

# Did You Really Beat the Market? A Practical and Parsimonious Approach to Evaluating Risk-Adjusted Performance

David J. Moore

College of Business, California State University, Sacramento, CA, USA

Email: [djmpd@csus.edu](mailto:djmpd@csus.edu)

**How to cite this paper:** Moore, D.J. (2021) Did You Really Beat the Market? A Practical and Parsimonious Approach to Evaluating Risk-Adjusted Performance. *Journal of Mathematical Finance*, 11, 554-577.  
<https://doi.org/10.4236/jmf.2021.113031>

**Received:** August 2, 2021

**Accepted:** August 28, 2021

**Published:** August 31, 2021

Copyright © 2021 by author(s) and Scientific Research Publishing Inc.  
This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).  
<http://creativecommons.org/licenses/by/4.0/>



Open Access

## Abstract

This study contributes to the literature by modifying and recasting the Modigliani and Modigliani *M-Squared* risk-adjusted performance measure in a practical setting. Specifically, rather than combine the risk-free asset (Treasury Bill) with the portfolio under consideration to match *market* risk, this study combines the risk-free asset with a levered (or unlevered) market ETF to match *portfolio* risk. In so doing, this study addresses the question: Could an investor have earned higher returns with the same risk (standard deviation) using a simple combination of the risk-free asset and a readily available levered (or unlevered) market ETF? The study also addresses the impact of the context where one captures return measurements on outperformance conclusions. Although this study focuses its analysis on in-sample descriptive statistics, the new Risk-Equivalent Excess Performance measure and contextualization provide a basis for future out-of-sample inferential analysis.

## Keywords

Return, Risk, Risk-Adjusted, Benchmark, CAPM, Market Efficiency

## 1. Introduction

This study presents a simple framework to evaluate claims of superior investment performance. As a finance professor, I often hear claims of superior returns and therefore claims of superior skill from “investment managers” in the guise of students, faculty colleagues, family, friends, and strangers. I have heard such market beating claims regarding stocks, bonds, foreign exchange, real estate, cryptocurrency, and even baseball cards. The frequency of such claims increases during bull markets. This is consistent with Hoffman and Post [1] who find that individual investor belief in skill increases with recent returns and that their be-

lief is not impacted by market returns. In other words, bias clouds vision of the possibility that recent high returns are a result of high market returns in general and not skill.

Outperformance claims also appear in national media. For instance, the April 2021 *Wall Street Journal* article titled “The pandemic year’s top stock-fund managers” reports manager Dennis P. Lynch of the Morgan Stanley Inception fund (ticker: MSSGX) earned a 12-month net return of 273% as of March 31, 2021 [2].

Astonishingly, the word “risk” is not mentioned in the Wall Street Journal article at all. This would not surprise Mark Hebner, founder and president of Index Fund Advisors, who in a January 2016 *Money Management Executive* column [3] states: “But here’s the next number that I’ve never seen in the press: volatility or the deviation.”

We live in an era with increased FinTech (Financial Technology) adoption and, unfortunately, a proliferation of misinformation sources. As evidence of FinTech’s increased adoption, Robinhood’s July 1, 2021, S-1 filing reveals 18 million accounts as of March 31, 2021 [4]. Meanwhile, misinformation regarding returns and risk is on the rise fueled by posts from non-professionals (and professionals) on Facebook, Twitter, Reddit, etc. Together, increased FinTech adoption and misinformation proliferation lead to bubbles like the recent “meme stock” craze (see “\$26 Billion Gone! ‘Meme Stock’ Crash Erases Nearly Half of Gains” [5]).

Undoubtedly, many of Robinhood’s millions of account holders are new to financial markets. The majority of those lured into the “meme stock” craze likely lost their investment—or more as many used borrowed funds to purchase stocks. As evidence of borrowing to purchase “meme stocks”, see “Robinhood claims it ONLY forced the sale of GameStop shares if they were bought with borrowed funds” [6] that states: “At one point, an estimated half of Robinhood’s 13 million users owned some GameStop stock.” This study serves as a counterbalance to the misinformation that abounds in social media and the omission of important performance measurement context in mainstream media.

Shifting the focus to the academic literature, the sheer volume devoted to risk measurement indicates risk is critical for performance reporting. For instance, Cogneau and Hubner [7] perform a census of over 100 risk-adjusted performance measures in the academic literature. For now, the introductory quote in Chapter 9 of Altman [8], originally from Schoolman *et al.* [9] serves as a compass for the current study:

“Good answers come from good questions not from esoteric analysis.”

In the context of assessing market-beating performance claims, the [hopefully] good question I address in this study is:

*The Moore Performance Question:* Could an investor have earned higher returns with the same risk (standard deviation) using a simple combination of the risk-free asset and a readily available levered (or unlevered) market

ETF?

The *Moore Performance Question* takes into account two key factors that are often overlooked in outperformance claims: risk and market (benchmark) returns.

While risk is the primary focus of this study, I must note risk is just one of many critically important factors to consider when assessing performance. **Table 1** lists what I term the *nine ingredients of valid performance measurement*. I discuss these in greater detail in Sections 2.2 and 3.

In sum, the importance and contribution of this study are twofold. First, the nine ingredients of valid performance measurement serve as a reminder for practitioners (veteran and novice alike) and academics to be mindful of the full context of performance reporting. Second, the new Risk-Equivalent Excess Performance measure provides a straightforward measure of value-added performance inclusive of the nine ingredients.

The remainder of the paper is as follows. Section 2 presents a literature review reflective of the starting point of this study and the nine ingredients of valid performance measurement. Section 3 describes the approach this study takes in addressing performance measurement concerns, the sample construction, summary statistics, and construction of the *REEP* measure. Section 4 presents results and Section 5 concludes.

## 2. Literature Review

### 2.1. The Starting Point: Modigliani and Modigliani *RAP* Measure

This study begins with the Modigliani and Modigliani [10] *M-2* (also called Risk-Adjusted Performance or *RAP*) measure. Modigliani and Modigliani construct *RAP* by combining the risk-free asset (Treasury Bill) with the portfolio under consideration to match *market* risk. Higher *RAP* corresponds to a higher ranking. In contrast, the measure of this study, Risk-Equivalent Excess Performance or *REEP*, combines the risk-free asset with a levered (or unlevered) market ETF to match *portfolio* risk.

The study of Cogneau and Hubner [7] notes the *RAP* measure is a linear function of the Sharpe Ratio and therefore shares its disadvantages. Specifically, Cogneau and Hubner state the Sharpe Ratio (and by extension *RAP*) 1) does not quantify value-added in that it only ranks funds, 2) produces rankings affected by the choice of the risk-free rate, 3) is suitable for investors who invest in only one fund, 4) is subject to sampling error in the standard deviation calculation, and 5) presumes normality while most fund returns are not normally distributed. I refer to this set of observations as the *Cogneau-Hubner Critique*.

**Table 1.** The nine ingredients of valid performance measurement.

1. Gains and losses	4. Taxes	7. Sample size
2. Cash	5. Risk	8. Appropriate benchmark
3. Fees and costs	6. Time period	9. Luck vs. skill

The Risk-Equivalent Excess Performance measure of this study addresses the bulk of the *Cogneau-Hubner Critique*. First, *REEP* is a direct measure of value added over an alternative and easily constructed portfolio. Second, results are less sensitive to the choice of risk-free asset because 1) *REEP* is based on gross returns not excess returns and 2) actual borrowing costs are inherent in the construction of the measure.

Regarding one-fund investing, *REEP* shares the same disadvantage as *RAP* in that it considers funds in isolation. However, the primary focus of this study is not to rank funds, but rather to evaluate claims of outperformance. The same can be said in the use of standard deviation as a measure of risk. Regarding non-normal fund returns, this leaves room for future research using alternative non-parametric measures of risk. Regardless, as we shall see, the *REEP* measure constructed in Section 3 does indeed answer the *Moore Performance Question*.

## 2.2. Ingredients of Valid Performance Measurement

This section discusses performance measurement ingredients and their importance in the context of extant literature. It is important to bear in mind that although all ingredients are critical, not all are present in performance measurements obtained from acquaintances, the media, or even academic literature. However, by the end of this section readers will be more aware (or reminded) of important pieces of information necessary for valid performance measurement claims.

**Ingredient 1: Gains and losses.** The first step to valid performance measurement is procurement of accurate returns that include both gains and losses. I emphasize *both gains and losses* because, as Thaler [11] shows in his seminal work on consumer choice, humans suffer from “mental accounting.” In the context of reporting returns from their investments, it is much like reporting “winings” from a casino—many often neglect money they lost and speak only of their winning bets.

**Ingredient 2: Cash.** Related to the “mental accounting” phenomenon pointed out by Thaler, some investment managers report returns “net of cash.” In a May 2019 *Bloomberg* article [12], Warren Buffett criticized the practice of reporting net of cash returns stating “It makes their return look better if you sit there a long time in Treasury Bills ... It’s not as good as it looks.” As a simple illustration, consider hypothetical performance reported by two different investment managers. Presume fund manager A held 50% of assets under management in cash and the other 50% in equities that earned 12% over a year. Now presume manager B held 5% of their assets in cash and the other 95% in equities that earned 8% over the same year. **Table 2** presents the illusion of “net of cash” return reporting.

