The Journal of FINANCE
The Journal of THE AMERICAN FINANCE ASSOCIATION

THE JOURNAL OF FINANCE • VOL. LXX, NO. 3 • JUNE 2015

Change You Can Believe In? Hedge Fund Data Revisions

ANDREW J. PATTON, TARUN RAMADORAI, and MICHAEL STREATFIELD*

ABSTRACT

We analyze the reliability of voluntary disclosures of financial information, focusing on widely-employed publicly-available hedge fund databases. Tracking changes to statements of historical performance recorded between 2007 and 2011, we find that historical returns are routinely revised. These revisions are not merely random or corrections of earlier mistakes; they are partly forecastable by fund characteristics. Funds that revise their performance histories significantly and predictably underperform those that have never revised, suggesting that unreliable disclosures constitute a valuable source of information for investors. These results speak to current debates about mandatory disclosures by financial institutions to market regulators.

IN JANUARY 2011, THE SECURITIES and Exchange Commission (SEC) proposed a rule requiring U.S.-based hedge funds to provide regular reports on their performance, trading positions, and counterparties to a new financial stability panel established under the Dodd-Frank Act. A modified version of this proposal was voted for in October 2011, and was phased in starting late 2012. The rule requires detailed quarterly reports (using new Form PF) for 200 or so large hedge funds, those managing over U.S.$1.5 billion, which collectively account for over 80% of total hedge fund assets under management (AUM); for smaller hedge funds, the reports are less detailed, and are required only annually. The rule states clearly that the reports will only be available to the regulator, with no provisions regarding reporting to funds' investors. Nevertheless, hedge funds argued against the adoption of the rule, citing concerns that the government regulator responsible for collecting the reports could not guarantee that their contents would not eventually be made public.[1]

*Patton is at the Department of Economics, Duke University, and Oxford-Man Institute of Quantitative Finance. Ramadorai is at Saïd Business School, Oxford-Man Institute, and CEPR. Streatfield is at Saïd Business School and Oxford-Man Institute. We thank Michael Brandt, Oliver Burkart, John Campbell, Evan Dudley, Cam Harvey, Jakub Jurek, Byoung Kang, Robert Kosowski, Bing Liang, Terry Lyons, Colin Mayer, Narayan Naik, Ludovic Phalippou, Istvan Nagy, Chris Schwarz, Neil Shephard, and seminar participants at the Oxford-Man Institute of Quantitative Finance, the Paris Hedge Fund Conference, Saïd Business School, Sydney University, UC-San Diego, UNC-Chapel Hill, and Vanderbilt University for comments and suggestions. Sushant Vale provided excellent and dedicated research assistance.

[1] See SEC press releases 2011-23 and 2011-226, available at http://www.sec.gov/news/press.shtml. For response from the hedge fund industry, see "Hedge Funds Gird to Fight Proposals on Disclosure," *Wall Street Journal*, February 3, 2011.

DOI: 10.1111/jofi.12240

The economic theory literature almost uniformly predicts that providing more information to consumers is welfare enhancing (an early example is Stigler (1961); see also Jin and Leslie (2003, 2009) and references therein). Hedge funds, however, are notoriously protective of their proprietary trading models and positions, and generally disclose only limited information, even to their own investors. One important piece of information that many hedge funds do offer to a wider audience is their monthly investment performance. This information (as well as information on fund characteristics and AUM)[2] is self-reported by thousands of individual hedge funds to one or more publicly available databases. Under the 3(c)1 and 3(c)7 exemptions to the Investment Company Act, disclosing past performance and fund size to publicly available databases is thought to be one of the few channels that hedge funds can use to market themselves to potential new investors (see Jorion and Schwarz (2010), for example). As a result, these databases are widely used by researchers, current and prospective investors, and the media.

In this paper, we closely examine hedge fund disclosures to these publicly available databases, and provide empirical evidence to underpin the current debate on hedge fund disclosure regulation. We are particularly interested in whether these voluntary disclosures by hedge funds are reliable guides to their past performance. We attempt to answer this question by tracking changes to statements of performance in these databases recorded at different points in time between 2007 and 2011. In each "vintage" of these databases,[3] hedge funds provide information on their performance from the time they began reporting to the database until the most recent period. We find evidence that, in successive vintages of these databases, older performance records (going as far back as 15 years) of hedge funds are routinely revised. This behavior is widespread: 49% of the 12,128 hedge funds in our sample revised their previous returns by at least 0.01% at least once, nearly 30% of funds revised a previous monthly return by at least 0.5%, and over 20% revised a previous monthly return by at least 1%. These are very substantial changes, comparable to or exceeding the average monthly return in our sample period of 0.62%.

While positive revisions are also commonplace, negative revisions are more common and larger when they occur, that is, on average, initially provided returns present a more rosy picture of hedge fund performance than final performance figures. This suggests that prospective investors could be wooed into making decisions based on initially reported histories that are then subsequently revised. Moreover, the revisions are not random. Indeed, we find that information on the characteristics and past performance of hedge funds can predict their propensity to revise. For example, funds-of-hedge-funds and hedge funds in the Emerging Markets style are significantly more likely to

[2] Note that the information provided does not include the holdings or trading strategies of the fund.

[3] This has links to the "real-time data" literature in macroeconomics; see Croushore (2011) for a recent survey.

revise their histories of returns than Managed Futures funds. Larger funds, more volatile funds, and less liquid funds are also more likely to revise.

Several characteristics of revising funds suggest the nature of incentives that may drive revising behavior. For example, a fund experiencing a change in management company or manager is 10% more likely to revise its past returns, holding all else constant. Following such events, we hypothesize that new management might be interested in a "fresh start," revamping the accounting, marking-to-market, auditing, and compliance practices of their newly acquired funds, thus resulting in a sequence of revisions to past returns.[4] Another important characteristic associated with revising behavior is the presence of a high-water mark in the fund. Managers may have greater incentives to revise past returns *downwards* (or simply to correct previous valuation errors only in the positive direction) when they are well below their high-water marks, so as to reset the level at which they begin earning performance fees. Consistent with this explanation, we find that funds with a high-water mark are 13% more likely to revise than those without a high-water mark. Moreover, when funds with a high-water mark revise returns, their average return revision is −62 basis points. In contrast, funds without a high-water mark provision have average return revisions of +40 basis points. This allows for a refinement of our finding that the unconditional average return revision is negative: funds with an incentive to revise returns below high-water marks revise *downwards* on average, whereas funds without high-water marks revise returns *upwards*, making past returns appear higher in subsequent revisions.

To provide a concrete example of the sort of revising behavior to which we refer, consider the (anonymized but true) case of Hedge Fund X, which was incorporated in the early 1990s. The fund began reporting to a database four months following inception, and a year after inception it reported AUM in the top quintile of all funds. In the mid 2000s, the fund experienced a troubled quarter and saw its AUM halve in value. It then ceased reporting AUM figures. The fund's performance recovered, and during the last quarter of 2008 it reported a particularly good double-digit return, putting it in the top decile of funds. However, a few months later this high return was revised downward significantly, into a large negative return. A similar pattern emerged later that year, when a previously reported high return was adjusted substantially downward in a later vintage, along with two other past returns. A further sequence of poor returns was then revealed, and the fund was finally reported as closed in mid 2009.

The example provided above suggests that revisions might be useful signals of fund quality to investors, that is, they may reflect adverse selection problems embedded in voluntary disclosures of financial information. It is also possible, of course, that revisions are innocuous despite being systematically associated with particular fund characteristics. For example, they may simply be corrections of earlier mistakes, and therefore contain no information about future

[4] While this may be well intentioned, any such changes to preexisting practices may also indicate the presence of poor preexisting operational controls within the fund.

fund performance. However, such corrections would have to be substantial, as we find that simple errors such as digit transpositions and decimal point errors make up only a negligible fraction of the revisions observed in our sample.

To further investigate the information content of revisions, at each vintage of data we categorize hedge funds into those that have revised their return histories at least once (revisers) and the remainder (nonrevisers). We find that the future performance of revising funds is significantly lower than that of nonrevising funds, and that there is a far greater risk of experiencing a large negative return when investing in a revising fund. Moreover, we find that revisers are significantly more likely to cease reporting to a database, a signal that is correlated with liquidation.

Put differently, this analysis reveals in real-time that funds with unreliable reported returns are likely to underperform in the future. This finding is virtually unchanged by adjusting for risk using various models, is not greatly affected by varying the size threshold for detecting significant revisions, is stronger for revisions pertaining to periods far back in time and for funds with higher levels of asset illiquidity, and is robust to various other changes in parameter values. Results from these robustness checks also show that performance differentials between revisers and nonrevisers are higher for more illiquid funds, but they are by no means restricted to these funds.

Our analysis suggests that mandatory, audited disclosures by hedge funds, such as those adopted by the SEC in 2011, could be beneficial to investors and not just regulators, and contributes to a growing list of examples highlighting the benefits of an independent auditor or regulator for financial institutions. For example, Daníelsson et al. (2001) note that, under Basel II, European banks were given a choice: they could use a standardized model to measure their risk exposures (used in setting their capital requirements) or they could use their own in-house models. The in-house models were subject to audit by the banking regulator, but, owing to the complexity of each bank's models, it is not clear whether the regulator could properly monitor their effectiveness. After the financial crisis, it was noted both in the press and the finance literature that these models appear to have underestimated the true risk of many banks' positions.

The remainder of the paper is structured as follows. In Section I, we review related literature. In Section II, we describe the data and explain how we determine revisions. Section III outlines our methodology. We present our main empirical results in Section IV and robustness checks in Section V. Section VI concludes. An Internet Appendix contains additional analyses. [5]

I. Related Literature

Several previous authors have noted problems with self-reported hedge fund returns. The fact that hedge fund managers voluntarily disclose returns to

[5] The Internet Appendix is located in the online version of the article on the *Journal of Finance* Web site.

hedge fund databases means that they are able to choose if and when to start reporting, and when to stop reporting. This leads to substantial biases not seen in traditional data sets, such as listed equities or registered mutual funds. Ackermann, McEnally, and Ravenscraft (1999), Fung and Hsieh (2000, 2009), and Liang (2000) provide an overview of these biases such as survivorship, self-selection, and backfill.

Self-reporting also leads to the possibility of using different models to value assets, as well as to the possibility of earnings smoothing. For example, Getmansky, Lo, and Makarov (2004) document high serial correlation in reported hedge fund returns relative to other financial asset returns, and consider various possible explanations for this observation, including underlying asset illiquidity. Asness, Kraill, and Liew (2001) note that the presence of serial correlation leads reported returns to appear less risky and less correlated with other assets than they truly are, thus providing an incentive for hedge fund managers to intentionally "smooth" their reported returns, a form of earnings management for the hedge fund industry. Cassar and Gerakos (2011) match due diligence reports with smoothing measures, and find that smoother returns are associated with managers who have greater discretion in sourcing the prices used to value the fund's investment positions. Bollen and Pool (2008) extend Getmansky, Lo, and Makarov (2004) to consider autocorrelation patterns that change with the *sign* of the return on the fund, with the hypothesis being that hedge fund managers have a greater incentive to smooth losses than gains; they find supportive evidence. This finding is reinforced using a different approach in Bollen and Pool (2009), who document that there are substantially fewer reported monthly returns that are small and negative than one might expect. When aggregating to bimonthly returns no such problem arises, suggesting that the relative lack of small negative returns in the data is caused by temporarily overstated returns. Jylhä (2011) extends Bollen and Pool's (2009) work on misreporting by conditioning the search for pooled distribution discontinuities on various fund attributes. In a recent study, Bollen and Pool (2012) propose a variety of "flags" for potential fraudulent activity based just on reported returns, and link these to an indicator for whether the fund has been charged with legal or regulatory violations.

Agarwal, Daniel, and Naik (2011) find evidence that hedge funds tend to underreport returns during the calendar year, leading to a spike in reported returns in December that cannot be explained using risk-based factors (a similar result for quarter-end returns for mutual funds can be found in Carhart et al. (2002)). The motivation for doing so is that hedge funds are paid incentive fees once a year based on annual performance. At higher frequencies, Patton and Ramadorai (2012) find that estimated hedge fund risk exposures appear to be highest at the beginning of the month, and lowest just prior to end-of-month reporting periods.

