Surprise, Endowments Have Significantly Outperformed the 60/40 Benchmark

Hossein Kazemi

Isenberg School of Management & The CAIA Association

Final Version September 2020

For Updated Results see the Appendix

Introduction

Several recent papers have claimed that endowments have significantly underperformed the "passive" 60/40 equity/bond investment strategy. Further, they have identified allocations to alternative asset classes as the primary reason behind this underperformance. In this concise note, I demonstrate that these results are not robust and entirely dependent on the assumptions made. Also, the results are likely flawed because they ignore some of the realities of managing an endowment (e.g., an endowment may have as much as 10% in cash over extended periods). The findings and conclusions of this note are summarized below:

- The claim that endowments have underperformed the 60/40 portfolio is a function of using an improper benchmark. Using a globally diversified "60/40 passive" benchmark that accounts for the cash holdings of endowments, I show that endowments have outperformed the benchmark since 2000. More specifically, different cohorts of endowments have outperformed the benchmark by 2.2%-3.8% per year since 2010.
- Endowments may have underperformed the 60/40 benchmark employed in these studies, despite their allocations to alternatives. Since 2000, a diversified portfolio of alternatives has *outperformed* the US 60/40 benchmark over every 10-year window. In fact, the figures suggest that endowments should have increased their allocations to alternatives. If there is any underperformance, the blame lies somewhere else.
- Some papers use Sharpe style analysis to measure the alphas of endowments. While this approach is quite useful in identifying the investment style of an endowment on an ex-post basis, it cannot and should not be used to create a benchmark to measure skill. As mentioned by the CFA Institute, a benchmark should be specified in advance based on the knowledge available at that time. This means the US or global 60/40 portfolio could be used as a benchmark. However, benchmarks created after the fact using the ex-post performance of the manager should not be used to evaluate the manager. In a sense, using Sharpe style to evaluate a manager involves using the manager's skill to create a benchmark, which is used to measure her skills.
- Some endowments lack the needed skills and connections to construct an alternative portfolio that could deliver the "average" return of hedge fund and private equity indices. However, this should not lead one to recommend that endowment managers must dump alternatives in favor of the 60/40 portfolio, especially now with rates at historical lows and equity valuations at historical highs. Endowment managers should be urged to use these indices as a rough benchmark. If their alternative portfolios have underperformed these benchmarks in the past and are likely to do so going forward, then they should reconsider their approaches to managing their alternative assets.

- In the updated results displayed in Appendix B, I have replicated the approach and the benchmarks used in Ennis (2020a). It turns out the estimated alphas that appear in that paper become positive if 2019 data is added. That is, using the same rolling 10-year approach used in Ennis (2020), I obtain positive alphas for all three cohorts of endowments. Also, there is a positive relationship between estimated alphas and the size of endowments' allocations to alternatives and the size of the endowment.

I want to focus on a recent paper by Hammond (2020), which uses the S&P500 Index and Barclays US Aggregate Bond Index to construct the 60/40 portfolio. Some of the results presented here will have some bearings on the recent papers by Ennis (2020a and 2020b), which blame alternatives for the alleged underperformance. Hammond (2020) assumes that back in the 1980s, the 1990s, and so on, investing in US indices using a specific rebalancing method was the proper passive investment strategy. Let us look at these assumptions one at a time

