_{t=0}^{T-1} \frac{(1 + r_{p,t+1})^{1-\gamma}}{1 - \gamma} \left( \frac{1}{T} \right)
\]

\[
r_{p,t+1} = \sum_{i=1}^{N_t} \left( \bar{\omega}_{i,t} + \frac{1}{N_t} \theta' x_{i,t} \right) \cdot r_{i,t+1}
\]

- \(\theta\) has dimension \(K\), the number of characteristics.
- BSV use size, book-to-market, and momentum, and find that the optimal portfolio outperforms the market, and the equally-weighted market.
- Extremely parsimonious relative to mean-variance optimization, which suffers notoriously from estimation risk.
Data

- Start forming optimal portfolios on December 31, 1959.
  - Sample consists of all stocks with no missing returns from January 1955 through December 1959, and valid Compustat data.
- On every month-end date, the sample contains $N_t K + 1$ vectors $\gamma_{i,t}$.
  - $\gamma_{i,t} = (x_{i,t}, r_{i,t+1})$. (No look-ahead bias.)
- Use the $\theta$ vector just obtained to form out-of-sample portfolio for each of the next 12 months.
Data (continued)

- Rolling Protocol: Drop 12 months in 1960, so that there are always 180 months used to estimate $\theta$.
- Updating Protocol: Add next 12 months. Sample sized used in estimating $\theta$ grows from 180 months to 624 months.
- Note that stocks can come into the sample at any time—in both the in- and out-of-sample sets.
- Exclude smallest 10% of eligible stocks pre-1975 and 20% post-1975.
Data – Characteristics

- **Size**: Number of outstanding shares (CRSP) × closing price—both month-end before portfolio formation.
- **Book-to-market**: Fiscal year-end book value from Compustat – lagged 6 - 18 months. Time-matched to market cap.
  - Book value: Total assets + deferred taxes + ITC - total liabilities - preferred equity.
- **Momentum**: Compounded return from month -12 to -1.
- **Beta**: OLS coefficient of excess stock return on excess market return in months -59 - 0.
- **Lagged same-month return**:
  - \( r_{t-11} \)
  - \( \sum_{i=1}^{5} r_{t-(12i-1)} \cdot \frac{1}{5} \)
One of the brilliant insights of BSV is to standardize and normalize the coefficients.

- Makes the portfolio weight constraint natural.
- Allows direct comparison of characteristics’ effects on weights.
- Also means that the characteristics are not different from shrunken characteristics.
Number of stocks in sample—by month
An interesting aspect of the BSV portfolio selection algorithm is that there is no likelihood function (or errors).

Construct sampling distributions by drawing (with replacement) from \( \Upsilon_{i,t} \) to construct a pseudo-sample of the same term and size as in the data. We then select the optimal portfolios (to obtain \( \theta \)) at the beginning of each oos year. Repeat 10,000 times to construct a sampling distribution of \( \theta \) and functions of interest of \( \theta \) and the data (e.g., alpha, Sharpe ratio, and CE).

I believe that this approach requires that the investor have a dogmatic prior (not exactly clear what that is).
Exploit our empirical design

- Average $\theta$ coefficients on each characteristic in isolation.
- Time series behavior of these coefficients.
- Substitutability / Complementarity between characteristics when used in combination.
- Difference between rolling and updating on coefficients.
- Effects of $\gamma$ on coefficients.
- Small sample properties.
This plot shows the bootstrap distributions for the coefficients on the stock's size from Model 5 (top panel) and Model 8 (lower panel). The box and whiskers plot shows the 95%ile range (whiskers), the interquartile range (box), and median (bar inside box) for the coefficient in each year. The sample estimate is shown as a red X.
Summary of Size $\theta$

- By itself and in conjunction with mom, btm, and lagged same-month returns:
  - Significantly negative for $\gamma = 3$.
  - Significantly positive for $\gamma \geq 15$.
  - Not much difference between rolling and updating protocols.

- When beta is in characteristic set:
  - Significantly negative for all $\gamma$.
  - Rolling shrinks $\theta$ for $\gamma = 3$.
  - Bootstrap standard deviations are twice as high than when beta is not.
  - Sample bias much larger ($\sim 1 \sigma$) than when beta is not.

