Alpha and Performance Efficiency of Ivy League Endowments: Evidence from Dynamic Exposures

Apollon Fragkiskos
MPI

Sean Ryan
MPI

Michael Markov
MPI

Introduction

As of June 2017, the total AUM of university and college endowments was estimated in excess of $560 billion.[1] These funds are used to generate income in order to satisfy current and future operational expenses of their affiliated universities, with annual effective spending rates over 4 percent.[2] They have received much attention in the recent past for their superior investment returns compared to other institutional investors [Lerner, Schoar, and Wongsungwai, 2007]. With a combined $124 billion AUM[3], the Ivy League endowments[4] demand a lot of attention, given their impressive track record of mostly double-digit returns over the past two and half decades (Exhibits 1 and 2, next page), showing that the group has outperformed both historical payouts of 5 percent, even after an average inflation of 3 percent is added to those, and a 60-40 portfolio[5] by a wide margin.

Ivy League fund performance has been associated with their increasing allocations to private asset classes (real estate, private equity and hedge funds). These funds find it easier to invest in such assets, as they can afford managers and consultants with great expertise (Dimmock and Stephen 2012). Yale, in particular, held 69 percent of its assets in 2006 in real estate, private equity and hedge funds [Lerner et al. 2007]. As documented in the literature, such heavy weightings toward private or alternative asset classes largely explains why Ivy League endowments have enjoyed large positive returns in the past.

However, even the largest endowment funds were not immune to the recent financial crisis. The Ivy group experienced losses that exceeded 20 percent in 2009. Such losses have important policy implications because university endowments are typically forced to reduce payout rates during negative financial shocks [Brown et al. 2010]. Significant losses also put into question the endowment model as advocated by David Swensen, known as the 'Yale model'.[6]

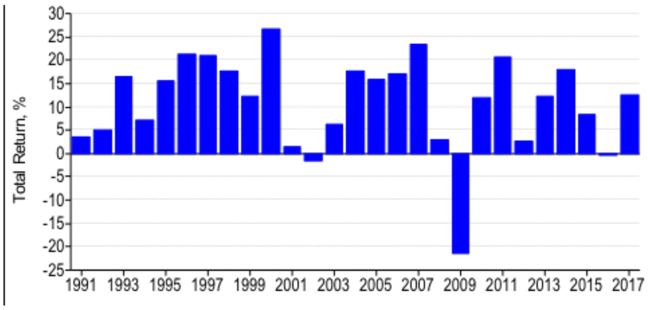


Exhibit 1: Annual Returns of Ivies Shown by Fiscal Year
Source: Nacubo
Note: The returns displayed for each year correspond to fiscal years that end June 30 as opposed to December31. This applies to all subsequent charts that display endowment returns

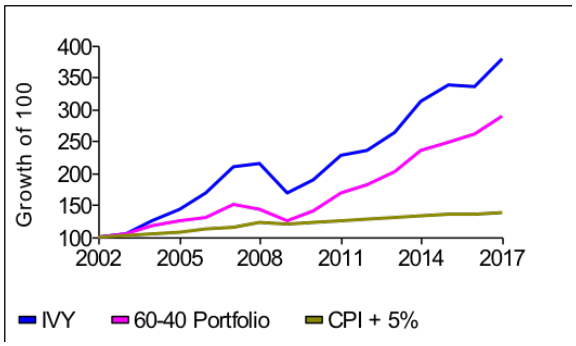


Exhibit 2: Cumulative Growth of Ivies vs. 60-40 Portfolio Shown by Fiscal Year
Source: Nacubo, Bloomberg

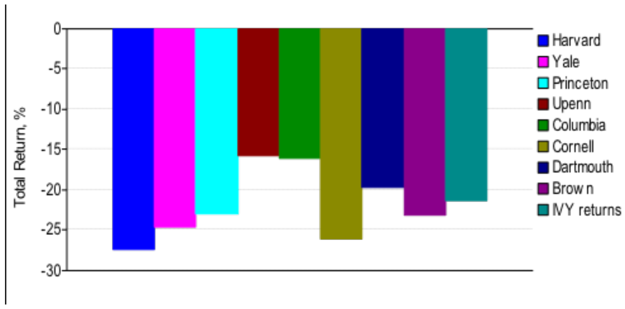


Exhibit 3: Performance of Ivies During Fiscal Year 2009
Source: Endowment Annual Reports

Motivation

The publicly available asset allocations that the funds disclose tend not to use traditional asset class breakdowns. They typically group investments by other attributes such as "Independent Return," "Real Assets" and "Absolute Return." This type of non-standardized disclosure makes endowment performance comparisons difficult. The fact that alternative investments occupy a large portion of Ivy League asset allocations further exacerbates this problem. In this paper we seek to shed light on the alpha-generating abilities of Ivy League endowments and the financial risks[7] they assume over time in order to evaluate their performance efficiency. The results will be derived at the individual endowment level and then aggregated bottom up in order to arrive at conclusions about Ivy League endowments as a group. We will also supply results based on the Ivy index for reference and particularly given that this index allows us to go back further in history.

Our analysis is based on investment returns experienced by the funds, which are reported on an annual basis. While these go back to 1988 across endowments, broken down by size of AUM and type of institution (public or private)[8], publicly available returns on individual Ivy League endowments are harder to obtain and typically only go back a few years. This data limitation makes returns-based analysis using traditional methods such as the one put forward by Sharpe (1992), very challenging.

To overcome this limitation, we perform style analysis based on market indices that correspond to allocations made by endowments (Private and Public Equity, Fixed Income, Commodities, Hedge funds). We used a Dynamic Style Analysis model (Markov et al., 2004) that enables the calculation of alpha and betas using small data sets. Our approach allows us to create a portfolio of indices that mimics the returns of the funds.

Even though the endowment funds provide only annual data, the factor-mimicking portfolio is available at a higher frequency[9] than annual. This provides us with sufficient observations to infer the aggregate risk of specific endowment funds.[10]

We present results on alpha related to manager selection and market timing, alpha against common public indices, risk and performance efficiency in terms of Sharpe ratio.

Existing work

Alpha

There have been many studies that analyze endowment returns. Lerner, Schoar and Wang (2008) find that endowments earn strong excess returns relative to S&P 500. Chen (2016) finds that larger endowments, proxied by Harvard, Yale and Princeton, earn returns that are 8 percent higher than the smallest endowments. This is partially due to their ability to absorb fixed costs associated with illiquid asset classes that have higher expected returns, and partially due to their informational advantage as they have more money to hire the best talent. Brown, Garlappi, and Tiu (2010) used reported asset allocation weights and benchmark returns and found that the average endowment earns a negligible alpha. In particular, timing and security selection explained 14.59 percent and 8.39 percent of the variation of each endowment's returns, whereas asset allocation explained 74.42 percent of it. Barber and Wang (2013) analyzed grouped endowment returns for the Ivy and top SAT schools by regressing their annual returns against common benchmarks. They found no evidence of manager selection, timing, and tactical asset allocation abilities.

Risk

There are not many studies that calculate the investment risks endowments take. Chen (2016) reports that larger endowments may have lower risk aversion and thus be willing to invest more of their wealth to riskier asset classes. Using reported allocations and using market indices as proxies, he calculates the risk each endowment takes. Although this approach doesn't account for the fact that the same asset class will have a different risk profile across endowments, it provides a good estimate of the risks endowments take in a boom period. The author finds that when the market is in boom, the larger the endowment, the higher the return it achieves. This disappears or even reverses in high-risk regimes, indicating that larger endowments take on more risk and that may or may not result in higher returns, depending on the regime. It also becomes evident that larger endowments allocate more into riskier asset classes. In terms of Sharpe ratio, the larger endowments do not display any advantage over smaller ones and, in fact, in some cases they show a disadvantage.

Endowment data

Endowment returns are reported annually, usually during September, three months after the end of the fiscal year. Individual endowment return series were mainly collected from annual reports that endowments publish over the years on their own websites. For our analysis, these go back to 2003.[11]

Exposures

Analyzing endowment returns presents many challenges due to aggregate changes in allocations to major asset classes mentioned above. Such changes may be due to a different perception of expected returns to each asset class as well as changes in the risks that endowments are willing to take. If endowment funds face non-tradable risks for example, then they will choose portfolios that best hedge those risks. In other words, high standard deviation of non-financial income is associated with safer portfolios (Dimmock and Stephen, 2012). Credit constraints, amount of research taking place in the university and a large proportion of university revenues coming from endowments all result in safer portfolios. And while the need for regular cash flows to affiliated universities means that liquidity is a concern, universities with greater selectivity that can raise tuition at will or universities with a high ratio of donations to fund size do not face large liquidity constraints, allowing them to invest capital in illiquid private asset classes.