1. Truly passive strategies should have considered global market-cap-weighted equity and bond indices. An endowment manager is not fulfilling her fiduciary responsibilities by ignoring the diversification benefits of global investing. Diversification is called the only free lunch in financial markets for good reasons. Endowments have accepted these and have always allocated to global bonds and equities. Even if they did not, the proper passive benchmarks for endowments with broad mandates should include globally diversified portfolios. If endowments choose to target global allocations that are below or above the global benchmarks, then that should be considered when measuring an endowment manager's skill.
2. One could argue that a passive strategy is actually the buy and hold strategy rather than one that rebalances every month or every year. If we are choosing an active strategy of rebalancing, then why not daily or 5-year rebalancing? As we will see, it will make a difference. The timing of rebalancing could make a difference, as well. For instance, if an endowment does annual rebalancing, ex-post, the result will be different depending on when the rebalancing is performed within each year.
3. The paper assumes that an endowment's funds are invested the moment a donor contributes to the university or the college. Also, it assumes that the entire endowment is unrestricted and is available to be invested. Of course, endowments have many restrictive funds, and the investment process may take several weeks or months. NACUBO survey of endowments reports that more than 50% of their assets have some restrictions. Some donations are made in the form of securities or other assets. These contributions cannot be invested immediately. Most donations have restrictions regarding the use of the assets, which impact how those funds are deployed. In short, an endowment is not an unrestricted pile of cash that can be fully invested under a uniform asset allocation strategy. Further, NACUBO reports that endowments hold around 2.5%-4.1% in cash. These facts cannot be ignored.
4. Most papers use indices that are not investable products. While fees have declined in recent years, they were meaningful in the 1980s and 1990s. On the other hand, while endowments report net of external fees performance, they ignore the effects of internal fees.

One important message of this note is that because we have so few observations regarding endowments' performance, their estimated risk-adjusted returns change drastically depending on the period selected, the length of the window used to estimate the parameters, and the benchmarks selected. In short, the results are not robust. Therefore, researchers should be very careful in drawing

strong conclusions from their results. This word of caution applies to the results reported here, too. They could change drastically, depending on the period and the length of the window selected to estimate the parameters. Finally, I estimated the alphas of endowments using the US benchmarks while applying the Sharpe style approach. The estimates were either positive or insignificantly different from zero during the most recent periods, reinforcing my point that the results reported here or in other papers are not robust.

Performance versus Proper Benchmarks

So, what happens if one makes the correct assumptions (I am ignoring #4)? Below is a table from Hammond (2020). Somewhat similar results are presented in Ennis (2020a) and Ennis (2020b), where US aggregate bond rather than the global bond index is used. The global index has underperformed the US index by about 30bp each year since 1990.

Exhibit 1: Performance of Endowments and the US 60/40 Benchmark

	Endowments			Benchmark
	Small Cohort	Average	Larger Cohort	60/40 US Indices
FY 1990-1999	12.3	12.9	14.0	14.6
FY 2000-2009	3.9	4.0	6.1	1.5
FY 2010-2019	7.7	8.4	9.0	10.5

Source Hammond (2020), Exhibit 21, Page 17

Given these results, it appears that the endowment managers have done a poor job during the last 30 years. But have they? Also, what is the reason for this underperformance of the US 60/40 portfolio? Is because endowments have allocated too much to alternatives?

Below is the same table where I have added four columns using MSCI World Free and Barclays Global Aggregate Total Return Unhedged Indices to construct the global 60/40 portfolio, and a portfolio of alternative assets (see Appendix A for details). I have ignored the cash holding issue for now and have assumed that the entire fund is invested.

Exhibit 2: Performance of Endowments, Various 60/40 Benchmarks & Alternatives

	Endowments			US Indices	Global Indices			Alternative Assets Portfolio
	Small Cohort	Average	Large Cohort	60/40	60/40 Monthly Rebalancing	60/40 Buy & Hold	60/40 5-Year Rebalancing	Quarterly Rebalancing
FY 1990-1999	12.3	12.9	14.0	14.6	10.8	10.6	9.8	21.4
FY 2000-2009	3.9	4.0	6.1	1.5	1.6	1.3	3.1	8.9
FY 2010-2019	7.7	8.4	9.0	10.5	7.7	7.4	6.1	10.9

Source Hamond (2020), Bloomberg, and Author Calculations

Two important conclusions can be drawn from this table:

- Endowments have significantly outperformed a globally diversified benchmark. This can be seen by comparing the figures under Global Indices with the performances of endowments. Almost every group of endowments has outperformed the global benchmarks over every subperiod.
- Endowments have underperformed the US 60/40 benchmark despite their allocations to alternative assets. This can be seen by comparing the performance of the US 60/40 portfolio with that of the Alternative Assets Portfolio. Over every subperiod, the alternative portfolio has outperformed the US 60/40 portfolio. Clearly, alternatives are unlikely to be a source of underperformance. Perhaps, endowments should increase their allocations to alternatives!

