Asset Allocation and Bad Habits

Andrew Ang, Amit Goyal, and Antti Ilmanen

Prof. Andrew Ang is at Columbia Business School, New York (United States).

Prof. Amit Goyal is at Université de Lausanne (Switzerland).

Antti Ilmanen is Managing Director, AQR Capital Management, based in London (United Kingdom).

This article documents the "bad habits" of investors in asset allocation practices. Whereas financial markets exhibit momentum over multi-month horizons but more reversion to the mean over multi-year horizons, many investors act like momentum investors even at these longer horizons. Both these patterns are well known anecdotally but have not been well documented statistically, especially together. This article therefore addresses two empirical questions. First, How do funds reallocate based on past returns? The authors provide direct evidence using the CEM Benchmarking data on pension fund target allocations over a 22-year period. Second, What are momentum/reversal patterns in financial markets returns? Evidence is provided using more than a century of data. Merging the findings from the two data sets provides evidence consistent with the premise that investors chase returns over multi-year horizons, which is likely to hurt their long-run performance. However, the statistical evidence on pro-cyclical multi-year asset allocations and multi-year mean reversion patterns in asset-class returns is on the borderline of statistical significance.

Keywords: Asset Allocation, Mean Reversion, Momentum Investing, Pension Fund, Return Attribution

Attributing Investment Returns

At its simplest level, the process of pension fund investment management involves (1) drafting an investment policy statement, (2) defining asset allocation, and (3) selecting investment managers in specified mandates. There is an active debate as to the relative contributions of Steps 2 and 3 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 in 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) find that pension funds' active returns are roughly equally driven by asset allocation, market timing, and security selection.

While there have been many studies of return attribution, almost none have directly examined the impact of past returns on asset allocation. This article is an attempt to fill this gap, focusing on how asset allocations depend on past performance.[1]

Many pension funds rebalance their asset-class allocations regularly to specific target weights, such as the conventional 60% stocks and 40% bonds. But there is anecdotal evidence that funds may let their allocations drift with relative asset-class performance. This may reflect passive buy-and-hold policies, 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 – that is, to 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 of past performance and the high costs of transition. Ang and Kjaer (2011) argue that one of the biggest vices of long-term investors is pro-cyclical 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, in the aggregate, earn lower dollar-weighted returns than time-weighted returns is indicative of ill-timed investor flows (Dichev 2007; Friesen and Sapp 2007). Institutional investors' fire-and-hire decisions at the manager level have also been shown to lose value over time (Goyal and Wahal 2008).

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

A simple numerical example illustrates our basic idea. Imagine that there are three years (Year 0, Year 1, and Year 2) and two asset classes (stocks and bonds). Returns on stocks are 10%, 10%, and 5% for the three years, while 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% × 5% + 33% × 10% = 6.7%. If the fund knew ahead of time about the reversal in Year 2, it would revise its asset allocation to 33% stocks and 67% bonds, earning 33% × 5% + 67% × 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 beyond one year has a deleterious effect.

Despite well-known anecdotal evidence, we do not have enough long-term data to make statistically conclusive statements about multi-year investor behavior. Some types of pro-cyclical flows are especially hard to discern. Consider changes made in investor benchmarks, where new asset classes tend to be added after multi-year rallies. Because decisions to change the benchmark are rarely evaluated over time, such pro-cyclicality 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 the consequent long-run losses. We use annual data from CEM Benchmarking on the evolving asset allocations of American pension funds from 1990 through 2011 to provide direct evidence on return-chasing behavior at the 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 pro-cyclical 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.

Key Findings

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 the one-year horizon, and instead the typical pension fund keeps chasing returns over multi-year horizons, to the detriment of the institution's long-run wealth. However, the statistical significance of these findings is weak.

The studies closest to ours in terms of scope are those by Heisler et al. (2007) and Stewart et al. (2009), who show that investment products that receive contributions later underperform products that experience withdrawals over one to five years. Post-flow underperformance is due more to product (manager) selection than to category (asset class) reallocation, but both activities detract value. The authors' data are based on the 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 Andonov et al. (2012), who use CEM Benchmarking data to study North American institutional plans' asset-class allocations and returns over time. The focus of these studies is to study scale economies in asset management; Andonov et al. (2012) also decompose the return into its components of asset allocation, market timing, and security selection. However, neither study looks at multi-year return chasing or relates the findings to medium-term reversal patterns in asset returns.

Data Description

CEM Benchmarking Inc. 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 peer comparisons in costs, asset allocations, and performance. Voluntary reporting raises the issue of selection bias, but earlier studies have found no evidence of such bias in fund returns (Bauer, Cremers, and Frehen 2010; Andonov et al. 2012). We have data on 978 different pension funds over the 22-year period from 1990 to 2011; this study focuses on 573 American pension funds, but preliminary analysis suggests similar results for the global universe. The funds vary in size, with median assets under management (AUM) near $3 billion and mean AUM near US$10 billion. The participating funds hold 30%–40% of AUM by American pension funds and about 4% of American equity market capitalization (Andonov et al. 2012).