These considerations will result in time varying exposures against factor sets consisting of major asset classes such as the ones we use in this paper. This means that traditional methods of regression analysis such as the one put forward by Sharpe (1992) are not well suited given they assume constant exposures over the period analyzed. A way to get around this is to perform rolling regressions over shorter windows within the entire analysis period. Given the available returns are limited to only twelve annual observations, this is not a viable approach. To alleviate such concerns, we use a dynamic modeling technique called Dynamic Style Analysis (DSA) that is designed to work with scarce data and allows us to detect the dynamics of asset-based exposures (Markov et al., 2004).

The set of indices we used to explain the return series of each fund was formed based on common asset classes disclosed in endowments' annual reports. Any particular endowment portfolio may have (small) investments outside this set or may target different types of private investments than the ones corresponding to the indices chosen above. The indices we used, however, correspond to the largest percentage allocations reported by endowments and serve as a comprehensive set based on which risk and return can be evaluated. The fact that we are using the same set of indices across funds also enables us to have a common base for comparison.

For real estate we used the Cambridge Associates Real Estate index. This index represents an aggregate of individual commercial property returns based on properties owned by funds that institutional investors invest in, such as closed end funds, commingled funds and funds that are of sufficient size.[12] For hedge funds we used the Eurekahedge 50 index.[13] This index avoids the selection and instant history bias of the commonly used HFRI Fund Weighted Composite Index used in many studies, contains limited survivorship bias and is comprised of funds with top AUM. This makes it more applicable as a benchmark for hedge fund investments made by large institutional investors vs. an index that includes small funds. As a result, that index should represent more closely the institutional investor experience. For buyout and venture capital we used the corresponding indices from Cambridge Associates,[14] which are constructed based on the underlying cash flows and Net Asset Values provided by the general partners. Cambridge Associates obtains data from limited partners and general partners who have raised or are raising capital. Therefore, it may be biased toward well-performing funds, which may reduce the calculated fund alphas. However, given the large coverage of the database, this bias is likely to be low. Since we did not have many data points available for regression analysis, we used a portfolio with equal weights to the buyout and venture capital indices in our regressions. Given we didn't have access to a private natural resources index, we used a public commodity index as a proxy. For the rest of the public factors, we used indices that correspond to asset classes endowments invest in.

Public Equity	MSCI World
Real Estate	Cambridge Associates Real Estate
Private Equity	50% Cambridge Associates Private Equity + 50% Cambridge Associates Venture Capital
Commodities	Bloomberg Commodity
Bonds and Cash	Bloomberg Barclays U.S. Aggregate Bond Index
Hedge Funds	Eurekahedge 50

Exhibit 4: Index Proxies by Asset Class

We follow Sharpe's original approach by constraining the coefficients of the regression to be positive and add up to one. The budget constraint essentially assumes that there is no implied leverage in the aggregate holdings of the fund compared to the indices used to analyze its returns. In the case of endowment funds, we don't know whether the managers they invest in take on positions that are more levered than the indices being employed in the analysis. The allocation to each asset class may also be more concentrated than the indices we have used or the funds may invest in riskier stocks or bonds than what the indices hold. Alternatively, funds may invest in long/short equity strategies, which can have an aggregate market exposure close to zero.[15] Although in the aggregate level we do not expect our exposures to act as leveraged as a whole, the analysis will show factor exposures that may differ from actual holdings as some asset classes may be more or less leveraged than others. The constraints we apply also have the effect of increasing stability and mitigate multicollinearity which is particularly useful in the presence of limited data.[16] We do not attempt to quantify any currency hedging that may take place against the equity or real estate portion of the portfolio, assuming that all foreign exposures are unhedged.

Model strength

Given the limited data availability, we use a powerful technique that avoids overfitting in order to calibrate the time varying properties of the model. This is based on MPI's proprietary cross validation statistic, called predicted R-squared.[17] Similar to R-squared, predicted R-squared is also used as an indicator of a model's explanatory power.

Endowment	R^2	Predicted R^2
Harvard	96.9	96.9
Yale	99.3	99.3
Princeton	99.4	99.4
Penn	97.7	97.7
Columbia	97.8	97.8
Cornell	98.0	98.0
Dartmouth	96.6	96.5
Brown	98.9	98.9

Exhibit 5: R-Squared

In our analysis, the predicted R-squared values are all high and above 96 percent, as shown in Exhibit 5, suggesting a high explanatory power.

Another indicator of a model's explanatory power is whether the fund and factor mimicking portfolio, including the alpha component, move together. In table A of the Appendix, we can see that each fund's cumulative growth is closely tracked.

The exposures estimated were also compared against the annual reports from the funds which show asset class breakdowns and were found to be very similar, providing further support to our results.

Exposure portfolios

Exhibit 6 shows the dynamic factor exposures obtained by our model,[18] and Exhibit 7 groups and displays the overall exposure to alternatives. What is immediately obvious is the large,[19] and for the most part increasing, exposure to alternatives across most endowments, driven mainly by increasing exposures to private equity.

Endowment performance

The performance of the Ivy League funds over the period analyzed has been rather impressive, with all of them beating the 60-40 portfolio. We hereby shed light on the ways these funds have been able to generate such returns.

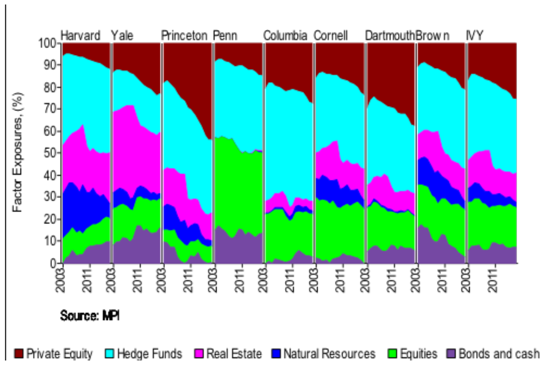


Exhibit 6: Factor Exposures of Ivies Shown by Fiscal Year
Source: MPI

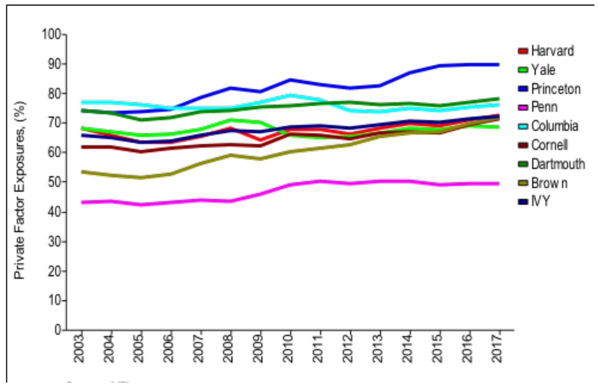


Exhibit 7: Exposures of Ivies to Private Investments Shown by Fiscal Year
Source: MPI

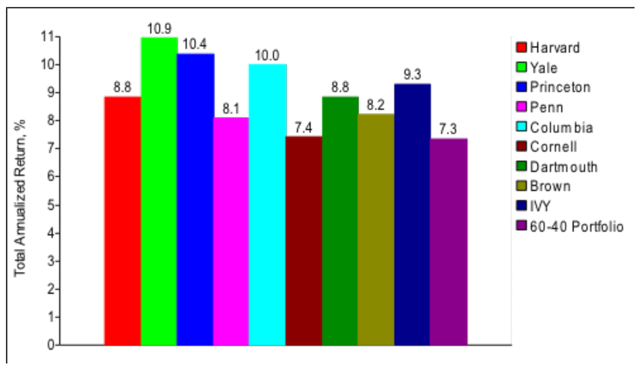


Exhibit 8: Total Annualized Return of Ivies and 60-40 Portfolio
Source: Nacubo

Timing

The first return component that we look at is called timing, and it is commonly used to measure the effectiveness of the portfolio's allocation decisions against a benchmark. In the absence of a well-defined benchmark[20] for the types of allocations the funds follow, we have used the average exposures of each mimicking portfolio over the analysis period as the benchmark. In order to calculate timing, we compared the returns of the mimicking portfolios against their average. This comparison provides us a measure of the return that was generated from shifts in asset class exposures over time. (Exhibit 9, next page) shows the annualized timing returns for the endowments. The timing returns are all small and negative, indicating that funds likely do not engage in market timing when measured on an annual basis and against the indices we use.[21]

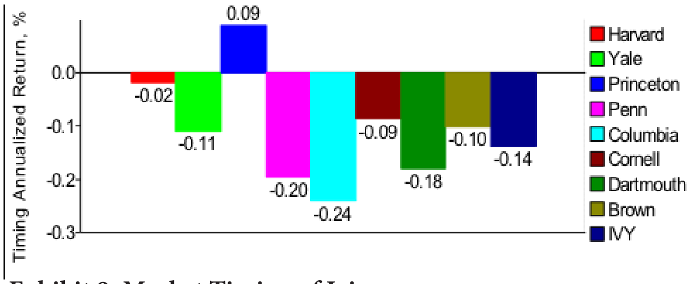


Exhibit 9: Market Timing of Ivies
Source: MPI

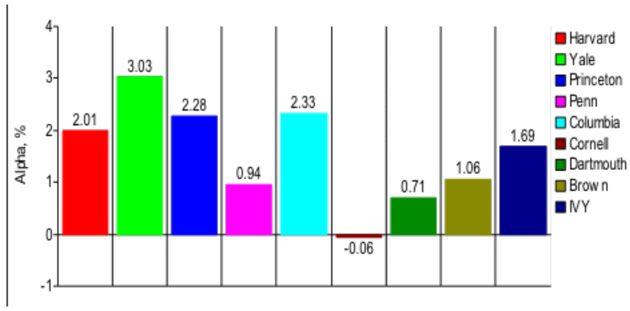


Exhibit 10: Private Alpha of Ivies
Source: MPI

Private alpha

We now turn to the question of whether endowment fund managers have superior alpha generation capabilities. This can arise from security selection, if the endowment manages their own investments, or selection generated by managers of the funds in which the endowment invests, in which case it speaks of the ability of the endowment team to select skilled managers.[22] For the rest of the document we will call this type of alpha private alpha.

