SAME SAME BUT DIFFERENT
Most of us will be familiar with the experience of driving in a large metropolitan area. If you are familiar with the city, you can typically estimate with some precision how long it will take to drive from point A to point B.

On average it may take forty-five minutes to commute from home to work if you leave at 7am. Sometimes construction, weather, or some unknown factor causes the trip to take an hour. On other days the trip takes just twenty minutes. Sometimes there is an obvious reason for the trip to take more or less time than usual, like unusual weather, construction, or the mid-summer lull. Far more often, the difference in timing seems completely random.

For this reason, when experienced drivers estimate driving times they usually think in terms of a range. It’s sensible to expect the commute to take between thirty minutes and one hour most of the time. This allows drivers to set expectations that include an error term. If you have an 8am meeting at the office, you’d better leave at 7am to ensure that you get there on time.

While it seems natural to think that driving in traffic is a highly random phenomenon, many investors find it hard to translate this way of thinking to their investments. For example, many strategies are offered in different formats. We offer our Adaptive Asset Allocation strategy at several volatility targets, with different amounts of leverage, and in several different structures. We also run it on both futures and ETFs.

The goal of this article is to illustrate how seemingly inconsequential changes to the trading mechanics of a strategy, which have little impact on the long-term expected performance, can have a material impact on results in the short-term.

To explore this concept we will examine the results of simulations based on the exact same underlying strategy, with exactly the same universe of investments, but where a change is made to just one minor variable. Specifically, we will see how small differences in rebalance frequency can have negligible impact on long-term results, but can lead to performance differences of 10 percentage points or more over one year observation horizons.
DATA

We use daily total return data for the following asset classes from 1991 through July 2018:

- US stocks - VTI ETF extended with S&P 500 Index
- Canadian equities – EWC ETF extended with S&P TSX Composite
- European stocks - VGK ETF extended with S&P Europe BMI
- Asian stocks - VPL ETF extended with S&P Asia Pacific BMI
- Emerging stocks - VWO ETF extended with S&P Emerging BMI
- Treasuries - IEF ETF extended with S&P US Treasury 7-10 Year TR Index
- T-Bonds - TLT ETF extended with S&P US Treasury Bond 20+ Year TR Index
- Commodities - DBC ETF extended with Deutsche Bank Liquid Commodity Index
- REITs - IYR ETF extended with Cohen & Steers US Realty Shares, Inc. Class I
- International REITs - RWX ETF extended with Cohen & Steers Int’l Realty Shares, Inc.
  Class I and Morgan Stanley International Real Estate Fund (MSUAX)
- Gold - GLD ETF extended with continuous gold futures

Table 1: Performance summary table. Simulated results.

<table>
<thead>
<tr>
<th>Inception</th>
<th>Commodities</th>
<th>Canada</th>
<th>Gold</th>
<th>Treasuries</th>
<th>REITs</th>
<th>IntREITs</th>
<th>T-Bonds</th>
<th>Europe</th>
<th>Asia</th>
<th>US Stocks</th>
<th>Emerging</th>
</tr>
</thead>
<tbody>
<tr>
<td>1991-01-02</td>
<td>4.88%</td>
<td>7.99%</td>
<td>4.24%</td>
<td>6.24%</td>
<td>7.48%</td>
<td>5.40%</td>
<td>7.58%</td>
<td>8.05%</td>
<td>3.82%</td>
<td>10.50%</td>
<td>5.23%</td>
</tr>
<tr>
<td>1992-01-02</td>
<td>4.88%</td>
<td>7.99%</td>
<td>4.24%</td>
<td>6.24%</td>
<td>7.48%</td>
<td>5.40%</td>
<td>7.58%</td>
<td>8.05%</td>
<td>3.82%</td>
<td>10.50%</td>
<td>5.23%</td>
</tr>
<tr>
<td>1993-10-08</td>
<td>4.88%</td>
<td>7.99%</td>
<td>4.24%</td>
<td>6.24%</td>
<td>7.48%</td>
<td>5.40%</td>
<td>7.58%</td>
<td>8.05%</td>
<td>3.82%</td>
<td>10.50%</td>
<td>5.23%</td>
</tr>
<tr>
<td>1994-12-29</td>
<td>4.88%</td>
<td>7.99%</td>
<td>4.24%</td>
<td>6.24%</td>
<td>7.48%</td>
<td>5.40%</td>
<td>7.58%</td>
<td>8.05%</td>
<td>3.82%</td>
<td>10.50%</td>
<td>5.23%</td>
</tr>
</tbody>
</table>

