

Asset Allocation and Bad Habits

Professor Andrew Ang
Columbia Business School
Email: aa610@columbia.edu
Tel: +1 212 854 9154

Professor Amit Goyal
Université de Lausanne
Email: amit.goyal@unil.ch
Tel: +41 21 692 3676

Antti Ilmanen, Ph.D.
AQR Capital Management
Email: antti.ilmanen@aqr.com
Tel: +44 7887 475 184

April 2014

Preliminary – comments welcome

Abstract

This article concentrates on documenting investors' 'bad habits' in asset allocation practices. The prior is that one of the biggest vices of long-term investors is procyclical investing at multi-year horizons — ill-timed flows into and out of good investments hurt investor's performance. Financial markets exhibit momentum over multi-month horizons but more mean-reversion over multi-year horizons. However, many investors act like momentum investors even at these longer horizons. These patterns are anecdotally well-known but not well statistically documented, especially together. Thus, we address the following two empirical questions:

1. How do funds reallocate based on past returns? We provide direct evidence using the CEM Benchmarking data on pension fund target allocations over a 22-year period.
2. What are momentum/reversal patterns in returns? We provide evidence using more than a century of financial market data.

The twin data are qualitatively consistent with the priors that investors' chasing returns over multi-year horizons is likely to hurt their long-run performance. However, statistical evidence on procyclical multi-year asset allocations or on multi-year mean reversion patterns in asset class returns is only borderline significant.

1. Introduction

Pension funds are one of the most important participants in global financial markets. After many years of relative neglect, the academic literature has started focusing attention on the characteristics of this part of the finance industry. A simplified process of investment management at pension funds involves two broad steps. The first is drafting an investment policy statement / defining asset allocation. The second is selecting investment managers in specified mandates. Blake et al. (2012) document a secular increase in this second aspect of decentralized investment management and Goyal and Wahal (2008) study the efficacy of the process of selecting external investment managers. There is an active debate in the literature about the relative contributions of these two steps to the excess returns earned by pension funds. While Brinson, Hood, and Beebower (1986), Brinson, Singer, and Beebower (1991), and Blake, Lehmann, and Timmermann (1999) argue that asset allocation is important for explaining the time-series variation in returns, Ibbotson and Kaplan (2000) and Kritzman and Page (2002, 2003) show that security selection is more important in explaining the cross-sectional variation in returns earned by pension funds. Andonov, Bauer, and Cremers (2012, henceforth ABC) find that pension funds' active returns are roughly equally driven by asset allocation, market timing and security selection. There are, thus, many papers that study return attribution. At the same time, there are almost no papers that directly study the impact of past returns on asset allocation. Our paper attempts to fill this gap, focusing on the dependence of asset allocations on past performance.¹

Many pension funds rebalance their asset class allocations regularly to some target weights, such as the conventional 60% stocks and 40% bonds. Yet, there is anecdotal evidence that funds may let their allocations drift with relative asset class performance. This might just reflect passive buy-and-hold policies, or a desire to maintain asset class allocations near market-cap weights, or more proactive return chasing. We focus on the last possibility.

Both retail and institutional investors are anecdotally known to chase returns – buy into recent and longer term winners – whether asset classes or managers. And many lack patience when facing a few years of underperformance even if they are aware of the limited predictive ability in past performance and the high transition costs. Ang and Kjaer (2011) argue that one of the biggest vices of long-term investors is procyclical investing at multi-year horizons — ill-timed flows into and out of good investments can make the investor's performance poor.² Several studies suggest that the “bad habit” of multi-year return chasing damages investors' long-run wealth. For example, the fact that investors earn lower dollar-weighted returns than time-weighted returns is indicative of ill-timed investor flows (Dichev (2007) and Friesen and Sapp (2007)). Institutional investors' fire-

¹ There are studies of other determinants of asset allocation. For instance, Dyck and Pomorski (2011) document positive scale economies in asset management. This is, however, disputed by ABC (2012) who argue that larger pension funds would be better off by investing more in passive mandates. The role of size in alpha generation is thus, unclear. Even if larger funds generate better performance, it is not obvious that their experience can be replicated (Lerner, Schoar, and Wang (2008)). Regulations play another important role in asset allocation. Rauh (2006) and Addoum, van Binsbergen, and Brandt (2010) find that corporate pension fund investment policy is dictated by mandatory funding rules. It is also important to note that while asset allocation is based in principle on expected returns, in practice, past returns guide expectations about the future.

² The other issue is the misalignment of incentives between managers and investors (Lakonishok, Shleifer, and Vishny (1992)).

and-hire decisions at the manager level have also been shown to lose value over time (Goyal and Wahal (2008)).

There might be nothing inherently bad in return chasing – financial markets tend to exhibit momentum over multi-month horizons (Jegadeesh and Titman (1993), Moskowitz, Ooi, and Pedersen (2012), Asness, Moskowitz, and Pedersen (2013)). Such return persistence makes trend-chasing profitable, if not overwhelmed by trading costs. However, when it comes to multi-year returns, financial markets are more likely to exhibit mean-reversion than continuation tendency (DeBondt and Thaler (1985), Poterba and Summers (1988), Cutler, DeLong, and Summers (1991), Ilmanen (2011), and Zakamulin (2013)). Yet, too many investors act like momentum investors even at these longer horizons. Institutional investors might be especially procyclical at three- to five-year horizons, reflecting their typical performance evaluation periods. Conventionality pressures may exacerbate procyclical flows as some investors chase their peer allocations and not just returns.

A simple numerical example can help illustrate our basic idea. Imagine that there are three years: Year 0, Year 1, and Year 2. There are two asset classes of stock and bond. The returns on stocks are 10%, 10%, and 5% for the three years while the returns on bonds are 5%, 5%, and 10%. These patterns are supposed to mimic momentum in the first year and reversal in the second year. Further assume that the weight in an asset class is directly proportional to the returns earned in that asset class. This means that at the end of year 0, the weights in stocks and bonds are 67% and 33%, respectively (in proportion to returns of 10% and 5% observed at the end of year 0). There is no change in weights at the end of Year 1. Given these weights, the return of the fund is $67\% \times 5\% + 33\% \times 10\% = 6.7\%$. If the fund knew about the reversal that was going to happen in Year 2, the fund would have revised its asset allocation to 33% stocks and 67% bonds and earned $33\% \times 5\% + 67\% \times 10\% = 8.4\%$. Thus, the loss from return chasing is 1.7%. While this example is stylized and relies on unrealistic assumptions such as perfect foresight, the general message is that return chasing at horizons of beyond one year has a deleterious effect.

Despite well-known anecdotal evidence we lack long enough histories for making statistically conclusive statements of multi-year investor behavior. Some types of procyclical flows are especially hard to discern. Consider changes made in investor benchmarks where new asset classes tend to get added after multi-year rallies. Because decisions to change the benchmark are rarely evaluated over time, such procyclicality can be missed, however prevalent it is.