In addition, returns arising from missing factors in the model could also impact private alpha. As Barber and Wang (2013) show in the case of regression with constant coefficients, any bias in the calculated alpha is given by the below relationship:

$$E(\alpha) = \beta_n \alpha_0$$

where α_0 is the alpha that the missing factor produces when regressed against the factors we used in the model and β_n is the beta of the fund to the missing factor when all other factors are included in the regression. A symptom of such a missing factor is low R^2. Given our set of indices spans the vast majority of asset classes that the Ivy League endowments invest in and that the R^2 of all regressions are high, it seems very unlikely that we have omitted a significant factor. Even if there is a missing factor, the beta against it should be small, resulting in a small alpha bias as the above equation.

While the aggregate index leads one to believe that Ivies do not generate superior returns, which is in agreement with the findings of Barber and Wang (2013), the results at the endowment level are mixed. Some endowments achieve private alphas above 2 percent, for the most part statistically reliable, while others do not.

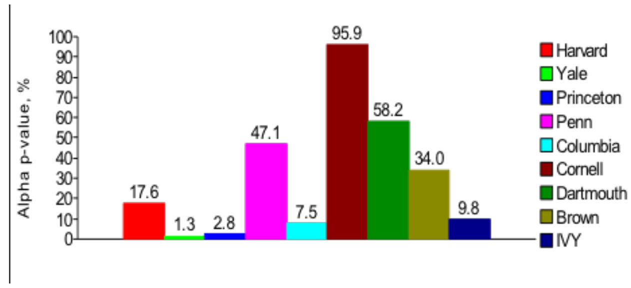


Exhibit 11: Private Alpha P-Values of Ivies
Source: MPI

Replication

To further evaluate the alpha-generating capabilities of the Ivy League funds, it would be of interest to find out if the factor-mimicking portfolios displayed in Exhibit 6 are able to generate similar returns out of sample.

Data

In order to have as much out-of-sample data as possible, the data we use for this exercise extends as far back as we can go depending on its fund's history, deviating from using a common and recent time period. This still leaves us with only a few data points to calculate out of sample statistics for some funds as shown in Exhibit 12. Given we do not have long history available for all funds, we also analyze the combined returns across all Ivy League funds as reported in Exhibit 1.

Endowment	History (number of years)	Out of sample (number of years)
Harvard	38	28
Yale	22	12
Princeton	18	8
Penn	32	22
Columbia	16	6
Cornell	18	8
Dartmouth	22	12
Brown	16	6
Ivy	28	18

Exhibit 12: Available History and Data Points for out-of-sample Estimations

Process

The replication takes place as follows. We use the first 10 years of an endowment's history to regress its returns against the public and private asset classes reported in Exhibit 4. Using the factor exposures obtained that correspond to the last year of the in-sample time period, we estimate the replicating portfolio returns over the next year by multiplying those exposures by the realized factor returns over the next year. We proceed repeating the same steps for each year going forward, each time using a 10-year rolling historical window for the estimation of the factor exposures.[23]

Exposures

Given the portfolios we are trying to replicate contain rather stable asset allocations, an indication that replication is reliable and avoids data overfitting is whether the exposures calculated at each time step are erratic or not. This has to be taken in context of the limited data we are working with, which will introduce sampling variability to a certain degree. Therefore, while we expect a rolling exposure calculation to produce more volatile exposures than a calculation that spans a longer time period, the change in exposures observed should not result in unreasonable turnover. We demonstrate what happens to the factor exposures of the replicating portfolio of the Ivy League index as an example. These are shown in Exhibit 13, which also shows for comparison the factor exposures of an in-sample analysis of the index over the entire period.[24] The exposures between the index and its replicating portfolio are similar to each other and consistent with the individual endowment exposures of Exhibit 6.

In Exhibit 14, we display the style dispersion for each endowment and replicating portfolio. For the most part, the style dispersion of the replicating portfolio is well controlled in relation to the in-sample dispersion.

Tracking

To first get a sense of whether endowments add value on top of their factor-mimicking portfolios, we examine how close each replicating portfolio tracks its corresponding endowment.

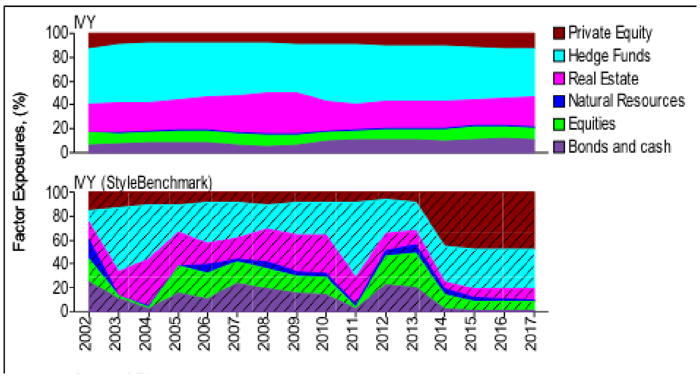


Exhibit 13: Factor Exposures of the IVY Index and its Replicating Portfolio (Style Benchmark) Shown by Fiscal Year
Source: MPI

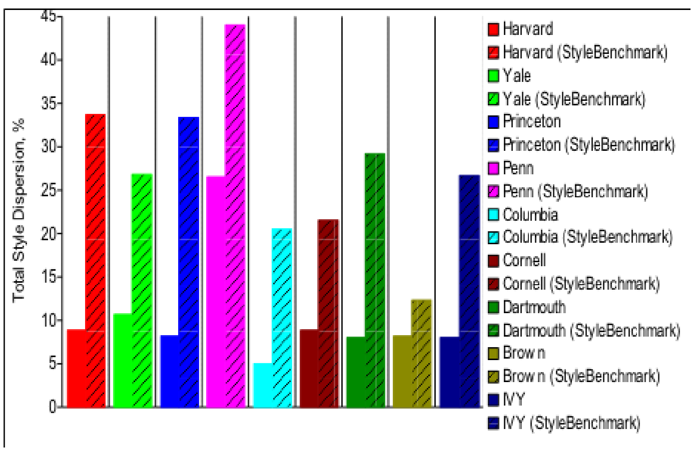


Exhibit 14: Style Dispersion of Ivies and their Replicating Portfolios.
Source: MPI

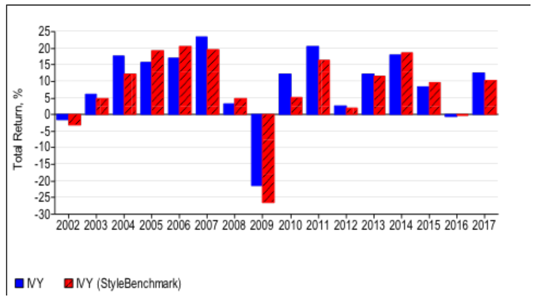


Exhibit 15: Yearly Returns of the Ivy League Index and its Replicating Portfolio Shown by Fiscal Year
Source: MPI

While a tight replication is not necessary, we should not observe deviations noticeably larger than the deviations of the in-sample factor-mimicking portfolios against their corresponding endowments.

Taking the Ivy League again as example, Exhibit 15 shows that the yearly returns between the index and its replicating portfolio are very close.

To quantify tracking across all endowments, in the next two charts we display statistics that summarize how well the endowment returns are being tracked. In Exhibit 16 we observe that the replicating portfolios achieve tracking errors that range between 1.9 percent and 4.8 percent. For the most part the results indicate close tracking, given the average standard deviation for the time period analyzed across all endowments is around 11 percent. With the exception of Cornell, the out of sample R^2 achieved across endowments in (Exhibit 17, next page) is at high to very high levels, in accordance to the R^2 values from Exhibit 5.

