5-2020

An Examination Of Performance Amongst Us Public Pension Plans

Richard Kaser
University of Pennsylvania

An Examination Of Performance Amongst Us Public Pension Plans

Keywords
pension, retire, assets

Disciplines
Business

An Examination of Performance Amongst US Public Pension Plans

By

Richard Kaser

An Undergraduate Thesis submitted in partial fulfillment of the requirements for the

JOSEPH WHARTON SCHOLARS

Faculty Advisor:

Itamar Drechsler

Associate Professor, Finance Department

THE WHARTON SCHOOL, UNIVERSITY OF PENNSYLVANIA

MAY 2020

1 Introduction

Pension plans play an integral role in many retirees' lives in the United States. As of 2018, approximately 30.2 million American workers participated in some form of a defined benefit plan (Pension Rights Center, "How…"). Further, a disproportionate number of these workers are public employees, accounting for nearly half of these participants. Thus, public pension plans are amongst the most influential investment institutions in the country. To put their size into perspective, they account for approximately $4.5 trillion in assets (as of Q2 2009), while the total US stock market and bond market are valued at approximately $30 trillion and $40 trillion, respectively (NASRA, "Public…"). Additionally, as of 2018, approximately 21 million employees and retirees participated in state and local government pension plans (Urban Institute, "State…"). Generally, then, the management of assets under the control of these plans is of concern to many. The question then arises as to whether each plan manages its assets optimally, given its characteristics and the context in which it operates. In order to answer this, one must understand the impact of both these characteristics and context on performance. However, this thought is of little relevance unless certain plans consistently outperform others. Thus, the importance of analyzing performance persistence amongst US public pension plans becomes evident.

Historically, external asset managers hired by pension plans (hereby referred to as pension funds) have been the focus of performance persistence research in relation to pensions. Much of this research has sought to provide guidance regarding allocation of capital between active and passive strategies. Only recently has work concerning the performance persistence of the portfolios of pension plans themselves been published. Moreover, these papers investigate pension systems outside of the US; these plans are not entirely similarly structured, nor do they operate in entirely similar markets. As such, the persistence of performance of public pension plans in the US remains in question, and this paper aims to examine this issue.

The academic literature on performance persistence amongst financial portfolio managers is expansive, yet not without its limitations. For any given investment style or asset class, neither the degree to which managers' performance persists from period to period nor, therefore, an explanation for this phenomenon is entirely agreed upon.

Many instrumental papers investigate performance persistence within the mutual fund space, as this data is relatively accessible. For example, a 1997 study by Mark M. Carhart finds that, upon review of a data set free of survivorship bias, mutual fund performance tends to persist under certain conditions (Carhart 1997). By creating portfolios of one-year return-sorted mutual funds at the beginning of each calendar year 1963-1993, he reveals significant return spreads between the portfolios over the year following formation. Further, this persistence appears to be explained primarily by investment expenses and common stock factors, namely size. Using similar methodologies, he also finds that sorting based on past performance over longer intervals doesn't provide further insight. However, sorting based on past three-year four-factor model alpha leads to longer-term persistence in alpha, largely explained by expenses. This model includes factors for excess return of a value-weighted aggregate market proxy, size, value, and 1-year momentum. Although, the use of the same model to sort and estimate performance presents potential bias.

However, the evidence for mutual fund performance persistence is largely limited to shorter intervals under certain conditions, as is the case with the aforementioned Carhart study. Similar results are found in papers by Lu Zheng in 1999, Martin J. Gruber in 1996, and Nicolas P.B. Bollen and Jeffrey A. Busse in 2004. Further, in a 1995 study, Burton G. Malkiel examines the tendency of "winners" (defined as funds to have had an above-median return over the prior calendar year) to repeat as winners (Malkiel 1995). He finds significant evidence of the tendency of winners to repeat in the 1970s, though insignificant in the 1980s. Again, this study is confined to a one-year predictor and horizon. Additionally, using recent data and a methodology similar to that used by Carhart, a 2016 study by Morningstar finds evidence of short term mutual fund performance persistence, though no evidence of persistence over a longer time horizon (Brian and Li 2016).

There is empirical evidence in more comparable investment spaces, namely UK pension funds. To provide a distinction, "pension fund managers" typically refers to investment managers employed by pension plan trustees to manage money. Thus, this generally refers to a segment of a plan portfolio as opposed to the portfolio in its entirety, which is the object I focus on.

In a 1997 paper, Gavin Brown, Paul Draper, and Eddie McKenzie use a transition matrix technique in which they track year-to-year movement of funds between performance quartiles throughout several time periods between 1986 and 1992 (Brown et al. 1997). The authors define performance using both total and market exposure-adjusted returns. Generally, this study finds that funds in upper performance quartiles are more likely to remain in their given quartile from one period to another than if this were determined by chance, providing some evidence for persistence in UK pension fund returns. Further, a 2005 paper by Ian Tonks evaluates the performance spreads of UK pension managers ranked by performance over prior quarterly, annual, and 3-year evaluation periods during 1983-1997 (Tonks 2005). The author also compares the observed distribution of win-loss combinations across periods to distributions expected in the absence of persistence. Performance is defined as abnormal return with respect to the Fama-French three-factor model, though performance is also estimated using an additional momentum factor as a secondary consideration. This investigation finds that there is considerable persistence on a yearly horizon and lesser persistence on longer horizons. The evidence for this is weaker once momentum is accounted for in the performance model.

Within the US pension space, a 1992 study by Josef Lakonishok, Andrei Shleifer, Robert W. Vishny, Oliver Hart and George L. Perry find no annual persistence, though some biannual persistence in performance among a set of 769 tax-exempt equity pension funds over 1983-1989 (Lakonishok et al 1992). These authors use a transition matrix technique very similar to that used by Brown, Draper, and McKenzie. However, this study is limited in comparison due to its confinement to equities. Some theory suggests that, structurally, US public pension plan portfolios, as opposed to the managed funds investigated in previous pension fund studies, may be particularly prone to performance persistence. Notably, in a 2004 paper, Jonathan B. Berk and Richard C. Green propose a model reconciling the rational nature of investors and their performance-chasing tendencies (Berk and Green 2004). In this model of mutual funds, investors direct their resources toward managers that they expect to produce positive net abnormal returns. However, there are costs associated with scale of operations for managers, and these costs increase with scale. Thus, through directing flows in and out of funds, investors allocate capital such that each expected net abnormal return is zero, as benchmark outperformance is absorbed by costs of scale of operations and fees. Such a concept provides a framework through which rational investors, performance-chasing behavior, and lack of performance persistence coexist. This theory is important to the pension fund space in that US public pension plans are not open investment vehicles. Capital flows are largely determined by employment and the associated retirement benefits. Therefore, it is possible that these portfolios are not subject to the same fair pricing imposed on funds under the model proposed by Berk and Green.

As for literature to which this idea might apply, there has been little persistence research concerning pension entities that likely exhibit similar flow determinants to those of US public pension plans. However, in 2012, Xiaohong Huang and Ronald J. Mahieu investigated persistence amongst Dutch pension plans and found no significant evidence. More specifically, in the Netherlands, employees are legally required to participate in a pension plan particular to their industry, and these pension plans are required to report risk-adjusted returns in the form of a legally-mandated annual z-score measure (Huang and Mahieu 2012). This z-score is calculated as net return minus net return on a prespecified benchmark, normalized by the total portfolio's risk. Given its calculation relative to a pension plan's ex ante benchmark, this measure largely reflects the ability of a pension plan to implement its investment strategies as opposed to the efficiency of these strategies themselves. Given poor performance by this measure (performance in the bottom 10% over a five-year period), plans' participants are free of legal requirement to participate. Thus, in this system, flows can be directly influenced by performance. Huang and Mahieu analyze performance persistence amongst 57 of these plans from 1998 through 2006 by essentially regressing plans' z-score measures at time t on z-score measures at time t-1 as well as performing a Spearman rank correlation test. In this second method, the authors determine a correlation coefficient between z-scores in each year and the next, ultimately finding none to be significantly different from zero. However, this might be expected, given the construct of the z-score as well as the legal framework through which Dutch pension plans' capital flows can be directly influenced by performance.

Generally, then, the performance persistence literature has yet to cover the US public pension plan space, and, theoretically, results need not translate across different types of entities. Given the structural differences between each type of entity explored in the papers mentioned above and public pension plans in the US, further investigation is required.

2 Data

In order to expand upon this topic, performance data for US public pension systems is necessary. Conveniently, the Center for Retirement Research at Boston College, the Center for State and Local Government Excellence, and the National Association of State Retirement Administrators provide a public database of hundreds of variables for US state and local public pension plans from 2001 through 2018. The 194 plans in this data set account for 95% of all public pension members and assets in the US. For most plan-years (defined as a given plan in a given year), annual return, investment, actuarial assessment, liability, employee, and beneficiary data are provided. Though, for each plan year, returns can be reported as either net or gross figures. Due to this inconsistency in reporting, much of my analysis focuses on subgroups of plans that consistently report using the same convention, as will be elaborated on. For certain plan-years, return data is not available, reducing the total number of plan-years of use to 2,700.

For a subset of 208 of these plan-years, returns are provided as estimated figures based on change in market assets, annual contributions, and annual expenses. However, within this set, significant outliers bring the accuracy of these estimations into question. Notably, outliers exceed 100% returns for certain plan-years (for which approximate equity percentage is within reasonable bounds), leading me to believe that the calculation of these return estimates can be influenced by flow data. Generally, given this lack of credibility and the size of the remaining sample, I choose to omit these observations from my analysis.

In order to calculate excess returns and adjust these for risk, I rely on data sets provided by Kenneth French containing monthly Fama-French Three Factor returns, momentum factor returns, and US Treasury bill rates. To correctly adjust annual plan returns for risk, I create annual factor and US Treasury bill returns, matching the fiscal year-end date of each plan-year to the year-end date of the factor and US Treasury bill returns.

Further, I investigate a set of fee data for these plans over the same time period, acquired from public Comprehensive Annual Financial Reports. Generally, this reporting is inconsistent, and the fee data is useful for broad observations of particular subsets separated by reporting convention.