- Overall:
  - Appeal of small stocks declines monotonically in $\gamma$. 

This plot shows the bootstrap distributions for the coefficients on the stock's book-to-market ratio from Model 5 (top panel) and Model 9 (lower panel). The box and whiskers plot shows the 95th percentile range (whiskers), the interquartile range (box), and median (bar inside box) for the coefficient in each year. The sample estimate is shown as a red X.
Summary of Book-to-Market $\theta$

- Significantly positive for all $\gamma$, all characteristic sets.
- Substitutability with beta (Significantly lower in Model 2 than in Model 1, for all $\gamma$.)
- Complementarity with lagged same-month returns (Significantly higher in Model 4 than Model 1, for all $\gamma$.)
- No complementarity or substitutability with momentum and size.
- Rolling and updating very similar.
- Exposure to value monotonically declining in $\gamma$—although it flattens out at high $\gamma$ at a significantly positive value.
- (Recalling that characteristics are standardized), the effect is larger than size, momentum and beta.
This plot shows the bootstrap distributions for the coefficient on the stock's momentum from Model 6 (top panel) and Model 7 (bottom panel). The box and whiskers plot shows the 95%ile range (whiskers), the interquartile range (box), and median (bar inside box) for the coefficient in each year. The sample estimate is shown as a red X.
Momentum $\theta$: $\gamma$ 3 and $\gamma$ 5

This plot shows the bootstrap distributions for the coefficient on the stock's momentum from Model 6, gamma 3 (top panel) and Model 6, gamma 5 (bottom panel). The box and whiskers plot shows the 95th percentile range (whiskers), the interquartile range (box), and median (bar inside box) for the coefficient in each year. The sample estimate is shown as a red X.
Summary of Momentum $\theta$

- Weakly complementary to size, book-to-market, and beta.
- Weakly substitutable with lagged same-month returns.
- Model 2 significantly higher than Model 7, for all $\gamma$, rolling and updating.
- Exposure to momentum monotonically declining in $\gamma$. 
This plot shows the bootstrap distributions for the coefficient on the stock’s monthly beta from Model 10 (top panel) and Model 2 (bottom panel). The box and whiskers plot shows the 95%ile range (whiskers), the interquartile range (box), and median (bar inside box) for the coefficient in each year. The sample estimate is shown as a red X.
Summary of Beta $\theta$

- Remarkably constant across $\gamma$.
- Always significantly negative (all $\gamma$, characteristic sets, and Rolling / Updating).
- Strong complementarity with size.
- The *complementarities* shrink in $\gamma$.
- (As we see on the previous slide), *Complementarities* vary widely over time.
This plot shows the bootstrap distributions for the coefficient on the stock’s same-month return last year from Model 3 (top panel) and Model 4 (bottom panel).

The box and whiskers plot shows the 95% range (whiskers), the interquartile range (box), and median (bar inside box) for the coefficient in each year. The sample estimate is shown as a red X.
This plot shows the bootstrap distributions for the coefficient on the stock's same-month return averaged over the previous 5 years from Model 3 (top panel) and Model 4 (bottom panel). The box and whiskers plot shows the 95%ile range (whiskers), the interquartile range (box), and median (bar inside box) for the coefficient in each year. The sample estimate is shown as a red X.
θ on both same-month return characteristics

This plot shows the bootstrap distributions for the coefficient on the stock's same-month return from last year from Model 5—Rolling protocol (top panel) and average same-month return over the preceding 5 years from Model 5—Rolling protocol (bottom panel). The box and whiskers plot shows the 95% range (whiskers), the interquartile range (box), and median (bar inside box) for the coefficient in each year. The sample estimate is shown as a red X.
Sum of $\theta$ on both same-month return characteristics: Updating vs. Rolling

This plot shows the bootstrap distributions for the sum of the coefficient on the stock's lagged same-month return and the moving average of the same-month return over the previous 5 years from Model 5, with coefficient of relative risk aversion $= 3$. The top panel is from the updating protocol and the bottom panel is from the rolling protocol. The box and whiskers plot shows the 95%ile range (whiskers), the interquartile range (box), and median (bar inside box) for the coefficient in each year. The sample estimate is shown as a red X.
Summary of lagged same-month return $\theta$s