The basic information in CEM Benchmarking surveys includes the fund portfolio's *actual* and *policy* weights as well as its realized returns. An actual weight is the average weight in an asset class during the year, while the policy weight is the weight in an asset class dictated by strategic policy portfolio considerations. Policy weights are available on a calendar-year-end basis and are the target weights for the next calendar year.

The weights and returns are available at a quite detailed level (e.g., weights in internal passive large-cap American equities). Because our calculations do not require such detailed information, we 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 parentheses next to each of the nine asset classes:

1. Domestic equity (CRSP value-weighted)
2. International equity (MSCI AC World ex-US)
3. Domestic fixed income (Barclays US aggregate)
4. International fixed income (Barclays Global aggregate ex-US, currency-hedged)
5. Real estate (NCREIF)
6. Private equity (Cambridge Associates)
7. Hedge funds (HFRI)
8. Commodities (S&P GSCI commodities)
9. Cash (US T-bill)

The data can be split in different ways; for instance, we can consider separating international equity and fixed income into their developed markets and emerging markets components. This could be interesting, given institutional investors' recent movement into emerging markets. However, the lack of reliable long-term data on these sub-classes prevents us from attempting this finer classification.

Our chosen benchmarks are those that best reflect asset-class returns for the corresponding asset class. At the same time, these choices must to some extent remain ad hoc and, therefore, debatable. For instance, NAREIT may arguably be a better benchmark than NCREIF for some investors who do not invest in the private real estate market. We leave exploration of these fine distinctions to future research.[3]

Strategic Target Asset Allocations

Figure 1 sets out the time series of cross-sectional averages of actual and policy weights to the nine asset classes. A noteworthy 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),[4] but no such mechanical relation is expected ex ante for policy weights. It is therefore more useful to conduct our analysis using policy weights. This approach also has two other advantages. First, policy weights are strategic targets deliberately chosen by pension funds' Boards, so they can be more unambiguously interpreted as active decisions; therefore, studying policy weights' dependence on past market returns more closely matches our objective of analyzing the bad habits of institutional return chasing. Second, the data on realized (actual) holdings in our data set are not year-end weights but average weights, and the date mismatch between calendar-year returns and average-year actual weights further complicates analysis and interpretation of results calculated using actual weights.

As Figure 1 shows, 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. Figure 1A shows policy weights for equities rising from 54% in 1990 to a peak of 61% in 1999–2001 before falling to 46% in 2011. Fixed-income weights fell from 33% in 1990 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% in 1990 to 6% in the late 1990s before rising to 16% in 2011. Large changes in policy allocation during the sample period reveal that characterizing pension funds as rebalancing to fixed 60/40 stock/bond weights is a gross oversimplification of industry practices.

Actual weights exhibit similar time paths (see Figure 1A), 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.

The patterns of equity and fixed-income asset allocations shown in Figure 1B reveal that the broad equity allocation conceals a dramatic reduction in home bias. Policy weights on international equities rose fairly consistently, from 9% in 1990 to 21% in 2011, while the share of American equities hovered around 40%s until 2005, then fell sharply to 25% by 2011, indicating a significant decline in home bias. Fixed-income assets remained significantly (~90%) home-biased, reflecting the fact that institutional liabilities are mainly dollar-denominated.

Drilling into alternatives in Figure 1C reveals gradual shifts over time: 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 to private equity and real estate have consistently exceeded the target allocations, averaging about 2% each. The private equity overweight grew to >5% in 2011 (with actual allocation exceeding 10%), perhaps not an altogether deliberate choice by pension funds but a result of commitments made during the earlier "good years."

Figure 1: Cross-sectional average actual (Act) and policy (Pol) allocations to various asset classes by 537 American DB pension plans, 1990–2011, for (A) all asset classes; (B) domestic and international equity and fixed-income classes; and (C) alternative asset classes.