3 Methodology

In order to evaluate performance, I use abnormal returns with respect to three separate models: the Capital Asset Pricing Model (CAPM), the Fama-French Three Factor Model (three-factor model), and a model that modifies the three-factor model by including a momentum term (four-factor model). In choosing these models, I defer to much of the past performance persistence research (Carhart 1997; Tonks 2005). I estimate coefficients for the three models first via regression over the entire timespan of the data. The single-factor, three-factor, and four-factor models are of the following forms for plan i, respectively (note that these returns are expectations in reality, though this has been excluded for the sake of notation):

$$r_i - r_f = \beta_{i,r_{MKT}-r_f}(r_{MKT} - r_f) + \alpha_i$$

$$r_i - r_f = \beta_{i,r_{MKT}-r_f}(r_{MKT} - r_f) + \beta_{i,r_{SMB}}(r_{SMB}) + \beta_{i,r_{HML}}(r_{HML}) + \alpha_i$$

$$r_i - r_f = \beta_{i,r_{MKT-r_f}}(r_{MKT} - r_f) + \beta_{i,r_{SMB}}(r_{SMB}) + \beta_{i,r_{HML}}(r_{HML}) + \beta_{i,r_{MOM}}(r_{MOM}) + \alpha_i$$

I define performance with respect to a model as the alpha estimate in coefficient estimates of these models. For a particular plan-year, I estimate performance by adjusting returns by the returns on the appropriate factors multiplied by their corresponding coefficients, as illustrated in the following equations for performance of plan i in year t with respect to the CAPM, three-factor model, and four-factor model.

$$\alpha_{i,t,CAPM} = r_{i,t} - r_{f,t} - \beta_{i,r_{MKT-r_f}}(r_{MKT,t} - r_{f,t})$$

$$\alpha_{i,t,3-Factor} = r_{i,t} - r_{f,t} - \beta_{i,r_{MKT-r_f}}(r_{MKT,t} - r_{f,t}) - \beta_{i,r_{SMB}}(r_{SMB,t}) - \beta_{i,r_{HML}}(r_{HML,t})$$

$$\alpha_{i,t,4-Factor} = r_{i,t} - r_{f,t} - \beta_{i,r_{MKT-r_f}}(r_{MKT,t} - r_{f,t}) - \beta_{i,r_{SMB}}(r_{SMB,t}) - \beta_{i,r_{HML}}(r_{HML,t}) - \beta_{i,r_{MOM}}(r_{MOM,t})$$

One method by which I approach evaluating performance persistence is to examine the relationship between plan performance in year t and t+1. Specifically, taking all qualifying plan-years, meaning those that have return data for a particular year and the next, I plot performance in t+1 on performance in t and estimate a linear relationship, evaluating the significance of the regression coefficient estimates. The results of this method, however, are susceptible to being strongly influenced by single plan-year outliers. In order to minimize this effect, in each qualifying period, meaning a period in which return data exist for plans in that period and the next, I form equal-weighted portfolios corresponding to quantiles of performance in period t. I then evaluate the relationship of the performance of these portfolios from one period to the next in lieu of that of individual plans.

In this case, it is useful to eliminate the influence of future data on the information on which I sort. In order to address this, I perform further analysis in which I estimate coefficients with respect to the three pricing models for each plan using data from 2001 through period t for all years $t \in$ [2001, 2018], which I will refer to as to as rolling coefficients. I then estimate plan-year performance, referred to as rolling performance, using the above equations, though substituting rolling coefficients for those estimated using the entire span of the data.

Naturally, I extend the aforementioned regression method, which is particular to one-year formation periods and one year of subsequent observations, to different formation period and performance horizon lengths. Specifically, I vary either length from one to five years. In doing so, I estimate performance for a particular period as the arithmetic average rolling performance over that period. This alters the set of qualifying periods in each test, restricting the observations to periods for which return data is available for the combined length of the formation period and performance horizon.

Further, I employ a more categorical approach as an alternative measure of persistence, namely tracking the performance quantiles that performance-sorted portfolios fall into following the initial formation period. Following a similar method to the transition matrix techniques used in previous studies of performance persistence, I create matrices in which each row i represents a performance quantile in the formation period, and each column j represents a performance quantile over the performance horizon (Brown et al. 1997; Lakonishok et al. 1992). Each portfolio, formed following the same procedure as the portfolios used in regression, is counted in the appropriate entry. Thus, for matrix A, entry A_{11}, for example, represents the total number of portfolios that were in the highest-performing quantile over the formation period and remained in the highest-performing quantile over the subsequent performance period. The statistical significance of deviations of observed entries from the values of entries expected at random are then used as a measure of performance persistence. In essence, the larger the entries surrounding the main diagonal of a given matrix, the more likely plans are to remain in similar performance quantiles from year to year.

As mentioned, the returns provided by the public plan data set can either be reported as net of fees or gross. Across plans, this reporting style varies, and, in some cases, for a given plan, convention can vary from year to year. Further, in some plan-years, it is not made clear as to whether the reported returns are net of fees or gross. Therefore, I apply each of the described tests to three subsets of data: aggregate, gross, and net, where the aggregate set contains all reported return data, the gross set consists of data for plans that are never known to report net figures, and the net set consists of data for plans that are never known to report gross figures. Given the relative importance of net returns, I present all results corresponding to the net subgroup; though, aggregate results are included in the Appendix. Using these subsets, I largely filter out the effects of comparing returns that follow different reporting conventions. However, this presents a self-selection issue in that plans that perform particularly well over some time span may choose to report net figures, while the opposite may be true for poor performers. Though, the minority of plans (39 out of 150 plans for which returns are reported, excluding estimated data) change convention within 2001-2018, and these are excluded from the gross and net subsets.

Finally, considering the importance of fees to performance, I analyze a separate set of fee data for the pension plans found in the public plan data set. This data is drawn from the same sources as the public plan data, namely legally-mandated Comprehensive Annual Financial Reports. Notably, US public pension plan reporting regarding fees is inconsistent across plans and years. Thus, I use this data in a general sense, finding broad measures such as average fee as a percentage of assets. Further, given the reporting differences, I perform these analyses on two separate subsets: plan-years for which management fees are explicitly reported and plan-years for which management and performance fees are explicitly reported, where performance fees are defined as fees explicitly dependent on management performance. Note that this could include carry.

4 Results

The described analyses were first performed with factor coefficients gathered from regressions for each plan using return data for the entire span of 2001-2018. However, an issue arises from this: look-ahead bias in plan-year performance estimates. Using the entire span of the return data to estimate plans' factor coefficients forces plan-year performance estimates, which are calculated using the previously provided equations and necessarily rely on these sensitivity estimates, to indirectly incorporate future performance data, complicating tests of persistence. However, as will be discussed in further detail, I continue with this analysis as well as a potential solution and will compare the results and issues that either method presents.

A summary of the regression parameter estimates obtained using the full span of the return data for each plan-year follows:

CAPM Parameter Estimation (for Net Subset)

Parameter Estimate	$\alpha_{i,CAPM}$	$\beta_{i,r_{MKT}-r_f}$
Minimum	-0.0182	0.452
Maximum	0.0231	0.766
Median	0.0125	0.658
Mean	0.0111	0.648
SE(Mean)	0.0011	0.008
Standard Deviation	0.0083	0.059

Three-Factor Model Parameter Estimation (for Net Subset)

Parameter Estimate	$\alpha_{i,3-Factor}$	$\beta_{i,r_{MKT}-r_f}$	$\beta_{i,r_{SMB}}$	$\beta_{i,r_{HML}}$
Minimum	-0.0159	0.495	-0.191	-0.084
Maximum	0.0197	0.794	0.258	0.165
Median	0.0098	0.672	0.025	0.053
Mean	0.0081	0.659	0.012	0.050
SE(Mean)	0.0010	0.007	0.011	0.007
Standard Deviation	0.0073	0.054	0.080	0.050

Four-Factor Model Parameter Estimation (for Net Subset)

Parameter Estimate	$\alpha_{i,4-Factor}$	$\beta_{i,r_{MKT}-r_f}$	$\beta_{i,r_{SMB}}$	$\beta_{i,r_{HML}}$	$\beta_{i,r_{MOM}}$
Minimum	-0.0158	0.496	-0.098	-0.170	-0.152
Maximum	0.0217	0.795	0.262	0.168	0.144
Median	0.0076	0.674	0.045	0.057	0.046
Mean	0.0066	0.659	0.033	0.051	0.042
SE(Mean)	0.0010	0.008	0.010	0.008	0.008
Standard Deviation	0.0069	0.056	0.070	0.058	0.055

Upon inspection, the market sensitivity estimated by any of the three models (with means of 0.648, 0.659, and 0.659 for the CAPM, three-factor model, and four-factor model, respectively, and corresponding median values of 0.658, 0.672, and 0.674) is somewhat unsurprising. The public pension plan data set also provides an approximation of each plan-year's portfolio split between equity and fixed-income, and the equity percentage averages 53.2% and 52.2% by median and mean, respectively, across all net subset plan-years for which these figures are available. Naturally, such a statistic is likely to loosely track market sensitivity estimates, likely underestimating the true exposure, as the fixed-income portions of the portfolios are not strictly risk-free and may contribute to the total portfolios' market sensitivities. Of more interest, perhaps, are the results for the estimates of size, value, and momentum sensitivities. I find that US public pension plans tend not to load heavily on size, value, or momentum. The coefficient estimates for these factors (by both the three-factor and four-factor models for size and value) are modest in terms of mean and median and, for the vast majority of plans, lacking in statistical significance.

Further, the alpha estimates are notable. I find mean annual alpha estimates of 1.11%, 0.81%, and 0.66% for the CAPM, three-factor, and four-factor models, respectively, with corresponding median estimates of 1.19%, 0.98%, and 0.76%. All three models' results suggest that US public pension plan management possesses substantial skill. One alternative explanation is that the plans' return data is generally reported inaccurately with respect to time. If this were true, and returns lagged or led the reported figures in reality, analysis of this form would tend to underestimate exposure to the various risk factors, leading to large alpha estimates. However, upon running the analysis with added lag and lead terms for each risk factor, I find the coefficient estimates of these terms to be slightly negative, though statistically insignificant. Using these coefficient estimates, I measure plan-year performance as previously described, adjusting plan-year returns by exposure to risk factors corresponding to the three models of interest.

As mentioned, one method of ridding the results of look-ahead bias is to calculate rolling factor coefficients for each plan-year. Following this protocol, the plan-year performance estimates no longer incorporate future return data in any way. However, a larger, more fundamental statistical issue is worsened by this solution. Specifically, the linear models used to obtain factor coefficients, largely overfit even when fit to the entire span of the 2001-2028 data, become particularly overfit. This directly results from a tightening of restrictions on qualifying plan-years. With rolling data, for example, I estimate models over 2001-2002 to obtain backward rolling factor coefficients in order to estimate plan-year performance for plans with return data for 2002. Simply stated, the span of the data is bounded by the year for which the rolling regression factor coefficients are used to estimate performance.

