Aggregate Alpha in the Hedge Fund Industry: A Further look at Best Ideas*

F. Amir-Ghassemi, A. Papanicolaou†‡ & M. Perlow§

December 3, 2021

Abstract

This article is an examination of stockpicking behavior of nearly 1,500 hedge funds using regulatory mandated position-level data from the SEC (Form 13F). Using data from June 1999 to Dec 2018, abnormal excess alpha is found on both a gross and dollar basis. Breaking the twenty-year sample into two periods, the authors note a significant decline in gross alphas after the 2008 Global Financial Crisis. In contrast, dollar alphas remain economically and statistically significant. This finding coincides with an increase in aggregate assets in the post-crisis period, suggesting asset growth may be impeding gross alphas. To test this thesis, the authors analyze the 'Best Ideas' within manager portfolios. They find no significant difference between the alphas generated by managers' Best Ideas and the rest of their portfolios, suggesting asset growth is not a significant determinant of alpha deterioration. These findings broadly contrast with prior studies conducted on mutual funds, suggesting differences in portfolio construction and incentive effects.

1 Introduction

Our paper seeks to analyze the stockpicking skill of hedge fund managers. We do this by examining three primary hypotheses: first, we test if aggregate skill exists across the entire asset class. Second, we evaluate if this skill has evolved over time. Third, we assess if hedge fund managers exhibit asset gathering behavior similar to that found in the mutual fund space.

To test these hypotheses, we look at the position-level transparency of hedge funds through the Security and Exchange Commission's ("SEC") mandated quarterly Form 13F filings. This approach

*The authors benefited from a set of individuals for providing valuable assistance. We'd like to extend thanks to: Lu Zhang for kindly furnishing factor data, to Miguel Antón, Jonathan Berk, and Christian Lundblad for taking the time to speak, and to Brian Healy and Tray Spilker for providing valuable edits and comments on earlier drafts.

†Department of Mathematics, North Carolina State University, Campus Box 8205, Raleigh, NC 27695 *apapani@ncsu.edu*

‡Andrew Papanicolaou is partially supported by NSF grant DMS-1907518.

§Mssrs Amir-Ghassemi and Perlow are Partners at Epsilon Asset Management and can be reached at *research@epsilonmgmt.com*

neutralizes many biases associated with voluntary self-disclosure of returns by private investment partnerships. Moreover, by building up from positions instead of analyzing reported returns, we can dissect their portfolios. This allows us to conduct tests which address core theories relating to investment management skill.

The analysis in this paper has two parts: In the first part we provide a baseline understanding of abnormal returns associated with the long-equity portfolio positions of hedge fund managers across 20 years: from 1999 to 2018. Our approach is to construct aggregate composites of domestic equity-oriented common stocks from hedge fund 13F filings, representing trillions of dollars of reported assets. We simulate profit and loss associated with these holdings and calculate traditional gross alphas. We supplement this analysis with a dollar alpha calculation to add a different lens for measuring skill. Our principal findings across these composites demonstrate gross alphas of 13 and 20bps per month, respectively, for the two composite indices proposed in this paper. Across our universe of funds, dollar alphas had a median of \$0.63 million per month and an average of \$2.7 million per month.

We extend this analysis by splitting the time series into two roughly equal periods, centered around the Global Financial Crisis of 2008. In the later period (2009-2018), aggregate assets are demonstrably larger than the former period (1999-2009). We find that gross alphas were closer to 17 - 30bps per month in the former period and 5 - 9bps per month in the later period. This represents a statistically significant decline. Perhaps more importantly, the gross alpha generated in the later period was statistically insignificant. To infer the causes of gross alpha decline, we look to dollar alphas, which were similarly more robust in the former period (\$5.5 mm avg. per month) than the later (\$1.3mm avg. per month). Unlike gross alphas, dollar alphas were statistically significant in the Post period.

A central point of our inquiry is understanding the greater decline in gross alpha versus dollar alpha. We hypothesize asset gathering as the culprit. To test this hypothesis, we analyze the holdings of these hedge fund managers using the Best Ideas methodology of [Cohen et al., 2010]. Best Ideas provide an empirical method to test if increased asset levels are contributing to gross alpha decline. Our analysis of Best Ideas suggests that the decline in gross alpha in the Post period is not principally driven by asset gathering.

The paper will be organized as follows: Section 2 provides the theoretical motivation behind our approach. Section 3 presents a walkthrough of our data source and the important implications behind its choice. Section 4 defines notations and methods for constructing portfolios. This will lead to the analysis in Section 5. Section 6 concludes. The appendix provides added detail on methodology, while our online code repository[1] provides various other analyses performed for robustness.

2 Theoretical Motivation

A wide breadth of research on investment management skill is guided by the principle of the arithmetic of active management ([Sharpe, 1991]), which posits that the cross section of excess returns generated by active managers will sum to zero. This is revelatory in part because it calls to question the relatively

[1]available at `https://github.com/bestideas-APP/best_ideas`

expensive pursuit of active management.

There has been empirical evidence to support Sharpe's theory. [Fama and French, 2010] conducted a careful examination of decades of reported mutual fund returns, finding evidence to support Sharpe. At the time, mutual funds were a pluralistic share of all equity investors. The authors' analysis supported many of their core beliefs pertaining to market efficiency. Their work characterizes the distribution of alphas across these funds as indifferent from zero, calling into question the very existence of skill.

A strand of thinking in the literature pioneered by [Berk and Green, 2004] questions whether 'zero sum alpha' necessarily implies an absence of skill. Berk and Green present a model theorizing the existence of investment management skill within Sharpe's zero-sum equilibrium. They posit that gross alpha – net of fees and transaction costs – is zero-sum, but dollar alpha – where skill is defined by dollar valued added – exists. [Berk and Van Binsbergen, 2015]'s comparable analysis of mutual fund managers found significant and persistent dollar alphas associated with mutual funds. Their results provided a counterpoint to [Fama and French, 2010]'s gross alpha analysis.

Within [Berk and Green, 2004]'s model, dollar alphas can exist while gross alphas are zero due to the competition of asset pricing. Put simply, capital allocators will bid up skilled investment managers to the point where gross alpha is priced competitively, much like the stock of a superior company trading at a higher multiple than a peer. The increase in asset price is not solely caused by increased management fees, reducing the net of fee alpha. It can also be a function of excess asset accumulation by the manager. In their model, such a manager diverts incremental capital inflows into "closet-indexing" portfolio exposure. Because of this, Berk and Green's model hypothesizes that fund manager portfolios demonstrate some mixture of a *Skill Portfolio* and a *Filler Portfolio*.

This brings us to Best Ideas as a measure of such an asset-gathering phenomenon. The working paper [Cohen et al., 2010] developed methodologies to empirically test Berk and Green, and they did so across mutual fund manager portfolio positions. They demonstrate four methods to separate Best Ideas from the rest of a fund's portfolio, the simplest of which being the security with the highest weight after removing the market portfolio's respective weight. They posit that the security with the highest active weight corresponds with the highest future alpha. This approach is a way to separate the Filler of those managers who grow their fund size beyond their strategy's alpha-generating capacity. The authors found statistically significant alphas associated with these Best Ideas relative to the rest of their portfolio across mutual fund managers.

3 Data Choice

Our findings are built upon position-level transparency from 13F data. This transparency captures quarterly position-level holdings across any entity which owns in excess of $100mm of Section 13(f) securities[2].

These securities can be summarized as marketable equities, publicly-traded real estate investment

[2]The SEC does not provide an explicit definition of 13F securities and instead provides a list of all eligible securities which must be filed here: https://www.sec.gov/divisions/investment/13flists.htm

trusts, certain corporate bonds, and exchange-traded options. We source our 13F data from Novus Partners, Inc., an analytics provider with a specialization in 13F data. Refer to 7.1 for more details on the cleaning techniques and breadth of this unique dataset.

We source individual security pricing data from CRSP (Center for Research in Security Prices), covering relevant equity-like instruments traded on NYSE, NASDAQ, and AMEX. Securities are matched to the Form 13F holdings of each hedge fund. This allows us to represent an investment manager's Section 13(f) portfolio ("Portfolio") as of quarter-end with security pricing to simulate fund-level profit and loss (see Section 4.1).