Earlier studies have shown only indirect evidence on multi-year return chasing in institutional asset allocations – and on consequent long-run losses. We use annual data from CEM Benchmarking on evolving U.S. pension funds' asset allocations 1990-2011 to provide direct evidence on return-chasing behavior at asset class level over multi-year horizons. We also document evidence on financial markets' multi-month momentum and multi-year reversal patterns. By contrasting evidence of multi-year procyclical institutional allocations with findings of multi-year return reversals in many financial assets, we hope to make at least some investors remedy their bad habits.

Our key findings are easily summarized. Pension funds in the aggregate do not recognize the shift from momentum to reversal tendencies in asset returns beyond one-year horizon. Instead, a typical

pension fund keeps chasing returns over multi-year horizons, to the detriment of the institutions long-run wealth. However, the statistical significance is weak.

The studies closest to ours in terms of scope are those by Heisler et al. (2007) and Stewart et al. (2009). These authors show that investment products that receive contributions later underperform products that experience withdrawals over one to five-years. Post-flow underperformance is more due to product (manager) selection than category (asset class) reallocation, but both activities detract value. Their data are based on PSN database of institutional products, not on actual institutional allocation data.

The studies closest to ours in terms of data requirements are those by Dyck and Pomorski (2011) and ABC (2012). These authors use CEM Benchmarking data to study North American institutional plan asset class allocations and returns over time. The focus of these studies is to study scale economies in asset management. The latter study also decomposes the return into its components of asset allocation, market timing, and security selection. However, neither of these studies looks at multi-year return-chasing or relates their findings to medium-term reversal patterns in asset returns.

The rest of the paper proceeds as follows. Section 2 describes the data, including limitations, and offers preliminary evidence on pension funds' asset class allocations over time. Section 3 provides both time-series and cross-sectional evidence on asset allocation and its dependence on past returns at various horizons. Section 4 documents empirical patterns in asset class returns (strong multi-month momentum patterns followed by weaker reversal tendencies). Section 5 interprets the results, contrasting the differential response of institutional asset allocations and asset class returns to return shocks, before offering concluding remarks.

2. Data Description

CEM Benchmarking Incorporated collects pension fund data through yearly questionnaires, with the broadest coverage in North America. Pension funds participate in these surveys mainly to gain information about their peer comparisons in terms of costs, asset allocations, and performance. Voluntary reporting raises the question of selection bias, but earlier studies find no evidence of such bias in fund returns (Bauer, Cremers, and Frehen (2010) and ABC (2012)). We have data on 978 different pension funds over a time period of 1990 to 2011 (22 years). In this study, we focus on 573 U.S. pension funds, but preliminary analysis suggests similar results for the global universe. The funds have varied sizes, with median near \$3 billion and average near \$10 billion. The participating funds hold 30-40% of the assets under management by U.S. pension funds and hold about 4% of U.S. equity market capitalization (ABC (2012)).

The basic information in CEM Benchmarking surveys contains, amongst other variables, the fund portfolio's *actual* and *policy* weights as well as realized returns. The actual weights are the average weight in an asset class during the year while the policy weights are the weights in an asset class dictated by strategic policy portfolio considerations. These policy weights are available on a calendar-year-end basis and refer to the target weights for the next calendar year.

The weights and returns are available at a quite detailed level (for example, weights in internal passive large-cap U.S. equities). For our purposes, we do not need information at such refined level. We, therefore, first aggregate the portfolio weight data to the level of nine broad asset classes. For returns, each pension fund can choose its own asset-class benchmarks. However, we choose to standardize the benchmarks and use the common benchmarks listed in parenthesis next to the asset class in the following list.

- i. Domestic equity (CRSP value-weighted)
- ii. International equity (MSCI AC World ex-U.S.)
- iii. Domestic fixed income (Barclays U.S. aggregate)
- iv. International fixed income (Barclays Global aggregate ex-U.S., currency-hedged)
- v. Real estate (NCREIF)
- vi. Private equity (Cambridge Associates)
- vii. Hedge funds (HFRI)
- viii. Commodities (S&P GSCI commodities)
- ix. Cash (U.S. T-bill)

It is possible to consider different splits of the data. For instance, we can consider separating international equity and fixed income further into their developed markets and emerging markets components. This is conceivably interesting as the last few years have seen a large movement from institutional investors into emerging markets. However, lack of long-term reliable data on these sub-classes (needed in Section 4) precludes us from attempting this finer classification.

Our chosen benchmarks best reflect the asset class returns for the corresponding asset class. At the same time, these choices must remain, to some extent, ad-hoc and, therefore, debatable. For instance, NAREIT may arguably be a better benchmark for some investors who do not invest in private real estate market. We leave explorations of these fine distinctions to future research.

It is important to understand a few limitations of the data. The main one is that while the cross-sectional coverage of the data is good, the data are available only at annual frequency. This means that we do not have enough power in many of the time-series tests described below. Another limitation is that although we have actual weights in different asset classes, these are average weights during the year and not year-end weights. This is one reason to favor reported policy weights, which are year-end weights and thus better matched with realized returns over a calendar year.

We could conceivably do the analysis for fewer than nine asset classes. For instance, it is possible to aggregate real estate, private equity, commodities, and hedge funds into an alternate asset class. However, the characteristics of these asset classes, as well as institutional allocation patterns to them, are so disparate that we deemed it worthwhile to not do so. In our data, across all fund-years, 89% of the policy weight allocations are in four asset classes, domestic/international equity and domestic/international fixed income, and the data quality is best for them. We will, therefore, conduct our main analysis only for these asset classes. It is possible to collapse these asset classes into two: stocks and bonds. However, the variation between domestic and international allocations is interesting, so we keep these separate.

Recall that we also have access to funds' self-reported benchmark returns in each asset class as well as fund reported overall policy return. In the majority of cases, the fund reported benchmark returns are very close to our selected benchmarks' returns. However, in some cases, and especially for a few asset classes such as real estate and cash, there are noticeable differences. It is not clear whether this is due to data errors or due to the specific nature of investments by a fund that necessitates a specific benchmark. We opt for our generic benchmarks in the analysis below but note that this does not cause many differences in our main asset classes of equity and fixed income.