A summary of the rolling regression parameter estimates follows:

Two-Year Rolling CAPM Estimates (for Net Subset)

Parameter Estimate	$\alpha_{i,CAPM}$	$\beta_{i,r_{MKT}-r_f}$
Minimum	-0.4322	-2.205
Maximum	1.0765	6.024
Median	0.1401	1.421
Mean	0.1471	1.274
SE(Mean)	0.0417	0.228
Standard Deviation	0.2948	1.580

Four-Year Rolling Three-Factor Estimates (For Net Subset)

Parameter Estimate	$\alpha_{i,3-Factor}$	$\beta_{i,r_{MKT}-r_f}$	$\beta_{i,r_{SMB}}$	$\beta_{i,r_{HML}}$
Minimum	-0.0208	0.225	-0.370	-0.334
Maximum	0.1011	0.793	0.713	0.138
Median	0.0221	0.623	0.080	-0.014
Mean	0.0259	0.616	0.063	-0.029
SE(Mean)	0.0037	0.016	0.030	0.012
Standard Deviation	0.0269	0.116	0.211	0.086

Five-Year Rolling Four-Factor Estimates (For Net Subset)

Parameter Estimate	$\alpha_{i,4-Factor}$	$\beta_{i,r_{MKT}-r_f}$	$\beta_{i,r_{SMB}}$	$\beta_{i,r_{HML}}$	$\beta_{i,r_{MOM}}$
Minimum	-0.0265	-0.218	-1.025	-0.821	-0.749
Maximum	0.2217	1.608	0.423	0.335	0.251
Median	0.0431	0.722	-0.196	-0.002	0.087
Mean	0.0482	0.677	-0.201	-0.030	0.038
SE(Mean)	0.0049	0.033	0.027	0.023	0.025
Standard Deviation	0.0351	0.237	0.195	0.161	0.176

In the tables above, summary statistics for the rolling parameter estimates are shown. For each model, results for regressions run on returns through the first year for which the number of data points exceeds the number of risk factors are included. These are chosen in order to highlight the overfitting of models fit to short spans of return data. Regressions run on any shorter span of data will result in coefficient estimates of zero on one or more risk factors, as an exact least squares solution does not require all factor coefficients to be nonzero. Turning to the risk factor coefficient estimates, the market sensitivities estimated over these shortest data spans are not entirely consistent with those fit to the entire span. Specifically, the means of those estimated by the CAPM and three-factor models are significantly different from those estimated using the full span of the data; though, of course, this could also be attributed to actual changes in the market sensitivities of public pension plan portfolios over time. However, both of these results show underestimates of market sensitivity relative to the full-data regressions while, as shown in the table below, the average estimated equity percentage of public pension plans in the net subset actually declines over time. This suggests that actual portfolio characteristics have moved in the opposite direction, providing some evidence that the observed differences between the model estimates using shortened data and the full regressions can be attributed to the severe overfitting. The most easily observed

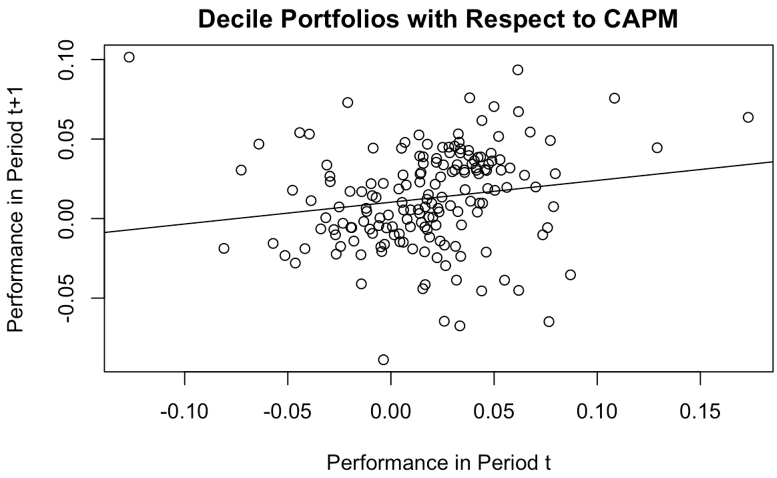
Mean Estimated Equity Proportion of Portfolio Over Time (Net Subset)

Year	2001	2002	2003	2004	2005	2006	2007	2008	2009
Est. Equity %	57.2	54.7	56.9	59.0	58.2	57.9	58.1	52.5	50.4
Year	2010	2011	2012	2013	2014	2015	2016	2017	2018
Est. Equity %	49.6	49.7	47.9	49.2	48.6	47.8	47.5	49.6	47.5

results of the overfitting are the vastly different alpha estimates between the different data spans. For the CAPM estimates, annual alphas range from -43.22% to 107.65%; for the three-factor model, they range from -2.08% to 10.11%; for the four-factor model, the alpha estimates take on values from -2.65% to 22.17%. Clearly, these extreme values are unrealistic, and mean and median annual alpha estimates are significantly higher than the corresponding estimates generated using the full span of return data. As for the additional risk factor coefficients in the three-factor and four-factor models, they take on wider ranges as well, though their mean and median values are not as drastically affected.

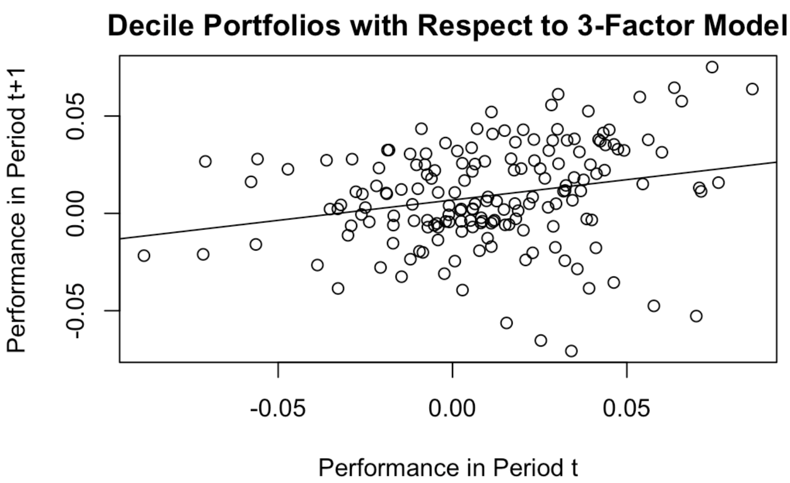
Generally, the three-factor and four-factor models used in either case are overfit, given that I use three and four explanatory variables to model four and five observations in the most restrictive cases and 18 observations in the most generous case. Statistically, this is well beyond reasonable bounds. As for the CAPM, estimating parameters using the longest few data spans is not unreasonable. In all, while both full regression and rolling regression techniques are flawed, they present a tradeoff between look-ahead bias and the severity of overfitting. It is therefore important to keep both the inevitable overfitting issue and this additional tradeoff in mind when considering the following analysis of performance.

Now, I turn to the persistence test method in which I form portfolios sorted on prior plan performance and regress future portfolio performance on this prior data. Specifically, for formation periods of one, two, three, four, and five years, I form quartile and decile portfolios for performance with respect to each of the three models. Consider first the results obtained using the parameter estimates for the full span of the data, beginning with formation period and performance horizon lengths of one year, highlighted in the three figures below. Evaluating for the net subset, while the coefficient values are positive for prior period performance for all three models, I find them to be insignificant, indicating a lack of persistence from period to period for this particular formation period and performance horizon length. The same is true when portfolios are formed by quartile as opposed to decile. Altering the formation period length and holding the performance horizon constant at one year, I find mixed results. For a formation period length of two years, the regression estimates are significant for all models and both quantile types. However, from the two longest formation period lengths (see columns three and four of row one in the tables below), the effect is insignificant. When I extend both the formation period and performance horizon to the same length, I find significant coefficients on past performance for lengths of two, four, and five years. Tables containing the entire set of coefficients on past performance are found below (quartile results are found in the Appendix).



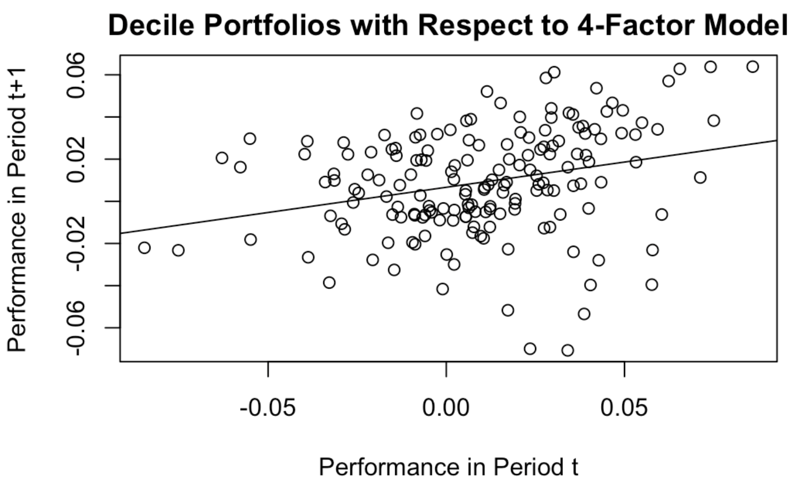
Decile Portfolio Performance in *t+1* (CAPM)	
(Intercept)	0.01***
	(0.00)
Performance in *t*	0.14*
	(0.06)
R^2	0.03
Adj. R^2	0.02
Num. obs.	170
RMSE	0.03

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$



Decile Portfolio Performance in *t+1* (3-Factor)	
(Intercept)	0.01^{***}
	(0.00)
Performance in t	0.21^{**}
	(0.06)
R^2	0.06
Adj. R^2	0.05
Num. obs.	170
RMSE	0.03

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$



Decile Portfolio Performance in *t+1* (4-Factor)	
(Intercept)	0.01^{***}
	(0.00)
Performance in t	0.24^{***}
	(0.06)
R^2	0.08
Adj. R^2	0.07
Num. obs.	170
RMSE	0.02

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

Past Performance Coefficient Estimates (CAPM Performance Decile Sorting; *p<.05)

	1-Yr. Formation Period	2-Yr. Formation Period	3-Yr. Formation Period
1-Yr. Horizon	0.137	0.379*	0.465*
2-Yr. Horizon	0.242*	0.350*	0.413*
3-Yr. Horizon	0.227*	0.319*	0.186
4-Yr. Horizon	-0.005	0.285*	0.339*
5-Yr. Horizon	0.175*	0.265*	0.147*

	4-Yr. Formation Period	5-Yr. Formation Period
1-Yr. Horizon	0.496	0.483
2-Yr. Horizon	0.433*	0.410*
3-Yr. Horizon	0.348*	0.353*
4-Yr. Horizon	0.343*	0.312*
5-Yr. Horizon	0.303*	0.309*

Past Performance Coefficient Estimates (3-Factor Performance Decile Sorting; *p<.05)

	1-Yr. Formation Period	2-Yr. Formation Period	3-Yr. Formation Period
1-Yr. Horizon	0.209	0.410*	0.509*
2-Yr. Horizon	0.223*	0.402*	0.455*
3-Yr. Horizon	0.216*	0.351*	0.220
4-Yr. Horizon	0.171*	0.318*	0.373*
5-Yr. Horizon	0.161*	0.285*	0.312*

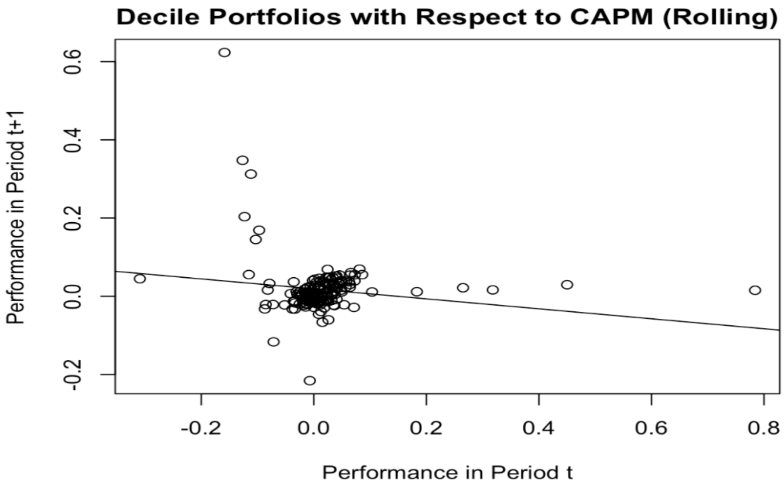
	4-Yr. Formation Period	5-Yr. Formation Period
1-Yr. Horizon	0.510*	0.526
2-Yr. Horizon	0.475*	0.450*
3-Yr. Horizon	0.430*	0.374*
4-Yr. Horizon	0.354*	0.335*
5-Yr. Horizon	0.322*	0.303*