Excess return

The excess returns achieved by the funds are generally in line with their in-sample alphas from Exhibit 10. The excess return of Yale, the top performer, for example, against its replicating portfolio is 2.69 percent. By comparison, the average alpha reported in Exhibit 10 is 3.03 percent. Despite their recent 2017 under performance compared to the rest of Ivy endowments, Yale emerges on the top, followed by Harvard. The only endowments that break this rule are Columbia and Cornell, which show excess returns that flip sign in relation to the in-sample alphas. For both of these endowments, as per Exhibit 12, we had less than half the in-sample points available for excess return estimation though so the results are not directly comparable, rather they supplement each other.

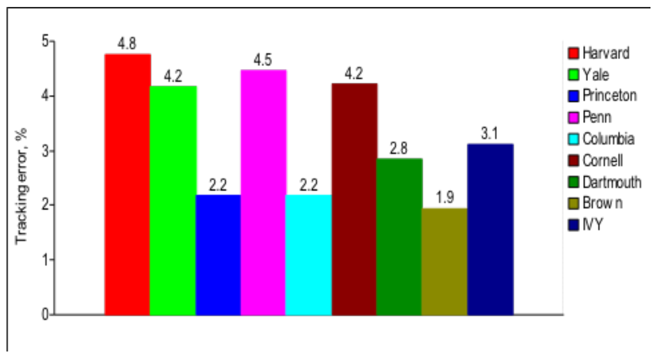


Exhibit 16: Tracking Error of Replicating Portfolios
Source: MPI

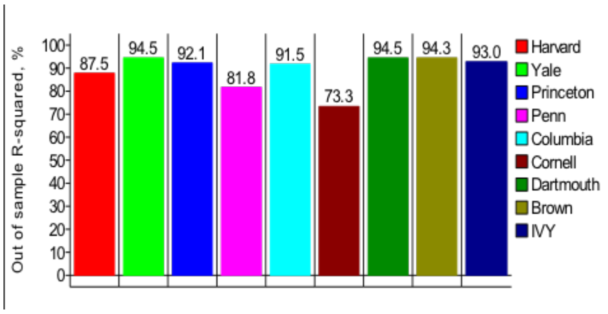


Exhibit 17: Out of Sample R^2 of Replicating Portfolios
Source: MPI

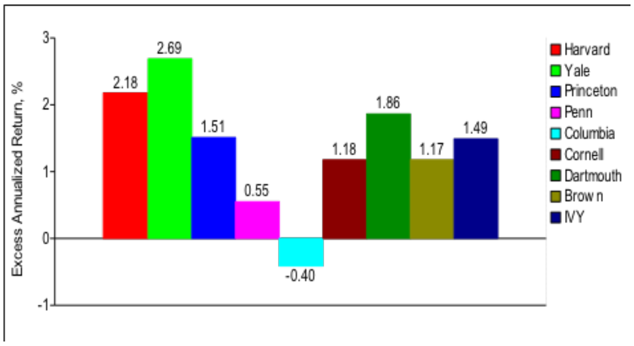


Exhibit 18: Excess Returns over Replicating Portfolios
Source: MPI

Unlike the in-sample results, in this case, none of the p-values of the excess returns are statistically significant. As shown in Appendix B, the excess returns are not just very volatile, but in many cases, they flip signs from one year to the next. This indicates that none of the Ivy League funds achieve reliably positive excess returns against the asset classes chosen.

Public alpha

Some of the asset classes endowments invest in, such as buyout, venture capital and hedge funds, have the potential of generating alpha against public indices, as shown in Exhibits 20 and 21 for buyout and venture capital. Although Exhibit 10 provides some clues of how true this is, for a more direct comparison against public indices, in particular, we omitted those from the analysis and reran the results. For the rest of the document we will refer to this alpha as public alpha.

The results were run with the same parameters as the original regression.[25] Previous studies (Woodward 2004) have found that the returns of private equity and real estate may depend on return lags that go beyond one year in the past. We attempted to include a one-year lag of each of the public indices, but this did not materially affect the results, therefore, the one-year lags were excluded.

Most funds now display large positive and significant public alphas. Interestingly, the rank of endowments in relation to their private alphas from Exhibit 10 is very close to the rank in relation to Exhibit 22. The public alphas of endowments have gone up about 2 percent in relation to the private alphas from Exhibit 10.

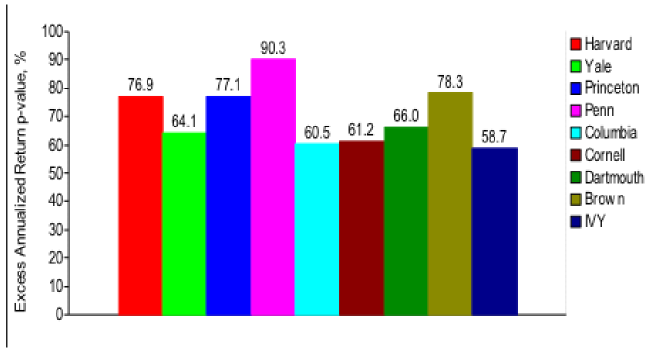


Exhibit 19: *p*-values of Excess Returns
Source: MPI
Note: The p-values have been produced using a two-tailed test

Beta	Annual alpha	Reference
2.2	1%	Buchner (2014)
2.2 – 2.4	8.3% - 8.6%	Axelson et al (2013)
1.39	-2%	Ang et al (2013)
1.4	0%	Franzoni et al (2012)
0.94	1.6%	Driessen et al (2008)
1	-3%	Phalippou et al (2009)
1	-0.1%	Jegadeesh et al (2009)

Exhibit 20: Alpha and Beta of Buyout

Beta	Annual alpha	Reference
2.6	3.5%	Buchner (2014)
2.57	-8.3%	Driessen et al (2008)

Exhibit 21: Alpha and Beta of Venture Capital

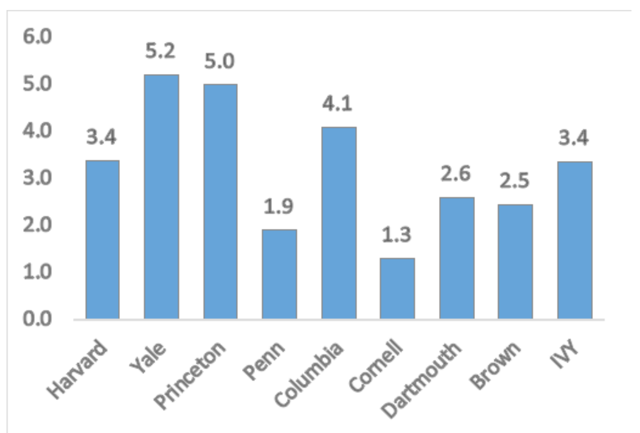


Exhibit 22: Public Alpha of Ivies
Source: MPI

Public alpha attribution

To understand which of the private assets mostly contributes towards the high public alphas observed in Exhibit 22, we perform alpha attribution. We first note that a fund's return may be decomposed as follows:

$$Y_{fund} = \alpha_0 + \Sigma b_{public} X_{public} + \Sigma b_{private} X_{private} \quad (1)$$

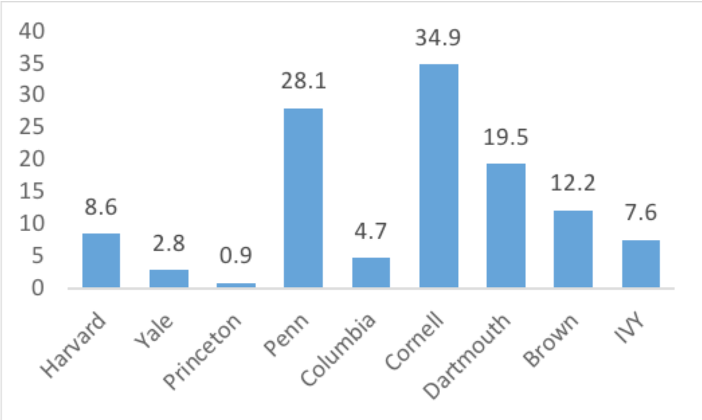


Exhibit 23: Public Alpha P-Value of Ivies
Source: MPI

Exhibit 24: Public Alpha of Private Indices
Source: MPI

where Y and X denote return time series and α0 is the private alpha. We can also decompose each of the private series as follows:

$$X_{\text{private},i} = \alpha_i + \Sigma b_{public} X_{public} \qquad (2)$$

where αi is the public alpha of each private series. From (1) and (2), as an approximation, the public alpha of a fund will then be given by:

$$\text{Public alpha: } \alpha_0 + \Sigma b_{\text{private}} \alpha_i \qquad (3)$$

This alpha has been calculated more accurately in Exhibit 22 based on the below regression:

$$Y_{\text{fund}} = \alpha_1 + \Sigma b_{\text{public}} X_{\text{public}} \qquad (4)$$

We call the difference between the public alphas as calculated in (3) vs (4) as idiosyncratic alpha, α_u:

$$\alpha_u = \alpha_1 - \alpha_0 - \Sigma b_{\text{private}} \alpha_i \qquad (5)$$

This term is still part of a fund's overall ability to deliver alpha against public indices but we do not know which of the private factors, if any, is responsible for this portion of the alpha. This term should generally be small. We can now attribute the public alpha calculated in (4) to its constituents:

$$\alpha_1 = \alpha_0 - \Sigma b_{\text{private}} \alpha_i + \alpha_u \qquad (6)$$

We now proceed to regress each of the private indices against the set of public indices in order to obtain their public alphas, shown in Exhibit 24. The regression is performed using the same parameters as the regression in Exhibit 22 to ensure consistency.