Figure 1 shows the time-series of cross-sectional averages of actual weights and policy weights to these nine asset classes. A somewhat striking feature of the data is that actual weights and policy weights move so closely together. We expect some mechanical relation between realized returns and actual weights (unless funds practice frequent rebalancing to fixed asset class weights). To see this, denote funds by $i = 1, \dots, N$, time by $t = 1, \dots, T$, and asset class by $a = 1, \dots, A$. Let year-end weights be $w_{i,a,t}$, and returns are $R_{i,a,t}$. Denote total fund returns by $R_{i,t}$. Construct what the weights would have been due to just return realization and call these as passive weights, $w_{i,a,t}^{\text{passive}}$. It is easily shown that:

$$w_{i,a,t}^{\text{passive}} = w_{i,a,t-1} \times \frac{1 + R_{i,a,t}}{1 + R_{i,t}}.$$

The active weights $w_{i,a,t}^{\text{active}}$ are then given by:

$$w_{i,a,t}^{\text{active}} = w_{i,a,t} - w_{i,a,t}^{\text{passive}}.$$

Thus, if the funds do not consciously rebalance their portfolio, the year-end weights will inherit the relative asset class returns realized during the year. No such mechanical relation is expected ex-ante for policy weights. It will be, therefore, more useful to conduct the analysis for policy weights. There are two other advantages in using policy weights. First, policy weights are strategic targets deliberately chosen by pension fund boards, so they can be more unambiguously interpreted as active decisions. Thus, studying policy weights' dependence on past market returns corresponds more closely to our objective of analyzing the bad habits of institutional return chasing. Second, as mentioned before, we do not have year-end weights but rather average weights for realized (actual) holdings. The date mismatch with calendar-year returns further complicates the analysis and interpretation of results using actual weights.

Figure 1 shows that policy weights (strategic target asset class allocations) are 57% for equities, 32% for fixed income, 9% for alternatives and 2% for cash when averaged across all funds (equally-weighted) over the 1990-2011 period. Panel A shows that policy weights for equities rose from 54% to 61% peak in 1999-2001 before falling to 46% in 2011. Fixed-income weights fell from a third to 29% in 2004-2006 before rising to 35% in 2011, and cash weights had a similar U-shaped time profile. Alternatives weights fell from 10% to 6% in late 1990s before rising to 16% in 2011. Large policy allocation changes during the sample period reveals that characterizing pension funds as rebalancing to fixed 60/40 stock/bond weights is a gross simplification of industry practices.

Actual weights exhibit similar time paths but pension funds consistently underweighted equities and fixed income while overweighting alternatives and cash. Most deviations were modest; the largest deviation was alternatives 7-9% overweight in 2008-2011.

Drilling into equity and fixed-income assets in Panel B of Figure 1, it is worth highlighting that the broad equity allocation conceals a dramatic reduction in home bias. Policy weights on international equities rose pretty consistently from 9% in 1990 to 21% in 2011, while the share of US equities hovered in 40% until 2005 before falling sharply to 25% in 2011. This development indicates a significant decline in home bias. (Fixed-income assets remained significantly (~90%) home-biased, in line with institutional liabilities being mainly dollar-denominated.)

Drilling into alternatives in Panel C of Figure 1 reveals gradual shifts over time as private equity has overtaken real estate as the most popular alternative asset (policy weights 5.3% and 4.7% in 2011), while hedge fund and commodity weights have risen from close to zero to 4.2% and 1.6%, respectively. Curiously, pension funds' actual allocations on private equity and real estate have consistently exceeded the target allocations (averaging about 2% each). The private equity overweight grew to over 5% in 2011 (actual allocation exceeding 10%), perhaps not an altogether deliberate choice by pension funds but a result of commitments made during the earlier "good years."

3. How Do Funds Reallocate Based On Past Returns?

Time-series regressions

The key question here is how the average fund changes its policy weights based on past returns. Therefore, we regress average industry policy weights in an asset class on past returns. In other

words, we define average weights as $\bar{w}_{a,t}^{\text{policy}} = \frac{1}{N} \sum_{i=1}^N w_{i,a,t}^{\text{policy}}$ and run variants of the following time-series regression for each asset class:

$$\bar{w}_{a,t}^{\text{policy}} = \alpha_a + \sum_{j=0}^3 \beta_j R_{\text{bench}}_{a,t-j} + e_{a,t}. \quad (1)$$

As mentioned previously, we use common benchmark returns as explanatory variables. We estimate equation (1) separately for each asset class, as well as for pooled versions of asset classes. The pooled regression is run with asset-specific intercepts, α_a . To make best use of our limited data, we focus mostly on the pooled specification that includes only equity and fixed income. Note that, in addition to lagged returns, we want to use returns at time t as we have noticed that policy weights move together with actual weights. Thus, funds seem to be shifting policy weights in response to past returns, apart from any mechanical effect of returns.

Panel A of Table 1 reports the results of regression (1) where Newey-West adjusted t -statistics are in parenthesis below the coefficients. Since there are only 22 years of data, we also use the following specification in Panel B with only two independent variables that still capture the essence of the relationship between weights and past returns.

$$\bar{w}_{a,t}^{\text{policy}} = \alpha_a + \beta_0 R_{\text{bench}}_{a,t} + \beta_1 R_{\text{bench}}_{a,t-1:t-3} + e_{a,t}. \quad (2)$$

We also remedy the problem of limited time-series mildly by padding benchmark returns for three years before 1990. In other words, we use the average weights from 1990-2011 and benchmark returns (whenever available) from 1987-2011 to predict these weights. Increase in the number of observations should increase the power of our tests and lead to higher statistical significance, if it exists.

Table 1 shows that weights are, in general, positively related to contemporaneous returns as well returns over the last three years for each of the main asset class as well as for the pooled specification. The results on individual asset classes are weaker; the adj- R^2 is even negative for some asset classes such as international equity and international fixed income. Part of the reason for relatively poor fit for individual asset classes is that a regression to explain slow-moving asset class weights with volatile returns will always have low power. A more important reason is that we need to look at re-allocation across asset classes rather than just allocation within an asset class.

These considerations lead to focusing more on pooled results which also use our limited data more effectively. For the pooled four-asset case, the coefficients on $\text{Return}(t)$ and $\text{Return}(t-1:t-3)$ are 0.044 (t -statistic = 1.43) and 0.175 (t -statistic = 1.69), respectively. However, the statistical significance is low (t -statistics rarely exceed two and the small number of observations can even lead to negative adj- R^2 s for some single asset classes).³ Since one standard deviation of annualized returns corresponds to roughly 10% (it is around 20% for equity and around 5% for fixed income), the economic magnitude of weights in response to one standard deviation move in each-year past returns is around 0.6%.

One issue that can be debated is whether to use levels or changes of policy weights as the dependent variable. In unreported results, we find that the adj- R^2 goes up for each of the first four asset classes but decreases for the pooled version. Correspondingly, the statistical significance of coefficients goes up in regressions for individual asset classes but decreases in the pooled regression.

