

Outperforming the Market: Portfolio Strategy Cloning from SEC 13F Filings

Jan L. Schroeder (jan@sec-api.io)*

Abstract

Can mirroring of hedge fund strategies lead to market-outperforming returns? Our findings demonstrate that cloned portfolios in the top quartile, derived from SEC Form 13F filings, replicate the funds' performances and exceed the S&P 500 index by 24.3% on an annualized risk-adjusted basis. Analyzing over 150,000 portfolios between 2013 and 2023, we compare original versus replicated strategies across twelve performance metrics. Except for annualized volatility, maximum drawdown and tracking error, we demonstrate that cloned portfolios rebalanced on the disclosure date of filings, successfully mirror the performance of the original funds, including both market-underperforming and -outperforming funds.

Keywords: Alpha Cloning, Portfolio Replication, SEC 13F Filings, Institutional Investment Managers, Quantitative Analysis of Fund Performances and Portfolio Cloning

JEL Classification: G11, G12, G14, G23

Acknowledgements: We would like to thank Jonathan Lewellen, Professor of Finance at Tuck School of Business at Dartmouth, Nikolai Roussanov, Professor of Finance at Wharton, University of Pennsylvania, and Peter Posch, Professor of Finance at TU Dortmund for their invaluable comments and feedback.

*Jan L. Schroeder is a Research Fellow at the Department of Finance, TU Dortmund, Germany, and the founder and CEO of SEC-API.io. The authors have no competing interests to declare. All errors are our own.

1 Introduction

Inspired by Frank et al. (2004) and Verbeek & Wang (2013), we analyze the unexplored viability of replicating the investment strategies of top-performing institutional investment managers, commonly known as hedge funds, to determine whether such replication can emulate the performance of the original funds and ultimately achieve returns that surpass the performance of the SP500 benchmark index. By creating cloned portfolios from the holdings data disclosed in SEC 13F filings, we endeavor to mimic the performance characteristics of the source funds. Our analysis encompasses over 150,000 portfolios, while investigating more than 13,000 institutional investment managers from 1998 to 2023.

1.1 Registered Investment Companies vs Institutional Investment Managers

Clarifying the distinctions between registered investment companies and institutional investment managers—and by extension, mutual funds and hedge funds—is essential before discussing our paper. Mutual funds are classified as registered investment companies (RICs) under Section 3(c)(1) and Section 3(c)(7) of the Investment Company Act of 1940. There are three types of RICs: open-end management investment companies, closed-end management investment companies, and unit investment trusts (UITs). Open-end funds, commonly known as mutual funds, continuously offer new shares and redeem existing shares at their net asset value. They do not trade on exchanges like publicly listed companies or ETFs; instead, investors can purchase and redeem shares directly from the fund. An example of an open-end fund is the Vanguard Federal Money Market Fund (ticker: VMFXX). Hedge funds, while also being investment companies, are exempt from the above regulations and are classified as institutional investment managers (IIMs). They are primarily regulated under the Securities Exchange Act of 1934, particularly Section 13(f). An IIM is defined as an entity that exercises investment discretion over an account holding equity securities with an aggregate fair market value of at least \$100 million on the last trading day of any month of a calendar year. Unlike mutual funds, hedge funds are not identified by ticker symbols. Although all hedge funds with assets under management (AUM) exceeding \$100 million are IIMs, not all IIMs are hedge funds. Form ADV distinguishes seven types of IIMs: hedge funds, liquidity funds, private equity funds, real estate funds, securitized asset funds, venture capital funds, and private funds.

For readability, we will use the term “mutual fund” to refer specifically to “open-end management investment companies,” a type of registered investment company. When we refer to “hedge funds” or simply “funds,” we mean all “institutional investment managers”

that are regulated under Section 13(f) and thus disclose their holdings through Form 13F filings on the SEC EDGAR system. Our analysis focuses solely on institutional investment managers.

Both mutual funds and hedge funds must disclose their holdings publicly on a quarterly basis through the SEC’s EDGAR system. Mutual funds use Form N-PORT from 2019 (previously N-Q), while hedge funds use Form 13F (13F-E until 1999, and 13F-HR starting in 1998). It is important to note that hedge funds do not disclose regulated information such as their gross market values and portfolio holdings except through the SEC’s EDGAR system and FINRA. Consequently, any database that claims to provide monthly performance data on hedge funds is inherently unreliable, as there is no verifiable source of truth for such data. The only dependable sources are the quarterly published 13F filings and the annual Form ADV filings. Therefore, researchers should exercise caution and avoid relying on databases that provide hedge fund performance data with a resolution higher than quarterly. Our analysis exclusively focuses on Form 13F data, providing insights into hedge funds, rather than Form N-PORT data associated with mutual funds.

1.2 Prior Research

We analyze the viability of replicating the buy and sell activities of over 150,000 hedge fund portfolios to compare the performance of original and cloned portfolios across twelve performance and risk factors. Our study diverges from existing research by focusing exclusively on institutional investment managers, excluding mutual funds, and replicating entire portfolios on a quarterly basis rather than using subsets or factor-based replicas. Previous research on performance replication falls into three categories: hedge fund performance cloning, mutual fund performance cloning, and straightforward fund performance evaluation without cloning.

Table 1 provides an overview of previous research on hedge fund performance cloning, detailing each study’s sample size (number of funds analyzed), historical data period, cloning strategy employed, and authors.

Interestingly, all prior research in the hedge fund performance cloning category focuses on replicating specific aspects of portfolios, either through deriving particular factors (Hasanhodzic & Lo, 2007; Hayes, 2012; V. Weber & Peres, 2013; Hayes & Ba, 2014) or by copying only selected parts of portfolios (Antón et al., 2010; Bollen & Fisher, 2013; Agarwal et al., 2013; Aragon et al., 2013; DiPietro, 2019; Fleiss et al., 2021; Brown et al., 2021; S. S. Cao et al., 2021). Our paper is the first to replicate entire hedge fund portfolios comprehensively. While this approach is novel for hedge funds, similar strategies have been applied to mutual funds with success (Frank et al., 2004; Verbeek & Wang, 2013), demonstrating that disclosed

Sample Size	Data History (Years)		Cloning Strategy	Authors
1,610	1986 to 2005	(19)	Factor models using six factors to replicate individual fund exposures to produce linear clones	Hasanhodzic & Lo (2007)
1,662	1983 to 2018	(35)	Portfolio creation based on “best ideas” positions of hedge funds	Antón et al. (2010)
6	1990 to 2009	(19)	Estimate factor exposures of hedge fund indices to construct factor-based clones	Hayes (2012)
13	1994 to 2012	(18)	Clone returns using futures contract returns as explanatory variable	Bollen & Fisher (2013)
942	1999 to 2007	(8)	Return evaluation of positions disclosed after confidential treatment expires	Agarwal et al. (2013)
7,000	2006 to 2012	(6)	Clone returns using investible risk factors derived from futures contract prices	V. Weber & Peres (2013)
250	1999 to 2006	(7)	Individual position performance based on confidential treatment filings	Aragon et al. (2013)
21	Unknown		Returns captured using a beta regime-switching model	Hayes & Ba (2014)
1,331	2013 to 2020	(7)	Machine learning based to construct portfolios from a set of funds	DiPietro (2019)
1,331	2013 to 2020	(7)	Machine learning based factor models used to replicate fund returns	Fleiss et al. (2021)
1,500	2004 to 2017	(13)	Individual positions of funds selected based on crowding metrics	Brown et al. (2021)
643	2003 to 2017	(14)	Copycatted trades identified from digital footprints left on SEC EDGAR website	S. S. Cao et al. (2021)
13,035	2013 to 2023	(11)	Time-lagged replication of entire hedge fund portfolios	Our paper (2024)

Table 1: Overview of Prior Work in Hedge Fund Performance Cloning.

This table summarizes various studies on hedge fund performance replication, detailing the sample size, data history, cloning strategy employed, and the authors. The studies range from using factor models and subset portfolio creation based on “best ideas” positions to machine learning-based constructions and beta regime-switching models. Our study extends this body of work by implementing time-lagged replication of entire hedge fund portfolios.

holdings by active mutual funds can be effectively used by copycats to achieve comparable or superior net returns. Our methodology is primarily inspired by Frank et al. (2004) and Verbeek & Wang (2013), who examined the implications of mandatory portfolio disclosure by mutual funds, particularly the feasibility of “copycat” funds replicating disclosed portfolios to attain similar returns. These studies created hypothetical “copycat” funds that adjusted their holdings with each semi-annual and quarterly disclosure, respectively, revealing efficiencies in mimicking disclosed strategies. It’s crucial to distinguish between factor cloning and portfolio cloning. Prior research on hedge fund strategy cloning is primarily focused on factor-based cloning strategies. For example, Hasanhodzic & Lo (2007) developed “linear clones” to evaluate how well six distinct factors could predict the funds’ expected returns. Other examples of factor-based approaches include Hayes (2012), V. Weber & Peres (2013), Hayes & Ba (2014), DiPietro (2019), and Fleiss et al. (2021). Portfolio cloning can be further categorized into entire portfolio cloning and selected position cloning. While Frank et al.

(2004) and Verbeek & Wang (2013) are examples of cloning entire mutual fund portfolios, no such work has been undertaken for hedge funds. All prior research on hedge fund portfolio cloning focuses on selected position cloning (Agarwal et al., 2013; Antón et al., 2010; Aragon et al., 2013; Brown et al., 2021; S. S. Cao et al., 2021). The landscape of prior research focusing on the evaluation of fund performances and comparing performances across different types of funds, rather than replicating such performances through copycat portfolios, is much larger than the replication studies. Examples include comparing hedge fund and mutual fund performance (Ackermann et al., 1999; Griffin & Xu, 2009; Agarwal et al., 2009; Lewellen, 2011) and analyzing specific performance characteristics (Malkiel, 1995; Fung & Hsieh, 2000, 2001; Agarwal & Naik, 2003; Martin & Puthenpurackal, 2008; Sandvik et al., 2011; Sias et al., 2016; Amir-Ghassemi et al., 2022; S. Cao et al., 2021; Brown et al., 2021; Lazo-Paz et al., 2023).

Our paper ventures into an area not extensively explored previously due to several reasons. Anderson & Brockman (2018) shed light on inconsistencies within Form 13F institutional holdings, highlighting discrepancies between reported holdings and the SEC’s official list, including errors in market valuations. Therefore, one significant barrier has been the inaccuracies within the 13F filings themselves, which necessitate outlier detection algorithms and considerable time investment to clean and standardize the data for analysis. Furthermore, accurately mapping CUSIPs to ticker symbols, especially in a manner that avoids survivorship bias, poses a problem given that CUSIPs and tickers change over time and may not always have a direct equivalent. Securing historical pricing data that is free from survivorship bias presents another hurdle, as does obtaining accurate pricing for companies that have experienced changes in their ticker symbols or CUSIPs. These factors contribute to the complexity of assembling a comprehensive dataset that accurately reflects historical market movements and institutional behaviors. Lastly, the computational power and memory required to simulate the performance of over 150,000 portfolios spanning more than a decade is substantial. This computational demand has likely deterred prior attempts to undertake such an analysis.

2 Data, Hypothesis & Methods

2.1 Data

Our data collection and consolidation process leverages multiple Application Programming Interfaces (APIs) for the acquisition of unbiased datasets. We use data that is free from survivorship bias, incorporating both historical hedge fund information and security pricing data that account for all entities throughout the time period studied, regardless of their current status.