- Effect is very large for low $\gamma$.
- Appeal of predictability diminishes sharply in $\gamma$.
- Substitutability with Momentum, and complementarity with: size, BtM, and beta.
- (As seen on the last slide), Rolling is much more volatile than Updating.
- (As seen on the last slide), Rolling subject to much larger sample bias than Updating.
Short positions by month, Updating protocol

Model 1
1.0 2.0 3.0 4.0
Model 2
1 2 3 4 5
Model 3
0 1 2 3 4
Model 4
1 2 3 4 5 6 7 8
Model 5
Year
1980 1990 2000 2010
0 2 4 6 8

Color Code: gamma = 3 gamma = 5 gamma = 10 gamma = 15 gamma = 20
Short positions by month, Updating protocol

Introduction
Characteristics and the Cross Section
Contributions
Results

BSV
Data
Bootstrap
Detailed Results
θ
Leverage
Out-of-sample
Alpha
Sharpe Ratio
Certainty Equivalent
Factor exposures
Month-by-Month

What next?
Leverage generally declines in $\gamma$ (as expected).

Individual Characteristics:
  - Momentum: Leverage peaks in late 1998, declines in last three years.
  - BtM: Highest of 4 traditional characteristics—declining over time.
  - Size: Largest temporal changes.
  - Lagged same-month return: Higher than other 4 characteristics; declines significantly over time.
  - More stable for persistent characteristics: mom, BtM, size; large month-to-month variation for more volatile characteristics.

Combinations:
  - From Models 4 and 5 reaches high of 8 in first three years, declines over time.
As BSV note, not surprising that conditioning on size, BtM, and momentum do not produce significantly positive FFC alpha.

Therefore surprising that conditioning on beta does.
- By itself, 8 bps (significantly > 0).
- Case of $\gamma = 3$, Updating, adding to: size, BtM, and mom increases alpha from 6 (insig.) to 60 bps (sig.)
- Same case, adding beta adds 60 bps to all other characteristics (not sig.)

Points of Reference:
- Value-Weighted Index: 3 ($-2, 8$).
- Value-Weighted Index: 7 ($5, 10$).
Unconditional FFC alpha

- **Lagged same month returns** produce significant oos alpha.
- Case of $\gamma = 3$, Updating, alpha is 162 (119, 208) bps per month.
- (Same case): Adding size, mom, and BtM to $r_{t-11}$ and $\bar{r}_{t-11}$ increases alpha to 210 (147, 281) bps.
- Maximum bootstrap mean of the oos alpha across 100 cases: $\gamma = 3$, Rolling format, all 6 characteristics: 334 (251, 428) bps. (Sample alpha: 45% annual.)
- Not surprisingly, alpha declines monotonically in $\gamma$: $\gamma = 15$, Rolling, all 6 characteristics: 78 (58, 101) bps. (Sample alpha: 10.5% annual.)
Sharpe Ratio

- Unlike alpha, SR does not diminish in $\gamma$.
- In all 10 cases, maximized using all 6 characteristics (Model 5). 1 exception: $\gamma = 3$, Updating: excluding beta does just as well.
- Maximized: $\gamma = 5$, Updating, (All 6 characteristics): Sample: 1.25; B-S: 1.19 (1.10, 1.29), (annualized).
- Rolling, $\gamma = 5$, All 6 characteristics: Sample: 1.16; B-S: 1.12 (1.01, 1.23).
- Reference:
  - Value-Weighted Index: 0.49 (0.45, 0.53).
  - Equally-Weighted Index: 0.57 (0.56, 0.59).
Certainty Equivalent

- Interesting that the \textit{out-of-sample} alpha and Sharpe ratios of the optimized portfolios are so large, since neither was maximized in-sample.

- Here, we have 456 months of out-of-sample returns, that we use to measure the average utility. CE return is the constant rate of return on the portfolio that would provide the same average utility.

- This is what was maximized in-sample.
Certainty Equivalent

- **Benchmarks for $\gamma = 3$:**
  - Value-Weighted Index: Sample: 75, B-S: 75 (70, 80) bps per month.
  - Equally-Weighted Index: Sample: 86, B-S: 86 (83, 88) bps per month.