We explored many other specifications and came up with broadly similar results – positive coefficients on past returns but limited statistical significance. As an example of a robustness check, first notice that, since the sum of weights has to be one, the left-hand side of equation (1) is constrained while the right-hand side returns are not. We change the specification slightly by using relative benchmark returns. We calculate average policy returns across the funds to get an weighted-average benchmark return, $\bar{Rb}_t = \frac{1}{N} \sum_{i=1}^N R_{i,t}^{\text{policy}}$. We then run the following regression:

$$\bar{w}_{a,t}^{\text{policy}} = \alpha_a + \sum_{j=0}^3 \beta_j (R_{\text{bench}}_{a,t-j} - \bar{Rb}_{t-j}) + e_{a,t}. \quad (3)$$

We can run this regression only over the CEMB sample of 1990-2011. The results of this specification are available upon request. The broad evidence is similar (pooled specification for levels gives stronger results but for changes in weights the results are weaker).

³ The adj- R^2 for the pooled specification is high due to asset-specific intercepts (which allow the regression to explain a large part of cross-asset variation in weights).

Overall, there is evidence that asset class policy weights are positively related to past returns of even three years ago. The evidence seems economically strong. However, the statistical significance is low, most likely due to the limited length of the data series.

Panel regressions

One can perhaps gain more power by using the entire cross-section of funds. We run the following panel regression:

$$w_{i,a,t}^{\text{policy}} = \delta_{1,i} + \delta_{2,a} + \delta_{3,t} + \sum_{j=0}^3 \gamma_j R_{i,a,t-j} + u_{i,a,t}, \quad (4)$$

where δ_1 , δ_2 , and δ_3 are dummies for funds, asset class, and time, respectively. Thus, the heterogeneity in the sample due to different funds, asset classes, and years is partly accounted for by the fixed effects. The inference is then done on γ coefficients. Note that the explanatory variables are not the benchmark returns but actual returns realized in asset class a by pension plan i . The standard errors are double-clustered for fund- and time-effects following Pedersen (2009). We run regression (3) separately for domestic equity, for pooled panel of four asset classes, as well as all asset classes.

The results are reported in Table 2 (coefficients on dummies are not reported). Coefficients on returns are positive but the statistical significance varies – many t -statistics exceed two, others fall below it.⁴ In our preferred specification for four asset classes, the coefficients on $\text{Return}(t)$ and $\text{Return}(t-1)$ are positive although statistically significant. This shows that pension funds increase policy weights in response to recently realized returns. At the same, we find that the coefficients on $\text{Return}(t-2)$ and $\text{Return}(t-3)$ are positive and statistically significant. This means that pension funds weight allocations react positively even to returns realized a few years ago.

We also run panel regressions for sub-samples of corporate and public pension funds, and small and large pension funds. There is some evidence that corporate funds engage in more return chasing than do public funds. However, the differences across different kinds of pension funds are not statistically significant.

4. What Are Momentum/Reversal Patterns In Asset Class Returns?

Turning to the analysis of momentum and reversal patterns in asset class returns, we use return data dating back to 1900, when available. We study multi-country composite of non-U.S. equities (equally weighted with a gradually growing country universe), U.S. equities, MSCI EAFE equity index, multi-country composite of non-U.S. government bonds (equally weighted with a gradually growing country universe), Barclays Aggregate U.S. bond index, and S&P GSCI commodity index. All data are provided by AQR Capital.

⁴ We also run Fama and MacBeth (1973) regressions as an alternative to panel regressions. The coefficients on contemporaneous returns and their first three lags are 0.141, 0.114, 0.035, and 0.080 for the sample of first four asset classes with asset-specific intercepts. All coefficients are positive, consistent with multi-year return chasing, similar to Table 2 but with larger coefficients for the first years.

The reasons for using such long data histories are twofold. First, any attempt to capture statistical patterns of multi-year autocorrelation patterns requires decades of data to have even a double-digit number of independent observations. Shorter windows would likely lead to spurious relations, especially if multiple coefficients are estimated, given the relatively weak momentum/reversal patterns in financial markets. Second, autocorrelation patterns over the 1990-2011 period would not only suffer from a short sample period but also from a look-ahead bias. Investors know the in-sample autocorrelation patterns only with hindsight – their real-time expectations on asset return persistence (which influenced their asset allocations decisions through the past twenty years) were likely informed by historical patterns observed during the previous decades. We alleviate (but do not eliminate) the look-ahead bias by using data starting much before 1990. At the same time, there is a concern of selection bias from using such long histories of data, for example due to exclusion of markets such as Russia or China from early parts of the sample. The implications of such exclusions on momentum and reversal patterns for broad asset classes are not immediately obvious and we hope that this does not lead to any systematic biases in our analysis below.

Figure 2 shows for each asset the average autocorrelation of lagged monthly returns in the previous 12 months (Y1), and in months lagged by 13-24 (Y2), 25-36 (Y3), 37-48 (Y4), 49-60 (Y5). As already recognized twenty years ago (e.g., Cutler, DeLong, and Summers (1991)), many financial assets exhibit momentum or continuation tendencies over multi-month horizons (up to a year) and somewhat less consistent multi-year reversal patterns. The average autocorrelations are positive for the most recent year and mainly negative or near zero for the preceding years.

Since multi-month momentum is the stronger empirical effect, it is possible that too fast portfolio rebalancing (which will be hurt by momentum) is as detrimental a habit to institutional investors as is too persistent return-chasing (which will be hurt by long-term mean reversion). However, our annual asset allocation data is too coarse to study within-year rebalancing. It is worth noting that when we have only access to annual return data, the distinction between one-year momentum and multi-year reversals loses some sharpness as the within-first-year momentum is partly concealed and partly offsets second-year reversal tendencies.

To capture return dependencies over multiple periods, we need partial autocorrelations or multiple regression coefficients instead of these simple autocorrelations. We provide such evidence in the next section.

5. Interpretations and Concluding Remarks

Finally, we contrast the results in Sections 3 and 4, and create a stylized impulse-response graph which tracks the cumulative impact of a return shock at time (year) t on the policy weights and future returns of the same asset class through the next few years.

Recall from Section 3 that high returns at time t lead pension funds to allocate more weight to the asset class at t but apparently also later, say at time $t+3$. In contrast, observed return autocorrelations in Section 4 suggest that high returns at t lead to high returns at $t+1$ but to low returns from $t+2$ to $t+4$. The multiple regression on asset returns that we want to run is:

$$R_t = \mu + \rho_1 R_{t-1} + \rho_2 R_{t-2} + \rho_3 R_{t-3} + \rho_4 R_{t-4} + u_t. \quad (5)$$

We expect coefficient ρ_1 to be positive but coefficients ρ_2 , ρ_3 , and ρ_4 to be low or negative. Table 3 shows the results for a pooled regression of the annual returns of four major asset classes between 1900-2011: U.S. equities, non-U.S. equities, U.S. Treasuries, non-U.S. government bonds (data starting in 1933). We will use the coefficients on the four lagged returns in the next section (impulse-response function): **0.034**, **-0.154**, **0.076**, **-0.060**.⁵ Note that the first coefficient on previous-year return likely reflects a mixture of positive autocorrelations over a few months and negative autocorrelations beyond them. The positive coefficient on the third year is surprising but statistically insignificant. Indeed, only the negative coefficient on the second year is statistically significant.⁶

Smart funds recognizing these autocorrelation patterns should optimally increase policy weights at time t (to benefit from anticipated high returns at $t+1$) and optimally reduce weights at $t+1$ to $t+3$ (to benefit from anticipated low returns from $t+2$ to $t+4$). Admittedly, strictly reading Table 3 suggests that actual returns are positive at $t+3$ and negative at $t+2$ and $t+4$, but it may well be reasonable to average/smooth these results and view all return responses ($t+2$, $t+3$, $t+4$) as mildly negative. The regression test that is suggested for policy weights is then the following:

$$w_t = c + \beta_0 R_t + \beta_1 R_{t-1} + \beta_2 R_{t-2} + \beta_3 R_{t-3} + e_t. \quad (6)$$

A smart fund would see positive β_0 and negative β_1 to β_3 . Results in Tables 1 and 2 indicate, however, that all of these coefficients are positive.