We source factor data for the [Carhart, 1997] as well as [Fama and French, 2015] asset-pricing models from Ken French's website. Factor data for the [Hou et al., 2015] model was kindly furnished by Lu Zhang. For the sake of brevity, we will present results across [Fama and French, 2015] going forward, with ancillary factor model results in the online supplement.[3]

4 Methodologies

4.1 Profit and Loss Calculation

We assume for the sake of analysis that each portfolio represents a complete projection of the aggregate manager holdings, and that trades associated with these portfolio positions occur at quarter-end with pricing from CRSP and quantities from the updated Form 13Fs. As such, portfolio constituents are calculated at each quarter-end. Assuming buy-and-hold, we compute daily returns inter-quarter using derived daily weights and total returns from CRSP to generate gross return,

$$R_t^g = \sum_{k=1}^{N} w_{k,t-1} r_{kt},$$

where $w_{k,t-1}$ is the end-of-day (EOD) weight of security k at $t-1$ and r_{kt} is the total return of security k on day t.

Daily return calculation is used to retain fidelity in graphing on our online appendix. We then convert all daily return streams to a monthly frequency to match that of our factor models. It is important to note that we calculate portfolios as of quarter-end despite the 45 day "lag" period associated with the release of 13F filings. For this reason, these are not investable portfolios.

While this buy-and-hold return calculation approach may seem inhibiting, it provides several advantages in measuring skill across hedge funds. We can frame these advantages through [Stulz, 2007]'s decomposition of the deficiencies of traditional return analysis on hedge fund firms: survivorship bias, issues of return smoothing, and autocorrelation of returns. Incubation bias can also be added to the list of deficiencies. For more color on these specific issues, see Appendix 7.2.

[3] All t-statistics are adjusted using [Newey and West, 1987] with a lag length of 9 months.

4.2 Aggregate Alpha Measurements

To perform an analysis of aggregate hedge fund manager gross alphas, we need to denote the subset of assets under consideration as well as the data underlying.

Let A represent the subset of securities across CRSP, defining our aggregate security universe,

$$A := \Big\{ \text{ all NYSE, NASDAQ, and AMEX stocks } \Big\} \ .$$

Let HF be the set of hedge funds who filed a Form 13F.

$$HF := \Big\{ \text{ all managers filing a 13F} \Big\} \ .$$

Let M_i represent a 13F from manager i from HF,

$$M_i := \Big\{ \text{ all stocks in manager } i\text{'s 13F } \Big\} \ .$$

Letting N_k denote the total number of shares outstanding for each $k \in A$, let N_{ik} denote the total number of shares held by manager i in stock k. Letting S_k denote the price of stock k, and letting $N_k^{\mathrm{hf}} = \sum_i N_{ik}$ (i.e. the total number of shares outstanding across our universe of hedge fund firms), the following are the portfolios we will use for benchmarking:

$$\bar{w}_k = \frac{N_k S_k}{\sum_{\ell \in A} N_\ell S_\ell} \qquad \forall k \in A \tag{1}$$

$$\widetilde{w}_{ik} = \frac{N_{ik} S_k}{\sum_{\ell \in M_i} N_{i\ell} S_\ell} \qquad \forall k \in M_i, \forall i \in HF \tag{2}$$

$$\bar{w}_k^{\mathrm{hf}} = \frac{N_k^{\mathrm{hf}} S_k}{\sum_{\ell \in A} N_\ell^{\mathrm{hf}} S_\ell} \qquad \forall k \in A \ . \tag{3}$$

To summarize, formula (1) represents that market capitalization weight of stock k relative to all NYSE, NASDAQ, and AMEX stocks, (2) denotes the capitalization weight of stock k relative to the rest of manager i's portfolio, and (3) represents that aggregate dollar value of stock k across all hedge fund managers' portfolios divided by total dollar value filed across all stocks.

Our empirical gross alpha findings are presented across two aggregate composites, representing an asset-weighted and a fund-weighted representation of all hedge fund data capture. The asset weighted composite is given by (3) that we will refer to as HFU_{AW}. We also construct a composite representation of aggregate hedge fund behavior where every fund has equal weight in the aggregate, debiasing the size consideration of M_i. We refer to this as HFU_{FW}:

$$\bar{w}_k^{\mathrm{hf}^*} = \frac{1}{|HF|} \sum_{i \in HF} \frac{N_{ik} S_k}{\sum_{\ell \in M_i} N_{i\ell} S_\ell} \qquad \forall k \in A \tag{4}$$

where $|HF|$ represents the number of managers in the set HF.

4.3 Hedge Fund Manager Dollar Alphas Methodology

To provide a different lens for evaluating fund manager skill, we measure dollar alphas. We follow [Berk and Van Binsbergen, 2015] in defining gross alpha as

$$\alpha_{it}^g \equiv R_{it}^g - R_{it}^B,$$

where R_{it}^{g} indicates manager i's total gross return and R_{it}^{B} represents the benchmark for manager i at time t.

To estimate R_{it}^{B}, we regress each manager's simulated monthly return against [Fama and French, 2015] separately and derive a benchmark return comprised of weighted factor portfolios. Next, we obtain realized value added (or dollar alpha) by

$$V_{it} \equiv q_{i,t-1}(R_{it}^{g} - \hat{R}_{it}^{B}),$$

where $q_{i,t-1}$ is manager i's simulated assets under management (AUM) at $t-1$, adjusted by inflation. In quarterly periods, AUM is represented by the sum of filed eligible assets whereas intra-quarter months simulate NAVs using updated security pricing. We then estimate value added for each fund by

$$\hat{S}_i = \sum_{t=1}^{T_i} \frac{V_{it}}{T_i},$$

where T_i represents the number of periods fund i existed. Finally, we estimate the mean of the distribution

$$\bar{S} = \frac{1}{N}\sum_{i=1}^{N} \hat{S}_i, \tag{5}$$

and the time-weighted mean

$$\bar{S}_W = \frac{\sum_{i=1}^{N} T_i \hat{S}_i}{\sum_{i=1}^{N} T_i}. \tag{6}$$

4.4 Methods for Isolating Best Ideas

We further our examination of aggregate hedge fund alphas by analyzing their Best Ideas. This approach is inspired by applying the model of [Berk and Green, 2004] on hedge fund portfolio positions, as originally performed by [Cohen et al., 2010]. We isolate Best Ideas using 9 distinct methodologies. As mentioned, methods r_{1-4} are inspired from [Cohen et al., 2010]'s initial implementation of Best Ideas on mutual funds as an empirical application of [Berk and Green, 2004]'s model. R_{5-7} are derived from [Shumway et al., 2009], and r_{8-9} from practitioner's view on portfolio analysis. The following is a brief summary of our nine methodologies. For a detailed breakdown of these methodologies, please see Appendix 7.4.

Run	Tilt Metric (λ_{ik})	Description
r_1	$w_{ik} - \bar{w}_k$	Best ideas (Market Tilt)
r_2	$\sigma_k^2(w_{ik} - \bar{w}_k)$	Best Ideas (Variance-adjusted Market Tilt)
r_3	$w_{ik} - \widetilde{w}_{ik}$	Best Ideas (Intra-Cap Weight Tilt)
r_4	$\sigma_k^2(w_{ik} - \widetilde{w}_{ik})$	Best Ideas (Variance-Adjusted Intra-Cap Weight Tilt)
r_5	$(\Sigma w_i)_k$	Best Ideas (Mean-Variance Tilt)
r_6	$(\Sigma(w_i - \bar{w}))_k$	Best Ideas (Cap-Weight Adjusted Mean-Variance Tilt)
r_7	$(\Sigma(w_i - \widetilde{w}_{ik}))_k$	Best Ideas (Intra-Cap Weight Adjusted Mean-Variance Tilt)
r_8	w_{ik}	Best Ideas (Portfolio Weight / Zero Tilt)
r_9	$w_{ik} - \bar{w}_k^{\mathrm{hf}}$	Best Ideas (HFU$_{AW}$ Tilt)

Table 1: The 9 methodologies we employ to identify best ideas.

where w_{ik} represents the weight of security k in manager i's portfolio, $\bar{w}_k$ represents security k's weight in the market portfolio, σ_k^2 represents security k's idiosyncratic variance, $\widetilde{w}_{ik}$ represents security k's weight if manager i's portfolio is re-weighted according to each stock's market capitalization, Σ represents a shared covariance matrix, and $\bar{w}_k^{\mathrm{hf}}$ represents security k's weight in the HFU_{AW} portfolio.