We have not considered the impact of cash holdings. Comparing endowments' performance with a 60/40 portfolio assumes that the entire fund is fully invested at all times and that there are no cash holdings or any restrictions on funds. So, what happens when the effects of cash holdings are considered?

I assume that endowments hold about 4% of their holdings in short-term US Treasury Bills. The following table reports those results.

Exhibit 3: Performance of Endowments and Various Equity/Bonds/Cash Benchmarks

	Endowments			US Indices	Global Indices and T-Bills		
	Small Cohort	Average	Large Cohort	60/40	Monthly Rebalancing*	Buy & Hold*	5-Year Rebalancing*
FY 1990-1999	12.3	12.9	14.0	14.6	10.6	10.4	9.3
FY 2000-2009	3.9	4.0	6.1	1.5	1.6	1.4	2.2
FY 2010-2019	7.7	8.4	9.0	10.5	7.4	7.2	5.2

Source Hamond (2020), Bloomberg, and Authors Calculations

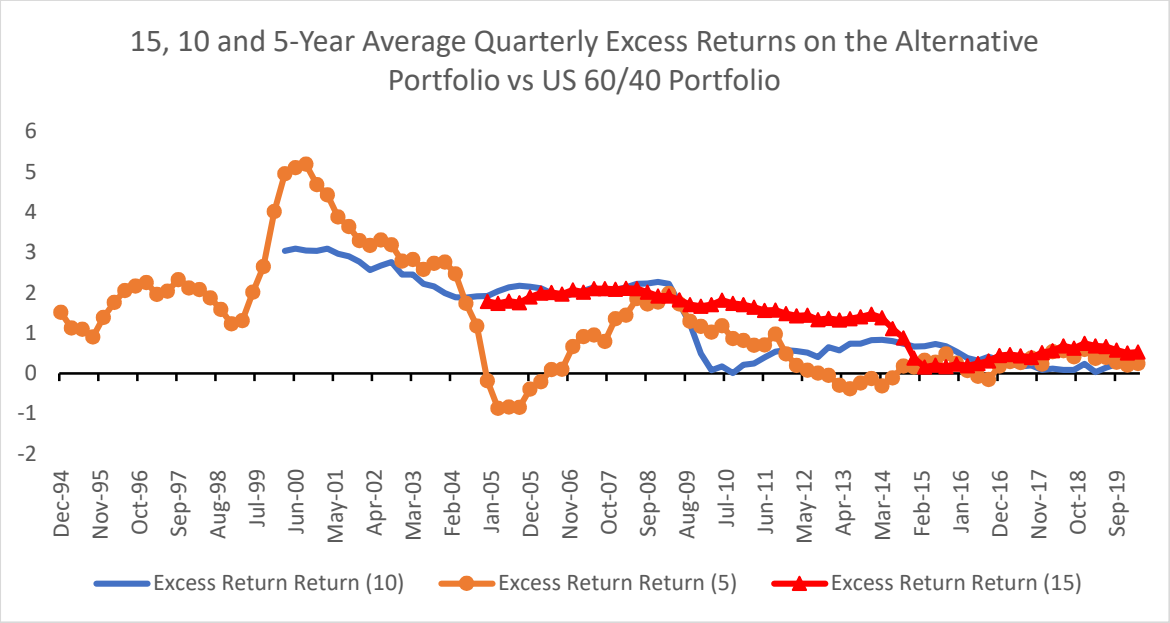
* Portfolios are 57.6/38.4/4, which means 4% in T-Bills

Well, the numbers speak for themselves. I think endowment managers deserve a raise. Using the global benchmark and assuming 5-year rebalancing, endowments have outperformed the benchmark by 2.2%-3.8% per year since 2010. By the way, for those who would argue that the under-diversified US markets were the right benchmark, I would counter why not use the Apple stock as a benchmark?