We sourced structured fund holdings data from SEC-API.io by utilizing their 13F Institutional Ownership API¹. This platform enabled us to systematically retrieve and organize holdings information from 13F-HR filings submitted to the SEC’s EDGAR system. Additionally, to facilitate the tracking of securities, we employed SEC-API.io’s Mapping API², which translates CUSIP identifiers into EDGAR Central Index Keys (CIK) and ticker symbols, including instances of ticker symbol changes. The extraction of gross market values reported in 13F filings was carried out through SEC-API.io’s Filing Download API³, which allows for the efficient bulk retrieval of 13F cover pages. Historical equity pricing data was sourced from AlgoSeek⁴, specifically their equity prices database and Security Master File. These resources provided not only historical adjusted and unadjusted price data for all holdings of each fund, but also facilitated the mapping of CUSIP identifiers to AlgoSeek’s internal security IDs over time, ensuring accurate price tracking even when a company has undergone ticker changes.

13F filings include two core files that are of interest to our analysis. The first file represents the cover page detailing the number of holdings and the aggregated value of reported assets, among others. The second file is the holdings information table attached to each 13F filing. It lists each invested company or instrument, specifying the security type (e.g., stock, note), quantity held, and position value. Figure 1 illustrates an example of a cover page of Ritholtz Wealth Management’s (CIK 1633901) 13F filing, while Figure 2 shows an example of a 13F information table.

¹<https://sec-api.io/docs/query-api/13f-institutional-ownership-api>

²<https://sec-api.io/docs/mapping-api>

³<https://sec-api.io/docs/sec-filings-render-api>

⁴<https://www.algoseek.com/products.html\#us.equity.market.data>

Report Summary:

Number of Other Included Managers: 1
Form 13F Information Table Entry Total: 753
Form 13F Information Table Value Total: 2,577,869,068
(round to nearest dollar)

Figure 1: Ritholtz Wealth Management’s 13F Cover Page Sample.

This image showcases a sample cover page from a 13F filing by Ritholtz Wealth Management, highlighting the disclosed number of holdings (753) and the aggregate portfolio value denominated in USD (\$2.58 billion). The cover page of the filing can be accessed through the provided link in the footnote⁵.

NAME OF ISSUER	TITLE OF CLASS	CUSIP	FIGI	VALUE (to the nearest dollar)	SHRS OR PRN AMT	SH/ PRN	PUT/ CALL	INVESTMENT DISCRETION	OTHER MANAGER	VOTING AUTHORITY		
3M CO	COM	88579Y101	BBG001S5T7X2	2,144,228	19,614	SH		SOLE		97	0	19,518
ABBOTT LABS	COM	002824100	BBG001S5N9M6	4,077,602	37,046	SH		SOLE		76	0	36,969
ABBVIE INC	COM	00287Y109	BBG002SY4RZ3	5,636,477	36,371	SH		SOLE		372	0	35,999
ACACIA RESH CORP	ACACIA TCH COM	003881307	BBG001SJLJT9	92,445	23,583	SH		SOLE		0	0	23,583
ACCENTURE PLC IRELAND	SHS CLASS A	G1151C101	BBG001SCXK90	1,693,320	4,826	SH		SOLE		138	0	4,687
ACME UTD CORP	COM	004816104	BBG001S5NCC0	223,066	5,205	SH		SOLE		0	0	5,205
ADOBE INC	COM	00724F101	BBG001S5NCQ5	11,026,958	18,483	SH		SOLE		86	0	18,397
ADTALEM GLOBAL ED INC	COM	00737L103	BBG001S8M5S3	972,086	16,490	SH		SOLE		175	0	16,315

Figure 2: Detail of Ritholtz Wealth Management’s Holdings from a 13F Filing

This image is a detailed excerpt from Ritholtz Wealth Management’s 13F information table file, showcasing the firm’s holdings. It includes columns for the issuer name, class of securities, identifiers such as CUSIP and FIGI, the total value of each holding rounded to the nearest dollar, and the number of shares held. The type of security is also indicated. Access to the complete source filing can be obtained through the provided link⁶.

2.2 Hypothesis Formulation & Testing

Null Hypothesis (H0): Cloned portfolios, sourced from 13F filings of institutional investment managers with a track record of market outperformance on a risk-adjusted basis from 2013 to 2023 and rebalanced on the filing disclosure date rather than quarter end date, do not match or exceed the performance of their source funds or the market benchmark.

Alternative Hypothesis (H1): Cloned portfolios that are rebalanced on the filing disclosure date, reflecting the strategies of institutional investment managers with historical

⁵https://www.sec.gov/Archives/edgar/data/1698218/000108514624000293/xslForm13F_X02/primary_doc.xml

⁶https://www.sec.gov/Archives/edgar/data/1698218/000108514624000293/xslForm13F_X02/infotable.xml

market outperformance on a risk-adjusted basis, duplicate or surpass the market performance achieved by the source funds.

To test these hypotheses, we conduct an analysis of the descriptive statistics for the differences in performance metrics between the cloned and original funds. We compare the distribution of each factor’s values for the cloned funds against those of the original funds. Using the Alpha factor as an illustrative case, we compute the difference (delta) between the cloned and original fund alphas to form a pairwise delta distribution. We then calculate and evaluate the descriptive statistics of this distribution, including the mean and the 25th, 50th, and 75th percentiles. For factors where higher values indicate superior performance, such as the Sharpe ratio and CAGR, we reject the null hypothesis for that factor if all calculated statistical metrics (mean, 25/50/75th percentiles) of the pairwise delta distribution are zero or positive. Conversely, for factors where lower values are preferable, such as maximum drawdown and annualized volatility, the null hypothesis for that factor is rejected if all metrics are zero or negative. The overall null hypothesis is rejected if the delta distributions for at least 8 out of the 12 factors show that the cloned portfolios perform at least as well as, or better than, the original portfolios.

To examine whether a delta distribution aligns with a normal distribution, we employ a Q-Q (quantile-quantile) plot. This graphical tool plots the actual quantiles from our dataset against the expected quantiles of a standard normal distribution. Ideally, if the data is normally distributed, the plot points should align closely with the diagonal 45-degree reference line. Any significant deviations from this line may indicate non-normality in the data distribution. For a more formal test of normality, we apply the Shapiro-Wilk test, which calculates a test statistic and an associated p -value. A p -value greater than 0.05 typically suggests that the dataset does not significantly diverge from normality, while a p -value below this threshold indicates non-normal distribution characteristics in the data. We conduct Levene’s test and an F-test to verify the assumption of equal variances across the delta distributions, which is a critical prerequisite for selecting the appropriate statistical follow-on test. Both tests serve to determine if multiple groups have similar levels of variance, with a high p -value supporting the hypothesis of equal variances. The Wilcoxon signed-rank test, a non-parametric method, is utilized to ascertain whether the delta distributions of each factor significantly deviate from zero. Given the real-world nature of financial data, it is reasonable to presume that the distributions of factors for both original and cloned funds – and consequently their delta distributions – are not normally distributed. Additionally, the assumption of equal variances and the criteria for independence are not satisfied. The connection between the cloned and original portfolios complicates the assumption of independent observations, a key requirement for several statistical tests, including the Mann-Whitney U test. Cloned

portfolios, by design, substantially duplicate the composition of their source portfolios. As a result, their performance is intrinsically tied to the performance of the shared holdings. This replication process generates a dependency, as changes in the value of shared positions simultaneously affect both the cloned and original portfolios.

Inspired by the methodologies of Brinson et al. (1986), Ackermann et al. (1999), Hayes (2012), and Agarwal et al. (2013), we have selected the S&P 500 index (ticker: SPY) as the benchmark for measuring the performance of the funds. Our analysis encompasses 12 factors, all annualized based on the approach of Paleologo (2021). In contrast to Verbeek & Wang (2013), our study addresses several key areas they missed: We consider volatility and other risk metrics, not just gross and net returns, our analysis goes beyond mean results to include the distribution of outcomes and we evaluate risk-adjusted alpha returns to provide a more comprehensive assessment of performance.

Factor and Description	Sources
Alpha: Fund's ability to outperform the SPY on a risk-adjusted basis. Higher = superior. $\alpha = R_p - (R_f + \beta(R_b - R_f))$ Where: • α is the annualized alpha of the portfolio based on quarterly returns. • R_p represents the annualized return of the portfolio, calculated as $R_p = \frac{1}{n} \sum_{i=1}^n r_{p_i} \times 4.$ • R_b represents the annualized return of the benchmark, calculated as $R_p = \frac{1}{n} \sum_{i=1}^n r_{b_i} \times 4.$ • r_{p_i} and r_{b_i} represent the quarterly returns of the portfolio and SPY, respectively, of the i-th quarter. • R_f is the risk-free rate of 2%. • β is the beta of the portfolio.	Jensen (1968), Ackermann et al. (1999), Fama & French (2004)
Beta: Fund's market volatility in relation to SPY. A value closer to 0 suggests lower market-related volatility and is considered more favorable. $\beta = \frac{\text{Cov}(R_p, R_b)}{\text{Var}(R_b)}$ Where: • β is the beta of the portfolio. • R_p represents the quarterly returns of the portfolio. • R_b represents the quarterly returns of SPY. • $\text{Cov}(R_p, R_b)$ is the covariance between the portfolio returns and the SPY returns. • $\text{Var}(R_b)$ is the variance of the SPY returns.	Jensen (1968), Blume (1975), Wallerstein et al. (2009), Sandvik et al. (2011)

Factor and Description	Sources
Arithmetic Mean Return (AMR): Fund's average annual return. Higher = superior. $\text{AMR} = \left(\frac{1}{n} \sum_{i=1}^n x_i \right) \times 4$ Where: • x_i represents the quarterly returns of the portfolio. • n is the number of quarters the fund reported 13F filings that can be cloned.	Mitchell & Pulvino (2001), O'Shaughnessy (2005), Wallerstein et al. (2009)
Geometric Mean Return (GMR): Fund's compound average annual return. Higher = superior. $\text{GMR} = \left(1 + \frac{1}{n} \sum_{i=1}^n x_i \right)^4 - 1$ Where x_i and n are the same as for AMR.	
Cumulative Average Growth Rate (CAGR): Fund's mean annual growth rate. Higher = superior. $\text{CAGR} = \left(\prod_{i=1}^n (1 + r_i) \right)^{\frac{4}{n}} - 1$ Where: • r_i represents the portfolio's quarterly returns. • n is the number of quarterly returns. • $\prod$ denotes the product over all quarterly returns.	Ackermann et al. (1999)
Excess Return: Difference between the fund's annualized return and that of the SPY benchmark. Higher = superior. $\text{Annualized Excess Returns} = R_p - R_b$ Where R_p and R_b are the same as for alpha.	Antón et al. (2010), Hayes (2012), Bollen & Fisher (2013)
Tracking Error: Standard deviation of the fund's annualized excess returns relative to SPY. Lower tracking error reflects more consistent performance in relation to the benchmark. $\text{Annualized Tracking Error} = \sigma_{(R_p - R_b)} \times \sqrt{4}$ Where: • R_p represents the individual quarterly returns of the portfolio. • R_b represents the individual quarterly returns of the benchmark. • $\sigma_{(R_p - R_b)}$ is the standard deviation of the difference between the portfolio returns and the benchmark returns. • The time factor for annualization is $\sqrt{4}$ to account for quarterly returns.	
Information Ratio: Annualized excess return divided by the tracking error to evaluate the consistency of outperformance relative to the benchmark. A higher information ratio is indicative of a fund manager's efficient active management. $\text{Information Ratio} = \frac{\text{Excess Return}}{\text{Tracking Error}}$	