4.5 Portfolio Disaggregation Techniques

We measure 3 variants of alphas associated with manager portfolios: (1) the standalone Best Idea of each manager (which we define as *Best Ideas*), (2) a market-neutral portfolio that is long the Best Idea of each manager while short the Rest of the manager's pro-rated portfolio constituents (which we define as *Best minus Rest*), and (3) the portfolio positions excluding Best Ideas (which we define as *Rest Ideas*). Our goal is to thoroughly examine potential differences in gross alpha characteristic across the portfolio. Following [Berk and Green, 2004], if there is evidence of asset gathering, there should be evidence of an alpha generative sleeve within a manager's portfolio contrasting with what is otherwise passively invested 'filler'. See Appendix 7.5 for detail on the formulation of these three methodological composites.

4.6 Pre / Post GFC Analysis

The abnormal return demonstrated previously across 20 calendar years of analysis is further decomposed into sub-periods, which we define as PreGFC and PostGFC. We do so to understand if the behavior has changed over time. We think this is a particularly important question given the size of aggregate reported assets has grown nearly 700% during our sample period.

We choose the Great Financial Crisis of 2008 - 2009 as a point of demarcation. June 2009 was chosen as the partitioning date because (a) it represents an approximate midpoint in the twenty-year time-series of data (120 months prior, 114 months after); (b) it arguably represented a structural shift across hedge fund managers and their risk behaviors, especially as it pertains to counterparty risk and use of leverage; (c) we observe the regulatory assets captured in the later period coincides with an elevated level of aggregate assets, as compared to the earlier years in the sample set.

5 Results & Analysis

5.1 Gross Alpha Decomposition of Aggregate Level Hedge Fund Portfolios: Total Period, Pre, and Post

We begin by measuring gross alphas and dollar alphas across our entire set of managers, for our entire time series. To measure gross alphas, we generate two composites across our Hedge Fund Universe ('HFU'): an asset-weighted composite ('HFU_{AW}') and a fund-weighted composite ('HFU_{FW}') defined in equations (3) and (4), respectively. HFU_{FW} treats each fund as an equal contributor within the composite, while HFU_{AW} aggregates all assets filed across all funds. The latter skews results towards (against) funds who file larger (smaller) reported asset levels. Table 2 contains summary statistics for the filing firms comprising our composites.

Table 3 shows the gross alpha estimates for regressions against our HFU_{FW} and HFU_{AW} composites. As a group, hedge fund managers generate abnormal returns of 20bps per month on a fund-weighted basis across our selected factor model and 13bps per month on an asset-weighted basis. These alphas are statistically significant at the 5% level. A market beta close to 1.0 for our composites aligns with our treatment of each hedge fund managers' filed portfolio as a fully invested long-only portfolio. It is noted that results are stronger when we treat firms equally, negating the impact of firm size. The economic return generated by the fund-weighted composite exceeds the asset-weighted composite by 12bps per month. In total, hedge fund stockpicks generate statistically significant gross alphas.

Moving to Table 4, we continue by decomposing our time-series into sub-periods of PreGFC (June 1999 – June 2009) and PostGFC (July 2009 – December 2018). We do so to understand if the behavior of hedge fund security selection has changed over time. We choose the Global Financial Crisis of 2008 - 2009 as an anchor for doing so. We chose this because it represents an approximate midpoint in the twenty-year time-series of data (120 months prior, 114 months after).

As indicated in Table 2, aggregate reported assets increase by nearly \$1.5 trillion in the PostGFC period from 2009 – 2018 (from \$680 billion to \$2,120 billion), as compared to a \$360 billion increase from 1999 – 2008 (from \$320 billion to \$680 billion). To control for some abnormal asset changes in 2008, the average annual reported assets in PreGFC were \$700 billion compared to \$1,800 billion in PostGFC. Table 4 demonstrates the regression results split between the periods.

The PreGFC period saw a mean monthly return of 29bps per month, as the period had two severe market pullbacks (the 2000-2002 tech bubble/market recession and the 2008 Global Financial Crisis). What is perhaps remarkable is the 31bps of monthly gross alpha generated by our fund-weighted composite, besting the actual economic return. We see the same dynamic on the asset-weighted composite, where the mean monthly return was 9bps per month, while the gross alpha was 17bps per month. Receiving a superior alpha term compared to economic return is a great prospect for investors.

The PostGFC period saw a statistically significant decline in gross alpha from the PreGFC period, even as mean monthly returns soared with a rising bull market. For HFU_{FW}, the mean monthly return was 119bps per month in PostGFC, while gross alpha was just 9bps per month; barely statistically significant. Things are worse for the asset-weighted composite which overweights larger filing entities. The mean monthly return was 116bps per month, but the gross alpha was a statistically insignificant 5bps per month. In short, our alphas were far less impressive in the PostGFC, especially when you allow the size of the firm to impact the assessment of alpha.

5.2 Dollar Alpha Decomposition of Aggregate Level Hedge Fund Portfolios: Total Period, Pre, and Post

As explained in the methodology section, a dollar alpha decomposition seeks to measure the excess dollar contribution generated by a fund manager's portfolio. This can be a powerful lens for analyzing investment management skill against the backdrop of asset growth. For the purpose of distilling our twenty-year analysis into an average monthly dollar figure, we follow [Berk and Van Binsbergen, 2015] in inflation-adjusting the dollar contributions across our time series. Unlike with gross alphas, we calculate dollar alpha at the firm level to control for the asset-level disparity across these firms.

Our principal findings for dollar alphas are presented in Table 6. When measuring dollar alphas, we find hedge funds extract on average \$2.7 million in monthly alphas from the market through their long equity stockpicks. The median value is approximately \$0.6 million, demonstrating skew. Roughly 73% of eligible reporting firms generated positive dollar alphas during the entire analysis period. This is a formidable amount of dollar alpha and is statistically significant at the 1% level.

Similar to our gross alpha measurements, we compute dollar alphas across the PreGFC and PostGFC periods. We also see a disparity in dollar alphas extracted across both periods. In the PreGFC period, hedge funds averaged $5.5 million of monthly alpha, and roughly 87% of funds generated positive dollar alpha. The median dollar alpha was also quite pronounced, approximately $1.9 million per month.

There is a noted decline in the PostGFC period. The dollar alphas drop to $1.3 million per month, and the median falls perilously close to zero ($0.2 million). Just 57% of filing entities generated dollar alpha in the Post period. Unlike with gross alphas, this dollar alpha value remains statistically significant. In the following section, we investigate the drivers of this asymmetric decline. We do so by comparing managers' best ideas to the rest of their portfolio positions.

5.3 Best Ideas: A Test of Portfolio Saturation

To summarize our results above, hedge funds as an asset class have been able to demonstrate statistically significant gross and dollar alphas associated with their domestic equity stockpicks. The period from June 1999 - June 2009 saw more robust gross and dollar alphas than the period following (July 2009 – December 2018). However, gross alphas declined worse than dollar alphas. What could be driving this? A source of difference can be the impact of size on the calculations. As an example, a fund that is able to apply a strategy which generates a fixed $50 million of alpha per annum will see its gross alpha deteriorate as its assets grow. We test if this dynamic is at work with the aggregate hedge fund industry.

As we alluded to earlier, the aggregate asset levels of hedge fund 13F filings differ dramatically from PreGFC to PostGFC. The average annual asset footprint across hedge funds in the Pre period was approximately $700 billion (see Table 2). This contrasts to average annual assets in the Post period of approximately $1,800 billion, or roughly a 250% increase.

One way to normalize this is to compare the total dollar amount of dollar alpha extracted by these hedge fund firms. If our example of a firm with a fixed capacity for dollar alpha creation holds across all firms, we would expect dollar alphas to maintain some nominal level even as gross alphas decline. During the 120 month PreGFC period, the entirety of filing hedge funds extracted over $500 billion of alpha from the market, even while average aggregate assets were approximately $700 billion. Moving to PostGFC, during this 114 month period, the asset class extracted just $135 billion of alpha from the market, even as average assets swelled to $1,800 billion. This suggests that size may not be the only culprit.

In testing this, we revert to the model of [Berk and Green, 2004]. It hypothesizes that fund managers attract inflows after a demonstration of past skill. Their strategies are unable to linearly scale, so a portion of incremental capital flow will be passively invested. Within this dynamic, gross alpha will decline across the portfolio, but there will still be positions which represent the initial portfolio characteristic which exhibited positive gross alpha. If we can dissect a portfolio, we may be able to demonstrate this phenomenon. The tool we use to do is the so-called Best Ideas approach of [Cohen et al., 2010].

