1 Introduction

In recent years, the rise of commission-free trading platforms has profoundly reshaped retail investor behavior and sparked growing interest among scholars in behavioral finance. One of the most prominent examples is Robinhood, a mobile-first brokerage that gained widespread popularity for eliminating trading fees and offering a highly gamified user experience. The platform attracted a large number of retail investors, particularly young and inexperienced individuals[1].

An important development in the empirical literature was the release of Robintrack, an open-source dataset that tracks the number of Robinhood users holding individual stocks over time. This data, collected via Robinhood's public API, provides a rare opportunity to directly observe the trading dynamics and portfolio shifts of real retail investors. The dataset can be downloaded from https://robintrack.net/.

Several recent studies, including [Fedyk, 2024] and [Welch, 2022][2], have leveraged the Robintrack dataset to examine retail investor performance. Their findings suggest that Robinhood investors, contrary to popular belief, exhibited strong market timing and outperformed passive benchmarks. In particular, these papers report significant cumulative returns and positive alpha using standard factor models.

In this paper, we revisit these claims by constructing an alternative methodology for portfolio formation based on the same dataset. Specifically, we analyze whether the returns of Robinhood users' favorite stocks exhibit stochastic dominance over benchmark indices. Our goal is to offer a more nuanced assessment of whether retail investors truly generate abnormal returns or whether previous results may be driven by sample selection or methodological choices.

2 Robinhood Dataset

2.1 Description of the Dataset

The dataset records the **number** of Robinhood users holding at least one share of 8,619 securities, with observations taken hourly. Following [Welch, 2022] and [Fedyk, 2024], we aggregate this data on a daily basis by selecting the last observation of each trading day.

[1] According to Robinhood's IPO filing, the typical user on the platform is 31 years old, with an average account balance of approximately \$3,500. Notably, around half of the platform's users are investing for the first time.

[2] it must be noted that the former explicitly follows the method of the latter

The sample spans from February 5, 2018, to August 13, 2020, covering 818 days. Note that the dataset includes non-trading days and contains some missing observations.

Since the dataset only provides the number of investors per security, we cannot track individual holdings, monetary amounts, or share quantities. Moreover, buy/sell flows are unobservable; however, we can approximate them using changes in the number of holders.

We merge this dataset with CRSP to obtain market-level information and later construct a benchmark index. The resulting dataset contains 7,613 unique securities, substantially more than in [Fedyk, 2024] and [Welch, 2022], who restrict their analysis to U.S. common stocks only. Details of the data cleaning procedure are provided in Appendix B.

In terms of security types, common stocks represent 57.3% of the dataset, while ETFs and other funds account for 26.3% and 8.7%, respectively. Structured products, REITs, and ADRs constitute the remaining share. When classifying by market capitalization, stocks dominate with 83.1%, followed by ETFs (9.3%) and other funds (3.3%).

The total number of open positions on any given day is calculated as the sum of users holding at least one share across all securities—i.e., a row-wise sum across the dataset.

Market data for each security was retrieved from CRSP[3] via WRDS. Out of the full universe, 8,099 securities were available in CRSP, as it includes only U.S.-listed assets.

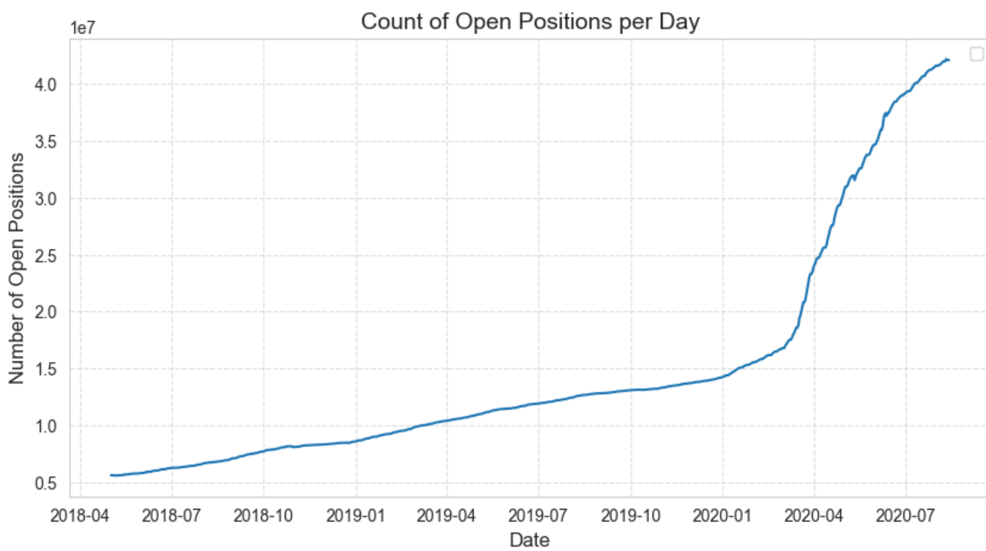


Figure 1: Daily count of open Robinhood positions, May 2018-August 2020.

[3]The Center for Research in Security Prices (CRSP), based at the University of Chicago, provides high-quality historical market data widely used in finance research and investment analysis.

The figure above shows the daily count of open positions on Robinhood from April 2018 to mid-2020. We observe a steady increase in user participation, with a sharp acceleration beginning in early 2020. This surge coincides with the onset of the COVID-19 pandemic, likely driven by a combination of heightened market volatility, increased retail interest, and fiscal stimulus payments.

3 Building the Robinhood Portfolio

As explained above, the biggest limitation of the Robintrack dataset is that it counts the number of users holding a certain security and doesn't provide any information on the amount invested in a particular security.

This section presents a comprehensive framework for constructing and evaluating such a portfolio using distinct weights methods and assumptions. We and the other authors adopt two distinct approaches to compute value and returns. By comparing these approaches, we can better understand how different assumptions about retail investor behavior influence our assessment of the Robinhood crowd's investment performance.

3.1 Weights Methods

[Fedyk, 2024] and [Welch, 2022] use the same approach to build the performance of the Robinhood crowd (or "reference index"): they build daily weights and then apply the weights from the previous day to daily stock returns, directly building portfolio returns.

First, it is necessary to define how those weights are computed. They define two different types of weights, although they yield similar findings in their analysis.

The first method is the "dollar method", which assumes that every investor represents an equal dollar amount investment in the stock.

$$w_{i,t}^{\text{dollar}} = \frac{N_{i,t}}{\sum_j N_{j,t}} \tag{1}$$

where $w_{i,t}$ is the Robinhood portfolio weight of security i at time t and $N_{i,t}$ is the number of investors in security i at time t.

Alternatively, they define the "share method", where each Robinhood investor in a stock represents a one share investment in that stock.

$$w_{i,t}^{\text{share}} = \frac{N_{i,t} \cdot P_{i,t}}{\sum_j N_{j,t} \cdot P_{j,t}} \tag{2}$$

where $P_{j,t}$ is the price of stock j at time t.

The approach we developed differs on how the weights are applied. Nonetheless, we need to define how the components of the portfolio are weighted:

$$w_{i,t}^{\text{mine}} = \frac{N_{i,t}}{\sum_j N_{j,t}} \tag{3}$$

where $w_{i,t}$ is the Robinhood portfolio weight of security i at time t and $N_{i,t}$ is the number of investors in security i at time t. This is identical to 1 but is being listed for clarity.

3.2 Alternative Methodologies for Constructing Portfolio Returns

3.2.1 Fedyk's Approach

The aggregate Robinhood Portfolio returns are derived in [Fedyk, 2024] and [Welch, 2022] by multiplying weights by their daily returns[4], assuming that the weights, however computed, represent a certain share of wealth in a stock held by Robinhood crowd.

From now on, we will define this method as the "Fedyk" method.

Defining Returns Returns differ mathematically based on which weights are applied to returns.