**Table 2.** Net of cash reporting illustration.

Manager	Net of cash return	Inclusive of cash (true) return
A	12%	$0.50 \times 0\% + 0.50 \times 12\% = 6.0\%$
B	8%	$0.05 \times 0\% + 0.95 \times 8\% = 7.6\%$

Clearly, Manager B earned more money for their clients (7.6%) than Manager A (6.0%) even though Manager A may claim higher net of cash returns (12% vs. 8%). Buffett further states “Firms will include money that’s sitting in Treasury Bills waiting to be deployed when charging management fees, but will exclude it when calculating a so-called internal rate of return, the performance measure in which most funds are judged [12].” Thus, returns inclusive of cash positions (and the associated zero return of that portion of the portfolio) is requisite for valid performance measurement.

**Ingredient 3: Fees and costs.** The same 2019 *Bloomberg* article states “Buffett has a consistent history of blasting asset managers for charging high management fees and collecting performance fees on gains that sometimes don’t beat broader markets.” The sentiment is echoed by Christopher Ailman, Chief Investment Officer of CalSTRS, the second largest pension fund in the United States. In a June 2021 interview with *CNBC* [13] Ailman states: “Our active managers in US equities added value, and they did produce alpha, but not after fees.”

“Mental accounting” manifests with regards to fees and costs in the form of placing costs associated with investing in a separate ledger. For instance, a person claiming to earn superior returns with rental real estate property may exclude the following costs from quoted returns: closing costs, maintenance, property taxes, income taxes, vacancies, non-payment of rent, tenant procurement costs, eviction costs, insurance, and in the extreme-litigation costs.

Similarly, a cryptocurrency “investor” may exclude trading commissions, fees to the cryptocurrency exchange to convert back to US dollars, opportunity cost of time spent away from full-time job, memberships or subscriptions to data feeds or trading education seminars, and of course taxes. Similar costs occur for stock, bond, and FX traders as well—and may be neglected in the media and conversation with investors claiming they beat the market.

**Ingredient 4: Taxes.** Active strategies involve more transactions than passive indexing. In particular, active strategies may execute more short-term trades and thereby expose investors to higher short-term gains taxes [14]. Thus, the presence of higher taxation for more frequent trading associated with active management must be considered when comparing results to passively managed instruments with infrequent trading.

As an illustration, presume short-term gains are taxed at 35% and long-term gains at 15%. Presume returns from the active strategy are  $R_{\text{active}}$  and returns from the passive strategy are  $R_{\text{passive}}$ . In order to have comparable after-tax returns, the following inequality shows the active manager’s returns need to exceed the passive manager’s by over 30%.

$$R_{\text{active}}(1-0.35) > R_{\text{passive}}(1-0.15)$$

$$R_{\text{active}} > 1.3077R_{\text{passive}}$$

**Ingredient 5: Risk.** Risk measurement receives extensive attention in the aca-

demographic literature with over 100 different measures [7]. Authors Modigliani and Modigliani [10] address risk with a straightforward question: “do returns adequately compensate us for the risk what we bear?” In this study I address risk via a similar yet distinct question: Could an investor have earned higher returns with the same risk (standard deviation) using a simple combination of the risk-free asset and a readily available levered (or unlevered) market ETF? Section 3 describes the approach this study uses to address these questions.

**Ingredient 6: Time period.** “Past performance is no guarantee of future results” is a common refrain in fund prospecti and advertisements. Two realities contribute to this prophecy: 1) time-varying investing environment and 2) time-varying management. Regarding the time-varying investing environment, returns relative to a benchmark in the 1960-1970 time period may have little to no relation to the 2020-2030 time period. Market, economic, political, geopolitical, and why not, climate change impacts vary from one decade to the next. As such, one must be mindful of projecting returns from a period that may be different on many important dimensions than some other or future period.

Regarding management changes, mutual funds have management changes over time. Management turnover is inherently higher in Student Investment Funds (SIFs) where managers are technically students under the supervision of faculty. As such, SIF management teams vary from semester to semester. As new management inherits the assets of the old, should they count returns of inherited assets as if they chose them? Managers subject to self-attribution bias [1] would tend to include returns from inherited assets and attribute to themselves if positive while exclude returns from inherited assets and attribute to prior managers if negative.

**Ingredient 7: Sample size.** Naturally, larger sample sizes are better. In the absence of large samples researchers often turn to simulation (see examples in Ingredient 9 below). In order to make out-of-sample inferences, particularly for application of the Central Limit Theorem, sample sizes must be sufficiently large. However, in conjunction with Ingredient 6 (Time Period), the sample must also be representative of the population. For instance, a 20 year sample from 1930 to 1950 may not be representative of the population of returns from 1980-2020 (or 2020-2040).

Author Muralidhar [15] modifies several risk-adjusted performance measures to account for varying histories. Unfortunately, these modifications still do not completely address the representativeness requirement of the Central Limit Theorem. This shortcoming is amplified in the context of small samples.

**Ingredient 8: Appropriate benchmark.** Should the returns of a small cap growth fund be compared to the S&P 500, a large cap blend index? Obviously not. In the Wall Street Journal Article mentioned in the introduction [2], 3 of the “pandemic top 5” performing funds are small cap growth funds while the other two are broad cap growth. Unfortunately, the returns for all 5 funds in that article are compared to that of the S&P 500 index in the absence of risk measures.

While selection of an appropriate benchmark is pervasive in academic literature and finance texts, it sometimes eludes our friends outside the ivory towers and the media (see Hebner's critique [3] of media omitting "risk" when presenting performance numbers).

In a strict sense, the market portfolio is a market capitalization weighted portfolio of all risky assets around the world. Unfortunately such a portfolio is unobservable [16]. However, Doeswijk *et al.* [17] construct an index of market portfolio returns through extensive data collection. The authors note that the tests of Stambaugh [18] found exclusion of assets such as bonds and residential real estate from the market portfolio had little impact on CAPM inferences. Yet, Doeswijk *et al.* [17] do note that "certain asset pricing applications" do necessitate a broader market portfolio representation than just the S&P 500.

**Ingredient 9: Luck vs. skill.** Presume all previous ingredients are accounted for in an outperformance claim. This final ingredient, luck vs. skill, is perhaps the most difficult to prove. And again, regarding Ingredient 6 (time period), can one guarantee their skill will persist in the future? Nevertheless, here I present four broad approaches to quantifying luck vs. skill in the extant literature:

**Ex-post analysis of volatility.** Treynor and Mazuy [19] examine the volatility of fund returns in declining market vs. rising market return time frames. They define a skillful fund manager as one who successfully anticipates market declines (increases) and shifts their portfolios to less (more) volatile securities. Treynor and Mazuy find no evidence that mutual fund managers outguessed the market.

**"The Fundamental Law of Active Management."** Grinold [20] introduces "The Fundamental Law of Active Management" which leads to a series of equations in Grinold and Kahn [21] that relate *ex-ante* information ratios to managerial skill and breadth of investments. However, Goodwin [22] notes that *ex-ante* information ratios and breadth measures are difficult to estimate making the Grinold-Kahn equations less operational.

**Bootstrap simulation.** Fama and French [23] perform bootstrap simulations on a cross section of 660 to 3156 mutual funds over the 273 calendar months from January 1984 to September 2006. The authors test if the distribution of simulated cross-sectional alpha has any observations in the tails. The authors find few funds produced benchmark-adjusted returns net of costs. Furthermore, I will add that the ability to identify the few benchmark beating funds in advance is elusive and undocumented.

**Generalized Binomial Distribution (GBD) simulation.** Bhootra *et al.* [24] employ GBD simulation to identify whether or not observed persistence of mutual funds in the top 25% of returns can occur via chance. Using a sample of 981 mutual funds over the 1995-2009 period, the authors find evidence that more funds achieve persistence in the top 25% than would be predicted by chance. While the results are promising in that they confirm the presence of skill in the mutual fund industry, the results are still subject to Ingredient 6: Time period. A



process that worked in the 1995-2009 time period has no guarantee of working in the 2021-2034 time period. Furthermore, Bhootra *et al.* document the presence of skill *ex-post* with no mechanism to identify persistent top 25% performers in advance.

Collectively, the extant literature covers the ingredients for valid performance measurement and developing risk-adjusted measures. This study extends that literature stream by summarizing the relevant factors of valid performance measurement and developing a parsimonious and practical measure that addresses the *Cogneau-Hubner Critique* of the Modigliani and Modigliani [10] *RAP* measure.

### 3. Data & Methodology

#### 3.1. Addressing Performance Measurement Concerns

The previous section detailed nine distinct ingredients or considerations for valid performance measurement. This section details the approaches used in this study to ensure validity of performance measurements herein.

1) *Gains and losses*. Returns obtained from Bloomberg L.P. [25] and the local pension fund include both gains and losses.

2) *Cash*. Returns obtained from Bloomberg and the local pension fund include cash holdings.