Given coefficients ρ and β , we can plot the impulse response to, say, a 10% shock in returns at time t . For returns, we use the regression coefficients from the pooled equity-and-bond momentum/reversal analysis in Table 3, while for policy weights, we use the coefficients from the cross-sectional Fama-MacBeth regression for the pooled equity-and-bond asset allocation analysis in Table 2. Our impulse-response differs from a standard function as we use coefficients from two equations (5) and (6) estimated separately. Thus, the dependence between the two equations comes from the effect of common independent variables. In particular, we give a shock of 10% to Return(t) and let this shock permeate through equations (5) and (6) independently.

The impulse-response is shown in Figure 3. The arrows highlight a time shift needed while reading this figure. The fact that policy weights go up at time t is good as high weights benefit from high momentum returns at time $t+1$. However, the fact that the policy weights are high at time $t+1$ is bad as positions suffer from return reversals at time $t+2$, and so on.

Figure 3 summarizes our key finding of the bad habit of excessively procyclical pension fund asset allocations. Pension funds in the aggregate do not recognize the shift from momentum to reversal tendencies in asset returns beyond one-year horizon. Instead, a typical pension fund keeps chasing

⁵ The four coefficients are 0.089, -0.208, 0.156, and -0.055 if we use the sample period 1990-1989 and, thus, only the information available for investors before our asset allocation data begins. These coefficients are broadly similar to those for the full sample in Table 3.

⁶ The weakness of statistical evidence corresponds to weaker mean-reversion in multi-year market returns (time-series relation) than in multi-year security selection (cross-sectional relation).

returns over multi-year horizons, to the detriment of the institutions long-run wealth. We hope this evidence will help at least some pension funds to reconsider their asset allocation practices.

References

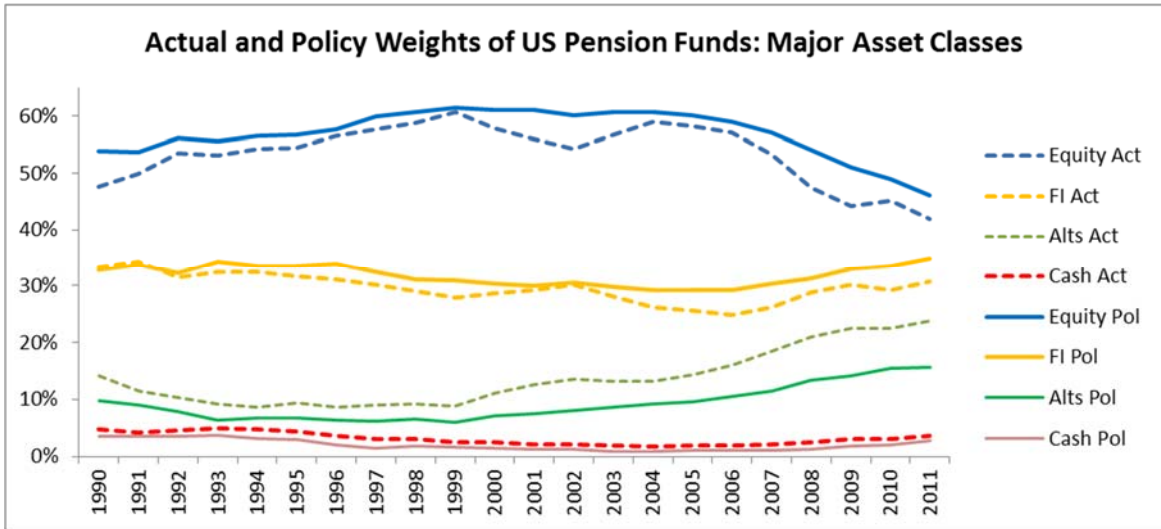
- Addoum, Jawad M., Jules H. van Binsbergen, and Michael W. Brandt, 2010, “Asset Allocation and Managerial Assumptions in Corporate Pension Plans,” Working paper.
- Andonov, Aleksandar, Rob Bauer, and Martijn Cremers, 2012, “Can Large Pension Funds Beat the Market? Asset Allocation, Market Timing, Security Selection, and the Limits of Liquidity,” Working paper.
- Ang, Andrew, and Knut N. Kjaer, 2011, “Investing for the Long Run,” in Tomas Franzen, ed., *A Decade of Challenges: A Collection of Essays on Pensions and Investments*, 94–111.
- Asness, Clifford S., Tobias J. Moskowitz, and Lasse H. Pedersen, 2013, “Value and Momentum Everywhere,” *Journal of Finance* 68, 929–985.
- Bauer, Rob, Martijn Cremers, and Rik Frehen, 2010, “Pension Fund Performance and Costs: Small Is Beautiful,” Working paper.
- Blake, David, Bruce N. Lehmann, and Alan Timmermann, 1999, “Asset Allocation Dynamics and Pension Fund Performance,” *Journal of Business* 72, 429–461.
- Blake, David, Alberto Rossi, Allan Timmermann, Ian Tonks, and Russ Wermers, 2012, “Decentralized Investment Management: Evidence from the Pension Fund Industry,” forthcoming in *Journal of Finance*.
- Brinson, Gary P., L. Randolph Hood, and Gilbert L. Beebower, 1986, “Determinants of Portfolio Performance,” *Financial Analysts Journal* 42, 39–48.
- Brinson, Gary P., Brian D. Singer, and Gilbert L. Beebower, 1991, “Determinants of Portfolio Performance II: An Update,” *Financial Analysts Journal* 47, 40–48.
- Cutler, David M., James M. Poterba, and Lawrence H. Summers, 1991, “Speculative Dynamics,” *Review of Economic Studies* 58, 529–546.
- DeBondt, Werner F.M., and Richard Thaler, 1985, “Does Stock Market Overreact,” *Journal of Finance* 40, 793–805.
- Dichev, Ilia D., 2007, “What Are Stock Investors’ Actual Historical Returns? Evidence From Dollar-Weighted Returns,” *American Economic Review* 97, 386–401.
- Dyck, Alexander, and Lukasz Pomorski, 2011, “Is Bigger Better? Size and Performance in Pension Plan Management,” Working Paper.
- Friesen, Geoffrey, and Travis Sapp, 2007, “Mutual Fund Flows and Investor Returns: An Empirical Examination of Fund Investor Timing Ability,” *Journal of Banking & Finance* 31, 2796–2816.