Factor and Description	Sources
Sharpe Ratio: A metric for risk-adjusted return, computed as the fund's annualized return over its annualized volatility with a risk-free rate set at 2%. Higher = superior. $\text{Sharpe Ratio} = \frac{R_p - R_f}{\sigma_p}$ • R_p is the annualized return of the portfolio, calculated as $R_p = \frac{1}{n} \sum_{i=1}^n r_{p_i} \times 4$. • R_f is the risk-free rate of 2%. • σ_p is the annualized standard deviation of the portfolio returns, calculated as $\sigma_p = \sigma \times \sqrt{4}$. • r_{p_i} represents the individual portfolio returns. • σ represents the standard deviation of the portfolio returns.	Sharpe (1994), Ackermann et al. (1999), Lo (2002), Hasanhodzic & Lo (2007), Wallerstein et al. (2009)
Sortino Ratio: Focuses on downside volatility with a minimum acceptable return set at 2%. Higher Sortino Ratios indicate better downside risk management. $\text{Sortino Ratio} = \frac{R_p - \text{MAR}}{\sigma_{\text{downside,annualized}}}$ Where: • R_p is the annualized return of the portfolio, calculated as $R_p = \frac{1}{n} \sum_{i=1}^n r_{p_i} \times 4$. • r_{p_i} represents the quarterly portfolio returns. • MAR is the annualized minimum acceptable return of 2%. • $\sigma_{\text{downside,annualized}}$ is the annualized downside deviation, calculated as $\sigma_{\text{downside,annualized}} = \sigma_{\text{downside}} \times \sqrt{4}$. • σ_{downside} is the standard deviation of the negative excess returns.	Sortino & Price (1994), Pedersen & Satchell (2002), Chaudhry & Johnson (2008)
Volatility: Standard deviation of the fund's quarterly returns, providing an understanding of return variability. Lower annualized volatility represents more stable performance. $\sigma_{\text{annualized}} = \sigma \times \sqrt{4}$ Where: • σ is the standard deviation of the returns. • $T_{\text{return period}} = 4$ for quarterly returns.	Hasanhodzic & Lo (2007), Markowitz & Blay (2013), Brown et al. (2021)
Maximum Drawdown per Quarter (Worst Single Quarter): The largest peak-to-trough loss in the fund's value over each quarter. A lower maximum drawdown is indicative of lower risk and better performance during downturns. $\text{Max Drawdown} = \min(R_{\text{quarterly}})$ Where: • $R_{\text{quarterly}}$ represents the series of quarterly returns. • $\min(R_{\text{quarterly}})$ denotes the minimum value in the series of quarterly returns.	Ackermann et al. (1999), Markowitz & Blay (2013), Brown et al. (2021)

2.2.1 Methodology

Our analysis starts with the computation of performance metrics for each fund, utilizing the gross market value (GMV) of their holdings as indicated on the 13F cover pages. To ensure the integrity of our dataset, we excluded funds with less than six consecutive 13F filings, thereby confirming that each included fund was operational and complied with SEC disclosure regulations for a minimum of one year. Additionally, we corrected identifiable inaccuracies within the 13F disclosures. Such corrections were necessary for instances where funds reported GMVs in incorrect magnitudes – e.g., denominating dollar values in millions instead of thousands – or inaccurately swapped the figures for the number of holdings and GMVs. Funds that omitted a quarterly report or failed to disclose their GMV in any 13F filing were also excluded from our analysis.

To calculate Alpha and Beta, we utilized the percentage change in reported GMVs across each reporting period to represent a fund’s quarterly returns. Specifically, quarterly returns were computed for the periods of January 1st to March 31st (Q1), April 1st to June 30th (Q2), July 1st to September 30th (Q3), and October 1st to December 31st (Q4) annually. It is crucial to note that fund returns were calculated based on these fixed quarterly periods, rather than the intervals between 13F disclosures, to prevent inaccuracies in performance measurement that could skew Alpha and Beta calculations. The S&P 500 index, represented by the ticker SPY, served as the benchmark for these calculations, with SPY’s quarterly returns computed over the identical periods. To accommodate quarters beginning or ending on weekends or public holidays, we adjusted the starting or ending price of SPY by carrying forward the last available closing price. For instance, if Q2 begins on a Saturday, we use the previous Friday’s adjusted closing price as the starting point for Q2’s return calculation:

$$\text{Q2 Return (SPY)} = \left(\frac{\text{Adj Close (June 30th)} - \text{Adj Close (April 1st)}}{\text{Adj Close (April 1st)}} \right)$$

Here, the adjusted closing price for April 1st equals that of March 31st if the markets were closed on April 1st.

To replicate a fund’s portfolio, we gathered all relevant holdings data (CIKs, CUSIPs, reported value, and number of shares for each position) from the 13F information tables, referencing the public disclosure date of the filing rather than the period-end date. Holdings deemed unreplicable – such as options, warrants, or notes, which either lack necessary disclosure for replication (e.g., strike prices and expiration for options) or are insufficiently liquid for trading simulation – were excluded from the cloning process. The replication of a portfolio was then straightforwardly achieved by duplicating all its clonable holdings.

We conducted backtesting simulations that involved executing buy and sell orders for

each holding based on their reported shares at the disclosure date, using unadjusted closing prices. For instance, if Fund A reported holding 100 shares of Stock S in Q1 and then reported an increase to 150 shares in Q2 on the disclosure date T , indicating a purchase of 50 shares, we simulated buying an additional 50 shares of S at its unadjusted closing price on date T . Conversely, if Fund A's holdings of Stock S decreased from 100 shares in Q1 to 50 shares in Q2, we simulated selling 50 shares at Stock S's unadjusted closing price on date T . Following Verbeek & Wang (2013) approach, we compute the quarterly return of a cloned portfolio as the return of a simulated buy-and-hold portfolio that invests in the most recently disclosed equity positions and is rebalanced at the next 13F disclosure date. The quarterly return QR for a cloned portfolio targeting fund j is defined as:

$$QR_t^j = \sum_{i=1}^N \tilde{w}_{i,t-1}^j R_{i,t}$$

Where $R_{i,t}$ denotes the return on the equity position i , and the value weights are given by

$$\tilde{w}_{i,t-1}^j = \frac{N_{i,t-1}^j P_{i,t-1}}{\sum_{i=1}^N N_{i,t-\tau}^j P_{i,t-1}}$$

Where $N_{i,t-\tau}^j$ denotes the number of shares of stock i held by fund j disclosed in the fund's most recent 13F filing at time $t-\tau$, and $P_{i,t-1}$ is the stock price at the date when the portfolio was previously rebalanced, i.e. approximately three months prior to t . Unlike the original portfolios, which calculate quarterly returns based on the calendar quarter's first and last day, cloned portfolios' returns were calculated based on the disclosure dates to avoid data leakage or foresight bias, as cloned positions could not have been known at the quarter's end.

To evaluate our hypothesis, we calculated the distributions of all factors for both original and cloned portfolios, categorizing the number of funds within specific ranges (e.g., 10 funds with an alpha of 0.2%, 18 funds with an alpha of -0.03%). Given that the distributions did not adhere to normality and equal variance as well as independence was not observed, we opted for the Wilcoxon Signed-Rank Test to assess the statistical significance of differences in all factor distributions between the original and cloned portfolios. This approach ensures a rigorous examination of the performance differential without the constraints of parametric assumptions.

2.2.2 Limitations of Our Methodology

One significant constraint of using 13F data in our analysis is its exclusion of cash flow activities, such as cash inflows from additional capital raised by funds or cash redemptions. In scenarios where a fund acquires extra securities through new capital, our method replicates these purchases without accounting for the cloned portfolio’s cash levels. Adapting our cloning strategy to scenarios with less or limited capital would necessitate a recalibration of portfolio weights, possibly incorporating fractional shares for smaller positions to maintain proportional accuracy. Contrary to Kacperczyk et al. (2006) and Verbeek & Wang (2013), our analysis does not incorporate trading costs, such as execution costs for buying or selling shares on exchanges like NASDAQ or NYSE. This exclusion is based on the significant decline in execution costs over the past decades and the availability of market-making executions, where entities can earn fees rather than incur costs by buying at the bid and selling at the ask. Additionally, we do not account for liquidity constraints. Our methodology involves buying or selling the delta of positions in a single day at the closing price, without considering the spread or liquidity. This approach may not accurately reflect the true nature of executing large-scale positions, as funds must adhere to internal risk parameters that limit the proportion of average daily dollar volume they can trade. Consequently, if a position size exceeds these thresholds, a fund would typically scale in or out of the position over several days to weeks, depending on its size. 13F filings are disclosed after the quarter ends, typically a few days to several weeks later. Consequently, our analysis compares the quarterly returns of source portfolios with the lagged quarterly returns of cloned portfolios, leading to a slight mismatch in the return calculation periods. This discrepancy introduces a potential distortion, as the return periods for the original and cloned portfolios do not align perfectly. However, we anticipate that any resulting error will remain relatively stable across the board due to the regulatory consistency in the timing of 13F filings. Given our extensive sample size of over 150,000 portfolios, we expect that such discrepancies will not significantly affect the overall results of our study.

We observed that some funds invest in highly correlated or identical securities, such as allocating 3% of the fund’s GMV to the SP500 ETF (SPY) and another 2% to a similar index under a different ticker (IVV). Similarly, investments in NASDAQ 100 (QQQ) alongside its largest constituents like APPL, AMZN, and META result in overlapping exposures. While such strategies might be rational for large institutional managers adhering to specific risk parameters, they are less efficient for cloning with significantly less capital. In these instances, our cloning algorithms could be refined to filter out redundancies, reallocating capital from omitted positions to enhance the portfolio’s diversification and efficiency.

Additionally, the selective inclusion of source positions in the cloned portfolio introduces

a potential selection bias. This bias may reflect the underlying criterion (tradeability at the time of 13F disclosure) that could inadvertently misalign the cloned portfolio’s performance with the strategic choices made for the source portfolio. Our criteria for excluding positions were solely based on the tradability of a security, assessed through its linkage to a ticker on a U.S. exchange, which may influence the comparability and performance outcomes of our cloned portfolios.

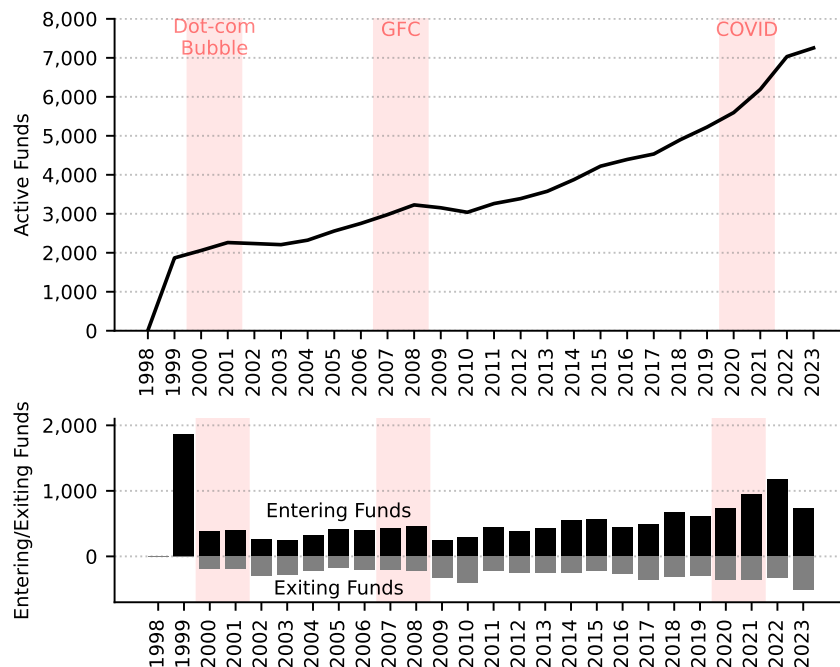
In practice, we successfully replicated a high-performing portfolio with over 700 holdings, scaling down its GMV from \$7 billion to \$700K for simulation in a paper trading account through Interactive Brokers’ APIs. However, manually replicating and rebalancing portfolios, especially those with over 100 holdings – which accounts for 50% of all funds – every quarter is not feasible without automated software assistance, as we’ve developed for this research.