- **Benchmarks for $\gamma = 5$:**
  - Value-Weighted Index: Sample: 53, B-S: 53 (47, 58) bps per month.
  - Equally-Weighted Index: Sample: 54, B-S: 54 (51, 56) bps per month.

- **Benchmarks for $\gamma = 10$:**
  - Value-Weighted Index: Sample: -10, B-S: -12 ($-18$, $-5$) bps per month.
  - Equally-Weighted Index: Sample: -45, B-S: -46 ($-50$, $-41$) bps per month.
Certainty Equivalent: Inference

- Inference:
  - Which characteristics lead to a significant increase in CE?
  - Is this effect robust across $\gamma$?
  - Do these effects parallel effects on alpha and SR?

- We confirm the BSV findings: Model 1:
  - $\gamma = 3$:
    - Updating: B-S 2.5%ile: 98 (EWI 97.5%ile: 88).
    - Rolling: B-S 2.5%ile: 51.
  - $\gamma = 5$:
    - Updating: B-S 2.5%ile: 66 (VWI 97.5%ile: 58).
    - Rolling: B-S 2.5%ile: 48.
  - $\gamma = 10$:
    - Updating: B-S 2.5%ile: -8 (VWI 97.5%ile: -5).
    - Rolling: B-S 2.5%ile: -15.
Certainty Equivalent: Using a higher $\gamma$

Model 1, Updating Protocol, $\gamma$ 3 investor.

<table>
<thead>
<tr>
<th>$\gamma$ Used</th>
<th>B-S 2.5%ile</th>
<th>B-S Median</th>
<th>B-S 97.5%ile</th>
</tr>
</thead>
<tbody>
<tr>
<td>EWI</td>
<td>83.38</td>
<td>85.79</td>
<td>88.27</td>
</tr>
<tr>
<td>3</td>
<td>97.72</td>
<td>120.93</td>
<td>144.38</td>
</tr>
<tr>
<td>5</td>
<td>135.11</td>
<td>150.94</td>
<td>168.64</td>
</tr>
</tbody>
</table>

Model 1, Updating Protocol, $\gamma$ 5 investor.

<table>
<thead>
<tr>
<th>$\gamma$ Used</th>
<th>B-S 2.5%ile</th>
<th>B-S Median</th>
<th>B-S 97.5%ile</th>
</tr>
</thead>
<tbody>
<tr>
<td>VWI</td>
<td>47.37</td>
<td>52.47</td>
<td>57.60</td>
</tr>
<tr>
<td>5</td>
<td>66.09</td>
<td>81.77</td>
<td>97.42</td>
</tr>
<tr>
<td>10</td>
<td>86.96</td>
<td>97.99</td>
<td>109.89</td>
</tr>
</tbody>
</table>
CE: Size, BtM, mom individually

- Momentum in isolation:
  - For all $\gamma$ values and both protocols, B-S mean CE is less than preferred index (difference not significant).
  - Ex: 95% bounds for $\gamma = 3$, Updating: 71 (57, 83), (recall EWI: 86 (83, 88).)

- Size in isolation:
  - For $\gamma = 3, 5, 10$, both protocols, B-S mean CE is significantly less than preferred index.
  - Ex: 95% bounds for $\gamma = 3$, Updating: (68, 78), (recall EWI 2.5%ile: 83).
  - For $\gamma = 15, 20$, both protocols, not worse than index, but not significantly better.
Book-to-Market in isolation:
- For $\gamma = 3, 5, 10$, both protocols, CE is significantly larger than preferred index.
- Ex: $\gamma = 3$, Updating: 107 (93, 122), Rolling: 113 (100, 128).
- For $\gamma = 15$, Rolling protocol significantly better than EWI, but not updating.
- For $\gamma = 20$, Not sig. different from VWI.
- Additional example: $\gamma = 10$: Updating: 5 ($-6, 16$), Rolling: 16 ($6, 25$), (VWI: $-12 (-18, -5)$.