Formally the returns of the Robinhood portfolio, using the dollar method, are defined as:

$$r_t^{\text{dollar}} = \sum_{i=1}^{N} w_{i,t-1}^{\text{dollar}} \cdot r_{i,t} \tag{4}$$

where $r_{i,t}$ is the realized simple return on security i at time t, and $w_{i,t-1}^{\text{dollar}}$ is the dollar weight as defined in equation 1 of security i, for the previous day to avoid look-ahead bias.

Alternatively, they derive returns for the Robinhood portfolio using the share method in the following manner:

$$r_t^{\text{share}} = \sum_{i=1}^{N} w_{i,t-1}^{\text{share}} \cdot r_{i,t} \tag{5}$$

where $r_{i,t}$ is the realized simple return on security i at time t, and $w_{i,t-1}^{\text{share}}$ is the share weight as defined in equation 2 of security i, for the previous day to avoid look-ahead bias.

[4]returns are computed directly by CRSP and are adjusted for dividends, e.g. if $P_0 = 10$ and $D_1 = 5$ and $P_1 = 5$ returns would be 0%

Defining Value Although Fedyk and Welch never define the value of their portfolio in their papers, we can derive the value of the Robinhood Portfolio computed according to their method from compounded returns.

The value of the Robinhood portfolio using the dollar method can be defined as follows:

$$V_T^{\text{dollar}} = V_0 \prod_{t=1}^{T} (1 + r_{\text{dollar},\, t}) \tag{6}$$

where V_0 can be assumed equal to 1 without loss of generality and $r_{\text{dollar},\, t}$ is derived from equation 4.

Using the share method instead, the value of the Robinhood portfolio can be defined as follows:

$$V_T^{\text{share}} = V_0 \prod_{t=1}^{T} (1 + r_{\text{share},\, t}) \tag{7}$$

where V_0 can be assumed equal to 1 without loss of generality and $r_{\text{share},\, t}$ is derived from equation 5.

3.2.2 My Approach

On the other hand, I first compute the value of the Robinhood portfolio by doing a weighted sum of the prices of the securities in the dataset. Conceptually, this represents the portfolio of an investor who decides to allocate a certain number (or percentage) of shares to each security. I will call this this method "Mine" or simply the method built on prices.

We can therefore define the value of the Robinhood Portfolio as follows:

$$V_t^{\text{mine}} = \sum_{i=1}^{N} w_{i,t-1}^{\text{mine}} \cdot P_{i,t} \tag{8}$$

where $w_{i,t-1}^{\text{mine}}$ is the weight of security i computed according to 3 at time $t-1$ and $P_{i,t}$ is the price of security i at time t.

I then track the evolution of the value of the portfolio as defined in 8 to compute returns.

$$r_t^{\text{mine}} = \frac{V_t^{\text{mine}}}{V_{t-1}^{\text{mine}}} - 1 \tag{9}$$

3.3 Capturing the Persistence of Investor Composition

Although both approaches ultimately yield a time series of Robinhood portfolio returns, there is a fundamental difference in what these return paths represent.

In the method used by [Fedyk, 2024] and [Welch, 2022], the portfolio is effectively rebalanced every day to reflect the current composition of investor popularity. Each day's return is computed based on that day's weights and the corresponding daily stock-level returns. This provides a valid snapshot of the average return generated by the stocks held on a given day.

However, this approach does not preserve the economic exposure that investors accumulate through time. A stock that was extremely popular for several days but declines in popularity just before a price spike will have minimal influence on the portfolio's return when that spike occurs. Only the weights at time $t-1$ affect the return at time t[5], so the model captures immediate sentiment shifts but not the cumulative effects of holding positions over time.

In contrast, the methodology I propose (9) applies weights to stock prices and computes returns from changes in total portfolio value. This implies that a stock that was heavily weighted yesterday continues to influence portfolio performance today, even if its popularity has declined. The return reflects both the dynamics of price changes and the path dependency of investor composition.

As a result, my method embeds the effects of investor flows, popularity shifts, and concentration in the actual evolution of portfolio value. The cumulative performance is not a sequence of disconnected daily snapshots, but a reflection of how crowd behavior builds, persists, and unwinds over time.

Conceptually, this distinction is important when studying behavioral dynamics. Retail investor behavior, particularly on platforms like Robinhood, is driven not only by cross-sectional preferences at a point in time but also by persistent patterns of attention, sentiment, and herding. A portfolio that evolves with these behavioral shifts provides a more realistic measure of the actual wealth path experienced by retail investors, rather than an idealized, continually rebalanced index.

In this sense, computing returns from the portfolio value offers a more structurally consistent and behaviorally meaningful representation of the Robinhood crowd's investment trajectory.

4 Comparing Returns and Distribution measures

The biggest difference does not appear when using different kinds of weights (”dollar” or ”share” method) but rather when building the portfolio from prices or returns. Moreover,

[5] Previous day's weights are taken to prevent look-ahead bias

Fedyk and Welch build their portfolio only using common American stocks (share code 10 or 11). In my final analysis I look at all types of securities but significant differences emerge even when using the same sample. Additionally, by recreating the method employed by the other authors we can analyse its return when dealing with all kinds of securities. In section 4.2 I compare returns using only common stocks or the full sample of securities, in the other sections of this paper I will use the full sample unless explicitly stated.

As the other authors have claimed in their papers, the Portfolio built directly from returns had a significantly higher cumulative return compared to the market and positive alpha.

I will proceed to analyse in more detail the distribution of returns of different Robinhood portfolios, showing that my method depicts a far less rosy picture of the ”Robinhood strategy”.

Moreover, [Fedyk, 2024] has analysed extensively the differences between the portfolio obtained using the share method and the dollar method. We'll focus on the dollar method since it is the same approach I use to compute weights.

4.1 Constructing Moving Averages

Defining log returns allows us to simply compute moving averages, showing the profitability of the Robinhood portfolio at different time frames. Conceptually, the value of an n-day moving average on a given date T represents the return an investor would have earned by initiating the position at the open of day $T-n$ and holding it continuously up to close of day T. More rigorously, for a given horizon of n days, the log return over the period is:

$$r_n = \sum_{t=T-n+1}^{T} r_t = \ln\left(\prod_{t=T-n+1}^{T} R_t\right) \tag{10}$$

where r_t are daily log returns and R_t are daily gross returns.

4.2 Retail Performance Under Different Samples

4.2.1 Rolling Returns Across Investment Horizons

Common Stocks Only At short horizons (5-30 days), the two Robinhood portfolios (Fedyk and Mine) display highly similar dynamics, with both closely tracking market indices and exhibiting bursts of volatility during periods of market stress, particularly around the onset of the COVID-19 crash. This suggests that in the very short term, retail investors tend to move in tandem with broader market trends, with limited divergence in return profiles across methodologies. However, at the 120-day horizon, both Robinhood portfolios outperform the S&P 500 and a World ETF[6].

This finding is consistent with the idea that many retail investors, especially on Robinhood, engaged in "buy-the-dip" behavior during the COVID-19 crash. Their increased exposure to beaten-down or speculative stocks during the downturn appears to have been rewarded in the subsequent rebound. Importantly, this also coincides with a period of explosive growth for the platform itself, which may have amplified attention and capital inflows into popular names.

Nonetheless, this post-crash outperformance comes after a prolonged period of clear underperformance. Prior to March 2020, both Robinhood portfolios consistently lag behind the benchmark indices, with my method in particular reflecting substantial drawdowns and poor stock selection.

What differentiates the two methods most clearly is the strength of the post-COVID recovery. While Fedyk's method shows a relatively steady climb, my price-based approach rebounds even more sharply after March 2020. This reflects the fact that, under my methodology, investor positions are not rebalanced away from prior favorites. As a result, stocks that surged after the crash contributed disproportionately to the portfolio's recovery.

In sum, the 120-day results provide two key insights: first, that retail traders on Robinhood did benefit from post-crisis market dynamics, and second, that the magnitude and nature of this benefit depends heavily on the modeling approach, particularly when capturing persistence in portfolio composition.