3) *Fees and costs*. Returns from mutual funds obtained from Bloomberg are based on net asset value (NAV) which is net of fees and costs. Return data for the local pension fund are in both gross and net terms as are the Student Investment Fund returns.

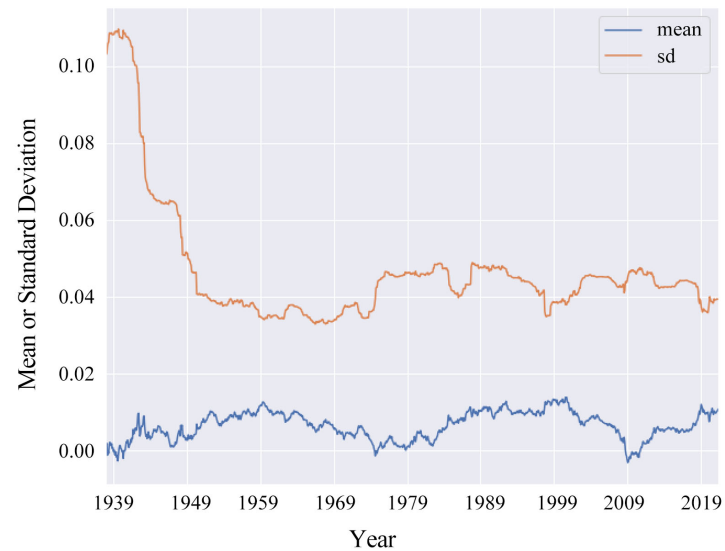
4) *Taxes*. To abstract from taxes, this study presumes assets are held in a non-taxable or tax-deferred account. This is the case for both the local pension fund and the Student Investment Fund and could be the case for the other settings (e.g., IRA, 401k, and 403b accounts).

5) *Risk*. This study follows Modigliani and Modigliani [10] and others using standard deviation as the risk measure. This study also modifies the risk-adjusted performance measure of Modigliani and Modigliani [10]. More on this in Section 3.4.

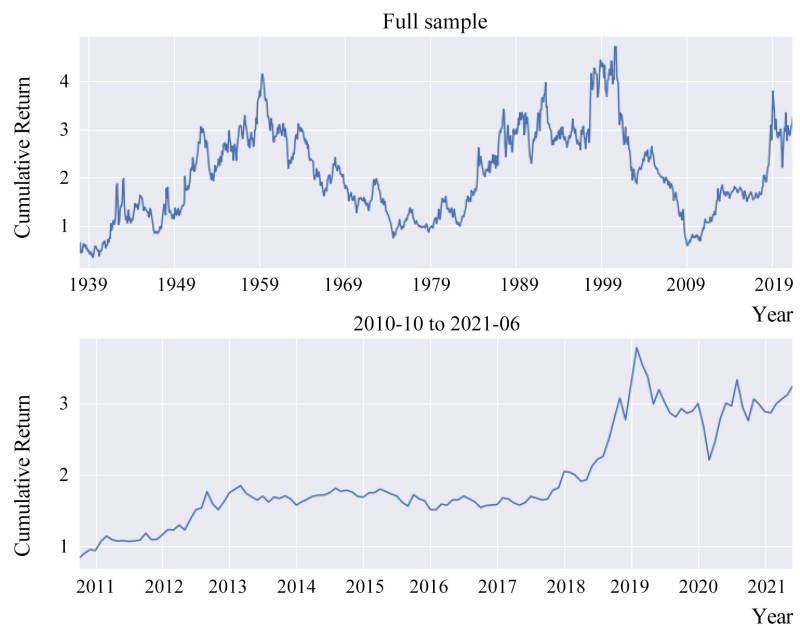
6) *Time period*. Section 2.2 suggests the time-varying investment environment could nullify out-of-sample inferences. To illustrate the time-varying investment environment, **Figure 1** and **Figure 2** present the rolling 10-year (120 month) mean and standard deviation and rolling 10-year (120 month) cumulative return for the S&P 500 index, respectively. Both figures illustrate significant volatility in average monthly returns, monthly standard deviation, and cumulative 10 year returns. The bottom panel of **Figure 2** highlights how we have been in a bull run for more than a decade while the top panel reveals bull runs historically precede bear markets.

However, this study focuses on evaluation of claims *during a specific time period* and thereby does not make any out-of-sample claims. In the process, this





**Figure 1.** Rolling 120 month mean and standard deviation for S&P 500 from 1937-12 to 2021-06.



**Figure 2.** Rolling 120 month cumulative return for S&P 500 from 1937-12 to 2021-06.

study raises awareness of the time-varying investment environment in US equity markets and that out-of-sample results could vary substantially from in-sample results.

Another time period consideration is the time-varying manager scenario. To illustrate performance measurement in the context of inherited holdings, I utilize the first and last trade dates for the two most recent managers (Z and M) of the Student Investment Fund used in this study. As such, I conduct analysis of SIF performance in four time periods shown in **Table 3**.

**Table 3.** Student Investment Fund management changes.

Era	Time frame
1.	Inception (2010.11) to Manager Z first trade (2019.11)
2.	Manager Z first trade (2019.11) to Manager M first trade (2020.12)
3.	Manager M first trade (2020.12) to present (2021.06)
4.	Full sample (2010.11 to 2021.06)

7) *Sample size*. From a statistical perspective in a financial return context, the population includes returns we have observed (e.g., [Figure 1](#) and [Figure 2](#)) and future returns we have yet to observe (future return graph not available). Having surrendered the focus to in-sample descriptive statistics, I alleviate the pressure to have a large dataset or one representative of the population. However, future research with a larger sample or simulation can contribute to the literature.

8) *Appropriate benchmark*. Allow me to quote a 2019 Consortium for Data Analysis in Risk participant and retired U.C. Berkeley Finance professor:

“Market efficiency is not an absolute truth. Market efficiency points to a benchmark to judge performance.”

The Modigliani and Modigliani [\[10\]](#) Risk-Adjusted Performance (*RAP* or *M-squared*) measure, and the Risk-Equivalent Excess Performance (*REEP*) measure of this study, share the sentiment of that quote. Market efficiency suggests an efficient frontier exists that is a linear combination of the risk-free asset (e.g., Treasury Bills) and a value-weighted portfolio of all traded securities in the market (a market portfolio often approximated by the S&P500). In both *RAP* and *REEP* measures, the “appropriate benchmark” is a portfolio on the efficient frontier.

For example, consider one of the small cap growth funds mentioned in “The pandemic year’s top stock-fund managers” [\[2\]](#). If markets are [reasonably] efficient, and the S&P 500 [reasonably] approximates the market portfolio, then Modigliani and Modigliani [\[10\]](#), this study, and numerous others, use an “appropriate benchmark” to judge performance.

9) *Luck vs. skill*. Luck vs. skill is both difficult to quantify [\[22\]](#) and project into the future (Ingredients 6, 7, and 9). While Fama and French [\[23\]](#) and Bhootra *et al.* [\[24\]](#) find some evidence that skill exists in mutual fund management, Bessembinder [\[26\]](#) points out that identifying managers with such skill reliably in advance is still unresolved. As such, I save luck vs. skill analysis for future research that may utilize the Risk-Equivalent Excess Performance measure developed herein.

### 3.2. Sample Construction

**Table 4** provides a brief description of the data used in this article. All data are obtained from Bloomberg L.P. [\[25\]](#) with the exceptions of Funds C and Cn, obtained from a local pension plan. I compute the expense ratio associated with

Fund Cn by subtracting the mean of returns net of fees (Fund Cn) from the mean of gross returns (Fund C). This amounts to 0.24%.

Fees are paid differently for the Student Investment Fund of this study than other funds in the sample. SIF fees are not withdrawn from the SIF account. Rather, fees are paid by the provider of the funds to a separate College of Business account. So, unlike Fund C where gross and net returns are known while the expense ratio is extrapolated, in Fund S the expense ratio and gross returns are known while net of fee returns are extrapolated. Therefore, I obtain the SIF returns net of fees (Fund Sn) by subtracting the monthly expense ratio ( $0.50\% \div 12$ ) from the monthly gross returns of Fund S.

One might question the inclusion of a single stock, AAPL, in the list of financial instruments in this study. Surprisingly, and against the advice of most (if not all) finance academics and practitioners, some hold an undiversified portfolio of just AAPL. For example, the founder of Chewy, Ryan Cohen, reportedly placed the bulk of the proceeds from Chewy's sale to Petsmart in just two stocks: AAPL and WFC (Wells Fargo). Thus, I included AAPL to determine if Cohen, and others who are "all in" on AAPL, would have been better off in US Treasuries and a levered ETF.

