Aalto University

OPEN ACCESS

This is an electronic reprint of the original article.
This reprint may differ from the original in pagination and typographic detail.

Bollen, Nicolas P. B.; Joenvaara, Juha; Kauppila, Mikko

Hedge Fund Performance

Published in:
Financial Analysts Journal

DOI:
10.1080/0015198X.2021.1921564

Published: 01/01/2021

Document Version
Peer-reviewed accepted author manuscript, also known as Final accepted manuscript or Post-print

Please cite the original version:
Bollen, N. P. B., Joenvaara, J., & Kauppila, M. (2021). Hedge Fund Performance: End of an Era? *Financial Analysts Journal*, *77*(3), 109-132. https://doi.org/10.1080/0015198X.2021.1921564

This material is protected by copyright and other intellectual property rights, and duplication or sale of all or part of any of the repository collections is not permitted, except that material may be duplicated by you for your research use or educational purposes in electronic or print form. You must obtain permission for any other use. Electronic or print copies may not be offered, whether for sale or otherwise to anyone who is not an authorised user.

Hedge Fund Performance: End of an Era?

NICOLAS P.B. BOLLEN*, JUHA JOENVÄÄRÄ† and MIKKO KAUPPILA‡

ABSTRACT

This paper documents a decline in aggregate hedge fund performance over the past decade. We test whether a set of prediction models can select subsets of individual funds that buck the trend and subsequently outperform. Two of the predictors reliably pick funds that lower the volatility and raise the Sharpe ratio of a multi-asset class portfolio relative to a stock/bond portfolio over the full 1997–2016 sample. Hedge fund allocations reduce volatility across two sub-periods but fail to improve the Sharpe ratio from 2008 onwards. Potential explanations for the erosion of hedge fund performance are explored.

JEL Classifications: G11, G12, G14, C31

Keywords: hedge fund performance

*Nicolas P.B. Bollen (corresponding author), Owen Graduate School of Management, Vanderbilt University, nick.bollen@vanderbilt.edu; †Juha Joenväärä, Aalto University School of Business, juha.joenvaara@aalto.fi; ‡Mikko Kauppila, University of Oulu, mikko.kauppila@oulu.fi. The authors thank Antti Ilmanen, Olga Kolokolova, Peter Nyberg, Russ Wermers, and seminar participants at Boston College, Northeastern University, Vanderbilt University, the 2017 Chicago Alternative Investments Research Conference, the Office of Financial Research, the 2018 Midwest Finance Association Meeting, and the 10th Annual Hedge Fund and Private Equity Research Conference.

Electronic copy available at: https://ssrn.com/abstract=3034283

Hedge Fund Performance: End of an Era?

A growing chorus disparages the performance of hedge funds.[1] Some cite Warren Buffet's successful million-dollar bet in 2007 that the S&P 500 would earn higher returns over the next 10 years than a portfolio of funds selected by Protégé Partners.[2] Academic evidence is provided by Dichev and Yu (2011), who arrive at a similar conclusion when comparing hedge funds in aggregate to the S&P 500, Bali et al. (2013), who find that only two of 11 hedge fund indices are superior to the S&P 500 using utility-based performance metrics, and Sullivan (2021), who reports that the alpha of a broad hedge fund index flipped from positive to negative in the decade following the 2008 financial crisis. Given the roughly $5 trillion (Barth et al. (2021)) currently invested in hedge funds, the implications of the purported decline in their performance are critically important for the financial health of their investors including endowments and pension plans.

In this paper, we use data on individual funds to address three related questions. First, has aggregate hedge fund performance generally declined over time, as claimed by the popular press? Second, do any of the prediction models developed in the academic literature enable investors to pick a subset of funds that perform well out-of-sample? Third, what are the likely causes of any decline in aggregate hedge fund performance, and what does this suggest for hedge fund investment going forward?

We measure hedge fund performance using the union of six commercial databases of returns over two sub-periods demarcated by the December 2007/January 2008 breakpoint. This date coincides with the timing of Warren Buffet's bet and more importantly defines two sub-periods that both contain a full stock market cycle. An equally weighted hedge fund index features a cumulative return of 225% from 1997 through 2007, far outpacing an equally weighted stock and bond portfolio, which generates a cumulative return of 125%. In contrast, the hedge fund portfolio earns just 25% from 2008 through 2016, whereas the stock and bond portfolio returns 70% despite the large drawdown in equity markets during the financial crisis. Analyses of the cross-sectional distribution of individual hedge funds also indicate a marked decline in performance over time. The percentage of funds with positive and significant Fung and Hsieh (2004) alpha, for example, drops from about 20% to 10% over the sample, while the percentage with significantly negative

[1] See "A losing bet: Hedge funds haven't delivered on their promise," *Economist*, 7 May 2016.

[2] The index earned 7.1% annually versus 2.2% for the five hedge funds selected. See "Warren Buffet: Why index funds trump hedge funds," *Kiplinger*, 8 January 2018.

Electronic copy available at: https://ssrn.com/abstract=3034283

alpha increases from 5% to, at times, about 20%. In sum, we confirm reports that hedge fund performance has weakened considerably over the past decade.

We next determine whether investors can use any of a set of predictor variables, developed in prior work, to select subsets of hedge funds that perform satisfactorily out-of-sample despite the general decline over time. Our tests of performance predictability measure the realizable benefit of investing in hedge funds, taking into account real-world constraints and well-known database biases. Most importantly, we simulate random selection of 15 funds from the top quintile as ranked by the predictor variables and repeat the process 1,000 times. In this way, we can assess the probability that an investor who uses a particular predictor variable to select a realistic number of funds would earn returns superior to those generated by an alternative. We measure the benefit of an allocation to hedge funds by comparing the performance of an equally weighted portfolio consisting of the S&P 500 and the Vanguard Total Bond Market Index fund (VBTIX), hereafter the "stock/bond" portfolio, to that of a multi-asset class portfolio that includes a 20% allocation to hedge funds selected by the competing predictor variables.[3] Given the diversification potential of a hedge fund allocation, we compute two utility-based measures to assess the value added for risk-averse investors.

Over the full sample, two of the predictors generate a significant increase in the multi-asset class portfolio's Sharpe ratio relative to that of the stock/bond portfolio: the alpha from the Fung and Hsieh (2004) seven-factor benchmark and the Macro timing measure of Bali et al. (2014). In both cases, the improvement in Sharpe ratio is achieved by a substantial reduction in the volatility of the overall portfolio. Consequently, these predictors also significantly improve the utility achieved by a risk-averse investor when risk aversion is set sufficiently high. When we repeat the analysis over the two sub-periods, however, the improvement in Sharpe ratio is only present in the 1997–2007 sub-period. In the 2008–2016 sub-period, the hedge fund allocation generally reduces volatility but lowers average return as well, leaving the Sharpe ratio unchanged. This result indicates that an overall deterioration in hedge fund performance cannot be overcome by using the predictors.

[3] We have executed all analyses using other allocations within the stock/bond portfolio and obtained qualitatively identical results. Similarly, as discussed later in the paper, we demonstrate robustness to the allocation weight on hedge funds in the multi-asset class portfolio.

Electronic copy available at: https://ssrn.com/abstract=3034283

We perform a variety of robustness tests. First, we vary the allocation weights in the multi-asset class portfolio to see whether our results depend on the 20% allocation to hedge funds we use in the baseline analysis. Over the full sample, and both sub-periods, the optimal allocation to hedge funds is close to the 20% allocation we use in our main analysis, indicating that our results are robust to varying the allocation weights. Second, we examine the forecasting power of combinations of the predictors. To avoid data-snooping, we simply consolidate ranks of all predictors related to timing skill, all other predictors, and a combination of all seven predictors. We repeat all of the analyses described above. Over the full sample and both sub-periods, all combinations select funds that significantly lower overall portfolio volatility. All combinations result in portfolio Sharpe ratios higher than the passive benchmark in the 1997−2007 sub-period. None of the combinations, however, are able to pick funds for an allocation that raise the overall portfolio's Sharpe ratio over the 2008–2016 sub-period.

We conclude our study by testing a number of alternative explanations for the general decline in hedge fund performance. Two involve potential database problems. Backfilling superior returns when funds initially report to a database may inflate the earlier sub-period relative to the later sub-period. Conversely, when successful funds decide to stop reporting, future superior returns are censored from the database and this may deflate the later sub-period relative to the earlier sub-period. We test both effects and conclude they cannot explain our results.

Next, we consider three explanations related to economic mechanisms. First, anecdotal evidence in the financial press suggests that central bank interventions following the 2008 financial crisis have created distortions in asset markets, e.g., increased correlations across risky assets and muted volatility, which render many hedge fund strategies difficult to implement. Second, increased regulatory oversight from the 2010 Dodd–Frank reforms has imposed new compliance costs and potentially chilled some profitable hedge fund trading and reporting activity. (Cumming et al. (2017), Dimmock and Gerken (2016), and Honigsberg (2019)). Third, performance-sensitive capital flows and the publication of relevant research documenting successful hedge fund strategies may have eroded of alpha opportunities. (Fung et al. (2008) and Cao and Velthuis (2017)). These three possible economic explanations generate testable hypotheses related to the timing of the decline in hedge fund performance. We find that the structural breaks in the forecasting power of each predictor coincide with either the onset of the financial crisis or the enactment of the Dodd-

Electronic copy available at: https://ssrn.com/abstract=3034283

Frank regulation. This result suggests that the drop is due in part to increased regulation and changes in market characteristics associated with central bank stimulus activity.

To the extent that the impact of heightened regulatory scrutiny and central bank market intervention is likely to persist, investors might recalibrate expectations for future hedge fund performance downward. Does this mean they should also reduce hedge fund allocations? Our results indicate that especially risk averse investors can continue to justify a modest allocation to alternatives like hedge funds for their diversification benefit. Furthermore, like most studies, we measure performance and performance persistence at annual time-scales. Longer periods may be more appropriate for some strategies, such as investing around themes like technological disruption, climate change, and demographic trends. We leave analysis of performance measurement at longer horizons for future research.

1. Performance Prediction

This section provides details of our analytical framework for testing hedge fund performance predictability. In the context of the seminal work of Berk and Green (2004), the search for performance predictability might seem futile. In their model, some mutual fund managers have the ability to generate abnormal returns, thereby attracting capital from rational investors. Under the assumption of decreasing returns to scale, though, subsequent net-of-fee returns match those of a passive benchmark. Most empirical studies of mutual fund performance, including, for example, Carhart (1997), find little evidence of persistence, consistent with the main theoretical prediction of Berk and Green (2004). In contrast, prior research, including Kosowski et al. (2007) and Jagannathan et al. (2010), document performance persistence in hedge funds. There are at least two reasons why the Berk and Green (2004) model may not fully characterize hedge fund returns, both of which are rooted in differences in the institutional structure of mutual funds and hedge funds. First, the opacity of hedge fund trading strategies complicates learning about managerial ability.[4] Second, the performance fee paid to hedge fund managers aligns their interests with those of their

[4] Though the Fung and Hsieh (2004) seven-factor model has become the standard for estimating risk exposures and managerial performance, in most hedge funds more than half of the variation in returns is left unexplained. This makes establishing an appropriate benchmark difficult.

Electronic copy available at: https://ssrn.com/abstract=3034283

investors.[5] The Glode and Green (2011) model features both of these elements and demonstrates that hedge fund performance can persist in equilibrium.

A standard test for the ability to predict performance computes the values of predictor variables and records realized fund performance in non-overlapping periods. The researcher then determines whether values of the predictor variables in a ranking period are related to realized fund performance in a subsequent holding period by forming portfolios of funds by sorting on one or more of the predictor variables (we use quintile portfolios). The information content of a predictor is typically measured by the spread between the subsequent performance of the top and bottom quintiles.[6] This type of analysis overstates the benefit for investors because they cannot short the bottom quintile of funds; hence, we focus on the performance of only the top quintile.[7] The top quintile portfolios also overstate the investor experience for three reasons. First, an investor can typically invest in only a small number of funds rather than all funds in the top quintile, and so will typically experience higher risk due to less diversification across funds. Second, to the extent that top quintile performance is driven by a subset of superior funds, some of which may be closed to new investment, investors may earn lower average returns. Third, performance measured from returns reported to databases may be biased upwards. We explicitly address these concerns as described next.

To reflect constraints on the number of funds an investor can choose, we examine portfolios with 15 funds selected from a given quintile, motivated by a 2018 JP Morgan survey of institutional investors, including endowments, pension funds, insurance companies, and family offices, who hold on average somewhere between 15 and 20 hedge funds.[8] We simulate 1,000 draws of 15 funds and compute performance statistics for each iteration. Our simulation-based framework to assess the performance of an investment in hedge funds has two benefits. First, we can generate useful

[5] Agarwal et al. (2009) show a positive relationship between the dollar value of a manager's performance-based compensation and subsequent hedge fund returns.