- Goyal, Amit, and Sunil Wahal, 2008, “The Selection and Termination of Investment Management Firms by Plan Sponsors,” *Journal of Finance* 63, 1805–1847.
- Heisler, Jeffrey, Christopher Knittel, John Neumann, and Scott Stewart, 2007, “Why Do Institutional Plan Sponsors Fire Their Investment Managers?” *Journal of Business and Economic Studies* 13, 88–115.
- Ibbotson, Roger G., and Paul D. Kaplan, 2000, “Does Asset Allocation Policy Explain 40, 90, or 100 Percent of Performance?,” *Financial Analysts Journal* 56, 16–19.
- Ilmanen, Antti, 2011, *Expected Returns*, Wiley, Chichester, U.K..
- Jegadeesh, Narasimhan, and Sheridan Titman, 1993, “Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency,” *Journal of Finance* 48, 65–91.
- Kritzman, Mark, and Sébastien Page, 2002, “Asset Allocation Versus Security Selection: Evidence from Global Markets,” *Journal of Asset Management* 3, 202–212.
- Kritzman, Mark, and Sébastien Page, 2003, “The Hierarchy of Investment Choice,” *Journal of Portfolio Management* 29, 11–23.
- Lakonishok, Josef, Andrei Shleifer, and Robert Vishny, 1992, “The Structure and Performance of the Money Management Industry,” *Brookings Papers: Microeconomics*, 339–379.
- Lerner, Josh, Antoinette Schoar, and Jianlan Wang, 2008, “Secrets of the Academy: The Drivers of University Endowment Success,” *Journal of Economic Perspectives* 22, 207–222.
- Moskowitz, Tobias J., Yao Hua Ooi, and Lasse H. Pedersen, 2012, “Time Series Momentum,” *Journal of Financial Economics* 104, 228–250.
- Pedersen, Mitchell A., 2009, “Estimating Standard Errors in Finance Panel Data Sets: Comparing Approaches,” *Review of Financial Studies* 22, 435–480.
- Poterba, James M., and Lawrence H. Summers, 1988, “Mean Reversion in Stock Prices: Evidence and Implications,” *Journal of Financial Economics* 22, 27–59.
- Rauh, Joshua D., 2006, “Investment and Financing Constraints: Evidence from the Funding of Corporate Pension Plans,” *Journal of Finance* 61, 33–71.
- Stewart, Scott, John J. Neumann, Christopher R. Knittel, and Jeffrey Heisler, 2009, “Absence of Value: An Analysis of Investment Allocation Decisions by Institutional Plan Sponsors,” *Financial Analysts Journal* 65, 34–51.
- Zakamulin, Valeriy, 2013, “Secular Mean Reversion and Long-Run Predictability of the Stock Market”, Working paper.

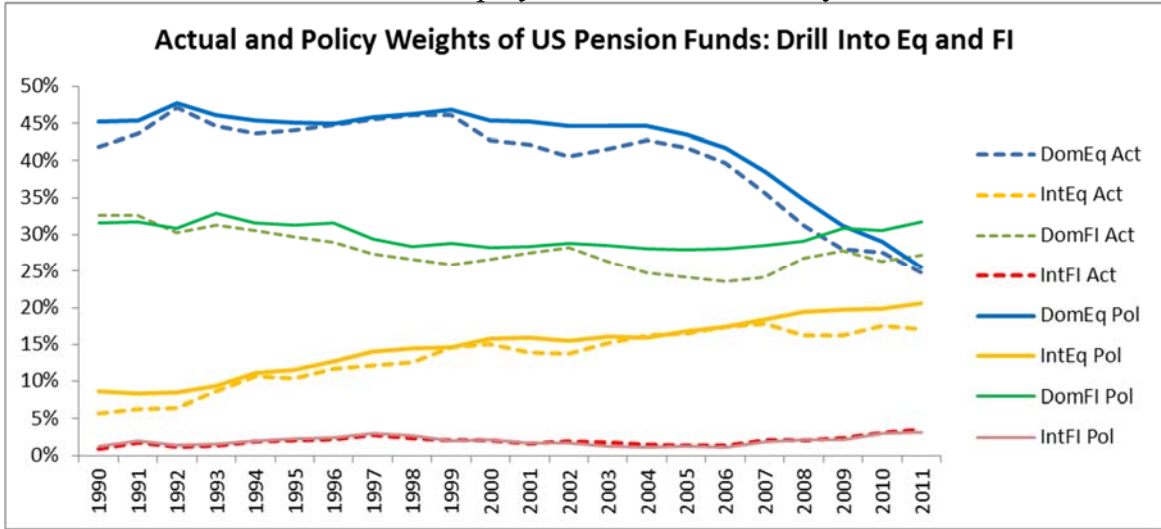
Figure 1: Actual and Policy Allocations to Various Asset Classes

We first calculate the cross-sectional average allocation across all the funds and then plot the time-series of these averages. Panel A shows allocation to all asset classes, Panel B shows allocations only to domestic/international equity and fixed-income asset class, and Panel C shows allocations to alternative asset classes. We show both the actual (Act) weights as well as policy weights (Pol). The policy weights are year-end weights while the actual weights are average weights during the year. The data include 573 defined-benefit U.S. pension plans. The sample period is 1990 to 2011.