Others have looked at 13-F filings by hedge funds to uncover evidence of unreliable voluntary disclosure, for example, Cici, Kempf, and Puetz (2011) find evidence that these filings often appear to be valued at prices different from prevailing closing prices in CRSP, Ben-David et al. (2011) present evidence that hedge funds appear to increase holdings of illiquid stocks at critical reporting valuation dates, and Agarwal, Daniel, and Naik (2011) find that hedge funds are the greatest users of confidentiality provisions to delay reporting of sensitive positions in 13-F filings. Ben-David et al. (2013) While our paper is related to this stream of research, the new empirical phenomenon we document might be better labeled "history management"—with closer parallels to earnings restatements than to earnings management (see Dechow, Geb, and Strand (2010) for a comprehensive review of the accounting literature on the subject).

The literature on hedge funds has also considered the role of mandatory disclosures for hedge funds. For a unique, and brief, period in 2006 before the rule was vacated, the SEC required hedge funds to disclose a variety of information such as potential conflicts of interest and past legal and regulatory problems. These Form ADV disclosures were designed to deter fraud, or control operational risk more generally. Brown et al. (2008, 2012) report evidence that these mandatory disclosures of information related to operational risk were beneficial to investors. The authors find that the information in these disclosures enabled investors to select managers that went on to have better performance, and that conflicts identified in the Form ADV filings were correlated with other flags for operational risks.

Our analysis of changes in the reported histories of hedge fund returns is also related to Ljungqvist, Malloy, and Marston (2009), who study changes in the I/B/E/S database of analysts' stock recommendations. These authors document that up to 20% of matched observations are altered from one database to the next, using annual vintages of the I/B/E/S database from 2001 to 2007. Like us, they find that these revisions are not random: recommendations that were further from the consensus, or from "all star" analysts, were more likely to be revised than others, and undoing these changes reduces the persistence in the performance of analyst recommendations. While the focus of these authors is primarily to illuminate problems of replicability in academic research, our concerns run deeper on account of the environment of limited disclosure for hedge funds. This environment generates a greater reliance on self-reported hedge fund data. We demonstrate that hedge fund return revisions could skew allocations by investors reliant on the initial return presented. Moreover, the significantly lower future returns and greater downside risks in troubled times experienced by funds with unreliable disclosures suggests the issue that we identify represents a source of risk to hedge fund investors, and quite possibly a broader systemic risk.

Finally, it is worth noting that information on the trading strategies and positions of hedge funds also has implications for how they are compensated. Foster and Young (2010) show theoretically the difficulty of devising a performance-based compensation contract for hedge fund managers that rewards skilled managers but not unskilled managers. With only return histories made available for performance evaluation, unskilled managers can mimic skilled managers arbitrarily well simply by taking on an investment with a small probability of a large crash. Foster and Young (2010) argue that transparency of positions, not just performance, is needed to separate skilled managers from unskilled managers.

II. Data

A. Consolidated Hedge Fund and Fund-of-Fund Data

We employ a large cross-section of hedge funds and funds-of-funds over the period from January 1994 to May 2011, which is consolidated from data in the TASS, HFR, CISDM, Morningstar, and BarclayHedge databases. Appendix A contains details of the process followed to consolidate these data. The funds in the combined database come from a broad range of vendor-classified strategies, which are consolidated into 10 main strategy groups: Security Selection, Macro, Relative Value, Directional Traders, Funds-of-Funds, Multi-Process, Emerging Markets, Fixed Income, Managed Futures, and Other (a catch-all category for the remaining funds).[6] The set contains both live and dead funds. Returns and AUM are reported monthly, and returns are net of management and incentive fees.

B. Hedge Fund Database Vintages

Hedge fund data providers update their databases from time to time. These updates not only include the incremental changes since the previously published version, but also the entire history of returns for each fund. This allows us to compare reported histories across vintages of these databases at various points in time. We compare a total of 40 vintages of the different databases between July 2007 and May 2011.[7] At each of these vintages $v \in \{1, 2, \ldots 40\}$, we track changes to returns for all available databases. Not every database is updated with the same periodicity, and in those cases the newer vintage is simply set to the previous one, thus forcing zero detected changes.

We apply standard filters to the data before analysis. First, we remove 82 funds with very large or small returns to eliminate a possible source of error (truncating between monthly return limits of -90% and $+200\%$).[8] Second, we remove 186 funds that report data only quarterly. Third, we remove funds with insufficient return histories (less than 12 months) and missing fund level data (such as no "Strategy" or "Offshore" indicators recorded). Fourth, as less than one-third of Morningstar funds passed these quality filters, we remove the remaining 832 Morningstar funds to ensure sufficient depth by database. The final cleaned data set contains 18,382 unique hedge funds. Of these funds,

[6] The mapping between these broad strategies and the detailed strategies provided in the databases is reported in the Internet Appendix.

[7] Vintages were collected in July 2007, and then monthly from January 2008 to May 2011, with February and November 2009 omitted due to data download errors.

[8] Although -100 would be a natural choice, we used -90 to specifically remove cases in which data providers use large negative returns as place holders for missing observations.

Table I

Summary Statistics: Data Set

This table shows summary statistics for our sample of hedge funds, with time-series statistics in Panel A computed using only the May 2011 (final) vintage of the 40 vintages of data that we capture. AUM refers to assets under management. Panel A shows broad statistics on returns and AUM, Panel B shows the strategies into which the funds are classified, and Panel C shows the databases from which the funds are sourced.

Panel A: Fund Summary Statistics

Num. Funds	Average Fund AUM U.S.$MM	Average Fund Return	Average Fund History Length (years)
12,128	138.25	0.618	6.133

Panel B: Fund Strategies

	Fund Count	Count%
Security Selection	1,762	14.53%
Macro	685	5.65%
Relative Value	191	1.57%
Directional Traders	1,503	12.39%
Funds-of-Funds	3,822	31.51%
Multi-Process	1,371	11.30%
Emerging Markets	612	5.05%
Fixed Income	597	4.92%
Managed Futures	1,444	11.91%
Other	141	1.16%
Total	12,128	100.00%

Panel C: Funds by Database

	Fund Count	Count%
TASS	4,585	37.81%
HFR	2,983	24.60%
CISDM	1,106	9.12%
BarclayHedge	3,454	28.48%
Total	12,128	100.00%

12,128 report returns to two or more vintages of our databases, and these funds comprise the final sample of hedge funds that we employ in our analysis.

Table I shows descriptive characteristics of the sample of 12,128 funds. (A corresponding table for the complete set of 18,382 funds is available in the Internet Appendix.) On average, funds report for six years, have U.S.$138 million in assets, and generate returns of approximately 0.62% per month. Just under one-third of them are Funds-of-Funds, with Security Selection and Directional Traders being the predominant hedge fund strategies represented in the data. Approximately one-third of the funds are from the TASS database, with the

CISDM database accounting for the smallest share of the four databases represented in our final sample, at just under 10% of funds.

C. Changes: Revisions, Deletions, and Additions

We compare return histories across successive vintages and group changes into three categories: additions, deletions, and revisions. To shed light on these categories, consider $Ret_{i,t,v}$, the return for fund i at time t reported in vintage v of the database. We drop i and t for ease of exposition, and let $v-1$ indicate the previously available vintage for the database in which the fund's data were reported (this may not necessarily be immediately one vintage prior as not all databases update simultaneously).

An "addition" implies that a return is added to the fund's history in a later vintage, that is, Ret_{v-1} was not in the database, but Ret_v is present. Clearly there are legitimate circumstances in which this would happen, such as when a new fund launches, or when new return updates are provided for months between dates when the two vintages were captured. To rule these cases out when counting additions, we exclude all fund launches (i.e., cases in which the entire fund history appears in a vintage), and exclude return months within 12 months of the prior vintage $v-1$ (to avoid picking up late reporting).[9] In contrast, "deletion" implies that a return goes missing between vintages, that is, Ret_{v-1} was reported but Ret_v was not.

The third category, "revisions" consists of cases in which both Ret_{v-1} and Ret_v are available but are not equal to each other. These revisions constitute the main focus of our analysis. As mentioned above, we filter out small changes (less than one basis point) that may be attributable to rounding, and for our main analysis we focus on revisions related to returns that are over three months old, and do not count as revisions those pertaining to more recent returns. The motivation for this filter is that most hedge fund databases report returns that are net of fees, and since hedge fund fees are most often linked to performance, recent returns may be subject to innocuous revisions arising from this source. We discuss this difference further below.

Table II shows the prevalence of these three types of changes to funds' return histories. Fully 49% of the 12,128 fund sample has one of the three types of changes described above (labeled "Any Change"). Of these, revisions of preexisting data are the most frequent, at 45%, followed by deletions at 8%, and additions at 3%. (Some funds have multiple types of changes, and so the sum

[9] For example, consider the case in which vintage $v-1$ for a fund was captured in June 2009, and this vintage shows fund histories up to February 2009. The next vintage v is captured in August 2009 and this vintage shows fund histories up to July 2009. We would disregard any additions of data occurring after June 2008 when computing additions for this fund. So, for example, if March 2009 and April 2009 returns are missing in $v-1$ but present in v, these months would not be counted as additions, to ensure that we do not capture late updates of returns by the fund's manager to the database provider. Our focus for additions is on backfilling of past history rather than short-term lags in fund reporting. See Aragon and Nanda (2011) on strategic reporting delays for poor returns.

Table II

Summary Statistics on Return Changes across Vintages

This table shows summary statistics of changes in returns (additions, deletions, and revisions) between successive vintages. Panel A shows counts of the three types of changes separately, as well as “Any Change.” Panel B shows the proportion of revising funds with at least one revision that is at least as large as the size thresholds listed. Panel C shows various percentiles of (positive, negative, and net) revisions, and their absolute values. Panel D shows the proportions of revising funds with at least one revision that relates to a return that is at least as old as the “recency” thresholds listed. Panel E explores potential reasons for innocuous revisions, namely, sign changes, decimal errors, and digit transpositions.

Panel A: Changes Breakdown at Fund Level

	Fund Count	Any Change Count	Deletions Count	Additions Count	Revisions Count
Funds	12,128	5,938	976	363	5,446
% of Funds		49.0%	8.0%	3.0%	44.9%

Panel B: Size of Revisions

		Revisions Count			
	Fund Count	at least 0.01%	at least 0.1%	at least 0.5%	at least 1%
Funds	12,128	5,446	4,718	3,363	2,581
% of Funds		44.9%	38.9%	27.7%	21.3%

Panel C: Summary Statistics for the Distribution of Revisions

	Revisions	Absolute Revisions	Positive Revisions	Negative Revisions
Count	63,791	63,791	31,039	32,752
Mean	−0.029	0.908	0.904	−0.912
Median	−0.020	0.140	0.140	−0.140
95th perc	1.860	3.800	3.776	−0.020
5th perc	−1.957	0.020	0.020	−3.816

Panel D: Recency of Revisions

		Minimum Recency of Revisions Count			
	Fund Count	1 or More Months	More than 3 Months	More than 6 Months	More than 12 Months
Funds	12,128	6,891	5,446	4,340	3,423
% of Funds	100.0%	56.8%	44.9%	35.8%	28.2%
Revisions		87,461	63,791	51,426	43,192
% of Revisions (base)		137.11%	100.00%	80.62%	67.71%

Panel E: Potentially Innocuous Revisions

	Reviser Count	Sign Change	Decimal Place	Digit Transposition	Sign, Decimal, or Transpose
Funds	5,446	154	63	211	390
% of Funds	44.9%	1.27%	0.52%	1.74%	3.22%
Revisions	63,791	179	405	250	834
% of Revisions	100%	0.28%	0.63%	0.39%	1.31%

of the individual categories is greater than the "Any Change" proportion.) The large percentage of funds with revisions demonstrates that this is a widespread problem: funds that have had at least one change in their reported history manage around 46% of the average total AUM in the hedge fund universe (this number peaks at $1.8 trillion in June 2008).