Using a higher $\gamma$ does not have much effect on the CE from size, book-to-market, and momentum individually.
CE: Size, BtM, mom

- Does adding mom and size to BtM enhance CE—beyond using only BtM?
  - No. The difference between CE from Models 1 and 9 is not significantly different in any of the 10 cases.
  - Example: $\gamma = 3$: Updating: BtM only: 107 (93, 122), size, BtM, and mom: 124 (98, 144).
  - Example: $\gamma = 3$: Rolling: BtM only: 113 (100, 128), size, BtM, and mom: 85 (51, 114).
  - Example: $\gamma = 10$: Updating: BtM only: 5 (−6, 16), size, BtM, and mom: 5 (−8, 19).
  - Example: $\gamma = 10$: Rolling: BtM only: 16 (6, 25), size, BtM, and mom: 0 (−15, 14).
  - BtM only does as well with rolling as with updating. Rolling much worse when size included.
  - Transactions costs and leverage are much higher when mom is included in characteristic set.
Does adding mom and size to BtM enhance CE—beyond using only BtM?

- Using a higher $\gamma$ adds to momentum’s appeal.
- Adding size and momentum to book-to-market affords significantly higher CE for the $\gamma$ 3 investor who uses the $\gamma$ 5 optimal portfolio.
Beta

- **In isolation:**
  - Not much difference between updating and rolling protocols.
  - For $\gamma = 3$, CE significantly less than EWI.
  - For $\gamma = 5$, CE not significantly different from VWI.
  - For $\gamma = 10$, CE significantly higher than VWI.
  - For $\gamma > 10$, CE not significantly different from VWI.
  - Similarly the $\gamma = 5$ investor using the optimal portfolio conditioned only on beta for the $\gamma = 10$ investor dominates the indices.

- **Complementing BtM, size, mom:**
  - In 9 of 10 cases, adding beta to these three does not significantly change the CE.
  - The one case where it does: $\gamma = 10$, Rolling Protocol: size, BtM, and mom: 0 ($-15, 14$), add beta: 30 ($14, 43$). (This is the maximum CE for the $\gamma = 10$ investor.)
In isolation:
- Big difference between rolling and updating protocols.
- Ex: $\gamma = 3$: Updating: 50 ($-13, 88$) (not sig. different from EWI).
  Rolling: -336 ($-1984, -3$).
- Updating: Not sig. different from higher index for $\gamma = 3, 5$.
  Sig. less than higher index for $\gamma = 10, 15, 20$.

Complementarities and substitutabilities
- Rolling protocol, Adding to other characteristics lowers CE:
  - Significantly for $\gamma \leq 10$.
  - Insignificantly for $\gamma > 10$.
- Updating protocol: For all $\gamma$ no sig. change. Always reduces precision.
  - Ex.: $\gamma = 3$, Updating: mom only: 71 ($57, 83$); add lsmr: 11 ($-137, 76$).
Lagged same-month return

- This characteristic is most positively affected by using a higher $\gamma$.
- Both individually and in combination.
  - For both the $\gamma$ 3 investor using $\gamma$ of 5 to optimize, and the $\gamma$ 5 investor using $\gamma$ of 10 to optimize, individually the same-month return characteristics can be used to significantly increase CE above the indices.
  - Additionally, this characteristic–while deleterious to CE when using your own $\gamma$–is very efficacious in combination with other characteristics when using higher $\gamma$.

Model 5, Updating Protocol, $\gamma$ 3 investor.

<table>
<thead>
<tr>
<th>$\gamma$ Used</th>
<th>B-S 2.5%ile</th>
<th>B-S Median</th>
<th>B-S 97.5%ile</th>
</tr>
</thead>
<tbody>
<tr>
<td>EWI</td>
<td>83.38</td>
<td>85.79</td>
<td>88.27</td>
</tr>
<tr>
<td>3</td>
<td>2.58</td>
<td>138.63</td>
<td>206.27</td>
</tr>
<tr>
<td>5</td>
<td>192.40</td>
<td>225.35</td>
<td>259.61</td>
</tr>
</tbody>
</table>
Model 5, Updating Protocol, $\gamma$ 5 investor.

<table>
<thead>
<tr>
<th>$\gamma$ Used</th>
<th>B-S 2.5%ile</th>
<th>B-S Median</th>
<th>B-S 97.5%ile</th>
</tr>
</thead>
<tbody>
<tr>
<td>VWI</td>
<td>47.37</td>
<td>52.47</td>
<td>57.60</td>
</tr>
<tr>
<td>5</td>
<td>-11.48</td>
<td>87.10</td>
<td>131.46</td>
</tr>
<tr>
<td>10</td>
<td>120.87</td>
<td>145.03</td>
<td>166.13</td>
</tr>
</tbody>
</table>
Unconditional market loading

- \{\text{Mom, size, BtM}\}: Insig. different from 1 for all \(\gamma\), and both protocols.