Panel A: All Asset classes



Panel B: Equity and Fixed Income only



Panel C: Alternative asset classes only

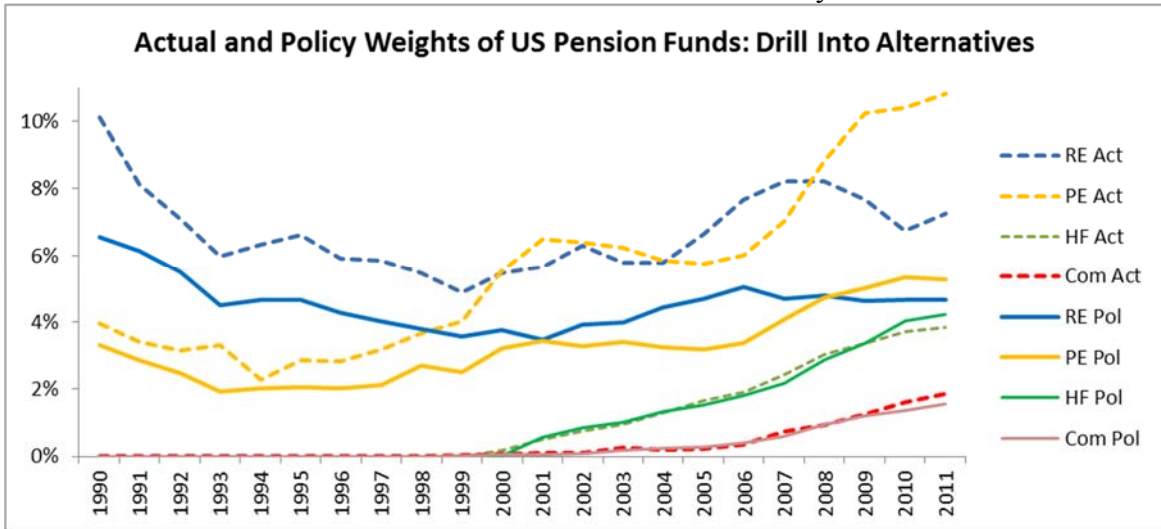


Figure 2: Momentum and Reversal Patterns in Financial Markets Over A Century

We study multi-country composite of non-U.S. equities (equally weighted with a gradually growing country universe), U.S. equities, MSCI EAFE equity index, multi-country composite of non-U.S. government bonds (equally weighted with a gradually growing country universe), Barclays Aggregate U.S. bond index, and S&P GSCI commodity index. The figure shows for each asset the average autocorrelation of lagged monthly returns in the previous 12 months (Y1), and in months lagged by 13-24 (Y2), 25-36 (Y3), 37-48 (Y4), 49-60 (Y5). The sample period is indicated next to each asset class.

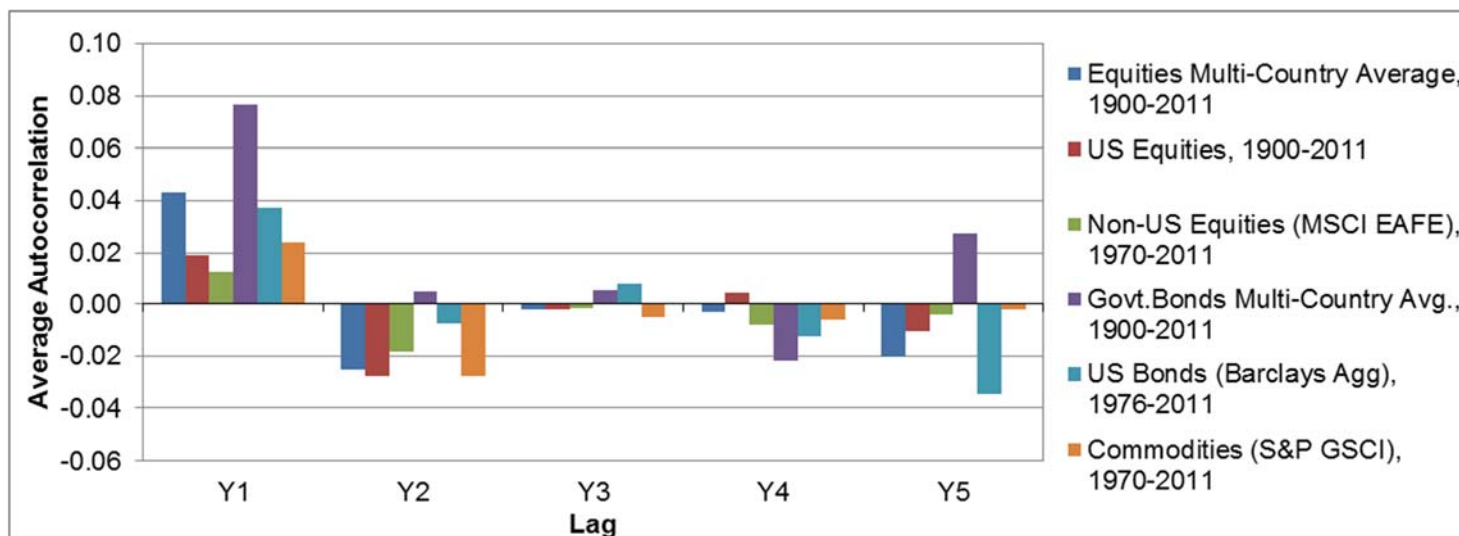


Figure 3: Stylized impulse-response on future returns and future policy weights

We plot the impulse-response functions from the following two regressions:

$$R_t = \mu + 0.034R_{t-1} - 0.154R_{t-2} + 0.076R_{t-3} - 0.060R_{t-4} + u_t$$

$$w_t = c + 0.052R_t + 0.054R_{t-1} + 0.083R_{t-2} + 0.082R_{t-3} + e_t$$

The coefficients for the return regression are taken from Table 3 while the coefficients for the weights regression are taken from the bold-face specification for equity and fixed-income cross-sectional regression specification in Panel B of Table 2.

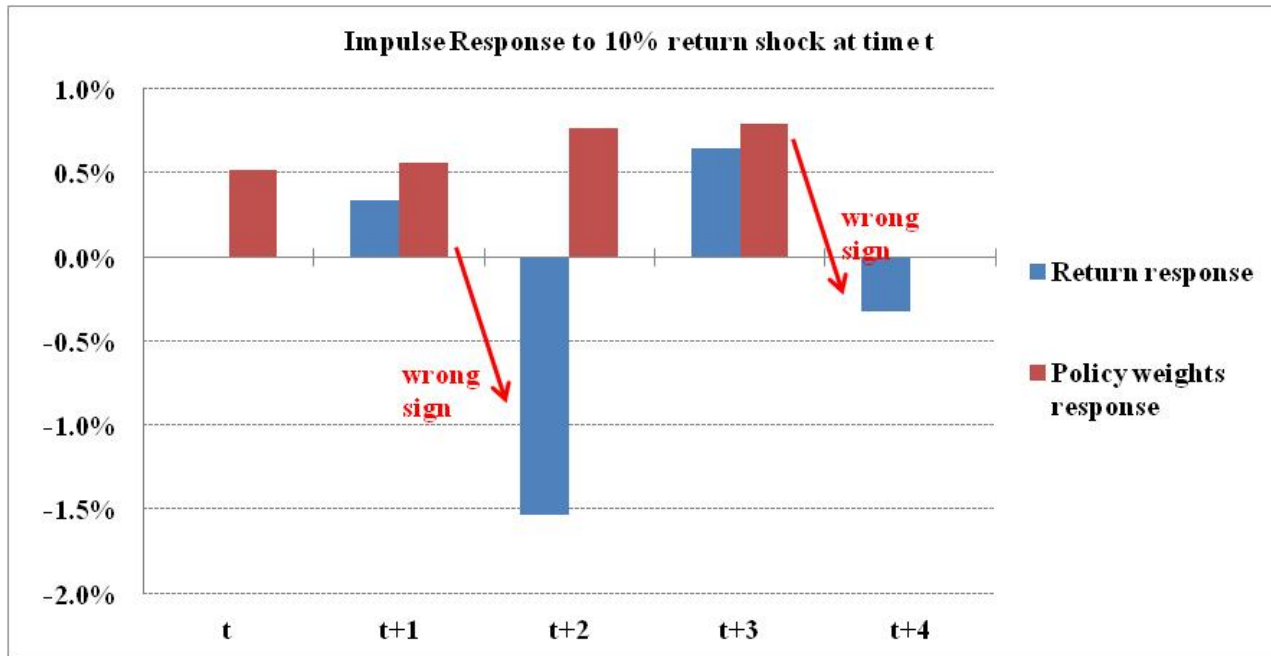


Table 1: Time-series regressions of policy weights on lagged returns

We run the following time-series regression of the average policy weight on contemporaneous and lagged returns.