Panels B and C of Table II report summary statistics on the size of revisions in our sample. We observe that 45% (6,906 funds) of funds revise their returns at least once by at least one basis point, while 28% of funds revise at least once by at least 50 basis points. Panel C reveals that the mean absolute revision is 91 basis points. To provide an appropriate comparison, the mean monthly return across hedge funds is 62 basis points, as reported in Table II, that is, lower than the mean absolute monthly revision. The revisions that we detect are therefore substantial.

Panel D of Table II reports on the "recency" k of the revisions that we detect in our data, defined as the difference between the date of the return and the date at which a revision was detected. For example, if the return for January 2008 was revised between the December 2008 and January 2009 vintages of data, then this revision would have $k = 12$ months. Each of the columns in Panel D shows the proportion of revising funds remaining once we exclude revisions near the vintage date (e.g., our main analyses are for $k > 3$, where we ignore revisions of returns that occur within three months of the date of the return). As we increase k, the proportion of funds that are flagged as having revised their returns declines, from 57% in total before any k filter is imposed, down to 28% when we ignore any revision within a year of the return date. Almost one half of the return revisions in our sample relate to returns that are more than 12 months in the past. Presaging results from later in the paper, it seems unlikely that these revisions are merely corrections of data entry errors, or a simple consequence of illiquid positions being marked-to-market.

Panel E of Table II attempts to determine whether the revisions that we find in our data are mainly attributable to common data entry errors. We consider three such errors: sign changes (where the revised return is identical to the original return except for the sign), decimal place errors (where the revised return differs from the original return by a factor of 0.01, 0.1, 10, or 100), and transposition errors (where adjacent digits in the original return are transposed in the revised return). We find that these contribute only a negligible fraction of the observed revisions—only 3.2% of funds have one of these types of errors, compared with the 44.9% of funds that have revised their returns at least once. Thus, these common types of data entry errors do *not* appear to be the primary source of the return revisions that we uncover in our data.

Figure 1 provides a different view of the prevalence of revisions. Panel A of the figure shows that, conditioning on a fund experiencing at least one revision, roughly 20% of the funds revise returns in fewer than 5% of the total number of vintages in which they are present in the data set. However, a little over a quarter of the funds in the sample revise returns in 10% to 20% of the vintages in which they are present, and roughly 12% of the sample funds revise returns in over half of the vintages in which they are present. Panel B of the figure

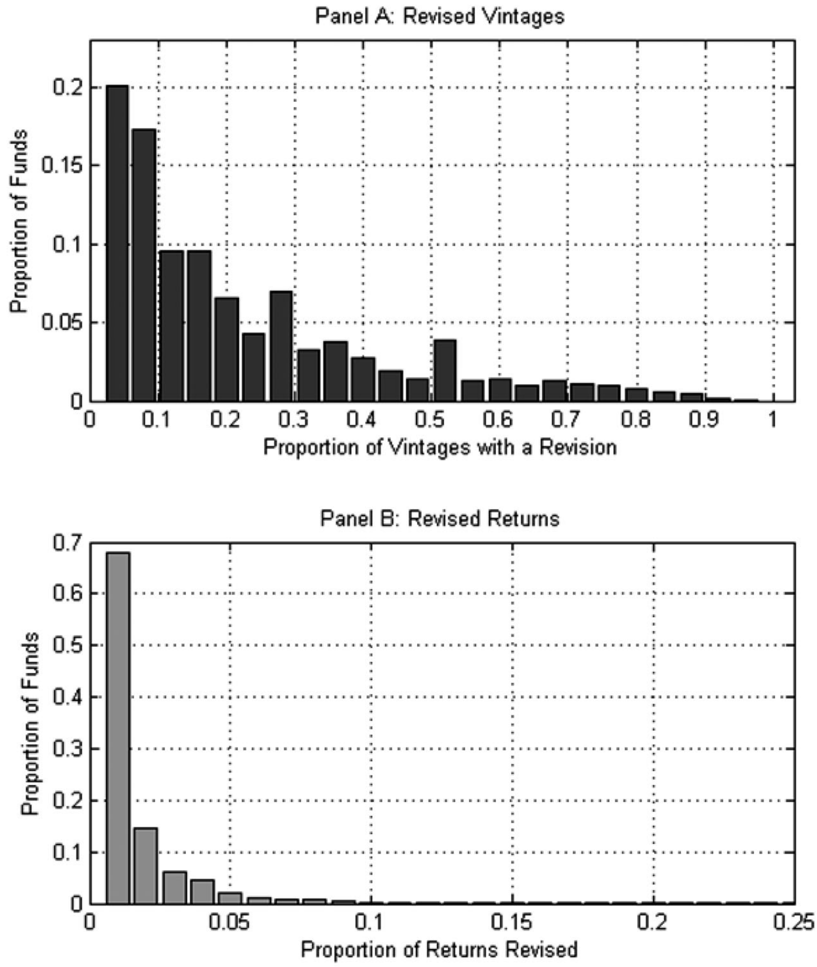

Figure 1. Histograms of the prevalence of revisions, across funds with at least one revised return. Panel A presents a histogram, across funds, of the proportion of vintages in which a given fund appears that contain at least one revised return. The left-most bar indicates that 20% of funds have revised at least one return in 4% of the data vintages in which they appear. Panel B presents a corresponding histogram for individual monthly returns that have been revised. In both panels only funds that have revised at least one return are included in the sample.

shows that the revisions often pertain to a relatively small fraction of the total available return history of the fund: roughly 70% of the funds revise 1% or less of their total return histories. However, there is a relatively long tail to this distribution, with a small fraction of funds revising over 5% of their total history and a handful of funds revising over 15% of their total return history.

In the Internet Appendix, we also show the prevalence of return revisions by strategy, which reveals that, while there is a degree of heterogeneity across strategies, even relatively liquid strategies such as Managed Futures and

Global Macro have a substantial fraction of funds that revise data. We study the determinants of revising behavior, including strategy affiliations, in more detail using a probit model as described in Section III A.

D. Hedge Fund Return Factors

To make appropriate risk adjustments in analyzing portfolio performance for the revising and nonrevising funds, we calculate alphas via the widely used Fung and Hsieh (2001) seven-factor model for hedge fund returns. The Fung–Hsieh factors have been shown to have considerable explanatory power for hedge fund and fund-of-fund returns. They comprise four market-related factors, namely, an equity market factor (S&P 500), an equity size factor (Russell 2000 less S&P 500), a bond market factor that uses a constant-maturity-adjusted 10-year Treasury bond yield less the short-term Treasury rate, and a bond credit-spread factor using the change in Moody's BAA credit spread over a constant-maturity-adjusted 10-year Treasury bond yield, as well as three trend-following strategy factors formed from excess returns on portfolios of lookback straddle options for bonds (PTFSBD), currencies (PTFSFX), and commodities (PTFSCOM).[10] We use tradeable versions of the bond market and bond credit spread factors to facilitate cleaner interpretations of the alpha in these models. In robustness checks we also add an eighth factor to the Fung–Hsieh set, namely, MSCI Emerging Market index returns, and we employ the Fama–French–Carhart and Pástor–Stambaugh models as alternative risk adjustment models.

III. Methodology

We begin by documenting the characteristics of funds prone to return history changes, focusing our analysis on the most prevalent category of changes, namely, revisions. This analysis helps us shed light on the incentives for funds to engage in revising behavior. We then analyze the determinants of the size and sign of revisions, documenting the differences between initially perceived and final histories. This analysis improves our understanding of how an investor using the database would see different pictures of hedge fund performance if he or she had employed different vintages of the data. Finally, we form portfolios of reviser and nonreviser funds to determine the information content of revisions for future performance and shortfalls.

A. Which Funds Revise?

We estimate a fund-vintage-level probit regression. The dependent variable is a revision indicator, $Rev_{i,v}$, for fund i at vintage v, which takes the value of

[10] Data for the trend-following factors can be found on David Hsieh's Web site (http://faculty.fuqua.duke.edu/∼ dah7/HFRFData.htm). Datastream and the Federal Reserve Web site are sources for the equity and bond factors, respectively.

one for any fund that experiences a revision of returns between two successive vintages of data, and zero otherwise. The explanatory variables include a lag of the dependent variable to investigate whether revisions are autocorrelated across vintages, that is, whether funds that have revised returns in the past are likely to do so again in the future, and a number of fund characteristics measured at vintage $v-1$, which are described below, and collectively denoted by the vector $X_{i,v-1}$: [11]

$$Rev_{i,v} = \alpha + \gamma Rev_{i,v-1} + X'_{i,v-1}\beta + u_{i,v}. \tag{1}$$

The vector $X_{i,v-1}$ of right-hand-side variables includes several variables based on return and AUM histories. First, we employ AUM to study whether changes are more likely to occur for larger or smaller hedge funds, ranking funds by their AUM computed using vintage $v-1$.[12] Second, we use the average of all available returns and recent (past 12-month) returns for each fund, again computed using data from vintage $v-1$, to capture whether weaker performing funds resort to changes to recast their histories. Third, we use the standard deviation of all available returns to capture whether funds with more volatile returns experience pressure to delete or recast disappointing performance. Fourth, we use a measure of return smoothing suggested by Getmansky, Lo, and Makarov (2004), namely, the first-order autocorrelation coefficient of all available returns.

In all cases in which we employ cross-sectional ranks, these are standardized between zero and one. Fifth, we include a variable that computes the number of returns in fund i's history up to vintage v. This variable controls for the purely mechanical possibility that, if there is a small fixed chance of a data capture error, then a longer return history provides more exposure to return revisions. Of course, this variable also measures the age of a fund, so it has multiple interpretations.

In addition to these variables computed from return and AUM histories, we include strategy fixed effects in our specifications to control for the possibility that differences in volatility and liquidity occasioned by the use of these different strategies, as well as differential access to information about these strategies (e.g., underlying returns for obscure investments by Emerging Markets funds may be difficult to independently verify) might lead to differences in the propensity to alter data. We further include database fixed effects since the controls (e.g., the verification of returns preloading) implemented by each database vendor may vary, and thus influence the propensity for changes. Next, we employ an indicator for whether the fund is offshore or onshore, as funds in

[11] Standard errors are clustered by vintage to control for the possibility that there are certain periods in which unexplained revisions are more likely to be prevalent. The Internet Appendix also presents results that explain the prevalence of additions, deletions, and "any change," a catch-all category encompassing all three types of changes.

[12] We do not attempt to match the timing of AUM revisions to return revisions in this paper. However, Streatfield (2012), using a slightly different sample, finds AUM revisions occur at a similar incidence to returns (36.5% on 18,381 funds).

offshore jurisdictions may be subject to less scrutiny, and we condition on the lockup restrictions imposed by the fund on its investors, as these restrictions provide liquidity safeguards for the fund manager but also may allow managers to hide from the reputational consequences of changing data within the lockup period. We also include an indicator for whether the fund has a hurdle rate provision, or any audit information available in the database.[13]

Finally, two fund characteristics deserve special mention, as they help us better understand the incentives for fund managers to revise returns. The first is a dummy variable that indicates whether a given fund experienced a change of management company or a change of manager. Inclusion of this variable allows us to explore the possibility of an "operational risk" (in the sense of Brown et al. (2008)) explanation for revisions, focusing specifically on mergers, management changes, and fund takeovers.[14] Following such events, we hypothesize that new management might potentially be interested in a "fresh start," revamping the accounting, marking-to-market, auditing, and compliance practices of their newly acquired funds.

The second characteristic we include is a dummy that takes the value of one if a fund has a high-water-mark provision. This variable relates to a second possible explanation for revisions, namely, the potential reduction in high-water marks associated with retrospective negative return revisions. Managers may have greater incentives to revise past returns downwards when they are well below their high-water marks, so as to reset the level at which they begin earning performance fees. We defer further discussion of these variables to our discussion of the estimation results.[15]

B. Determinants of the Size and Direction of Revisions

Having determined which funds revise, we next turn to understanding the impact of revising history on the historical performance record of funds. We do so by comparing the initially reported return for fund i in month t with the same fund-month return seen in the last database vintage in which it appears. This analysis attempts to answer the following question: if an investor only looked at a return expressed by the fund's portfolio manager the first time it was made public, how does this differ from what the investor might see in the database at the last available vintage?