- Conditioning on beta:
  - In isolation:
    - Sig. \(< 0\) for all \(\gamma\) and both protocols.
  - Adding to other characteristics:
    - All cases: lowers market loading significantly.
    - Ex.: \(\gamma = 3\), Updating:
      \{\text{Mom, size, BtM}\}: 1.07 (1.00, 1.14)
      \{\text{Mom, size, BtM, beta}\}: 0.30 (0.13, 0.47)

- Conditioning on lagged same-month returns:
  - In isolation:
    - Sig. \(> 1\) for all \(\gamma\) and both protocols, except \(\gamma = 20\), Updating.
    - Shrinking in \(\gamma\).
Unconditional market loading

- Conditioning on lagged same-month returns:
  - Adding to other characteristics:
    - Adding to \{mom\} and \{mom, size, BtM\}: increases significantly for \(\gamma \leq 15\)
    - No sig. effect for \(\gamma = 20\)
      - Both protocols.
    - Adding to \{mom, size, BtM, beta\}: No sig. effect (all \(\gamma\), both protocols).

- Conditioning on \{mom\}:
  - Slightly > 1 for \(\gamma \leq 10\), both protocols.
  - Insig. different from 1 for \(\gamma = 15, 20\).

- Conditioning on \{size\}:
  - Updating:
    - All insig. different from 1.
  - Rolling:
    - All sig. > 1.
Unconditional market loading

- Conditioning on \{BtM\}:
  - Sig < 1 in all 10 cases.
  - Ex.: $\gamma = 3$, Updating protocol: 0.93 (0.88, 0.98)
Unconditional SMB loading

- **Conditioning on \{\text{beta}\}**: 
  - Sig. < 0 in all 10 cases.
  - Ex.: \( \gamma = 3 \), Updating protocol: 
    \(-0.69 \) (\(-0.80\), \(-0.59\)).

- **Conditioning on \{\text{size}\}**: 
  - Sig. > 0, for \( \gamma = 3 \).
  - Ex.: Updating Protocol: 
    0.62 (0.38, 0.85).
  - Sig. < 0 for \( \gamma \geq 10 \).
  - Ex.: \( \gamma = 10 \), Updating Protocol: 
    -0.18 (\(-0.25\), \(-0.10\)).

- **Conditioning on \{\text{BtM}\}**: 
  - Sig. > 0 for \( \gamma \geq 5 \)
  - Ex.: Updating Protocol, \( \gamma = 3 \): 
    0.25 (0.15, 0.35).
  - Insig. different from 0 for \( \gamma \geq 10 \).
Unconditional SMB loading

- Conditioning on \{mom\}:
  - Sig. $< 0$ for $\gamma \geq 10$.
  - Sig. $> 0$ for $\gamma \leq 5$, Rolling protocol
    Ex.: $-0.69 (-0.80, -0.59)$.

- Conditioning on \{size\}:
  - Sig. $> 0$, for $\gamma = 3$.
  - Ex.: Updating Protocol:
    $0.62 (0.38, 0.85)$.
  - Sig. $< 0$ for $\gamma \geq 10$.
  - Ex.: $\gamma = 10$, Updating Protocol:
    $-0.18 (-0.25, -0.10)$.

- Conditioning on \{BtM\}:
  - Sig. $> 0$ for $\gamma \geq 5$
  - Ex.: Updating Protocol, $\gamma = 3$
    $0.25 (0.15, 0.35)$.
  - Insig. different from 0 for $\gamma \geq 10$. 

Unconditional SMB loading
Unconditional SMB loading

- Conditioning on $\{r_{t-11}, \bar{r}_{t-11}\}$:
  - Sig. < 0 for $\gamma \geq 10$.
  - Sig. > 0 for $\gamma \leq 5$, Rolling protocol
    Ex.: -0.69 ($-0.80, -0.59$).