[6] I've used Vanguard's VOO for the S&P500 and Vanguard's VT as a World Equity ETF

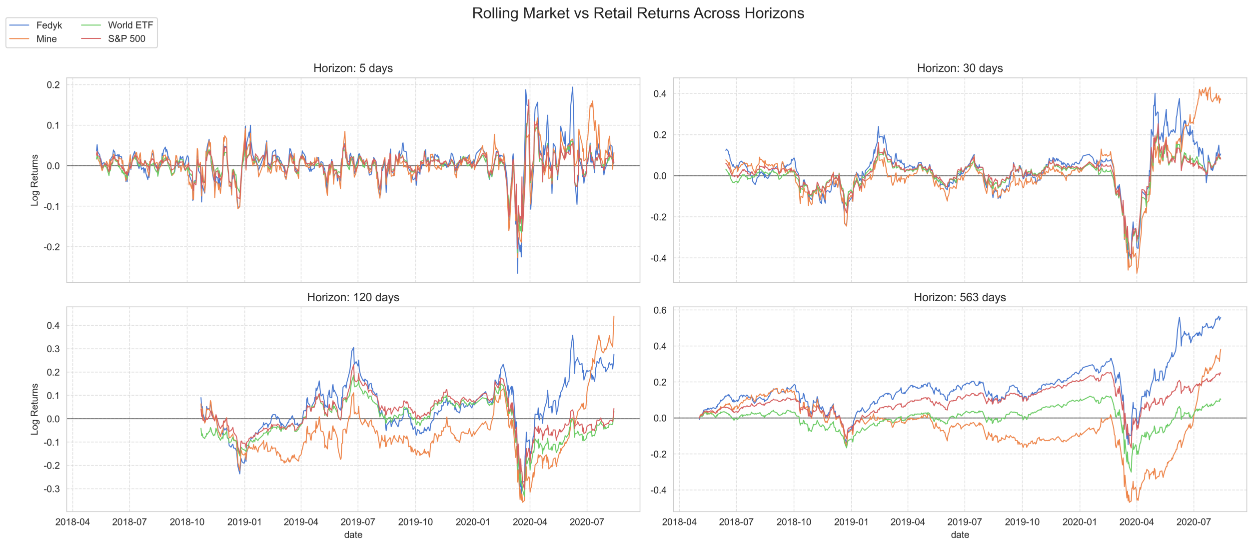


Figure 2: Rolling log-returns of Robinhood and market portfolios at four investment horizons.

Full Universe of Securities We now extend the analysis to all securities in the dataset, including ETFs, REITs, and structured products.

Expanding the sample, we still observe that short-term movements (5-30 days) remain closely correlated across all portfolios, with limited divergence in returns or volatility between methods.

However, the performance gap widens at longer horizons. In contrast to the stock-only case, where both Robinhood portfolios outperformed after the crash, here the drawdowns, particularly in the price-based portfolio, are deeper and more persistent.

Low performance spans the entire pre-COVID period: my portfolio underperforms continuously throughout 2019, while Fedyk's stays closer to the benchmarks but still lags in absolute terms.

At the 563-day horizon, both Robinhood portfolios end below the S&P 500 and the World ETF. My portfolio, while showing a stronger post-crash rebound, barely catches up to the S&P 500 by the end of the sample, and only due to its heavier exposure to post-crash winners. Fedyk's method performs slightly better but remains clearly below the benchmark, reversing the apparent outperformance seen in the stock-only sample.

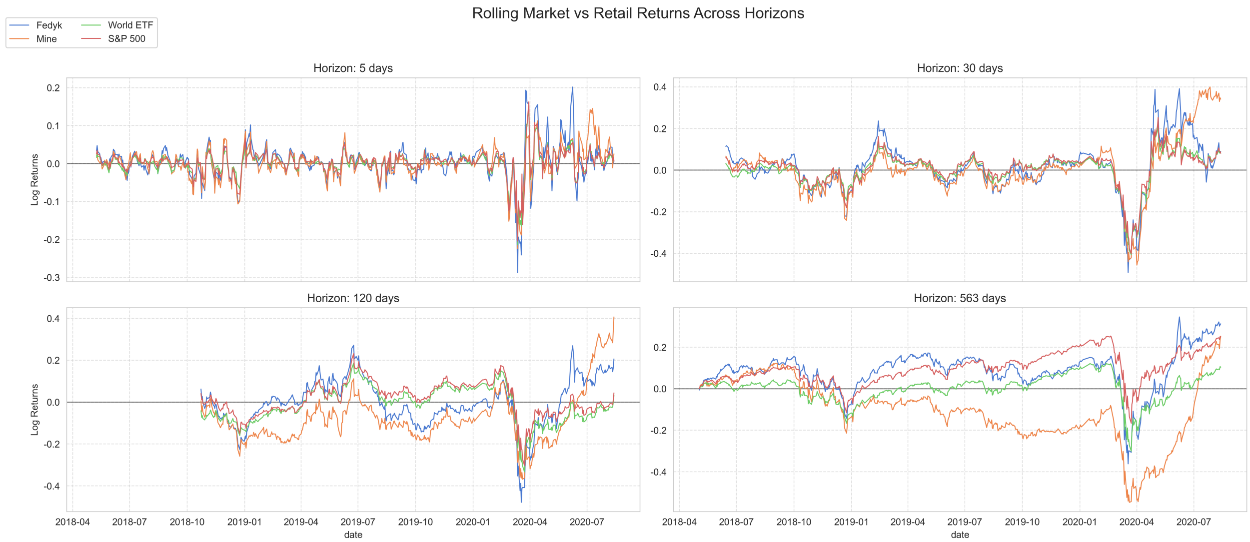


Figure 3: Rolling log-returns of Robinhood and market portfolios at four investment horizons.

A detailed description of the distributions can be found in table 2.

4.2.2 Distribuional insights

Common Stocks Only The distribution of returns across horizons provides further evidence on the differences between the Robinhood portfolios and the benchmarks.

At short horizons (5-30 days), all portfolios are tightly centered around zero, with relatively similar dispersion. Both Robinhood portfolios display slightly higher volatility than the S&P 500 and the World ETF, with standard deviations of approximately 0.045 for five-day returns, compared to about 0.03 for the benchmarks. However, mean returns remain close to zero for all portfolios, and short-term behavior shows limited divergence across methods.

At the 120-day horizon, differences become more evident. While Fedyk's portfolio maintains a positive mean log return of approximately 0.05, my price-based portfolio exhibits a negative mean return of about -0.06. The median is also much worse, at about -0.14 versus -0.02. This shift is reflected in the distribution shapes, with my portfolio showing a wider left tail and greater dispersion relative to both Fedyk's method and the benchmarks.

At the 563-day horizon, the gap further widens. The price-based portfolio remains centered around negative returns, with a mean log return of approximately -0.04, while the rebalanced portfolio achieves a positive mean return of 0.17. In contrast, the S&P 500 maintains a strong positive outcome, with a mean log return close to 0.1.

Despite the strong rebound observed after the COVID-19 crash, the cumulative performance of the price-based portfolio remains significantly weaker, often with higher standard deviation.