[13] Underlying databases differ in the types and level of information they provide, with some providing the date of last audit, others providing annual audit flags, and yet others providing auditor names. Our indicator takes the value of one if any audit information is available for the fund, and zero otherwise. The Internet Appendix contains descriptive statistics for several of these variables.

[14] Brown et al.'s (2009) "omega" score is an indicator of operational risk associated with fund failure, and is correlated with fund characteristics. Problem funds in their study used less leverage, had a lower incidence of high-water marks, and had lower lockups. In contrast, our revisers have higher lockups, and, in particular, higher high-water marks.

[15] A theoretical model of the "optimal" amount of misreporting, in terms of the incentives to honestly report versus those to over- or underreport returns, may shed some light on the trade-offs managers face, and is left for future research.

In particular, we condition the return differences occasioned by revisions on various fund characteristics and period fixed effects. The dependent variable in these regressions is the average difference, for all years in which a fund experienced return revisions, between the final set of annual returns provided by a fund and the first set of annual returns provided for the same fund-years. For example, if a fund initially reported 6% average annual return for year t, and at the final vintage this average stood at 4%, then the return difference variable would be -2%.

In these specifications, we only include periods in which the fund had at least six months of return observations to reduce noise in the dependent variable. We explain both the absolute value of all such differences as well as the signed revisions on the independent variables. Period dummies include crisis dummies for the 1998–1999 period, the 2000–2001 period, and the 2008–2009 period. Several of the remaining regressors are described earlier, with three new additions, namely, the rank of flows experienced by the fund relative to all other funds in the same year, the management fee, and the incentive fee of the fund.

C. Are Revisions Informative about Future Performance?

Our final question is whether knowing that a fund has revised its past performance conveys useful information about its future performance. The null hypothesis here is that these revisions are innocuous and provide no information about future returns. One alternative is that they are an indicator of either poor operational controls or of dishonesty, both of which provide negative information about revising funds (as in Brown et al. (2008)). Another possibility is that revisions are a sign of honesty, in the sense that revisers "fess up" to past mistakes. In this case, we might expect performance to be higher for revisers than nonrevisers.

To consider these hypotheses rigorously, we employ two methods to determine the performance differentials between revising and nonrevising funds. Our first approach is to form portfolios of the returns of funds based on their revising behavior, allocating funds to one of two groups: "reviser" funds that have revised at least once, and "nonreviser" funds that have had no revisions up until a given vintage. At the first vintage, by definition, all funds are nonrevisers. At each subsequent vintage, once we observe revising behavior, we allocate funds into these two groups, moving several funds from the nonreviser portfolio to the reviser portfolio at each step. Once a fund is categorized as a reviser, we track all its subsequent returns in the reviser portfolio.

Note that this is a real-time strategy. Consider the example of a fund making its first ever return revision, say of its previously reported January 2007 return, in the August 2008 database vintage. Once we detect this historical return revision, we immediately classify the fund as a reviser. The reviser portfolio will then include the fund's returns from September 2008 until the end of our sample period, and the nonreviser portfolio will no longer track its returns from September 2008 onwards. Thus, at each time period, the nonreviser portfolio contains funds that have never revised data in any previous vintages, although it could contain funds that are yet to be identified as revisers. Within each portfolio, we weight all monthly returns of funds equally, computing a time series of portfolio returns.[16] We can then look at whether there are differences in the returns of reviser and nonreviser portfolios, and risk-adjust these return differences in various ways.

We also use the cessation of reporting to a database as a sign of future performance—a key, though not the sole, reason for this is fund liquidation. We compute the liquidation probabilities for revisers and nonrevisers, at horizons ranging from 6 to 30 months. Given the turbulent period that our sample covers, we compute these probabilities starting from six different dates (June 2008–December 2010).[17]

IV. Results

A. Which Funds Revise?

Table III shows the results of estimating the probit regression equation (1) for revisions. (The results for other change types, including whether a fund made any one of the three different types of changes, can be found in the Internet Appendix.) These regressions present the marginal effects of each continuous right-hand-side variable, that is, the change in probability in the dependent variable that results from an infinitesimal change in each of these variables. For dummy variables, such as *Offshore*, the effect is captured for the discrete change in the variable from zero to one.

Table III reveals that asset size, prior-year return rank, and return autocorrelation are significantly positive determinants of a fund's propensity to report a change in history.[18] Ang, Gorovyy, and van Inwegen (2011) show that hedge fund leverage is negatively related to fund return volatility and size. Taken together with the results from the probit, this suggests that leverage is very likely lower for funds with a greater propensity to revise. This evidence appears quite similar to the finding in Brown et al. (2008) that leverage is lower for "problem" funds than for "nonproblem" funds.

The indicator for whether the fund revised returns in the previous vintage is highly significant, revealing that some funds are regular revisers of their returns. The number of returns present for a fund has a significant effect on the propensity to make a revision, although this could simply be a mechanical effect as described above. Turning to the strategy indicators, Funds-of-Funds show the highest chance of reporting changes, which is perhaps unsurprising, as Fund-of-Fund performance numbers are a function of underlying hedge fund

[16] In Section V C we use the median returns of the reviser and nonreviser funds to address concerns about outliers driving the results, and show that this is not an issue in our sample.

[17] For example, the liquidation probabilities for both revisers and nonrevisers are much higher in the period starting December 2008 than in the period starting June 2009.

[18] Although these marginal effects are focused on the median rank, we confirm in the Internet Appendix that these effects are present when considering other quantiles.

Table III

Probit Regression for Revisions

This table shows the marginal effects from a probit regression on fund vintage data. The dependent variable takes the value of one if a fund revised data between vintage $v - 1$ and vintage v. The independent variables are: average returns across all dates up to $v - 1$, past 12 month average returns, average AUM, standard deviation of returns, and autocorrelation of returns, all measured as ranks relative to the other funds in the data; the number of return observations in the return history of the fund; a dummy variable that takes the value of one if the fund experienced a data revision in the prior vintage; a dummy variable that takes the value of one if the fund is located offshore; a total restrictions variable (measured as the sum of the reported lockup and redemption notice periods); a dummy variable for any information pertaining to audits in any of the databases; dummies indicating whether the fund has a high-water mark, hurdle rate provision, or a change in management company or fund manager in prior vintages. We also include database and strategy fixed effects in the regressions. d*F*/d*x* shows the change in the independent variable for a discrete change in any independent dummy variable from zero to one, and the slope at the mean for continuous independent variables. Robust standard errors control for heteroskedasticity, and cluster by vintage. *, **, and *** denote significance at the 10%, 5%, and 1% level, respectively.

	d*F*/d*x*	*Z*-Statistics
Avg. AUM (Rank) ($v - 1$)	0.033***	(7.024)
Avg. Ret (Rank) ($v - 1$)	0.007	(1.608)
Prior Year Avg. Return (Rank) ($v - 1$)	0.038***	(4.960)
Ret. Std. (Rank) ($v - 1$)	0.003	(1.035)
Return Autocorrelation (Rank) ($v - 1$)	0.014***	(4.057)
Return History Length ($v - 1$)	0.000*	(1.896)
Prior Vintage Revision Indicator	0.246***	(11.495)
Offshore	−0.006***	(−3.100)
Total Restrictions	0.002***	(4.881)
Audit Flag	0.021***	(6.749)
Management Fee	0.002**	(2.520)
Incentive Fee	−0.000**	(−2.093)
Hurdle Rate Provision ($v - 1$)	0.011***	(4.172)
Mgmt. Company or Manager Change ($v - 1$)	0.004	(0.996)
High-Water Mark ($v - 1$)	0.097***	(3.384)
Database Fixed Effects		
HFR	0.009**	(2.517)
CISDM	−0.056***	(−5.478)
BarclayHedge	0.028***	(2.740)
Strategy Fixed Effects		
Macro	0.025***	(7.739)
Relative Value	0.008	(1.601)
Directional Traders	−0.008***	(−2.794)
Funds-of-Funds	0.049***	(6.473)
Multi-Process	0.010***	(3.950)
Emerging Markets	0.004	(1.069)
Fixed Income	0.008***	(2.808)
Managed Futures	0.039***	(6.058)
Other	0.016***	(3.493)
N	328,633	
Pseudo-R^2	0.171	

performance numbers, suggesting that their revisions may simply be a function of revisions in the hedge funds that they hold. [19]

An increase in the total restrictions (lockup plus redemption notice period) on removing capital from the fund has a significantly positive effect on the propensity to report changes in histories. This may be correlated with greater asset illiquidity, as suggested by Aragon (2007), or constitute evidence that having a "longer period in which to hide" prior to withdrawals by investors shields funds from the adverse consequences of revisions.

The presence of audit information, reflected in the audit flag, has a large, significantly positive coefficient. At first glance this seems counterintuitive, as one might expect that funds not subject to audits would have more latitude to change returns. However, auditing could trigger corrections in returns; alternatively frequent changes in returns could prompt investors to press for funds to undergo audits.

This audit result is similar to the result we find for changes in management company or fund manager. We find that a fund experiencing a management change is roughly 10% more likely to revise its past returns, holding all else constant. This result is strongly statistically significant and economically important, and provides evidence in support of the "manager change" hypothesis outlined earlier, which holds that new management might be interested in a "fresh start," revamping the accounting, marking-to-market, auditing, and compliance practices of their newly acquired funds, thus triggering a set of revisions to past returns. Furthermore, this result is not just driven by a small set of funds—over the sample period, 21% (13%) of revising (nonrevising) funds experienced a change in management company or fund manager. Note that these are downward-biased estimates, as only roughly 50% of the sample funds sourced from various databases record any manager name information at all.

Finally, we find that funds with a high-water mark are 10% more likely to revise than those without a high-water mark. This is an important finding to which we return below.

B. Determinants of the Size and Direction of Revisions

We now turn to explaining the size and direction of revisions. As a first step, we take all 5,446 revising funds and construct a portfolio using their reported returns. We then report these returns using two different sets of data, namely, the very first vintage of returns for each fund, and the last vintage available for these funds, once the impact of all revisions has been incorporated. We plot the returns on this portfolio in Figure 2. While the first vintage appears in July 2007, revisions occur across the entire possible range of return history from 1994 to 2011, and hence the figure plots these two alternative reported histories.

[19] In the Internet Appendix, we present results corresponding to Table III with Funds-of-Funds removed from the sample. The results are very similar and all of the main conclusions hold.

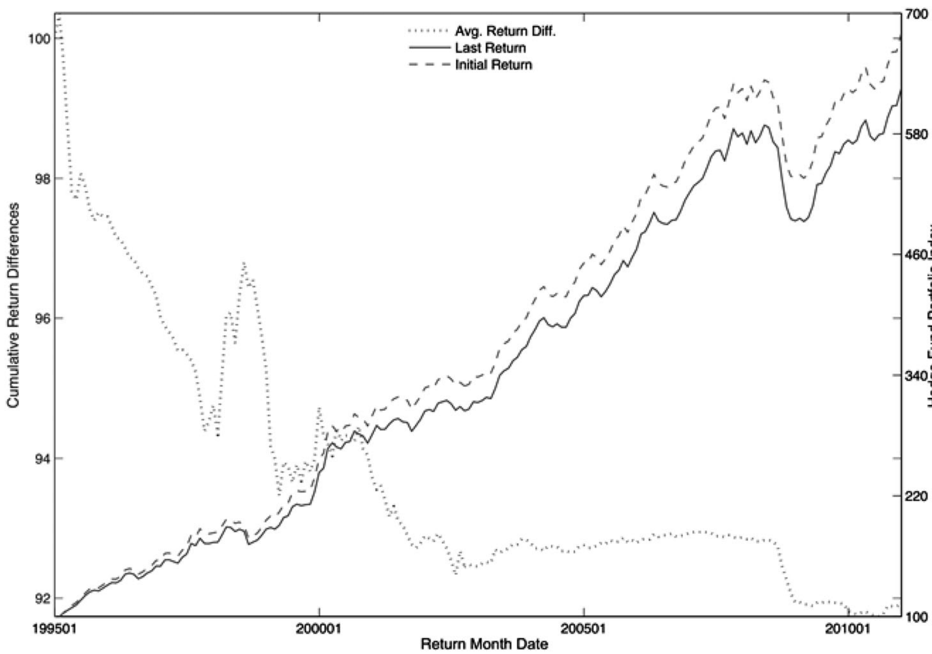

Figure 2. Cumulative differences between last and initial returns. The figure shows the cumulative average return differences between the last reported return at the most recent available vintage (denoted Last) and the first time the return is reported in a database (denoted Initial) for reviser funds. The picture plots the performance histories that would have been seen initially, versus that seen once the impact of all revisions has been taken into account. The index is set to 100 on December 31, 1994.