**Table 4.** Data series descriptions.

Series	Ticker	Name	MarketCap	Strategy	Expense Ratio	Inception
Fund W1	MSSGX	MSIF Inception Portfolio	Small	Growth	1.00	1989-11-01
Fund W2	JMIGX	Jacob Discovery Fund	Small	Growth	2.00	1997-12-31
Fund W3	BPTRX	Baron Parners Fund	Broad	Growth	1.32	1992-01-31
Fund W4	HDPMX	Hodges Fund	Broad	Blend	1.16	1992-10-09
Fund W5	DMCRX	Driehaus Micro Cap Growth Fund	Small	Growth	1.43	2013-11-18
Fund L	LVPSX	ProFunds Large-Cap Value	Large	Value	1.78	2002-10-01
Fund A	ARKK	ARK Innovation ETF	Broad	Blend	0.75	2014-10-31
Fund C	n/a	One fund within local pension plan	Broad	Blend	0.00	2010-01-01
Fund Cn	n/a	Fund C net of fees	Broad	Blend	0.24	2010-01-01
Fund D	n/a	Student D 1 Year Upside Picks	Broad	Growth	n/a	2020-03-13
Fund S	n/a	Student Investment Fund	Large	Blend	0.00	2010-10-31
Fund Sn	n/a	Fund S net of fees	Large	Blend	0.50	2010-10-31
AAPL	AAPL	Apple Inc. common stock	Large	Blend	n/a	1980-12-12
SPX	SPX	S&P 500 Index	Large	Blend	n/a	1927-12-30
Fund 1X	IVV	iShares Core S&P 500 ETF	Large	Blend	0.03	2000-05-19
Fund 2X	SSO	ProShares Ultra S&P 500 ETF	Large	Blend	0.90	2006-06-21
Fund 3X	UPRO	ProShares UltraPro S&P 500	Large	Blend	0.92	2009-06-25
Rf	GB1M	One month treasury bill rate	n/a	n/a	n/a	1989-12-01

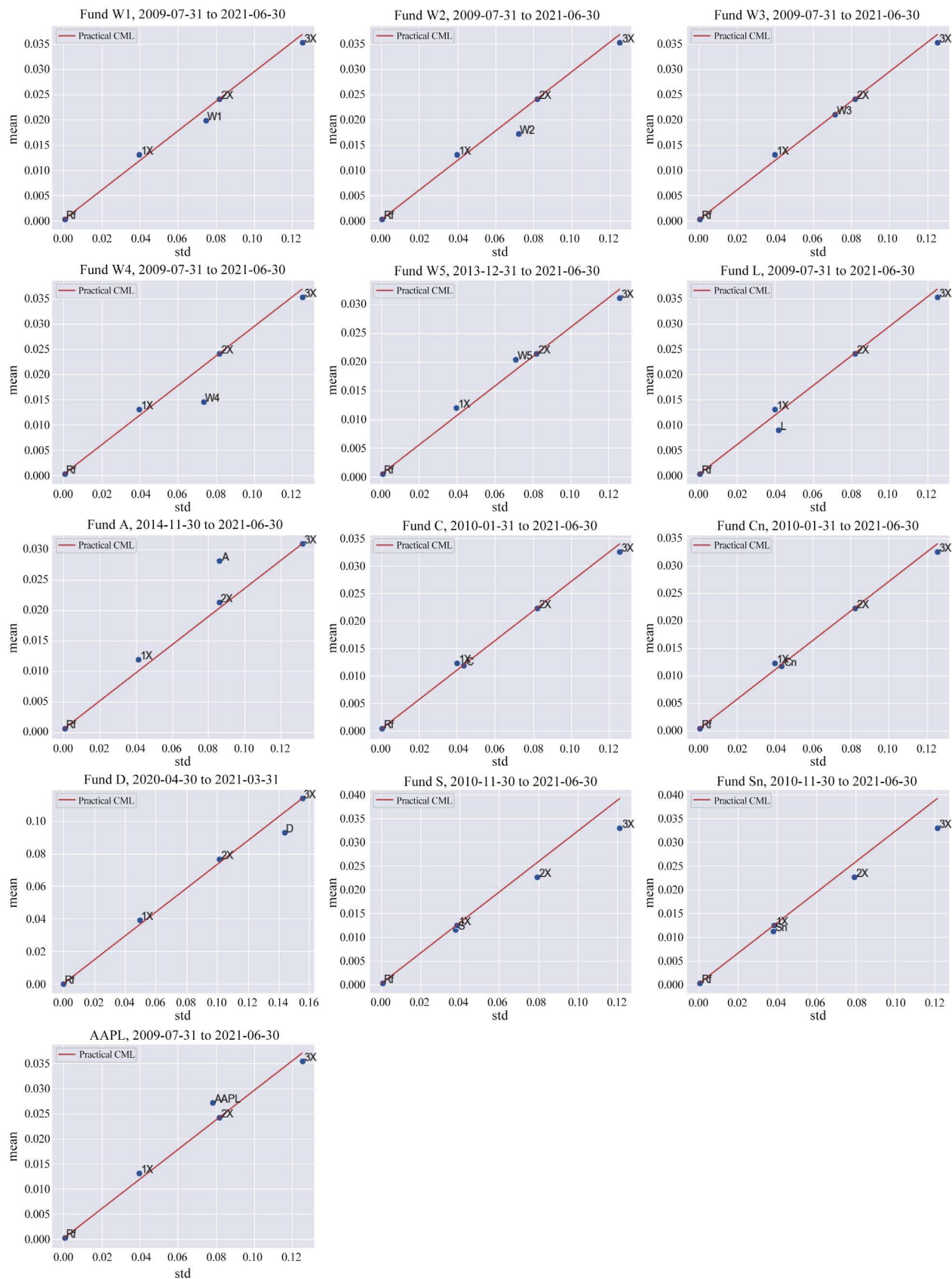
Since this study is an in-sample assessment of the *Moore Performance Question*, missing data issues are mitigated. To extrapolate results out-of-sample, or to make comparisons between financial instruments, one must deal with varying fund return data availability. Such considerations are left for future research. As such, this study does not fill missing data with any values. Rather, it focuses on the data that are available. This focus is evident in the following section that explicitly lists start and end dates for each time series.

### 3.3. Summary Statistics

**Table 5** presents summary statistics for monthly returns. The table illustrates the diversity in data availability (start and end dates) as indicated in the last two columns. **Figure 3** visualizes the summary statistics and plots what I call a “practical” Capital Market Line (CML). The Practical CML combines the risk-free asset (one month Treasury Bill) with one of three ETFs: the unlevered S&P500 (Fund 1X: IVV), the 2X levered S&P 500 (Fund 2X: SSO), or the 3X levered S&P 500 (Fund 3X: UPRO). The decision rule on which fund to use for the market portfolio is as follows:

**Table 5.** Monthly gross return summary statistics.

	count	mean	std	min	50%	max	start	end
Fund W1	362	0.0135	0.0710	−0.2132	0.0143	0.3453	1991-05-31	2021-06-30
Fund W2	281	0.0135	0.0747	−0.2238	0.0120	0.2510	1998-02-28	2021-06-30
Fund W3	353	0.0162	0.0751	−0.3414	0.0180	0.3281	1992-02-29	2021-06-30
Fund W4	344	0.0105	0.0712	−0.3330	0.0120	0.2445	1992-11-30	2021-06-30
Fund W5	91	0.0204	0.0709	−0.2135	0.0161	0.2184	2013-12-31	2021-06-30
Fund L	224	0.0061	0.0441	−0.1735	0.0109	0.1265	2002-11-30	2021-06-30
Fund A	80	0.0282	0.0861	−0.1818	0.0153	0.2575	2014-11-30	2021-06-30
Fund C	138	0.0119	0.0436	−0.1562	0.0154	0.1331	2010-01-31	2021-06-30
Fund Cn	138	0.0117	0.0436	−0.1564	0.0152	0.1328	2010-01-31	2021-06-30
Fund D	12	0.0931	0.1438	−0.1883	0.1094	0.2808	2020-04-30	2021-03-31
Fund S	128	0.0116	0.0380	−0.1136	0.0161	0.1279	2010-11-30	2021-06-30
Fund Sn	128	0.0112	0.0380	−0.1140	0.0157	0.1274	2010-11-30	2021-06-30
AAPL	463	0.0238	0.1272	−0.5774	0.0252	0.4540	1982-12-31	2021-06-30
SPX	1122	0.0064	0.0538	−0.2994	0.0091	0.3914	1928-01-31	2021-06-30
Fund 1X	253	0.0069	0.0430	−0.1663	0.0124	0.1268	2000-06-30	2021-06-30
Fund 2X	180	0.0156	0.0896	−0.3493	0.0250	0.2518	2006-07-31	2021-06-30
Fund 3X	144	0.0354	0.1252	−0.4813	0.0475	0.3717	2009-07-31	2021-06-30
Rf	379	0.0021	0.0018	−0.0000	0.0017	0.0068	1989-12-31	2021-06-30



**Figure 3.** Summary stats in mean-std space.

$$sd_i > sd_{2X} \rightarrow \text{use Fund 3X}$$

$$sd_i > sd_{1X} \rightarrow \text{use Fund 2X}$$

$$sd_i \leq sd_{1X} \rightarrow \text{use Fund 1X}$$

where  $sd_i$  is the standard deviation of the fund (or stock) of interest and  $sd_{kX}$  is the standard deviation of the respective S&P 500 ETF. Note that in order to plot 1X, 2X, and 3X on the same graph the sample is limited by the younger of the three funds (Fund 3X, UPRO, with a 2009-06-25 inception date).