Was it Ever Rational to Replace Alternatives with the US 60/40 Portfolio?

As mentioned in the introduction, some recent papers (e.g., Ennis 2020a and 2020b) have blamed allocations to alternatives as to the primary reason for the alleged underperformance of endowments. Let us see if, since 2000, there was ever a time that statistical evidence indicated that alternatives should have been abandoned in favor of a US 60/40 portfolio. Given the results presented above, there is no point in considering the proper benchmark of the global 60/40 portfolio as endowments have significantly outperformed this benchmark since 2000.

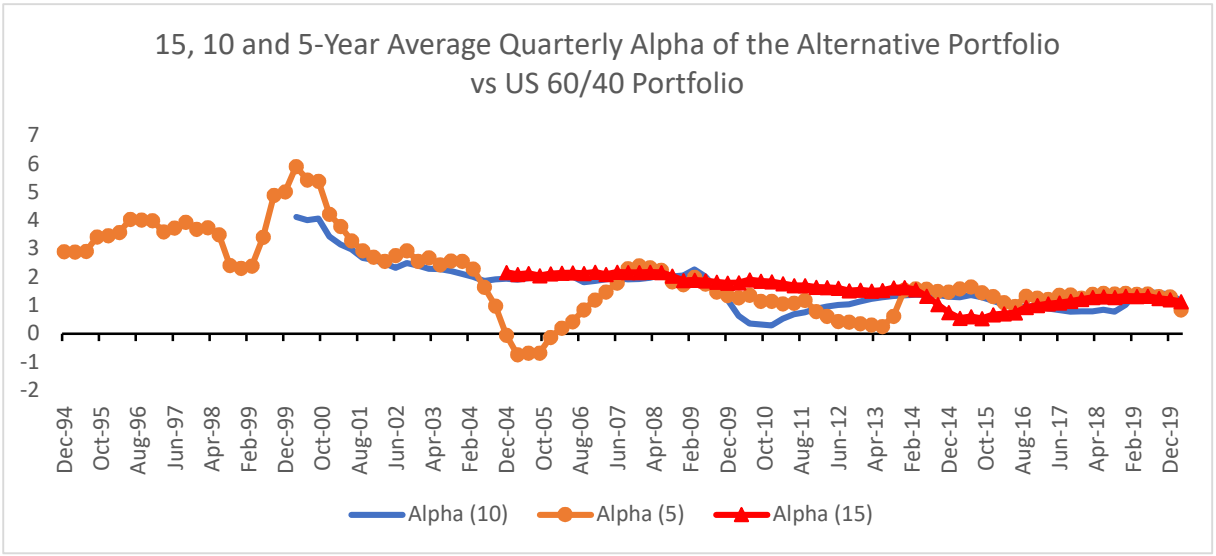
First, I examine the rolling 5-, 10-, and 15-year average *excess* quarterly return of the alternative portfolio versus the US 60/40 portfolio.



Sources: CISDM, Cambridge Associates, Bloomberg, and Authors Calculations

On a 5-year rolling basis, there were a few occasions that the average excess return became negative. On the other hand, both the 10-year and the 15-year average excess returns never dropped below zero. Given that the average life a typical venture or private equity fund is about seven years, and the very long-term time horizons of endowments, I would argue that the 10-year and 15-year averages are more relevant. While the simple excess return is intuitive, it does not adjust for the riskiness of the alternative portfolio. Below, I present two sets of results that adjust for risk and volatility.