The figure shows clearly that the cumulative difference between final and initial returns has a significant negative trend. Thus, what a prospective investor infers about fund performance apparently depends on when he or she sees it and (especially in periods of stress, as we shall see later) final reported performance is significantly lower than initially reported performance. This suggests the danger of prospective investors being wooed into making decisions based on initially reported histories that are then subsequently revised.

While it is tempting to infer a great deal from this plot, it is consistent with multiple possibilities. The first is dishonesty, within performance initially reported to be higher than the true realization in order to increase commitments to funds, and then subsequently revised back once sufficient time has elapsed. A second possibility is that changes in management or auditors, as detected in Table III, lead to reevaluations of accounting techniques and past reported performance figures, and, in turn, to significant revisions to previously optimistic assessments. Third, fee revisions may cause a chain of net asset revaluations with consequences for older performance numbers, a possibility that we take into account below. Fourth, illiquidity could cause original estimates to be revised when valuations are finally realized. However, the revisions pertain to

Table IV

High-Water Marks and Revisions

This table examines the relationship between revisions and the presence of a high-water mark provision. Panel A conditions revising behavior on the presence of a high-water mark. For example, there are 7,977 funds with a high-water mark, and the proportion of revisers in this group is 49.35%. Panel B shows the sign and size of the average revision conditional on the presence of a high-water mark, averaged separately across positive and negative revisions, as well as across all revisions. *, **, and *** denote significance at the 10%, 5%, and 1% level, respectively, for tests of difference in means.

Panel A: Propensity to Revise Conditional on a High-Water Mark

	Fund Count	Reviser Count	% of Category
All Funds	12,128	5,446	44.90%
High-Water Mark	7,977	3,937	49.35%
No High-Water Mark	4,151	1,509	36.35%
Difference			13.00%***

Panel B: Size of Revision Conditional on a High-Water Mark

	Average Size of Revision		
	Positive Revision	Negative Revision	Net Revision
High-Water Mark	2.465	−3.483	−0.618
No High-Water Mark	4.033	−3.092	0.397
Difference			−1.015***

periods many years in the past (in some cases, up to 15 years), making it hard to explain all revisions as a consequence of later marking-to-market, and, even if the illiquidity explanation is the proximate cause, there is clearly a significant positive bias in initial estimates.

Yet another possibility is that valuation errors of both types occur, but fund managers may have greater incentives to correct them downwards rather than upwards. In particular, acknowledging overestimation of past returns may allow managers to revise historical high-water marks down, allowing the earlier collection of incentive fees, while acknowledging underestimation of past returns would require payments to investors (without even accounting for high-water marks).[20]

We explore this final explanation, which gains support from the higher propensity of funds with high-water mark provisions to revise as detected in Table III, further in Table IV, which focuses on the relationship between revisions and the existence of a high-water mark provision in the fund. The table shows that, when funds with a high-water mark revise returns, their average return revision is −62 basis points, in contrast with funds without a high-water mark, whose average return revision is +40 basis points, a difference of over 100 basis points. This important result adds more subtlety to the finding in

[20] We thank Istvan Nagy for suggesting this explanation.

Figure 2 that the average revision across all funds is negative. In particular, when we condition on the presence of a high-water mark in the fund, we find that funds with an incentive to revise returns below high-water marks revise *downwards* on average, whereas funds without high-water marks revise returns *upwards*, making past returns appear higher in subsequent revisions.

Our next step, as described in Section III, is to construct calendar-year returns for any fund-year that contains at least one revised return using both initial and final reported data, and to explain the difference between the two (i.e., final less initial), using a number of variables. Panel A of Table V, which analyzes the absolute value of these differences, shows that return revisions are on average large. Moreover, these revisions are larger in absolute value during crises, with all three of the crisis dummy variables having significantly positive coefficients. Of these, the very largest revisions pertain to the 1998–1999 crisis period, adding 1.58% to the already large baseline revision. This is followed by the 2000–2001 NASDAQ crisis period with roughly 77 basis points per annum, and the most recent crisis, with 68 basis points per annum.

Turning to the remaining fund characteristics, it appears that offshore funds have larger absolute revisions, in line with our conjecture that potentially weaker enforcement in such jurisdictions may lead to more important revisions. Perhaps surprisingly, funds with audit information appear to be associated with revisions that are larger in absolute value, suggesting that at least some revisions may be occasioned by the enhanced scrutiny of recent audits or the appointment of a new auditor. In keeping with this result, Jylhä (2011) finds that funds with prominent auditors have more misreporting discontinuities, although Liang (2003) finds no such evidence in his earlier study of the auditing of TASS returns. Finally, the table shows that smaller funds and those with high incentive fees have larger revisions, which is consistent with greater incentives for dishonesty, as well as with the possibility of larger revaluations when fee structures change.

Panel B of Table V explains return differences, rather than their absolute values, and finds that, during crisis periods, particularly the 2000–2001 and 2008–2009 periods, revisions are significantly negative, meaning that the initially reported return tends to be revised downwards in subsequent vintages of the database, as seen earlier. The table also shows that large funds with high management fees tend to make upward revisions.

We now turn to evaluating the predictive content of revisions, constructing portfolios of revisers and nonrevisers as successive vintages reveal their identities.

C. The Future Performance of Revisers and Nonrevisers

Figure 3 plots the cumulative performance of the reviser and nonreviser portfolios constructed as described in Section III C. Panel A shows that the returns of the revisers are appreciably lower than those of nonrevisers. This difference is economically substantial with a cumulative difference of 12.4%

Table V

Explaining Revision Return Differences

This table conditions the return differences occasioned by revisions on various fund characteristics and period fixed effects. The dependent variable is the average difference, for all years in which a fund experienced return revisions, between the final set of annual returns provided by a fund and the first set of annual returns provided by the same fund for the same year. For example, if fund X initially reported a 6% average annual return for year t, and at the final vintage this average stood at 4%, then the return difference variable would be −2%. We only include periods in which the fund had at least six months of return observations to reduce the noise in the dependent variable. Panel A takes the absolute value of all such differences as the dependent variable, and Panel B conditions the signed revisions on the independent variables. Period dummies include crisis dummies for the 1998 to 1999 period, the 2000 to 2001 period, and the 2008 to 2009 period. The remaining regressors have been described earlier in these tables, with three new additions, namely, the rank of prior flows and returns experienced by the fund relative to all other funds in the same year, as well as the management fee, and the incentive fee of the fund. t-statistics, shown in parentheses, are robust to heteroskedasticity and clustered at the fund level. *, **, and *** denote significance at the 10%, 5%, and 1% levels respectively.

Panel A: Absolute Value of Differences				
	Coeff	t-Stat	Coeff	t-Stat
Constant	1.170	(20.277)***	1.252	(5.345)***
Crisis dummy1: 1998–99	1.580	(2.891)***	1.577	(2.919)***
Crisis dummy2: 2000–01	0.770	(2.435)**	0.744	(2.368)**
Crisis dummy3: 2008–09	0.677	(8.330)***	0.669	(8.174)***
Offshore			0.300	(2.695)***
Total Restrictions			−0.022	(−1.251)
High-Water Mark or Hurdle			−0.206	(−1.609)
Audit			0.356	(2.431)**
Management Fee			0.028	(0.284)
Incentive Fee			0.022	(2.795)***
Asset $t-1$ rank			−1.122	(−5.462)***
Return prior year $t-1$ rank			−0.295	(−1.859)*
Flow prior year $t-1$ rank			0.062	(0.462)
N	7,628		7,628	
Adjusted R^2	0.012		0.026	
Panel B: Return Differences				
Constant	−0.007	(−0.129)	−0.149	(−0.725)
Crisis dummy1: 1998–99	−0.139	(−0.216)	−0.164	(−0.253)
Crisis dummy2: 2000–01	−0.809	(−2.403)**	−0.819	(−2.445)**
Crisis dummy3: 2008–09	−0.375	(−4.412)***	−0.370	(−4.348)***
Offshore			−0.133	(−1.503)
Total Restrictions			0.009	(0.527)
High-Water Mark or Hurdle			−0.129	(−1.114)
Audit			−0.038	(−0.294)
Management Fee			0.155	(1.873)*
Incentive Fee			−0.001	(−0.108)
Asset $t-1$ rank			0.256	(1.742)*
Return prior year $t-1$ rank			0.117	(0.719)
Flow prior year $t-1$ rank			−0.176	(−1.140)
N	7,628		7,628	
Adjusted R^2	0.003		0.004	

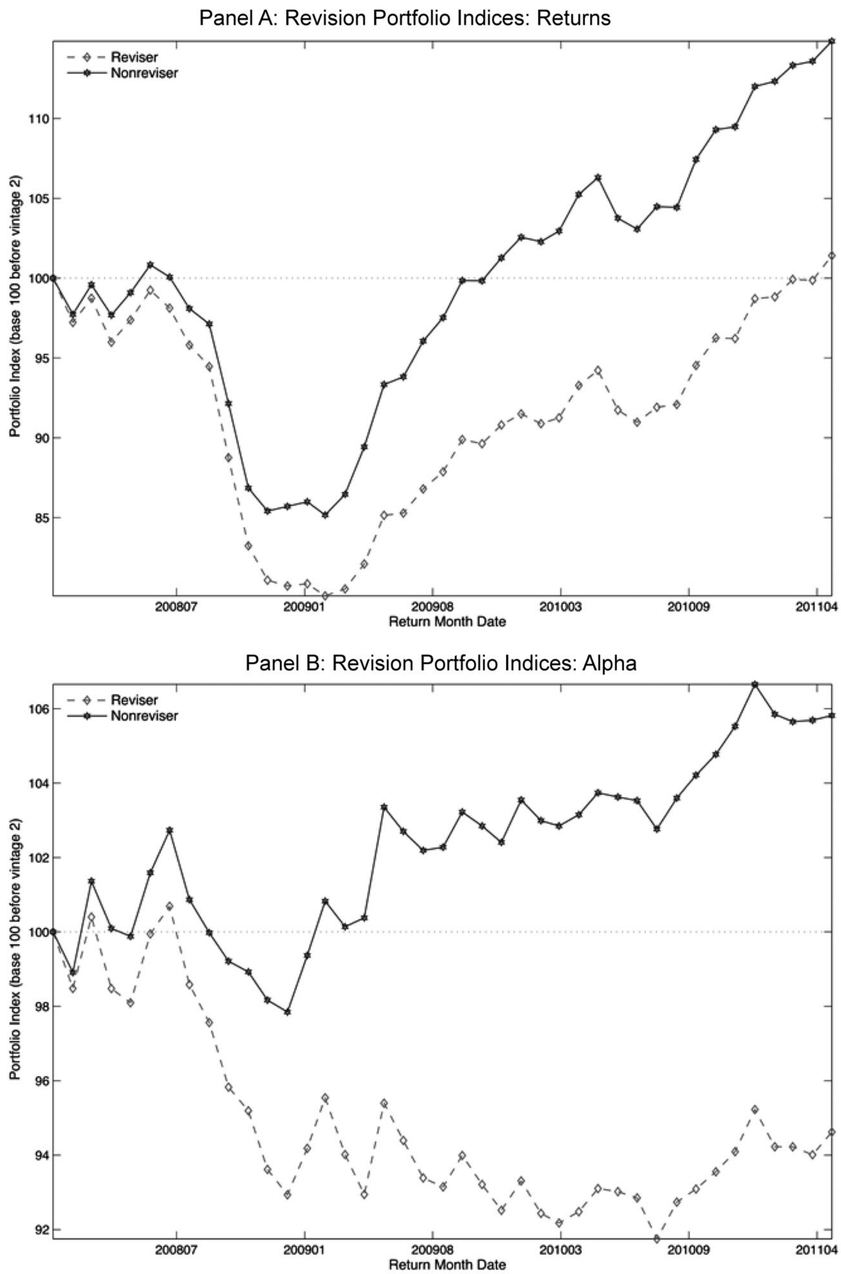

Figure 3. Portfolio performance—revisers and nonrevisers. This figure shows the cumulative performance of reviser and nonreviser portfolios. The nonreviser portfolio is comprised of funds that never revise between vintages plus the early records of funds before they become revisers. For example, if a fund first revises at vintage v, it will be included in the nonreviser portfolio prior to that vintage; once it joins the reviser portfolio it is removed from the nonreviser portfolio. The index is set to 100 at December 31, 2007, just before the second vintage starts. Equal-weighted returns are employed in Panel A, and Panel B plots cumulative alpha + epsilon using the Fung–Hsieh seven-factor model.