[6] Alternatively, the relation between the ranking period predictor and the holding period performance can be measured in a cross-sectional regression. See, for example, Goetzmann and Ibbotson (1994).

[7] Investors who use the predictors to avoid the worst funds do in effect benefit, but they ultimately need to invest somewhere.

[8] See "2018 Institutional Investor Survey" JP Morgan Capital Advisory Group. In addition, evidence in Teo (2013) and Brown et al. (2012) indicates a portfolio of 15 funds achieves the majority of possible diversification benefit. Consistent with these results, we find, in unreported analysis, that funds in the 15-fund portfolios we create feature average correlation of 0.27. Furthermore, we find qualitatively similar results when increasing the portfolios to 30 funds.

Electronic copy available at: https://ssrn.com/abstract=3034283

insights by explicitly incorporating some of the constraints and frictions that affect an investor's realized performance. Second, the simulations provide a new and informative test for statistical significance: we compute not only the average performance of a hedge fund selection strategy but also its empirical *p*-value, based on the percentage of simulations in which it outperforms an alternative. This *p*-value has an intuitive interpretation as the probability that an investor benefits from a hedge fund allocation.[9]

For those funds that close to new investment at some point in the sample, we obtain the dates on which this occurs. In our analysis, we only select from those funds that are still open to new investment.

To reverse database biases, we adjust reported returns in two ways. Getmansky et al. (2004) show that, especially for funds with illiquid holdings, reported returns feature time series properties consistent with managerial conservatism in updating portfolio values, resulting in reported returns that likely have lower volatility than unobserved true portfolio returns. For this reason, we "de-smooth" returns following the procedure in Getmansky et al. (2004). In addition, as studied in Aiken et al. (2013), databases feature a type of censoring when funds stop reporting prior to periods of poor performance, resulting in upward bias. We follow Liang and Park (2010) and identify failure when a fund has a negative six-month average return and a negative 12-month change in AUM prior to the cessation of reporting. We employ a typical approach to reverse the resulting database bias by appending to a failing fund's return history one ad-hoc additional loss of 50%.

2. Data and Predictor Variables

As noted by Agarwal et al. (2009) and Joenväärä et al. (2019), only a minority of hedge funds report their returns to multiple databases, so we consolidate six major commercial databases (BarclayHedge, EurekaHedge, eVestment, HFR, Lipper TASS and Morningstar). This should not only increase the power of our tests, but also decrease the possibility that some documented results are driven by database-specific selection biases. We harvest variables common to all databases

[9] Jackwerth and Slavutskaya (2016) also conduct a simulation exercise in their study of the benefit of hypothetical allocations to hedge funds for a sample of UK pension funds. For each pension fund, they simulate a hedge fund allocation by selecting a single subset of hedge funds at random and measure the change in the pension fund's performance. In contrast, we repeat the random selection many times to construct a distribution of outcomes that permits a rich characterization of the investor experience.

Electronic copy available at: https://ssrn.com/abstract=3034283

(e.g., compensation structure, share restrictions, fund domicile, and investment style), eliminate duplicate share classes, and consolidate funds appearing in two or more databases by selecting the version with the longest time series. We use the global hedge fund universe and exclude funds of hedge funds. Our data consist of static fund characteristics, and monthly time series of USD-converted net-of-fees returns and assets under management (AUM) from January 1994 through December 2016.

We use May/June 2017 snapshots of the commercial databases, but to minimize the effect of strategic reporting delays (Aragon and Nanda (2017)), we only include observations until December 2016. To reduce survivorship bias, we start using data from EurekaHedge and eVestment in 2003 and 2010, respectively, and use data from other databases starting in 1994. These starting years eliminate suspiciously low annual attrition rates early in the database histories indicative of survivorship bias. To alleviate backfill bias, we remove the first 12 return observations of each fund. As mentioned, since we interested in studying whether an investor can use predictors to successfully select funds, we determine whether and when each fund becomes closed to new investment, and only permit new investment in funds that are still open. Finally, as in Joenväärä, Kosowski, and Tolonen (2019), to minimize the possibility that our results are driven by small funds that are available only to family and friends, we retain a fund's history starting from the time its AUM first reaches 20 million USD.[10] This leaves us with 12,173 funds, of which 4,431 (36.4%) are still reporting as of December 2016, and 7,742 (63.6%) have stopped reporting.

We replicate seven previously proposed predictors based on a fund's historical returns, capital flows, managerial compensation contract, and other relevant variables. Table 1 describes the predictor variables and their hypothesized performance implications.[11] We divide the measures into three categories: two skill categories (broad and timing skills) and one incentive category.

The first category consists of two broad skill measures: the *t*-value of the widely used alpha from the Fung and Hsieh (2004) seven-factor model ("FH alpha") and the strategy distinctiveness index ("SDI") from Sun et al. (2012). The two broad skill variables have an intuitive interpretation, so that higher values refer to higher skill. The second category consists of four timing skill measures: market timing ("Market timing") following Treynor and Mazuy (1966); volatility timing

[10] We show later in the paper that our results are robust to increasing this AUM requirement.

[11] Technical details are available from the authors.

Electronic copy available at: https://ssrn.com/abstract=3034283

("Volatility timing") following Chen and Liang (2007); liquidity timing ("Liquidity timing") following Cao et al. (2013); and macroeconomic timing ("Macro") following Bali et al. (2014). All of these measures are based on regressions, and we use the *t*-value of the relevant timing coefficient as the measure. Based on the existing literature, we expect all timing skill measures except Volatility timing to predict higher future performance; we expect lower Volatility loading to predict higher future performance, since a skillful manager should lower their market exposure during times of high volatility as in Busse (1999). The third category consists of an incentive measure that estimates the increase in a manager's compensation from increased fund returns: the option delta ("ΔOption") proposed by Agarwal et al. (2009) calculated on a dollar-per-dollar basis. We use the algorithm of Feng (2011) to track per-investor NAV and HWM at a monthly frequency; importantly, if a fund does not have a high-water mark provision, then we reset HWM to the NAV at each year-end. Based on agency theory, we expect the incentive measure to predict higher performance.

To ensure uniformity, we calculate all measures using a 24-month rolling window, except for the incentive measure, which is based on the full history until a given month.[12] Due to our 12-month backfill correction and 24-month data requirement, we study ex post returns over the period January 1997 through December 2016.

3. Results

3.1 Aggregate Hedge Fund Performance

This subsection addresses the paper's first main question: whether overall hedge fund performance has declined substantially over time as reported by the popular press. Figure 1 shows the cumulative return of the stock/bond portfolio, the cumulative return of an equally weighted portfolio of all hedge funds available each month, and the aggregate capital flow into hedge funds. Capital flow is defined as cumulative dollar flow as a percentage of beginning-of-period industry assets under management.

[12] There are two main reasons why we use a 24-month rolling window in estimating our predictors. First, a longer window would potentially introduce both reverse survivorship bias (Linnainmaa (2013)) and multiperiod sampling bias (Fung and Hsieh (2000)). Second, prior literature has documented that hedge funds' risk exposures varies from month to month (Bollen and Whaley (2009)) and even within months (Patton and Ramadorai (2013)). Hence, a longer window would likely result in larger estimation error for many of the predictors.

Electronic copy available at: https://ssrn.com/abstract=3034283

In the early years of the sample, the dot-com bubble drives the stock/bond portfolio to deliver higher returns than hedge funds as shown in Figure 1A. From 2000 onwards, however, the hedge funds far outpace the stock/bond portfolio generating about 225% total return vs. 125% over this period. Capital flow from investors rises along with cumulative returns: their parallel slopes suggest no evidence of industry-level decreasing returns to scale over this sub-period, which we test more formally in Section 4. Figure 1B illustrates a dramatic reversal over the latter period of January 2008 through December 2016. The two return series track each other quite closely during the financial crisis years, but starting from mid-2011 they diverge, with the hedge funds delivering about 25% total return vs. 70% for the stock/bond portfolio. In terms of raw returns, then, it does appear that the relative performance of hedge funds vis à vis standard assets has changed substantially over the past decade as reported widely by the popular press. Note also that capital flow is negative in the early part of the period, consistent with a flight to safety. Thereafter, capital flow is positive but orders of magnitude smaller than in Figure 1A again inconsistent with an industry-level decreasing returns to scale argument. Figure 2 the compares risk-adjusted performance of the two return series by computing their Sharpe ratios using rolling 24-month estimation windows. The same qualitative result holds: hedge funds were superior for much of the earlier half of the sample, the two series are similar during the financial crisis years, and the stock/bond portfolio significantly outperforms since about mid-2011.

Table 2 presents annualized summary statistics of the individual hedge funds in our sample along with the stock/bond portfolio and a portfolio consisting of a 20% allocation to all available hedge funds, 30% allocation to the S&P 500, and the remaining 50% invested in the VBTIX, with weights rebalanced annually. The difference between the stock/bond portfolio and the 20/30/50 portfolio measures whether an institutional investor is better off with or without a hedge fund allocation. Panel A shows results for the first sub-period of January 1997 through December 2007. The portfolio with a 20% hedge fund allocation generates slightly higher average return than the stock/bond portfolio, 7.9% versus 7.6%, and substantially lower standard deviation, 5.5% versus 7.4%. Consequently, the 20% hedge fund allocation raises the Sharpe ratio from 0.54 to 0.79. The superior performance of hedge funds over this period helped fuel the tremendous growth of the industry as illustrated in Figure 1A. The median individual hedge fund features both substantially higher average return of 9.7% and higher volatility of 9.9% than the stock/bond portfolio, indicating the benefit of an allocation to hedge funds on a risk-adjusted basis may only be achievable by

Electronic copy available at: https://ssrn.com/abstract=3034283

investing in a large portfolio of funds to diversify their risk. We determine whether a realistically sized portfolio of hedge funds can reliably deliver a benefit to investors in Subsection 3.2.

Note the wide dispersion in hedge fund performance: Sharpe ratios are 0.22 and 1.17 at the 25th and 75th percentiles, respectively, motivating the goal of predicting individual hedge fund performance. The stock/bond portfolio's Sharpe ratio of 0.54 is slightly lower than the median individual hedge fund. In contrast, the stock/bond portfolio's annualized Fung and Hsieh (2004) seven-factor alpha of 0.34% is far below the 4.36% generated by the median hedge fund. These conflicting results illustrate the difficulty in selecting the appropriate performance metric. One could interpret the 4.36% alpha as a risk-adjusted performance measure if one believes any variation unexplained by the seven-factor model is diversifiable. However, as shown by Bollen (2013), a substantial fraction of the unexplained variation appears to be common across funds and hence is likely the result of one or more omitted risk factors. Given this ambiguity, in subsequent analyses we continue to report the seven-factor alpha but focus attention on alternative measures of performance that do not require an explicit specification of risk factors, including the Sharpe ratio and utility-based certainty equivalents.

Panel B shows results using data from January 2008 through December 2016. Though the stock/bond portfolio features a 1.3% drop in average return and a slight increase in volatility, its Sharpe ratio rises substantially from 0.54 to 0.77 due to the dramatic reduction in interest rates in the post-financial crisis years. The 20% allocation to hedge funds now lowers the average return of the stock/bond portfolio from 6.3% to 5.2%. Though volatility is also reduced from 7.9% to 6.4%, the substantial drop in returns leaves the Sharpe ratio essentially the same with or without the hedge fund allocation. It is quite likely that an allocation to a smaller number of hedge funds would lower the Sharpe ratio below that of the stock/bond portfolio since there would be less diversification across the hedge funds. Even at the 75th percentile, individual hedge funds fail to match the Sharpe ratio of the stock/bond portfolio. The gap in performance has led to a reduction in the allocation to hedge funds by a number of institutional investors including CALPERs, as well as some compression in hedge fund fees.

Another way to assess hedge fund performance is to compare individual funds to an appropriate benchmark. Following Fung et al. (2008), we split funds into two subsets based on whether a fund's CAPM beta is statistically significant at the 10% level, measured over a fund's

Electronic copy available at: https://ssrn.com/abstract=3034283

fully history in the 1996–2016 period. In Table 3, we compare the "Have Beta" funds to the S&P 500 and the "Zero Beta" funds to three-month Libor.[13] In Panel A, the average "Have Beta" Sharpe ratio is 0.76 in the earlier period relative to the S&P 500's 0.44. In contrast, the S&P 500 delivers 0.51 in the later period in Panel B, compared to 0.26 for the average hedge fund. The median fund's average return drops from 10.6% to 3.1% over the two periods, whereas the S&P 500's average return drops much more modestly from 10.2% to 8.2%. Clearly, funds with equity exposure suffer substantially in the later period. The "Zero Beta" hedge funds also feature lower returns, with the average fund dropping from 7.8% to 3.7%. Other metrics for the "Zero Beta" funds are more stable, however: the excess return relative to three-month Libor for the average fund is 3.7% in the earlier period and 3.0% in the later period. The alpha of the median fund actually increased slightly from 2.78% to 3.18%. That said, institutional investors such as pension funds and University endowments set average return targets to meet their spending obligations, so the drop in average returns of the "Zero Beta" funds is still problematic.