Figure A

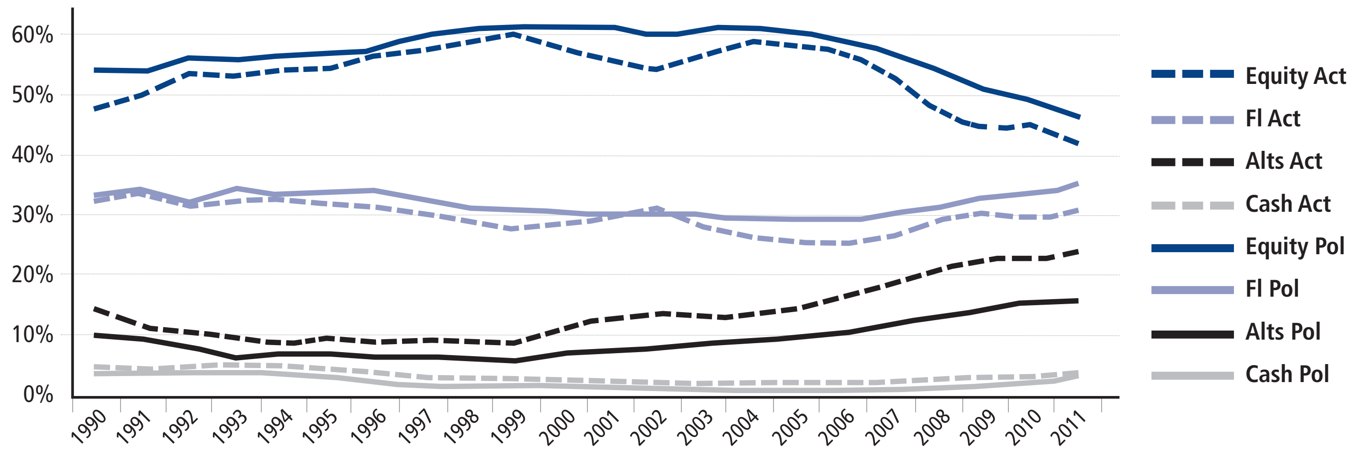


Figure B

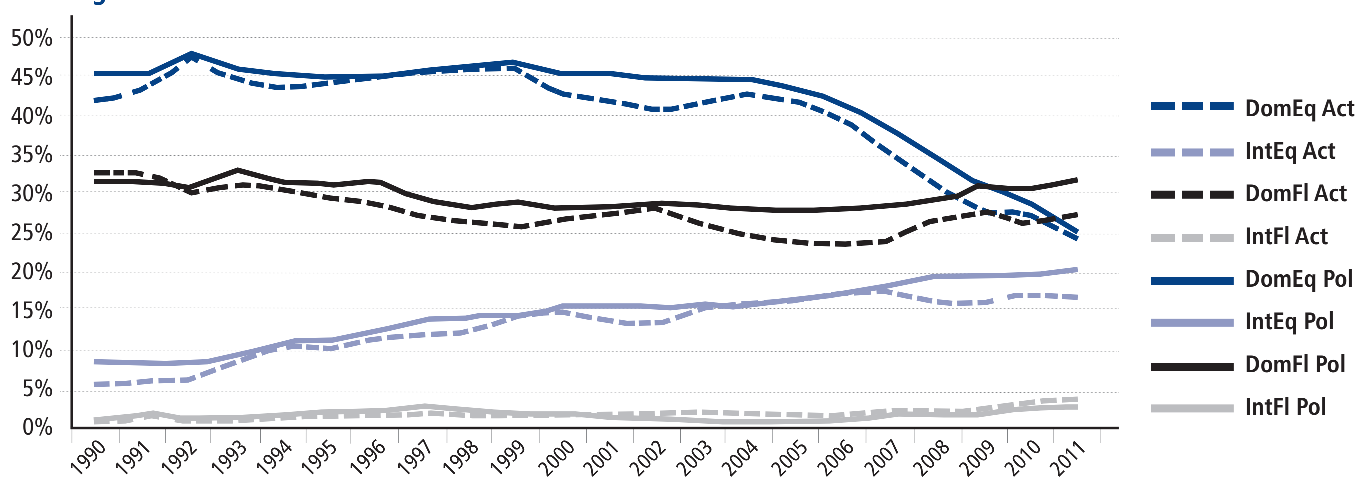
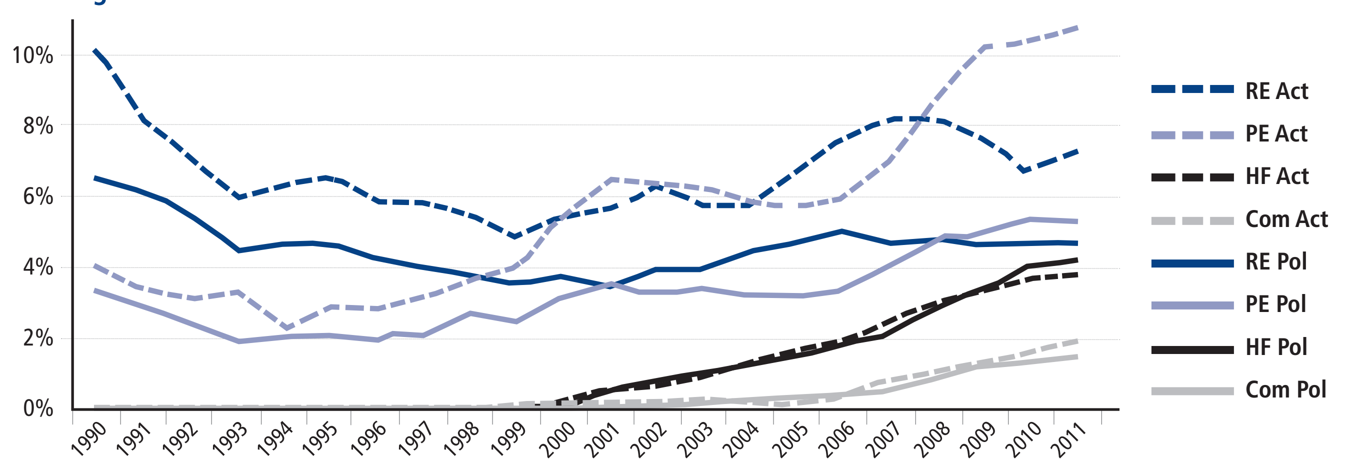