The public alpha attribution results are displayed in Exhibits 25 and 26, omitting the idiosyncratic alpha for ease of comparison as that turned out to be quite small, compared to the public alpha.

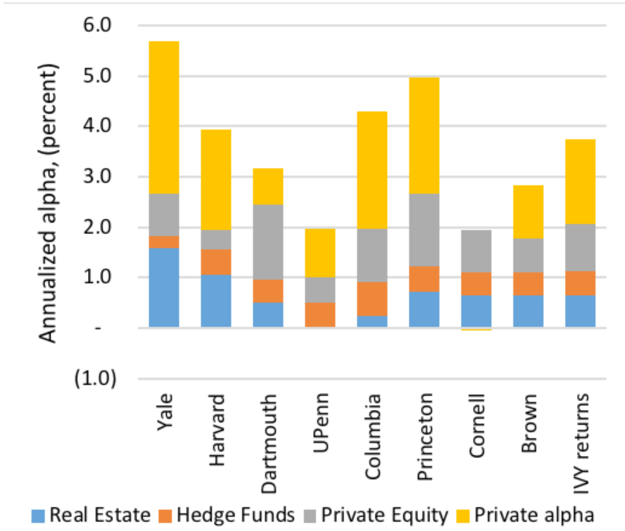


Exhibit 25: Public Alpha Attribution by Ivy
Source: MPI

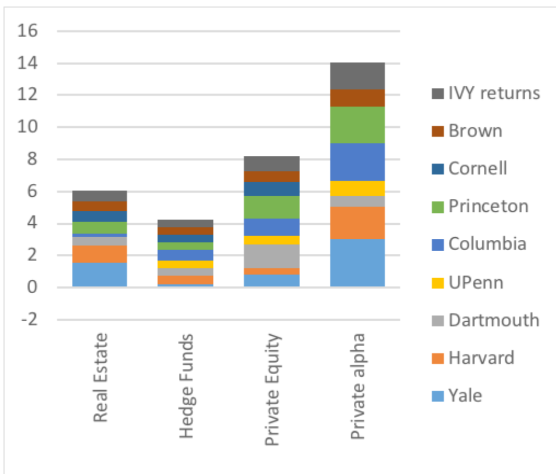


Exhibit 26: Public Alpha Attribution by Asset Class and Type of Alpha
Source: MPI

First, we observe that a good portion of endowment public alpha is due to their private alpha. This is particularly true for Yale, Harvard, Columbia and Princeton, in accordance with Exhibit 10. Cornell seems to be the only exception here with a negative private alpha, similar to a negative overall alpha from Exhibit 10. Second, we observe that, overall, it is private equity and real estate that contribute most to public alpha, with hedge funds having a smaller contribution. Finally, we observe that the public alpha contribution from hedge funds is consistent throughout all endowments, indicating no particular manager selection advantage among endowments. Private equity and real estate, on the other hand, show a less consistent view, with Yale and Harvard gaining a lot of alpha from real estate while the rest of endowments gained more from private equity.

Endowment risk

The question we are now trying to answer is how much risk funds take to achieve the high returns and alphas we saw in the previous section. To do this, we use the fund exposures and risk of the underlying factors over the period examined.

Once a factor-mimicking portfolio is constructed for each fund, the task of calculating the risk of that portfolio becomes essentially an exercise of calculating risk for each individual asset class within the factor-mimicking portfolio. While this is rather straight-forward for public asset classes, there are many challenges associated with this for real estate, private equity and venture capital.

Given there is no secondary market for the real estate properties held by real estate funds or private companies found inside private equity or venture capital funds, valuations (NAVs) for these entities are based on appraisals that are typically backward looking [Geltner, Jenkinson et al.]. This introduces smoothing in the time series of returns, which causes variance estimates to be biased toward zero. In commercial real estate, for example, there are infrequent purchases or sales of individual properties and valuations must be inferred based on recent sales of comparable properties, historical trends or current operating income (Geltner 1993). In venture capital, companies are valued every year or two when the time comes to negotiate new funding. Between such events, prices are typically carried forward or are a mix of recent and less recent company valuations. Buyouts are even more difficult to evaluate. There are few comparable transactions between when a company is bought and sold. So, just like in real estate, valuations are mainly based on appraisals, which may be quarterly or less frequent, and returns will certainly be backward looking and will avoid any transaction outliers. The lack of market-based valuations on a regular basis results in stated returns for these asset classes being artificially smooth.

One way to estimate the risk of these asset classes is to find out the factors they are exposed to along with the factor exposures. It would then be a matter of estimating risk for those factors as we describe further in the paper for the funds we analyze. A popular method that is followed by academics (Geltner 1992) and practitioners (Kinlaw et al. 2013), is to apply a de-smoothing algorithm that recovers the true volatility of the series. Since the approach based on factors leaves a large portion of the variance of each index unexplained,[26] we have chosen to apply the de-smoothing approach to the indexes in the factor-mimicking portfolio in order to estimate the risk of a specific fund.

As described in Geltner (2003), this method is based on an assumption that the effect of appraisal-based valuation is such that the observed, smoothed return of an aggregate private index is partially due to the true, de-smoothed and unobserved return of that index and partially due to past observed returns.[27]

$$r^*_t = w_0\, r^D_t + \varphi_1 r^*_{t-1} + \varphi_2 r^*_{t-2} + \ldots$$

Where

r^*_t = the publicly reported index return for year t

r^D_t = the de-smoothed return

$$w_0 + \varphi_1 + \varphi_2 + \ldots = 1$$

In order to uncover the de-smoothed return to calculate its risk, we can regress the observed series against one or more of its lags.

$$r^*_t \sim \alpha + \varphi_1 r^*_{t-1} + \varphi_2 r^*_{t-2} + \ldots + u_t$$

The constant is there in order to fully separate the calculated betas from the idiosyncratic term, $w_0\, r^U_t$. It contains uncaptured effects as well as any trend present in the residuals. We assume that any uncaptured effects are small, effectively making α part of the de-smoothed series: $w_0\, r^U_t = \alpha + u_t$. Based on the coefficients we find, we can solve for the de-smoothed series:

$$r_t = [r^*_t - \varphi_1 r^*_{t-1} - \varphi_2 r^*_{t-2} - \ldots] / w_0$$

Where

$$w_0 = 1 - \varphi_1 - \varphi_2 - \ldots$$

De-smoothing takes place based on quarterly returns calculated over the time period June 2002 and June 2017, resulting in 60 quarterly data points.

Real estate

For the period examined, we used one index lag as it was enough to remove the serial correlation present in the original index.[28] The beta against the first lag was 0.68 with a highly significant p-value of 0.00000025. The de-smoothed index is plotted against the original index in (Exhibit 27, next page). We can see that they are both very close to each other, with the de-smoothed index having much higher risk with an annualized standard deviation of 25.27 percent compared to 11.05 percent of the original index. We also observe that the trough for the de-smoothed index took place in December 2008 as opposed to December 2009 for the original index.

Compared against the Case Shiller HPI index, which reached its lowest level in April 2009, it looks like the de-smoothed trough of the real estate index is a more realistic representation of what actually took place.

Exhibit 27: Original vs. De-smoothed Cambridge Associates Real Estate Index
Source: Cambridge Associate, MPI

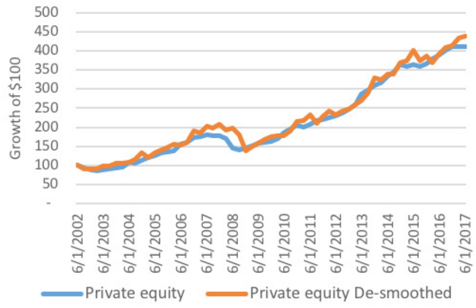


Exhibit 28: Original vs. De-smoothed Private Equity Portfolio
Source: Cambridge Associate, MPI

Private equity

For the private equity portfolio we have created, we also find one lag most suitable.[29] We obtain a beta coefficient 0.4712 with a highly significant p-value 0.0068 percent. The de-smoothed portfolio tracks the original very close but with a higher annualized standard deviation of 13.61 percent compared to 8.26 percent of the original portfolio.