As a final illustration of the general trend in hedge fund performance, we plot in Figure 3 the percentage of individual hedge funds with significant positive or significant negative alpha measured over a rolling 24-month window as assessed using a two-sided 10% significance level. Approximately 20% of funds deliver significant positive alpha until about 2008, when the fraction drops roughly in half. In contrast, about 5% of funds feature significant negative alpha until 2011, and since then the fraction bounces around but peaks above 20% several times. This result shows that in the more recent sub-period, a substantial number of individual funds significantly underperform at any point in time, again motivating the task of predicting hedge fund performance with the goal of avoiding this left tail. We turn next to this task.

3.2 Predicting Hedge Fund Performance

This subsection addresses the paper's second main question: whether prediction models developed in prior literature enable investors to select subsets of individual funds that subsequently outperform. If so, then some investors may still benefit from an allocation to hedge funds despite their general decline in performance described above.

[13] Note that the alpha of the S&P 500 is not measured since it is one of the Fung and Hsieh factors and that we treat the three-month Libor as a risk-free asset and so only report its average value.

Electronic copy available at: https://ssrn.com/abstract=3034283

3.2.1 Full Sample Evidence

We first determine the extent to which predictors can identify top-performing funds over the full sample. Since investors do not hold hedge funds in isolation, we frame the analysis of performance prediction in the context of a multi-asset class portfolio. We measure the benefit of an allocation to hedge funds by comparing the performance of the stock/bond portfolio to the performance of a portfolio in which the weights are 30% S&P 500, 50% the bond index fund VBTIX, and 20% hedge funds. We construct portfolio returns representing a buy and hold investment, i.e., weights evolve through time taking into account the return of each asset. Naturally, portfolio performance will depend on the specific allocation weights chosen by a given investor, so we test for robustness later in the paper. To shed more light on the benefit to a risk-averse investor, we compute two types of utility-based measures of performance.[14] The first is the MPPM of Goetzmann et al. (2007), which can be interpreted as the minimum incremental return that a risk-averse investor would accept to exchange a risky investment, e.g., a multi-asset class portfolio, for the risk-free asset.[15] For our implementation, we assume a risk-aversion level of 3, which, as argued by Goetzmann et al., is a realistic assumption in the sense that the CRSP VW index is historically optimal for risk-aversion levels between 2 and 4 depending on the time period. The second, derived from the framework in Fleming et al. (2001), can be interpreted as the maximum fee a risk-averse investor would pay in order to switch from one risky asset to another. For this reason, we label it the "Delta" method. The Delta is useful in our context to compare the portfolios with and without an allocation to hedge funds. For the Delta, we show results for three different risk aversion levels to illustrate how the benefit of hedge funds can vary across investors.

Table 4 summarizes the performance of portfolios with a 20% hedge fund allocation in which hedge funds are selected annually from the top quintile as ranked by each of the seven predictors. As mentioned, since investors cannot hold all funds in the top quintile, we report the average performance of 1,000 selections of 15 funds. This avoids overstating the diversification benefit of hedge funds. To gauge the information content of the predictors, we also draw 1,000 sets

[14] An additional reason for measuring performance by quantifying the utility of return series is to avoid potential factor model misspecification as noted in a performance prediction context by Carhart (1997). Suppose, for example, that a factor model omits a relevant risk factor. To the extent that exposure to the omitted risk factor is serially correlated, the alpha will be as well.

[15] As noted by Bali et al. (2013), a benefit of the MPPM over standard performance measures is that it incorporates non-normalities in hedge fund return distributions.

Electronic copy available at: https://ssrn.com/abstract=3034283

of 15 funds at random ("Random") from all funds. For average return, standard deviation, Sharpe ratio, alpha, and MPPM, Table 4 lists differences between the performance statistics of portfolios with a 20% allocation to hedge funds to those of the stock/bond portfolio with no hedge fund allocation. We test for significant differences using empirical *p*-values, i.e., the percentage of simulated portfolios with a hedge fund allocation that generate performance statistics above or below those of the stock/bond portfolio. We use two-sided significance levels, requiring that a simulated selection strategy deliver higher performance in at least 95% or at most 5% of the simulations to achieve a 10% significance level. In practice, an investor might use a looser definition of significance, but we adhere to common scientific levels. The Delta measure is the difference in certainty equivalents of the two portfolios by construction. Selection strategies are sorted by the difference in Sharpe ratios.

In all cases, the allocation to hedge funds reduces the average return, and for six of the fund selection strategies the difference is statistically significant. The random selection of funds, for example, reduces average return by 0.76% annually relative to the portfolio with no hedge fund allocation. Reductions in standard deviation are more substantial, however, and significant in all cases. The random selection of hedge funds results in an overall portfolio with a Sharpe ratio about equal to that of the portfolio with no hedge fund allocation, indicating that even naïve selection delivers a diversification benefit that offsets a reduction in average return. For the other sorts, the allocation to hedge funds also increases Sharpe ratio, albeit modestly. Fung and Hsieh (2004) alpha and Macro timing are the only predictors which significantly improve the Sharpe ratio, in both cases by 0.13, relative to the portfolio without a hedge fund allocation. This result indicates that some of the performance predictors are able to select subsets of funds that reliably improve the return characteristics of an investor's overall portfolio.

The MPPM of the portfolio with no hedge fund allocation is 4.03%, indicating investors with risk aversion of 3, the value we use for the MPPM, would be willing to pay a maximum of 4.03% to switch to it from the risk-free asset. None of the predictors are able to direct an allocation that provides a significantly higher MPPM. The sort on alpha, for example, generates an MPPM just about equal to that of the portfolio with no hedge fund allocation. That said, the evaluation of MPPM hinges on the assumed risk aversion level of 3. Investors with higher levels of risk aversion would value more highly the risk-reducing property of hedge funds. We illustrate this point using the Delta measures of incremental utility. None of the predictors generate overall portfolios that

Electronic copy available at: https://ssrn.com/abstract=3034283

are significantly preferred to the stock/bond portfolio at risk aversion levels of $\gamma = 1$ or $\gamma = 5$. For $\gamma = 10$, the hedge fund portfolios based on the alpha and Macro sorts feature Delta values of about 1%. This result indicates that an extremely risk-averse investor would be willing to pay an annual incremental fee of up to 1% to switch from the stock/bond portfolio to the portfolio with a 20% allocation to hedge funds drawn from the top quintiles from these predictors.

3.2.2 Time Variation in Performance Predictability

The results in Subsection 3.2.1 reflect performance over the full out-of-sample period of 1997–2016. A motivation for our study is to assess whether predictor variables can select successful funds given the relatively poor performance over the past decade, however, so we now split the sample into two sub-periods to test for temporal robustness.

Table 5 shows differences between the summary statistics of the stock/bond portfolio and the multi-asset class portfolio with a 20% allocation to hedge funds over the two sub-periods. Panels A and B list the predictors in order as determined by the difference in Sharpe ratios. Panel A shows results for the 1997–2007 period during which the stock/bond portfolio delivers a Sharpe ratio of 0.54. All of the selection strategies significantly raise the Sharpe ratio, in all cases by substantially lowering standard deviation. The top two predictors, Macro timing and Fung and Hsieh (2004) alpha, for example, both result in an overall portfolio average return indistinguishable from that of the stock/bond portfolio but with standard deviation roughly 2% lower. Note that risk reduction is a general feature of hedge funds over the earlier period since the random selection of funds also reduces overall portfolio volatility significantly.

Panel B shows results for the 2008–2016 period. Here again the hedge fund allocations reduce standard deviation significantly, illustrating the robustness of top quintile hedge funds as diversifying tools. However, now the hedge fund allocations substantially reduce average returns, which offsets the benefit of lower standard deviation, so that in no case does the hedge fund allocation significantly increase the Sharpe ratio. Differences between the Sharpe ratio of the stock/bond portfolio and the multi-asset class portfolios are insignificant, indicating that the tradeoff between reduced average return and standard deviation is acceptable using this measure.

Our simulated selection strategies require many decisions regarding which funds are in the acceptable set each year. We test for robustness to varying these decisions by focusing on the Fung and Hsieh (2004) alpha predictor, as it works best over the full sample and is arguably the most

Electronic copy available at: https://ssrn.com/abstract=3034283

widely used performance measure. In Table 6, we report how performance of the multi-asset portfolio varies when (a) restricting investment to only those funds with statistically significant alpha in the ranking period, and (b) restricting investment to only those funds with AUM larger than our base case 20 million USD. The former is an example of a variation in methodology and the latter is a practical concern related to the capacity of funds to accommodate the needs of larger institutional investors. We also list the original results from the sort on alpha in Tables 4 and 5 for comparison. Panels A through C show that over the full sample and both sub-periods, restricting funds to those with ranking period alpha significant at the 5% level modestly improves performance, suggesting that the extra precision helps identify more persistent skill. In all cases, though, the qualitative results are the same across all restrictions, indicating that our results are robust to the significance and AUM filters.[16]

To provide more general insight regarding the ability of the predictors, we measure the performance of portfolios with hedge fund allocations directed by combinations of the individual predictor variables. Combinations we consider include all timing predictors (Liquidity, Macro, Market, and Volatility), all non-timing predictors (Alpha, SDI, Option Delta), and all seven predictors together. To combine predictors, we compute a fund's fractional rank, based on a cross-sectional sort on a given measure, and then average a fund's fractional ranks across the predictors used. There are many other combinations one could consider, however, as cautioned by Foster et al. (1997) and Harvey et al. (2016), a more comprehensive search over possible combinations raises the risk of data-snooping.[17]

Table 7 shows the results sorted by the resulting Sharpe ratios. Panel A reflects the full sample, whereas Panels B and C show the two sub-periods. In all cases, the non-timing predictors perform the best. Over the full sample, the combination of non-timing predictors directs a hedge fund allocation that lowers average returns by 72 basis points annually and standard deviation by over 2%, resulting in a Sharpe ratio that is 0.13 higher than the passive benchmark. Comparing

[16] In response to a referee question, we also considered accommodating binding capacity constraints by measuring the benefit of investing in top-quintile actively managed equity mutual funds, as ranked by Carhart (1997) alpha, instead of hedge funds. We find the same qualitative results with actively managed mutual funds: a modest benefit in the earlier period (though not as large as with hedge funds) but no benefit in the later period. This is consistent with results in Choi and Zhao (2021). In the interest of brevity, we do not report the results, but they are available on request.

[17] In unreported analysis, we considered several other combinations of predictors, including composite rankings from the top individual predictors in the first half of the sample, which are tested out of sample in the second half. Results are qualitatively unchanged.

Electronic copy available at: https://ssrn.com/abstract=3034283

results in Panels B and C, it is clear the hedge fund allocation only increases the Sharpe ratio in the earlier 1997–2007 sub-period. Note, however, that in both sub-periods, the allocation to hedge funds significantly lowers portfolio standard deviation.

The results above show that a 20% allocation to hedge funds selected from the top quintile as ranked by all of the predictors provides investors with a significant risk reduction vehicle. Portfolio volatility is lower with the allocation over the full 1997–2016 period as well as both sub-periods. The diversification benefit of these hedge fund allocations is notable given evidence of increasing correlations among standard asset classes as studied by Cotter et al. (2018). In this context, the risk reduction provided by a hedge fund allocation is especially valuable to investors, even with the accompanying drop in overall portfolio average returns. We show that diversification benefit is a general feature of hedge funds as it is achievable through random selection.

The Fung and Hsieh (2004) alpha and Macro timing predictors provide valuable incremental information by helping investors to reliably choose funds for a 20% allocation that raises a multi-asset class portfolio's Sharpe ratio over the full sample. In sub-sample analysis, however, we show that the benefit is only present in the 1997–2007 period, and that no predictor can boost portfolio performance relative to the stock/bond portfolio over 2008–2016.