5.4 Gross Alpha Decomposition of Best Ideas: Total Period

We present Best Idea gross alphas in Table 6. Across our nine Best Idea methodologies and our entire time period, hedge funds generated between 21 – 33bps of monthly gross alphas through their Best Ideas. Several of these Best Idea methodologies generated statistically insignificant alphas due to the volatility of their returns. These tended to be the methodologies with volatility-weighted tilt metrics (i.e., r_2, r_4, r_6, r_7 in Table 1). While these significant alpha seems impressive, remember that the HFU_{FW} composite comprised of all filed hedge fund holdings generates comparable alphas (20bps per month). To better understand if the Best Ideas of managers generate substantially different alphas from the Rest of their portfolio positions, we must look to a market neutral construct of Best minus Rest.

We present Best minus Rest market-neutral portfolio regressions in Table 7. Across these composites, the alphas deteriorate measurably. Starting with the economic returns of these market neutral constructs, returns varied from -3bps per month to 13bps per month. Meanwhile, the Fama-French 5 factor model alphas are between 2bps per month and 21bps per month. While some of those figures seem meaningful, none of the alphas were statistically significant at the 5% level. This indicates no statistical difference in alpha characteristic between hedge fund best ideas and the rest of their portfolio positions. For robustness, we tested varying intervals of Best Ideas in formulating these composites (presented in Table 8). This takes not only the sole top position of each fund manager in constructing a composite of aggregate Best Ideas, but the top 3 or 5. We find similar results, demonstrating no statistically significant alphas.

Just to hammer the point home, we look at the Rest Ideas as standalone portfolios. This allows us to understand if these non Best Idea positions behave similar to a closet-indexed portfolio (see Table 9). Instead, we find statistically significant alphas across each of our nine Rest portfolios across the entire time-period. Similar to our Best minus Rest tests, we examine increasing intervals of excluded Best Ideas in formulating Rest Idea composites. Results hold strong through the exclusion of 20 Best Ideas from each manager.

This suggests that a hypothesis wherein hedge fund portfolio alphas have deteriorated due to asset accumulation is unsupported, which we conclude because their 'rest' positions of hedge fund portfolios are indeed alpha generative exposure.

5.5 Best minus Rest Idea Gross Alphas: PreGFC vs PostGFC

To punctuate our findings above, we analyze the alpha characteristics of Best minus Rest market-neutral composites in the Pre and PostGFC periods. Between the two periods, assets increase dramatically and a disparity between gross and dollar alpha exists. If this disparity is solely a function of asset gathering, we would expect to see Best Idea alphas differing from Rest Idea alphas in the later period. We present these findings in Table 10.

Similar to the Total period, in the PreGFC period where asset levels were lower and gross alphas more meaningful, Best minus Rest constructs again generate statistically insignificant alphas. This would follow if there was no need for firms to asset gather in this earlier period.

The true test is in the PostGFC period. Our findings are much the same as the total period: statistically insignificant alphas associated with most of the Best minus Rest Ideas of hedge fund firms. The one outlier finding is with r_8, our measurement which simply looks at securities with the largest portfolio weight in each manager's filings. This method does not remove the market's corresponding weight in determining Best Ideas, and generates statistically significant spread (approx. 7bps per month).

All of this suggests that the deterioration of alphas across hedge funds during the PostGFC periods is not simply a function of asset gathering.

6 Conclusions

We find evidence of aggregate stockpicking ability by analyzing the long domestic equity holdings of a very large population of hedge fund managers. These results vary depending on aggregation methodology but represent 13 or 20bps of abnormal return per month over a twenty-year sample. While return simulation has its limitations, our results are less prone to many biases associated with self-reported fund returns.

We further decompose this twenty-year result into two nearly equal periods centered around the Global Financial Crisis. In the first period we find much stronger gross alphas compared to the second period, which saw a statistically significant decline. To better understand this trend, we look at the dollar alpha generated by fund managers: again, over the total period, then split into the two sub-periods. Here we too find a decline in dollar alphas from Pre to Post. Nonetheless, dollar alpha generated in the Post period remains statistically significant.

We hypothesize the difference in decline may be a function of growth of the asset class, which grew from an average total asset base of approximately \$700 billion in the Pre period to \$1,800 billion in the post period. However, our empirical tests seem to suggest that asset growth is not the driver of alpha deterioration. First, we saw a nominal decline in inflation-adjusted dollar alpha generated across all funds: from \$500 billion in the Pre period to \$135 billion in the post period. Next, we analyze the Best Ideas of these fund managers to test if fund managers were stuffing their portfolios with index-hugging exposure. We found constant alpha characteristics across the portfolio positions (i.e., between their 'Best Ideas' and the 'Rest' of their portfolio positions) of hedge fund managers across both time periods.

A few takeaways from this analysis: if the decline in hedge fund dollar alpha is not caused by asset gathering, it may be caused by other factors. First, the Global Financial Crisis arguably represented a psychological shift across hedge fund managers and their risk behaviors. We saw structural change to key counterparties of these firms (e.g., prime brokers) and the composition of their investors (increased institutionalization). These changes might have impacted the efficacy of hedge fund stockpicking. Additionally, alphas generated in the recent period may simply be a reflection of increased market efficiency. This would suggest a more competitive environment for extracting dollar alpha from the market.

Even with a 250% increase in aggregate asset footprint, we do not find evidence to suggest asset gathering behavior. A reason may be incentives. Given the incentive fee portion of a hedge fund manager's compensation can provide limitless compensation (like a call option), the restricting element of finding pricing equilibrium - as posited by Berk and Green's model - may not be applicable to this subset of managers. Put more simply, hedge fund firms have competing incentives between asset gathering (maximizing management fees) and outperformance (maximizing incentive fees).

To extend our findings, we suggest the following as areas of further inquiry:

- a further examination of hedge fund skill on this dataset through bootstrap, a la [Fama and French, 2010], [Kosowski et al., 2006], and [Harvey and Liu, 2018];
- an event-based analysis meant to deeply examine the time-series decline of alpha generated by hedge fund stockpicking; and
- a further examination of the population of hedge fund managers, grouped by characteristics (e.g., size) or historical behaviors (e.g., perceived skill), in order to better decompose a heterogeneity of behaviors, a la [Patel and Spilker, 2020].

References

[Ackermann et al., 1999] Ackermann, C., McEnally, R., and Ravenscraft, D. (1999). The performance of hedge funds: Risk, return, and incentives. *The journal of Finance*, 54(3):833–874.

[Berk and Green, 2004] Berk, J. B. and Green, R. (2004). Mutual fund flows and performance in rational markets. *Journal of Political Economy*, 112:1269–1295.

[Berk and Van Binsbergen, 2015] Berk, J. B. and Van Binsbergen, J. H. (2015). Measuring skill in the mutual fund industry. *Journal of Financial Economics*, 118(1):1–20.

[Carhart, 1997] Carhart, M. (1997). On persistence in mutual fund performance. *The Journal of Finance*, 52:57–82.

[Cohen et al., 2010] Cohen, R. B., Polk, C., and Silli, B. (2010). Best ideas. *Available at SSRN 1364827*.

[Evans, 2010] Evans, R. B. (2010). Mutual fund incubation. *The Journal of Finance*, 65(4):1581–1611.

[Fama and French, 1992] Fama, E. and French, K. (1992). The cross-section of expected stock returns. *The Journal of Finance*, 47(2):427–465.

[Fama and French, 2010] Fama, E. F. and French, K. R. (2010). Luck versus skill in the cross-section of mutual fund returns. *The journal of finance*, 65(5):1915–1947.

[Fama and French, 2015] Fama, E. F. and French, K. R. (2015). A five-factor asset pricing model. *Journal of financial economics*, 116(1):1–22.

[Getmansky et al., 2004] Getmansky, M., Lo, A. W., and Makarov, I. (2004). An econometric model of serial correlation and illiquidity in hedge fund returns. *Journal of Financial Economics*, 74(3):529–609.

[Harvey and Liu, 2018] Harvey, C. R. and Liu, Y. (2018). False (and missed) discoveries in financial economics. *Available at SSRN 3073799*.

[Hou et al., 2015] Hou, K., Xue, C., and Zhang, L. (2015). Digesting anomalies: An investment approach. *The Review of Financial Studies*, 28(3):650–705.