Drawdown

Now that we have obtained the true series for each of the private indices, we can proceed with the risk estimation of the endowment funds. We start with the maximum drawdown, a commonly used statistic that measures the largest cumulative loss over a time period. It is often thought of as a tail-risk measure that offers insights into the magnitude of potential cumulative losses. Annual performance reporting tends to smooth out the performance pattern hiding the actual "investor pain" intra-year, something that quarterly factor-mimicking returns are now able to indicate. We follow the approach developed in Li et al. (2012), where reported monthly hedge fund returns were used to infer and project daily intra-month performance. In this case, we use the quarterly index returns to calculate intra-year performance. For real estate, buyout and venture capital, we used the de-smoothed indices since they more accurately represent the true drawdowns that these asset classes experienced.[30]

To make sure that we don't overestimate or underestimate the calculated drawdowns based on the factor-mimicking portfolios, we take an extra step to ensure that the total return from each factor-mimicking portfolio at each fiscal year equals the reported annual endowment return. To do this, we add a constant to each quarterly return of a factor-mimicking portfolio within a given fiscal year, such that the geometrically compounded factor-mimicking portfolio return equals the reported annual Ivy return. For the 15 fiscal years we examine, this means that we end up with 15 constant terms per each factor mimicking portfolio.

Exhibit 29 shows the maximum quarterly drawdowns of the endowments' factor portfolios compared to a 60-40 portfolio (orange bars). For comparison, we also supply the drawdowns based on annual returns (blue bars).

These drawdowns are significantly lower than if one were to calculate drawdowns based on the reported annual fund returns, highlighting the importance of our approach. They are also more severe than the 60-40 portfolio's drawdowns, indicating that endowments may take on considerably more tail risk.

Standard deviation

Exhibit 30 shows the annualized standard deviation of the various funds over the initial period. Having removed the bulk of serial correlation that the original indices displayed means that the factor-mimicking series are nearly i.i.d.

Similar to drawdowns, all endowments exhibit higher risk than the 60-40 portfolio, hinting that this may be the reason for the high returns.

Sharpe ratio

With better estimates of the endowments' risk using quarterly data and accounting for illiquid, appraisal-based investments, we are able to compare the endowments' risk-adjusted returns as

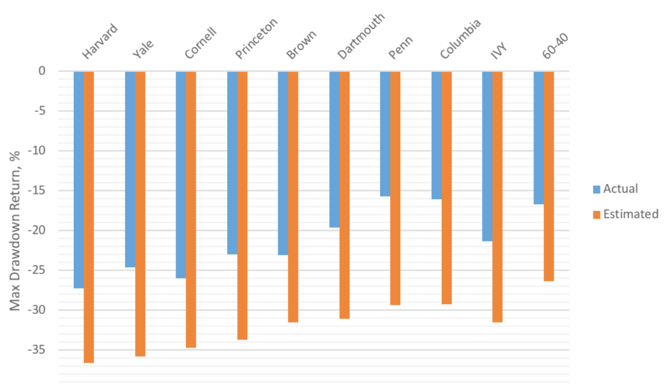


Exhibit 29: Actual and Estimated Max Drawdowns of Ivies
Source: MPI

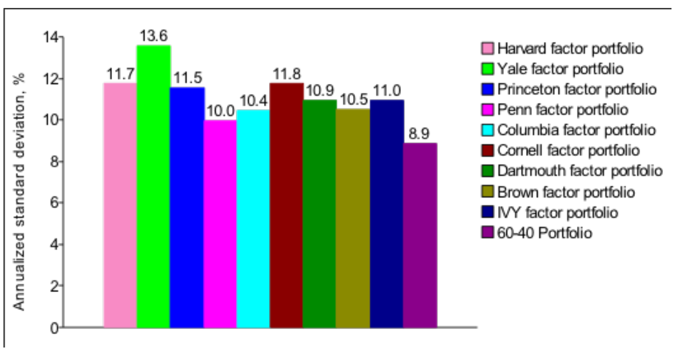


Exhibit 30: Estimated Annualized Standard Deviation of Ivies
Source: MPI

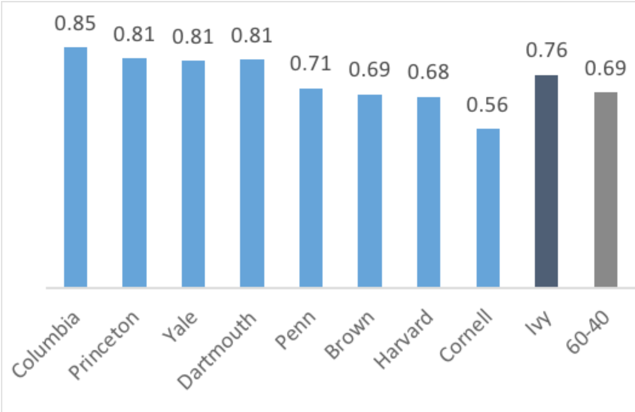


Exhibit 31: Sharpe Ratios of IVY Factor Portfolios
Source: MPI

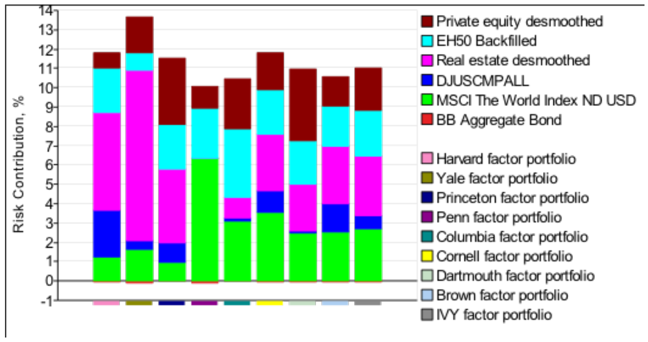


Exhibit 32: Estimated Systematic Risk Contribution of Ivies
Source: MPI

calculated using Sharpe ratio. The risk portion of the calculation is based on the risk of the factor-mimicking portfolios whereas the performance portion is based on the adjusted factor-mimicking portfolio returns as described in the Drawdown section. For the risk-free rate we used the three-month Treasury Bill Index. Our focus is on the ex-post Sharpe ratio measure over the period analyzed.[31]

The results are interesting. Despite the difference in total returns or alpha-generating capabilities, the dispersion in performance efficiency among the various funds is rather small. Moreover, the Sharpe ratios achieved by the Ivy endowments are all very close and span the 60-40 Sharpe ratio. From a performance efficiency standpoint and over the period examined, there doesn't seem to be any particular advantage in having endowments invest in private asset classes or in their manager selection ability.

The results are also somewhat different from the private alphas observed in Exhibit 10. For example, Yale is able to achieve a higher private alpha than Columbia, but Columbia has the higher Sharpe ratio. This may seem counter-intuitive at first. The explanation here is that, although Yale achieves higher alpha, that alpha is not high enough to achieve higher Sharpe ratio than Columbia.[32]

Indeed, from the risk contribution analysis in Exhibit 32 we observe that Yale's systematic risk is quite a bit larger than Columbia's, due to the increased real estate exposure. In the same sense, most endowments do not earn large enough alphas to be more efficient than a 60-40 portfolio which, although by construction achieves no alpha against public indices, exhibits much lower risk as per Exhibit 30.

Conclusion

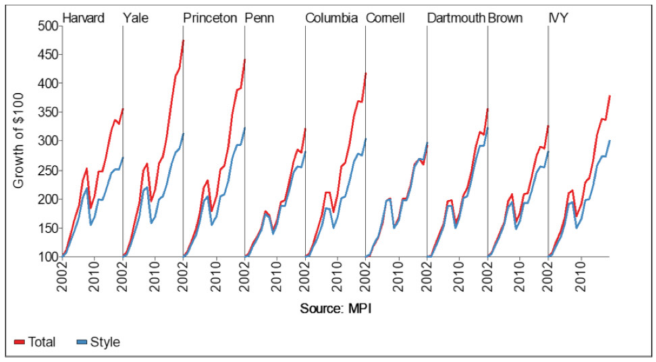
This paper analyzed the historical performance of the Ivy League endowments over the fiscal period 2003–2017. We used a factor model that includes public and private benchmarks representing stocks, bonds, commodities, real estate, hedge funds and private equity. To take into account the distinct dynamics that each fund exhibits, we performed regressions using a proprietary Dynamic Style Analysis model. This dynamic model enables us to calculate market timing and more accurate alphas than if we were to assume constant exposures. We find that Ivy League endowments likely do not engage in market timing of significance or possess alpha-generating market timing abilities relative to the complete set of factors that explain the funds' performance. The in-sample private alphas seem to be noticeable and significant for some Ivies, indicating that some possess manager selection abilities. When tested out of sample, however, we found a similar magnitude in their excess returns over factor-mimicking portfolios, yet not significant, casting doubt on their alpha generation capabilities. When analyzed against public indices, we find that in-sample alphas go up by about 2 percent vs. when using all indices, indicating that the decision to invest in private asset classes does produce additional alpha. Alpha attribution analysis showed that public alpha is mainly due to private alpha. Public alpha contribution from asset classes came mainly from private equity and real estate as opposed to hedge funds.