| Compound Return | 4.88%       | 7.99%    | 4.24%  | 6.24%      | 7.48% | 5.40%    | 7.58%   | 8.05%  | 3.82%| 10.50%    | 5.23%    |
| Volatility      | 1.94%       | 20.73%   | 15.98% | 6.18%      | 23.11%| 17.12%   | 11.72%  | 20.19% | 20.19%| 17.70%    | 24.28%   |
| Sharpe Ratio    | 0.24        | 0.37     | 0.21   | 0.67       | 0.35  | 0.31     | 0.5     | 0.38   | 0.18 | 0.53      | 0.28     |
| Max Drawdown    | -74.00%     | -60.80%  | -45.60%| -10.40%    | -74.10%| -73.60%  | -26.60% | -63.60%| -55.50%| -55.40%   | -67.70%  |
| Positive Rolling Yrs | 62.06 | 69.06 | 62.17 | 81.8 | 75.8 | 66.27 | 79.81 | 70.09 | 59.95 | 81.53 | 66.45 |

Source: Analysis by ReSolve Asset Management. Data from CSI Data and S&P. Simulated and hypothetical data. Past performance is no guarantee of future results.
We investigate the long-term and short-term performance dispersion for a simple version of our Adaptive Asset Allocation strategy. At each rebalance period, portfolios are formed by maximizing the expected Sharpe ratio of the portfolio, where returns are estimated based on momentum characteristics. Covariances are estimated based on the sample covariance matrix of returns in the recent past.

We run simulations where portfolios are rebalanced at each daily frequency between 1 and 20 days. Some simulations rebalance every day; others are rebalanced every second day; still others are rebalanced every 20 days; and every frequency in between. In addition, at each rebalance period the portfolios are scaled to target a 10% annualized volatility (0.63% daily standard deviation) based on the same covariance estimate that we use to form portfolios. We cap total exposure at 200%. Table 3 describes the results of our twenty simulations executed over the period 1991 - June 2018.

---

1 Specifically, at each rebalance we form an ensemble of portfolios optimized on momentum estimated from many randomly sampled sets of lookback windows. We estimate covariance using an Exponentially Weighted Moving Average of returns, with a decay factor that is drawn randomly and varies between 0.93 and 0.97.
LONG-TERM RESULTS

Table 3: Performance of Adaptive Asset Allocation strategy at 10% target volatility, rebalanced at frequencies between 1 and 20 days, sorted by Sharpe ratio. Simulated results.

<table>
<thead>
<tr>
<th></th>
<th>Worst</th>
<th>5th %ile</th>
<th>25th %ile</th>
<th>Median</th>
<th>75th %ile</th>
<th>95th %ile</th>
<th>Max</th>
<th>Multi</th>
</tr>
</thead>
<tbody>
<tr>
<td>Compound Return</td>
<td>14.11%</td>
<td>14.16%</td>
<td>14.59%</td>
<td>14.89%</td>
<td>14.99%</td>
<td>14.66%</td>
<td>15.07%</td>
<td>14.61%</td>
</tr>
<tr>
<td>Volatility</td>
<td>10.54%</td>
<td>10.49%</td>
<td>10.55%</td>
<td>10.50%</td>
<td>10.25%</td>
<td>9.85%</td>
<td>9.90%</td>
<td>10.10%</td>
</tr>
<tr>
<td>Sharpe Ratio</td>
<td>1.29</td>
<td>1.29</td>
<td>1.32</td>
<td>1.35</td>
<td>1.39</td>
<td>1.42</td>
<td>1.45</td>
<td>1.38</td>
</tr>
<tr>
<td>Max Drawdown</td>
<td>-19.32%</td>
<td>-19.52%</td>
<td>-18.95%</td>
<td>-22.85%</td>
<td>-20.74%</td>
<td>-18.36%</td>
<td>-18.53%</td>
<td>-19.20%</td>
</tr>
<tr>
<td>Positive Rolling Yrs</td>
<td>89.00</td>
<td>88.81</td>
<td>90.54</td>
<td>89.85</td>
<td>90.00</td>
<td>90.26</td>
<td>91.06</td>
<td>90.37</td>
</tr>
</tbody>
</table>

Source: Analysis by ReSolve Asset Management. Data from CSI Data and S&P. Refer to methodology description above and in footnote 1 above. Simulated and hypothetical data. Past performance is no guarantee of future results.