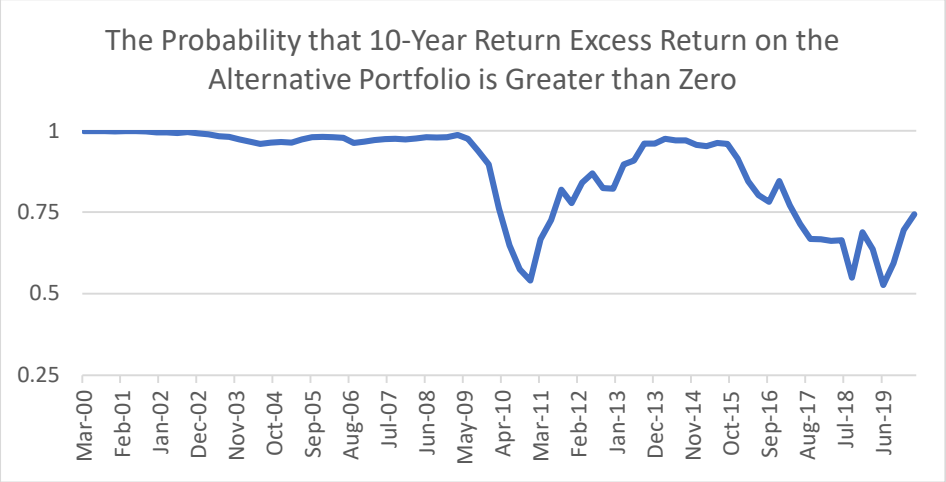
The following chart displays the 5-, 10-, and 15-year alphas of the alternative portfolio versus the US 60/40 portfolio.



Sources: CISDM, Cambridge Associates and Authors Calculations

We can see that there was only one short period in 2004-2005 that the alpha estimated using a 5-year window became negative. Again, at no point, the alphas measured over rolling 10-year or 15-year windows came close to zero. By the way, please note that these are quarterly alphas.

Finally, let us consider an endowment manager that, at the end of each quarter, looks at the 10-year excess return of the alternative portfolio versus the US 60/40 benchmark and asks a simple question. Given the excess return observed over the last ten years, what is the probability that the actual mean excess return on the alternative portfolio is greater than zero? These probabilities are displayed below:



Sources: CISDM, Cambridge Associates and Authors Calculations

We can see that at no point in time since 2000 this probability dropped below 50%. That is, there was no point during this period that a "rational" endowment manager would have looked at the 10-year return on alternatives and would have seen any evidence that the alternative portfolio is likely to underperform the US 60/40 portfolio.

In the face of this evidence, why would an endowment manager give up on alternatives? The estimated odds have always been in favor of the alternative portfolio. Of course, these odds would increase significantly if the proper global benchmark were used.

Caveats

The analysis of the previous section was conducted using CISDM Equally Weighted Hedge Fund Index, Cambridge Associates Venture Index, and Cambridge Associates Private Equity Index. These indices are not investable, and their performances are sure to differ from the performances of the alternative buckets of many endowments. Also, we can see that while the alternative portfolio has maintained its alpha, its excess raw return has declined in recent years.

There is no argument that some endowments lack the needed skills and connections to construct an alternative portfolio that, at the minimum, would deliver the "average" return represented by these indices. However, the sound recommendation to the entire endowment managers community is not to dump alternatives in favor of the 60/40 portfolio, especially now with rates at historical lows and equity valuations at historical highs. The sound recommendation is to urge endowment managers to use these indices as a rough benchmark. If their alternative portfolios have underperformed these benchmarks in

the past and are likely to do so going forward, then they should reconsider their approach to manager selection and allocation in the alternative space.

Appendix A

The alternative portfolio was created using equal allocations to private equity, venture capital, and hedge funds (represented by Cambridge Associates and CISDM Equal Weighted Index). The traditional benchmark was created using a 60% allocation to the S&P 500 Index and a 40% allocation to the Barclays US Aggregate Bond Index. All portfolios were rebalanced each quarter.

The alpha of the alternative portfolio was calculated using a simple regression

$$R_t(A) - R_t(TBill) = \alpha + \beta \times (R_t(B) - R_t(TBill)) + e_t$$

Here, $R_t(A)$ is the quarterly return on the alternative portfolio, $R_t(B)$ is the quarterly return on the benchmark, $R_t(TBill)$ is the 30-day TBill rate, and e_t is the error term. This regression was run using the rolling windows to calculate the 5-, 10-, and 15-year alphas.