3 Results & Discussion

3.1 Fund Lifespans and Survival Rates

We present the results of our analysis on the total number of active funds from 1998 to 2023 in Figure 3. This analysis allows for discerning trends in fund creation, offering insights into whether the asset management industry is experiencing growth or decline. Over the past 26 years, from 1998 to 2023, our analysis reveals a dynamic landscape in the fund management industry, involving 13,035 hedge funds. Notably, 2023 emerged as a landmark year, showcasing the highest number of active funds, with a remarkable surge in new entrants from under 500 in 2016 to over 1,000 by 2022. This period witnessed the greatest influx of newly registered funds, highlighting a significant shift in the industry’s composition. The year 2022 stood out not only for its record-setting new fund registrations but also for signaling a pronounced market adjustment. The repercussions of this strong market correction became particularly evident in 2023, a year that saw over 500 funds withdrawing from the investment arena, surpassing the exit rates observed in the aftermath of the Global Financial Crisis (GFC) in 2009 and 2010. This trend underscores a notable correlation between market downturns or recessions and an increase in the number of funds ceasing operations, with the peak of exits often coinciding with the transition from a bear to a bull market. This suggests that many funds tend to exit the market just as conditions begin to improve. For clarity, “entering funds” are identified as those reporting a Form 13F-HR for the first time, whereas “exiting funds” are recognized as those with previous 13F-HR filings but none in the given year. Given the SEC’s mandate requiring quarterly disclosures (four times per year), the absence of any filings from a fund within a year is a strong indicator of its operational

cessation.



Year	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
Active Funds	8	1,868	2,058	2,262	2,236	2,208	2,323	2,558	2,753	2,978	3,228	3,154	3,038
New Funds	8	1,860	388	398	271	252	331	418	395	433	466	253	291
Exiting Funds	0	0	198	194	297	280	216	183	200	208	216	327	407

Year	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023
Active Funds	3,262	3,389	3,577	3,876	4,221	4,393	4,533	4,901	5,224	5,596	6,191	7,032	7,254
New Funds	448	379	436	552	564	445	497	677	618	734	949	1,175	739
Exiting Funds	224	252	248	253	219	273	357	309	295	362	354	334	517

Figure 3: Trends in Active, Entering, and Exiting Investment Managers (1998–2023).

The top plot tracks the annual count of active investment managers filing at least one Form 13F with the SEC, identified by unique CIK numbers, from 1998 to 2023. Highlighted periods indicate significant market events: the Dot-com Bubble, the Global Financial Crisis (GFC), and the COVID-19 pandemic. The corresponding bar chart below shows the dynamics of market entry and exit among funds. “Entering Funds” denotes new managers filing a 13F for the first time, while “Exiting Funds” represents those that did not file a 13F in a given year after having filed in the preceding year. The two accompanying tables quantify the visual data, summarizing the number of active, new, and exiting funds per year, providing a detailed numerical basis for the graphical trends observed above.

Figure 5 illustrates the survival rates and operational longevity of funds from 1998 to 2023. We quantify the duration of activity for each fund by tracking the number of years it reported Form 13F filings to the SEC via EDGAR. This step involves plotting the data on a histogram where the x -axis represents the number of active years, and the y -axis shows the

count of funds for each duration of activity. Utilizing the entire period from 1998 to 2023, we create 26 distinct buckets to categorize the funds based on the number of years they were operational. This approach allows us to visualize the distribution of fund lifespans across the observed timeframe. Descriptive statistics capture the central tendency and dispersion of fund lifespans. These statistics include the mean (average number of active years), standard deviation (variation in lifespan), and the first, second (median), and third quantiles.

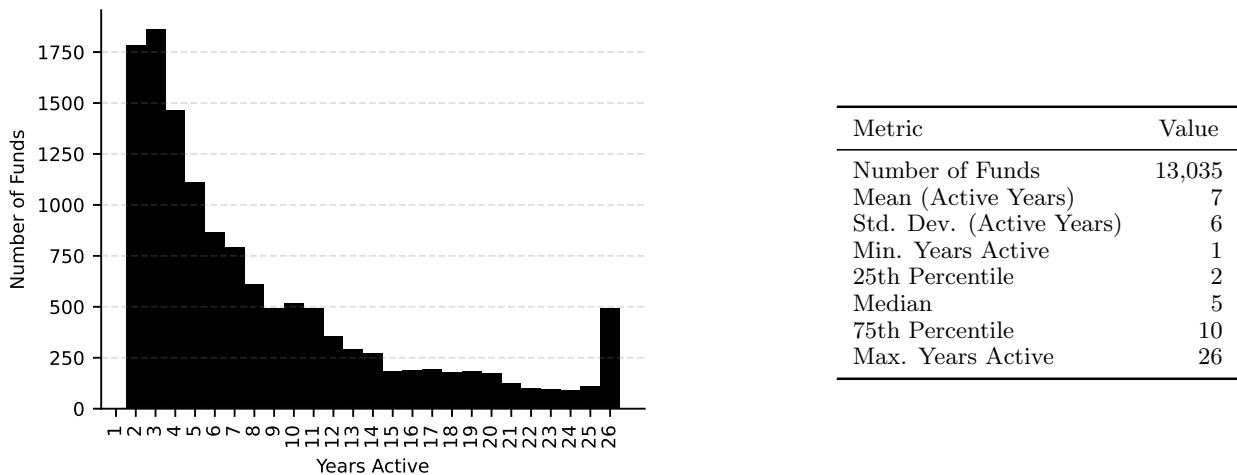


Figure 4: Fund Life Span Distribution and Descriptive Statistics (1998–2023).

The histogram displays the distribution of fund longevity, detailing the number of years funds remained active out of the 13,035 that have filed Form 13F-HR with the SEC from 1998 to 2023. Each fund’s activity duration is identified by its unique CIK, with a fund counted as active in any year it filed a 13F-HR form. The accompanying table provides descriptive statistics on fund survival, revealing that the median operational lifespan of a fund is only 5 years. Notably, about 3.8% of the funds exhibit exceptional longevity, operating for over 26 years, demonstrating significant endurance in the investment landscape.

Our survival rate analysis, encompassing 13,035 funds operating between 1998 and 2023, reveals an average fund lifespan of 7 years and a median lifespan of 5 years. This contrasts with Ackermann et al. (1999), who analyzed 547 hedge funds from 1987 to 1995 and found a mean fund age of 5.22 years and a median age of 4.01 years. Comparing their results with our more recent data, we conclude that hedge funds have successfully extended their lifespan by an average of 2 years, with a median increase of 1 year. In contrast to the hedge fund industry, the mutual fund industry demonstrates longer lifespans. Kacperczyk et al. (2006) found that mutual funds had a mean age of 13 years and a median age of 8 years during the period from 1984 to 2002. Similarly, Verbeek & Wang (2013) analyzed 3,046 mutual funds from 1985 to 2008 and reported a mean age of 12.63 years and a median age of 7.75 years. These comparisons highlight that mutual funds typically enjoy longer operational periods compared

to hedge funds. The standard deviation of 6 years highlights a significant variance in fund lifespans, indicating that while some funds operate for a brief period, others demonstrate much longer endurance. A closer look at the distribution shows that the minimum active duration is just 1 year, reflecting the challenges some funds face in sustaining operations. The 25th percentile, at 2 years, implies that a quarter of the funds ceased filings within two years of operation. At the higher end, the 75th percentile reveals that three-quarters of the funds were active for 10 years or less, and 25% of all funds were active for 11 years and more, underscoring that a significant portion of the sector does not surpass a decade of activity. Remarkably, the maximum lifespan extends to 26 years, showcasing that just under 500 funds have managed to navigate the asset management industry successfully over the entire period analyzed.

3.2 Evolution of Asset Values

In this part, we delve into the progression of asset values from 2013 to 2023 by analyzing the holdings values reported in 13F filings and report our findings in Figure 4. This analysis aims to shed light on the performance of institutional portfolios over time, revealing how they've adapted to shifting market conditions.

Figure 4 presents the time series of the aggregated gross market values of all portfolios in USD reported quarterly by investment managers, delineating the standard reporting periods that conclude at the end of March, June, September, and December. The trend from the beginning of 2013 to the onset of 2020 is characterized by a relative stagnation in holding values, suggesting a period of flat growth across the combined portfolios of the investment industry. During this timeframe, the growth in total holdings value does not exhibit significant upward momentum. A pivotal shift in the trajectory of holding values is marked by the COVID-19 pandemic. From 2020 to the end of 2021, there is an escalation in holding values, reflecting perhaps a reactive surge in investment during the pandemic period. In the fourth quarter of 2021, the aggregated assets reached an unprecedented peak, surpassing \$60 trillion. This surge is subsequently countered by a downturn that coincides with the market correction in 2022, indicating a retraction of holding values in response to changing economic conditions.

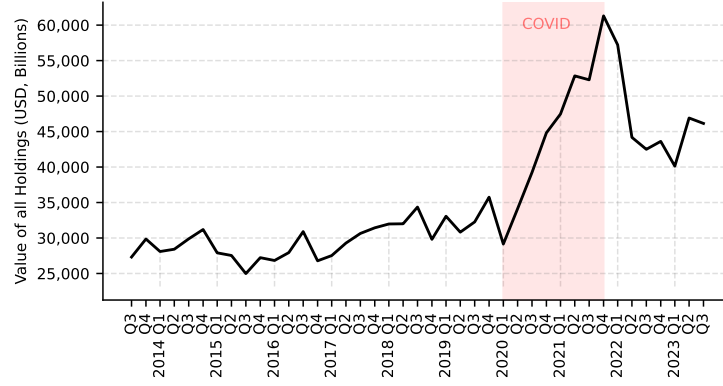


Figure 5: Aggregate Gross Market Value of 13F Portfolios (Q3 2013 – Q3 2023).

This time series graph traces the aggregate gross market value (GMV) of all portfolios reported in 13F filings from the third quarter of 2013 to the third quarter of 2023, measured in USD billions. Quarterly GMVs remained relatively stable from Q3 2013 until the end of Q1 2020, fluctuating between \$25 trillion and \$35 trillion. However, during the COVID-19 period, marked by unprecedented quantitative easing measures from the Federal Reserve, the GMVs soared, reaching a peak of over \$60 trillion. This spike reflects the infusion of liquidity into the markets, leading to an all-time high in the aggregated market value of the portfolios. The timeline demonstrates a retraction in 2022, indicative of the market correction and reduction in combined holdings across reported funds.

Year	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023
N	3,243	3,600	3,879	3,971	4,195	4,563	4,800	5,179	5,798	6,424	6,661
Mean	9,206	8,666	6,857	6,745	7,493	6,533	7,448	8,654	10,573	6,790	6,928
Median	497	466	413	413	461	361	414	459	475	323	326
Std. dev.	84,340	80,370	46,367	48,468	64,454	58,665	70,888	84,746	108,680	77,890	81,822
25th	205	195	170	177	193	157	186	202	219	149	152
75th	1,883	1,709	1,463	1,505	1,510	1,205	1,302	1,417	1,538	1,004	1,017
Min	0.1	0.1	0.1	0.1	0.1	0.1	0.2	0.1	0.1	0.1	0.1
Max	3,005,627	3,393,645	1,487,144	1,792,633	2,286,847	2,225,644	2,932,328	3,447,554	4,404,582	3,672,651	4,070,961

Table 3: Descriptive statistics of the gross market values (GMV) in millions of dollars per fund from 2013 to 2023.