Figure C



Do Funds Reallocate Based on Past Returns?

To answer our next key question – How does the average fund change its policy weights based on past returns? – we first look at time-series regressions. We regress average industry policy weights in an asset class on past returns; in other words, we define average weights as $\bar{w}^{\text{policy}}_{a,t} = \frac{1}{N}\sum_{i=1}^{N} w^{\text{policy}}_{i,a,t}$ and run variants of the following time-series regression for each asset class:[5]

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

Table 1A reports the results of Equation (1), showing Newey–West adjusted t-statistics in parentheses below the coefficients.

Since we only 22 years' worth of data, we also use the following specification in Table 1B, which has only two independent variables that still capture the essence of the relationship between weights and past returns:

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

Finally, we remedy the problem of limited time-series to some extent by padding benchmark returns for three years before 1990 – in other words, by using the average weights from 1990 to 2011 and benchmark returns (whenever available) from 1987 to 2011 to predict these weights. Increasing the number of observations should increase the power of our tests and lead to higher statistical significance.

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

Panel A

	Dom Eq	Int Eq	Dom FI	Int FI	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)
Adjusted 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

	Dom Eq	Int Eq	Dom FI	Int FI	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)
Adjusted R^2	16.38	−9.60	14.74	−2.36	10.49	13.99	76.43	18.17	17.12	**94.81**	97.29

Dom Eq = domestic equity; Int Eq = international equity; Dom FI = domestic fixed income; Int FI = international fixed income; RE = real estate; PE = private equity; HF = hedge funds; Com = commodities; First 4 = Dom Eq, Int Eq, Dom FI, Int FI.

**We run the following time-series regression of the average policy weight on contemporaneous and lagged returns.The regression is run separately for each of the nine asset classes as well as for pooled versions of the first four and all nine asset classes. The pooled 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 previous three years:*

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

We report the slope coefficients and their Newey–West corrected t-statistics (with three lags) in parentheses. The average policy weights are cross-sectional averages of 573 American pension funds over the 1990–2011 period; benchmark returns are padded for three years before 1990, so the sample period is 1987–2011. Our preferred specification (Dom Eq, Int Eq, Dom FI, Int FI) is shown in bold.

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

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

These considerations lead us to focus more on pooled results, which also allows us to use our limited data more effectively. For the pooled four-asset case (Dom Eq, Int Eq, Dom FI, Int FI), the coefficients on Return (t) and Return (t–1:t–3) are 0.044 (t = 1.43) and 0.175 (t = 1.69), respectively. However, the statistical significance is low (most values of t are <2, and the small number of observations can even lead to negative adjusted R^2 values for some single asset classes).[6] Since one standard deviation of annualized returns corresponds to roughly 10% (~20% for equity, ~5% for fixed income), the economic magnitude of weights in response to a move of one standard deviation in each-year past returns is around 0.6%.[7]

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, but the statistical significance is low, most likely because of the limited length of the data series.

Using the Entire Cross-Section of Funds

Perhaps more explanatory power can be gained by using the entire cross-section of funds. We run the following panel regression:

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

where δ_1, δ_2, and δ_3 are dummies for funds, asset class, and time, respectively. Thus, the heterogeneity in the sample resulting from different funds, asset classes, and years is partly accounted for by the fixed effects. The main coefficients of interest are the γ's. Note that the explanatory variables are not 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 the pooled panel of four asset classes, and for all asset classes.

The results are reported in Table 2 (coefficients on dummies are not reported). Coefficients on returns are positive, but statistical significance varies: while many t-statistics exceed 2, others fall below it.[8] In our preferred specification (Dom Eq, Int Eq, Dom FI, Int FI), the coefficients on Return (t) and Return (t–1) are positive, though not statistically significant. This shows that pension funds increase policy weights in response to recently realized returns. At the same time, the coefficients on Return (t–2) and Return (t–3) are positive and statistically significant, which means that pension funds' weight allocations react positively even to returns realized a few years ago.