3.2.3 Robustness to Varying Allocation Weights

Differences across investors in risk aversion, investment horizon, wealth levels, and other parameters will affect optimal allocation weights across the assets in their portfolios. We test for robustness in our portfolio analysis by varying hedge fund allocations from 0% to 50%, with stock and bond allocations selected each year to maximize the overall portfolio Sharpe ratio. Optimization is based on past sample moments estimated using expanding windows for means and rolling 36-month windows for variances and correlations, thus, the out-of-sample period is from January 2000 through December 2016.[18]

Table 8 reports average performance measures when selecting hedge funds from the top quintile as sorted by Fung and Hsieh (2004) alpha.[19] For the first three columns in each panel, we

[18] To improve the estimates of risk premia for stocks (S&P 500) and bonds (VBTIX), we estimate them using monthly returns starting from February 1972. The bond returns are spliced such that starting from October 1995 we use the VBTIX index, and before October 1995 we use its benchmark index (Bloomberg-Barclays US Aggregate Bond Index).

[19] Results using other predictors are available from the authors on request.

Electronic copy available at: https://ssrn.com/abstract=3034283

test for significant differences between the performance of the 0% hedge fund allocation and the others. For the fourth column, we test whether alpha is significantly different from zero. Panel A1 shows results for the full sample. The 25% hedge fund allocation generates the highest Sharpe ratio of 0.96 compared to 0.88 for the 0% allocation, though the difference is not significant using the empirical *p*-value. Panel B1 shows results using random hedge fund selection. Here the Sharpe ratio is monotonically decreasing in the hedge fund allocation. This result is consistent with Dichev and Yu (2011) and Bali et al. (2013), who document that hedge funds generally have underperformed standard assets, and demonstrates that the alpha criterion has valuable information about subsequent hedge fund performance since it enables an investor to avoid selecting funds that decrease performance. In Panel A2, the 12.5% hedge fund allocation generates a Sharpe ratio of 1.04 over the 2000–2007 period and this is significantly higher than the 0.87 of the 0% allocation. Panel B2 shows that random selection of hedge funds provides no significant benefit. Panel A3 shows that, over the 2008–2016 period, portfolios with hedge fund allocations between 12.5% and 25% generate Sharpe ratios almost identical to that of the 0% allocation. In contrast, in Panel B3, random selection of hedge funds significantly lowers the average Sharpe ratio for allocations 25% and higher. This result indicates that the alpha criterion does retain its information content over both sub-periods, but that hedge fund performance over the latter sub-period is insufficient to improve the performance of a stock/bond portfolio. In addition, the results in Table 8 justify our use of a 20% allocation to hedge funds in the other analyses.

4. Causes of the Decline in Hedge Fund Performance

Our analysis indicates that an allocation to top-quintile hedge funds as selected by all predictors would have significantly increased the performance of a multi-asset class portfolio relative to a stock/bond portfolio over the 1997–2007 sub-period, but provided no benefit over the 2008–2016 sub-period. Summary statistics in Table 2 preview this result, wherein the average individual hedge fund features a substantially higher Sharpe ratio than the stock/bond portfolio in the earlier sub-period, but a substantially lower Sharpe ratio thereafter.

What might explain the pronounced decline of overall hedge fund performance? This Section considers a variety of database-related and economic explanations.

Electronic copy available at: https://ssrn.com/abstract=3034283

4.1 Database-related Explanations

Two aspects of the available self-reported hedge fund return data may cause a spurious decline in observed aggregate performance. Jorion and Schwarz (2019) and Joenväärä et al. (2019) show that backfill may become more pronounced over time as more funds initiate reporting and import a return history. Performance estimated in earlier years will then become progressively more biased upwards. As mentioned in Section 2, we address backfill bias by dropping the first 12 months of a fund's history. We test whether a more rigorous correction affects our results by implementing the procedure proposed by Jorion and Schwarz: dropping all observations prior to the date a fund is added to a database. In unreported analysis, we find that aggregate hedge fund performance is somewhat weaker in both the 1997–2007 and 2008–2016 sub-periods after using the alternate correction for backfill. However, the difference between the two sub-periods is almost identical to our original results, indicating that backfill bias is not the cause of the decline.

Backfill is a form of selection bias: managers tend only to import histories with good performance, effectively censoring weak returns from the database resulting in an upward bias. Conversely, some managers with superior performance opt to cease reporting as they reach capacity and no longer need to advertise their track record. Consequently, these managers are censoring high returns in the more recent years of a database, resulting in a downward bias. Taken together, these two managerial decisions could cause the observed decline in performance over time.

To assess whether this mechanism could cause a decline in aggregate hedge fund performance, we examine all cessations of fund histories prior to the end of our data. We then determine the frequency with which these funds are superior, defined as those with positive and significant changes in Fung and Hsieh (2004) alpha and AUM over the prior 24 months.[20] This decision rule is analogous to the Liang and Park (2010) rule we use to determine whether a fund that ceases reporting failed and for which we append a negative 50% return as described in Section 1. In unreported analysis, we find that the average annual percentage of funds that fail increases from 3% to 5.4% over the two sub-periods, whereas the percentage of funds that cease reporting with superior performance declines from 0.7% to 0.4%. This result suggests that the censoring of

[20] We used a number of alternative performance criteria and obtained similar results.

Electronic copy available at: https://ssrn.com/abstract=3034283

superior performance is likely to be less of a factor in the later period and so cannot explain the decline we document.

As a caveat, some managers may opt never to report to a database in the first place, and this could affect our assessment of aggregate performance. Existing studies provide mixed evidence regarding the relationship between the decision to never report and performance, however, so it is unclear whether this phenomenon could affect our inference. Aiken et al. (2013) identify a set of individual funds that never report to commercial databases by studying a sample of registered funds of funds that must report their holdings. They find that the funds that never report feature average performance inferior to reporting funds. Agarwal et al. (2013) construct a sample of non-reporting funds by identifying a set of hedge funds that disclose their equity holdings in SEC Form 13-F and which never appear in commercial databases. They find no significant difference between the performance of non-reporting funds and other funds. Edelman et al. (2013) hand collect data on the largest hedge fund firms and conclude that the censoring of non-reporting good and bad performers neutralizes any bias. In contrast, Barth et al. (2021) find that non-reporting hedge funds, identified by studying filers of the new SEC Form PF, significantly outperform funds in the commercial databases.

4.2 Economic Explanations

At least three economic explanations for a decline in hedge fund performance are possible. First, the more recent period is characterized by a historic period of intervention by central banks around the world that some believe has resulted in increased correlation and depressed levels of volatility in many asset markets.[21] Increased correlation is often linked to synchronized entry to and exit from risky assets, i.e., "risk-on risk-off," in response to positive and negative signals, respectively, regarding stimulus actions. To illustrate, Figure 4 includes two time-series plots of U.S. equity market correlation over the 1997–2016 sample period. Figure 4A shows the average correlation of all common stocks in the CRSP database estimated using daily returns over the prior six months. The dashed lines correspond to the averages over our two sub-periods. The stock-level average correlation doubled from 0.15 to 0.30. Figure 4B shows corresponding results for ten S&P

[21] See for example "Lockstep stock moves lock up investors" 12/27/12 and "Hedge fund's assets fall by 95%" 6/28/15 in the Wall Street Journal.

Electronic copy available at: https://ssrn.com/abstract=3034283

industry indexes: the average increased by about 25% from 0.54 to 0.69. In both cases, the increases are highly significant.

We measure correlation across risky assets more broadly by computing the correlation between the returns of the passive benchmark and those of the top and bottom quintiles, respectively, as sorted by the Fung and Hsieh (2004) alpha. We use a 24-month rolling window and plot the two time series in Figure 5. Over the two sub-periods defined using the December 2007/January 2008 breakpoint, the average correlation of the top quintile portfolio and the passive benchmark increased from 0.55 to 0.61. More prominently, there is a five-year stretch of elevated correlations running from mid-2008 through mid-2013. Here the top-quintile features a correlation with the passive benchmark of 0.70 and above. Market commentators describe how economic conditions post-financial crisis have hampered the ability of hedge fund managers to sustain prior levels of performance. High correlations, for example, make security selection and long-short strategies more difficult to execute as many assets are affected contemporaneously by stimulus activity.[22]

Second, following the financial crisis and a number of well-known cases of managerial misconduct in the hedge fund industry, regulators have increased scrutiny, culminating in the 2010 passage of the Dodd–Frank Act. Cumming et al. (2017) argue that the additional regulation has had a negative impact on fund performance due to the increased cost of compliance. They report a median cost of $150,000 annually per fund, however, which would have a negligible direct impact on fund alpha. Additional regulation can affect observed fund performance in other ways. Dimmock and Gerken (2016) and Honigsberg (2019) show that various measures of misreporting decline after increases in regulation and this could worsen observed performance. If fund managers smooth returns less intensively, for example, then reported volatility would increase and Sharpe ratios would decrease. Additionally, the increased regulation could have a chilling effect on insider trading and other trading violations, which would reduce actual fund performance.

Third, increased competition among hedge fund managers could erode fund level performance. One channel for this effect is described in the literature on decreasing returns to scale in hedge fund management, as in Aragon et al. (2014), who study differences in regulation

[22] Increased correlation across risky assets may explain why Sullivan (2021) finds a decline in the standard error of alpha in an equity hedge fund index, which he labels "active risk," since the returns of funds' equity allocations are more fully explained by systematic risk.

Electronic copy available at: https://ssrn.com/abstract=3034283

governing onshore and offshore funds. Only offshore funds are permitted to advertise, and this results in fund flow that is more sensitive to performance, driving erosion of competitive advantage and weaker performance over time than onshore funds. A second channel through which competition can erode performance is the publication of research describing successful strategies, as studied by McLean and Pontiff (2016). The resulting copycat trading could eliminate the sources of abnormal returns.

We begin our investigation of these explanations by testing whether hedge funds in our sample feature performance that declines with fund size, consistent with the decreasing returns to scale assumption of Berk and Green (2004) and Pástor and Stambaugh (2012). Prior research finds supportive evidence, including Fung et al. (2008) and Cao and Velthuis (2017). Our test design is a panel regression using the recursive demeaning procedure developed by Pástor et al. (2015), who show it is necessary to correct the bias that otherwise would arise due to a mechanical contemporaneous correlation between changes in fund size and unexpected fund returns. The dependent variable on date t is fund abnormal return defined by the Fung and Hsieh (2004) factor model using all available fund returns up to date t. The independent variables are the natural log of fund size in millions of AUM at date $t-1$ as well as lagged industry size, measured as aggregate hedge fund AUM at date $t-1$ scaled by global equity market capitalization. Table 9 lists coefficient estimates and corresponding t-statistics. Panel A shows results for all funds. The coefficient on lagged industry size is significant and roughly –0.28, meaning for every 1% increase in relative industry size fund monthly alpha drops by 28 basis points. Panel B shows results for equity-oriented funds only. The coefficient on lagged industry size is about three times as large, though the significance is somewhat weaker. In neither case is the coefficient on lagged fund size significant. Thus, consistent with the results in Cao and Velthuis (2017), we find evidence of decreasing returns to scale at the industry level, which may partly explain the decline in hedge fund performance in our sample.

The other economic explanations described above generate testable hypotheses related to the date when performance declines. The explanation based on central bank intervention implies the decline should occur shortly after November 25, 2008, when the Federal Reserve announced QE1, and should be common for all predictors. The explanation based on regulatory reform also implies a common date for all predictors but sometime after the enactment of the Dodd–Frank reforms on July 21, 2010. The explanation based on academic research implies the decline of the

Electronic copy available at: https://ssrn.com/abstract=3034283

success of an individual predictor should occur soon after the relevant academic study were made public. In an effort to judge which of these explanations is most important, therefore, we determine for each predictor the date that best demarcates the sample into two distinct periods of performance. If we find a common date near the end of 2008, then the central bank intervention explanation would be supported. If we find a common date sometime after the middle of 2010, then the regulatory reform explanation would be supported. If we find dates that vary across the predictors and correspond to the publication dates of relevant academic research, then the publication explanation would be supported.

To proceed, we first record returns each month of top quintile funds as ranked based on the past 24 months using each of the predictors.[23] Next, following Bollen and Whaley (2009), we determine the optimal switch point in a regression framework by searching over all possible switch dates and selecting the one which maximizes the regression adjusted R^2. We conduct the analysis two ways for robustness. First, we simply regress returns on an intercept and an indicator variable that equals one after the switch point – here the coefficient on the indicator variable measures changes in average return. Second, we include the Fung and Hsieh (2004) seven factors – here the coefficient on the indicator variable measures changes in alpha. In almost all cases, the coefficient on the post-switch indicator variable is negative and statistically significant as expected. More importantly, the switch points appear to cluster quite closely in time. Figure 6 plots the switch points for the change in Fung and Hsieh (2004) alpha along with the dates on which predictors associated with academic research were posted to SSRN or published in the case of Fung and Hsieh (2004). The results for switch points of average returns are qualitatively identical and omitted for brevity. The performance of five of the seven predictors were best defined by a switch point in early 2008. The other two predictors were best defined by May 2011. This degree of clustering suggests that the explanation based on the publication of academic research is likely not relevant. Further, since the two clustering dates correspond to the central bank intervention and regulation explanations, respectively, these both are supported by the analysis.