[Kosowski et al., 2006] Kosowski, R., Timmermann, A., Wermers, R., and White, H. (2006). Can mutual fund "stars" really pick stocks? new evidence from a bootstrap analysis. *The Journal of finance*, 61(6):2551–2595.

[Malkiel and Saha, 2005] Malkiel, B. G. and Saha, A. (2005). Hedge funds: Risk and return. *Financial analysts journal*, 61(6):80–88.

[Newey and West, 1987] Newey, W. K. and West, K. D. (1987). Hypothesis testing with efficient method of moments estimation. *International Economic Review*, pages 777–787.

[Patel and Spilker, 2020] Patel, N. and Spilker, H. (2020). Alpha by affiliation. working paper, preprint available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3518906.

[Sharpe, 1991] Sharpe, W. F. (1991). The arithmetic of active management. *Financial Analysts Journal*, 47(1):7–9.

[Shumway et al., 2009] Shumway, T., Szefler, M., and Yuan, K. (2009). The information content of revealed beliefs in portfolio holdings. *January) University of Michigan working paper*.

[Stulz, 2007] Stulz, R. M. (2007). Hedge funds: Past, present, and future. *Journal of Economic Perspectives*, 21(2):175–194.

Table 2: Summary Statistics for Aggregate SEC Form 13F Filing Data Capture for Hedge Fund Firms Through Time Aggregate Unique Firms = 1,489 over time.

	Number of Live Funds	Average Fund Size ($MM)	Total Assets ($MM)	Mean Number of CRSP Securities by Firm
1999	199	1,609	320,196	116
2000	241	1,398	336,838	119
2001	291	1,209	351,763	114
2002	344	986	339,227	108
2003	433	1,260	545,385	116
2004	516	1,470	758,681	118
2005	620	1,577	978,048	116
2006	735	1,707	1,254,289	113
2007	827	1,719	1,421,579	100
2008	842	808	679,916	81
2009	818	1,255	1,026,404	96
2010	853	1,431	1,221,022	97
2011	867	1,396	1,210,247	93
2012	879	1,656	1,456,003	93
2013	915	2,118	1,937,883	95
2014	961	2,279	2,190,038	93
2015	976	2,158	2,106,630	93
2016	967	2,142	2,070,968	94
2017	981	2,542	2,493,342	95
2018	927	2,296	2,128,155	91

Table 3: Summary Statistics for HFU_{FW} and HFU_{AW}

FF5 represents $r_t - rf_t = \alpha + bRMRF_t + sSMB_t + hHML_t + pRMW_t + cCMA_t + \epsilon_t$

RM is the return on a value-weighted market portfolio of the CRSP universe of stocks, and RF is the 1-month Treasury bill rate. The construction of SMB and HML follow [Fama and French, 1992] while RMW and CMA follow [Fama and French, 2015]. HFU_{FW} represents a fund-weighted composite of all assets filed by hedge fund managers, given by equation (4). HFU_{AW} represents an asset-weighted composite of all assets filed by hedge fund managers, given by equation (3). The period is June 1999 through December 2018 (n = 234). T-statistics are found below the point estimates and are adjusted using [Newey and West, 1987] with a lag length of 9 months.

	Mean	$\widehat{\alpha}$	$\widehat{b}$	$\widehat{s}$	$\widehat{h}$	$\widehat{p}$	$\widehat{c}$
Panel A: Time Series of Aggregate Hedge Funds (Asset Weighted + Fund Weighted)							
HFU_{FW}	0.0073	0.0020	1.0238	0.3177	0.0109	-0.0366	0.0443
		3.60	56.32	13.99	0.23	-1.26	0.83
HFU_{AW}	0.0061	0.0013	1.0119	0.2539	-0.0575	-0.0572	0.0197
		2.61	46.54	7.79	-1.84	-1.50	0.31

Table 4: Summary Statistics for HFU_{FW} and HFU_{AW} split into Pre and Post GFC

FF5 represents $r_t - rf_t = \alpha + bRMRF_t + sSMB_t + hHML_t + pRMW_t + cCMA_t + \epsilon_t$

RM is the return on a value-weighted market portfolio of the CRSP universe of stocks, and RF is the 1-month Treasury bill rate. The construction of SMB and HML follow [Fama and French, 1992] while RMW and CMA follow [Fama and French, 2015]. HFU_{FW} represents a fund-weighted composite of all assets filed by hedge fund managers, given by equation (4). HFU_{AW} represents an asset-weighted composite of all assets filed by hedge fund managers, given by equation (3). The period is split into two distinct periods: PreGFC which is June 1999 through June 2009 (n = 120), and PostGFC, which is June 2009 through December 2018 (n = 114). T-statistics are found below the point estimates and are adjusted using [Newey and West, 1987] with a lag length of 9 months.

	Mean	$\widehat{\alpha}$	$\widehat{b}$	$\widehat{s}$	$\widehat{h}$	$\widehat{p}$	$\widehat{c}$
Panel A: Time Series of Fund-Weighted (HFU_{FW}) Hedge Funds							
PreGFC	0.0029	0.0031	1.0612	0.3164	-0.0174	0.0013	0.1094
		3.51	29.49	8.83	-0.22	0.020	1.66
PostGFC	0.0119	0.0009	1.0095	0.3046	0.0059	-0.1834	-0.1046
		2.18	123.22	19.19	0.29	-7.30	-2.75
Two Sample T-Test on $\widehat{\alpha}$ (Pre v Post): 24.59							

	Mean	$\widehat{\alpha}$	$\widehat{b}$	$\widehat{s}$	$\widehat{h}$	$\widehat{p}$	$\widehat{c}$
Panel B: Time Series of Asset-Weighted (HFU_{AW}) Hedge Funds							
PreGFC	0.0009	0.0017	1.0475	0.3037	-0.0953	0.0042	0.0508
		2.08	28.90	7.76	-2.22	0.10	0.59
PostGFC	0.0116	0.0005	1.0143	0.1436	0.0004	-0.1077	-0.0577
		1.19	169.36	9.16	0.02	-5.74	-2.00
Two Sample T-Test on $\widehat{\alpha}$ (Pre v Post): 14.24							

Table 5: Dollar Alpha Summary Statistics

Per [Berk and Van Binsbergen, 2015]: For every fund in our database, we estimate the monthly value added in excess of [Fama and French, 2015], $\hat{S}_i$. The Cross-Sectional mean, standard error, t-statistic and percentiles are the statistical properties of this distribution. Percent with less than zero is the fraction of the distribution that has value added estimates less than zero. The Cross-Sectional Weighted mean, standard error and t-statistic are computed by weighting by the number of periods the fund exists, that is, they are the statistical properties of $\bar{S}_W$ defined by Equation (6). The numbers are reported in $ millions per month. The total period is June 1999 through December 2018 (n = 234), which is split into two distinct periods: PreGFC which is June 1999 through June 2009 (n = 120), and PostGFC, which is June 2009 through December 2018 (n = 114). We require each fund have at least 8 quarters of filing data to be included.

	Total Period	PreGFC	PostGFC
Panel A: Dollar Alphas for Hedge Fund Universe (FF5)			
Cross-Sectional Mean	2.6814	5.6792	1.2588
Standard Error of the Mean	0.2595	0.4891	0.2675
t-Statistic	10.3344	11.6109	4.7056
1st Percentile	-7.8064	-5.4442	-13.1798
5th Percentile	-1.9564	-1.1540	-4.8978
10th Percentile	-0.8220	-0.1589	-2.2235
50th Percentile	0.6399	1.9499	0.1713
90th Percentile	7.5496	14.7353	5.6300
95th Percentile	15.1762	26.1923	10.5848
99th Percentile	35.5121	57.1291	28.3647
Percent with less than zero	27.05%	12.35%	42.57%
Cross-Sectional Weighted Mean	3.6448	7.0120	1.4848
Standard Error of the Weighted Mean	0.2938	0.5445	0.3002
t-Statistic	12.4072	12.8788	4.9452
No. of Funds	939	575	754

Table 6: Summary Statistics for Best Idea Portfolios

FF5 represents $r_t - rf_t = \alpha + bRMRF_t + sSMB_t + hHML_t + pRMW_t + cCMA_t + \epsilon_t$

RM is the return on a value-weighted market portfolio of the CRSP universe of stocks, and RF is the 1-month Treasury bill rate. The construction of SMB and HML follow [Fama and French, 1992]) while RMW and CMA follow [Fama and French, 2015]. Best Idea methods (r_{1-9}) are detailed in Table 1 and composite weights described in Section 4.4. The period is June 1999 through December 2018 (n = 234). T-statistics are found below the point estimates and are adjusted using [Newey and West, 1987] with a lag length of 9 months.