Table 2: Panel Regressions of Policy Weights on Lagged Returns*

	Dom Eq	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

Dom = domestic; Eq = equity; FI = fixed income; Int = international; Nobs = number of observations.

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

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

The regression is run separately for domestic equity, for pooled versions of domestic and international equity and fixed income, and for 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. Average adjusted R_2 and average number of observations are reported in last row below each specification. The sample includes 573 American pension funds over the 1990–2011 period. Our preferred specification (Dom Eq, Int Eq, Dom FI, Int FI) is shown in bold.

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

What Are the 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, to study a multi-country composite of non-American equities (equally weighted, with a gradually growing country universe), American equities, MSCI EAFE equity index, a multi-country composite of non-American government bonds (equally weighted, with a gradually growing country universe), Barclays Aggregate US bond index, and the S&P GSCI commodity index. All data are provided by AQR Capital.

There are two reasons for using such long data histories. First, any attempt to capture statistical patterns of multi-year autocorrelations requires decades of data to achieve even a double-digit number of independent observations; shorter windows are likely to produce spurious relationships, 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 suffer not only from the short sampling period but also from a look-ahead bias. Investors know the in-sample autocorrelation patterns only with hindsight; their real-time expectations on the persistence of asset returns, which influenced their asset allocation decisions through the past 20 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 going back to 1900.

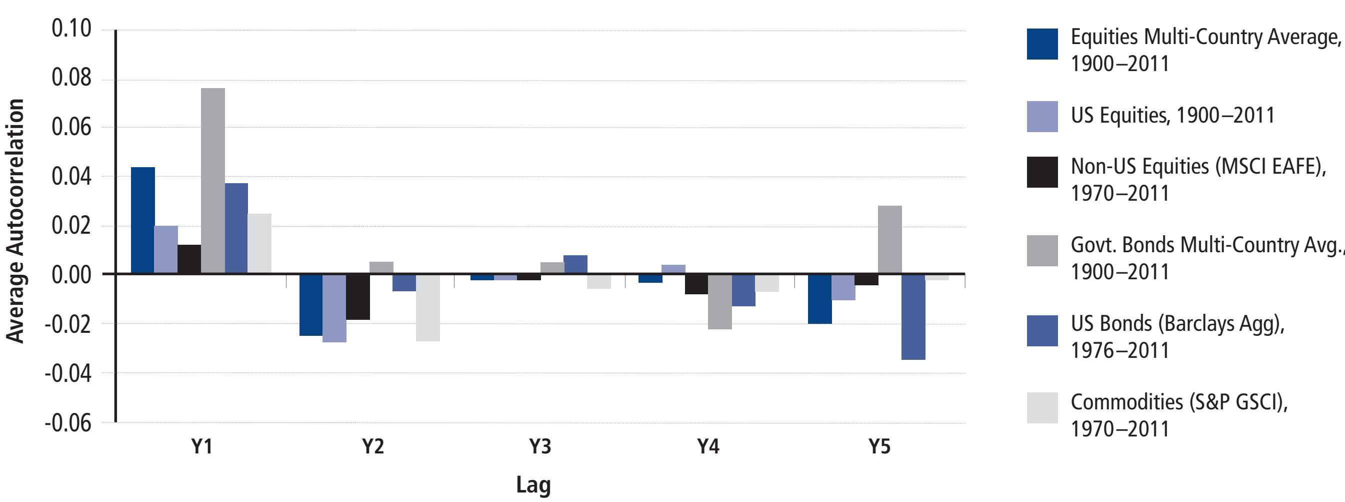
At the same time, concerns about selection bias arise from using such long histories of data; for example, markets such as Russia or China are excluded from the early parts of the sample. The implications of such exclusions on momentum and reversal patterns for broad asset classes are not immediately obvious, however, and we hope that they do not create 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 four previous periods, going back 13–24 months (Y2), 25–36 months (Y3), 37–48 months (Y4), and 49–60 months (Y5). As was already recognized 20 years ago (e.g., Cutler, Poterba, 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-rapid portfolio rebalancing (which will be hurt by momentum) is as detrimental to institutional investors as too-persistent return chasing (which will be hurt by long-term mean reversion). However, our annual asset allocation data are too coarse to allow us to study within-year rebalancing. It is worth noting that when we only have 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. These are provided in the next section.

Figure 2: Momentum and Reversal Patterns in Financial Markets over a Century



For each asset class, the figure shows the average autocorrelation of lagged monthly returns in the previous 12 months (Y1) and in periods lagged by 13–24 months (Y2), 25–36 months (Y3), 37–48 months (Y4), and 49–60 months (Y5). The sample period is indicated next to each asset class.

Implications of Our Findings

We combine our findings on pension funds' behavior and financial markets' behavior by creating an impulse-response graph that tracks the cumulative impact of a return shock at year t on the policy weights and future returns of the same asset class through the next few years.