In sum, our analyses in this subsection provide mixed support for a decreasing returns to scale explanation for the decline in hedge fund performance and no support for a publication effect.

---

[23] We have also conducted the analysis using the spread between top and bottom quintiles and obtained qualitatively similar results.

Electronic copy available at: https://ssrn.com/abstract=3034283

We do find some evidence consistent with the notion that heightened regulatory oversight has hurt performance, which constitutes a long-lasting effect. Our findings most strongly support the hypothesis that central bank intervention has affected markets in ways that have impeded the ability of hedge fund managers to sustain their performance in the first half of our sample. We leave further study of the macro-determinants of aggregate hedge fund performance for future work.

5. Summary

Judged by historical standards, passive investments have generated an extremely high reward-to-risk ratio post-financial crisis. Some investors in high-fee alternative assets including hedge funds and private equity have been disappointed by the performance of their funds by comparison and have undertaken fundamental changes to their allocations.[24] These events motivate our study of hedge fund performance. We address three main questions. First, has the performance of hedge funds in aggregate declined over the past decade as reported in the popular press? Second, can investors use any of the large number of predictor variables drawn from existing research to identify subsets of funds that reliably outperform? Third, what is the most likely explanation for any aggregate decline in hedge fund performance, and what does it imply about future investment allocations?

Using a broad sample of hedge funds consolidated from six commercial databases, we verify a substantial drop in performance relative to standard asset classes when splitting our sample using a December 2007/January 2008 breakpoint. The decline in performance is notable both in a value weighted index of all hedge funds each month, as well in the cross-sectional distribution of individual hedge funds.

We test whether investors can improve the performance of their portfolio by using predictor variables to select a subset of funds for a 20% allocation to hedge funds. We measure the realizable benefit of an investment in hedge funds by simulating an investment program that randomly selects 15 funds from the top quintile as ranked by each of the predictors. In addition, we control for biases in hedge fund return data by de-smoothing reported returns and applying a delisting adjustment to funds that stop reporting. Over the full 1997–2016 sample, almost all predictors result in a multi-

[24] See "Will Public Pensions Regret Dumping Hedge Funds?" Institutional Investor, 15 February 2017.

Electronic copy available at: https://ssrn.com/abstract=3034283

asset class portfolio with significantly lower volatility than that of a stock/bond portfolio with no hedge fund allocation. Furthermore, seven of the predictors improve the Sharpe ratio. A random selection of funds also reduces portfolio volatility but does not increase the Sharpe ratio, indicating that a large number of the predictors add value.

The diversification benefit of hedge funds is present in two subsamples of the data, 1997–2007 and 2008–2016. However, in the latter period, a significant drop in average return accompanies the risk reduction, so that the Sharpe ratios of the resulting portfolios are no different from that of a stock/bond portfolio with no hedge fund allocation. While investors with above-average risk aversion may be able to generate incremental utility with a hedge fund allocation, especially when selected using the alpha and Macro timing predictors, the benefit is due solely to a diversification benefit that typically comes at a cost of lower average returns. We conduct numerous robustness tests and in no case do we overturn the result that hedge fund performance has been a significant drag on portfolio performance in the 2008–2016 period.

We shed light on reasons for the decline in the benefit offered by hedge funds. We find some evidence of decreasing returns to scale at the industry level. We also identify the dates which best demarcate the decline in performance of funds in the top quintile as ranked by each of the predictors. The optimal switch points cluster during the depths of the 2008–2009 financial crisis as well as the passage of the Dodd-Frank reforms. This result is consistent with another explanation for the decline in hedge fund performance: an economic environment inhospitable to many hedge fund strategies driven by central bank interventions and the impact of heightened regulatory scrutiny. To the extent that these effects are likely to persist, hedge funds in aggregate may not be able to achieve the same level of success going forward that fueled their rise in the mid 1990's to the mid 2000's. However, since we find that hedge funds constitute a reliable diversifying vehicle throughout our sample, more risk averse investors can continue to justify a modest allocation to alternatives like hedge funds for their diversification benefit.

Electronic copy available at: https://ssrn.com/abstract=3034283

References

Agarwal, V., N. Daniel, and N. Naik (2009). Role of managerial incentives and discretion in hedge fund performance. The Journal of Finance 64, 2221-2256.

Agarwal, V., V. Fos, and W. Jiang (2013). Inferring reporting biases in hedge fund databases from hedge fund equity holdings. Management Science 59, 1271-1289.

Aiken, A., C. Clifford, and J. Ellis (2013). Out of the dark: Hedge fund reporting biases and commercial databases. Review of Financial Studies 26, 208-243.

Aragon, G., B. Liang, and H. Park (2014). Onshore and offshore hedge funds: Are they twins? Management Science 60, 74-91.

Aragon, G. and V. Nanda (2017). Strategic delays and clustering in hedge fund reported returns. Journal of Financial and Quantitative Analysis 52, 1-35.

Bali, T., S. Brown, and K. Demirtas (2013). Do hedge funds outperform stocks and bonds? Management Science 59, 1887-1903.

Bali, T., S. Brown, and M. Caglayan (2014). Macroeconomic risk and hedge fund returns. Journal Financial Economics 114, 1-19.

Barth, D., J. Joenväärä, M. Kauppila, and R. Wermers (2021). The hedge fund industry is bigger (and has performed better) than you think. Working paper, Office of Financial Research.

Berk, J. and R. Green (2004). Mutual fund flows and performance in rational markets. Journal of Political Economy 112, 1269-1295.

Bollen, N. (2013). Zero-R^2 hedge funds and market neutrality. Journal of Financial and Quantitative Analysis 48, 519-547.

Bollen, N. and R. Whaley (2009). Hedge fund risk dynamics: Implications for performance appraisal. Journal of Finance 64, 985-1035.

Brown, S., G. Gregoriou, and R. Pascalau (2012). Diversification in funds of hedge funds: Is it possible to overdiversify? The Review of Asset Pricing Studies 2, 89-110.

Busse, J. (1999). Volatility timing in mutual funds: Evidence from daily returns. Review of Financial Studies 12, 1009-1041.

Cao, C. and R. Velthuis (2017). Are hedge fund capacity constraints binding? Evidence on scale and competition. Working paper, Pennsylvania State University.

Cao, C., Y. Chen, B. Liang, and A. Lo (2013). Can hedge funds time market liquidity? Journal of Financial Economics 109, 493-516.

Carhart, M. (1997). On persistence in mutual fund performance. The Journal of Finance 52, 57-82.

Cotter, J., S. Gabriel, and R. Roll (2018). Nowhere to run, nowhere to hide: Diversification in a flat world. Working paper, University College Dublin.

Choi, J., and K. Zhao (2021). Carhart (1997) Mutual Fund Performance Persistence Disappears Out of Sample. Critical Finance Review, forthcoming.

Cumming, D., N. Dai, and S. Johan (2017). Dodd-Franking the hedge funds. Journal of Banking & Finance, forthcoming.

Chen, Y. and B. Liang (2007). Do market timing hedge funds time the market? Journal of Financial and Quantitative Analysis 42 (4), 827-856.

Dichev, I. and G. Yu (2011). Higher risk, lower returns: What hedge fund investors really earn. Journal of Financial Economics 100, 248-263.

Dimmock, S., and W. Gerken (2016). Regulatory oversight and return misreporting by hedge funds. Review of Finance 20, 795-821.

Edelman, D., W. Fung, and D. Hsieh (2013). Exploring uncharted territories of the hedge fund Industry: Empirical characteristics of mega hedge fund firms. Journal of Financial Economics 109, 734-758.

Electronic copy available at: https://ssrn.com/abstract=3034283

Feng, S. (2011). A Comparison of hedge fund gross and net performance. Working Paper, Edinboro University.
Fleming, J., C. Kirby, and B. Ostdiek (2001). The economic value of volatility timing. The Journal of Finance 56, 329-352.
Foster, F., T. Smith, and R. Whaley (1997). Assessing goodness-of-fit of asset pricing models: The distribution of the maximal R^2. The Journal of Finance 52, 591-607.
Fung, W. and D. Hsieh (2000). Performance characteristics of hedge funds and commodity funds: Natural vs. spurious biases, Journal of Financial and Quantitative Analysis 35, 291-307.
Fung, W. and D. Hsieh (2004). Hedge fund benchmarks: A risk-based approach. Financial Analysts Journal, 65-80.
Fung, W., D. Hsieh, N. Naik, and T. Ramadorai (2008). Hedge funds: Performance, risk, and capital formation. Journal of Finance 63, 1777-1803.
Getmansky, M., A. Lo, and I. Makarov (2004). An econometric analysis of serial correlation and illiquidity in hedge-fund returns. Journal of Financial Economics 74, 529-609.
Glode, V. and R. Green (2011). Information spillovers and performance persistence for hedge funds. Journal of Financial Economics 101, 1-17.
Goetzmann, W., and R. Ibbotson (1994). Do winners repeat? Journal of Portfolio Management (Winter) 9-18.
Goetzmann, W., J. Ingersoll, M. Spiegel, and I. Welch (2007). Portfolio performance manipulation and manipulation-proof performance measures. Review of Financial Studies 20, 1503-1546.
Harvey, C. R., Y. Liu, and H. Zhu (2016). … and the cross-section of expected returns. Review of Financial Studies 29 (1), 5-68.
Honigsberg, C. (2019). Hedge fund regulation and fund governance: Evidence on the effects of mandatory disclosure rules. Journal of Accounting Research 57, 845-888.
Jackwerth, J. and I. Slavutskaya (2016). The total benefit of alternative assets to pension fund portfolios. Journal of Financial Markets 31, 25-42.
Jagannathan, R., A. Malakhov, and D. Novikov (2010). Do hot hands exist among hedge fund managers? An empirical evaluation. The Journal of Finance 65, 217-255.
Joenväärä, J., M. Kauppila, R. Kosowski, and P. Tolonen (2019). Hedge fund performance: What do we know? Critical Finance Review, forthcoming.
Joenväärä, J., R. Kosowski, and P. Tolonen (2019). The effect of investment constraints on hedge fund investor returns. Journal of Financial and Quantitative Analysis 54, 1539-1571.
Jorion, P., and C. Schwarz (2019). The fix is in: Properly backing out backfill bias. Review of Financial Studies 32, 5048-5099.
Kosowski, R., N. Naik, and M. Teo (2007). Do hedge funds deliver alpha? A Bayesian and bootstrap analysis. Journal of Financial Economics 84, 229-264.
Liang, B. and H. Park (2010). Predicting hedge fund failure: A comparison of risk measures. Journal of Financial and Quantitative Analysis 45, 199–222.
Linnainmaa, J. (2013). Reverse survivorship bias. Journal of Finance 68, 789-813.
McLean, R. and J. Pontiff (2016). Does academic publication destroy stock return predictability? Journal of Finance 71, 5-32.
Pástor, L. and R. Stambaugh (2012). On the size of the active management industry. Journal of Political Economy 120, 740-781.
Pástor, L., R. Stambaugh, and L. Taylor (2015). Scale and skill in active management. Journal of Financial Economics 116, 23-45.

Electronic copy available at: https://ssrn.com/abstract=3034283

Patton, A. and T. Ramadorai (2013). On the high-frequency dynamics of hedge fund risk exposures, Journal of Finance 68, 597-635.
Sullivan, R. (2021). "Hedge fund alpha: Cycle or sunset? Journal of Alternative Investments, forthcoming.
Sun, Z., A. Wang, and L. Zheng (2012). The road less traveled: Strategy distinctiveness and hedge fund performance. Review of Financial Studies 25, 96-143.
Teo, M. (2013). Diversification in hedge fund portfolios. BNP Paribus hedge Fund Centre, Hedge Fund Insights.
Treynor, J. and K. Mazuy (1966). Can mutual funds outguess the market? Harvard Business Review 44, 131-136.

Electronic copy available at: https://ssrn.com/abstract=3034283

Figure 1. Cumulative Returns and Hedge Fund Capital Flow

Figures 1A and 1B show the cumulative return of a stock/bond portfolio and an aggregate hedge fund portfolio, as well as aggregate capital flow into hedge funds, over two time periods. The stock/bond portfolio consists of 50% S&P 500 and 50% Vanguard Total Bond Market Index (VBTIX), rebalanced annually. The hedge fund portfolio is equally weighted each month. Hedge fund capital flow is defined as cumulative dollar flow into all hedge funds in the sample as a percentage of beginning-of-period assets under management.