The table shows the number of funds (“N”) each year, along with the mean, median, and selected percentiles of GMV. The statistics indicate a heavily right-tailed distribution with significant disparities between the mean and median GMVs, influenced by a small number of funds managing exceedingly large assets.

Year	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023
N	3,243	3,600	3,879	3,971	4,195	4,563	4,800	5,179	5,798	6,424	6,661
Mean	1,090	741	697	692	649	592	708	782	709	795	417
Median	103	101	94	98	101	93	99	105	107	94	92
Std. dev.	15,751	9,226	10,189	11,687	8,486	6,969	11,590	11,190	11,210	17,594	3,110
25th	41	39	34	36	36	33	37	41	41	35	36
75th	259	250	238	244	250	240	245	246	264	224	220
Min	1	1	1	1	1	1	1	1	1	1	1
Max	537,814	323,481	522,844	665,096	442,915	288,190	555,527	430,862	696,810	952,754	157,613

Table 4: Descriptive statistics for the number of holdings per fund from 2013 to 2023.

The table shows the number of funds (“N”) each year, alongside the mean, median, and selected percentiles for the number of holdings. The statistics reveal variations in portfolio sizes, with some funds holding a vast number of securities while others maintain a more concentrated portfolio, also indicating a heavily right-tailed distribution with significant disparities between the mean and median number of holdings per fund.

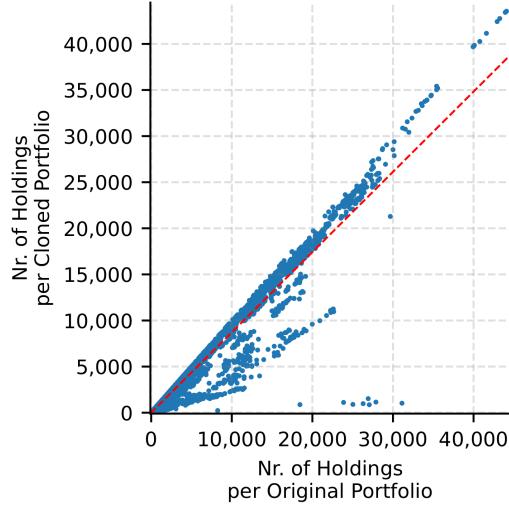
Table 3 and Table 4 reveal that the mean value is insufficient to accurately describe the GMV’s and number of holdings per fund. For instance, in 2023, the mean fund GMV was \$6.93 billion, whereas the median GMV was significantly lower at \$0.33 billion. The same heavy right-tail distribution is observed in the number of holdings, with the mean number of holdings per fund at 417 positions in 2023, while the median number of positions is just 92. This distortion arises because a small fraction of funds manage disproportionately large assets, with some exceeding \$40 billion. Between 2013 and 2023, 75% of all funds managed GMVs between \$1 billion and \$1.8 billion, while only 2.5% exceeded the \$32 to \$48 billion threshold. We also observed a downward trend in the median GMV, which declined from \$496 million in 2013 to \$326 million in 2023. This trend is mirrored in the top 2.5% of funds, where the threshold for inclusion in this elite group dropped from \$45 billion in 2013 to \$34 billion in 2023. These trends can be attributed to the doubling of the number of active funds from 2013 to 2023. Newer, younger funds typically start with less capital, and the number of new funds reached an all-time high in 2022 with 1,175 entries. Consequently, while the total GMV remained relatively stable over the decade, the increased number of funds meant that more entities shared the overall asset pool, except for the period affected by COVID-19 from 2020 to 2022.

3.3 Performance of Cloned vs Original Portfolios

To confirm the effectiveness of our replication process for portfolios from 2013 to 2023 and ensure a sufficiently large sample size for statistically significant conclusions, we analyzed the correlation between the number of holdings disclosed in 13F filings and those in the replicated portfolios. This comparison was visualized in a scatterplot in Figure 6, with the original portfolio holdings on the x -axis and the replicated portfolio holdings on the y -axis. A linear regression analysis yielded an R^2 value of 0.9535, indicating that 95.35% of the variance in the number of holdings between original and replicated portfolios can be explained by the linear relationship, with a statistically significant p -value demonstrating the regression’s reliability.

The regression’s standard error was notably low at 0.0005, and the slope of the regression line was 0.8705, closely approaching 1, alongside an intercept of 11.92. These figures collectively suggest a high degree of accuracy in the replication process across 161,408 portfolios examined.

Descriptive statistics provide further evidence of the accuracy of our replication process. Specifically, at the 25th percentile, the original portfolio contained 42 holdings, while their replicated counterpart had 38, indicating a slight deviation. Moreover, at the 75th percentile,



	Original	Cloned
Number of Portfolios	161,408	161,408
Mean (Holdings per Port.)	398	358
Std. Dev. (Holdings per Port.)	1,508	1,344
Min	3	3
25%	42	38
50%	102	95
75%	246	230
Max	44,083	43,546

Figure 6: Replication Accuracy of Portfolio Holdings from 13F Filings (2013–2023).

The left side of the figure presents a scatterplot comparing the number of holdings for each original portfolio, as reported in 13F filings by the source fund, with the number of holdings successfully replicated, across 161,408 portfolios between 2013 and 2023. The tight clustering around the linear regression line (in red), with an R^2 value of 0.9535, illustrates a high level of replication accuracy—indicating that 95.35% of the variance in the number of holdings is accounted for by the replication process. The slope of the regression line is 0.87, which is close to 1, suggesting that the cloned portfolios closely match their original counterparts in terms of the number of holdings. Outliers in the plot are clearly discernible and are due to holdings such as options and warrants, which were not replicated as part of our methodology. The right-hand table provides supporting descriptive statistics, highlighting that the median number of holdings in original portfolios is 102 compared to 95 in the cloned portfolios, further confirming the efficacy of the replication approach.

there was a noticeable, yet modest, difference with the original portfolio having 246 holdings compared to 230 in the replicated portfolio. This disparity, as anticipated, is due to the exclusion of certain assets such as options, notes, and warrants from the replication process.

The Spearman’s rank correlation coefficient calculation resulted in a value of 0.9926, accompanied by a statistically significant p -value of less than 0.01. This outcome underscores the efficacy of our cloning approach, revealing a near-perfect rank-order correlation between the original and replicated portfolios, which indicates a successful replication strategy.

3.4 Example of a Single Replication Process & Analysis

To illustrate and explain our methodology, we examine the case of Ritholtz Wealth Management (Ritholtz) with CIK 1698218, identified as one of the top-performing funds in our analysis. Ritholtz initiated its reporting with a 13F filing for Q2 2017 (April 1st to June 30th, 2017), submitted on August 4th, 2017. This date also marks our first replication of the fund’s holdings. Over 25 quarters, spanning from 2017 to 2023, the fund submitted 25

Form 13F filings, each revealing an updated portfolio.

For each new 13F disclosure, we analyzed the fund’s securities holdings (e.g., 200 shares of Apple) and calculated adjustments needed between the current and preceding filings to rebalance the replicated portfolio accordingly. Throughout this period, our replication focused solely on reproducible positions, excluding instruments like options.

The GMV of both the source and the replicated portfolios was determined for each reporting period. The GMV of the source portfolio was directly extracted from its 13F filing, while the GMV of the cloned portfolio was calculated by multiplying the number of shares by the unadjusted closing prices of the respective securities.

The bottom right time series plot in Figure 7 visually compares the normalized GMV of the original fund’s portfolios, the cloned portfolios, and the SPY benchmark, each standardized to start at \$100,000 in Q2 2017. This comparison effectively showcases the source fund’s consistent outperformance of the SPY benchmark throughout the analysis period and the cloned portfolio’s accurate reflection of the source fund’s performance trajectory.

The cloneability of the holdings is depicted in the bottom left dual bar chart, comparing the number of holdings in the original and cloned portfolios for each 13F filing. Apart from the COVID-19 period, the cloned portfolios matched the original in terms of the number of holdings, demonstrating our strategy’s efficacy.

Additionally, the asset allocation breakdown by security type is presented in the top right plot, highlighting the portfolio’s investment distribution across ETFs, US and Canadian stocks, CEFs, and ADRs. Notably, in the initial two years, the fund achieved benchmark outperformance primarily through a 95% allocation in ETFs, with the remainder in US stocks. This figure demonstrates our cloning strategy’s ability to accurately replicate diverse asset classes, thereby closely mirroring the performance of the source fund.

3.5 Factor Comparison for all Source and Cloned Portfolios

Upon constructing both the source and cloned portfolios, we calculated their respective performance factors as delineated in the Data & Methods section. Figure 8 serves as an exemplar for Ritholtz Wealth Management, exhibiting the performance and risk metrics of the original fund with those of its cloned counterpart. To derive these metrics for Ritholtz and its replicate, we conducted a backtest that involved simulating the trading activities of the cloned portfolios. This entailed executing buy and sell orders to rebalance the clone in tandem with the disclosure dates of Ritholtz’s subsequent 13F filings.

This methodology was consistently applied across the dataset to generate the suite of the 12 performance and risk factors for all original and cloned entities. Focusing on the

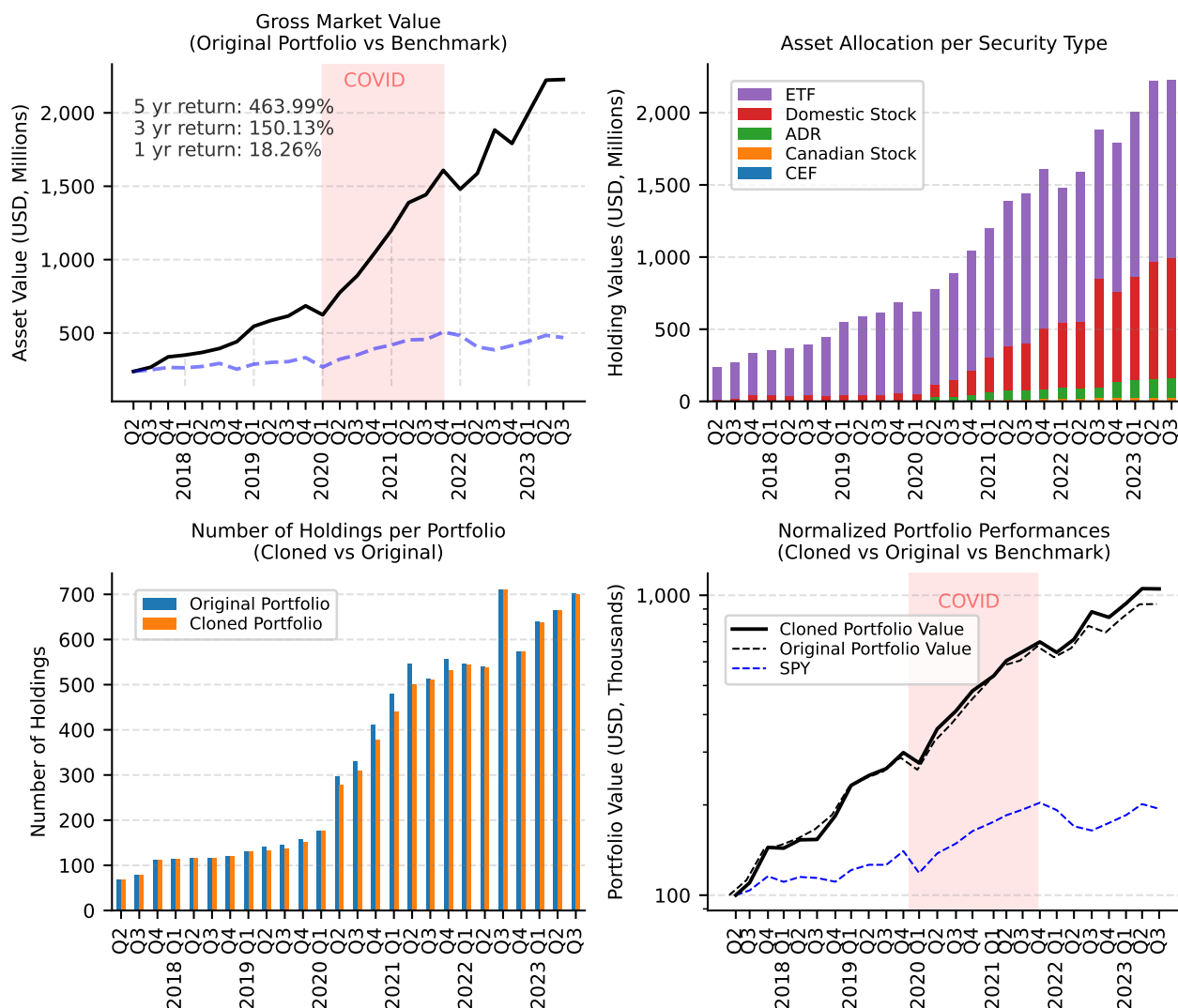


Figure 7: Representations of Original and Cloned Portfolio Metrics and Assets Allocation across Time.