- Conditioning on $\{\text{size}\}$:
  - Sig. > 0, for $\gamma = 3$.
  - Ex.: Updating Protocol:
    0.62 (0.38, 0.85).
  - Sig. < 0 for $\gamma \geq 10$.
  - Ex.: $\gamma = 10$, Updating Protocol:
    -0.18 ($-0.25, -0.10$).

- Conditioning on $\{\text{BtM}\}$:
  - Sig. > 0 for $\gamma \geq 5$
  - Ex.: Updating Protocol, $\gamma = 3$:
    0.25 (0.15, 0.35).
  - Insig. different from 0 for $\gamma \geq 10$. 
Unconditional HML loading

- Conditioning on \( \{ r_{t-11}, \bar{r}_{t-11} \} \):
  - Sig. \(< 0\) for \( \gamma \leq 10 \).
  - Insig. different from 0 for \( \gamma \geq 15 \)
    - Both protocols.
    - Ex.: \( \gamma = 3 \), Updating: -0.60 (-0.76, -0.46).

- Conditioning on \( \{ \text{size} \} \):
  - Sig. \( > 0\), for \( \gamma \leq 5 \).
  - Ex.: Updating Protocol, \( \gamma = 3 \):
    0.17 (0.10, 0.23).

- Conditioning on \( \{ \text{BtM} \} \):
  - Sig. \( > 0\) for all \( \gamma \), both protocols.
  - Large drop from \( \gamma = 3 \): 2.4, (2.1, 2.7) to \( \gamma = 10 \): 1.1, (1.0, 1.2).
  - Flat as \( \gamma \) increases above 10.
  - Protocols very similar.
Focus on $r_{t-11}, \bar{r}_{t-11}$

- Summarizing, optimal portfolios conditioned on lagged same-month returns for $\gamma = 3$, have:
  - 1.5 loading on market risk premium.
  - -0.3 loading on SMB.
  - -0.6 loading on HML.
  - 0.1 loading on MOM.

- So, it appears that the positive alpha on this strategy is distinct from the traditional sources of risk.

- Keloharju, Linnainmaa, and Nyberg (2013) suggest that the lagged same-month effect has to harness known risk factors.

- Intuition: stocks that did well last January are likely small stocks, and because they are small, they are also likely to do well next January.

- We examine this by analyzing the FFC regressions for each month of the year.
## FFC factors by month

<table>
<thead>
<tr>
<th>Factor</th>
<th>All months</th>
<th>Jan</th>
<th>Feb</th>
<th>Mar</th>
<th>Apr</th>
<th>May</th>
<th>Jun</th>
<th>Jul</th>
<th>Aug</th>
<th>Sep</th>
<th>Oct</th>
</tr>
</thead>
<tbody>
<tr>
<td>$R_m - r_f$</td>
<td>0.63** (0.21)</td>
<td>1.33 (0.86)</td>
<td>0.19 (0.67)</td>
<td>0.92 (0.65)</td>
<td>1.42* (0.62)</td>
<td>0.80 (0.63)</td>
<td>0.42 (0.57)</td>
<td>-0.02 (0.69)</td>
<td>0.20 (0.85)</td>
<td>-0.87 (0.77)</td>
<td>-0.05 (1.11)</td>
</tr>
<tr>
<td>SMB</td>
<td>0.29 (0.14)</td>
<td>1.50* (0.52)</td>
<td>1.04 (0.68)</td>
<td>0.32 (0.59)</td>
<td>0.07 (0.48)</td>
<td>0.39 (0.44)</td>
<td>0.92 (0.48)</td>
<td>-0.54 (0.46)</td>
<td>-0.16 (0.45)</td>
<td>0.16 (0.37)</td>
<td>-1.39* (0.56)</td>
</tr>
<tr>
<td>HML</td>
<td>0.36* (0.14)</td>
<td>1.09 (0.65)</td>
<td>0.70 (0.69)</td>
<td>1.04* (0.36)</td>
<td>0.66 (0.46)</td>
<td>0.02 (0.34)</td>
<td>-0.16 (0.44)</td>
<td>0.78 (0.49)</td>
<td>0.61 (0.41)</td>
<td>0.24 (0.41)</td>
<td>-0.57 (0.48)</td>
</tr>
<tr>
<td>MOM</td>
<td>0.64** (0.21)</td>
<td>-1.47 (0.97)</td>
<td>1.74* (0.70)</td>
<td>0.62 (0.58)</td>
<td>-1.04 (1.10)</td>
<td>-0.27 (0.64)</td>
<td>2.14** (0.65)</td>
<td>1.06 (0.47)</td>
<td>-0.34 (0.49)</td>
<td>1.96** (0.56)</td>
<td>0.59 (0.69)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Factor</th>
<th>Nov</th>
<th>Dec</th>
</tr>
</thead>
<tbody>
<tr>
<td>$R_m - r_f$</td>
<td>1.65* (0.74)</td>
<td>1.55* (0.55)</td>
</tr>
<tr>
<td>SMB</td>
<td>0.40 (0.45)</td>
<td>0.82* (0.36)</td>
</tr>
<tr>
<td>HML</td>
<td>-0.45 (0.54)</td>
<td>0.30 (0.46)</td>
</tr>
<tr>
<td>MOM</td>
<td>0.75 (0.80)</td>
<td>1.95* (0.68)</td>
</tr>
</tbody>
</table>
Factor exposures by month