Figure 1A. 1/1997–12/2007

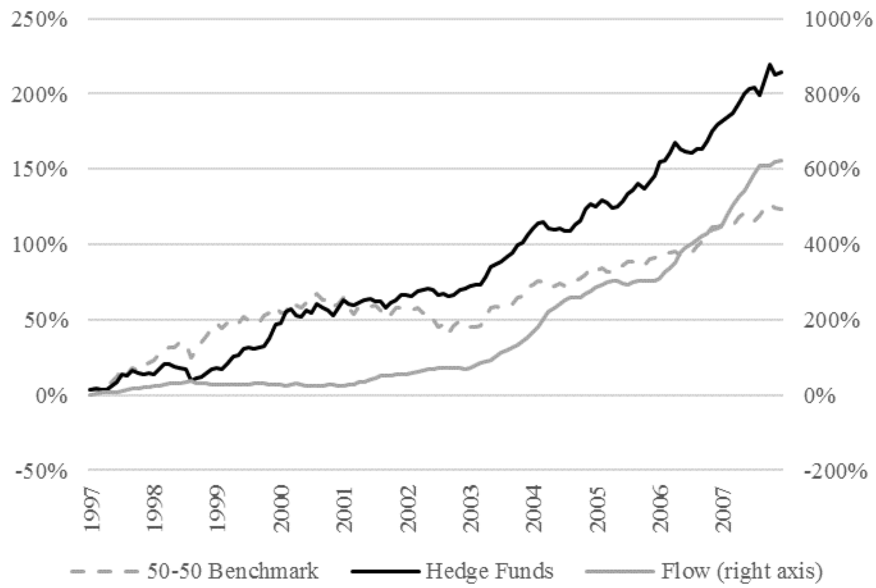
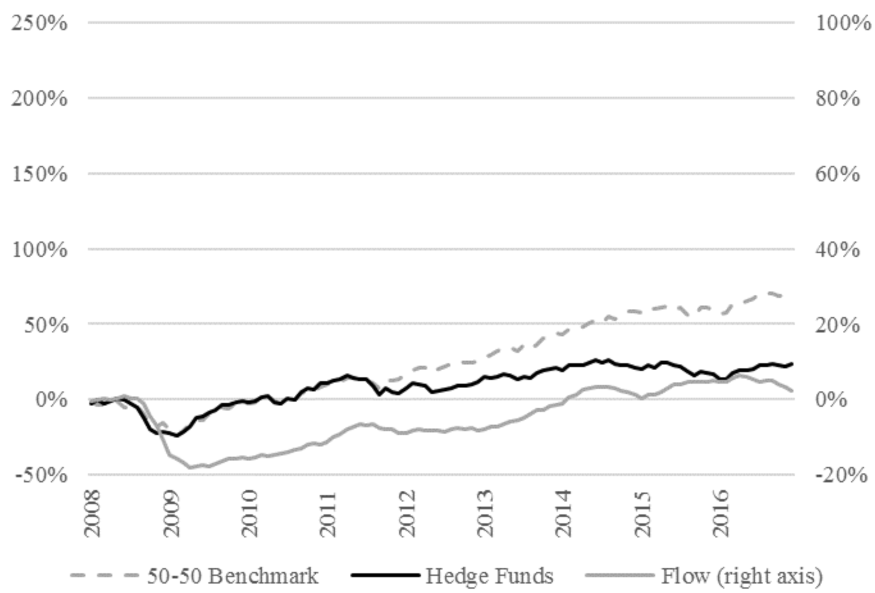


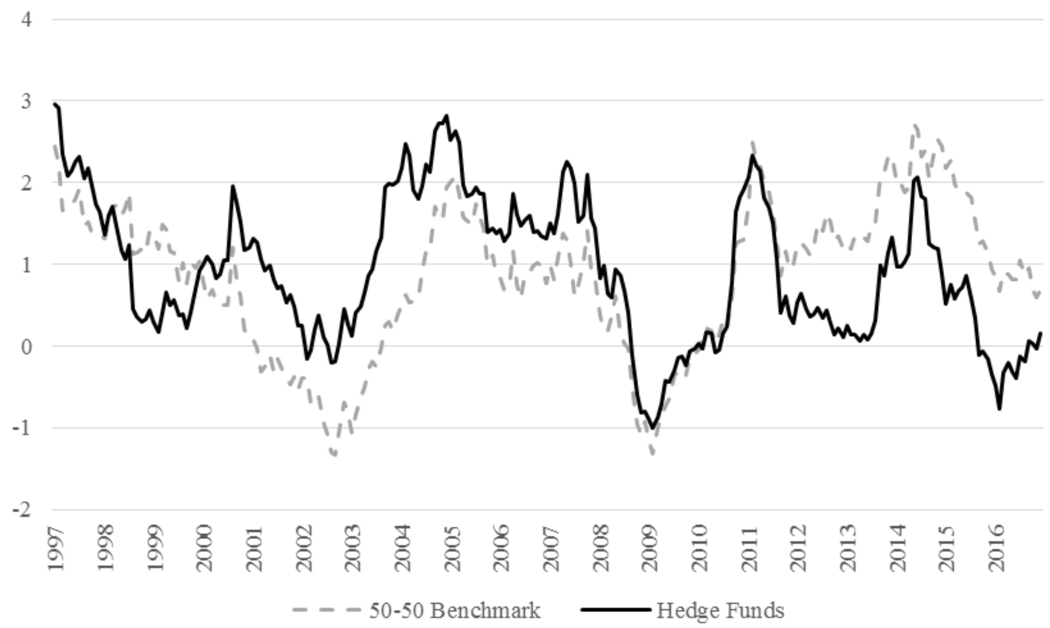
Figure 1B. 1/2008–12/2016



Electronic copy available at: https://ssrn.com/abstract=3034283

Figure 2. Sharpe Ratios

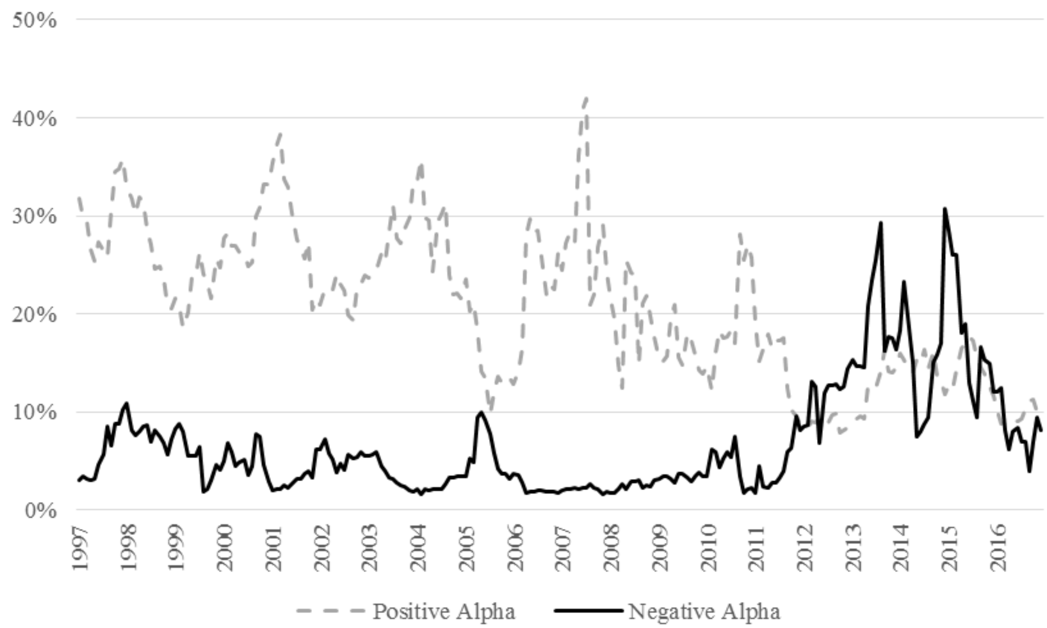
Figure 2 shows the Sharpe ratios of a stock/bond portfolio and an aggregate hedge fund portfolio. The stock/bond portfolio consists of 50% S&P 500 and 50% Vanguard Total Bond Market Index (VBTIX), rebalanced annually. The hedge fund portfolio return is equally weighted each month. The Sharpe ratio is a computed using a 24-month rolling window.



Electronic copy available at: https://ssrn.com/abstract=3034283

Figure 3. Hedge Fund Alpha

Figure 3 shows the percentage of hedge funds with positive or negative alpha significant at the 10% level. Alpha is estimated from a rolling 24-month estimation window.



Electronic copy available at: https://ssrn.com/abstract=3034283

Figure 4. Equity Correlations

Figure 4A shows the average pair-wise correlations of individual stocks. Figure 4B shows the average pair-wise correlations of ten S&P industry sectors. In both cases correlations are estimated using daily returns and a rolling 6-month estimation window. Dashed lines indicate sub-period averages.

Figure 4A. Stock Correlations

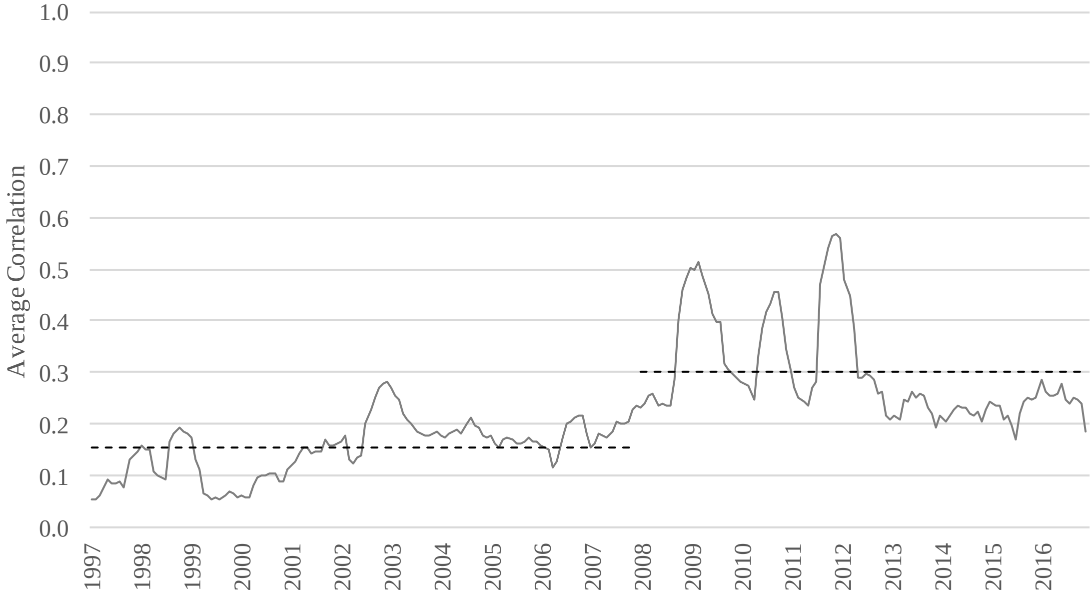
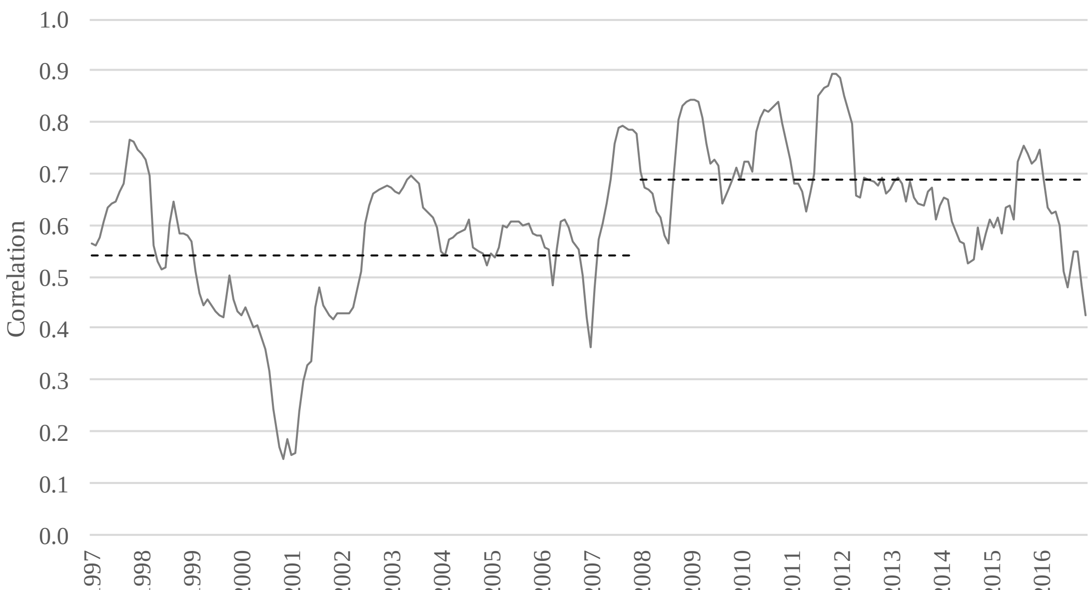