The figure features four visualizations that compare the performance and asset composition of source portfolios of Ritholtz Wealth Management (Ritholtz) and cloned counterparts, track their holdings over time, and benchmark them against the SPY index. Top Left: Showcases the time series of gross market value (GMV) of the original portfolios disclosed in 13F filings by Ritholtz versus the SP500 index (ticker: SPY), based on equal initial investments, that being the starting value of the source fund. This subplot, covering Q2 2017 to Q3 2023, reveals returns for 1-year (18.26%), 3-year (150.13%), and 5-year periods (463.99%), with values expressed in millions of USD. Top Right: Displays a stacked bar chart detailing quarterly asset allocation by security type within each source portfolio. Categories include ETFs, US and Canadian stocks, CEFs, and ADRs, with the total asset value per quarter matching the GMV from the top left plot. Bottom Left: A comparative bar chart illustrates the number of instruments (holdings) per portfolio as reported in each 13F filing, for both original and cloned portfolios. The chart traces the evolution from an identical count in Q2 2017 through a divergence starting in early 2019 due to unreplicable options, a widening gap during COVID, and a recent convergence in Q4 2022, where the cloned portfolio achieves 99.7% replication of the source's holdings. Bottom Right: This subplot presents a time series of normalized GMVs, each starting at \$100,000, for cloned and original portfolios alongside SPY, serving as a benchmark. It visually underscores the effectiveness of the cloning strategy in mirroring the original portfolios' GMV and, by extension, their performance, with both consistently outperforming SPY.

	Original	Cloned
General		
CIK	1698218	1698218
Asset Value (Most Recent)	\$2,227,118,451	\$2,215,443,852
Asset Value (Start)	\$238,244,000	\$211,002,396
# Most recent Holdings	702	700
Periods	25	25
First PeriodOfReport	2017-06-30	2017-08-04
% Holding Value invested in ETFs	55.31%	55.31%
Factors		
Beta	0.5805	0.8845
Alpha	30.71%	30.69%
Annualized Arithmetic Mean Return	38.92%	41.37%
Annualized Geometric Mean Return	44.98%	48.25%
Annualized Median Return	53.53%	49.58%
Annualized CAGR	42.99%	45.68%
Annualized Sharpe Ratio (RFR=2%)	2.0484	1.9070
Annualized Sortino Ratio (MAR=2%)	5.3719	6.0044
Annualized Volatility	18.02%	20.65%
Max DD (Qtr)	−8.98%	−7.76%
Annualized Excess Return	26.22%	29.55%
Tracking Error	8.46%	8.33%
Information Ratio	3.0993	3.5474
1 Yr Return	18.26%	19.21%
3 Yr Return	150.13%	155.17%
5 Yr Return	463.99%	584.55%
Total Return	834.81%	949.96%
Skewness (Qtr Returns)	−0.2730	0.2110
Kurtosis (Qtr Returns)	−0.0420	−0.2410

Figure 8: Comparative Performance Metrics of Original and Cloned Portfolio.

The table details metrics of the original fund, Ritholtz Wealth Management, and its cloned counterpart, illustrating the effectiveness of the cloning strategy. General information, such as recent and initial asset values, number of holdings, reporting periods, and percentage of holdings in ETFs, provides a foundation for comparison. Performance and risk factors, such as Sharpe (Risk Free Rate: 2%) and Sortino (Minimum Acceptable Return: 2%) ratios, are detailed. Notably, the cloned fund exhibits a marginally higher Alpha, annualized geometric mean return, and total return, demonstrating its capability to closely emulate and even slightly enhance the performance of the original portfolio. Risk metrics like annualized volatility and maximum quarterly drawdown (Max DD) are also compared, showing a modest increase in risk for the cloned fund. Additionally, the 1, 3, and 5-year returns highlight the cloned portfolio’s outperformance over time.

Alpha factor as an illustrative case, Figure 9 graphically contrasts the alpha distribution of all original funds with that of the cloned funds.

To address the null hypothesis—that portfolios replicated from 13F filings between 2013 and 2023, and rebalanced according to the disclosure dates, do not achieve the performance of the source funds or outperform the market—we employed the Wilcoxon Signed-Rank Test on the delta distributions. This test evaluates each factor’s delta distribution to ascertain statistically significant discrepancies from 0. The results, including the descriptive statistics detailing the differences for each factor across all funds, are systematically presented in Figure 12, providing a robust statistical foundation to reject the null hypothesis. Figure 10

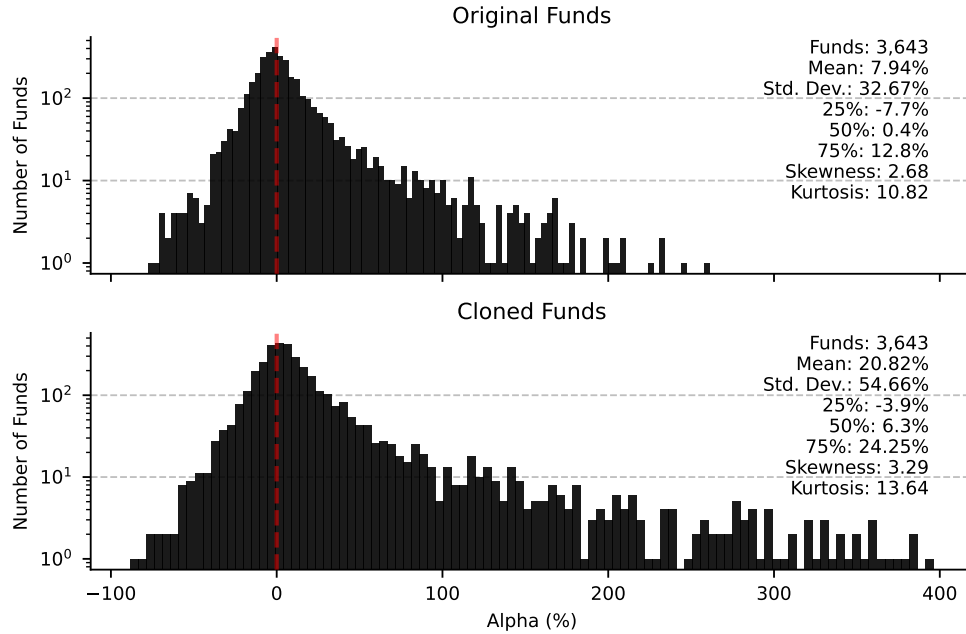


Figure 9: Alpha Distribution Comparison between Original and Cloned Funds.

The top graph illustrates the alpha distribution for 3,643 original funds from 2013 to 2023, with a mean of 7.94% and notable right skewness, as reflected by a skewness value of 2.68. The distribution exhibits heavy tails with a kurtosis of 10.82, indicating a propensity for extreme outliers. Quartile values highlight a wide dispersion of alpha, with the median fund slightly outperforming the SP500 benchmark by 0.4%. The bottom graph depicts the alpha distribution for the cloned funds, mirroring the same number of funds. The mean alpha significantly increases to 20.82%, and the distribution becomes more positively skewed (skewness of 3.29) and exhibits even heavier tails (kurtosis of 13.64) than the original funds. The cloned funds demonstrate a stronger median performance with an alpha of 6.3%, and a more substantial upper quartile performance, with 75% of funds attaining an alpha above -3.9%. Both distributions show that the cloned funds not only replicate the performance of the original funds but also amplify their positive traits, achieving higher mean returns and exhibiting a tendency towards higher outperformance, as indicated by their increased alpha values.

allows for a direct comparison between the factors of source and cloned funds with Figure 11 visualizing the correlation between each factor in a scatterplot.

Factor	Samples	Mean		Std. Dev.		Min		25%		50%		75%		Max	
		Original	Cloned	Original	Cloned	Original	Cloned	Original	Cloned	Original	Cloned	Original	Cloned	Original	Cloned
Reporting Periods	3,643	29.0	29.0	10.21	10.21	6.0	6.0	19.0	19.0	31.0	31.0	39.0	39.0	40.0	40.0
Recent Holdings	3,643	247.16	247.16	448.24	448.24	6.0	6.0	43.0	43.0	101.0	101.0	225.0	225.0	4,428.0	4,428.0
Recent Asset Value (\$M)	3,643	7,700.17	1,866.43	77,463.51	5,227.44	2.05	1.55	223.84	183.62	511.92	424.38	1,667.16	1,268.68	4,070,960.99	147,491.05
Alpha (%)	3,643	7.94	20.82	32.67	54.66	-77.4	-88.0	-7.7	-3.9	0.4	6.3	12.8	24.25	261.1	396.0
Beta	3,643	1.14	1.42	1.18	2.2	-6.88	-18.49	0.8	0.81	1.03	1.13	1.33	1.68	17.94	33.58
Ann. Arithmetic Mean Return (%)	3,643	23.1	38.02	33.82	60.64	-74.9	-72.3	6.1	10.3	14.5	20.4	28.0	40.2	262.6	710.8
Ann. Geometric Mean Return (%)	3,643	30.67	67.44	56.66	211.92	-56.3	-54.9	6.2	10.7	15.3	22.1	31.05	46.7	652.8	5848.0
Ann. Median Return (%)	3,643	14.73	13.96	18.4	20.55	-53.8	-67.8	4.9	3.3	13.5	12.9	23.0	23.2	214.3	237.1
CAGR (%)	3,643	9.99	13.08	19.01	23.12	-64.6	-62.8	0.75	1.2	8.8	11.6	17.8	21.9	121.2	220.2
Ann. Sharpe (RFR=2%)	3,643	0.4	0.48	0.49	0.46	-2.93	-2.29	0.17	0.26	0.4	0.46	0.66	0.7	4.13	3.5
Ann. Sortino (MAR=2%)	3,635	1.4	2.2	6.95	11.87	-2.87	-3.07	0.33	0.55	0.84	1.16	1.63	2.24	364.39	628.88
Ann. Volatility (%)	3,643	51.93	79.56	67.09	125.31	6.2	7.5	19.7	22.5	28.8	38.9	50.0	75.95	446.3	1324.0
Max Drawdown - Qtr Returns (%)	3,643	-35.53	-39.82	21.01	24.33	-91.44	-98.25	-46.96	-56.82	-28.81	-34.52	-20.1	-19.26	-0.04	0.67
Ann. Excess Return (%)	3,643	9.69	25.48	33.49	60.46	-88.2	-84.9	-6.7	-1.8	1.2	8.0	14.4	27.6	237.2	697.6
Tracking Error	3,643	23.33	38.07	34.23	63.08	1.4	0.8	6.3	8.9	11.9	17.8	22.9	36.85	224.9	661.7
Information Ratio	3,643	-0.08	0.38	1.4	1.11	-6.99	-6.46	-0.86	-0.18	0.1	0.52	0.77	0.97	10.19	6.44

Figure 10: Descriptive Statistics of Performance and Risk Metrics for Original and Cloned Funds.