	Mean	$\widehat{\alpha}$	$\widehat{b}$	$\widehat{s}$	$\widehat{h}$	$\widehat{p}$	$\widehat{c}$
				Panel A: Best Ideas			
r_1	0.0074	0.0024	0.9506	0.3278	-0.0686	0.0359	0.0137
		2.42	27.60	7.79	-1.04	0.60	0.11
r_2	0.0070	0.0031	1.1723	0.4983	0.0013	-0.7143	0.0728
		1.80	26.88	10.46	0.01	-5.50	0.50
r_3	0.0083	0.0028	0.9569	0.4612	-0.0578	-0.0107	0.0551
		2.86	26.10	10.90	-0.84	-0.15	0.42
r_4	0.0067	0.0021	1.1904	0.5616	0.0286	-0.6155	0.0835
		1.23	27.96	11.62	0.26	-4.77	0.56
r_5	0.0063	0.0030	1.5061	0.7000	0.1541	-1.3604	-0.0403
		0.83	13.70	6.04	0.63	-6.66	-0.13
r_6	0.0071	0.0030	1.5061	0.6909	0.1877	-1.2169	-0.0000
		0.86	14.07	5.63	0.78	-5.57	-0.00
r_7	0.0084	0.0033	1.038	0.8407	-0.0453	-0.6491	0.2608
		1.68	23.44	12.83	-0.36	-4.92	1.74
r_8	0.0073	0.0025	0.9682	0.2572	-0.1006	0.0561	-0.0318
		3.07	34.26	7.49	-1.81	1.15	-0.26
r_9	0.0071	0.0022	0.9635	0.284	-0.0708	0.0480	-0.0174
		2.55	32.46	7.96	-1.18	0.98	-0.15

Table 7: Summary Statistics for 'Best - Rest' Portfolios

FF5 represents $r_t = \alpha + bRMRF_t + sSMB_t + hHML_t + pRMW_t + cCMA_t + \epsilon_t$

RM is the return on a value-weighted market portfolio of the CRSP universe of stocks, and RF is the 1-month Treasury bill rate. The construction of SMB and HML follow [Fama and French, 1992]) while RMW and CMA follow [Fama and French, 2015]. Composite weights and $spread_i$'s described in Section 4.5, and are based on the Best Idea methods (r_{1-9}) detailed in Table 1. The period is June 1999 through December 2018 (n = 234). T-statistics are found below the point estimates and are adjusted using [Newey and West, 1987] with a lag length of 9 months.

	Mean	$\widehat{\alpha}$	$\widehat{b}$	$\widehat{s}$	$\widehat{h}$	$\widehat{p}$	$\widehat{c}$
				Panel A: Best Idea - Rest			
$spread_1$	0.0001	0.0004	-0.0835	0.0157	-0.0804	0.0878	-0.0598
		0.47	-3.53	0.34	-1.77	1.31	-0.52
$spread_2$	0.0002	0.0017	0.1423	0.1835	-0.0387	-0.7207	0.0633
		1.07	3.31	4.06	-0.40	-5.69	0.46
$spread_3$	0.0011	0.0009	-0.0775	0.1589	-0.0688	0.0321	-0.0101
		1.22	-3.16	3.70	-1.70	0.43	-0.09
$spread_4$	-0.0001	0.0006	0.1601	0.2484	-0.0072	-0.6157	0.0741
		0.40	3.9	6.00	-0.08	-4.93	0.52
$spread_5$	-0.0002	0.0021	0.4627	0.3601	0.1285	-1.3964	-0.0488
		0.57	3.96	3.18	0.57	-6.65	-0.15
$spread_6$	0.0005	0.0021	0.4655	0.3558	0.1602	-1.2473	-0.0092
		0.58	4.08	3.04	0.73	-5.51	-0.03
$spread_7$	0.0013	0.0016	0.0093	0.5177	-0.0667	-0.6483	0.2439
		0.90	0.22	7.41	-0.71	-5.12	1.64
$spread_8$	0.0000	0.0006	-0.0639	-0.0598	-0.1117	0.1103	-0.1083
		0.78	-3.07	-1.53	-2.74	1.95	-1.04
$spread_9$	-0.0003	0.0002	-0.0698	-0.0306	-0.0818	0.1021	-0.0928
		0.21	-3.32	-0.81	-2.09	1.75	-0.93

Table 8: Summary Statistics for 'Best - Rest' Portfolios - Top-m securities

FF5 represents $r_t = \alpha + bRMRF_t + sSMB_t + hHML_t + pRMW_t + cCMA_t + \epsilon_t$

RM is the return on a value-weighted market portfolio of the CRSP universe of stocks, and RF is the 1-month Treasury bill rate. The construction of SMB and HML follow [Fama and French, 1992]) RMW and CMA follow [Fama and French, 2015]. Composite weights and $spread_i$'s described in Section 4.5, and are based on the Best Idea methods (r_{1-9}) detailed in Table 1. The period is June 1999 through December 2018 (n = 234). T-statistics are found below the point estimates and are adjusted using [Newey and West, 1987] with a lag length of 9 months.

	Mean	$\widehat{\alpha}$	$\widehat{b}$	$\widehat{s}$	$\widehat{h}$	$\widehat{p}$	$\widehat{c}$
			Panel A: Top 3 Best Ideas - Rest				
spread_1	0.0001	0.0002	-0.0603	0.0104	-0.0758	0.1190	-0.0409
		0.24	-3.27	0.26	-2.00	2.19	-0.40
spread_2	-0.0000	0.0007	0.1239	0.2259	-0.0799	-0.4398	-0.0276
		0.71	3.78	6.45	-1.09	-6.87	-0.37
spread_3	0.0011	0.0007	-0.0570	0.1620	-0.0541	0.0682	-0.0130
		1.04	-2.59	4.22	-1.67	1.11	-0.13
spread_4	0.0005	0.0007	0.1233	0.2894	-0.0502	-0.3860	-0.0120
		0.76	3.94	9.33	-0.78	-7.27	-0.18
spread_5	-0.0002	0.0019	0.4725	0.2205	0.0978	-1.1057	-0.2307
		0.72	6.61	2.61	0.64	-7.16	-1.05
spread_6	0.0002	0.0019	0.4331	0.2602	0.1081	-1.0033	-0.1982
		0.75	6.47	3.24	0.74	-6.72	-0.94
spread_7	0.0008	0.0010	-0.0147	0.3657	-0.0850	-0.4463	0.2132
		0.82	-0.50	9.61	-1.45	-4.94	2.08
spread_8	-0.0003	0.0000	-0.0593	-0.0493	-0.0917	0.1150	-0.0715
		0.03	-3.65	-1.27	-2.57	2.32	-0.76
spread_9	-0.0004	-0.0002	-0.0517	-0.0259	-0.0823	0.1218	-0.0593
		-0.34	-3.22	-0.71	-2.38	2.46	-0.63

Continuation of Table 8.