Past Performance Coefficient Estimates (4-Factor Performance Decile Sorting; *p<.05)

	1-Yr. Formation Period	2-Yr. Formation Period	3-Yr. Formation Period
1-Yr. Horizon	0.239	0.472*	0.560*
2-Yr. Horizon	0.298*	0.440*	0.477*
3-Yr. Horizon	0.264*	0.364*	0.214*
4-Yr. Horizon	0.202*	0.332*	0.387*
5-Yr. Horizon	0.195*	0.284*	0.262*

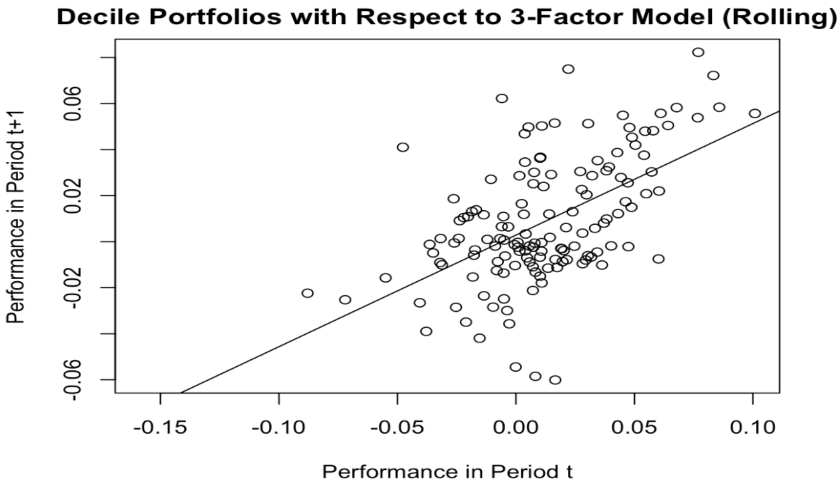
	4-Yr. Formation Period	5-Yr. Formation Period
1-Yr. Horizon	0.556*	0.618*
2-Yr. Horizon	0.475*	0.450*
3-Yr. Horizon	0.452*	0.436*
4-Yr. Horizon	0.406*	0.410*
5-Yr. Horizon	0.377*	0.364*

Comparing these results to those obtained using the rolling parameter estimates, I find fairly similar results. The relationship between portfolio performance in a given year and the next, as displayed in the three following figures, is not statistically significant for decile or quartile portfolios formed on performance estimated by the CAPM; though, the coefficient estimates of these relationships are significantly positive for the three-factor and four-factor models, indicating persistence for these formation period and performance horizon lengths. Extending the performance horizon to two years and holding the formation period to one year, I find the same to be true. This pattern holds as performance horizon is lengthened to three, four, and five years; in each case, evidence of persistence is found when persistence is evaluated using three-factor and four-factor models, though not when using the CAPM, regardless of quantile type. Lengthening the formation period, I find that, for formation periods of two years and one-year performance horizons, the rolling analysis yields significant, positive coefficients for performance regressed on sorting-period horizon regardless of quantile type or model selection. This trend too holds as formation period length is generalized; for formation periods between two and five years, I find significant, positive coefficients for past performance. Changing the two lengths in tandem, I find insignificant evidence of persistence in regressions for 3 year lengths when performance is estimated via the four-factor model and 5 year lengths when performance is estimated via either the three-factor or four-factor model. The coefficient estimates from all tests are provided below (quartile results are found in the Appendix).



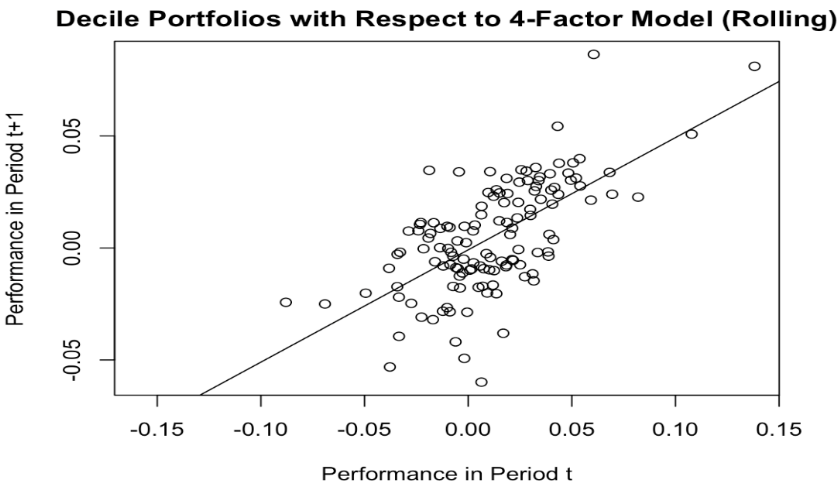
Rolling Decile Portfolio Performance in *t+1* (CAPM)	
(Intercept)	0.02^{***}
	(0.01)
Performance in t	
-0.13^{*}	
	(0.06)
R^2	0.03
Adj. R^2	0.02
Num. obs.	170
RMSE	0.07

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$



Rolling Decile Portfolio Performance in *t+1* (3-Factor)	
(Intercept)	0.00
	(0.00)
Performance in *t*	0.48***
	(0.06)
R^2	0.30
Adj. R^2	0.30
Num. obs.	140
RMSE	0.03

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$



Rolling Decile Portfolio Performance in *t+1* (4-Factor)	
(Intercept)	0.00
	(0.00)
Performance in t	0.50***
	(0.05)
R^2	0.41
Adj. R^2	0.41
Num. obs.	130
RMSE	0.02

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

Rolling Past Performance Coefficient Estimates (CAPM Performance Decile Sorting; *p<.05)

	1-Yr. Formation Period	2-Yr. Formation Period	3-Yr. Formation Period
1-Yr. Horizon	0.127	0.108*	0.253*
2-Yr. Horizon	-0.045	0.123*	0.238*
3-Yr. Horizon	-0.011	0.124*	0.130*
4-Yr. Horizon	0.005	0.111*	0.179*
5-Yr. Horizon	-0.006	0.098*	0.147*

	4-Yr. Formation Period	5-Yr. Formation Period
1-Yr. Horizon	0.331*	0.305*
2-Yr. Horizon	0.276*	0.267*
3-Yr. Horizon	0.242*	0.218*
4-Yr. Horizon	0.198*	0.178*
5-Yr. Horizon	0.236*	0.131*

Rolling Past Performance Coefficient Estimates (3-Factor Performance Decile Sorting; *p<.05)

	1-Yr. Formation Period	2-Yr. Formation Period	3-Yr. Formation Period
1-Yr. Horizon	0..485*	0.536*	0.481*
2-Yr. Horizon	0.404*	0.450*	0.382*
3-Yr. Horizon	0.351*	0.385*	0.276*
4-Yr. Horizon	0.299*	0.339*	0.333*
5-Yr. Horizon	0.266*	0.338*	0.332*

	4-Yr. Formation Period	5-Yr. Formation Period
1-Yr. Horizon	0.372*	0.296*
2-Yr. Horizon	0.312*	0.282*
3-Yr. Horizon	0.309*	0.205*
4-Yr. Horizon	0.264*	0.141
5-Yr. Horizon	0.191*	0.104

Rolling Past Performance Coefficient Estimates (4-Factor Performance Decile Sorting; *p<.05)

	1-Yr. Formation Period	2-Yr. Formation Period	3-Yr. Formation Period
1-Yr. Horizon	0.501*	0.461*	0.467*
2-Yr. Horizon	0.404*	0.398*	0.406*
3-Yr. Horizon	0.357*	0.360*	0.196
4-Yr. Horizon	0.315*	0.307*	0.272*
5-Yr. Horizon	0.287*	0.283*	0.262*

	4-Yr. Formation Period	5-Yr. Formation Period
1-Yr. Horizon	0.472*	0.351
2-Yr. Horizon	0.353*	0.155
3-Yr. Horizon	0.243*	0.080
4-Yr. Horizon	0.189*	0.038
5-Yr. Horizon	0.163*	0.104

Despite the significance of the regression coefficients in the majority of tests, generally, these results provide weak evidence for performance persistence, given the severity of overfitting and the look-ahead bias present in the analysis that incorporates risk-factor coefficients estimated using the entire span of the return data. While both the full regression and rolling regression methods can improve upon the other's weakness, in either case, the results are of little meaning and must be interpreted with these complications in mind.

Notably, evaluation of the regression results requires the adjustment of standard errors and, ultimately, the p-values corresponding to each coefficient estimate due to the use of overlapping data. Each observation of performance in one formation period and the subsequent performance horizon contains performance estimates that overlap with those of one or more other observations, depending on the combined length of the formation period and performance horizon. For example, the observations of portfolios created by sorting on performance in 2001 and observing performance in 2002 use performance estimates from 2001 and 2002. This 2002 information is also incorporated into the observations that sort on 2002 performance. Thus, autocorrelation is an issue in analysis of these observations. I confront this through the use of the Newey-West Estimator, which serves the purpose of correcting regression results in which autocorrelation is present in the residuals.

Additionally, the results obtained through the previously described transition matrix approach are of interest. These results not only provide a second, more categorical look at persistence, but also indicate at which end of the spectrum persistence, if any, tends to occur. Are the particularly high-performing plans continually doing so, or do the low-performers seem to struggle from period to period? First consider the set of matrices below; these correspond to the performance estimates generated using the risk-factor coefficients estimated from the entire span of the data (in simpler terms, not rolling estimates). Out of concern for spacing, I present the matrices involving quartile sorts and leave the matrices involving decile sorts for the Appendix. In reading these matrices, note that, for matrix A, entry A_{ij} in row i and column j represents the number of portfolios that fell in quantile i in the formation period and in quantile j over the performance horizon. Performance is decreasing in i and j, meaning that the top quantile in both the formation period and performance horizon corresponds to the upper-leftmost entry. I find that, for matrices corresponding to formation periods and performance horizons ranging from one year to two years, the majority of statistically significant elements lie near or on the main diagonal, indicating that relative performance in the formation period is more likely to be followed by similar relative performance over the performance horizon. In each matrix, at least one value on the main diagonal is significant. Additionally, it appears that most of this relative performance occurs near the bottom (in the sense of relative evaluation, not necessarily visually). These patterns hold as either period length is extended beyond two years. Though, evidence for particularly bottom-driven persistence does not appear when analyzing portfolios formed on deciles.