This table provides a side-by-side statistical summary of various performance and risk factors for 3,643 original and cloned funds for the period 2013 to 2023. It contrasts key metrics such as Alpha, Beta, annualized mean returns, Compound Annual Growth Rate (CAGR), Sharpe Ratio with a Risk Free Rate (RFR) of 2%, Sortino Ratio with a Minimum Acceptable Return (MAR) of 2%, volatility, maximum drawdown of quarterly returns, excess return, tracking error, and information ratio. The statistics underscore the cloned funds' enhanced mean Alpha and returns across different measures, a slightly elevated Beta, and increased volatility compared to original funds. Notable too is the cloned funds' higher maximum drawdown and excess return, with a tracking error and information ratio that suggest differentiated performance relative to the original funds.

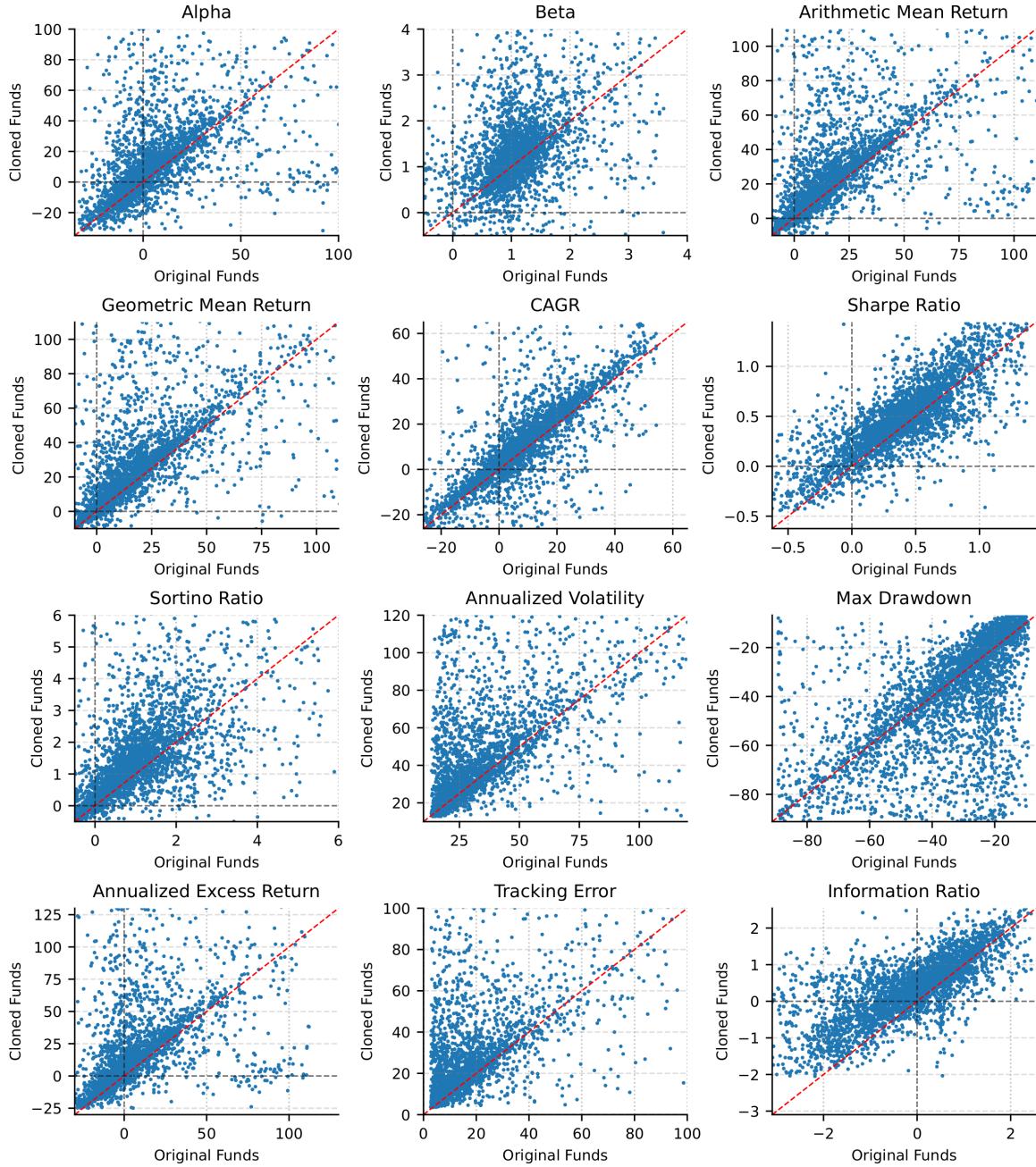


Figure 11: Correlation of Performance Metrics Between Original and Cloned Funds Across Various Factors.

Twelve scatterplots display the relationship between 3,643 original and cloned funds across a series of performance metrics, such as Alpha, Beta, and various types of returns, ratios, volatility, drawdown, and tracking error. Each plot on the x-axis represents a factor's value for the original funds, while the y-axis corresponds to the values for the cloned funds. The nearly linear patterns, highlighted by the red line of equality, illustrate a strong positive correlation between the original and cloned funds across most metrics. Notably, factors like Alpha, arithmetic mean return, Compound Annual Growth Rate and information ratio showcase a tight clustering around the line of equality, indicating that cloned funds often mirror the performance of their originals closely. The plots for Sharpe Ratio and Sortino Ratio also display close alignment, suggesting a consistent risk-adjusted return profile. Meanwhile, the graphs for annualized volatility and maximum drawdown reflect a greater variance, pointing to potential differences in risk behaviors. Overall, these visual comparisons affirm the cloning approach's ability to replicate fund performance effectively, with variations in certain risk-related factors.

Factor	<i>t</i> -stat	Samples	Mean	Std. Dev.	Min	25%	50%	75%	Max
Alpha	1,317,518.5***	3,643	9.40	30.57	−123.80	0.00	3.70	10.40	202.00
Beta	1,996,741.0***	3,643	0.20	0.96	−5.02	−0.15	0.09	0.46	6.56
Arithmetic Mean Return	1,054,798.0***	3,643	10.54	32.40	−111.70	0.90	3.60	10.30	213.80
Geometric Mean Return	1,063,496.5***	3,643	16.44	57.72	−179.10	1.00	4.00	12.10	458.00
CAGR	1,392,466.5***	3,643	2.65	10.33	−53.60	0.30	2.90	5.10	71.40
Sharpe Ratio	1,519,189.0***	3,643	0.08	0.24	−1.46	−0.02	0.08	0.20	1.49
Sortino Ratio	1,270,866.5***	3,643	0.49	1.28	−4.83	0.00	0.29	0.77	8.87
Annualized Volatility	1,492,135.0***	3,643	20.31	66.70	−279.20	−1.90	3.00	21.65	422.90
Max Drawdown	2,358,882.0***	3,643	−4.40	19.00	−74.86	−10.28	0.27	4.99	81.23
Annualized Excess Return	914,614.5***	3,643	11.45	32.55	−110.60	1.90	4.50	11.10	215.80
Tracking Error	1,056,415.0***	3,643	10.96	33.25	−142.60	0.00	2.10	11.60	216.80
Information Ratio	829,797.0***	3,643	0.46	0.74	−2.58	0.07	0.34	0.77	4.08

Figure 12: Statistical Analysis of Performance and Risk Metrics Delta Distributions (Cloned vs. Original Funds).

*This table delineates the delta distributions for various performance and risk metrics, where each metric for cloned funds is subtracted from that of the original funds. It presents the *t*-statistics from the Wilcoxon signed-rank test alongside highly significant *p*-values, indicating the strength of the evidence against the null hypothesis. The metrics where higher values indicate better performance (e.g., Alpha, CAGR, Sharpe Ratio) show positive median values, suggesting that cloned funds generally match or outperform the original funds. Conversely, for metrics where lower values are preferred (e.g., Volatility, Max Drawdown), the median values are less than or equal to zero, aligning with the desired outcomes. The results overall support the rejection of the null hypothesis, affirming that cloned portfolios – rebalanced on the disclosure date and modeled after institutionally successful strategies – do replicate or exceed the historical market performance of the original funds. Significance levels are denoted as follows: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$*

In analyzing the delta distribution of alpha, we observe that 50% of cloned portfolios have an alpha exceeding their source portfolios by at least 3.7%. Specifically, the top 25% of cloned funds, corresponding to the fourth quartile, demonstrate a remarkable alpha difference of 10.04% compared to their source portfolios, signifying that these clones achieve superior risk-adjusted returns.

The beta distribution shows a consistent median difference of 0.09, indicating stability across the cloned portfolios in terms of market sensitivity.

Further supporting the alpha findings, both the arithmetic and geometric mean returns in the third and fourth quartiles of cloned portfolios surpass those of the source portfolios. The arithmetic mean return at the 75th percentile is 10.3% higher, while the geometric mean return at the same percentile reaches 12.1% higher, indicating significant annualized return improvements in the top quartile clones. These observations are complemented by higher CAGR values, Sharpe, and Sortino ratios within the top quartile. Specifically, CAGR exceeds 5.1%, the Sharpe ratio surpasses 0.2, and the Sortino ratio exceeds 0.77, underscoring the clones' enhanced performance metrics.

However, it's noteworthy that cloned portfolios exhibit higher annualized volatility, with a median increase of 3%, and a greater maximum quarterly drawdown, with a median increase of 0.27%, suggesting increased risk in cloned portfolio strategies. Moreover, cloned portfolios demonstrate larger tracking errors with the median value of 2.1 indicating negative deviation from their source portfolios.

The outcomes of the Wilcoxon signed-rank test, indicated by the t-values and corresponding p -values in Figure 12, establish that all evaluated delta distributions significantly diverge from zero. This suggests that the performance of cloned funds is markedly different from that of the source funds. Specifically, nine out of the twelve examined performance and risk factors show that the cloned funds not only match the performance of the source funds but in several cases exceed it. This is evidenced by the positive delta values across the median and top quartile for these factors. However, for three factors—annualized volatility, maximum drawdown, and tracking error—the cloned funds present lower performance metrics.

Given these results, we confidently reject the null hypothesis. Instead, we accept the alternative hypothesis, affirming the premise that replicating the buy and sell activities of institutional investment managers based on their 13F filings on the date of disclosure can effectively mirror and surpass the performance of the original funds, thereby outperforming the S&P 500 benchmark.

4 Conclusion

We cloned 161,408 hedge fund portfolios from the period of 2013 to 2023, revealing that replicating the strategies of institutional investment managers from SEC 13F filings, rebalanced on the filing disclosure date, aligns with or exceeds the performance of the original funds. Performance and risk characteristics of cloned and original funds were evaluated across twelve dimensions. Our results confirm that cloned portfolios tend to replicate the superior performance of top-performing original funds, suggesting the efficacy of a straightforward cloning strategy in surpassing the market benchmark, particularly in terms of risk-adjusted returns. The upper quartile of original funds shows a minimum alpha of 12.8%, while the upper quartile of cloned counterparts exhibits a minimum alpha of 24.3%. This finding signifies that the top 25% of cloned funds do not merely track the performance of their source funds but also surpass the S&P 500 benchmark by a margin exceeding 24%, indicating substantial outperformance. However, despite the overall strong performance, the cloned funds exhibit increased annualized volatility and maximum drawdowns, indicating higher risk levels compared to their source funds. Additionally, the higher tracking error suggests a discrepancy in the performance consistency of cloned funds against the market benchmark. Our results on fund survival rates show a median operational lifespan of 5 years, with 3.4% of funds exceeding a 26-year lifespan, providing context for the durability of investment strategies over time. Furthermore, 75% of all funds manage portfolios smaller than \$2 billion, and the portfolio value distribution is heavily skewed by a small number of funds, with less than 2.5% of funds managing upwards of \$45 billion.