	Mean	$\widehat{\alpha}$	$\widehat{b}$	$\widehat{s}$	$\widehat{h}$	$\widehat{p}$	$\widehat{c}$
				Panel B: Top 5 Best Ideas - Rest			
$spread_1$	0.0001	0.0002	-0.0617	0.0011	-0.0768	0.0972	-0.0209
		0.31	-3.70	0.03	-2.17	1.86	-0.21
$spread_2$	0.0002	0.0009	0.1128	0.2142	-0.0723	-0.3825	-0.0991
		0.98	4.18	6.72	-1.09	-7.21	-1.55
$spread_3$	0.0011	0.0005	-0.0392	0.1629	-0.0387	0.0911	0.0009
		0.81	-2.13	5.21	-1.30	1.96	0.01
$spread_4$	0.0007	0.0008	0.1065	0.2903	-0.0483	-0.3168	-0.0491
		0.96	3.95	10.92	-0.88	-7.57	-0.88
$spread_5$	-0.0003	0.0017	0.4398	0.1913	0.0648	-0.9877	-0.221
		0.72	7.08	2.48	0.50	-8.08	-1.22
$spread_6$	0.0003	0.0017	0.4222	0.2117	0.0611	-0.8896	-0.1748
		0.80	7.08	2.90	0.50	-7.80	-1.02
$spread_7$	0.0008	0.0011	-0.0412	0.3448	-0.0800	-0.3662	0.1244
		1.38	-1.66	11.39	-1.44	-7.09	1.88
$spread_8$	-0.0003	0.0001	-0.0620	-0.0596	-0.0987	0.0948	-0.0547
		0.13	-4.03	-1.50	-2.71	1.84	-0.59
$spread_9$	-0.0003	-0.0000	-0.0584	-0.0350	-0.0812	0.0982	-0.0367
		-0.05	-3.91	-0.99	-2.44	1.98	-0.40

Table 9: Summary Statistics for Rest Idea Portfolios

FF5 represents $r_t - rf_t = \alpha + bRMRF_t + sSMB_t + hHML_t + pRMW_t + cCMA_t + \epsilon_t$

RM is the return on a value-weighted market portfolio of the CRSP universe of stocks, and RF is the 1-month Treasury bill rate. The construction of SMB and HML follow [Fama and French, 1992] while RMW and CMA follow [Fama and French, 2015]. Rest Idea methods are described in Section 4.5, and are based on the Best Idea methods (r_{1-9}) detailed in Table 1. The period is June 1999 through December 2018 (n = 234). T-statistics are found below the point estimates and are adjusted using [Newey and West, 1987] with a lag length of 9 months.

	r_1	r_2	r_3	r_4	r_5	r_6	r_7	r_8	r_9
Panel A: Rest Ideas ex Top Best Idea									
Mean	0.0073	0.0072	0.0072	0.0072	0.0073	0.0073	0.0073	0.0073	0.0073
$\widehat{\alpha}$	0.0019	0.0018	0.0019	0.0019	0.0019	0.0019	0.0020	0.0019	0.0019
	3.45	3.45	3.52	3.56	3.51	3.50	3.57	3.41	3.47
Panel B: Rest Ideas ex Top 5 Best Ideas									
Mean	0.0072	0.0073	0.0070	0.0071	0.0073	0.0073	0.0073	0.0074	0.0073
$\widehat{\alpha}$	0.0019	0.0019	0.0019	0.0019	0.0019	0.0019	0.0019	0.0019	0.0019
	3.32	4.00	3.59	4.01	3.57	3.61	3.70	3.24	3.36
Panel C: Rest Ideas ex Top 10 Best Ideas									
Mean	0.0073	0.0073	0.0067	0.0069	0.0073	0.0073	0.0072	0.0075	0.0075
$\widehat{\alpha}$	0.0020	0.0019	0.0018	0.0018	0.0018	0.0018	0.0019	0.0021	0.0021
	3.10	4.36	3.70	4.36	3.40	3.47	3.76	2.96	3.09
Panel D: Rest Ideas ex Top 15 Best Ideas									
Mean	0.0071	0.0073	0.0064	0.0067	0.0072	0.0072	0.0069	0.0075	0.0074
$\widehat{\alpha}$	0.0018	0.0019	0.0017	0.0017	0.0018	0.0018	0.0017	0.0020	0.0020
	2.60	4.38	3.80	4.43	3.14	3.28	3.52	2.53	2.73
Panel E: Rest Ideas ex Top 20 Best Ideas									
Mean	0.0071	0.0069	0.0061	0.0064	0.0072	0.0072	0.0069	0.0077	0.0075
$\widehat{\alpha}$	0.0020	0.0018	0.0016	0.0016	0.0018	0.0018	0.0019	0.0023	0.0022
	2.41	3.99	3.80	4.04	3.21	3.37	3.29	2.29	2.57

Table 10: Summary Statistics for Best minus Rest Idea Portfolios PreGFC vs. PostGFC

FF5 represents $r_t = \alpha + bRMRF_t + sSMB_t + hHML_t + pRMW_t + cCMA_t + \epsilon_t$

RM is the return on a value-weighted market portfolio of the CRSP universe of stocks, and RF is the 1-month Treasury bill rate. The construction of SMB and HML follow [Fama and French, 1992]) while RMW and CMA follow [Fama and French, 2015]. Best Idea methods (r_{1-9}) are detailed in Table 1 and composite weights described in Section 4.4. The period is June 1999 through December 2018 (n = 234). T-statistics are found below the point estimates and are adjusted using [Newey and West, 1987] with a lag length of 9 months. The total period is June 1999 through December 2018 (n = 234), which is split into two distinct periods: PreGFC which is June 1999 through June 2009 (n = 120), and PostGFC, which is June 2009 through December 2018 (n = 114).

	Mean	$\widehat{\alpha}$	$\widehat{b}$	$\widehat{s}$	$\widehat{h}$	$\widehat{p}$	$\widehat{c}$
			Panel A: Best Ideas - Rest (PreGFC)				
spread_1	0.0000	-0.0005	-0.0877	0.0451	-0.0969	0.1264	-0.0771
		-0.28	-2.02	0.71	-1.54	1.50	-0.55
spread_2	-0.0005	0.0032	0.1901	0.1794	-0.0975	-0.6730	0.0973
		1.21	3.15	2.71	-1.01	-3.95	0.62
spread_3	0.0018	0.0007	-0.0908	0.1718	-0.0770	0.0498	-0.0217
		0.48	-2.06	2.90	-1.52	0.50	-0.16
spread_4	-0.0005	0.0021	0.2162	0.2301	-0.0743	-0.5662	0.1197
		0.87	3.78	4.05	-0.82	-3.30	0.75
spread_5	-0.0011	0.0072	0.4750	0.3277	-0.1732	-1.2160	-0.0925
		1.47	2.26	1.87	-1.03	-4.02	-0.26
spread_6	0.0001	0.0070	0.5084	0.3086	-0.1246	-1.0572	-0.0245
		1.47	2.44	1.76	-0.76	-3.32	-0.07
spread_7	0.0048	0.0046	0.0636	0.5768	-0.1509	-0.5592	0.2488
		1.54	0.97	5.76	-1.26	-2.96	1.43
spread_8	-0.0006	-0.0004	-0.0679	-0.0300	-0.1346	0.1440	-0.1107
		-0.28	-1.62	-0.56	-2.11	1.94	-0.87
spread_9	-0.0006	-0.0007	-0.0724	-0.0101	-0.1009	0.1378	-0.1121
		-0.48	-1.77	-0.19	-1.65	1.72	-0.92

Continuation of Table 10.

	Mean	$\widehat{\alpha}$	$\widehat{b}$	$\widehat{s}$	$\widehat{h}$	$\widehat{p}$	$\widehat{c}$
			Panel B: Best Ideas - Rest (PostGFC)				
$spread_1$	0.0002	0.0010	-0.0685	-0.0402	-0.0857	-0.0033	-0.0094
		1.84	-3.88	-1.85	-3.55	-0.10	-0.21
$spread_2$	0.0009	0.0005	0.1059	0.2032	0.0626	-0.6641	0.0148
		0.31	1.76	3.36	0.41	-3.09	0.09
$spread_3$	0.0003	0.0009	-0.0527	0.1183	-0.0803	-0.0613	0.0045
		1.39	-3.58	4.58	-3.85	-2.17	0.10
$spread_4$	0.0002	-0.0004	0.1068	0.3092	0.1029	-0.5496	0.0106
		-0.23	1.98	5.48	0.67	-2.96	0.07
$spread_5$	0.0008	-0.0031	0.5074	0.3780	0.7139	-1.2094	-0.0172
		-0.79	5.71	2.40	3.42	-4.07	-0.06
$spread_6$	0.0010	-0.0026	0.4661	0.4292	0.6306	-1.1860	0.0351
		-0.69	5.40	2.58	2.96	-4.05	0.13
$spread_7$	-0.0023	-0.0021	0.0184	0.3616	0.1082	-0.5617	0.2379
		-1.23	0.33	5.23	1.26	-2.88	1.58
$spread_8$	0.0007	0.0012	-0.0474	-0.1169	-0.0711	0.0688	-0.1240
		2.49	-2.61	-4.69	-2.74	1.83	-2.33
$spread_9$	0.0001	0.0008	-0.0632	-0.0613	-0.0802	0.0367	-0.0285
		1.75	-3.86	-2.99	-3.58	1.13	-0.61

7 Appendix

7.1 On Data Curation

The use of Novus data allows us to expand the analyzable breadth of 13F data to nearly 20 years, versus what is more commonly available since the SEC transitioned its 13F filing data capture to machine-readable formatting. Novus curates an ontology to separate hedge fund filing firms from other entities. This allows our analysis to separate out tens of thousands of filing institutions not deemed hedge fund entities. Finally, in terms of the veracity of the holdings, Novus implements a data sanity check to look for mislabeling (e.g., incorrect CUSIP), incorrect unit denomination (e.g., filed unit millions instead of unit thousands), as well as other common errors. This allows for a rich and indicative dataset of hedge fund firm holdings dating back to 1999. Table 2 shows a representation of the number of unique institutions classified as Hedge Fund Managers, as well as the composite reported market value of their form 13F filings since June 1999.