emerging after just over three years.[21] This substantial return difference between the two portfolios may at first glance suggest that our classification of funds into revisers and nonrevisers has substantial predictive content. However, to better understand these differences, and to ensure that they are not simply driven by differences in the risk loadings or characteristics of funds, we need to risk-adjust (and potentially characteristic-adjust) these returns.[22]

Table VI presents results from a variety of models for risk-adjusting the return difference between the reviser and nonreviser portfolios, and shows that the findings above are robust to this choice. This table reports only the alpha from these regressions; the full set of results, including the coefficients on the various factors, are reported in the Internet Appendix. The alpha of the nonreviser–reviser difference from the Fung–Hsieh seven-factor model is 0.28% per month, or 3.3% per annum net of all fees and costs. We plot cumulative alpha (i.e., $\alpha + \varepsilon_t$ for each time-series portfolio regression) estimated using the Fung–Hsieh seven-factor model in Panel B of Figure 3, and find that it resembles the plot of raw returns: the nonrevisers consistently outperform the revisers. We also consider risk adjustment using the Fama–French three-factor model, as well as augmented variants that include momentum and liquidity factors, and find that the future poor performance of the "reviser" portfolio is not explained by these alternative models. Panels C through E consider various robustness checks, which are discussed in a separate section below.[23]

Having established that the reviser/nonreviser return differential is not explained by differences in exposure to risk factors, we next consider several possible drivers of this result. One inference is to consider revisions as a sign of dishonesty or poor operational controls within the fund. For example, when management changes occur in the fund (an important determinant of revising behavior), this could result in changes to operational controls going forward. While these may be put in place to generate better future performance, the very fact that changes may have been required highlights potentially important structural deficiencies in the fund's previous accounting practices that need to be remedied, and hence the presence of operational risk that may manifest itself in low future returns.

[21] Note that, even in the early periods of the out-of-sample period, we still have a substantial number of firms in the "reviser" portfolio, growing from 274 revising firms detected in the first month.

[22] In the Internet Appendix, we also plot cumulative flows for both reviser and nonreviser portfolios, using data from the final vintage. The reviser portfolio experiences significant outflows beginning in August and September 2008, during the Lehman collapse. The impact of large outflows and subsequent fire sales of fund assets might be one potential reason for the poor performance of the reviser portfolio (see Coval and Stafford (2007) and Jotikasthira, Lundblad, and Ramadorai (2012) for evidence of the importance of this mechanism). The flows may also simply be responding to poor performance, á la DeLong et al. (1990).

[23] The Internet Appendix presents results that correspond to Table VI but with Funds-of-Funds excluded. The risk-adjusted excess performance is smaller for non-Funds-of-Funds, around 0.24% per month compared with 0.28% per month; however, the difference in performance is still strongly statistically significant across all risk adjustment models.

Table VI

Do Revisions Predict Future Returns?

This table presents the estimated alpha from regressions of the difference in returns between the nonreviser and reviser portfolios over the 40 months from January 2008 to May 2011, on several different sets of factors, and conducts several robustness checks of the results. Panel A employs the Fung–Hsieh eight-factor model, and subsets of it. Panel B employs the Fama–French three-factor model, adds a momentum factor, and finally adds the Pástor–Stambaugh liquidity factor. Panel C shows the impact of using different size thresholds for flagging a revision as important, with the first column (1 bp) of Panel C reproducing the result from Panel A. Panel D shows the impact of using different "recency" thresholds for revisions, giving a "free pass" to revisions of recent returns. The second column (more than three months) of Panel D reproduces the result from Panel A. Panel E shows the significance of the differences in returns between the nonreviser and reviser portfolios using the portfolio's median return. Newey–West standard errors (with three lags) are employed. Regression alphas are shown, with t-statistics in parentheses beneath them. (Full estimation results are presented in the Internet Appendix.) *, **, and *** denote significance at the 10%, 5%, and 1% level, respectively.

Panel A: Return Differences (Fung–Hsieh Model)

	Constant	Market	FH 4	FH 7	FH 8
Alpha	0.309***	0.309***	0.277***	0.278***	0.279***
	(3.805)	(5.133)	(3.526)	(3.053)	(3.077)

Panel B: Return Differences (Fama–French Three Factors + Momentum + Pástor–Stambaugh Liquidity Model)

	FF3	FF3 + Mom	FF3 + Mom + Liquidity
Alpha	0.302***	0.276***	0.287***
	(3.777)	(4.596)	(4.973)

Panel C: Size of Revision (Fung–Hsieh Seven-Factor Model)

	Minimum Size of Revisions			
	1 bp	10 bp	50 bp	100 bp
Alpha	0.278***	0.292***	0.262***	0.250***
	(3.053)	(3.362)	(3.247)	(2.638)

Panel D: Recency of Revision (Fung–Hsieh Seven-Factor Model)

	Minimum Recency of Revisions			
	1 or More Months	More than 3 Months	More than 6 Months	More than 2 Months
Alpha	0.222***	0.278***	0.302***	0.255***
	(2.591)	(3.053)	(3.193)	(2.672)

Panel E: Regressions on Median Return Differences (Fung–Hsieh Seven-Factor Model)

	Constant	Market	FH 4	FH 7	FH 8
Alpha	0.207**	0.213***	0.196***	0.200***	0.203***
	(2.382)	(3.790)	(3.318)	(3.218)	(3.273)

If either dishonesty or poor operational controls were the driver of revisions, we might also expect to see differences in the tail risk of revisers relative to nonrevisers—the dramatic outflows from the reviser portfolio suggest that these differences may be stark. To test this conjecture, we employ a historical simulation method, where we estimate the bottom decile of performance from all returns seen from the beginning of the reviser portfolio up until each date, moving over time (this is done at the individual fund level within each of the portfolios). We also average the returns falling below these empirically computed decile thresholds to arrive at an expected shortfall measure.

Figure 4 plots these measures for the cross-section of underlying funds of the respective portfolios. We caution that this analysis is conducted on a relatively small sample of data, implying that our estimates of tail quantities are somewhat imprecise and hence these plots should be taken as suggestive. Nevertheless, the figure shows that the empirical bottom decile and the expected shortfall of the reviser portfolio are virtually always below the nonreviser portfolio over the entire period for both portfolio and cross-sectional measures. There is a dramatic divergence during the 2008–2009 crisis, with the empirical percentile and the expected shortfall collapsing in October and November 2008. While the tail risk of the revisers at the fund level recovers and appears quite similar to that of the nonrevisers in the more recent periods, this could be attributed to the weakest funds having been eliminated from the portfolio during the crisis period. Overall, this analysis suggests that investors are at greater downside risk when investing in funds that revise their returns. When we check the results using lower percentile thresholds, the conclusions are similar.

The recovery of the tail risk in the reviser portfolio toward the end of the sample period suggests that these funds might hold more illiquid assets in their portfolios, which simultaneously drives revisions, sharp declines in asset value, and subsequent recoveries. In this sense, we might simply be picking up differences in asset holdings. In the next section we explore this and other potential determinants of our findings.

Finally, we attempt to link our reviser flag with a more objective measure of future performance than self-reported returns. Table VII looks at liquidation probabilities of reviser and nonreviser funds through the probability that a fund will cease reporting to a hedge fund database. It should be noted that funds may cease reporting to a database for reasons other than fund liquidation, and so this analysis comes with a caveat. It nonetheless provides another piece of evidence about the future performance of reviser funds relative to nonreviser funds. Given the turbulent period that our sample covers, we compute these probabilities starting from six different dates (June 2008–December 2010), and for five horizons, ranging from 6 to 30 months. For example, in the period up to December 2008, 7,533 funds (2,140 revisers and 5,393 nonrevisers) report returns. Twelve months later, 26.5% of these funds had ceased reporting, with the revisers having a 32.1% liquidation rate, compared with 24.3% for nonreviser funds.

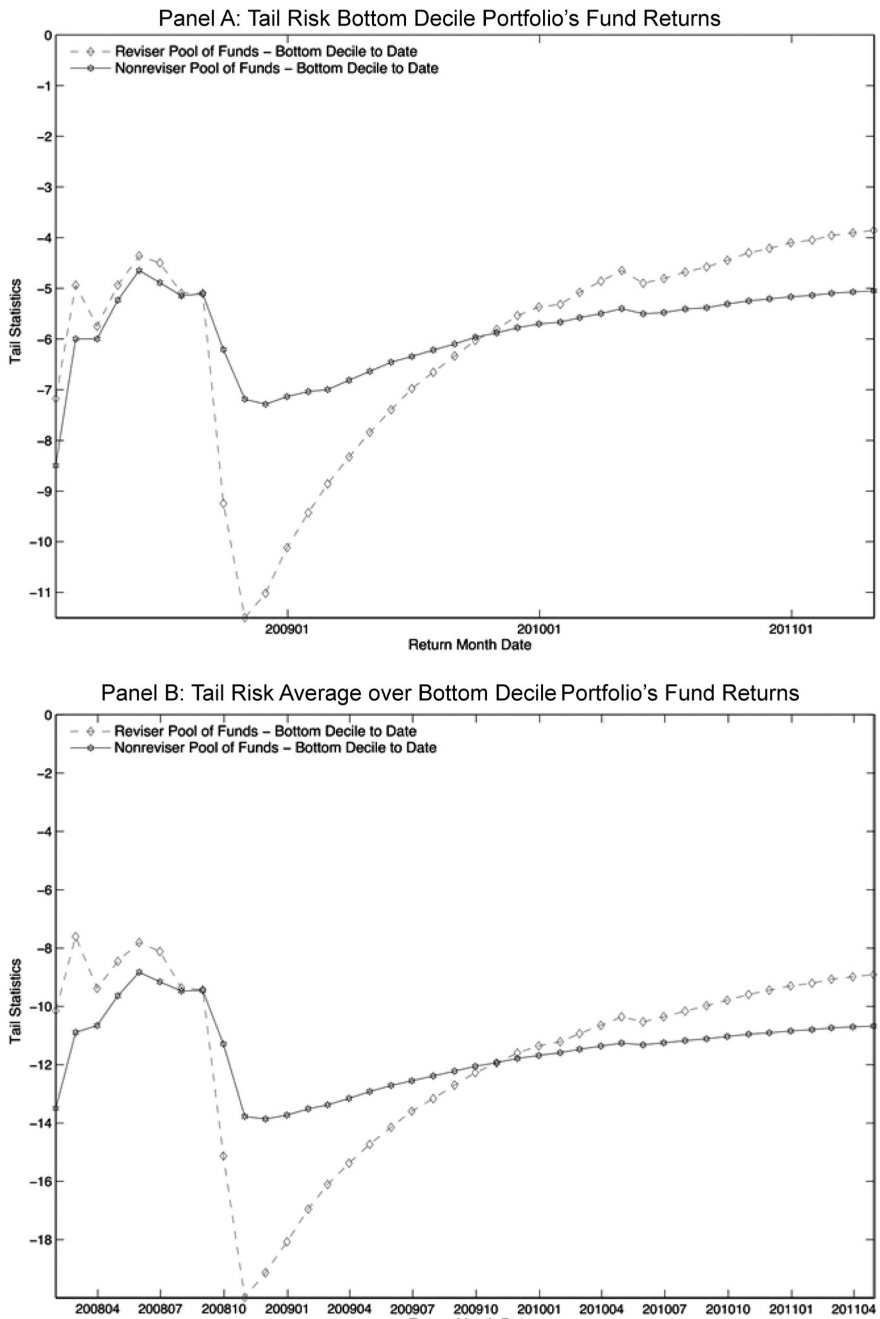

Figure 4. Tail risk percentiles for reviser and nonreviser portfolios. The figure shows the bottom decile tail statistics for the reviser portfolio and nonreviser portfolio. Panel A shows the empirical bottom decile for the portfolio fund returns using historical simulation. Panel B shows the average return of the portfolio fund returns in this bottom decile as a measure of expected shortfall.