Consistent with market efficiency, the majority of instruments (9 out of 13 or roughly 70%) are below the Practical CML (practical efficient frontier). However, Fund W5 (HDPMX), Fund A (ARKK), and AAPL are above the practical CML. It is worth noting that Fund W5 and Fund A are relatively new having the least amount of return history (sans Fund D). AAPL however, has a longer track record of being above the Practical CML (dating back to 2009-07-31, the start date limited by the Fund 3X inception).

Fund A (ARKK, the ARK Innovation ETF), managed by Cathie Wood, is an excellent example of digesting performance numbers with caution. By caution I mean in the context of the nine ingredients of Section 2.2, specifically sample size and time period. ARKK receives much attention in the press for market-beating performance yet is a relatively new fund started in late 2014. For instance, Seeking Alpha ranked ARKK third in its list of best innovation growth funds for 2021 [27]. de la Hoz [27] states “returns could be, and have been, outstanding.” Although returns were +150% in 2020, ARKK was down over 3% through July 2021 [28]. Skepticism in ARKK’s continued success is prevalent, to the extent that an anti-ARKK ETF (SARK for short ARKK) is in the works [28].

On to Apple Inc., which is substantially above the Practical CML. Apple has been around since 1980 and has its ups and downs. But for the past 12 years, Apple significantly outperformed the market. This relates to Bessembinder [26] who finds the bulk of US stock market gains are concentrated in the top 4% of listed companies while the remainder earn roughly the same as Treasury Bills. However, Bessembinder points out that the existence of persons able reliably identify such top performing stocks in advance is an open question.

One final observation before moving on to developing the Risk-Equivalent Excess Performance measure (which measures the distance from the Practical CML). Where the theoretical CML presumes borrowing and lending at  $R_f$  the Practical CML relies on ETF efficiencies (economies of scale, use of derivatives, etc.) to implement leverage at a much lower cost than many individual investors in a real-world setting.

In all sub-plots, which reflect varied time-frames, we see the cost of leverage increases with leverage. That is, a line between  $R_f$  and 1X will have a higher slope than a line between  $R_f$  and 2X which in turn will have a higher slope than a line between  $R_f$  and 3X. This is not surprising looking at the expense ratios for the market ETFs in Table 4 (0.03% for 1X, 0.90% for 2X, 0.92% for 3X).

### 3.4. Risk-Equivalent Excess Performance (REEP) Measure

Authors Modigliani and Modigliani [10] address performance measurement with a straightforward question: “do returns adequately compensate us for the risk what we bear?” In this study I address performance measurement via a similar yet distinct question: Could an investor have earned higher returns with the same risk (standard deviation) using a simple combination of the risk-free asset and a readily available levered (or unlevered) market ETF? **Figure 4** depicts the Risk-Adjusted Performance (RAP) measure of Modigliani and Modigliani [10] and the Risk-Equivalent Excess Performance (REEP) measure of this study.

In their RAP measure, Modigliani and Modigliani [10] de-lever (or lever) the portfolio under consideration to match the risk (standard deviation) of the market. In contrast, I lever (or de-lever) the *market* portfolio to match the risk of the portfolio under consideration to obtain the REEP measure. In other words, Modigliani and Modigliani [10] move the portfolio of interest to the market whereas REEP (Risk-Equivalent Excess Performance) meets the portfolio where it is. The following describes the construction of RAP and REEP measures.

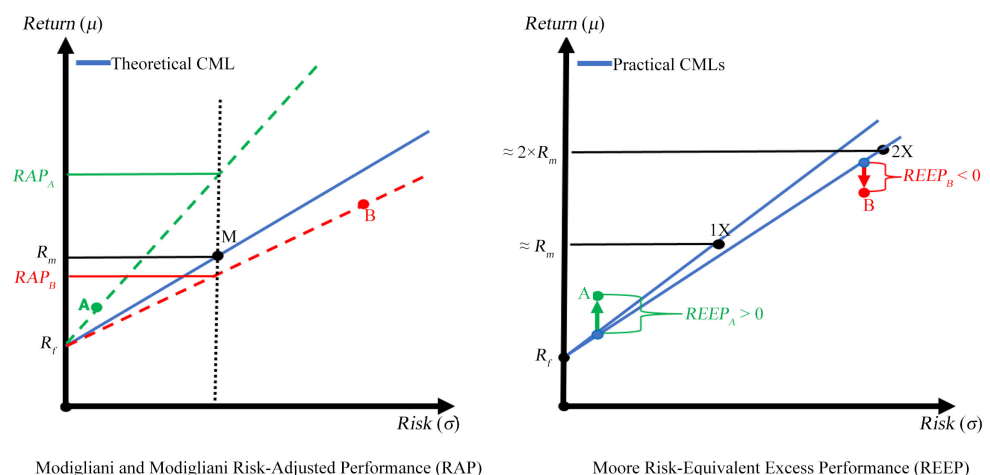
**Risk-Adjusted Performance (RAP).** Consider a portfolio under consideration  $P$  with mean return  $\hat{R}_p$  and standard deviation  $\sigma_p$ . Let  $\hat{R}_m$  and  $\sigma_m$  represent the mean and standard deviation of the market portfolio, respectively. Let  $R_f$  represent the return of the risk-free asset (1-month Treasury bill). The portfolio  $P$  can be levered (or de-levered) using the risk-free asset to construct a new portfolio  $P^*$  with the same standard deviation as the market. The leverage and mean return for such a portfolio is as follows:

$$\text{Leverage : } d_{rap} = \frac{\sigma_m}{\sigma_p} \quad (1)$$

$$\text{Mean return : } \hat{R}_{p^*} = d_{rap} \hat{R}_p + (1 - d_{rap}) R_f \quad (2)$$

$$RAP = \hat{R}_{p^*} \quad (3)$$

Authors Modigliani and Modigliani [10] refer to Equation (2) as RAP.



**Figure 4.** Risk-Adjusted Performance (RAP) vs. Risk-Equivalent Excess Performance (REEP).



*Risk-Equivalent Excess Performance (REEP)*. Consider the same portfolio under consideration, market portfolio, and risk-free asset of the *RAP* measure above. The *market* portfolio  $M$  can be levered (or de-levered) using the risk-free asset to construct a new portfolio  $P^{**}$  with the same standard deviation as the the portfolio under consideration. The leverage, mean return, and *REEP* measure are as follows:

$$\text{Leverage : } d_{reep} = \frac{\sigma_p}{\sigma_m} \quad (4)$$

$$\text{Mean return : } \hat{R}_{p^{**}} = d_{reep} \hat{R}_m + (1 - d_{reep}) R_f \quad (5)$$

$$REEP = \hat{R}_p - \hat{R}_{p^{**}} \quad (6)$$

Aside from the important difference in leverage (matching the market risk in *RAP* vs. matching the portfolio risk in *REEP*), *REEP* is an excess return measure. I construct *REEP* as such to provide an answer to the *Moore Performance Question* posed earlier in this study: Could an investor have earned higher returns with the same risk (standard deviation) using a simple combination of the risk-free asset and a readily available levered (or unlevered) market ETF? As such, the interpretation of *REEP* is straightforward:

- If  $REEP > 0 \rightarrow$  an investor could *not* have obtained higher return at the same level of risk as the portfolio under consideration by combining the risk-free asset and a market ETF.
- If  $REEP < 0 \rightarrow$  an investor *could* have obtained higher return at the same level of risk as the portfolio under consideration by combining the risk-free asset and a market ETF.

Now on to the results.

## 4. Results

### 4.1. Full Sample

For the context of computing *REEP*, I define the full sample as the time period from the first available monthly return of the youngest leveraged market ETF (Fund 3X: UPRO, 2009-07-31) to 2021-06-30. **Table 6** presents results for the full sample. The results are consistent with the findings in Section 3.3: only three portfolios have positive Risk-Equivalent Excess Performance (Fund W5, Fund A, and AAPL).

As an example of interpreting the results, look to Fund C and Fund Cn of **Table 6**. First, the risk (standard deviation) of Fund C exceeds that of the S&P500 (or the unlevered S&P500 ETF IVV) given the selection of benchmark 2X. Note, although the portfolio standard deviation  $sp$  is less than the market standard deviation ( $sm$ ),  $sm$  refers to the standard deviation of Fund 2X (SSO) not the S&P 500. Second, Fund C does not generate positive *REEP* before or after fees. Over the period of analysis, the pension fund would have had higher returns at the same level of risk by purchasing Treasury Bills and Fund 2X (SSO). Thus, the issue of underperformance is more than just fees.