7.2 On the comparative advantages of 13F data for analyzing hedge fund skill

13F data removes the survivorship and inclusion biases that afflict self-reported databases, as studied by [Malkiel and Saha, 2005] and [Ackermann et al., 1999]. Their estimates range from 100 - 400bps of annual performance inflation as a function of survivorship bias. This bias is neutralized because 13F data is mandated by a regulatory body. As such, our holistic representation of hedge fund managers may be more conclusive than a returns database, where secretive hedge fund managers may choose not to self-elect their returns. Additionally, by proxying a gross return rather than analyzing a net of fees return, we remove issues of return smoothing associated with incentive fee structures. Our analysis is also focused on marketable equities, which allows us to control for inappropriate risk-adjustment in measuring skill. Both of these benefits dampen autocorrelation in aggregate measurement, which makes the calculation of skill challenging through return databases. Beyond the deficiencies in running regressions on return databases, the subjectivity of asset valuation which comprise those returns – as noted by [Getmansky et al., 2004] – is particularly problematic for non-marketable securities; exposure to these assets is more common within hedge fund portfolios, as compared to mutual fund portfolios. Finally, incubation bias, as detailed by [Evans, 2010] may afflict hedge funds as much as mutual funds, thereby impacting the representativeness of return databases.

7.3 Fund Weighted / Asset Weighted Composite Methodology

7.4 Best Idea Portfolio Methodologies

7.4.1 Best Ideas (Methods r_m for $m = 1, 2, 3, 4$)

Initially following [Cohen et al., 2010], we begin with the managers overweight relative to the market portfolio. We let w_{ik} denote i^{th} fund's weight in the k^{th} asset, the tilt factor for r_1 is

$$\lambda_{ik} = w_{ik} - \bar{w}_k \ . \tag{7}$$

A best idea will have a λ_{ik} that is significantly different from zero. In [Cohen et al., 2010] it is argued from the point of view that mutual fund managers have 1 or 2 best ideas, hence the two stocks with the highest λ_{ik} are fund i's best ideas. Method r_2 adds theory to the intuitive motivation and calculates idiosyncratic risk over our standard measurement of CAPM β by taking the mean square error (σ_k^2) of the 60 day regression of excess returns for the security against all securities in set A defined in (1). Again, following [Cohen et al., 2010], we then scale r_1 by this measure of idiosyncratic risk,

$$\lambda_{ik} = \sigma_k^2(w_{ik} - \bar{w}_k) \ . \tag{8}$$

Similarly, r_3 and r_4 are calculated by replacing $\bar{w}_k$ with $\widetilde{w}_{ik}$ in (7) and (8). Ideally, we would subtract the relevant benchmark for each manager from the portfolio weight to determine an active weight, however this information is not available for each active manager. Instead, [Cohen et al., 2010] choose to use the intra-portfolio cap weight to achieve this goal, as shown in formula (2).

7.4.2 Best Ideas through Revealed Beliefs (Methods r_m for $m = 5, 6, 7$)

In [Shumway et al., 2009], the authors' approach inverts the mean-variance formula for the tangency portfolio, which returns an implied beliefs vector $\hat{\mu}_i$ given a shared covariance matrix. The covariance structure is modeled on stock returns using 53 factors, namely the [Fama and French, 1992] factors, adding Carhart's momentum factor, and lastly 49 industry portfolios from Ken French's website. The optimal portfolio weights for a mean-variance investor with risk aversion $\gamma > 0$ is given by

$$w = \frac{1}{\gamma}\Sigma^{-1}(\mu - r) \ . \tag{9}$$

Parameter γ is not readily available to us, but is not required for sorting the unnormalized tilt, r_5,

$$\lambda_{ik} = (\Sigma w_i)_k \ ,$$

for which the top values from the sort are the highest expected returns.

To retain consistency with [Cohen et al., 2010]'s assumptions, we produce Best Ideas through this methodology, relative to the market (r_6) as well as the intra-portfolio-weight (r_7).

7.4.3 Implied Beliefs (Methods r_m for $m = 8, 9$)

Finally, given that hedge funds may be agnostic towards the market as a benchmark given their compensation structures, we test the weight as reported on the manager's filing with no adjustment (r_8), as well as a Tilt metric anchored upon the excess weight of each manager relative to HFU_{AW} (r_9),

$$\lambda_{ik} = w_{ik} - \bar{w}_k^{\text{hf}}.$$

7.5 Portfolio Disaggregation Techniques

7.5.1 Best Ideas

To construct Best Idea Portfolios, we let $B_{im\lambda}$ indicate the top m best ideas of a fund i's portfolio with respect to tilt measure λ

$$B_{im\lambda} := \left\{ \text{portfolio weights of top } m \text{ best ideas in fund } i\text{'s portfolio with respect to tilt metric } \lambda \right\} .$$

The aggregate weight of the stock k in the top m best idea composite, with respect to tilt metric λ, is given by the following

$$w_{km\lambda} = \frac{1}{|HF|} \sum_{\ell \in HF} \frac{w_{\ell k}}{\sum_{z \in B_{\ell m\lambda}} w_z} \mathbf{1}_{\{k \in B_{\ell m\lambda}\}} \qquad \forall k \in A \ , \tag{10}$$

where $\mathbf{1}_{\{\ .\ \}}$ is an indicator function, $|HF|$ represents the number of managers in the set HF and m is the number of top of best ideas considered. Notably, if we are considering the top 1 idea for each manager, $B_{i1\lambda}$, then $\frac{w_{\ell k}}{\sum_{z \in B_{\ell 1\lambda}} w_z} \mathbf{1}_{\{k \in B_{\ell 1\lambda}\}} = 1$.

7.5.2 Best Minus Rest Ideas

In generating spread portfolios, [Cohen et al., 2010] decided to keep the proportion of the short portfolio constant and scale the non-Best Ideas up for each manager. This is done to see if there exists a spread in performance between the top m best ideas and the unperturbed remainder of the portfolio. Thus, the aggregate weight of the stock k in the top m best minus rest composite, with respect to tilt metric λ, is given by the following

$$w_{km\lambda} = \frac{1}{|HF|} \sum_{\ell \in HF} \left[\frac{w_{\ell k}}{\sum_{z \in B_{\ell m\lambda}} w_z} \mathbf{1}_{\{k \in B_{\ell m\lambda}\}} - \frac{w_{\ell k}}{\sum_{x \notin B_{\ell m\lambda}} w_x} \mathbf{1}_{\{k \notin B_{\ell m\lambda}\}} \right] \qquad \forall k \in A \ , \tag{11}$$

where $\mathbf{1}_{\{\ .\ \}}$ is an indicator function, $|HF|$ represents the number of managers in the set HF and m is the number of top best ideas considered.

7.5.3 Rest Ideas

To reinforce these findings, we also test standalone 'Rest Ideas' portfolios by increasing the interval of exclusion for 'Best Ideas'. Doing so furthers the robustness of our Best - Rest findings to ensure our spread results aren't a peculiar function of the top m ranked Best Ideas. This represents the scaled fund-weighted portfolio of the hedge fund universe excluding the top m Best Ideas. We define Rest Idea composite as the following

$$w_{km\lambda} = \frac{1}{|HF|} \sum_{\ell \in HF} \frac{w_{\ell k}}{\sum_{z \notin B_{\ell m\lambda}} w_z} \mathbf{1}_{\{k \notin B_{\ell m\lambda}\}} \qquad \forall k \in A \ , \tag{12}$$

where $\mathbf{1}_{\{\ .\ \}}$ is an indicator function, and $|HF|$ represents the number of managers in the set HF and m is the number of top of best ideas considered.