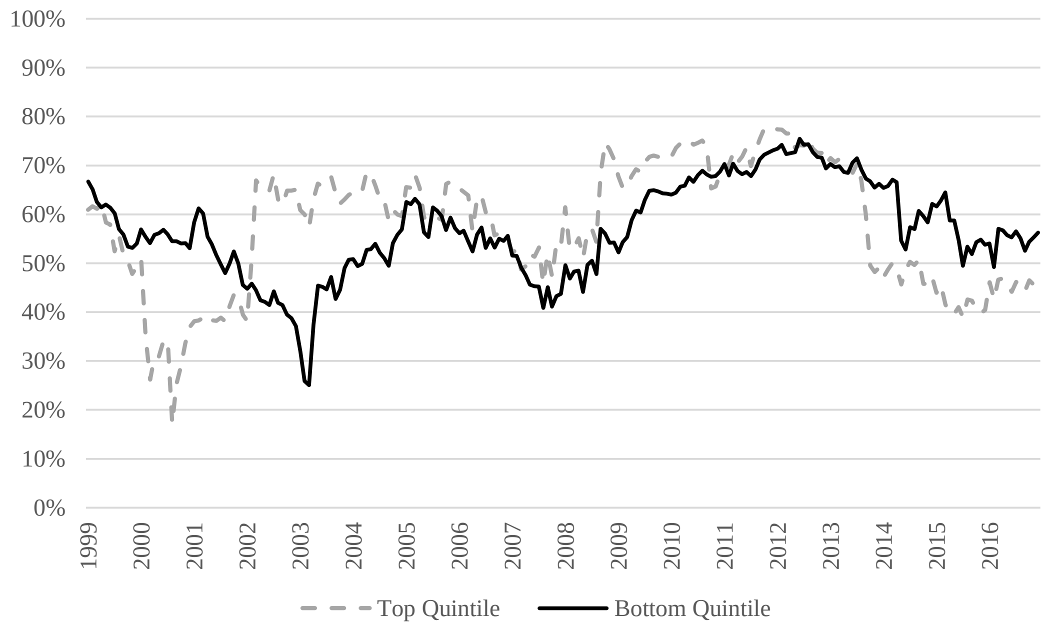
Figure 4B. Sector Correlations



Electronic copy available at: https://ssrn.com/abstract=3034283

Figure 5. Correlation between Hedge Fund Portfolios and the Benchmark Portfolio

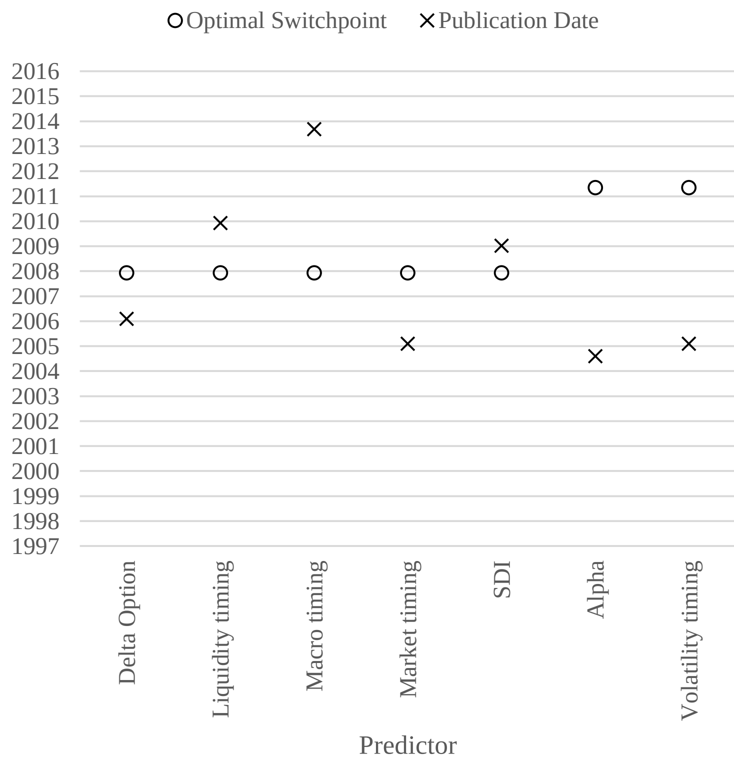
Depicted are rolling 24-month correlations between 15-fund portfolios drawn from either the top or bottom quintiles sorted by Fung and Hsieh (2004) alpha and the benchmark portfolio.



Electronic copy available at: https://ssrn.com/abstract=3034283

Figure 6. Optimal Switch Points of Alpha

Displayed for each predictor is the date (as a hollow circle) which optimally divides the sample into two regimes defined by the alpha of the top quintile. Also displayed is the date (as an “x”) on which the relevant papers were posted to SSRN, save for Alpha for which we use the publication date.



Electronic copy available at: https://ssrn.com/abstract=3034283

Table 1. Description of Hedge Fund Performance Predictors

Listed are the seven predictors of hedge fund performance used in our study, divided into three categories. Column "+/−" indicates whether we expect, based on existing literature, the measure to be related to higher (+) or lower (−) future performance. Predictors based on regression coefficients (including intercepts) are always precision-adjusted, i.e., we use the *t*-values of the coefficients. The dependent variable in all regressions is the fund excess return. We refer to the seven-factor model of Fung and Hsieh (2004) as FH factors. All measures are based on a 24-month rolling window, except for the incentive measure, which is based on a fund's full history until the ranking month.

Category	*Measure*	*+/−*	*Description*
Broad skill measures	Alpha	+	Intercept from a regression against FH factors.
	SDI	+	Strategy distinctiveness index of Sun et al. (2012).
Timing skill measures	Market	+	Market timing following Treynor and Mazuy (1966). Loading on squared S&P 500 excess return, controlling for FH factors.
	Volatility	−	Volatility timing following Chen and Liang (2007). Loading on the interaction of S&P 500 excess return and level of volatility (VIX), controlling for FH factors.
	Liquidity	+	Liquidity timing following Cao et al. (2013). Loading on the interaction of S&P 500 excess return and level of liquidity, controlling for FH factors.
	Macro	+	Macroeconomic timing skill of Bali et al. (2014). Loading on their macroeconomic uncertainty index.
Incentive measure	ΔOption	+	Dollar increase in the value of next year-end incentive options, per dollar increase in fund return, following Agarwal et al. (2009).

Electronic copy available at: https://ssrn.com/abstract=3034283

Table 2. Hedge Fund Performance

Listed in Panel A are annualized average return "Avg", standard deviation "Dev", Sharpe ratio "Sharpe", alpha from the Fung and Hsieh (2004) seven-factor model "Alpha", as well as skewness and kurtosis, based on monthly data from January 1997 through December 2007. Statistics are reported for (1) the stock/bond portfolio, (2) a portfolio of 20% hedge funds, 30% S&P 500, and 50% VBTIX, and (3) a consolidated database of hedge funds. The 20% allocation to hedge funds in the 20/30/50 portfolio is an equally weighted average of all available hedge funds at any point in time. Listed for hedge funds is the cross-sectional equally weighted average and the 25th, 50th, and 75th percentiles. Panel B lists corresponding statistics for the period January 2008 through December 2016.

Panel A. 1/1997–12/2007

	Stock/bond Portfolio	20/30/50 Portfolio	Hedge funds (N = 3,225) Average	25th	50th	75th
Avg	7.6 %	7.9 %	10.8 %	5.4 %	9.7 %	14.6 %
Dev	7.4 %	5.5 %	12.8 %	6.3 %	9.9 %	15.8 %
Skewness	-0.36	-0.37	-0.06	-0.48	-0.03	0.41
Kurtosis	0.25	0.33	2.30	-0.03	0.77	2.35
Sharpe	0.54	0.79	0.78	0.22	0.65	1.17
Alpha	0.34%	1.22%	5.01%	0.49%	4.36%	8.84%

Panel B. 1/2008–12/2016

	Stock/bond Portfolio	20/30/50 Portfolio	Hedge funds (N = 6,069) Average	25th	50th	75th
Avg	6.3 %	5.2 %	2.7 %	-0.8 %	3.2 %	6.8 %
Dev	7.9 %	6.4 %	14.7 %	8.3 %	12.6 %	18.3 %
Skewness	-0.85	-0.89	-0.18	-0.55	-0.15	0.25
Kurtosis	2.03	2.54	2.23	0.13	0.91	2.41
Sharpe	0.77	0.78	0.88	-0.08	0.24	0.61
Alpha	0.95%	0.52%	-2.26%	-7.44%	-1.18%	3.48%

Electronic copy available at: https://ssrn.com/abstract=3034283

Table 3. Performance of Have Beta and Zero Beta Hedge Funds

Listed in Panel A are annualized average return "Avg", standard deviation "Dev", Sharpe ratio "Sharpe", alpha from the Fung and Hsieh (2004) seven-factor model "Alpha", as well as skewness and kurtosis, based on monthly data from January 1997 through December 2007. Statistics are reported for (1) the S&P 500, (2) the subset of hedge funds in a consolidated database with significant CAPM beta "Have Beta", (3) three month Libor, and (4) the subset of hedge funds with insignificant CAPM beta "Zero Beta". Listed for hedge funds is the cross-sectional equally weighted average and the 25th, 50th, and 75th percentiles. Panel B lists corresponding statistics for the period January 2008 through December 2016.

Panel A. 1/1997–12/2007

		Have Beta hedge funds (N = 2,429)					Zero Beta hedge funds (N = 796)			
	S&P 500	Average	25th	50th	75th	3M Libor	Average	25th	50th	75th
Avg	10.2 %	11.9 %	6.5 %	10.6 %	15.8 %	4.1 %	7.8 %	3.1 %	6.8 %	11.1 %
Dev	14.6 %	13.3 %	6.9 %	10.6 %	16.6 %		10.4 %	4.5 %	8.1 %	13.1 %
Skewness	-0.53	-0.07	-0.48	-0.04	0.37		-0.03	-0.47	0.03	0.54
Kurtosis	0.67	1.82	-0.05	0.70	2.14		3.10	0.07	1.02	3.28
Sharpe	0.44	0.76	0.29	0.71	1.22		0.62	0.00	0.46	1.01
Alpha		5.60%	1.01%	4.91%	9.35%		3.40%	-1.05%	2.78%	7.00%

Panel B. 1/2008–12/2016

		Have Beta hedge funds (N = 4,905)					Zero Beta hedge funds (N = 1,164)			
	S&P 500	Average	25th	50th	75th	3M Libor	Average	25th	50th	75th
Avg	8.2 %	2.6 %	-1.1 %	3.1 %	6.7 %	0.7 %	3.7 %	0.0 %	3.7 %	7.3 %
Dev	15.7 %	15.3 %	9.2 %	13.7 %	19.2 %		11.0 %	5.4 %	8.7 %	13.3 %
Skewness	-0.70	-0.24	-0.58	-0.20	0.17		0.05	-0.32	0.11	0.53
Kurtosis	1.43	2.00	0.15	0.92	2.39		2.68	0.06	0.88	2.63
Sharpe	0.51	0.26	-0.09	0.21	0.56		0.91	-0.02	0.38	0.94
Alpha		-3.45%	-8.43%	-2.47%	2.36%		2.93%	-0.65%	3.18%	6.89%

Electronic copy available at: https://ssrn.com/abstract=3034283

Table 4. Economic Value of a Hedge Fund Investment in a Portfolio Context

Listed are differences between the annualized average return "Avg", standard deviation "Dev", Sharpe ratio "Sharpe", alpha from the Fung and Hsieh (2004) seven-factor model "Alpha", and the manipulation proof performance measure "MPPM" of a portfolio consisting of a 20% allocation to hedge funds, 30% to the S&P 500, and 50% to the Vanguard Total Bond Market Index (VBTIX) and those of a stock/bond portfolio. Listed also are measures of incremental utility offered by the portfolio over the stock/bond portfolio for three different risk aversion levels (γ = 1, 5, 10). Results are averaged across 1,000 simulations of a strategy of randomly selecting 15 funds each year from the top quintile formed by each of the predictors. Predictors are ranked by the differences in Sharpe ratio. Hedge fund returns are de-smoothed and adjusted for delistings. For the first five statistics, significance is determined by the percentage of simulations in which the portfolio's statistic is higher or lower than that of the stock/bond portfolio. For the last three columns, significance is determined by the percentage of simulations in which the incremental utility provided by the portfolio relative to the stock/bond portfolio is positive. *, **, and *** correspond to 10%, 5%, and 1% significance levels, respectively. The stock/bond portfolio features an average return of 7.02%, standard deviation of 7.58%, Sharpe ratio of 0.65, an alpha of 0.59%, and an MPPM of 4.03%.