**Table 6.** REEP calculations, full sample (2009-07-31 to 2021-06-30).

Portfolio~Benchmark	T	Start	End	sp	sm	d	REEP
Fund W1~2X	144	2009-07-31	2021-06-30	0.0746	0.0819	0.9108	−0.0022
Fund W2~2X	144	2009-07-31	2021-06-30	0.0720	0.0819	0.8788	−0.0041
Fund W3~2X	144	2009-07-31	2021-06-30	0.0715	0.0819	0.8732	−0.0001
Fund W4~2X	144	2009-07-31	2021-06-30	0.0734	0.0819	0.8960	−0.0072
Fund W5~2X	91	2013-12-31	2021-06-30	0.0709	0.0821	0.8633	0.0018
Fund L~2X	144	2009-07-31	2021-06-30	0.0420	0.0819	0.5130	−0.0036
Fund A~3X	80	2014-11-30	2021-06-30	0.0861	0.1318	0.6534	0.0078
Fund C~2X	138	2010-01-31	2021-06-30	0.0436	0.0821	0.5310	−0.0001
Fund Cn~2X	138	2010-01-31	2021-06-30	0.0436	0.0821	0.5313	−0.0003
Fund D~3X	12	2020-04-30	2021-03-31	0.1438	0.1553	0.9257	−0.0122
Fund S~1X	128	2010-11-30	2021-06-30	0.0380	0.0384	0.9879	−0.0007
Fund Sn~1X	128	2010-11-30	2021-06-30	0.0380	0.0384	0.9879	−0.0011
AAPL~2X	144	2009-07-31	2021-06-30	0.0781	0.0819	0.9540	0.0041

## 4.2. Pre-Pandemic Peak

In 2020 the S&P 500 peaked on 2020-02-19 at 3386 and bottomed on 2020-03-23 at 2237. Therefore I define the pre-pandemic peak period as the monthly returns from 2009-07-31 (first month of returns for youngest market ETF) to 2020-02-29. **Table 7** presents *REEP* calculations and **Figure 5** visualizes the summary statistics for the pre-pandemic peak period.

Like the full-sample in **Table 6**, three portfolios in the pre-pandemic period have generated positive *REEP*. However, Fund W3 in the pre-pandemic period replaces Fund W5 from the full sample.

## 4.3. Post-Pandemic 12 Month Bull

In 2020 the S&P 500 bottomed on 2020-03-23 at 2237. Therefore I define post-pandemic 12 month bull period as the monthly returns from 2020-04-30 to 2021-03-31. **Table 8** presents *REEP* calculations and **Figure 6** visualizes the summary statistics for the post-pandemic 12 month bull period.

Consistent the Wall Street Journal (WSJ) article praising pandemic performance, Funds W1-W5 all generated positive *REEP* during this period. In fact, of the 13 portfolios examined, only three did not generate positive *REEP*: Fund L, Fund D, and AAPL. Looking to **Table 4**, we see that Fund L has the second highest expense ratio at 1.78%. Given Fund L, a large-cap value fund, has a much higher cost than its benchmark (Fund 1X with a 0.03% expense ratio), it is not surprising that Fund L generated negative *REEP*. Fund D represents a concentrated position of 7 stocks that generated higher *return* than the market (Fund 1X)—but less than a combination of 93% in Fund 3X (**Table 8**, column *d*) and 7% in Treasury Bills.

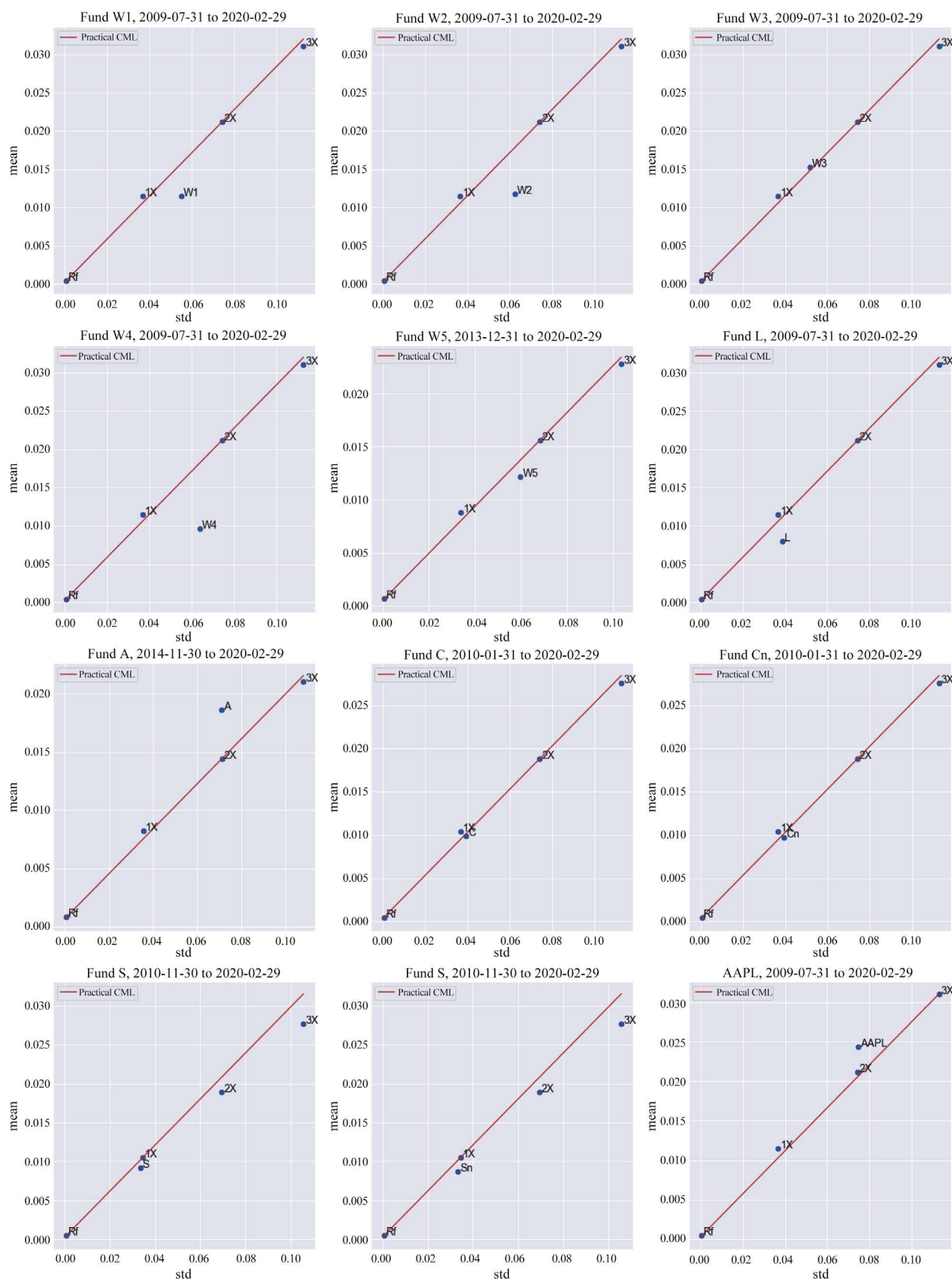
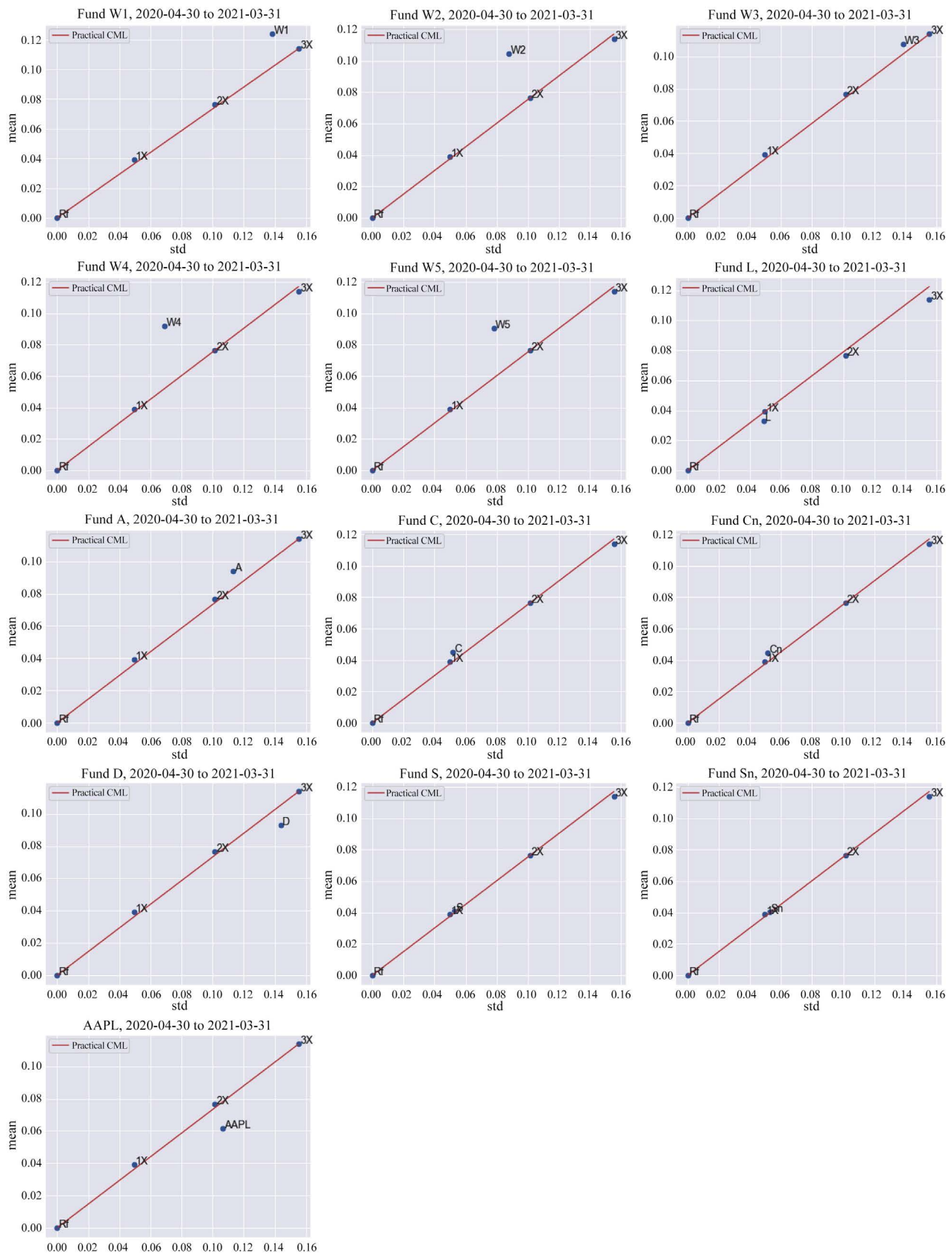