Given that cloned funds matched or outperformed their source funds in nine out of twelve factors, our study rejects the null hypothesis and demonstrates that replicating the holdings of institutional investment managers can indeed result in performance that is comparable to or even better than the original funds, thereby providing a potential avenue to outperform the S&P 500 benchmark.

References

- Ackermann, C., McEnally, R., & Ravenscraft, D. (1999). The performance of hedge funds: risk, return, and incentives. *The Journal of Finance*, 54(3), 833–874. doi: 10.1111/0022-1082.00129
- Agarwal, V., Boyson, N. M., & Naik, N. Y. (2009). Hedge funds for retail investors? An examination of hedged mutual funds. *Journal of Financial and Quantitative Analysis*, 44(2), 273–305. doi: 10.1017/s0022109009090188
- Agarwal, V., Jiang, W., Tang, Y., & Yang, B. (2013). Uncovering Hedge Fund Skill from the Portfolio Holdings They Hide. *The Journal of Finance*, 68(2), 739–783. doi: 10.1111/jofi.12012
- Agarwal, V., & Naik, N. Y. (2003). Risks and portfolio decisions involving hedge funds. *Review of Financial Studies*, 17(1), 63–98. doi: 10.1093/rfs/hhg044
- Amin, G. S., & Kat, H. M. (2003). Hedge fund performance 1990–2000: Do the “Money machines” really add value? *Journal of Financial and Quantitative Analysis*, 38(2), 251. doi: 10.2307/4126750
- Amir-Ghassemi, F., Papanicolaou, A., & Perlow, M. (2022, February). Aggregate alpha in the hedge fund industry: A further look at best ideas. *The Journal of Portfolio Management*, 48(3), 220–239. doi: 10.3905/jpm.2021.1.313
- Anderson, A., & Brockman, P. (2018). An Examination of 13F Filings. *Journal of Financial Research*, 41(3), 295–324. doi: 10.1111/jfir.12150
- Antón, M., Cohen, R. B., & Polk, C. (2010). Best Ideas. *Social Science Research Network*. doi: 10.2139/ssrn.1364827
- Aragon, G. O., Hertzel, M. G., & Shi, Z. (2013). Why Do Hedge Funds Avoid Disclosure? Evidence from Confidential 13F Filings. *Journal of Financial and Quantitative Analysis*, 48(5), 1499–1518. doi: 10.1017/s0022109013000483
- Blume, M. E. (1975). Betas and their regression tendencies. *The Journal of Finance*, 30(3), 785. doi: 10.2307/2326858
- Bollen, N. P. B., & Fisher, G. S. (2013). Send in the clones? Hedge fund replication using futures contracts. *The Journal of Alternative Investments*, 16(2), 80–95. doi: 10.3905/jai.2013.16.2.080
- Brinson, G. P., Hood, L. R., & Beebower, G. L. (1986). Determinants of portfolio performance. *Financial Analysts Journal*, 42(4), 39–44. doi: 10.2469/faj.v42.n4.39
- Brown, G. W., Howard, P., & Lundblad, C. (2021). Crowded trades and tail risk. *The Review of Financial Studies*, 35(7), 3231–3271. doi: 10.1093/rfs/hhab107
- Cao, S., Da, Z., Jiang, D., & Yang, B. (2021). The strategic use of 13F restatement by hedge funds. *Social Science Research Network*. doi: 10.2139/ssrn.3907560

- Cao, S. S., Du, K., Yang, B., & Zhang, A. L. (2021). Copycat skills and disclosure costs: Evidence from peer companies' digital footprints. *Journal of Accounting Research*, 59(4), 1261–1302.
- Chaudhry, A., & Johnson, H. L. (2008). The efficacy of the Sortino ratio and other benchmarked performance measures under skewed return distributions. *Australian Journal of Management*, 32(3), 485–502. doi: 10.1177/031289620803200306
- DiPietro, D. M. (2019). Alpha Cloning: using quantitative techniques and SEC 13F data for equity portfolio optimization and generation. *The Journal of Financial Data Science*, 1(4), 159–171. doi: 10.3905/jfds.2019.1.008
- Fama, E. F., & French, K. R. (2004). The capital asset pricing model: Theory and evidence. *The Journal of Economic Perspectives*, 18(3), 25–46. doi: 10.1257/0895330042162430
- Fleiss, A., Cui, H., Stoikov, S., & DiPietro, D. M. (2019). Constructing Equity Portfolios from SEC 13F Data Using Feature Extraction and Machine Learning. *The Journal of Financial Data Science*, 2(1), 45–60. doi: 10.3905/jfds.2019.1.022
- Fleiss, A., Kumaar, A., Rida, A., Shin, J., Lai, X., Fang, V., ... Li, A. (2021, November 8). *Deep reinforcement learning & feature extraction for constructing alpha generating equity portfolios*. (Available at SSRN: <https://ssrn.com/abstract=3958478> or <http://dx.doi.org/10.2139/ssrn.3958478>)
- Frank, M. M., Poterba, J. M., Shackelford, D. A., & Shoven, J. B. (2004). Copycat Funds: Information Disclosure Regulation and the Returns to Active Management in the Mutual Fund Industry. *The Journal of Law and Economics*, 47(2), 515–541. doi: 10.1086/422982
- Fung, W., & Hsieh, D. A. (2000). Performance characteristics of hedge funds and commodity funds: Natural vs. spurious biases. *Journal of Financial and Quantitative Analysis*, 35(3), 291. doi: 10.2307/2676205
- Fung, W., & Hsieh, D. A. (2001). The risk in hedge fund strategies: Theory and evidence from trend followers. *Review of Financial Studies*, 14(2), 313–341. doi: 10.1093/rfs/14.2.313
- Gabaix, X., Gopikrishnan, P., Plerou, V., & Stanley, H. E. (2003). A theory of power-law distributions in financial market fluctuations. *Nature*, 423(6937), 267–270. doi: 10.1038/nature01624
- Griffin, J. M., & Xu, J. (2009). How smart are the smart guys? A unique view from hedge fund stock holdings. *Review of Financial Studies*, 22(7), 2531–2570. doi: 10.1093/rfs/hhp026
- Groggel, D. J., & Conover, W. (2000). Practical nonparametric statistics. *Technometrics*, 42(3), 317. doi: 10.2307/1271101
- Hasanhodzic, J., & Lo, A. W. (2007). Can hedge funds be replicated? The linear case. *Journal of Investment Management*, 5(2), 5–45.

- Hayes, B. T. (2012). On the Market-Timing Ability of Factor-Based Hedge Fund Clones. *The Journal of Alternative Investments*, 15(1), 8–42. doi: 10.3905/jai.2012.15.1.008
- Hayes, B. T., & Ba, Y. A. (2014). Beta Regime-Switching Hedge Funds and their Clones. *The Journal of Alternative Investments*, 17(3), 87–110. doi: 10.3905/jai.2014.17.3.087
- Hubbard, J. W. (2008). Can Hedge-Fund Returns Be Replicated?: The Linear Case. *The C.F.A. Digest*, 38(1), 6–7. doi: 10.2469/dig.v38.n1.3
- Jensen, M. C. (1968). The performance of mutual funds in the period 1945-1964. *The Journal of Finance*, 23(2), 389–416. doi: 10.1111/j.1540-6261.1968.tb00815.x
- Kacperczyk, M., Sialm, C., & Zheng, L. (2006). Unobserved actions of mutual funds. *Review of Financial Studies*, 21(6), 2379–2416. doi: 10.1093/rfs/hhl041
- Kaizoji, T., & Miyano, M. (2016). Why does the power law for stock price hold? *Chaos, Solitons & Fractals*, 88, 19–23. doi: 10.1016/j.chaos.2016.03.015
- Lazo-Paz, R., Moneta, F., & Chincarini, L. B. (2023). Crowded spaces and anomalies. *Social Science Research Network*. doi: 10.2139/ssrn.4618248
- Lewellen, J. (2011). Institutional investors and the limits of arbitrage. *Journal of Financial Economics*, 102(1), 62–80. doi: 10.1016/j.jfineco.2011.05.012
- Lo, A. W. (2002). The statistics of Sharpe ratios. *Financial Analysts Journal*, 58(4), 36–52. doi: 10.2469/faj.v58.n4.2453
- Malkiel, B. G. (1995). Returns from investing in equity mutual funds 1971 to 1991. *The Journal of Finance*, 50(2), 549–572. doi: 10.1111/j.1540-6261.1995.tb04795.x
- Markowitz, H. M., & Blay, K. A. (2013). *Risk-Return Analysis: The Theory and Practice of Rational Investing (Volume One)*. McGraw-Hill Professional.
- Martin, G. S., & Puthenpurackal, J. (2008). Imitation Is the Sincerest Form of Flattery: Warren Buffett and Berkshire Hathaway. *Social Science Research Network*. doi: 10.2139/ssrn.806246
- Mitchell, M., & Pulvino, T. (2001). Characteristics of risk and return in risk arbitrage. *The Journal of Finance*, 56(6), 2135–2175. doi: 10.1111/0022-1082.00401
- O’Shaughnessy, J. P. (2005). *What Works on Wall Street*. McGraw Hill LLC.
- Paleologo, G. A. (2021). *Advanced Portfolio Management: A Quant’s Guide for Fundamental Investors*. John Wiley & Sons.
- Pedersen, C. S., & Satchell, S. E. (2002). On the foundation of performance measures under asymmetric returns. *Quantitative Finance*, 2(3), 217–223. doi: 10.1088/1469-7688/2/3/304

- Sandvik, S. H., Frydenberg, S., Westgaard, S., & Heitmann, R. K. (2011). Hedge fund performance in bull and bear markets: Alpha creation and risk exposure. *The Journal of Investing*, 20(1), 52–77. doi: 10.3905/joi.2011.20.1.052
- Schroeder, J., & Posch, P. N. (2023). *Anomalies, Trends and Patterns in Disclosure Activities: Understanding EDGAR*. Social Science Research Network. doi: 10.2139/ssrn.4664431
- Sharpe, W. F. (1994). The Sharpe ratio. *Journal of Portfolio Management*, 21(1), 49–58. doi: 10.3905/jpm.1994.409501
- Sias, R. W., Turtle, H. J., & Zykaj, B. B. (2016). Hedge fund crowds and mispricing. *Management Science*, 62(3), 764–784. doi: 10.1287/mnsc.2014.2131
- Sortino, F. A., & Price, L. N. (1994). Performance measurement in a downside risk framework. *The Journal of Investing*, 3(3), 59–64. doi: 10.3905/joi.3.3.59
- Verbeek, M., & Wang, Y. (2013). Better than the Original? The Relative Success of Copycat Funds. *Journal of Banking and Finance*, 37(9), 3454–3471. doi: 10.1016/j.jbankfin.2013.04.024
- Wallerstein, E., Tuchschnid, N. S., & Zaker, S. (2009). How do hedge fund clones manage the real world? *The Journal of Alternative Investments*, 12(3), 37–50. doi: 10.3905/jai.2010.12.3.037
- Weber, A. (2017). *Annual risk measures and related statistics*. Social Science Research Network. doi: 10.2139/ssrn.3054517
- Weber, V., & Peres, F. (2013). Hedge fund replication: putting the pieces together. *The Journal of Investment Strategies*, 3(1), 61–119. doi: 10.21314/jois.2013.024