I am aware that returns on alternative indices suffer from data smoothing. Even if one were to assume that the beta of the alternative portfolio was close 1 (a rather extreme assumption), the results would hold (this can be seen from the first graph of this section).

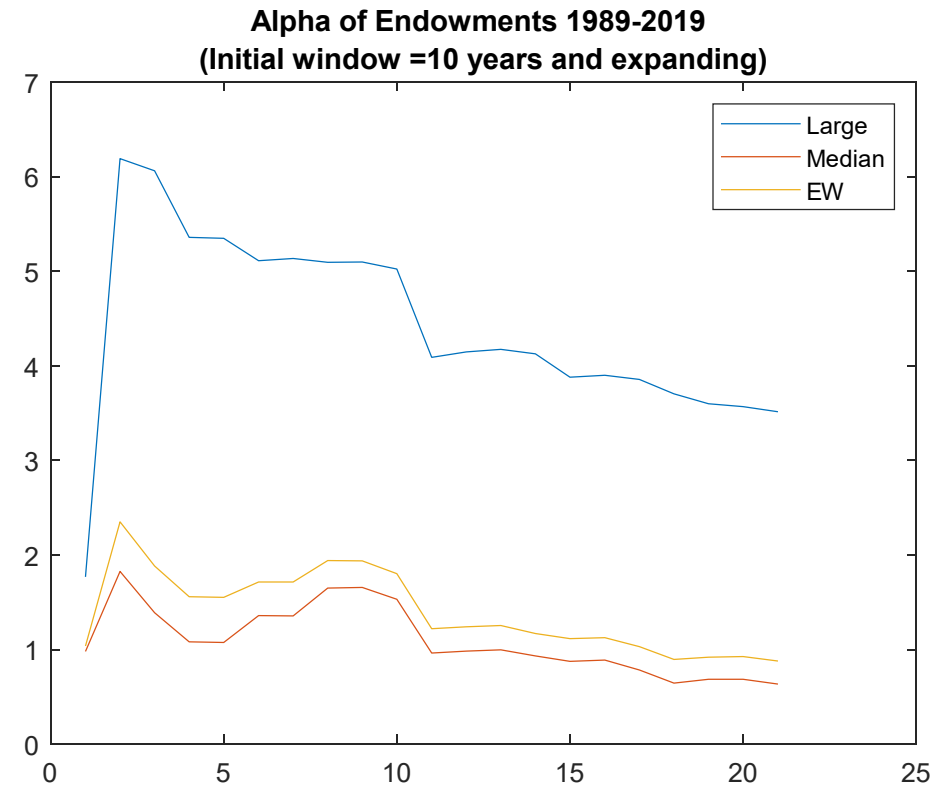
Appendix B