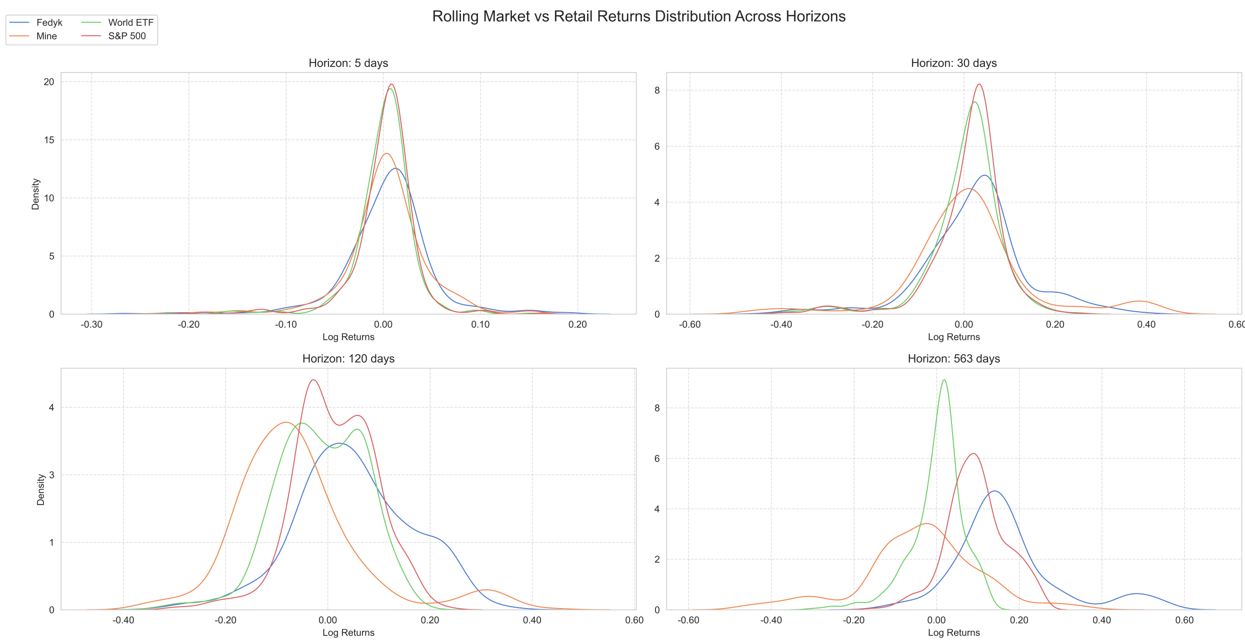


Figure 4: Distribution of rolling log-returns for retail and benchmark portfolios across investment horizons (stock-only sample)

A detailed description of the distributions can be found in table 1.

Full Universe of Securities Expanding the analysis to include all types of securities slightly alters the distributional patterns compared to the stock-only case.

At short horizons (5-30 days), return distributions remain centered around zero and maintain a similar dispersion across all portfolios. Standard deviations are comparable to the stock-only case, around 0.02 for both Robinhood portfolios at the daily level and 0.04 for five-day returns. Mean returns are very close to zero, confirming that short-term dynamics are largely unaffected by the broader sample.

At the 120-day horizon, however, differences become more evident. Both Robinhood portfolios exhibit flatter distributions with thicker left tails compared to the benchmarks. In particular, my price-based portfolio records a negative mean log return of approximately -0.07, while Fedyk's method also turns slightly negative at around -0.004, unlike in the stock-only case where it remained positive. The S&P 500 maintains a positive mean return over the same horizon, reinforcing the relative weakness of retail portfolios when more asset types are included.

At the cumulative horizon, the gap becomes substantial. My portfolio shows a strongly left-skewed distribution with a mean log return of approximately -0.10. Fedyk's portfolio, although positive at around 0.087, remains below the S&P 500, which achieves a mean return close to 0.1. The World ETF again displays lower returns, in line with its broader exposure to international markets.

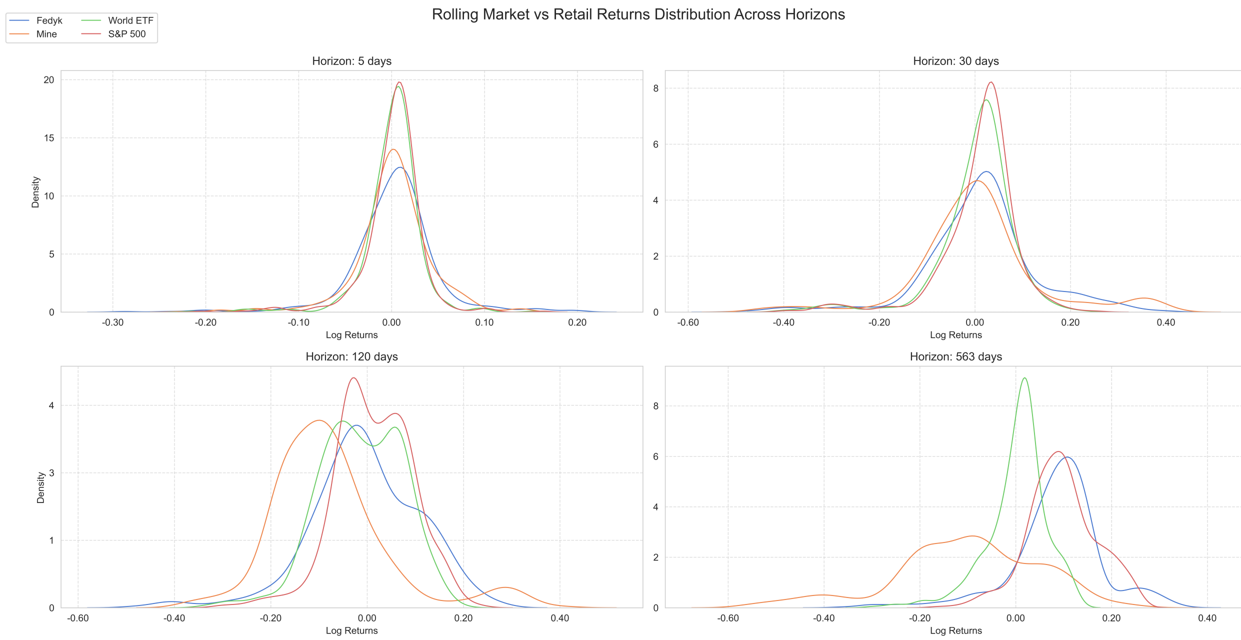


Figure 5: Distribution of rolling log-returns for retail and benchmark portfolios across investment horizons (complete sample)

Descriptive statistics supporting these findings are reported in Table 2.

4.3 Analysis of Short-Term Returns and the Impact of COVID

Comparing the daily returns for the whole period available we can already observe distinct behavior in terms of returns and distributions for the market and Robinhood indices. To test whether there are statistically significant differences between the mean daily returns of the Robinhood portfolios, the S&P 500 and the world ETF, I conduct an ANOVA test. The null hypothesis is that all portfolios have equal mean daily returns.

4.3.1 Covering the Whole Period

As expected from the results regarding the cumulative returns, mean daily returns are very similar for the Robinhood Portfolios. Fedyk's method has a mean of 0.000555, while the method based on prices has a mean of 0.000450. The ANOVA analysis conducted on these two portfolios, equal to a t-test in this case, has a p-value of 0.9273. This suggests that at the daily frequency the expected growth rate, i.e. the mean log returns, of the Robinhood portfolios are not statistically different.

In terms of standard deviation, the Robinhood portfolios have relatively larger values, around 0.02, while the market indices have values around 0.15. This is consistent with what can be seen from the distributions: retail portfolios have fatter tails while modal returns appear similar across all time series.

Conducting ANOVA tests on all possible combinations of these four timeseries, none has an acceptable p-value, results are in table 4. However, conducting a Fligner test to assess the difference in variances, the Robinhood portfolios show statistically significant different variances than the market proxies, results are in table 6.

Here below the plots of the distributions, while the returns are in table 2:

Figure 6: Distribution of daily log-returns for retail and benchmark portfolios (complete sample)

4.3.2 Excluding COVID

The results on daily returns, however, are probably impacted by the noise of the March 2020 crash. I run the same analysis filtering the dataframe up to February 3rd 2020, the date in which the pandemic was declared in the US.

In this case point estimates for means differ greatly even among Robinhood Portfolios, with the price-based approach yielding -0.000328 on average versus Fedyk's 0.000246. Nonetheless, also at this level returns are not statistically significant, results can be found in table 5.

In terms of standard deviation, the market indices show again clearly a different distribution, with variances being markably lower. This is corroborated by fligner tests, available in table 7.