Table VII

Liquidation Probabilities for Revisers and Nonrevisers

This table shows the liquidation probabilities of the reviser and nonreviser funds. Funds reporting returns are classified from the beginning of our vintage sample up to a given point in time (reported in the row headers) as revisers or nonrevisers, and this cohort is tracked over future six-month horizons until they stop reporting returns. Liquidation probabilities are calculated relative to the initial number of funds in the reviser cohort, reported in the column labeled "Fund Count." Liquidation probabilities are a weighted average across classification periods, based on fund count, as is the difference between the reviser and nonreviser average liquidation rates. The row labeled "Average All Funds" shows the average liquidation rate of the universe of funds. t-statistics of the difference in means are shown in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% level, respectively.

		Liquidation Probabilities: Months Ahead				
Classification Period	Fund Count	6	12	18	24	30
Revisers						
Up to Jun 2008	298	0.185	0.336	0.419	0.534	0.614
Up to Dec 2008	2,140	0.234	0.321	0.401	0.471	
Up to Jun 2009	2,251	0.115	0.219	0.314		
Up to Dec 2009	2,411	0.116	0.229			
Up to Jun 2010	2,445	0.133				
Up to Dec 2010	2,256					
Average		0.149	0.258	0.360	0.479	0.614
Nonrevisers						
Up to Jun 2008	8,577	0.138	0.308	0.374	0.428	0.516
Up to Dec 2008	5,393	0.176	0.243	0.301	0.419	
Up to Jun 2009	4,189	0.069	0.130	0.277		
Up to Dec 2009	3,773	0.054	0.213			
Up to Jun 2010	3,080	0.156				
Up to Dec 2010	2,306					
Average		0.124	0.242	0.330	0.425	0.516
Difference Revisers and Nonrevisers	0.025	0.016	0.030	0.054	0.098	
(t-statistic)		(4.236) ***	(2.304) **	(3.969) ***	(6.988) ***	(13.233) ***
Average All Funds		0.131	0.246	0.336	0.433	0.519
Difference as % All Funds Average	18.7%	6.7%	9.1%	12.5%	18.9%	

Averaging across the start date for the analysis, we find that the liquidation probabilities for revisers range from 15.7% to 61.4%, while the corresponding figures for nonrevisers are 11.9% and 51.6%. The difference between these probabilities ranges from 1.6% to as high as 9.8%, and is strongly statistically significant for all five horizons. As a proportion of the average liquidation probability at a given horizon, increases of this size represent an increase of up to 20% in the liquidation probability for reviser funds relative to nonreviser funds.

Detecting that a fund has revised one of its past returns helps us predict that it will significantly underperform funds that have never revised their returns, and significantly increases the probability that the fund will cease reporting to a database, potentially due to liquidation. The usefulness of the revision indicator in the future is, of course, susceptible to changes in investor and manager behavior: as investors become aware of the information content of this indicator, the incentives to revise past returns may change in turn.

V. Robustness Checks

In this section, we present results of a battery of robustness checks of our main empirical findings. The Internet Appendix presents additional robustness checks and analyses.

A. Varying the Minimum Size of the Revision

The first parameter that we vary is the minimum size of a change necessary for it to be labeled a revision. This test is one way to control for the possibility that our results are driven by the initial marking-to-market of illiquid assets. It also allows us to see if we can obtain stronger predictability signals by conditioning on larger revisions. Our main analysis uses a one basis point threshold for identifying revisions. Here we increase this threshold to 10, 50, and 100 basis points, in each case only classifying as revisions changes in returns across successive vintages that are greater than the threshold.

Panel C of Table VI reveals that the return differences reported in Panel A of the table persist, with the estimated monthly alphas across these thresholds ranging from 0.25% to 0.29%. Indeed, our results appear slightly stronger when we only consider funds with larger revisions in our set of revisers.

B. Varying the Minimum Age of the Revision

Our next robustness check is to give a "free pass" to revisions that occur close to the vintage date. As explained earlier, the recency, k, of a revision is the number of months between the return date and the date of the vintage in which the revision was observed. The parameter k is useful for evaluating various hypotheses. By setting k to be large, we can evaluate only those funds that revise "ancient history." Moreover, using a large k eliminates the incorporation of funds into the reviser portfolio that revised returns relatively quickly. In other words, we can give a free pass to small k revisers to allow for the possibility that funds may employ estimated returns for recent time periods, which could be revised due to accounting procedures or because of the revaluation of illiquid securities in light of more accurate information. The larger we set k, the less likely that we are picking up such revaluation revisions. In this robustness check, we consider both $k \geq 1$ to include all revisions, as well as $k > 3$ (our baseline in the paper), $k > 6$, and $k > 12$ to identify revisions older than one quarter, six months, and one year, respectively.

Panel D of Table VI shows that our results become slightly stronger as k increases, peaking at $k > 6$, and decreasing slightly for $k > 12$, but still higher than unrestricted $k \geq 1$. This suggests that revisions of very recent returns are generally more innocuous (in the sense that they do not help predict poor future returns) than revisions of older returns. It is worth noting here that we take additional care in two cases. First, for each k, we ensure that funds revising returns more recent than the threshold k are *not* included in the *nonreviser* portfolio, that is, they do not factor into any of our calculations, to ensure that we compare "true" nonrevisers with high-k revisers. Second, in any given vintage, we do not include funds in both reviser and nonreviser portfolios if they simultaneously conduct low- *and* high-k revisions.[24] This allows for the possibility of a benign AUM or valuation error found months ago that could, in some cases, cause a cascade of revisions. For example, an incorrectly processed corporate event in one of the equity holdings of the hedge fund could trigger off such a case. Despite these exclusions, high-k revisions are associated with significant return differentials between revisers and nonrevisers.

C. Controlling for Extreme Returns

One may wonder whether the poor future performance of hedge funds that have revised their returns is driven by a few extreme returns. To address this question, we consider the reviser/nonreviser performance differential using the *median* return for each of these groups rather than the mean return. Of course, the median return cannot be interpreted as the return on a portfolio of hedge funds (while the mean can), but it does shed light on the sensitivity of our results to rare large returns.

Panel E of Table VI shows that the risk-adjusted median return is slightly smaller than the risk-adjusted mean return (around 0.20% per month compared with 0.28%), but is strongly significant across all risk adjustment models. Thus, the negative future performance of revising funds is *not* attributable to the extreme poor performance of a few revising funds or, conversely, to the extreme high performance of a few nonrevising funds.

D. Two-Way Sorts on Fund Characteristics

In our earlier probit analysis, we find that reviser and nonreviser funds have different characteristics.[25] While the factor loadings of the return difference between these groups should capture such differences, we perform an additional test to check that our results are not driven by such characteristic-based differences. To do so, we double-sort by these characteristics and the reviser/nonreviser classification. We consider five such fund characteristics, three of which have been identified in the literature as relevant for expected returns, namely, the first autocorrelation of fund returns, which is a measure of the smoothness of the fund's returns á la Getmansky, Lo, and Makarov (GLM; 2004), the total lockup period imposed by the fund (see Aragon (2007)), and the

[24] Of course, if they only conducted a high-k revision in a subsequent vintage, they would then be included in the reviser portfolio.

[25] The Internet Appendix presents a formal comparison of key characteristics of reviser and nonreviser funds.

Table VIII

Robustness Checks: Liquidity and Fund Characteristics

This table conditions the results in Table VI on the cross-section of various fund characteristics. Sorting funds on specific characteristics, we split both revisers and nonrevisers into groups that are above (High) and below (Low) the cross-sectional median of all funds reporting in each period. These characteristics are: Rho1 (first return autocorrelation), the lockup period as at the last available vintage, fund size (AUM at the end of the prior period), Return Std (return standard deviation), and history length (the number of return observations in the return history of the fund). Returns are equally weighted within portfolios. Newey–West heteroskedasticity and autocorrelation-robust standard errors (with three lags) are employed. Regression betas are shown with t-statistics in parentheses beneath coefficients. *, **, and *** denote significance at the 10%, 5%, and 1% level, respectively.

	Alpha (Fung–Hsieh Seven-Factor Model)	
Characteristic	High	Low
Rho1	0.322***	0.107
	(−3.467)	(−1.275)
Lockup	0.367***	0.168*
	(−4.730)	(−1.718)
Fund Size	0.142**	0.522***
	(−2.166)	(−3.150)
Return Std	0.309***	0.286***
	(−2.633)	(−3.318)
History Length	0.120**	0.509***
	(−2.200)	(−3.474)

size of the fund, to control for the impact of capacity constraints (see Fung et al. (2008)). We also sort by the fund's total return volatility and history length (a measure of fund age).

Given the nature of the fund characteristics that we employ for these double-sorts, this analysis also allows us to investigate whether fund asset illiquidity (correlated with both the GLM measure and lockup periods, according to extant literature) helps explain the reviser–nonreviser difference. Specifically, if this were the case, we would expect to see no differences between revisers and nonrevisers within each portfolio of funds (independently) double-sorted by illiquidity proxies (autocorrelation, lockup, fund size), but pronounced differences across these illiquidity-sorted groups. If, however, we continue to see variation in reviser and nonreviser portfolio returns within these groups, this would suggest that the revisions provide orthogonal information to underlying asset illiquidity.[26]

The alphas of the return differences between reviser and nonreviser funds of these double-sorted portfolios are reported in Table VIII, and are all statistically significant with a single exception. We find that reviser–nonreviser differences

[26] Of course, if these proxies for illiquidity are not as good a measure of underlying asset illiquidity as our revisions measure, it is possible that the explanation might still apply. In that case, the interpretation is that we have found a better measure of asset illiquidity, although the other robustness checks (especially varying k) do not support this explanation.

are particularly stark among funds that have high return autocorrelation, but alphas for less-smooth (low Rho1) revisers are lower, but not significantly lower, than those for less-smooth nonrevisers. Getmansky, Lo, and Makarov (2004), for example, highlight that their measure of return smoothness could be on account of either true asset illiquidity or deliberate return-smoothing among funds—so our result that smooth return revisers have worse performance than smooth return nonrevisers may allow investors to discriminate between these two possibilities for observed return smoothness. We also find that small funds and young funds show stark differences between reviser and nonreviser portfolio returns. This suggests that, when revising behavior is detected in funds with relatively higher incentives to establish their reputations, it might well be construed as a particularly negative signal about their future return prospects.

In addition to the robustness checks described above, we conduct a series of other robustness checks that are described and presented in the Internet Appendix.

E. Comparison with Other Flags for "Problem" Funds

Bollen and Pool (2012) propose a variety of flags for potential fraudulent activity that are based on the statistical properties of reported returns and link these flags to an indicator for whether the fund has been charged with legal or regulatory violations. Accordingly, we next cross-tabulate our categorization of funds into revisers and nonrevisers with these statistical flags. To do so, we estimate the Bollen and Pool flags for each fund individually, and then aggregate funds into reviser and nonreviser groups and report the proportion of funds in each group that triggers each of these flags.

To implement these tests, we use the returns from the final vintage, and impose a minimum return history of 24 months. This reduces our sample of funds from 12,128 to 10,584. The tests we use are "Perc. Negative," "Count Zeros," "String," "Num. Pairs," "Perc. Repeats," "Uniform," "Benford," "Kink," "Index Rsq," "AR(1)," and "CAR(1)," as in Bollen and Pool (2012). The header to Table IX describes the construction of each of these flags. The table shows that five of these flags are triggered for a higher percentage of reviser funds than nonreviser funds, namely, the Perc. Negative, AR(1), Perc. Repeats, Count Zeros, and Kink flags. The remaining six flags, bar one, are not significantly different across revisers and nonrevisers (and three of them are triggered at a higher rate for nonrevisers than for revisers).[27]

To summarize, none of the Bollen–Pool (2012) flags contain exactly the same information as our categorization of funds into revisers and nonrevisers. The

[27] We find that nonrevisers trigger the Index Rsq flag more frequently than revisers. This flag associates strategy distinctiveness with potentially fraudulent activity (e.g., Madoff, whose claimed strategy did not match the returns generated). Our test results suggest that nonrevisers' returns are more idiosyncratic. Sun, Wang, and Zheng (2012) find that skilled managers are more likely to have distinct returns, so this could represent a quality effect in the nonrevisers group.