- So, there is significant variation across the months in FFC factors.
- Look at FFC regressions separately in each of the 12 months (38 observations).
- Conditioning on \( \{r_{t-11}, \bar{r}_{t-11}\} \), \( \gamma = 3 \), Updating:
  - HML loading:
    - Lowest loading in Nov: -2.5 (-3.16, -1.91).
    - Highest loading in Mar: 1.2 (0.65, 2.50).
  - We have 2 alphas:
    - From the full 456 month FFC regression.
    - Aggregating over the 12 month-of-the-year regressions.
  - The latter minus the former is the measure of bias due to ignoring month-dependent loadings.
  - This can also be decomposed into the bias due to each of the four factors.
  - We have bootstrap distributions on all of these functions.
Continuing with $\{r_{t-11}, \bar{r}_{t-11}\}$ $\gamma = 3$, Updating

- Unconditional alpha in this case: 162.3 (119.5, 207.9) (bps per month).
- Biases in unconditional alpha from:
  - $R_m - r_f$: 7.6 (1.2, 14.2).
  - SMB: 22.4 (13.4, 32.1).
  - MOM: 10.6 (14.2, 28.4).
  - **Total**: 61.8 (43.3, 81.5).
Biases in unconditional alphas

- Largest bias: $\gamma = 3$, Rolling, $x \in \{ \text{mom, BtM, size, } r_{t-11}, \bar{r}_{t-11} \}$
- Biases in unconditional alpha from:
  - $R_m - r_f$: 20.0 (5.1, 35.4).
  - SMB: 47.8 (26.2, 71.0).
  - HML: 23.4 (7.6, 40.0).
  - MOM: -7.2 (−32.01, 18.2).
  - Total: 84.0 (40.5, 129.3).
The information in lagged same-month returns can be used to increase expected returns.

Adding it to traditional characteristics increases expected returns and Sharpe ratios.

This information is useful to increase expected utility if the investor takes account of model uncertainty.

The alternative loss functions is another response to *pseudo-out-of-sample* evidence.

Keloharju, Linnainmaa, and Nyberg (2013) are correct--conditioning on this information increases exposure to traditional sources of systematic risk.
The sampling variation—in both $\theta$ coefficients and CE—varies widely across various $x$.

- This paper has used statistical comparisons.
- What are the normative implications for a risk-averse investor?
- Two approaches using machine learning (still no likelihood):
  - Bootstrap Aggregation (Bagging / Bragging).
  - Data Augmentation.
- Generally the DA approach is better than BA.
- Largest effect is on larger models (almost no effect on those with just one characteristic).
- Model 4 does beat Model 9 statistically using DA: M4 CE: 183 (127, 231); std: 146 (36, 213).
But even these gains in certainty equivalent pale in comparison to the gains achieved by using a higher \( \gamma \).

So the next step might be to combine these.

Also since using a higher \( \gamma \) works so well, one possibility is a lack of stationarity. But updating generally dominates rolling. This motivates using the machine learning / Data augmentation with the rolling procedure. (Trade-off estimation noise and non-stationarity.)