Here below the plots of the distributions, while the returns are in table 3:

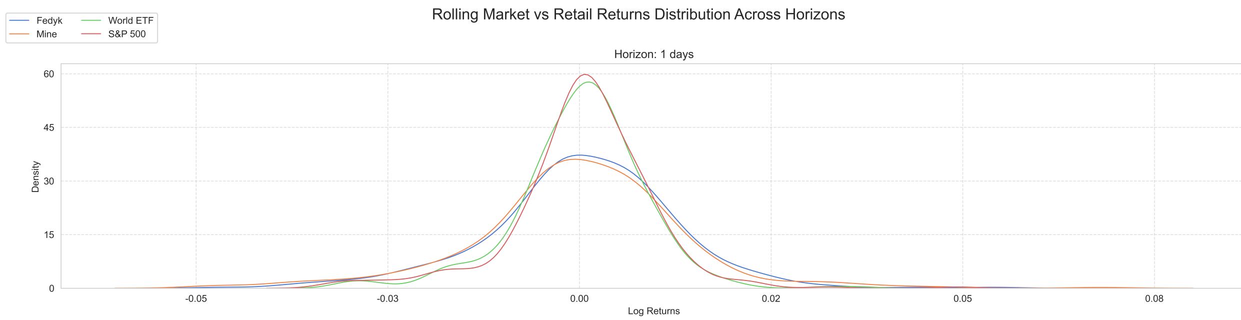


Figure 7: Distribution of daily log-returns for retail and benchmark portfolios (complete sample, excluding COVID)

4.4 Comments

Daily returns show a deceptively optimistic view, with the different series having statistically equal log-returns. However, the retail portfolios carry systematically higher day-to-day risk, and this compounds as underperformance at the cumulative level before the pandemic.

A one-way ANOVA on daily log returns fails to reject equality of means for any group of series, both including and excluding the pandemic.

However, looking at the risk dimension, standard deviations differ sharply. Fligner-Killeen tests reject homogeneity at any α level whenever a retail series is compared with a market proxy. As highlighted earlier, this can be easily seen from the shape of the distributions, with correspondingly fatter tails. In log space, larger dispersion automatically lowers long-run growth because every large swing pulls the cumulative sum away from the average trend. In other words, retail investors accept more noise than the market but earn no measurable extra drift in a single day.

Although ANOVA cannot reject equality between the Robinhood portfolios, the point estimates themselves are not identical (compare table 3). The gap in their mean returns, about -0.0006 log points, rapidly cumulates to -0.25 log points. The test lacks power because daily σ is large; the economic impact measured in compounding is evident.

Path dependence widens the difference. The price-weighted series 9 embeds a negative correlation between returns and remaining capital. When a stock falls, its contribution to the portfolio value shrinks, and therefore any subsequent rebound is applied to less capital, exactly how a normal portfolio would work. The return-weighted construction re-anchors weights every market close and therefore avoids this drag. Losses in early 2019 thus penalise the price-weighted portfolio twice.

5 Assessing Risk Preferences: A Dual-Criterion Approach

5.1 Setup and Definitions

The results of section 4 paint very different pictures of retail investors. Using the method set forth by [Welch, 2022] and [Fedyk, 2024], analyzed also in section 4.2.1, yields superior returns and similar drawdowns to the market. They also conclude that the Robinhood crowd has achieved positive alpha when analyzed under different factor models (table VII, IX, X in [Welch, 2022] and table 16 in [Fedyk, 2024]). These results appear to be in contrast with the existing literature on retail investing, most notably [Barber and Odean, 2000].

A more fundamental question, however, is whether those returns are attractive once investors' attitudes toward risk are taken into account. This section evaluates the Robinhood portfolio against its market benchmarks using two complementary criteria.

First, I adopt the constant-relative-risk-aversion (CRRA) framework, in line with the majority of asset pricing work.

$$U(W) = \begin{cases} \frac{W^{1-\gamma}-1}{1-\gamma}, \gamma \neq 1 \\ \ln(W), \gamma = 1 \end{cases} \tag{11}$$

By computing the expected utility of both the Robinhood and benchmark portfolios over a grid of possible risk aversion (γ) values, I identify the cutoff γ^* such that a representative CRRA investor is indifferent between the two. This delivers a concise, parametric summary of how risk preferences may shape portfolio choice.

However, this method cannot deliver precise estimates given the limited sample size. I therefore employ another method to directly estimate the risk aversion γ, following the Generalized Method of Moments (GMM) framework introduced by [Hansen and Singleton, 1982] to find the risk aversion that satisfies the intertemporal Euler Condition, aspresented also in [Cochrane, 2005].

$$\mathbb{E}\left[\beta \frac{U'(c_{t+1})}{U'(c_t)} R_{t+1}\right] = 1 \tag{12}$$

where R_{t+1} is the realized gross return on the asset at time $t+1$, β is discount factor, and $U'(\cdot)$ is the derivative of the utility function..

5.2 Expected Utility and Cutoff

5.2.1 Sampling Variability and Finite-Sample Limitations

Identifying a cutoff level of risk aversion provides a clear criterion for the CRRA utility-maximizing investor: it is the minimum γ at which an alternative strategy becomes preferred to the Robinhood portfolio. However, the limited size and noise of the sample imply wide conference intervals for the estimated expected utilities at different γ levels. In practice, this remains a useful conceptual framework to understand how risk preference and beliefs may affect portfolio choice, but in limited samples, its numerical outputs are more illustrative than definitive. However, this method yields precise results in a very specific case I will describe later.

Formally, we want to find γ^* defined as:

$$\gamma^* = \min\left\{\gamma_j : \mathbb{E}[U_p(\gamma_j)] \leq \mathbb{E}[U_m(\gamma_j)]\right\} \tag{13}$$

where $U_p(\cdot)$ is the utility of the Robinhood portfolio, while $U_m(\cdot)$ is the utility of the market portfolio.

In practice, we apply CRRA utility function (11) to each gross return observation and then takes the sample average. The resulting mean serves as the expected utility implied by the investor's revealed choices over the sample period. Wealth at time t is equal to gross returns, assuming the initial wealth W to be equal to one without loss of generality.

The main problem related to this approach is inherent to the sample we apply it to, having only 539 observations and inclusion of extreme events such as the COVID crash. These factors inflate the sample variability of our mean-utility estimates, so much that regardless of whether we use daily returns, various rolling-window horizons, or subperiods, the resulting confidence intervals remain prohibitively wide. Figure 8 illustrates this well:

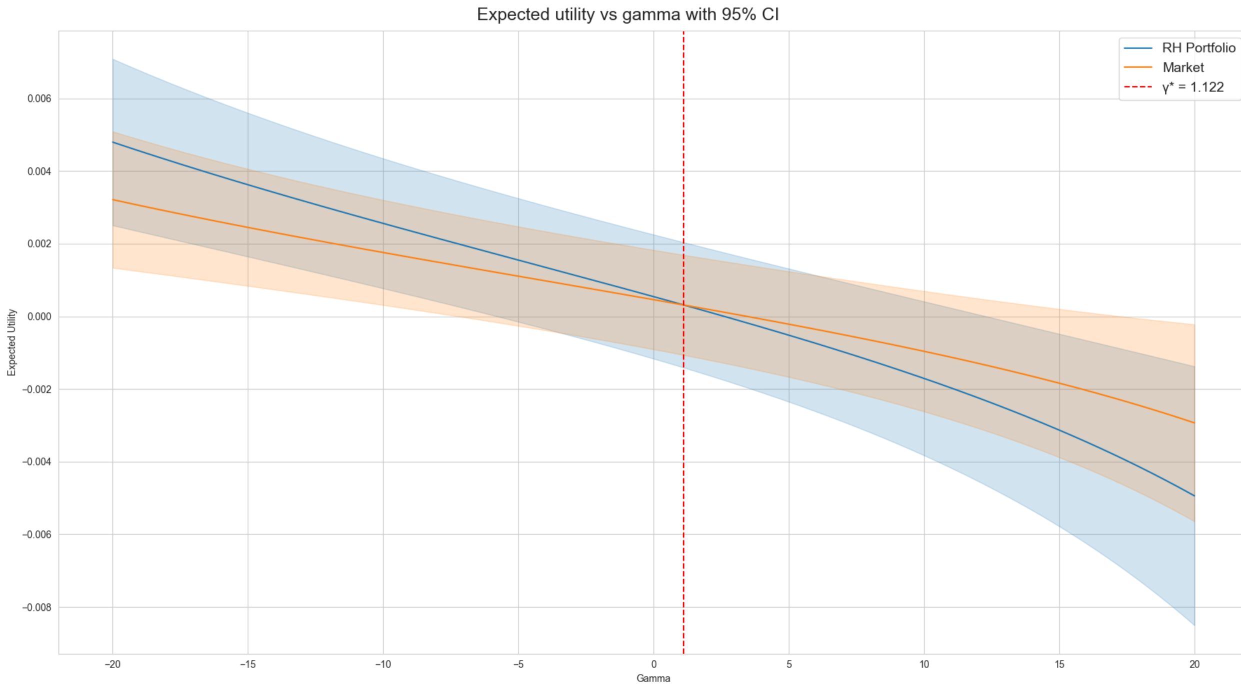


Figure 8: Utility for the Robinhood and Market portfolios evaluated over a grid of γ.

5.2.2 Augmented Return Sample via All Date-Pairs

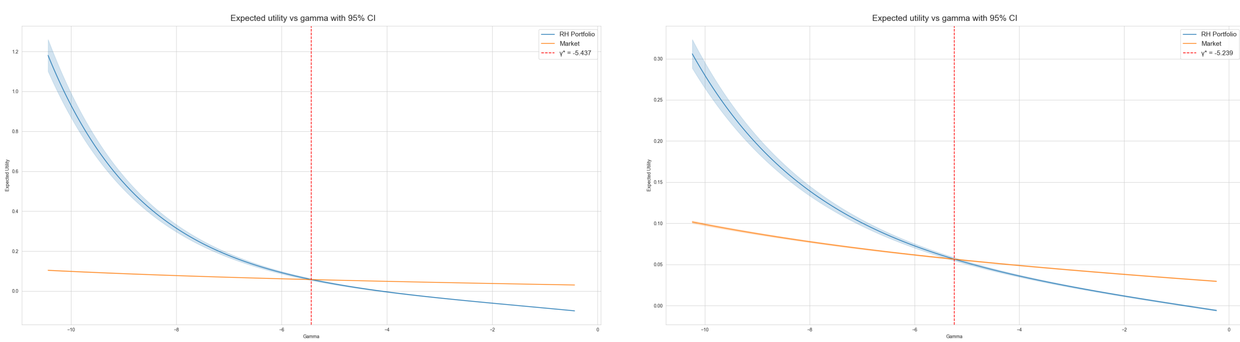
To overcome the limited information in only 539 daily observations, we construct an ”all-pairs” dataset that dramatically amplifies our effective sample. Specifically, let the trading-day indices in our original series be $1, 2, \ldots, T$. For every ordered pair of dates (i, j) with $1 \leq i < j \leq T$, we compute the cumulative excess return (net of the risk-free rate) between day i and day j as

$$R_{i,j} = \prod_{k=i+1}^{j} \Big(1 + r_k - r_{f,k}\Big), \quad 1 \leq i < j \leq T, \tag{14}$$

where r_k is the portfolio return on day k and $r_{f,k}$ is the daily risk-free rate. Using excess returns over the risk-free rate is necessary due to the changes in macroeconomic policy during the time of the sample. We treat each multi-day return $R_{i,j}$ as a separate outcome in the investor’s distribution of possible holding-period returns, we increase the number of observations from T to $T(T-1)/2$, which reduces sampling variability in our utility-based estimates. This ”all-pairs” approach preserves the time-ordering of returns while allowing us to evaluate expected utility cutoffs with far greater precision.

In practice, this results in finding a very low γ^* that satisfies equation 13, meaning that every investor with a reasonable risk-aversion parameter under CRRA utility would have a greater utility by investing in the market[7]. In some cases, particularly when ending the sample before the pandemic, the utility curves fitted on the "all-pairs" dataset simply do not intersect for a wide grid of possible risk aversion, and in fact we find also that in these cases the market stochastically dominates the robinhood portfolio (something which will be analyzed more in depth later).

Performing this exercise for the whole period on the two alternative Robinhood portfolios yields very telling results. In both instances the cutoff risk aversion as defined in 13 is negative; in particular, the portfolio I constructed has a cutoff risk aversion of -5.437 while the one built using the alternative method we obtain -5.239. The plot below shows this finding, only extremely risk loving investors would prefer the Robinhood portfolios over a diversified index.



(a) Robinhood Returns Built from Prices

(b) Robinhood Returns Built Following Fedyk

Figure 9: Expected Utility and Cutoff Risk-aversion for the Robinhood and Market Portfolio.

The only case in which γ^* is positive is when we compare the Fedyk Portfolio and the World equity ETF as a proxy for the market index, obtaining a value of 1.214. This implies that weakly risk-averse and risk-loving investors would have a greater utility by investing in this index. However, if we limit the analysis to the pre-COVID period, the utility of the World ETF strictly dominates for each $\gamma \geq -15.813$, highlighting that the former result might indeed be a result of the volatility experienced in early 2020.

[7]Both the risk-free rate and market returns are downloaded from Kenneth R. French data library https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html

5.3 Euler Condition Approach to CRRA Parameter Estimation

5.3.1 Bootstrap Methodology for GMM-Based γ Estimation

To address the limitations of our cutoff γ approach, we turn to the canonical Euler-equation estimator of risk aversion as presented in Chapter 1 of [Cochrane, 2005]. In that framework, the agent's first-order condition for optimal intertemporal consumption and portfolio choice requires 12 to hold.

Rewriting equation 12 and assuming CRRA utility we arrive at the following condition[8]:

$$\mathbb{E}\left[\beta\left(\frac{c_{t+1}}{c_t}\right)^{-\gamma} R_{t+1}\right] = 1 \tag{15}$$

Then, using gross portfolio returns as a proxy of consumption growth and letting $\beta = \frac{1}{1+\bar{r_f}}$ we derive the following GMM moment condition:

$$g(\gamma) = \mathbb{E}\left[\frac{R_{t+1}^{-\gamma}}{1+\bar{r}_f} R_{t+1} - 1\right] = 0 \tag{16}$$

I then use a root-solving algorithm to find the $\hat{\gamma}$ that satisfies equation 16.

Bootstrap Confidence Intervals. The GMM estimator described above yields a single point estimate $\hat{\gamma}$, but does not indicate its sampling precision. To quantify uncertainty, we apply a nonparametric bootstrap: we repeatedly resample our T return observations (with replacement), re-solve the GMM moment condition in each replicate, and record the resulting $\hat{\gamma}^{(b)}$ values. The empirical distribution of $\{\hat{\gamma}^{(b)}\}_{b=1}^{B}$ then provides a straightforward 95% confidence interval for γ. This approach requires no additional distributional assumptions and directly reflects the finite-sample variability of our risk-aversion estimate.

5.3.2 Empirical Estimates and Interpretation

Before describing the results a necessary premise has to be stated. A key condition for the validity of the nonparametric bootstrap is that the observations be independent and identically distributed (i.i.d.). In particular, resampling ”with replacement” implicitly treats each data point as a separate draw from the same underlying distribution. If the series exhibits serial dependence, as it would be if we used the rolling returns defined in earlier sections, then the standard bootstrap will underestimate sampling variability. To avoid this

[8] Derivation: $U'(x) = x^{-\gamma}$, hence $\frac{U'(c_{t+1})}{U'(c_t)} = \left(\frac{c_{t+1}}{c_t}\right)^{-\gamma}$

issue we resample the data on weekly and monthly frequencies to estimate risk aversion on different timeframes.