Figure 1: Growth of $1 invested in Adaptive Asset Allocation strategies rebalanced at frequencies from 1 through 20 days. Simulated results.
Over the 27 years covered in our simulation, the gap between the 5th and 95th percentile outcome was 16 basis points of Sharpe.

Some investors might be tempted to conclude that this is a meaningful difference. However, using the method from (Lo 2002) the standard error of the Sharpe ratio for our simulations is 20 basis points of Sharpe.

The average Sharpe ratio across the twenty tests is 1.36, suggesting that we would expect 95% of equivalent strategies to produce Sharpe ratios between 1.03 and 1.69. The realized Sharpe ratios from all of the simulations fall well inside this range, suggesting that we can’t reject the null hypothesis that all of the strategies are drawn from the same return distribution.

Indeed, a test of equal Sharpe ratios across all strategies, as described in (Pav 2016), produces a p-value of 0.51, making it clear that all of the strategies have produced statistically indistinguishable performance.

**SHORT-TERM RESULTS**

The analysis above established that varying the rebalance frequency from 1 through 20 days made no difference to the long-term performance of our strategy. Of course, very few investors are willing to invest in a strategy and walk away for 27 years.

Rather, most investors allocate to a strategy and then seek to evaluate performance over much shorter horizons. In our experience, the actual evaluation horizon for most investors is measured in months, not years.

And this is where the analysis of long-term results breaks down. The fact is, even for strategies with exactly the same expected long-term performance, the short-term can look astonishingly different. And those differences can be astonishingly large. Figure 2 plots the difference between the 95th and the 5th percentile compound returns for the 20 simulations across all rolling 252-day (1-year) periods.
The average difference between the best (95th percentile) and worst (5th percentile) simulation was 7.2 percentage points over any given 252-day period. The 95th percentile difference was 10.9. But even when the strategies were tracking closely (5th percentile difference) we still observed a 4.2 percentage point dispersion between the best and worst performing rebalance frequency.

Note that the difference between the best and worst simulations narrows over longer periods. Per Table 4 the average annualized return difference between best and worst is about 7 percent over rolling one-year periods, but shrinks to just 3 percent over five-years and 2.3 percent over ten-year horizons.
Table 4: Rolling differences in annualized compound returns between the best and worst performing Adaptive Asset Allocation strategies rebalanced at frequencies between 1 and 20 days, at various observation horizons.

<table>
<thead>
<tr>
<th></th>
<th>1-Year</th>
<th>3-Years</th>
<th>5-Years</th>
<th>10-Years</th>
<th>20-Years</th>
</tr>
</thead>
<tbody>
<tr>
<td>5th percentile</td>
<td>4.2%</td>
<td>2.7%</td>
<td>1.7%</td>
<td>1.3%</td>
<td>1.3%</td>
</tr>
<tr>
<td>Median</td>
<td>7%</td>
<td>3.9%</td>
<td>3%</td>
<td>2.3%</td>
<td>1.6%</td>
</tr>
<tr>
<td>95th percentile</td>
<td>10.9%</td>
<td>6.5%</td>
<td>4.5%</td>
<td>3.4%</td>
<td>2.1%</td>
</tr>
</tbody>
</table>

Source: ReSolve Asset Management. Simulated results.

Figure 3 plots the equity growth curves for the nine calendar years that exhibited the highest return dispersion amongst our 20 simulations. The largest dispersion was observed in 2004, followed closely by 2008. These are years with material intra-year drawdowns for all of the strategies, but where certain strategies happened to get very lucky or very unlucky with rebalance dates.

Figure 3: Performance of best and worst performing strategy permutations in years with the largest dispersion. Simulated results.
Source: ReSolve Asset Management. Each chart represents the growth of $1 for the worst and best performing strategy permutation each the calendar year. Simulated and hypothetical data. Past performance is no guarantee of future results.
CONCLUSION

The objective of this short article was to disabuse investors of the notion that short-term dispersion in results between similar strategies should prompt concern.

We simulated 20 Adaptive Asset Allocation strategies over a period of 27 years, which varied only by rebalance frequency, and demonstrated that the performance results were statistically indistinguishable over the full period. We then examined the results of the same simulations over rolling one-year windows, and showed that they often exhibit a large dispersion in performance over the short-term.

While some people may find this level of randomness uncomfortable, the good news is that investors should have no cause for concern if they observe short-term dispersion in results between similar strategies. Small differences in execution can cause material differences in short-term results purely due to good or bad luck. But over the long-term these differences will almost always average out.

REFERENCES