We find increasingly large exposures to private asset classes among most Ivy League funds, reaching as high as 90 percent. Although that helps them achieve high returns and alphas (against public indices), it doesn't come without risk. We used the factor-mimicking portfolios created by style analysis to calculate how much risk the funds take. For real estate and private equity, we applied a de-smoothing technique in order to overcome the staleness introduced by appraisal-based valuations and estimate the true risk of each index. Armed with proper exposures, frequent factor data and risk estimates, we then calculated drawdowns, standard deviations and Sharpe ratios. We find that all funds exhibited severe drawdowns during fiscal year 2009, and they all take on higher risk than the 60-40 portfolio.

Risk estimates have implications for Sharpe ratio where we find that the dispersion among endowment Sharpe ratios is small, spanning the Sharpe ratio of the 60-40 portfolio. The reason for this is that, although most endowments are able to achieve large alphas against public indices and overall high returns, the risk they may take is disproportionally higher than the alpha they may achieve in relation to the risk of the 60-40 portfolio.

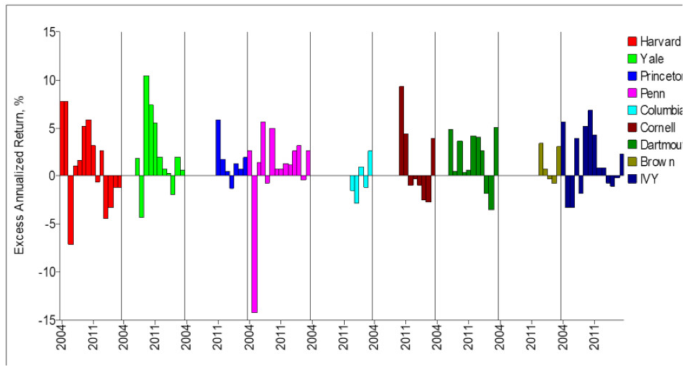
Our findings highlight important aspects for other institutional investors who may be looking for guidance from these large endowments that have access to elite alternative and private market fund managers. On the one hand, the high exposure to private asset classes does help achieve high returns, as evident by the Ivy endowments' past performance. On the other hand, however, such large allocations to alternatives are not guaranteed to achieve much higher alphas than public indices, as one needs to know what private investment to allocate to as well as what managers to pick. Private equity has done much better than hedge funds during the same timeframe, for example, but within private equity there may be a large dispersion in terms of alpha against public indices. Our results are in accordance with Lerner et al. (2008), who raise caution with respect to large private allocations. Even if high alphas are achieved, the question then becomes whether they are high enough to result in high Sharpe ratios. Investors, particularly those with liability and liquidity constraints and shorter time horizons than large, well-funded endowments, need to pay attention to the true risk they take by investing in private asset classes as such allocations may easily erode any alpha or return gain and result in Sharpe ratios similar or inferior to a 60-40 portfolio.

Appendix



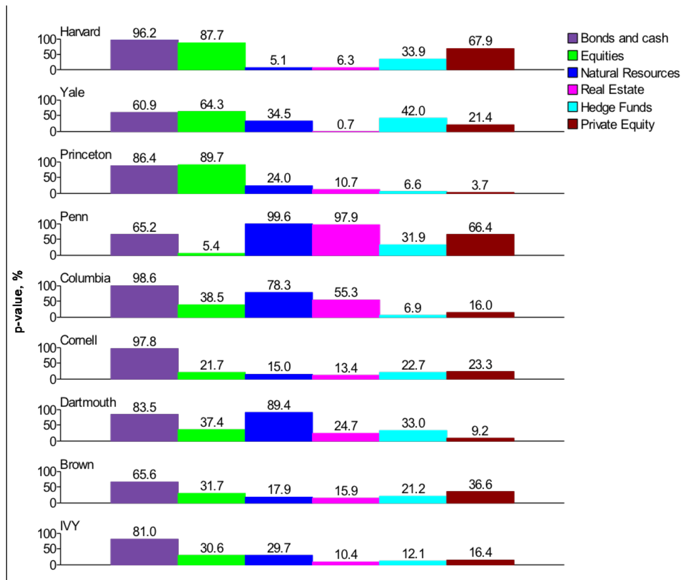
Appendix A

Factor mimicking portfolio tracking against endowment returns



Appendix B

Excess returns of factor mimicking portfolios



Appendix C

Figure C. Average p-values of the dynamic exposures shown in Exhibit 6

Endnotes

1. Based on the 2017 NACUBO study: https://www.nacubo.org/-/media/Nacubo/Documents/EndowmentFiles/2017-NCSE-Public-Tables--Number-of-NCSE-Participants--2.ashx?la=en&hash=936500535E61CA448ECD090E43B9E0733D9F4514

3. https://www.nacubo.org/-/media/Nacubo/Documents/EndowmentFiles/2017-NCSE-Public-Tables--Spending-Rates.ashx?la=en&hash=E1CE49F8E652B8C705F4414073B80E21DF1CFFEE

3https://www.nacubo.org/-/media/Nacubo/Documents/EndowmentFiles/2017-Endowment-Market-Values.ashx?la=en&hash=E71088CDC05C76FCA30072DA109F91BBC10B0290

4. https://en.wikipedia.org/wiki/Ivy_League

5.Consisted of 60% S&P 500 Index and 40% Bloomberg Barclays U.S. Aggregate Bond Index.

6. This refers to an asset allocation strategy that seeks to generate high returns by allocating to private equity and other alternative assets, having access to top performing managers and being broadly diversified.

7. By risk here we refer to the historical standard deviation that corresponds to each of the funds. This is purely an in-sample measure that treats the entire time series as identically independently distributed, ignoring volatility clustering or multiple regimes that may exist within the period examined. Had we had data of higher frequency it would have allowed for a more sophisticated risk analysis. Our measure however coincides with commonly used ways to evaluate risk for such funds, including variance for Sharpe ratio.

8. http://www.nacubo.org/Research/NACUBO-Commonfund_Study_of_Endowments/Public_NCSE_Tables/Total_Market_Value_of_Endowments.html

9. All indices have daily availability except indices used to factor mimic alternative asset classes, which are available monthly or quarterly.

10. We can still obtain standard deviations using annual data but this would result in estimates with very large variance (Kenney and Keeping, 1951) based on somewhat smooth return series.

11. Although there were many endowments that had returns prior to 2003, we constrained ourselves to using a common time period across all endowments in order to ensure a consistent comparison against their calculated alphas over the same time period.

12. Source: Cambridge Associates

13. The returns for this index prior to 2007 have been backfilled with the HFRI Fund Weighted Composite Index

14. We used preliminary data for the quarter ending on June 30, 2017, representing 61 percent of active funds updated compared to the prior quarter's NAV for US Buyout and 68 percent for US Venture Capital.

15. We acknowledge that long/short equity strategies or derivatives found in hedge funds may result in a negative market exposure. However, net exposure among long/short equity funds are typically positive, plus it is hard to think that endowment funds would take on negative overall market exposures.

16. Lobosco and DiBartolomeo (1997) find this to be the case on Sharpe regressions.

17. To calculate predicted R-squared, we re-estimate the model for a given fund while taking out one of the 15 annual return observations between 2003 and 2017 and then assessing the difference between the removed annual return observation and the estimated value for that observation made by the remaining observations. In this case, we re-estimate 15 different times taking out each observation in turn. We combine all 15 of the differences between estimated and actual returns to calculate the predicted R-squared.

18. In appendix C we display the p-values for each factor exposure

19. This is despite the fact that allocations to alternatives are generally reported to be less than the exposures we show here. The reason for this may have to do with the difference between exposures and allocations as explained further above. The lack of detailed data on the allocations that these funds follow makes it hard to reject or confirm the observed differences between exposures and allocations.

20. We could have used the 60-40 portfolio as a benchmark like we did in the introduction when we compared the Ivy League historical total returns. However, given the funds' exposures are very different from a 60-40 portfolio, we do not consider this portfolio to be appropriate for evaluation of the timing return component.

21. We can only evaluate timing on an annual basis. Although it is possible that funds may shift exposures more rapidly at a quarterly or higher frequency, this is unlikely given their size and policy.

22. Since the endowment fund returns we used are net of fees, to the extent that the funds invest in managers that trade public asset classes, given such asset classes in our regression use public indices where no fees are involved, high fees paid to those investment managers would have negative contribution to selection return.

23. Given the limited 10-year history of each window, we have dropped the alpha from the analysis so as to ensure we have as many degrees of freedom as possible.