Table IX

Fraud Flag Frequencies for Revisers and Nonrevisers

This table shows the proportions of reviser and nonreviser hedge funds that trigger the Bollen and Pool (2012) "performance flags." Returns are taken over the full history using the last available vintage. We require funds to have at least 24 months of returns. A 10% significance level was used for the Bollen and Pool (2012) tests. These tests are: Perc. Negative, triggered by low percentage of returns that are negative; AR(1), first-order return autocorrelation; Perc. Repeats, the proportion of returns that are repeated; Kink, a test of discontinuity at zero in distribution of fund returns; Count Zeros, the count of exactly zero returns; String, the count of the longest sequence of repeated data; Num. Pairs, the number of repeated blocks of length two, without counting overlaps; CAR(1), conditional serial correlation to check smoothing of losses, using the Fung–Hsieh seven-factor model for the unobserved return; Uniform, establishing whether the second digit of the value is uniformly distributed; Benford, establishing whether the second digit of the value follows Benford's Law for a second digit; and Index Rsq, the difference in relationship between the fund and its strategy peers. Critical values are obtained using a bootstrap procedure. The second-last column reports the difference in the proportion of funds that trigger a performance flag, and the final column reports p-values on these differences. *, **, and *** denote significance of the difference in proportions at the 10%, 5%, and 1% levels, respectively. The flags are sorted by the difference in proportions between reviser and nonreviser funds.

Flag	Reviser Funds ($N = 5{,}055$)	Nonreviser Funds ($N = 5{,}529$)	Difference	p-Value
Perc. Negative	0.359	0.251	0.108***	0.000
AR(1)	0.524	0.420	0.105***	0.000
Count Zeros	0.180	0.151	0.029***	0.000
Perc. Repeats	0.203	0.174	0.029***	0.000
Kink	0.211	0.185	0.026***	0.001
Num. Pairs	0.035	0.033	0.002	0.623
String	0.088	0.087	0.002	0.744
CAR(1)	0.127	0.126	0.001	0.864
Uniform	0.129	0.133	−0.004	0.547
Benford	0.106	0.114	−0.007	0.240
Index Rsq	0.156	0.206	−0.050***	0.000

"confusion matrix"[28] implied by the proportions in Table IX yields accuracy measures (which corresponds to a correlation measure) of between 0.42 and 0.54. Thus, funds identified as problem funds using the methods of Bollen and Pool (2012) have about 50% overlap with funds that we identify as revisers. Overall, our "reviser" category does correlate with some previously proposed flags, but it contains substantial unique information.

VI. Conclusions

This paper examines the reliability of voluntary disclosures of performance information by hedge funds. To do so, we track revisions to historical

[28] This matrix is used to compare two discrete classifications of a variable, in this case whether a fund is a "problem" fund or not. The "accuracy" measure is simply the sum of the proportions where the two classifications agree, and can be interpreted as a correlation measure for these classifications.

performance records by hedge funds in several publicly available hedge fund databases. We find evidence that, in successive vintages of these databases, older performance records (pertaining to periods as far back as 15 years) of hedge funds are routinely revised. These revisions are widespread, with nearly 50% of the 12,128 hedge funds in our sample (managing around 45% of average total assets) having revised their historical returns at least once. These revisions are not merely random reporting errors: they can be predicted in part using information on the characteristics and past performance of hedge funds, with larger, more volatile, and less liquid funds more likely to revise their returns. Initially reported performance track records present a far rosier picture of historical performance than track records that include all changes made in subsequent data vintages, especially for funds that have high-water mark provisions. Perhaps most interestingly, detecting that a fund has revised one of its past returns helps predict that it will subsequently underperform funds that have never revised their returns, and increases the probability that the fund will cease reporting to a database, potentially due to liquidation.

Recent policy debates on the pros and cons of imposing stricter reporting requirements on hedge funds have raised various arguments. The benefits of disclosure include market regulators having a better view of the systemic risks in financial markets, and investors and regulators being better able to determine the true risk-adjusted performance of the fund. The costs include the administrative burden of preparing such reports, and the risk of leakage of valuable proprietary information in the form of trading strategies and portfolio holdings. Our analysis suggests that mandatory, audited disclosures by hedge funds, such as those recently proposed by the SEC, would be beneficial to regulators. We believe that it would also be worth considering how these reporting guidelines, which currently only apply to funds' disclosures to regulators, could also apply to disclosures to prospective and current investors so as to help them make more informed investment decisions.

Initial submission: March 6, 2012; Final version received: October 9, 2014
Acting Editor: Williom Goetzmann

REFERENCES

Ackermann, Carl, Richard McEnally, and David Ravenscraft, 1999, The performance of hedge funds: Risk, return, and incentives, *Journal of Finance* 54, 833–874.

Agarwal, Vikas, Naveen D. Daniel, and Narayan Y. Naik, 2011, Do hedge funds manage their reported returns?, *Review of Financial Studies* 24, 3281–3320.

Agarwal, Vikas, Wei Jiang, Yuehua Tang, and Baozhong Yang, 2013, Uncovering hedge fund skill from the portfolio holdings they hide, *Journal of Finance* 68, 739–783.

Ang, Andrew, Sergiy Gorovyy, and Greg van Inwegen, 2011, Hedge fund leverage, *Journal of Financial Economics* 102, 102–126.

Aragon, George O., 2007, Share restrictions and asset pricing: Evidence from the hedge fund industry, *Journal of Financial Economics* 83, 33–58.

Aragon, George O., and Vikram Nanda, 2011, Strategic delays and clustering in hedge fund reported returns, Working paper, Arizona State University.

Asness Cliff, S., Robert Krail, and John M. Liew, 2001, Do hedge funds hedge? *Journal of Portfolio Management* 28, 6–19.

Ben-David, Itzhak, Francesco A. Franzoni, Augustin Landier, and Rabih Moussawi, 2013, Do hedge funds manipulate stock prices?, *Journal of Finance* 68, 2383–2434.

Berk, Jonathan B., and Richard C. Green, 2004, Mutual fund flows and performance in rational markets, *Journal of Political Economy* 112, 1269–1295.

Bollen, Nicolas P. B., and Veronika K. Pool, 2008, Conditional return smoothing in the hedge fund industry, *Journal of Financial and Quantitative Analysis* 43, 267–298.

Bollen, Nicolas P. B., and Veronika K. Pool, 2009, Do hedge fund managers misreport returns? Evidence from the pooled distribution, *Journal of Finance* 64, 2257–2288.

Bollen, Nicolas P. B., and Veronika K. Pool, 2012, Suspicious patterns in hedge fund returns and the risk of fraud, *Review of Financial Studies* 25, 2673–2702.

Brown, Stephen, William Goetzmann, Bing Liang, and Christopher Schwarz, 2008, Mandatory disclosure and operational risk: Evidence from hedge fund registration, *Journal of Finance* 63, 2785–2815.

Brown, Stephen, William Goetzmann, Bing Liang, and Christopher Schwarz, 2009, Estimating operational risk for hedge funds: The ω-score, *Financial Analysts Journal* 65, 43–53.

Brown, Stephen, William Goetzmann, Bing Liang, and Christopher Schwarz, 2012, Trust and delegation, *Journal of Financial Economics* 103, 221–234.

Carhart, Mark M., Jennifer N. Carpenter, Anthony W. Lynch, and David K. Musto, 2002, Mutual fund survivorship, *Review of Financial Studies* 15, 1439–1463.

Cassar, Gavin, and Joseph Gerakos, 2011, Hedge funds: Pricing controls and the smoothing of self-reported returns, *Review of Financial Studies* 24, 1698–1734.

Cici, Gjergji, Alexander Kempf, and Alexander Puetz, 2011, The valuation of hedge funds' equity positions, Working paper, College of William and Mary.

Coval, Joshua, and Erik Stafford, 2007, Asset fire sales (and purchases) in equity markets, *Journal of Financial Economics* 86, 479–512.

Croushore, Dean, 2011, Frontiers of real-time data analysis, *Journal of Economic Literature* 49, 72–100.

Daníelsson, Jón, Paul Embrechts, Charles Goodhart, Con Keating, Felix Muennich, Olivier Renault, and Hyun Song Shin, 2001, An academic response to Basel II, London School of Economics, Financial Markets Group Special Paper No. 130.

De Long, J. Bradford, Andrei Shleifer, Lawrence H. Summers, and Robert J. Waldmann, 1990, Positive feedback investment strategies and destabilizing rational speculation, *Journal of Finance* 45, 379–395.

Dechow, Patricia, Weili Geb, and Catherine Schrand, 2010, Understanding earnings quality: A review of the proxies, their determinants and their consequences, *Journal of Accounting and Economics* 50, 344–401.

Foster, Dean P., and H. Peyton Young, 2010, Gaming performance fees by portfolio managers, *Quarterly Journal of Economics* 125, 1435–1458.

Fung, William, and David A. Hsieh, 2000, Performance characteristics of hedge funds and commodity funds: Natural vs. spurious biases, *Journal of Financial and Quantitative Analysis* 35, 291–307.

Fung, William, and David A. Hsieh, 2001, The risk in hedge fund strategies: Theory and evidence from trend followers, *Review of Financial Studies* 14, 313–341.

Fung, William, and David A. Hsieh, 2009, Measurement biases in hedge fund performance data: An update, *Financial Analysts Journal* 65, 36–38.

Fung, William, David A. Hsieh, Narayan Y. Naik, and Tarun Ramadorai, 2008, Hedge funds: Performance, risk and capital formation, *Journal of Finance* 63, 1777–1803.

Getmansky, Mila, Andrew W. Lo, and Igor Makarov, 2004, An econometric model of serial correlation and illiquidity in hedge fund returns, *Journal of Financial Economics* 74, 529–609.

Jin, Ginger Zhe, and Phillip Leslie, 2003, The effect of information on product quality: Evidence from restaurant hygiene grade cards, *Quarterly Journal of Economics* 118, 409–451.

Jin, Ginger Zhe, and Phillip Leslie, 2009, Reputational incentives for restaurant hygiene, *American Economic Journal: Microeconomics* 1, 236–267.

Jorion, Philippe, and Christopher Schwarz, 2010, Strategic motives for hedge fund advertising, Working paper, University of California, Irvine.

Jotikasthira, Pab, Christian T. Lundblad, and Tarun Ramadorai, 2012, Asset fire sales and purchases and the international transmission of funding shocks, *Journal of Finance.* 67, 2015–2050.

Jylhä, Petri, 2011, Hedge fund return misreporting: Incentives and effects, Working paper, Imperial College London.

Liang, Bing, 2000, Hedge Funds: The living and the dead, *Journal of Financial & Quantitative Analysis* 35, 309–326.

Liang, Bing, 2003, The accuracy of hedge fund returns, *Journal of Portfolio Management* 29, 111–122.

Liang, Bing, and Hyuna Park, 2010, Predicting hedge fund failure: A comparison of risk measures, *Journal of Financial and Quantitative Analysis* 45, 199–222.

Ljungqvist, Alexander, Christopher Malloy, and Felicia Marston, 2009, Rewriting history, *Journal of Finance* 64, 1935–1960.

Patton, Andrew J., and Tarun Ramadorai, 2013, On the high-frequency dynamics of hedge fund risk exposures, *Journal of Finance* 68, 597–635.

Stigler, George J., 1961, The economics of information, *Journal of Political Economy*, 69, 213–225.

Streatfield, Michael P., 2012, Hedge funds: Fees, return revisions, and asset disclosure, Doctoral thesis, University of Oxford.

Sun, Zheng, Ashley Wang, and Lu Zheng, 2012, The road less traveled: Strategy distinctiveness and hedge fund performance, *Review of Financial Studies* 25, 96–143.

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's website:

Appendix SI: Internet Appendix.