Quartile Transition Matrices for 1-Yr. Formation Period, 1-Yr. Performance Horizon (CAPM, 3-Factor, 4-Factor performance left to right; *p<.05)

$$\begin{bmatrix} 8* & 2 & 1 & 6 \\ 2 & 8* & 7 & 0 \\ 5 & 2 & 8* & 2 \\ 2 & 5 & 1 & 9* \end{bmatrix} \begin{bmatrix} 7 & 5 & 2 & 3 \\ 3 & 4 & 7 & 3 \\ 3 & 4 & 7 & 3 \\ 4 & 4 & 1 & 8* \end{bmatrix} \begin{bmatrix} 7 & 5 & 3 & 2 \\ 3 & 4 & 7 & 3 \\ 3 & 4 & 7 & 3 \\ 4 & 4 & 1 & 8* \end{bmatrix}$$

Quartile Transition Matrices for 1-Yr. Formation Period, 2-Yr. Performance Horizon (CAPM, 3-Factor, 4-Factor performance left to right; *p<.05)

$$\begin{bmatrix} 7 & 2 & 3 & 4 \\ 3 & 7 & 4 & 2 \\ 5 & 2 & 7 & 2 \\ 1 & 5 & 2 & 8* \end{bmatrix} \begin{bmatrix} 6 & 5 & 2 & 3 \\ 4 & 5 & 6 & 1 \\ 4 & 4 & 8* & 0 \\ 2 & 2 & 0 & 12* \end{bmatrix} \begin{bmatrix} 6 & 5 & 4 & 1 \\ 8* & 4 & 3 & 1 \\ 2 & 5 & 6 & 3 \\ 0 & 2 & 3 & 11* \end{bmatrix}$$

Quartile Transition Matrices for 2-Yr. Formation Period, 1-Yr. Performance Horizon (CAPM, 3-Factor, 4-Factor performance left to right; *p<.05)

$$\begin{bmatrix} 6 & 5 & 1 & 4 \\ 3 & 7 & 5 & 1 \\ 5 & 1 & 9* & 1 \\ 2 & 3 & 1 & 10* \end{bmatrix} \begin{bmatrix} 9* & 1 & 3 & 3 \\ 2 & 8* & 4 & 2 \\ 4 & 4 & 5 & 3 \\ 1 & 3 & 4 & 8* \end{bmatrix} \begin{bmatrix} 8* & 1 & 5 & 2 \\ 5 & 5 & 5 & 1 \\ 2 & 7 & 5 & 2 \\ 1 & 3 & 1 & 11* \end{bmatrix}$$

Quartile Transition Matrices for 2-Yr. Formation Period, 2-Yr. Performance Horizon (CAPM, 3-Factor, 4-Factor performance left to right; *p<.05)

$$\begin{bmatrix} 7* & 2 & 3 & 3 \\ 2 & 5 & 6 & 2 \\ 4 & 5 & 5 & 1 \\ 2 & 3 & 1 & 9* \end{bmatrix} \begin{bmatrix} 7* & 3 & 4 & 1 \\ 5 & 3 & 6 & 1 \\ 2 & 8* & 4 & 1 \\ 1 & 1 & 1 & 12* \end{bmatrix} \begin{bmatrix} 9* & 1 & 4 & 1 \\ 4 & 7* & 4 & 0 \\ 2 & 4 & 6 & 3 \\ 0 & 3 & 1 & 11* \end{bmatrix}$$

Turning to the next set of matrices, which involve performance estimates determined using rolling risk factor sensitivities, I find very similar results, though the significant entries appear more evenly spread along the upper and lower ends of the main diagonals. Note that the sum of all entries in matrices corresponding to the models with more factors is lower; this results from the fact that the models produce performance estimates of zero for all plan-years that occur before the number of rolling observations surpasses the number of factors in the model, as the least squares solutions are guaranteed to be exact. Interestingly, in either case, model choice has little impact on the results. Neither the matrices corresponding to performance with respect to the CAPM, three-factor model, or four-factor model contain particularly many significant entries (for both the quartile and decile sets). Though, notably, the only matrix of this set with no significant entries along the main diagonal is the one-year formation period and one-year performance horizon matrix that uses CAPM coefficient-calculated performance. This is likely the result of the wide-ranging coefficient estimates observed earlier for the corresponding regressions, products of the severe overfitting, in turn. Again, these generally strong matrix results must be interpreted with overfitting and look-ahead bias (in the case of the first set) and severe overfitting (in the case of the second set) in mind. While suggestive, they don't provide a statistically robust sense of direction.

Quartile Rolling Transition Matrices for 1-Yr. Formation Period, 1-Yr. Performance Horizon (CAPM, 3-Factor, 4-Factor performance left to right; *p<.05)

$$\begin{bmatrix} 5 & 5 & 4 & 3 \\ 3 & 7 & 3 & 4 \\ 6 & 3 & 4 & 4 \\ 3 & 2 & 6 & 6 \end{bmatrix} \begin{bmatrix} 7* & 1 & 3 & 3 \\ 4 & 5 & 2 & 3 \\ 0 & 6 & 4 & 4 \\ 3 & 2 & 5 & 4 \end{bmatrix} \begin{bmatrix} 7* & 4 & 1 & 1 \\ 3 & 5 & 1 & 4 \\ 1 & 2 & 6* & 4 \\ 2 & 2 & 5 & 4 \end{bmatrix}$$

Quartile Rolling Transition Matrices for 1-Yr. Formation Period, 2-Yr. Performance Horizon (CAPM, 3-Factor, 4-Factor performance left to right; *p<.05)

$$\begin{bmatrix} 5 & 3 & 6 & 3 \\ 2 & 9* & 2 & 3 \\ 5 & 4 & 5 & 2 \\ 4 & 1 & 3 & 8* \end{bmatrix} \begin{bmatrix} 6* & 2 & 2 & 3 \\ 4 & 6* & 1 & 2 \\ 2 & 4 & 5 & 2 \\ 1 & 1 & 5 & 6* \end{bmatrix} \begin{bmatrix} 6* & 4 & 1 & 1 \\ 3 & 5 & 3 & 1 \\ 1 & 2 & 5 & 4 \\ 2 & 1 & 3 & 6* \end{bmatrix}$$

Quartile Rolling Transition Matrices for 2-Yr. Formation Period, 1-Yr. Performance Horizon (CAPM, 3-Factor, 4-Factor performance left to right; *p<.05)

$$\begin{bmatrix} 4 & 4 & 4 & 4 \\ 4 & 7 & 4 & 1 \\ 4 & 4 & 6 & 2 \\ 4 & 1 & 2 & 9* \end{bmatrix} \begin{bmatrix} 7* & 2 & 2 & 2 \\ 4 & 4 & 3 & 2 \\ 1 & 4 & 5 & 3 \\ 1 & 3 & 3 & 6* \end{bmatrix} \begin{bmatrix} 6* & 5 & 0 & 1 \\ 2 & 5 & 2 & 3 \\ 3 & 2 & 6* & 1 \\ 1 & 0 & 4 & 7* \end{bmatrix}$$

Quartile Rolling Transition Matrices for 2-Yr. Formation Period, 2-Yr. Performance Horizon (CAPM, 3-Factor, 4-Factor performance left to right; *p<.05)

$$\begin{bmatrix} 4 & 5 & 2 & 4 \\ 3 & 6 & 5 & 1 \\ 5 & 3 & 7* & 0 \\ 3 & 1 & 1 & 10* \end{bmatrix} \begin{bmatrix} 7* & 2 & 2 & 1 \\ 3 & 7* & 2 & 0 \\ 1 & 2 & 7* & 2 \\ 1 & 1 & 1 & 9* \end{bmatrix} \begin{bmatrix} 8* & 1 & 1 & 4 \\ 2 & 3 & 5 & 1 \\ 1 & 6 & 3 & 1 \\ 0 & 1 & 2 & 8* \end{bmatrix}$$

5 Fee Analysis

Additionally, I inspect the reported external managers' fees for the same set of plan-years, separating these figures into those explicitly reported as management fees, those explicitly reported as performance fees, and those reported under another title. Within this section, the results are no longer particular to the net subset. A more important distinction, given inconsistency across plan-years in fee reporting, is the convention of reporting. I separate the data into two sets: total fees for plan-years in which both management and performance fees are explicitly reported and management fees for plan-years in which at least management fees are explicitly reported. I then normalize these fees to annual percentages of plan assets, estimated by actuarial value.

I find that, for the management and performance fee-reporting group, US public pension plans pay an average of 0.542% of assets in external fees annually. The corresponding total fee average (including fees under titles other than management or performance) for the group that at least reports management fees explicitly is 0.33%. Paring this number down to only explicit management fees, I find that the average plan pays 0.301% of assets in explicit management fees annually. Inspecting further, I find no significant relationship between the fee-reporting convention and the annual return-reporting convention (net or gross).

While complete plan-average annual fee percentages range from 0.414% to 1.72% (with first and third quartile values of 0.648% and 1.067%, respectively), and percentages corresponding to explicit management fees range from .0244% to 4.59% (with first and third quartile values of 0.227% and 0.474%, respectively), it appears that certain plan-year investment variables have significant explanatory power with regard to this variation. Running a fixed-effects model for management fees on plan funding ratio (defined as actuarial assets divided by actuarial liabilities), total plan assets (expressed in millions of dollars), and rolling market sensitivity estimate (with respect to the CAPM) with dummy variables for year and plan identification, I find that funding ratio and total plan assets have significant relationships with explicit management fees (expressed as a percentage of total assets). Specifically, I find that lower funding ratios are associated with lower fees and larger assets totals are associated with higher fees (despite analyzing fees as a normalized figure). The details of the model are as follows:

Fixed Effects (Plan and Year) Model for Management Fees	
Funding Ratio	-0.00471^{***}
	(0.00034)
Total Assets	0.00004^{***}
	(0.00001)
Rolling Mkt. Beta	0.00007
	(0.00011)
R^2	0.15601
Adj. R^2	0.08059
Num. obs.	1476

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

Applying an equivalent test to the subset of plan-years for which both management fees and performance fees are explicitly reported, I find that, upon inclusion of fixed-effect terms for year and plan identity, the effects of funding ratio and total plan assets on fees are insignificant. This lack of statistical significance may be a product of the far smaller sample size of complete-reporting plan-years compared to the less strictly-defined management fee-reporting set (90 plan-years as opposed 1,578 plan-years, respectively). The details of this model are as follows:

Fixed Effects (Plan and Year) Model for Complete Fees	
Funding Ratio	−0.00688
	(0.00505)
Total Assets	−0.00000
	(0.00002)
Rolling Mkt. Beta	−0.00069
	(0.00149)
R^2	0.05970
Adj. R^2	-0.23068
Num. obs.	90

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

I additionally test fees as predictors of investment-related variables such as future market risk-taking, future risk-taking (as measured by excess return volatility), and future performance (with respect to all three models) via least-squares regression, though I find no significant results. The same is true of tests of these same investment-related variables as explanatory variables of fees: I find no significant coefficient estimates by least-squares.

6 Further Discussion

In all, the evaluation of performance persistence amongst US public pension plans appears largely unanswered due to limitations in the data. While my analysis points weakly toward persistence within this group, and, theoretically, pension plans' capital flows should not do away with net alpha, given their lack of relation to manager performance, the annual data is too coarse to avoid overfitting on a time span as short as 2001-2018. Mitigation of the overfitting by estimating plans' risk-factor sensitivities over the full span of the data, introduces look-ahead bias and fails to entirely solve the overfitting issue for any multifactor asset pricing model. To address the question in a more consequential manner, an annual data set spanning many more years or a set of monthly returns (neither of which exist or are publicly available, to my knowledge) would be required.

As with any such analysis, these results also present a joint-hypothesis issue. My tests of persistence are reliant on the acceptance of a particular asset-pricing model (in this case, I defer to prior research and choose three). Therefore, significant results can either be interpreted as evidence of the phenomenon of interest or, alternatively, failure of the underlying asset-pricing model. To disentangle these two hypotheses is impossible; to test one implies agreement with the other.