Recall from our fund behavior findings that high returns at time t lead pension funds to allocate more weight to the asset class not only at t but apparently also later – for instance, at time $t + 3$. In contrast, observed market return autocorrelations suggest that high returns at t are correlated with high returns at $t + 1$ but low returns from $t + 2$ to $t + 4$. The multiple regression that we run on asset returns is

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

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 and 2011: American equities, non-American equities, American Treasuries, and non-American government bonds (for which data start in 1933). We use the coefficients on the four lagged returns in the impulse-response function: 0.034, –0.154, 0.076, –0.060.[9] Note that the first coefficient on previous-year return likely reflects a mixture of positive autocorrelations over a few months and negative autocorrelations beyond that point. The positive coefficient on the third year is surprising but not statistically significant – indeed, only the negative coefficient on the second year achieves statistical significance.[10]

Table 3: Autocorrelations of Returns over a Century*

	ρ
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)
Adjusted R^2	3.9

*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: American equities, non-American equities, US Treasuries, and non-US government bonds. Sample period is 1900 to 2011, except for non-US 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.

These autocorrelation patterns should lead funds to increase policy weights at time t (to benefit from anticipated high returns at $t + 1$) and reduce them at $t + 1$ to $t + 3$ (to benefit from anticipated low returns from $t + 2$ to $t + 4$).[11] We therefore suggest the following regression test for policy weights:

$$w_t = c + \beta_0 R_t + \beta_1 R_{t-1} + \beta_2 R_{t-2} + \beta_3 R_{t-3} + e_t \qquad (5)$$

Optimal weightings would see positive β_0 and negative β_1 to β_3. The results shown in Tables 1 and 2 indicate, however, that in fact all four coefficients are positive.

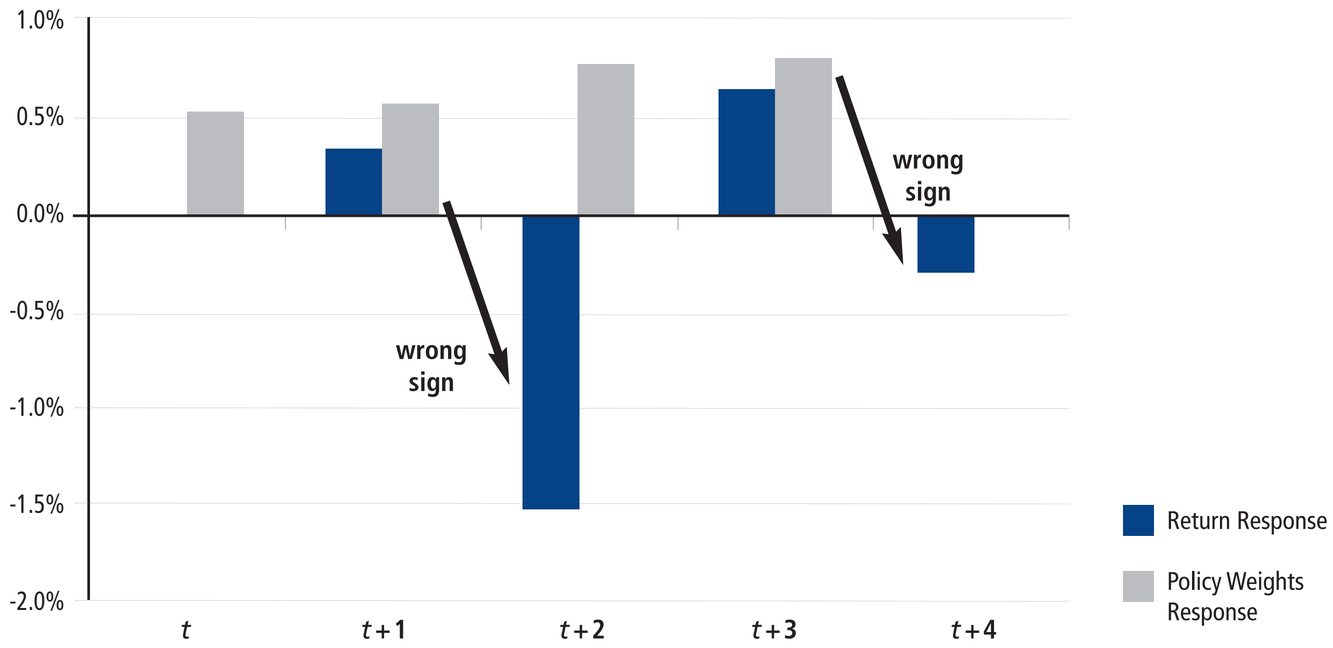
Given coefficients ρ and β, we can plot the impulse response to, say, a 10% shock in returns at time t. For returns, we use the regression coefficients from the pooled equity and bond momentum / reversal analysis shown in Table 3; for policy weights, we use the coefficients from the cross-sectional Fama–MacBeth regression (Fama and MacBeth 1973) for the pooled equity and bond asset allocation analysis shown in Table 2. Our impulse-response function differs from a standard function in that we use coefficients from Equations (4) and (5) estimated separately; thus, the dependence between the two equations comes from the effect of common independent variables. In particular, we apply a shock of 10% to Return (t) and let this shock permeate through Equations (4) and (5) independently.