1/1997–12/2016

Predictor	Avg	Dev	Sharpe	Alpha	MPPM	γ1	γ5	γ10
FH alpha	-0.37	-1.77***	0.13***	0.29***	-0.01	-0.25	0.24	0.88***
Macro timing	-0.42	-1.80***	0.13***	0.31***	-0.05	-0.30	0.19	0.84***
SDI	-1.23***	-2.50***	0.08	-0.12*	-0.74***	-1.07***	-0.42*	0.43*
Option delta	-0.98***	-2.11***	0.07	-0.13*	-0.56 **	-0.84***	-0.28	0.46*
Market timing	-0.68**	-1.44***	0.04	-0.05*	-0.37	-0.58*	-0.18	0.36
Volatility timing	-0.81**	-1.60***	0.04	-0.19	-0.47	-0.70**	-0.25	0.34
Random	-0.76*	-1.38***	0.02	-0.24	-0.47	-0.67*	-0.28	0.24
Liquidity timing	-0.91***	-1.50***	0.01	-0.31	-0.59 **	-0.81***	-0.39	0.17

Electronic copy available at: https://ssrn.com/abstract=3034283

Table 5. Benefit of an Allocation to Hedge Funds over Time

Listed are differences between annualized average return "Avg", standard deviation "Dev", Sharpe ratio "Sharpe", and alpha from the Fung and Hsieh (2004) seven-factor model "Alpha" of a portfolio consisting of a 20% allocation to hedge funds, 30% to the S&P 500, and 50% to the Vanguard Total Bond Market Index (VBTIX) and those of a stock/bond portfolio. Results are averaged across 1,000 simulations of a strategy of randomly selecting 15 funds each year from the top quintile formed by each of the predictors. Predictors are ranked by the differences in Sharpe ratio. *, **, and *** correspond to 10%, 5%, and 1% significance levels, respectively. In Panel A, the stock/bond portfolio features an average return of 7.61%, standard deviation of 7.36%, Sharpe ratio of 0.54, and an alpha of 0.34%. In Panel B, the stock/bond portfolio features an average return of 6.29%, standard deviation of 7.88%, Sharpe ratio of 0.77, and an alpha of 0.95%.

Panel A. 1/1997–12/2007

Predictor	Avg	Dev	Sharpe	Alpha
Macro timing	0.15	-2.02***	0.23***	0.84**
FH alpha	0.14	-1.91***	0.22***	0.65**
SDI	-0.63**	-2.51***	0.15**	0.23
Volatility timing	-0.13	-1.66***	0.14**	0.36
Market timing	0.00	-1.47***	0.13**	0.43
Liquidity timing	-0.07	-1.50***	0.13*	0.41
Random	-0.14	-1.58***	0.13*	0.33
Option delta	-0.75***	-2.29***	0.10*	-0.04

Panel B. 1/2008–12/2016

Predictor	Avg	Dev	Sharpe	Alpha
FH alpha	-1.00**	-1.64***	0.04	-0.06
Option delta	-1.27***	-1.93***	0.04	-0.22*
Macro timing	-1.12**	-1.59***	0.02	-0.26
SDI	-1.96***	-2.52***	0.00	-0.56
Market timing	-1.50***	-1.44***	-0.06	-0.66
Volatility timing	-1.63**	-1.58***	-0.06	-0.76
Random	-1.53**	-1.20***	-0.09	-0.89
Liquidity timing	-1.95***	-1.56***	-0.12	-1.19*

Electronic copy available at: https://ssrn.com/abstract=3034283

Table 6. Robustness of Alpha Predictor

Listed are differences between annualized average return "Avg", standard deviation "Dev", Sharpe ratio "Sharpe", and alpha from the Fung and Hsieh (2004) seven-factor model "Alpha" of a portfolio consisting of a 20% allocation to hedge funds, 30% to the S&P 500, and 50% to the Vanguard Total Bond Market Index (VBTIX) and those of a stock/bond portfolio. Results are averaged across 1,000 simulations of a strategy of randomly selecting 15 funds each year from the top quintile formed by Fung and Hsieh (2004) alpha. "($p > 0.05$)" and "($p > 0.10$)" indicate only those funds with alpha significant at the 5% or 10% level are selected. "(AUM > 50)" and "(AUM > 100)" indicate only those funds with AUM at least $50 million or $100 million are selected. Predictors are ranked by the differences in Sharpe ratio. *, **, and *** correspond to 10%, 5%, and 1% significance levels, respectively.

Panel A. 1/1997–12/2016

Predictor	Avg	Dev	Sharpe	Alpha
FH alpha ($p < 0.05$)	-0.23	-1.86***	0.17***	0.50***
FH alpha (AUM ≥ 50)	-0.36*	-1.78***	0.14***	0.29***
FH alpha (AUM ≥ 100)	-0.42**	-1.84***	0.13***	0.25***
FH alpha	-0.37	-1.77***	0.13***	0.29***
FH alpha ($p < 0.10$)	-0.41*	-1.74***	0.12***	0.23***

Panel B. 1/1997–12/2007

Predictor	Avg	Dev	Sharpe	Alpha
FH alpha ($p < 0.05$)	0.16	-2.14***	0.25***	0.80***
FH alpha (AUM ≥ 50)	0.19	-1.94***	0.23***	0.70***
FH alpha	0.14	-1.91***	0.22***	0.65**
FH alpha ($p < 0.10$)	0.00	-2.05***	0.21***	0.55***
FH alpha (AUM ≥ 100)	0.06	-1.96***	0.21***	0.56***

Panel C. 1/2008–12/2016

Predictor	Avg	Dev	Sharpe	Alpha
FH alpha ($p < 0.05$)	-0.71*	-1.60***	0.08	0.26***
FH alpha (AUM ≥ 100)	-1.01***	-1.73***	0.05	-0.04**
FH alpha	-1.00**	-1.64***	0.04	-0.06
FH alpha (AUM ≥ 50)	-1.02**	-1.65***	0.04	-0.09*
FH alpha ($p < 0.10$)	-0.92**	-1.44***	0.03	-0.07**

Electronic copy available at: https://ssrn.com/abstract=3034283

Table 7. Performance of Portfolios formed by Combined Predictors

Listed are differences between annualized average return "Avg", standard deviation "Dev", Sharpe ratio "Sharpe", and alpha from the Fung and Hsieh (2004) seven-factor model "Alpha" of a portfolio consisting of a 20% allocation to hedge funds, 30% to the S&P 500, and 50% to the Vanguard Total Bond Market Index (VBTIX) and those of a passive benchmark. Results are averaged across 1,000 simulations of a strategy of randomly selecting 15 funds each year from the top quintile subset formed by combinations of predictors in each category. Categories are ranked by differences in Sharpe ratio. Significance is determined by the percentage of simulations in which the portfolio's statistic is higher or lower than that of the passive benchmark. *, **, and *** correspond to 10%, 5%, and 1% significance levels, respectively.

Panel A. 1/1997–12/2016

Predictor	Avg	Dev	Sharpe	Alpha
Combined non-timing	-0.72***	-2.18***	0.13**	0.17***
Combined all	-0.67***	-2.01***	0.11**	0.17***
Combined timing	-0.66**	-1.56***	0.06	-0.05*

Panel B. 1/1997–12/2007

Predictor	Avg	Dev	Sharpe	Alpha
Combined non-timing	-0.16	-2.27***	0.21***	0.54***
Combined all	-0.09	-2.07***	0.20***	0.56***
Combined timing	-0.01	-1.59***	0.15**	0.48**

Panel C. 1/2008–12/2016

Predictor	Avg	Dev	Sharpe	Alpha
Combined non-timing	-1.41***	-2.12***	0.04	-0.25
Combined all	-1.38***	-1.99***	0.03	-0.29
Combined timing	-1.46***	-1.57***	-0.04	-0.64

Electronic copy available at: https://ssrn.com/abstract=3034283

Table 8. Optimal Portfolio Weights

Listed are annualized average return "Avg", standard deviation "Dev", Sharpe ratio "Sharpe", and alpha from the Fung and Hsieh (2004) seven-factor model "Alpha" of portfolios of hedge funds, equities (S&P 500), and bonds (Vanguard Total Bond Market Index). In Panel A, funds are selected annually from the top quintile as sorted by Fung and Hsieh alpha using a 24-month rolling window. In Panel B, funds are randomly selected annually. Hedge funds have an allocation as listed, and the other allocation weights are selected to maximize the ex post Sharpe ratio of the portfolio over the time periods listed. For the first three columns in each panel, tests for significant difference between the portfolio with 0% hedge fund weight and other allocations are indicated by *, **, and ***, corresponding to 10%, 5%, and 1% significance levels, respectively. In the fourth column, significance is tested versus a null of zero.

HF	Panel A1. Top Quintile Hedge Funds 1/2000–12/2016			
Weight	Avg	Dev	Sharpe	Alpha
0.0 %	4.86	3.70	0.88	0.55
12.5 %	5.03	3.58	0.95	0.78***
25.0 %	5.20	3.75	0.96	0.95**
37.5 %	5.36	4.27	0.89	1.10*
50.0 %	5.44	5.07***	0.77	1.15

HF	Panel B1. Randomly Selected Hedge Funds 1/2000–12/2016			
Weight	Avg	Dev	Sharpe	Alpha
0.0 %	4.86	3.70	0.88	0.55
12.5 %	4.72	3.71	0.83	0.39
25.0 %	4.55	4.26*	0.70	0.16
37.5 %	4.32	5.48***	0.52*	-0.16
50.0 %	4.12	7.00***	0.38**	-0.54

HF	Panel A2. Top Quintile Hedge Funds 1/2000–12/2007			
Weight	Avg	Dev	Sharpe	Alpha
0.0 %	6.13	3.37	0.87	1.02
12.5 %	6.64**	3.32	1.04**	1.38***
25.0 %	7.02*	3.54	1.09	1.67***
37.5 %	7.21	4.08*	1.02	1.84**
50.0 %	7.31	4.87***	0.90	1.97

HF	Panel B2. Randomly Selected Hedge Funds 1/2000–12/2007			
Weight	Avg	Dev	Sharpe	Alpha
0.0 %	6.13	3.37	0.87	1.02
12.5 %	6.43	3.38	0.96	1.26***
25.0 %	6.47	3.84	0.88	1.34*
37.5 %	6.19	4.95**	0.69	1.25
50.0 %	6.04	6.16***	0.54	1.28

HF	Panel A3. Top Quintile Hedge Funds 1/2008–12/2016			
Weight	Avg	Dev	Sharpe	Alpha
0.0 %	3.81	3.96	0.91	0.29
12.5 %	3.68	3.77	0.92	0.43
25.0 %	3.67	3.88	0.90	0.52
37.5 %	3.82	4.36	0.84	0.68
50.0 %	3.93	5.14*	0.74	0.69

HF	Panel B3. Randomly Selected Hedge Funds 1/2008–12/2016			
Weight	Avg	Dev	Sharpe	Alpha
0.0 %	3.81	3.96	0.91	0.29
12.5 %	3.26	3.96	0.77	-0.20
25.0 %	2.94	4.53	0.61*	-0.68
37.5 %	2.81	5.80***	0.46**	-1.18
50.0 %	2.66	7.52***	0.34**	-1.89

Electronic copy available at: https://ssrn.com/abstract=3034283

Table 9. Scale Effects

Listed are results of panel regressions in which the dependent variable is fund Fung and Hsieh (2004) alpha estimated over the 1997–2016 period. Standard errors are clustered by style and month and additionally clustered by fund using the Pástor et al. (2015) recursive demeaning procedure. Fund size is measured as log of millions of AUM. Industry size is aggregate hedge fund AUM in Panel A and aggregate AUM of equity-oriented hedge funds in Panel B. In both cases, industry size is scaled by global equity market capitalization. Tests for significance are indicated by *, **, and ***, corresponding to 10%, 5%, and 1% significance levels, respectively.

Panel A. All Funds	
Lagged Fund Size	0.002
t-statistic	0.82
Lagged Industry Size	-0.283***
t-statistic	-4.12

Panel B. Equity-oriented Funds	
Lagged Fund Size	0.006
t-statistic	1.37
Lagged Industry Size	-0.709*
t-statistic	-1.85

Electronic copy available at: https://ssrn.com/abstract=3034283