With regard to fees, I am able to get some sense of their magnitude and slight insight into its related metrics. Fee totals, which are difficult to aggregate, given the inconsistent nature of US public pension plan reporting, are also very difficult to put into perspective. In any given year, many plans pay hundreds of millions of dollars in external management fees alone. Across the entire US public pension system, this total sums to tens of billions for recent years. However, generally, it appears that, at the plan-level, paying external managers has little impact on after-fee performance. Going forward, a deeper causal analysis of the employment of external managers, along with expansion of performance persistence tests, given access to more granular return data, appear to be paths down which substantial questions may be answered.

References

Agarwal, Vikas and Narayan Y. Naik (2000). “Multi-Period Performance Persistence Analysis of Hedge Funds”. In: *Journal of Financial and Quantitative Analysis* 35.3, pp. 327–342. DOI: 10.2307/2676207. URL: https://www.cambridge.org/core/article/multiperiod-performance-persistence-analysis-of-hedge-funds/B39621DA520EC95E2E0BD29A594C834D.

Berk, Jonathan B and Richard C Green (2004). “Mutual Fund Flows and Performance in Rational Markets”. In: *Journal of Political Economy* 112.6, pp. 1269–1295. DOI: 10.1086/424739. URL: https://doi.org/10.1086/424739.

Bollen, Nicolas P. B. and Jeffrey A. Busse (2004). “Short-Term Persistence in Mutual Fund Performance”. In: *The Review of Financial Studies* 18.2, pp. 569–597. DOI: 10.1093/rfs/hhi007. URL: https://doi.org/10.1093/rfs/hhi007.

Brian, Alex and James Li (2016). *Performance Persistence Among U.S. Mutual Funds*. Tech. rep.

Brown, Gavin, Paul Draper, and Eddie McKenzie (1997). “Consistency of UK Pension Fund Investment Performance”. In: *Journal of Business Finance Accounting* 24.2, pp. 155–178. DOI: 10.1111/1468-5957.00100. URL: https://doi.org/10.1111/1468-5957.00100.

Carhart, Mark M. (1997). “On Persistence in Mutual Fund Performance”. In: *The Journal of Finance* 52.1, pp. 57–82. DOI: 10.2307/2329556. URL: http://www.jstor.org.proxy.library.upenn.edu/stable/2329556.

French, Kenneth (2020). *Kenneth French Data Library*. URL: http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html.

GRUBER, MARTIN J. (1996). “Another Puzzle: The Growth in Actively Managed Mutual Funds”. In: *The Journal of Finance* 51.3, pp. 783–810. DOI: 10.1111/j.1540-6261.1996.tb02707.x. URL: https://doi.org/10.1111/j.1540-6261.1996.tb02707.x.

How many American workers participate in workplace retirement plans?¡br¿ (July 2019). URL: https://www.pensionrights.org/publications/statistic/how-many-american-workers-participate-workplace-retirement-plans.

Huang, X. and R. J. Mahieu (2012). *Performance persistence of Dutch pension plans.* ID: alma9977048908003681.

KAPLAN, STEVEN N. and ANTOINETTE SCHOAR (2005). "Private Equity Performance: Returns, Persistence, and Capital Flows". In: *The Journal of Finance* 60.4, pp. 1791–1823. DOI: 10.1111/j.1540-6261.2005.00780.x. URL: https://doi.org/10.1111/j.1540-6261.2005.00780.x.

Lakonishok, Josef et al. (1992). "The Structure and Performance of the Money Management Industry". In: *Brookings Papers on Economic Activity.Microeconomics* 1992, pp. 339–391. DOI: 10.2307/2534766. URL: http://www.jstor.org/stable/2534766.

Malkiel, Burton G. (1995). "Returns from Investing in Equity Mutual Funds 1971 to 1991". In: *The Journal of Finance* 50.2, pp. 549–572. DOI: 10.2307/2329419. URL: http://www.jstor.org/stable/2329419.

Munnell, Alicia et al. (n.d.). *Public Plans Data.* URL: https://publicplansdata.org.

Public Pension Assets (2019). URL: https://www.nasra.org/content.asp?admin=Y&contentid=200.

State and Local Government Pensions (2018). URL: https://www.urban.org/policy-centers/cross-center-initiatives/state-and-local-finance-initiative/projects/state-and-local-backgrounders/state-and-local-government-pensions.

Tonks, Ian (2005). "Performance Persistence of Pension-Fund Managers". In: *The Journal of Business* 78.5, pp. 1917–1942. DOI: 10.1086/431447. URL: http://www.jstor.org/stable/10.1086/431447.

Zheng, Lu (1999). "Is Money Smart? A Study of Mutual Fund Investors' Fund Selection Ability". In: *The Journal of Finance* 54.3, pp. 901–933. DOI: 10.1111/0022-1082.00131. URL: https://doi.org/10.1111/0022-1082.00131.

7 Appendix

CAPM Parameter Estimation (Aggregate)

Parameter Estimate	$\alpha_{i,CAPM}$	$\beta_{i,r_{MKT}-r_f}$
Minimum	-0.0182	0.097
Maximum	0.0496	0.798
Median	0.0119	0.643
Mean	0.0121	0.626
SE(Mean)	0.0007	0.007
Standard Deviation	0.0082	0.080

Three-Factor Model Parameter Estimation (Aggregate)

Parameter Estimate	$\alpha_{i,3-Factor}$	$\beta_{i,r_{MKT}-r_f}$	$\beta_{i,r_{SMB}}$	$\beta_{i,r_{HML}}$
Minimum	-0.0159	0.104	-0.280	-0.105
Maximum	0.0492	0.825	0.352	0.236
Median	0.0092	0.652	0.025	0.047
Mean	0.0089	0.637	0.024	0.048
SE(Mean)	0.0006	0.006	0.007	0.004
Standard Deviation	0.0079	0.079	0.089	0.052

Four-Factor Model Parameter Estimation (Aggregate)

Parameter Estimate	$\alpha_{i,4-Factor}$	$\beta_{i,r_{MKT}-r_f}$	$\beta_{i,r_{SMB}}$	$\beta_{i,r_{HML}}$	$\beta_{i,r_{MOM}}$
Minimum	-0.0158	0.120	-0.268	-0.170	-0.520
Maximum	0.0473	0.825	0.291	0.241	0.174
Median	0.0073	0.655	0.046	0.052	0.045
Mean	0.0075	0.639	0.040	0.050	0.035
SE(Mean)	0.0007	0.006	0.007	0.005	0.006
Standard Deviation	0.0082	0.079	0.080	0.056	0.069

Two-Year Rolling CAPM Estimates (Aggregate)

Parameter Est.	$\alpha_{i,CAPM}$	$\beta_{i,r_{MKT-r_f}}$
Minimum	-1.2471	-6.226
Maximum	1.0765	6.024
Median	0.1598	1.289
Mean	0.1528	1.261
SE(mean)	0.0240	0.129
Standard Dev.	0.2913	1.558

Four-Year Rolling Three-Factor Estimates (Aggregate)

Parameter Estimate	$\alpha_{i,3-Factor}$	$\beta_{i,r_{MKT-r_f}}$	$\beta_{i,r_{SMB}}$	$\beta_{i,r_{HML}}$
Minimum	-0.0345	-0.112	-1.559	-0.341
Maximum	0.2382	0.816	0.713	0.264
Median	0.0231	0.621	0.058	-0.025
Mean	0.0314	0.601	0.025	-0.038
SE(mean)	0.0027	0.010	0.020	0.008
Standard Deviation	0.0325	0.125	0.240	0.100

Five-Year Rolling Four-Factor Estimates (Aggregate)

Parameter Estimate	$\alpha_{i,4-Factor}$	$\beta_{i,r_{MKT-r_f}}$	$\beta_{i,r_{SMB}}$	$\beta_{i,r_{HML}}$	$\beta_{i,r_{MOM}}$
Minimum	-0.0994	-1.084	-1.095	-1.268	-1.275
Maximum	0.3828	1.608	0.715	0.451	0.709
Median	0.0430	0.677	-0.163	-0.009	0.070
Mean	0.0490	0.632	-0.158	-0.045	0.018
SE(mean)	0.0038	0.020	0.017	0.016	0.017
Standard Deviation	0.0458	0.240	0.212	0.188	0.207

Past Performance Coefficient Estimates (CAPM Performance Quartile Sorting; *p<.05)

	1-Yr. Formation Period	2-Yr. Formation Period	3-Yr. Formation Period
1-Yr. Horizon	0.270	0.398*	0.503*
2-Yr. Horizon	0.248	0.375*	0.437*
3-Yr. Horizon	0.235*	0.344*	0.205
4-Yr. Horizon	0.199*	0.313*	0.374*
5-Yr. Horizon	0.194*	0.299*	0.153*

	4-Yr. Formation Period	5-Yr. Formation Period
1-Yr. Horizon	0.522	0.496
2-Yr. Horizon	0.462*	0.426*
3-Yr. Horizon	0.431*	0.374*
4-Yr. Horizon	0.378*	0.389*
5-Yr. Horizon	0.334*	0.330*

Past Performance Coefficient Estimates (3-Factor Performance Quartile Sorting; *p<.05)

	1-Yr. Formation Period	2-Yr. Formation Period	3-Yr. Formation Period
1-Yr. Horizon	0.224	0.435*	0.536*
2-Yr. Horizon	0.574*	0.428*	0.476*
3-Yr. Horizon	0.222*	0.377*	0.238
4-Yr. Horizon	0.206*	0.348*	0.403*
5-Yr. Horizon	0.171*	0.308*	0.314*

	4-Yr. Formation Period	5-Yr. Formation Period
1-Yr. Horizon	0.523	0.557*
2-Yr. Horizon	0.491*	0.472*
3-Yr. Horizon	0.443*	0.403*
4-Yr. Horizon	0.380*	0.367*
5-Yr. Horizon	0.345*	0.335*

Past Performance Coefficient Estimates (4-Factor Performance Quartile Sorting; *p<.05)

	1-Yr. Formation Period	2-Yr. Formation Period	3-Yr. Formation Period
1-Yr. Horizon	0.224	0.501*	0.581*
2-Yr. Horizon	0.307*	0.464*	0.490*
3-Yr. Horizon	0.268*	0.385*	0.221*
4-Yr. Horizon	0.178*	0.351*	0.406*
5-Yr. Horizon	0.204*	0.465*	0.278*

	4-Yr. Formation Period	5-Yr. Formation Period
1-Yr. Horizon	0.560*	0.633*
2-Yr. Horizon	0.519*	0.505*
3-Yr. Horizon	0.461*	0.465*
4-Yr. Horizon	0.424*	0.444*
5-Yr. Horizon	0.399*	0.394*

Rolling Past Performance Coefficient Estimates (CAPM Performance Quartile Sorting; *p<.05)

	1-Yr. Formation Period	2-Yr. Formation Period	3-Yr. Formation Period
1-Yr. Horizon	0.127	0.108*	0.253*
2-Yr. Horizon	-0.045	0.123*	0.238*
3-Yr. Horizon	-0.011	0.124*	0.130*
4-Yr. Horizon	0.005	0.111*	0.179*
5-Yr. Horizon	-0.006	0.098*	0.147*

	4-Yr. Formation Period	5-Yr. Formation Period
1-Yr. Horizon	0.331*	0.305*
2-Yr. Horizon	0.276*	0.267*
3-Yr. Horizon	0.242*	0.218*
4-Yr. Horizon	0.198*	0.178*
5-Yr. Horizon	0.236*	0.131*

Rolling Past Performance Coefficient Estimates (3-Factor Performance Quartile Sorting; *p<.05)