Table 8 reports our GMM estimates of the CRRA coefficient $\hat{\gamma}$ alongside 95% confidence intervals for both the portfolio built on prices (labelled ”Mine”) and the Fedyk benchmark. A few key observations emerge. At daily and weekly horizons, every point estimate of $\hat{\gamma}$ is not only strictly positive but also bounded away from zero by its confidence interval. This confirms that, under both specifications, the representative investor implied by the Euler-equation GMM is unequivocally risk averse over high-frequency returns.

Moving to monthly returns, the confidence bands widen substantially, now crossing zero, due to the much smaller number of non-overlapping observations. While the point estimates remain positive, their CIs reflect that, with only $\tilde{1}8$ monthly data points, we lack statistical power to reject mild risk neutrality at conventional levels.

Across all horizons, the ”Fedyk” market benchmark yields slightly higher $\hat{\gamma}$ values than our custom portfolio. This suggests that, to satisfy the Euler condition, a representative investor choosing the market must be marginally more risk averse than one choosing the alternative strategy, perhaps compensating for lower volatility or different return skewness.

The strictly positive daily and weekly estimates imply a genuine aversion to risk in the retail investor’s revealed preferences. In practical terms, these investors require compensation for variance in returns and would, under a fully rational CRRA framework, tilt away from high-variance holdings unless expected excess returns are sufficient.

These findings underscore that the Euler-GMM approach delivers clear, statistically significant evidence of risk aversion at the daily and weekly horizons—where our sample is large enough to pin down a strictly positive γ. It is precisely in these shorter-term windows that we can say with confidence that retail investors reveal genuine aversion to risk. As the horizon lengthens and observations dwindle, however, the resulting estimates become too imprecise to draw firm conclusions.

6 Stochastic Dominance Analysis and Behavioral Implications

In this final empirical exercise, we compare the full return distributions of the Robinhood portfolio and the market benchmark using nonparametric stochastic dominance tests. A second-order stochastic dominance (SSD) result in favor of the market would imply that *every* risk-averse investor (including but not restricted to all CRRA preferences with $\gamma > 0$)

strictly prefers the market's payoff to that of the Robinhood strategy.

6.1 Methodology

We first estimate the empirical cumulative distribution functions (CDFs) of daily returns for the two portfolios. Let

$$F_p(r) \;=\; \frac{1}{T}\sum_{t=1}^{T}\mathbf{1}\{r_{p,t} \le r\}, \quad F_m(r) \;=\; \frac{1}{T}\sum_{t=1}^{T}\mathbf{1}\{r_{m,t} \le r\}, \tag{17}$$

where $r_{p,t}$ and $r_{m,t}$ are the portfolio and market returns on day t.

To analyse whether the portfolio is preferred over the market we employ two different tests. First-order stochastic dominance (FSD) holds if $F_p(r) \le F_m(r)$ for all r; second-order stochastic dominance (SSD) holds if

$$G_p(r) \;\equiv\; \int_{-\infty}^{r} F_p(x)\,dx \quad \text{and} \quad G_m(r) \;\equiv\; \int_{-\infty}^{r} F_m(x)\,dx \tag{18}$$

satisfy $G_p(r) \le G_m(r)$ for all r.

6.2 Results

6.2.1 Rolling-Horizon SSD

Comparing Robinhood Portfolios Across non-overlapping horizons of 1, 5, 30, and 120 trading days—as well as the full 563-day sample—we compute for each return series the empirical CDFs $F_{\text{Mine}}(r)$ and $F_{\text{Fedyk}}(r)$, their integrals $G_{\text{Mine}}(r)$ and $G_{\text{Fedyk}}(r)$, and the fraction of evaluation points satisfying

$$G_{\text{Fedyk}}(r) \;\le\; G_{\text{Mine}}(r),$$

which indicates Fedyk's second-order stochastic dominance over the portfolio I built. This support fraction falls monotonically with the horizon: it exceeds 98 % at the 1- and 5-day horizons, drops below 1 % at 30 days, only rebounds to about 6 % at 120 days, and reaches 0 % over the full period. These findings show that Fedyk's SSD advantage becomes increasingly stronger at longer horizons, making it preferred by every risk-averse investor. This relation even becomes a first-order stochastic dominance when accounting for cumulative returns, as can be easily seen from plot.

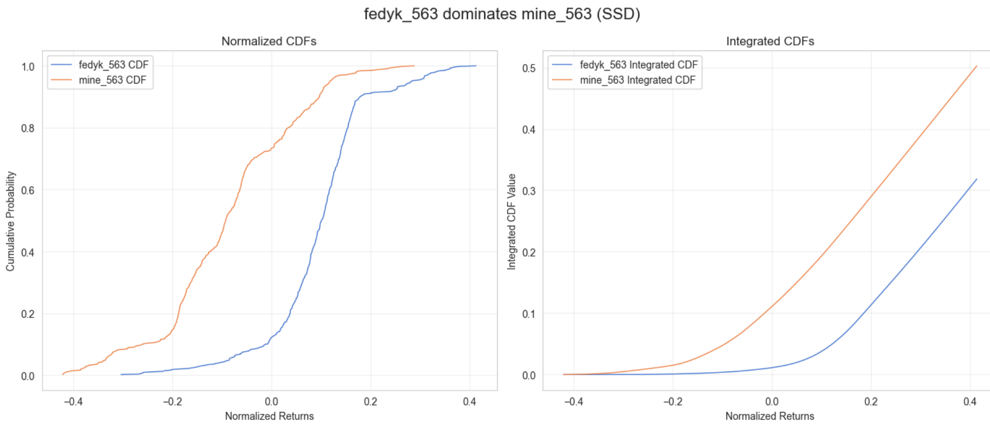


Figure 10: FSD Between the Robinhood Portfolios.

Comparing Robinhood to Broad-Market Proxies We use the Fedyk portfolio as a conservative benchmark for comparison. Its demonstrated dominance over our custom Robinhood strategy across horizons makes it an especially stringent benchmark: if our portfolio cannot overcome Fedyk's advantage, it would be even harder to outperform true market indices at longer holding periods.

Having established that the Fedyk portfolio second-order stochastically dominates our Robinhood strategy as the length of the horizon increases, we now benchmark against three widely-used market proxies: the S&P 500 ETF (VOO), the global all-world ETF (VT), and the Fama French market index (results for which are nearly identical to VOO).

At the 1- and 5-day horizons, all three market indices exhibit overwhelming SSD dominance over our Robinhood portfolio, each with over 97% of evaluation points satisfying

$$G_{\text{proxy}}(r) \leq G_{\text{RH}}(r).$$

By the 30-day horizon, this support fraction declines to approximately 92% for the S&P 500 and 90% for the all-world ETF, while the Fama French index mirrors the S&P 500's trajectory. Extending to 120 days, the S&P 500 retains complete dominance (100% of points), whereas the all-world ETF's dominance share falls to about 83%. Over the full 563-day sample, the S&P 500 continues to second-order dominate Robinhood, while the all-world ETF shows only weak support ($\approx$ 8%). Only the S&P 500 (VOO) and the Fama French market index preserve dominance in the long-run payoff distribution.

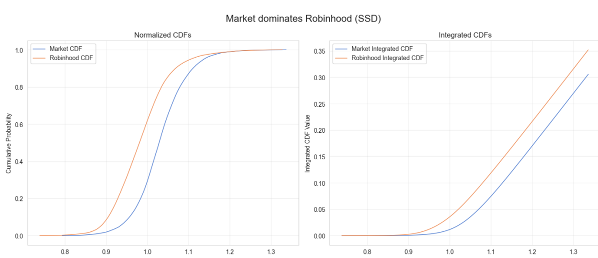
In other words, whether we compare the portfolio I built against the Fedyk portfolio or traditional market indices, the second-order stochastic dominance over our Robinhood strategy is nearly universal at intraday and weekly horizons. Our approach to building the Robinhood portfolio, fails SSD also against the World ETF at longer horizons.