Reconsidering Asset Allocation Practices

The impulse-response is shown in Figure 3. The arrows highlight a time shift that must be taken into account in interpreting this graph. The fact that policy weights go up at time t is good, as high weights benefit from high-momentum returns at time $t + 1$; on the other hand, the fact that the policy weights are high at time $t + 1$ is bad, as these positions will suffer from return reversals at time $t + 2$, and so on.

Figure 3 summarizes our key finding about the bad habit of excessively pro-cyclical asset allocation by pension funds. In the aggregate, pension funds do not recognize the shift from momentum to reversal tendencies in asset returns beyond a one-year horizon; instead, a typical pension fund keeps chasing returns over multi-year horizons, to the detriment of the institution's long-term wealth. We hope that this evidence will help at least some pension funds to reconsider their asset allocation practices.

Figure 3: Stylized Impulse-Response on Future Returns and Future Policy Weights



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

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

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

Coefficients for the return regression are taken from Table 3; coefficients for the weights regression are taken from the specification for the equity and fixed-income cross-sectional regression specification (shown in bold type) in Table 2.

Endnotes

1 Other determinants of asset allocation have been studied. For instance, Dyck and Pomorski (2011) document positive scale economies in asset management; however, their conclusions are disputed by Andonov et al. (2012), who argue that larger pension funds would be better off if they invested 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) have found that corporate pension funds' investment policies are 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 et al. 1992).

3 It is important to understand a few limitations of our data set. The main one is that while cross-sectional coverage is good, the data are available only at annual frequency, which means that we do not have enough power in many of the time-series tests described later in the article. Another is that although we have actual weights in different asset classes, these are average weights across 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.

4 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 $R_{i,a,t}$. Denote total fund returns by $R_{i,t}$. Construct what the weights would have been due to just return realization and call these passive weights, We can then easily show that $w_{i,a,t}^{\text{passive}}$.

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

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 a fund does not consciously rebalance its portfolio, the year-end weights will inherit the relative asset-class returns realized during the year.

5 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, αa. 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: funds seem to be shifting policy weights in response to past returns, apart from any mechanical effect of returns.

6 The adjusted R_2 for the pooled specification is high because asset-specific intercepts allow the regression to explain a large part of cross-asset variation in weights.

7 One disputed issue is whether to use levels of or changes in policy weights as the dependent variable. In unreported results, we find that the adjusted 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 have explored many other specifications, producing 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 must be 1, 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 and calculate average policy returns across the funds to get a weighted-average benchmark return,

$$\overline{Rb}_t = \frac{1}{N}\sum_{i=1}^{N} R_{i,t}^{\text{policy}} .$$

We then run the following regression:

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

We can run this regression only over the CEM Benchmark sample for 1990–2011, which contains all the required data. 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.

8 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 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 Year 1.

9 The four coefficients are 0.089, –0.208, 0.156, and –0.055 if we use the sample period 1900–1989, that is, only the information that was available to investors before our asset allocation data set begins. These coefficients are broadly similar to those for the full sample shown in Table 3.

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

11 Admittedly, a strict reading of 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 or smooth these results and view all return responses ($t + 2, t + 3, t + 4$) as mildly negative.

References

Addoum, Jawad M., Jules H. van Binsbergen, and Michael W. Brandt. 2010. "Asset Allocation and Managerial Assumptions in Corporate Pension Plans." Working paper. http://dx.doi.org/10.2139/ssrn.1710902

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. http://dx.doi.org/10.2139/ssrn.1885536

Ang, Andrew, and Knut N. Kjaer. 2011. "Investing for the Long Run." In A Decade of Challenges: *A Collection of Essays on Pensions and Investments*, ed. Tomas Franzen, 94–111. Andra AP-fonden, Second Swedish National Pension Fund.