24. We omit displaying the in-sample factor exposures for the first 10 years since we do not have available factor exposures during that period for the replicating portfolio. The in-sample style dispersion has been produced based on a regression that covers a time period that is 10 years longer than the replicating portfolio dispersion. Both dispersions are then calculated based on the weights of a common period that excludes the first 10 years.

25. This includes the same smoothness parameter found in the original regression. The reason for this is two-fold. On the one hand, we want to apply the same level of beta volatility among the two regressions to make them more comparable. We choose the original regression as the anchor on which we choose the optimal smoothness.

26. We followed the approach described in Woodward (2004) by regressing venture capital against six quarterly lags of Russell

2000 over the analysis period 2005-2016, using intercept and no constraints and got an R2 of 61.6 percent. Using the factor exposures from this approach to estimate risk would show much less of a variance than the index has. Doing the same for the private equity index resulted in a 68.7 percent R2 and a variance similar to the one obtained via the de-smoothing approach. Real estate does not have good enough market indices, that we are aware of, that explain most of its variance.

27. This is similar to equation (3) in Geltner (2003) but expanded to include more lags, as shown in equation (2) of Cho et al. (2001).

28. The Durbin-Watson statistic calculated against the residuals was found to be 2.36, indicating no statistical evidence of positive or negative serial correlation.

29. The Durbin-Watson statistic of the residuals based on 56 quarterly observations is 2.18 indicating no statistical evidence of positive or negative serial correlation.

30. Since the funds report returns in the same smoothed manner that the indices report, using the de-smoothed indices in the regression would have resulted in a mismatch. The only way to uncover the true drawdown is to use the original, smoothed, series in the regression and the de-smoothed series when calculating drawdowns.

31. The ex-ante Sharpe ratio requires expected values for the asset classes we looked at. Since this is very hard to obtain, we place our focus on the ex post measure.

32. Formally, assume a fund is given by: $y = \alpha + bx$. Assuming, for simplicity, that the risk-free rate has zero variance, the Sharpe ratio of this fund is given by: $SR = [a + b\text{mean}(x)] / b\sigma(x)$. If, among two funds, the systematic part increases in risk more than the alpha increase, then the Sharpe ratio will actually decrease.

References

Ang, Andrew, et al. "Estimating private equity returns from limited partner cash flows." (2014).

Axelson, U., M. Sorensen, and P. Stromberg. Alpha and Beta of Buyout Deals: A Jump CAPM for Long-Term Illiquid Investments. Working paper, 2014.

Barber, Brad M., and Guojun Wang. "Do (Some) University Endowments earn Alpha?." *Financial Analysts Journal* 69.5 (2013): 26-44.

Brown, Jeffrey R., et al. "How University Endowments Respond to Financial Market Shocks: Evidence and Implications." *The American Economic Review* 104.3 (2014): 931-962.

Brown, Keith C., Lorenzo Garlappi, and Cristian Tiu. "Asset Allocation and Portfolio Performance: Evidence from University Endowment Funds." *Journal of Financial Markets* 13.2 (2010): 268-294.

Buchner, Axel. "The Alpha and Beta of Private Equity Investments." Browser Download This Paper (2014).

Chen, Tuo. "Do the Rich Know Better?–University Endowment Return Inequality Revisited." (2016).

Cho, Hoon, Yuichiro Kawaguchi, and James D. Shilling. "Unsmoothing Commercial Property Returns: A Revision to Fisher–Geltner–Webb's Un-smoothing Methodology." T*he Journal of Real Estate Finance and Economics* 27.3 (2003): 393-405.

Das, Nandita. "Hedge Fund Classification Using K-means Clustering Method." 9th International Conference on Computing in Economics and Finance. 2003.

Dimmock, Stephen G. "Background risk and University Endowment Funds." *Review of Economics and Statistics* 94.3 (2012): 789-799.

Driessen, Joost, Tse-Chun Lin, and Ludovic Phalippou. A New Method to Estimate Risk and Return of Non-traded Assets from Cash Flows: The Case of Private Equity Funds. No. w14144. National Bureau of Economic Research, 2008.

Franzoni, Francesco, Eric Nowak, and Ludovic Phalippou. "Private Equity Performance and Liquidity Risk." *The Journal of Finance* 67.6 (2012): 2341-2373.

Fung, William, and David A. Hsieh. "Hedge-fund Benchmarks: Information Content and Biases." *Financial Analysts Journal* (2002): 22-34.

Geltner, David. "Estimating Market Values from Appraised Values without Assuming an Efficient Market." *Journal of Real Estate Research* (2009).

Gibson, Rajna, and Sébastien Gyger. "The style Consistency of Hedge Funds." *European Financial Management* 13.2 (2007): 287-308.

Jegadeesh, Narasimhan, Roman Kräussl, and Joshua Pollet. Risk and Expected Returns of Private Equity Investments: Evidence Based on Market Prices. No. w15335. National Bureau of Economic Research, 2009.

Kenney, J. F., and E. S. Keeping. "The Distribution of the Standard Deviation." *Mathematics of Statistics*, Pt 2 (1951): 170-173.

Kinlaw, Will, Mark Kritzman, and David Turkington. "Liquidity and Portfolio Choice: A Unified Approach." *The Journal of Portfolio Management* 39.2 (2013): 19-27.Lerner, Josh, Antoinette Schoar, and Jialan Wang. "Secrets of the academy: The Drivers of University Endowment Success." *The Journal of Economic Perspectives* 22.3 (2008): 207-222.

Lerner, Josh, Antoinette Schoar, and Jialan Wang. "Secrets of the academy: The Drivers of University Endowment Success." *The Journal of Economic Perspectives* 22.3 (2008): 207-222.

Lerner, Josh, Antoinette Schoar, and Wan Wongsunwai. "Smart Institutions, Foolish Choices: The Limited Partner Performance Puzzle." *The Journal of Finance* 62.2 (2007): 731-764.

Lo, Andrew W. "The Statistics of Sharpe Ratios." *Financial Analysts Journal* (2002): 36-52.

Lobosco, Angelo, and Dan DiBartolomeo. "Approximating the Confidence Intervals for Sharpe Style Weights." *Financial Analysts Journal* (1997): 80-85.

Li, D., Markov, M., and Wermers, R. "Monitoring Daily Hedge Fund Performance When Only Monthly Data is Available." *Journal of Investment Consulting*, Vol. 14, No. 1, 57-68, 2013. Winner of IMCA 2102 Best Academic Paper Competition.

Markov, Michael, Vadim Mottl, and Ilya Muchnik. "Principles of Nonstationary Regression Estimation: A New Approach to Dynamic Multi-factor Models in Finance." DIMACS, Rutgers Univ., Piscataway, NJ, Tech. Rep 47 (2004).

Phalippou, Ludovic, and Oliver Gottschalg. "The Performance of Private Equity Funds." *Review of Financial Studies* 22.4 (2009): 1747-1776.

Sharpe, William F. "Asset allocation: Management Style and Performance Measurement." *The Journal of Portfolio Management* 18.2 (1992): 7-19.

Woodward, Susan E. "Measuring Risk and Performance for Private Equity." Sand Hill Econometrics (2004).

Author Bios

Apollon Fragkiskos

MPI

Apollon heads MPI's research team, where his mandate is to guide and advance research in key competitive areas of the firm, such as style analysis, hedge fund replication, fund selection and portfolio construction. His focus is to create cutting edge analytics based on machine learning and financial techniques and provide innovative solutions to fund buyers, fund sellers and asset allocators. He is also actively engaged in publications of blogs, white papers and research.

Prior to joining MPI, Apollon was the head of the Research team within the risk analytics group of State Street Global Exchange. In that role, Apollon created, researched and led the implementation of key product enhancements such as returns based style analysis, hedge fund replication and macro-economic stress testing. Prior to that, he spent several years as a senior financial engineer in charge of fixed-income and derivatives pricing models, which formed the backbone of State Street's market risk management system.

Apollon holds a BSc in Physics from the University of Athens and an MSc in Financial Engineering from the Michael Smurfit School of Business.

Sean Ryan

MPI

Sean Ryan is a senior research analyst at MPI. His specific focuses include mutual and hedge fund analysis, endowment research, quantitative fund surveillance and other research projects. He holds a B.S. in Finance and MBA from Indiana University of Pennsylvania.

Michael Markov

MPI

Michael Markov is co-founder and chairman of MPI. In 1992, recognizing the power that quantitative analysis delivers investors, he led the development of the industry's first returns-based style analysis application based on Bill Sharpe's groundbreaking methodology. Continuing this spirit of innovation, he co-authored Dynamic Style Analysis, a significant advancement to legacy returns-based modeling techniques for investors seeking the most precise analysis of managed investment products and portfolios.

Being an industry expert in quantitative analysis, Michael is a frequent speaker at investment management forums around the globe on fund performance attribution, due diligence and monitoring, investment risk management and hedge fund analysis. His thoughts and opinions are regularly sought by leading financial press.