	1-Yr. Formation Period	2-Yr. Formation Period	3-Yr. Formation Period
1-Yr. Horizon	0..485*	0.536*	0.481*
2-Yr. Horizon	0.404*	0.450*	0.382*
3-Yr. Horizon	0.351*	0.385*	0.276*
4-Yr. Horizon	0.299*	0.339*	0.333*
5-Yr. Horizon	0.266*	0.338*	0.332*

	4-Yr. Formation Period	5-Yr. Formation Period
1-Yr. Horizon	0.372*	0.296*
2-Yr. Horizon	0.312*	0.282*
3-Yr. Horizon	0.309*	0.205*
4-Yr. Horizon	0.264*	0.141
5-Yr. Horizon	0.191*	0.104

Rolling Past Performance Coefficient Estimates (4-Factor Performance Quartile Sorting; *p<.05)

	1-Yr. Formation Period	2-Yr. Formation Period	3-Yr. Formation Period
1-Yr. Horizon	0.501*	0.461*	0.467*
2-Yr. Horizon	0.404*	0.398*	0.406*
3-Yr. Horizon	0.357*	0.360*	0.196
4-Yr. Horizon	0.315*	0.307*	0.272*
5-Yr. Horizon	0.287*	0.283*	0.262*

	4-Yr. Formation Period	5-Yr. Formation Period
1-Yr. Horizon	0.472*	0.351
2-Yr. Horizon	0.353*	0.155
3-Yr. Horizon	0.243*	0.080
4-Yr. Horizon	0.189*	0.038
5-Yr. Horizon	0.163*	0.104

Decile Transition Matrices for 1-Yr. Formation Period, 1-Yr. Performance Horizon (in order of CAPM, 3-Factor, 4-Factor performance; *p<.05)

$$\begin{bmatrix} 5* & 2 & 1 & 1 & 2 & 1 & 0 & 0 & 3 & 2 \\ 1 & 3 & 1 & 2 & 1 & 1 & 1 & 3 & 1 & 3 \\ 0 & 0 & 4* & 3 & 1 & 2 & 2 & 1 & 2 & 2 \\ 2 & 3 & 1 & 6* & 1 & 1 & 2 & 0 & 0 & 1 \\ 1 & 2 & 2 & 1 & 2 & 3 & 3 & 2 & 1 & 0 \\ 3 & 0 & 0 & 2 & 4* & 1 & 2 & 3 & 1 & 1 \\ 0 & 1 & 2 & 2 & 1 & 4* & 2 & 2 & 2 & 1 \\ 0 & 1 & 2 & 0 & 1 & 2 & 3 & 2 & 2 & 4* \\ 0 & 4* & 2 & 0 & 2 & 2 & 2 & 3 & 2 & 0 \\ 5* & 1 & 2 & 0 & 2 & 0 & 0 & 1 & 3 & 3 \end{bmatrix} \begin{bmatrix} 5* & 3 & 0 & 0 & 1 & 0 & 1 & 2 & 2 & 3 \\ 2 & 2 & 3 & 3 & 0 & 1 & 2 & 1 & 2 & 1 \\ 3 & 0 & 0 & 1 & 4* & 2 & 5* & 1 & 1 & 0 \\ 0 & 4* & 2 & 0 & 1 & 3 & 2 & 3 & 2 & 0 \\ 1 & 2 & 3 & 1 & 3 & 2 & 0 & 1 & 2 & 2 \\ 2 & 1 & 2 & 2 & 2 & 2 & 3 & 1 & 0 & 2 \\ 1 & 1 & 1 & 3 & 2 & 0 & 0 & 5* & 4* & 0 \\ 1 & 1 & 1 & 3 & 2 & 0 & 0 & 5* & 4* & 0 \\ 1 & 0 & 3 & 3 & 1 & 3 & 2 & 0 & 3 & 1 \\ 1 & 3 & 2 & 0 & 1 & 2 & 0 & 1 & 1 & 6* \end{bmatrix}$$

$$\begin{bmatrix} 5* & 2 & 1 & 0 & 2 & 0 & 2 & 1 & 1 & 3 \\ 1 & 2 & 0 & 4* & 4* & 2 & 1 & 1 & 1 & 1 \\ 2 & 3 & 3 & 1 & 0 & 1 & 3 & 3 & 0 & 1 \\ 3 & 2 & 3 & 2 & 1 & 2 & 0 & 1 & 1 & 2 \\ 2 & 1 & 1 & 4* & 1 & 4* & 1 & 0 & 0 & 3 \\ 0 & 1 & 2 & 4* & 4* & 1 & 3 & 1 & 1 & 0 \\ 2 & 1 & 1 & 0 & 0 & 3 & 2 & 5* & 2 & 1 \\ 0 & 0 & 2 & 2 & 2 & 1 & 1 & 3 & 6* & 0 \\ 0 & 3 & 2 & 0 & 3 & 2 & 2 & 1 & 3 & 1 \\ 2 & 2 & 2 & 0 & 0 & 1 & 2 & 1 & 2 & 5* \end{bmatrix}$$

Decile Transition Matrices for 1-Yr. Formation Period, 2-Yr. Performance Horizon (in order of CAPM, 3-Factor, 4-Factor performance; *p<.05)

3	5*	0	0	0	2	0	0	0	6*
3	2	3	3	1	0	0	2	2	0
1	2	4*	3	2	2	0	2	0	0
1	1	1	1	5*	4*	0	2	0	1
1	0	2	1	2	3	5*	0	2	0
1	0	2	3	1	4*	3	0	2	0
3	1	1	1	0	1	5*	2	1	1
1	3	2	1	1	0	0	5*	2	1
1	2	0	2	4*	0	1	2	3	1
1	0	1	1	0	0	2	1	4*	6*

4*	1	1	0	1	2	1	1	2	3
2	5*	3	1	0	0	0	3	1	1
2	0	2	5*	1	4*	0	2	0	0
1	0	3	1	2	3	3	0	2	1
2	1	3	2	4*	2	0	2	0	0
1	4*	2	3	1	0	1	2	1	1
1	4*	0	0	1	1	3	3	1	2
0	0	2	1	3	2	2	2	3	1
2	1	0	2	1	0	4*	1	4*	1
1	0	0	1	2	2	2	0	2	6*

5*	2	0	1	1	1	2	1	3	0
1	2	4*	1	3	2	2	0	0	1
2	4*	0	2	3	1	2	1	1	0
3	2	1	4*	1	0	1	2	1	1
3	2	3	2	0	2	0	0	2	2
0	1	4*	2	2	2	2	2	1	0
0	1	1	0	2	5*	2	2	2	1
1	0	1	1	3	1	2	2	2	3
0	1	2	0	0	2	1	5*	3	2
1	1	0	3	1	0	2	1	1	6*

Decile Transition Matrices for 2-Yr. Formation Period, 1-Yr. Performance Horizon (in order of CAPM, 3-Factor, 4-Factor performance; *p<.05)

2	3	2	1	0	2	0	0	2	4*
5*	3	1	0	1	1	1	1	3	0
5*	0	3	1	4*	1	0	2	0	0
0	1	2	2	2	2	3	1	1	2
1	2	3	2	1	4*	1	1	1	0
2	1	2	1	3	0	4*	1	2	0
0	1	1	1	2	2	4	2	2	1
0	0	1	5*	1	3	1	2	2	1
0	3	0	3	2	0	1	5*	1	1
1	2	1	0	0	1	1	1	2	7*

1	4*	2	1	3	0	0	0	2	3
4*	1	3	2	1	1	2	1	1	0
2	2	3	2	1	1	2	1	1	1
3	2	2	0	1	2	1	0	5*	0
0	1	1	3	2	6*	1	1	1	0
3	3	1	1	1	1	2	2	1	1
0	1	0	5*	3	2	2	2	0	1
0	0	2	0	3	3	2	2	3	1
2	2	2	1	1	0	2	1	2	3
1	0	0	1	0	0	2	6*	0	6*

2	4*	1	1	1	2	1	1	0	3
4*	3	3	2	1	0	0	3	0	0
2	1	3	3	3	2	1	0	1	0
2	1	2	1	0	2	2	3	0	3
2	3	2	2	0	1	4*	2	0	0
1	2	2	2	2	1	1	2	1	2
1	1	1	2	5*	2	0	1	2	1
1	1	1	0	2	2	4*	3	2	0
0	0	1	2	2	2	2	1	4*	2
1	0	0	1	0	2	1	0	6*	5

Decile Transition Matrices for 2-Yr. Formation Period, 2-Yr. Performance Horizon (in order of CAPM, 3-Factor, 4-Factor performance; *p<.05)

$$\begin{bmatrix} 4\ast & 3 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 4\ast \\ 2 & 1 & 4\ast & 2 & 1 & 1 & 2 & 0 & 1 & 1 \\ 3 & 4\ast & 3 & 0 & 1 & 2 & 0 & 2 & 0 & 0 \\ 0 & 1 & 1 & 3 & 3 & 1 & 1 & 3 & 2 & 0 \\ 0 & 2 & 2 & 4\ast & 0 & 2 & 3 & 2 & 0 & 0 \\ 1 & 3 & 1 & 2 & 4\ast & 0 & 3 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 2 & 5\ast & 3 & 0 & 2 & 0 \\ 1 & 0 & 1 & 0 & 2 & 2 & 0 & 5\ast & 1 & 3 \\ 1 & 0 & 1 & 2 & 1 & 1 & 2 & 1 & 5\ast & 1 \\ 2 & 1 & 1 & 0 & 0 & 1 & 1 & 1 & 2 & 6\ast \end{bmatrix} \begin{bmatrix} 2 & 4\ast & 0 & 1 & 1 & 0 & 4 & 1 & 1 & 1 \\ 2 & 1 & 4\ast & 4\ast & 1 & 1 & 1 & 1 & 0 & 0 \\ 3 & 2 & 4\ast & 1 & 1 & 1 & 1 & 0 & 0 & 2 \\ 1 & 0 & 2 & 3 & 1 & 4\ast & 1 & 0 & 3 & 0 \\ 1 & 3 & 2 & 1 & 0 & 1 & 4\ast & 3 & 0 & 0 \\ 4\ast & 0 & 1 & 3 & 3 & 1 & 1 & 2 & 0 & 0 \\ 1 & 1 & 1 & 1 & 4\ast & 4\ast & 0 & 2 & 0 & 1 \\ 0 & 2 & 1 & 1 & 3 & 2 & 2 & 1 & 0 & 3 \\ 0 & 2 & 0 & 0 & 1 & 1 & 1 & 2 & 5\ast & 3 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 3 & 6\ast & 5\ast \end{bmatrix}$$

$$\begin{bmatrix} 5\ast & 2 & 3 & 0 & 0 & 1 & 0 & 2 & 1 & 1 \\ 2 & 4\ast & 2 & 3 & 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 3 & 3 & 2 & 1 & 2 & 2 & 0 & 1 & 0 \\ 1 & 1 & 2 & 3 & 1 & 2 & 1 & 1 & 1 & 2 \\ 2 & 3 & 0 & 2 & 3 & 2 & 1 & 1 & 1 & 0 \\ 1 & 0 & 3 & 0 & 5\ast & 1 & 1 & 2 & 1 & 1 \\ 1 & 1 & 1 & 2 & 3 & 3 & 0 & 2 & 1 & 1 \\ 1 & 0 & 1 & 1 & 1 & 1 & 8\ast & 0 & 0 & 2 \\ 0 & 0 & 0 & 2 & 0 & 2 & 1 & 3 & 7\ast & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 3 & 2 & 8\ast \end{bmatrix}$$