Asness, Clifford S., Tobias J. Moskowitz, and Lasse H. Pedersen. 2013. "Value and Momentum Everywhere." *Journal of Finance* 68 (3): 929–85. http://dx.doi.org/10.1111/jofi.12021

Bauer, Rob, Martijn Cremers, and Rik Frehen. 2010. "Pension Fund Performance and Costs: Small Is Beautiful." Working paper. http://dx.doi.org/10.2139/ssrn.965388

Blake, David, Bruce N. Lehmann, and Alan Timmermann. 1999. "Asset Allocation Dynamics and Pension Fund Performance." *Journal of Business* 72 (4): 429–61. http://dx.doi.org/10.1086/209623

Brinson, Gary P., L. Randolph Hood, and Gilbert L. Beebower. 1986. "Determinants of Portfolio Performance." *Financial Analysts Journal* 42 (4): 39–44. http://dx.doi.org/10.2469/faj.v42.n4.39

Brinson, Gary P., Brian D. Singer, and Gilbert L. Beebower. 1991. "Determinants of Portfolio Performance II: An Update." *Financial Analysts Journal* 47 (3): 40–48. http://dx.doi.org/10.2469/faj.v47.n3.40

Cutler, David M., James M. Poterba, and Lawrence H. Summers. 1991. "Speculative Dynamics." *Review of Economic Studies* 58:529–46.

De Bondt, Werner F.M., and Richard Thaler. 1985. "Does the Stock Market Overreact?" *Journal of Finance* 40 (3): 793–805. http://dx.doi.org/10.1111/j.1540-6261.1985.tb05004.x

Dichev, Ilia D. 2007. "What Are Stock Investors' Actual Historical Returns? Evidence from Dollar-Weighted Returns." *American Economic Review* 97 (1): 386–401. http://dx.doi.org/10.1257/aer.97.1.386

Dyck, Alexander, and Lukasz Pomorski. 2011. "Is Bigger Better? Size and Performance in Pension Plan Management." Working paper. http://dx.doi.org/10.2139/ssrn.1690724

Fama, Eugene F., and James D. MacBeth. 1973. "Risk, Return and Equilibrium: Empirical Tests." *Journal of Political Economy* 81 (3): 607–36. http://dx.doi.org/10.1086/260061

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 (9): 2796–816. http://dx.doi.org/10.1016/j.jbankfin.2007.01.024

Goyal, Amit, and Sunil Wahal. 2008. "The Selection and Termination of Investment Management Firms by Plan Sponsors." *Journal of Finance* 63 (4): 1805–47. http://dx.doi.org/10.1111/j.1540-6261.2008.01375.x

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 (3): 16–19. http://dx.doi.org/10.2469/faj.v56.n3.2356

Ilmanen, Antti. 2011. *Expected Returns*. Chichester, UK: Wiley.

Jegadeesh, Narasimhan, and Sheridan Titman. 1993. "Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency." *Journal of Finance* 48 (1): 65–91. http://dx.doi.org/10.1111/j.1540-6261.1993.tb04702.x

Kritzman, Mark, and Sébastien Page. 2002. "Asset Allocation versus Security Selection: Evidence from Global Markets." *Journal of Asset Management* 3 (3): 202–12. http://dx.doi.org/10.1057/palgrave.jam.2240075

Kritzman, Mark, and Sébastien Page. 2003. "The Hierarchy of Investment Choice." *Journal of Portfolio Management* 29 (4): 11–23. http://dx.doi.org/10.3905/jpm.2003.319891

References (cont'd)

Lakonishok, Josef, Andrei Shleifer, Robert Vishny, Oliver Hart, and George L. Perry. 1992. "The Structure and Performance of the Money Management Industry." *Brookings Papers: Microeconomics* 1992:339–79. http://dx.doi.org/10.2307/2534766

Lerner, Josh, Antoinette Schoar, and Jianlan Wang. 2008. "Secrets of the Academy: The Drivers of University Endowment Success." *Journal of Economic Perspectives* 22 (3): 207–22. http://dx.doi.org/10.1257/jep.22.3.207

Moskowitz, Tobias J., Yao Hua Ooi, and Lasse H. Pedersen. 2012. "Time Series Momentum." *Journal of Financial Economics* 104 (2): 228–50. http://dx.doi.org/10.1016/j.jfineco.2011.11.003

Pedersen, Mitchell A. 2009. "Estimating Standard Errors in Finance Panel Data Sets: Comparing Approaches." *Review of Financial Studies* 22 (1): 435–80. http://dx.doi.org/10.1093/rfs/hhn053

Poterba, James M., and Lawrence H. Summers. 1988. "Mean Reversion in Stock Prices: Evidence and Implications." *Journal of Financial Economics* 22 (1): 27–59. http://dx.doi.org/10.1016/0304-405X(88)90021-9

Rauh, Joshua D. 2006. "Investment and Financing Constraints: Evidence from the Funding of Corporate Pension Plans." *Journal of Finance* 61 (1): 33–71. http://dx.doi.org/10.1111/j.1540-6261.2006.00829.x

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 (6): 34–51. http://dx.doi.org/10.2469/faj.v65.n6.4

Zakamulin, Valeriy. 2013. "Secular Mean Reversion and Long-Run Predictability of the Stock Market." Working paper. http://dx.doi.org/10.2139/ssrn.2209048