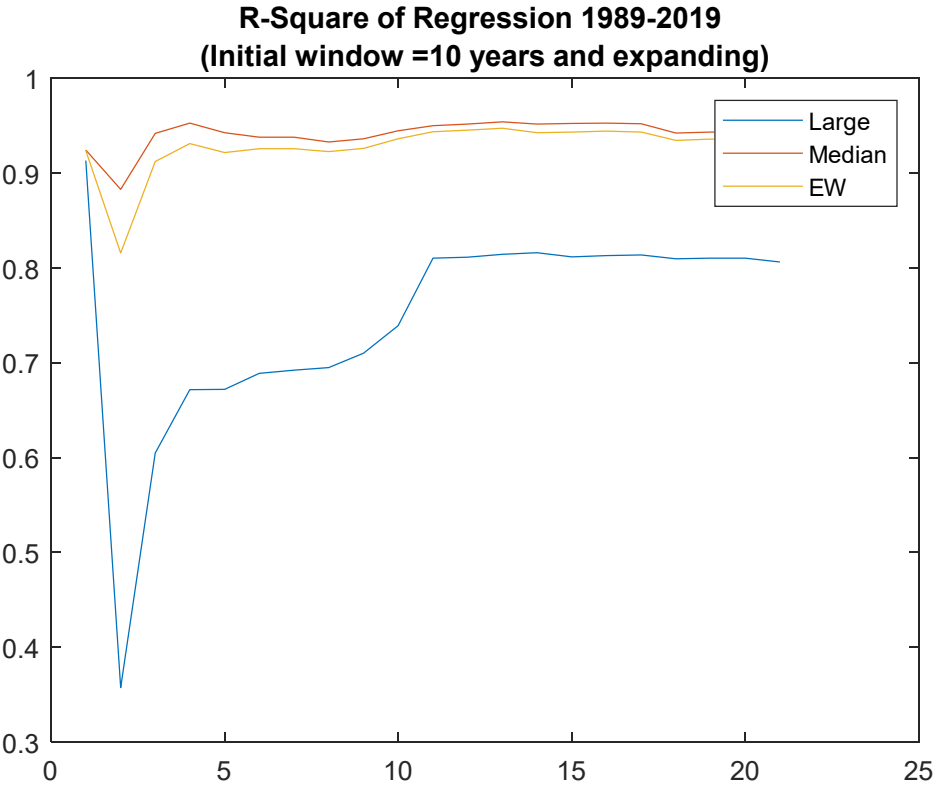
I attempted to replicate the exact procedure followed by Ennis (2020a). In particular, I ran the following regression

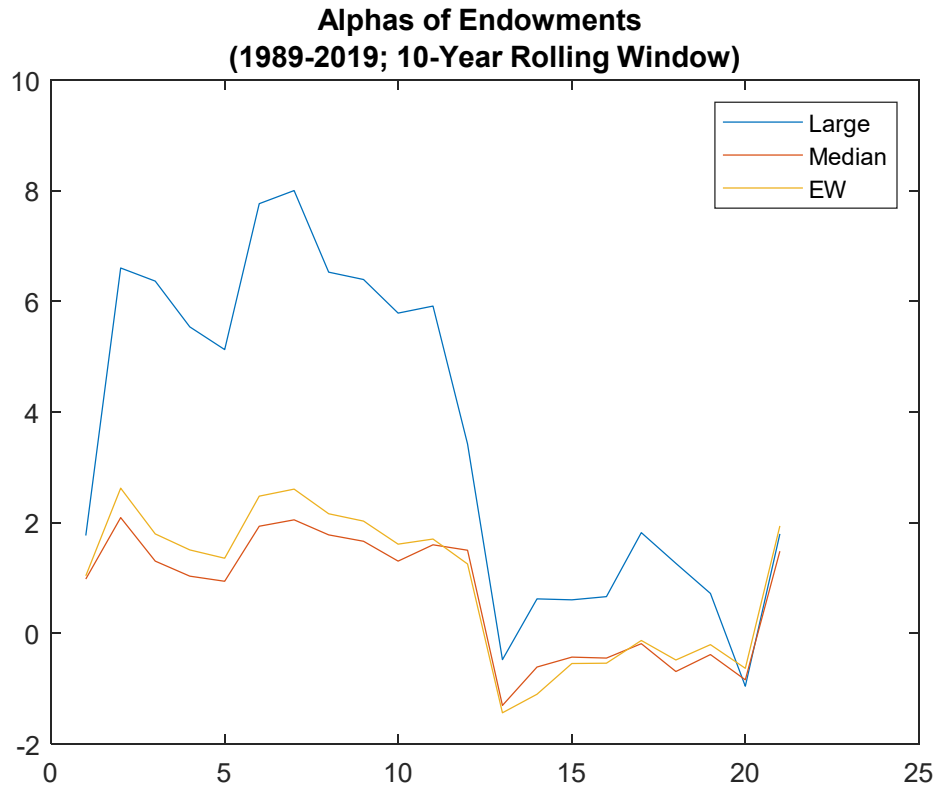
$$R_t(A) - R_f = \alpha + \beta_1 \times (R_t(MSCIxUS) - R_f) + \beta_2 \times (R_t(MSCI) - R_f) + \beta_2 \times (R_t(RUS) - R_f) + e_t$$