Figure 5. Pre-pandemic summary stats in mean-std space.



**Figure 6.** Post-pandemic 12 month bull summary stats in mean-std space.

**Table 7.** REEP calculations, Pre-pandemic (2009-07-31 to 2020-02-29).

Portfolio~Benchmark	T	Start	End	sp	sm	d	REEP
Fund W1~2X	128	2009-07-31	2020-02-29	0.0550	0.0743	0.7403	-0.0043
Fund W2~2X	128	2009-07-31	2020-02-29	0.0627	0.0743	0.8442	-0.0062
Fund W3~2X	128	2009-07-31	2020-02-29	0.0519	0.0743	0.6983	0.0003
Fund W4~2X	128	2009-07-31	2020-02-29	0.0640	0.0743	0.8613	-0.0086
Fund W5~2X	75	2013-12-31	2020-02-29	0.0597	0.0682	0.8757	-0.0016
Fund L~2X	128	2009-07-31	2020-02-29	0.0387	0.0743	0.5218	-0.0033
Fund A~2X	64	2014-11-30	2020-02-29	0.0711	0.0712	0.9990	0.0042
Fund C~2X	122	2010-01-31	2020-02-29	0.0394	0.0741	0.5320	-0.0004
Fund Cn~2X	122	2010-01-31	2020-02-29	0.0394	0.0741	0.5323	-0.0006
Fund S~1X	112	2010-11-30	2020-02-29	0.0334	0.0345	0.9700	-0.0010
Fund Sn~1X	112	2010-11-30	2020-02-29	0.0334	0.0345	0.9700	-0.0014
AAPL~3X	128	2009-07-31	2020-02-29	0.0748	0.1128	0.6624	0.0037

**Table 8.** REEP calculations, Post-pandemic bull (2020-04-30 to 2021-03-31).

Portfolio~Benchmark	T	Start	End	sp	sm	d	REEP
Fund W1~3X	12	2020-04-30	2021-03-31	0.1384	0.1553	0.8912	0.0222
Fund W2~2X	12	2020-04-30	2021-03-31	0.0879	0.1015	0.8659	0.0384
Fund W3~3X	12	2020-04-30	2021-03-31	0.1385	0.1553	0.8913	0.0061
Fund W4~2X	12	2020-04-30	2021-03-31	0.0692	0.1015	0.6824	0.0399
Fund W5~2X	12	2020-04-30	2021-03-31	0.0782	0.1015	0.7708	0.0317
Fund L~1X	12	2020-04-30	2021-03-31	0.0489	0.0496	0.9860	-0.0055
Fund A~3X	12	2020-04-30	2021-03-31	0.1132	0.1553	0.7288	0.0112
Fund C~2X	12	2020-04-30	2021-03-31	0.0515	0.1015	0.5077	0.0060
Fund Cn~2X	12	2020-04-30	2021-03-31	0.0515	0.1015	0.5078	0.0058
Fund D~3X	12	2020-04-30	2021-03-31	0.1438	0.1553	0.9257	-0.0122
Fund S~2X	12	2020-04-30	2021-03-31	0.0526	0.1015	0.5185	0.0012
Fund Sn~2X	12	2020-04-30	2021-03-31	0.0526	0.1015	0.5185	0.0007
AAPL~3X	12	2020-04-30	2021-03-31	0.1064	0.1553	0.6850	-0.0163

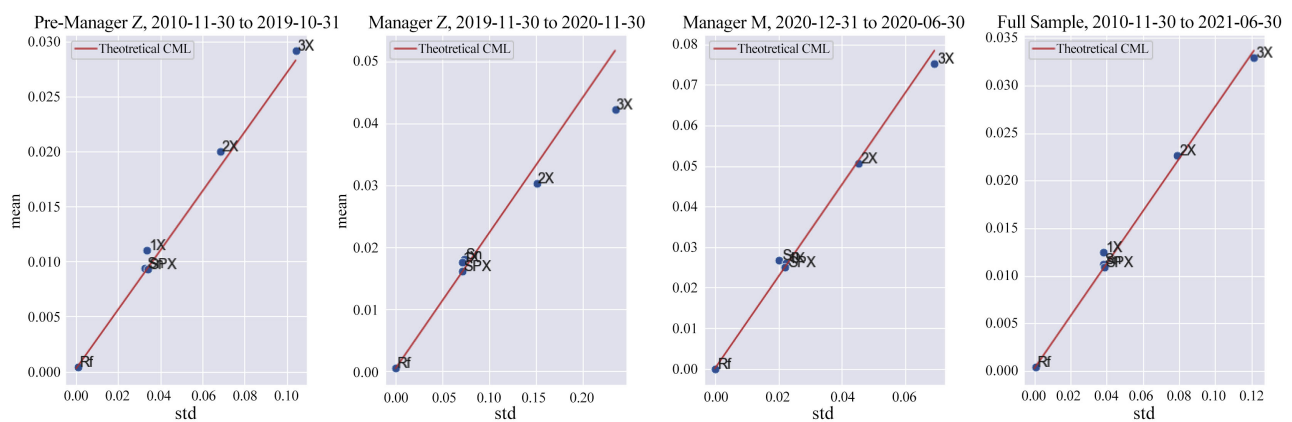
#### 4.4. Time-Varying SIF Management

**Table 9** presents the results for the varied management time periods. Bear in mind that the column for market standard deviation (*sm*) is the standard deviation corresponding to the selected benchmark (1X, 2X, or 3X). The Manager Z era is the only era where the portfolio risk (standard deviation) necessitated moving to a levered ETF (2X). However, Manager Z did earn the highest monthly *REEP* (0.0031 or 0.31%), albeit by a minuscule amount (0.0002, 0.02%, or 2 basis points) and by taking on more risk than stated by the fund's prospectus<sup>1</sup>.

<sup>1</sup>The prospectus states that risk should be in-line with the S&P 500. However, as shown in **Table 9**, the risk (standard deviation) exceeded that of the benchmark (1X) necessitating the use of the 2X levered fund as the benchmark.

**Table 9.** REEP calculations, Time varying SIF management.

	Portfolio~Benchmark	T	Start	End	sp	sm	d	REEP
Pre-Manager Z	Fund S~1X	108	2010-11-30	2019-10-31	0.0325	0.0337	0.9632	-0.0008
	Fund Sn~1X	108	2010-11-30	2019-10-31	0.0325	0.0337	0.9632	-0.0012
Manager Z	Fund S~2X	13	2019-11-30	2020-11-30	0.0736	0.1515	0.4856	0.0035
	Fund Sn~2X	13	2019-11-30	2020-11-30	0.0736	0.1515	0.4856	0.0031
Manager M	Fund S~1X	7	2020-12-31	2021-06-30	0.0203	0.0222	0.9125	0.0033
	Fund Sn~1X	7	2020-12-31	2021-06-30	0.0203	0.0222	0.9125	0.0029
Full sample	Fund S~1X	128	2010-11-30	2021-06-30	0.0380	0.0384	0.9879	-0.0007
	Fund Sn~1X	128	2010-11-30	2021-06-30	0.0380	0.0384	0.9879	-0.0011

**Figure 7.** SIF summary stats during different management regimes.

Again, Ingredients 6 (Time period) and 7 (Sample size) apply: we have very small sample sizes (13 months for Manager Z and 7 months for Manager M) from a rather unique time-period in history (COVID-19 era) that we hope is not representative of the population that includes future returns. Thus, any conclusions of superior performance should be taken with several grains of salt. This reiterates the need for future research that utilizes a larger sample or simulation.