Decile Rolling Transition Matrices for 1-Yr. Formation Period, 1-Yr. Performance Horizon (in order of CAPM, 3-Factor, 4-Factor performance; *p<.05)

$$\begin{bmatrix} 6* & 0 & 1 & 0 & 0 & 0 & 3 & 4* & 1 & 2 \\ 0 & 4* & 5* & 1 & 2 & 2 & 1 & 0 & 1 & 1 \\ 2 & 1 & 2 & 2 & 1 & 2 & 3 & 3 & 2 & 0 \\ 0 & 4* & 2 & 4* & 2 & 3 & 1 & 0 & 0 & 1 \\ 0 & 2 & 0 & 2 & 3 & 3 & 1 & 3 & 3 & 0 \\ 1 & 0 & 2 & 1 & 5* & 2 & 2 & 2 & 1 & 1 \\ 1 & 1 & 4* & 2 & 1 & 2 & 2 & 0 & 1 & 3 \\ 2 & 2 & 1 & 4* & 0 & 1 & 1 & 2 & 2 & 2 \\ 2 & 2 & 0 & 1 & 1 & 0 & 2 & 2 & 4* & 3 \\ 3 & 1 & 1 & 0 & 2 & 2 & 1 & 1 & 2 & 4* \end{bmatrix} \begin{bmatrix} 4* & 3 & 1 & 1 & 0 & 0 & 0 & 1 & 2 & 2 \\ 1 & 1 & 4* & 2 & 3 & 2 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 2 & 2 & 1 & 2 & 2 & 0 & 2 \\ 1 & 2 & 0 & 2 & 3 & 1 & 2 & 2 & 0 & 1 \\ 2 & 0 & 4* & 1 & 1 & 1 & 1 & 0 & 2 & 2 \\ 1 & 2 & 2 & 2 & 0 & 3 & 3 & 0 & 1 & 0 \\ 0 & 1 & 0 & 2 & 2 & 4* & 1 & 2 & 1 & 1 \\ 0 & 1 & 0 & 2 & 1 & 0 & 2 & 4* & 4* & 0 \\ 2 & 1 & 1 & 0 & 0 & 1 & 3 & 2 & 1 & 3 \\ 2 & 2 & 1 & 0 & 2 & 1 & 0 & 1 & 2 & 3 \end{bmatrix}$$

$$\begin{bmatrix} 3 & 2 & 4* & 1 & 0 & 1 & 0 & 0 & 1 & 1 \\ 4* & 1 & 1 & 0 & 1 & 3 & 0 & 0 & 2 & 1 \\ 2 & 2 & 0 & 3 & 2 & 3 & 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 2 & 3 & 1 & 0 & 2 & 1 & 1 \\ 0 & 4* & 0 & 2 & 1 & 0 & 2 & 1 & 1 & 2 \\ 0 & 0 & 2 & 2 & 3 & 3 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 7* & 2 & 1 & 1 \\ 0 & 1 & 2 & 1 & 1 & 1 & 0 & 3 & 2 & 2 \\ 2 & 1 & 1 & 1 & 1 & 0 & 4* & 1 & 0 & 2 \\ 1 & 1 & 2 & 1 & 0 & 0 & 0 & 2 & 4* & 2 \end{bmatrix}$$

Decile Rolling Transition Matrices for 1-Yr. Formation Period, 2-Yr. Performance Horizon (in order of CAPM, 3-Factor, 4-Factor performance; *p<.05)

4*	1	0	0	2	2	0	3	2	2
0	3	2	1	1	2	2	0	1	2
2	2	1	1	2	3	1	2	2	0
1	0	1	4*	2	2	1	2	0	1
1	0	3	4*	2	3	1	0	1	1
1	1	2	1	3	1	3	2	2	0
1	4*	1	1	0	2	3	1	1	2
3	2	1	1	1	0	2	3	1	2
0	3	2	2	0	1	1	2	3	2
3	0	1	1	1	0	2	1	3	4*

3	4*	0	1	2	0	0	1	1	1
2	3	1	2	2	0	0	1	1	1
0	2	3	2	1	2	1	2	0	0
2	0	1	3	1	3	0	0	2	1
2	0	4*	0	1	2	2	0	1	1
2	2	0	2	3	1	2	1	0	0
0	1	0	1	2	2	3	1	2	1
0	0	1	1	1	2	2	1	3	2
0	1	2	0	0	0	2	6*	1	1
2	0	1	1	0	1	1	0	2	5*

3	3	1	0	1	0	2	1	0	1
5*	1	1	0	0	0	1	2	1	1
2	0	2	4*	0	1	1	1	1	0
1	3	2	0	0	2	1	1	1	1
1	1	2	3	0	1	1	0	1	2
0	1	0	3	4*	1	2	0	1	0
0	0	0	0	4*	2	1	4*	1	0
0	1	1	1	1	2	0	1	5*	0
0	0	3	1	1	1	2	1	1	2
0	2	0	0	1	2	1	1	0	5*

Decile Rolling Transition Matrices for 2-Yr. Formation Period, 1-Yr. Performance Horizon (in order of CAPM, 3-Factor, 4-Factor performance; *p<.05)

$$\begin{bmatrix} 3 & 1 & 1 & 1 & 1 & 1 & 2 & 1 & 3 & 2 \\ 2 & 3 & 4* & 1 & 1 & 1 & 0 & 1 & 1 & 2 \\ 3 & 2 & 2 & 2 & 4* & 0 & 0 & 1 & 2 & 0 \\ 1 & 2 & 2 & 1 & 2 & 2 & 2 & 2 & 2 & 0 \\ 1 & 1 & 3 & 4* & 1 & 0 & 3 & 1 & 1 & 1 \\ 1 & 1 & 0 & 3 & 3 & 2 & 3 & 3 & 0 & 0 \\ 1 & 2 & 3 & 1 & 2 & 2 & 1 & 1 & 2 & 1 \\ 0 & 1 & 0 & 2 & 1 & 7* & 2 & 2 & 1 & 1 \\ 2 & 2 & 0 & 1 & 2 & 1 & 2 & 3 & 2 & 1 \\ 2 & 1 & 1 & 0 & 0 & 0 & 1 & 1 & 2 & 8* \end{bmatrix} \begin{bmatrix} 2 & 2 & 1 & 5* & 1 & 1 & 0 & 0 & 0 & 1 \\ 2 & 2 & 0 & 1 & 2 & 2 & 1 & 0 & 2 & 1 \\ 3 & 2 & 1 & 1 & 2 & 2 & 1 & 1 & 0 & 0 \\ 0 & 2 & 3 & 2 & 2 & 1 & 1 & 1 & 0 & 1 \\ 1 & 0 & 3 & 0 & 3 & 2 & 2 & 2 & 0 & 0 \\ 1 & 1 & 3 & 2 & 0 & 3 & 1 & 0 & 1 & 1 \\ 2 & 0 & 0 & 0 & 2 & 0 & 3 & 4* & 2 & 0 \\ 0 & 0 & 1 & 1 & 1 & 2 & 1 & 3 & 3 & 1 \\ 0 & 4* & 1 & 1 & 0 & 0 & 2 & 0 & 5* & 0 \\ 2 & 0 & 0 & 0 & 0 & 0 & 1 & 2 & 0 & 8* \end{bmatrix}$$

$$\begin{bmatrix} 1 & 2 & 2 & 2 & 0 & 0 & 2 & 0 & 0 & 3 \\ 3 & 2 & 1 & 1 & 3 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 2 & 2 & 0 & 2 & 3 & 0 & 1 & 0 \\ 2 & 1 & 0 & 1 & 2 & 2 & 0 & 1 & 3 & 0 \\ 2 & 1 & 1 & 1 & 2 & 2 & 2 & 0 & 0 & 1 \\ 0 & 1 & 1 & 2 & 5* & 0 & 0 & 1 & 0 & 2 \\ 2 & 1 & 2 & 0 & 0 & 3 & 2 & 1 & 1 & 0 \\ 1 & 1 & 2 & 0 & 0 & 1 & 0 & 5* & 2 & 0 \\ 0 & 0 & 0 & 3 & 0 & 1 & 2 & 1 & 4* & 1 \\ 0 & 2 & 1 & 0 & 0 & 0 & 1 & 2 & 1 & 5* \end{bmatrix}$$

Decile Rolling Transition Matrices for 2-Yr. Formation Period, 2-Yr. Performance Horizon (in order of CAPM, 3-Factor, 4-Factor performance; *p<.05)

$$\begin{bmatrix} 2 & 3 & 2 & 0 & 0 & 2 & 0 & 3 & 1 & 2 \\ 3 & 2 & 2 & 2 & 1 & 1 & 0 & 2 & 1 & 1 \\ 1 & 4\ast & 1 & 1 & 2 & 1 & 1 & 3 & 1 & 0 \\ 1 & 1 & 3 & 2 & 4\ast & 0 & 3 & 1 & 0 & 0 \\ 1 & 0 & 3 & 0 & 3 & 2 & 1 & 1 & 3 & 1 \\ 2 & 1 & 1 & 3 & 1 & 4\ast & 2 & 1 & 0 & 0 \\ 2 & 2 & 1 & 3 & 0 & 3 & 1 & 0 & 2 & 0 \\ 0 & 1 & 0 & 2 & 1 & 1 & 3 & 3 & 2 & 2 \\ 0 & 1 & 2 & 2 & 1 & 1 & 2 & 1 & 3 & 1 \\ 3 & 0 & 0 & 0 & 1 & 0 & 2 & 0 & 2 & 7\ast \end{bmatrix} \begin{bmatrix} 4\ast & 3 & 1 & 1 & 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 3 & 2 & 0 & 2 & 1 & 1 & 2 & 0 \\ 3 & 2 & 1 & 2 & 2 & 0 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 2 & 1 & 1 & 3 & 0 & 1 & 1 \\ 0 & 1 & 3 & 3 & 0 & 2 & 2 & 1 & 0 & 0 \\ 1 & 2 & 1 & 0 & 3 & 0 & 2 & 2 & 0 & 1 \\ 0 & 2 & 0 & 2 & 4\ast & 3 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 & 2 & 1 & 2 & 4\ast \\ 0 & 0 & 1 & 0 & 1 & 2 & 1 & 2 & 5\ast & 0 \\ 2 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & 2 & 6\ast \end{bmatrix}$$

$$\begin{bmatrix} 5\ast & 0 & 3\ast & 0 & 0 & 0 & 1 & 0 & 1 & 1 \\ 4\ast & 0 & 1 & 4\ast & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 3\ast & 3\ast & 1 & 0 & 2 & 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 3\ast & 1 & 3\ast & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 3\ast & 0 & 0 & 3\ast & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 6\ast & 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 4\ast & 1 & 1 & 1 & 1 & 1 & 0 \\ 0 & 2 & 3\ast & 0 & 0 & 0 & 1 & 2 & 2 & 1 \\ 0 & 0 & 0 & 0 & 2 & 1 & 1 & 2 & 1 & 4\ast \\ 0 & 1 & 0 & 1 & 0 & 0 & 2 & 1 & 2 & 4\ast \end{bmatrix}$$