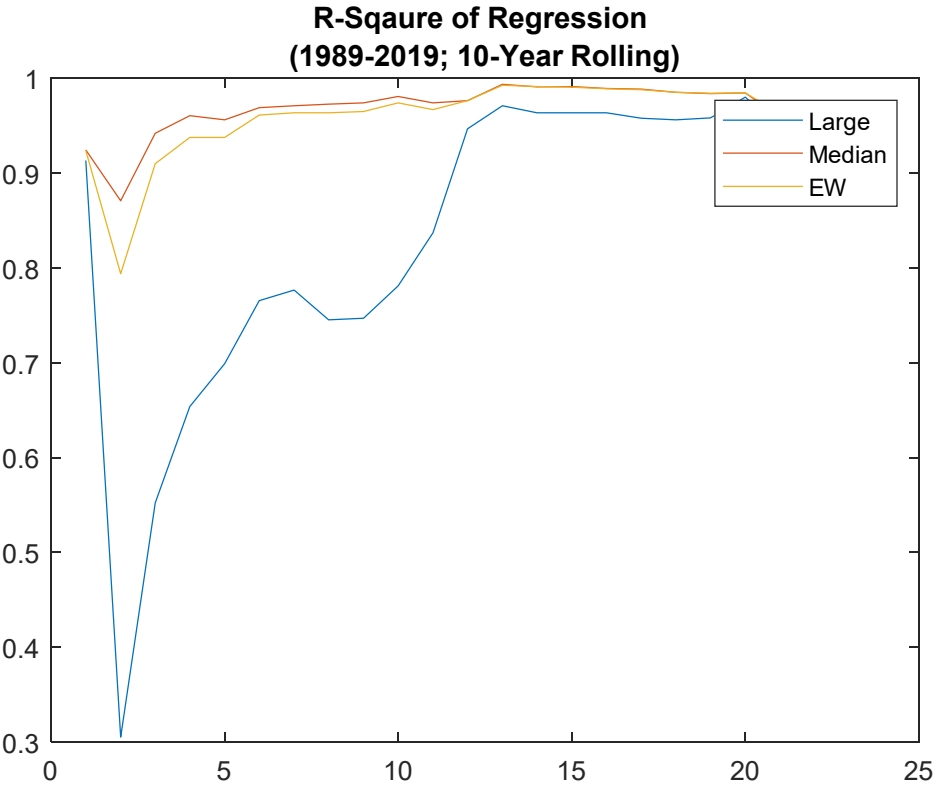
Here $R_t(MSCIxUS) =$ total return on MSCI Ex US index; $R_t(USAgg) =$ total return on US Agg index; $(R_t(RUS) =$ total return on Russell 3000 index and $R_t(TBill) =$ 30-day US TBill rate. These are the same benchmarks that are used in Ennis (2020).

Here are the results


Alpha of Endowments 1989-2019
(Initial window =10 years and expanding)
Large
Median
EW
0
1
2
3
4
5
6
7
0
5
10
15
20
25


R-Square of Regression 1989-2019
(Initial window =10 years and expanding)
Large
Median
EW
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
0
5
10
15
20
25


Alphas of Endowments
(1989-2019; 10-Year Rolling Window)
Large
Median
EW
10
8
6
4
2
0
-2
0
5
10
15
20
25


R-Sqaure of Regression
(1989-2019; 10-Year Rolling)
Large
Median
EW
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0
5
10
15
20
25

References

Hammond, Dennis, 2020, "A Better Approach to Systematic Outperformance? 58 Years of Endowment Performance," The Journal of Investing August 2020, 29 (5) 6-30

Ennis, Richard M., 2020a, "Institutional Investment Strategy and Manager Choice: A Critique," The Journal of Portfolio Management Fund Manager Selection 2020, 46 (5) 104-117

Ennis, Richard M., 2020b, "Endowment Performance," https://ssrn.com/abstract=3614875