$$\bar{w}_{a,t}^{\text{policy}} = \alpha_a + \sum_{j=0}^3 \beta_j R\text{bench}_{a,t-j} + e_{a,t}.$$

Regression is run separately for each of the nine asset classes as well as pooled versions of the first four and first nine asset classes. Pooled versions of the regressions have asset-specific intercepts. Panel B reports the results of the following specification in which we use only one variable corresponding to average returns over the last three-years.

$$\bar{w}_{a,t}^{\text{policy}} = \alpha_a + \beta_0 R\text{bench}_{a,t} + \beta_1 R\text{bench}_{a,t-1:t-3} + e_{a,t}.$$

We report the slope coefficients and their Newey-West corrected (with three lags) t -statistics in parentheses. The average policy weights are cross-sectional averages of 573 U.S. pension funds over 1990-2011. We pad benchmark returns for three years before 1990. Therefore the sample period is 1987-2011. Our preferred specification (with domestic and international equity and fixed income) is in bold-face.

Panel A											
	DomEq	IntEq	DomFI	IntFI	RE	PE	HF	Com	Cash	First 4	All
Return(t)	0.062 (1.72)	0.012 (0.41)	0.093 (2.19)	0.028 (1.02)	-0.036 (-1.27)	-0.028 (-2.92)	-0.031 (-7.24)	-0.006 (-1.49)	0.035 (0.21)	0.044 (1.43)	0.004 (0.27)
Return($t-1$)	0.063 (0.83)	0.027 (0.86)	0.143 (4.48)	0.022 (0.70)	-0.003 (-0.24)	-0.008 (-0.76)	-0.031 (-3.08)	-0.007 (-1.30)	0.191 (0.88)	0.055 (1.56)	0.010 (0.76)
Return($t-2$)	0.082 (0.89)	0.020 (0.72)	0.096 (1.83)	0.023 (0.97)	-0.001 (-0.05)	-0.013 (-1.05)	-0.039 (-2.38)	-0.007 (-1.30)	-0.256 (-1.03)	0.058 (1.53)	0.010 (0.69)
Return($t-3$)	0.142 (1.35)	-0.009 (-0.25)	0.027 (0.47)	-0.003 (-0.14)	-0.023 (-0.80)	-0.007 (-0.55)	-0.044 (-3.68)	-0.007 (-1.36)	0.408 (2.29)	0.062 (1.34)	0.013 (0.66)
adj- R^2	10.73	-19.19	11.00	-11.57	1.87	4.26	72.79	8.69	23.94	94.69	97.28

Panel B											
	DomEq	IntEq	DomFI	IntFI	RE	PE	HF	Com	Cash	First 4	All
Return(t)	0.063 (1.42)	0.009 (0.33)	0.090 (2.20)	0.029 (1.19)	-0.032 (-1.25)	-0.028 (-2.84)	-0.034 (-7.31)	-0.005 (-1.61)	0.004 (0.06)	0.044 (1.43)	0.002 (0.18)
Return($t-1:t-3$)	0.285 (1.29)	0.037 (0.53)	0.257 (2.85)	0.038 (0.61)	-0.022 (-1.00)	-0.029 (-1.20)	-0.111 (-3.86)	-0.020 (-1.50)	0.262 (1.89)	0.175 (1.69)	0.032 (0.76)
adj- R^2	16.38	-9.60	14.74	-2.36	10.49	13.99	76.43	18.17	17.12	94.81	97.29

Table 2: Panel regressions of policy weights on lagged returns

We run the following panel regression of the policy weight on contemporaneous and lagged returns.

$$w_{i,a,t}^{\text{policy}} = \delta_{1,n} + \delta_{2,a} + \delta_{3,t} + \sum_{j=0}^3 \gamma_j R_{i,a,t-j} + u_{i,a,t}$$

Regression is run separately for domestic equity, a pooled versions of the domestic and international equity and fixed income, and a pooled specification of all asset classes. We report only the slope coefficients on lagged returns together with their t -statistics in parentheses. Statistical significance is computed using double clustered errors for fund and time fixed effects. The average adjusted- R^2 and the average number of observations are reported in last row below each specification. The sample includes 573 U.S. pension funds over 1990-2011. Our preferred specification (with domestic and international equity and fixed income) is in bold-face.

	DomEq	EQ/FI (Dom/Int)	All
Return(t)	0.056 (0.50)	0.052 (1.75)	0.025 (1.22)
Return($t-1$)	0.038 (0.39)	0.054 (1.44)	0.025 (0.94)
Return($t-2$)	0.063 (0.96)	0.083 (2.14)	0.042 (1.47)
Return($t-3$)	-0.063 (-0.91)	0.082 (2.41)	0.063 (2.33)
R^2 /Nobs	75.5/ 1,516	59.8/ 4,831	72.3/ 7,182
Return(t)	0.039 (0.34)	0.048 (1.64)	0.025 (1.23)
Return($t-1:t-3$)	0.013 (0.21)	0.072 (2.48)	0.038 (2.69)
R^2 /Nobs	75.5/ 1,516	59.8/ 4,831	72.2/ 7,182

Table 3: Autocorrelations of Returns over a Century

We run the following pooled time-series regression of annual returns on lagged annual returns with asset-specific intercepts.

$$R_{a,t} = \mu_a + \sum_{j=1}^4 \rho_j R_{a,t-j} + u_{a,t}.$$

We include four major asset classes: U.S. equities, non-U.S. equities, U.S. Treasuries, non-U.S. government bonds. Sample period is 1900 to 2011 except for non-U.S. government bonds for which the data start in 1933. We report the autocorrelation coefficients with Newey-West adjusted t -statistics (with four lags) in parentheses.

Return($t-1$)	0.034 (0.55)
Return($t-2$)	-0.154 (-3.23)
Return($t-3$)	0.076 (1.17)
Return($t-4$)	-0.060 (-0.62)
adj- R^2	3.9