**Figure 7** visualizes the summary statistics associated with the Student Investment Fund returns net of fees during different management regimes. Three observations are of note. First, the market ETF (1X:IVV) is above the Theoretical CML in three out of four time frames and on the line in the third. This suggests the managers at iShares are doing a good job keeping costs down and in fact enhancing returns of clients vs. the benchmark index. Second, although the SIF portfolio (Sn) is above the Theoretical CML under both Manager Z and Manager M regimes, Manager M is further away from the CML indicating Manager M's superior performance with respect to the fund's stated benchmark (S&P 500 Index rather than any of the market ETFs). Finally, when looking at the full-sample (fourth chart on the right), the SIF portfolio net of fees lies on the Theoretical CML.

One last comment before moving on to the conclusion. As mentioned in Sec-



tion 3, this study does not address Ingredient 9: Luck vs. skill. Given the absence of a luck-vs-skill measure (combined with a small sample that is not representative of the population) one can not determine if the performance results from Managers Z and M are due to luck or skill.

## 5. Conclusions

In this study, I introduce a parsimonious and practical Risk-Equivalent Excess Performance (*REEP*) measure based on the well-known Modigliani and Modigliani [10] Risk-Adjusted Performance (*RAP*) measure. In addition, I highlight nine key ingredients of valid performance measurement: gains and losses, cash, fees and costs, taxes, risk, time period, sample size, appropriate benchmark, and luck vs. skill. I survey the literature relevant to these nine ingredients and discuss the approach of this study to address those concerns.

The results indicate few financial instruments (Fund W5, Fund A, and APPL) generate positive *REEP* over the 2009-07-31 to 2021-06-30 period. Even then, Fund W5 and Fund A had the smallest sample size (sans Fund D) of the sample. Furthermore, Fund W5 generated negative *REEP* in the pre-pandemic period. I also provide an example of how to address time-varying fund management when assessing performance. I accomplished this utilizing proprietary data from a university Student Investment Fund. Results affirm the need to be mindful of the nine ingredients of valid performance measurement in that there is insufficient evidence to conclusively determine a “winning” manager. Ultimately, the real winners are all students that participate in Student Investment Funds as they gain knowledge and skills useful in the workforce.

Rather than attempt to develop a method to reliably predict the future, *i.e.*, which financial instruments or managers will outperform the market on a risk-equivalent basis, this study examines claims of *ex-post* (observed) data. No one can reliably predict the future and therefore no one can reliably measure future performance. However, the new Risk-Equivalent Excess Performance measure and contextualization provide a basis for future out-of-sample inferential analysis. But, as always, “past performance is no guarantee of future results.”

## Acknowledgements

First, I would like to thank the reviewers for excellent feedback that improved this paper. I thank the local pension fund for providing data on one of their actively managed funds for this research. I thank “Student D” for developing a “one-year pandemic upside” portfolio and engaging in a friendly wager (that I ultimately lost due to omitting risk). I also thank the university that provided Student Investment Fund data and relevant information for this study.

## Conflicts of Interest

The author declares no conflicts of interest regarding the publication of this paper.



## References

- [1] Hoffman, A.O.I. and Post, T. (2014) Self-Attribution Bias in Consumer Financial Decision-Making: How Investment Returns Affect Individuals' Belief in Skill. *Journal of Behavioral and Experimental Economics*, **52**, 23-28.  
<https://doi.org/10.1016/j.socec.2014.05.005>
- [2] McGee, S. (2021) The Pandemic Year's Top Stock-Fund Managers. *Wall Street Journal Online*.  
<https://www.wsj.com/articles/the-pandemic-years-top-stock-fund-managers-11617555600>
- [3] Shilling, A. (2016) A Question for Active: Is It Luck or Skill? *Money Management Executive*, **24**, 1, 8-9.
- [4] Robinhood Markets, Inc. (2021) Form S-1 Registration Statement. United States Securities and Exchange Commission, Washington DC.
- [5] Krantz, M. (2021) \$26 Billion Gone! "Meme Stock" Crash Erases Nearly Half of Gains. *Investor's Business Daily*.  
<https://www.investors.com/etfs-and-funds/sectors/meme-stocks-billion-gone-meme-stock-crash-wipes-out-gains/>
- [6] Griffith, K. (2021) Robinhood Claims It Only Forced Sale of Gamestop Shares If They Were Bought with Borrowed Funds. *Dailymail*.  
<https://www.dailymail.co.uk/news/article-9205303/Robinhood-claims-forced-sale-GameStop-shares-bought-borrowed-funds.html>
- [7] Cogneau, P. and Hubner, G. (2009) The (More than) 100 Ways to Measure Portfolio Performance, Part 1: Standardized Risk-Adjusted Measures. *Journal of Performance Measurement*, **13**, 56-71.
- [8] Altman, D.G. (1991) *Practical Statistics for Medical Research*. Chapman and Hall, London.
- [9] Schoolman, H.M., Becketl, J.M., Best, W.R. and Johnson, A.F. (1968) Statistics in Medical Research: Principles versus Practices. *The Journal of Laboratory and Clinical Medicine*, **71**, 357-367.
- [10] Modigliani, F. and Modigliani, L. (1997) Risk-Adjusted Performance: How to Measure It and Why. *Journal of Portfolio Management*, **23**, 45-54.  
<https://doi.org/10.3905/jpm.23.2.45>
- [11] Thaler, R. (1980) Toward a Positive Theory of Consumer Choice. *Journal of Economic Behavior & Organization*, **1**, 39-60.  
[https://doi.org/10.1016/0167-2681\(80\)90051-7](https://doi.org/10.1016/0167-2681(80)90051-7)
- [12] Basak, S. (2019) Buffett Slams Private Equity over Inflated Returns, Debt. *Bloomberg.com*.  
<https://www.bloomberg.com/news/articles/2019-05-04/buffett-slams-private-equity-for-inflated-returns-debt-reliance>
- [13] Picker, L. (2021) Head of the Second-Largest U.S. Public Pension Fund Says Active Managers Rarely Added Value. *CNBC.com*.  
<https://www.cnbc.com/2021/06/28/christopher-ailman-of-calstrs-active-managers-rarely-added-value-above-fees.html>
- [14] Internal Revenue Service (2021, March) Topic No. 409 Capital Gains and Losses.  
<https://www.irs.gov/taxtopics/tc409>
- [15] Muralidhar, A.S. (2004) History Matters. *Derivatives Use, Trading, and Regulation*, **10**, 131-148.
- [16] Roll, R. (1977) A Critique of the Asset Pricing Theory's Tests Part I: On Past and

- Potential Testability of the Theory. *Journal of Financial Economics*, **4**, 129-176.  
[https://doi.org/10.1016/0304-405X\(77\)90009-5](https://doi.org/10.1016/0304-405X(77)90009-5)
- [17] Doeswijk, R., Lam, T. and Swinkels, L. (2020) Historical Returns of the Market Portfolio. *The Review of Asset Pricing Studies*, **10**, 521-567.  
<https://doi.org/10.1093/rapstu/raz010>
  - [18] Stambaugh, R.F. (1982) On the Exclusion of Assets from Tests of the Two-Parameter Model: A Sensitivity Analysis. *Journal of Financial Economics*, **10**, 237-268.  
[https://doi.org/10.1016/0304-405X\(82\)90002-2](https://doi.org/10.1016/0304-405X(82)90002-2)
  - [19] Treynor, J. and Mazuy, K. (1966) Can Mutual Funds Outguess the Market? *Harvard Business Review*, **44**, 131-136.
  - [20] Grinold, R.C. (1989) The Fundamental Law of Active Management. *Journal of Portfolio Management*, **15**, 30-37. <https://doi.org/10.3905/jpm.1989.409211>
  - [21] Grinold, R.C. and Kahn, R.N. (1992) Information Analysis. *Journal of Portfolio Management*, **18**, 14-21. <https://doi.org/10.3905/jpm.1992.409408>
  - [22] Goodwin, T.H. (1998) The Information Ratio. *Financial Analysts Journal*, **54**, 34-43.  
<https://doi.org/10.2469/faj.v54.n4.2196>
  - [23] Fama, E.F. and French, K.R. (2010) Luck versus Skill in the Cross-Section of Mutual Fund Returns. *The Journal of Finance*, **65**, 1915-1947.  
<https://doi.org/10.1111/j.1540-6261.2010.01598.x>
  - [24] Bhootra, A., Drezner, Z., Schwarz, C. and Stohs, M.H. (2015) Mutual Fund Performance: Luck or Skill? *International Journal of Business*, **20**, 52-63.
  - [25] Bloomberg, L.P. (2021) Stock, Treasury Bill, and Index Data.
  - [26] Bessembinder, H. (2018) Do Stocks Outperform Treasury Bills? *Journal of Financial Economics*, **129**, 440-457. <https://doi.org/10.1016/j.jfineco.2018.06.004>
  - [27] de la Hoz, J. (2021) Best Innovation Growth Funds for 2021. Seekingalpha.com.  
<https://seekingalpha.com/article/4413071-best-innovation-growth-funds-for-2021>
  - [28] Greifeld, K. (2021) Anti-Ark ETF to Bet against Cathie Wood's Flagship Fund. Bloomberg.com.  
<https://www.bloomberg.com/news/articles/2021-07-30/anti-ark-etf-to-bet-against-cathie-wood-s-flagship-fund>