6.2.2 All-Pairs Analysis

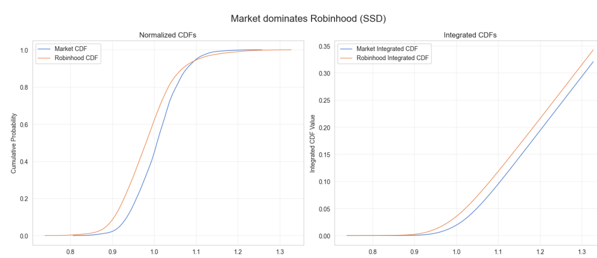
Our analysis now turns to the all-pairs dataset constructed according to Equation 14, which computes returns across all possible entry-exit point combinations. As previously established in Section 5.2.2, CRRA preference analysis indicates that only investors with extreme risk-seeking preferences would derive greater utility from the Robinhood portfolio compared to market indices.

These conclusions are strongly supported by stochastic dominance tests, with some cases demonstrating first-order dominance. When evaluating Fedyk's Robinhood portfolio against market proxies, we find the S&P 500 and Fama French market index exhibit clear second-order stochastic dominance. The world ETF (VT) achieves dominance over 96% of evaluation points, falling short of full dominance by only a marginal 4%. These results hold more strongly when using our alternative Robinhood portfolio construction method, where all three market proxies demonstrate complete second-order dominance.

The pre-pandemic subsample (through February 2020) reveals even more pronounced results in favor of market indices. Fedyk's portfolio fails second-order dominance under all market proxies during this period. Notably, the Fama French index achieves first-order stochastic dominance, while the S&P 500 and world ETF narrowly miss complete first-order dominance by merely 0.17% and 5.92% of evaluation points respectively.



(a) S&P 500



(b) World ETF

Figure 11: Second-order stochastic dominance tests comparing Fedyk's Robinhood portfolio to market indices (pre-pandemic period through February 3, 2020).

Our alternative Robinhood portfolio construction shows even stronger results, demonstrating first-order dominance across all three market proxies in this period. The economic implications of these findings are particularly striking. For any Robinhood investor selecting entry and exit points at random, the revealed preferences would appear economically irrational - the Robinhood portfolio either suffers first-order stochastic dominance or comes remarkably close to this threshold. The second-order dominance results further indicate that all risk-averse investors, regardless of their specific utility function form, would have been better served by market indices. These results hold particular significance when considering the pre-pandemic period, where market dominance appears most pronounced.

6.2.3 A Proof of Irrationality?

Our analysis reveals a fundamental contradiction in Robinhood investors' revealed preferences. Building on the risk aversion estimates from Section 5.3.2 (Table 8), where we found strictly positive γ values in 95% confidence intervals for daily and weekly returns, we now confront these findings with stochastic dominance tests. This creates an irreconcilable tension: while Robinhood investors exhibit risk-averse preferences through positive γ estimates, their chosen portfolios fail to satisfy basic rationality requirements when compared to market alternatives.

The contradiction emerges most sharply at weekly horizons. Despite estimated risk aversion parameters ($\hat{\gamma} > 0$) that should favor market indices under second-order stochastic dominance (SSD), we find that the Fama French market index demonstrates unambiguous SSD over both Robinhood portfolios, and S&P 500 and world ETF (VT) achieve SSD over 95-99% of evaluation points. This implies that even the most conservative specification of risk-averse preferences would favor market indices over Robinhood strategies, directly contradicting the investors' actual portfolio choices.

The tension persists across all horizons: Daily returns show 95-98% SSD support for market dominance, monthly returns reveal comprehensive market dominance (90.7-100% SSD support), and only a small fraction of monthly $\hat{\gamma}$ estimates dip below zero.

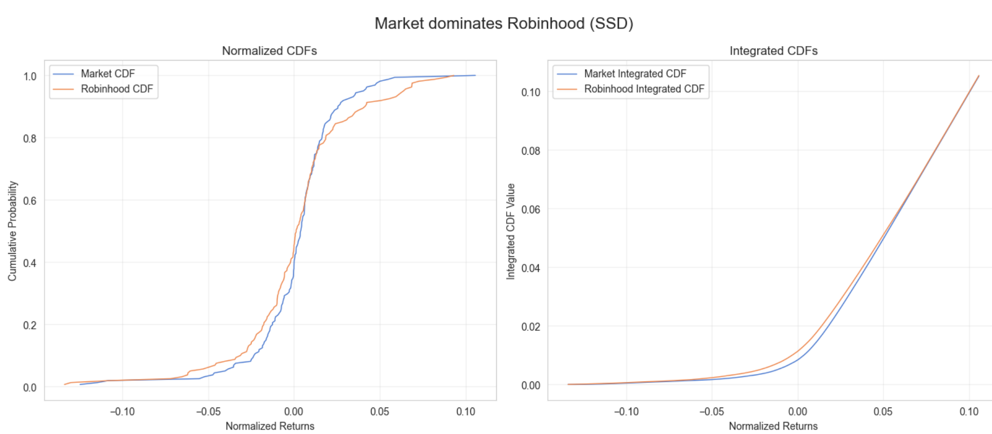


Figure 12: SSD of the Market over the Robinhood Portfolio.

These results present an inescapable conclusion: Robinhood investors' portfolio choices cannot be rationalized through standard risk-return tradeoffs. The combination of (1) positive estimated risk aversion and (2) market dominance across SSD tests creates a paradox that only resolves if we reject the premise of fully rational investors. Even under generous assumptions about return aggregation and bootstrap resampling (as detailed in Section 5.3.2), the evidence consistently challenges the rationality hypothesis.

Our findings suggest that non-standard preferences or behavioral factors, such as lottery-seeking behavior, attention-driven trading, or misunderstanding of statistical dominance, must drive these investment decisions. The persistence of this contradiction across multiple methodologies (CRRA estimation, SSD tests, and temporal subsamples) makes it particularly robust to alternative explanations.

6.3 Implications for Investor Behavior

Our nonparametric stochastic-dominance tests deliver a clear verdict: across multiple return horizons and market proxies, the broad market consistently second-order stochastically dominates the Robinhood strategy for short-term horizons, and at monthly and longer horizons the S&P 500 (VOO) in particular preserves that dominance where other proxies fade. Because SSD implies preference by every risk-averse investor (all CRRA utilities with $\gamma > 0$), these results extend and reinforce our earlier finding—based on Euler-GMM estimation—that retail investors exhibit positive risk aversion yet hold a portfolio that no rational, CRRA-maximizer would choose.

Taken together, the Euler-equation and stochastic-dominance evidence point to a fundamental disconnect between normative decision rules and observed retail behavior. Even though our bootstrap-corrected GMM estimates uncover strictly positive—and economically plausible—risk-aversion parameters at daily and weekly frequencies, the SSD tests show that any such risk-averse investor would uniformly prefer the market portfolio. The persistence of sub-optimal, under-diversified strategies therefore cannot be attributed to risk-aversion alone.

Instead, these findings call for behavioral explanations: investors may succumb to overconfidence in their timing ability, confirmation bias in interpreting past returns, or an illusion of control over individual trades. Such biases can sustain a self-confirming equilibrium, in which retail traders overweight episodes of success and dismiss losses as noise, reinforcing the belief that their custom portfolio offers an edge—even when the broad market outperforms on every risk-averse criterion.

Overall, our analysis underscores the dual power of parametric (Euler-GMM) and nonparametric (SSD) methods in asset-pricing research. The former quantifies the magnitude of revealed risk preferences, while the latter delivers a parameter-free, universal ranking of payoff distributions. By combining both approaches, we provide robust evidence that, despite genuine risk aversion, retail investors systematically deviate from the optimal market portfolio—highlighting the crucial role of behavioral frictions in real-world decision-making.

Appendix A Tables

Table 1: Descriptive Statistics for Daily and Rolling Returns
Note: Returns accounting only for investments in Stocks.

	count	mean	std	min	25%	50%	75%	max
Fedyk daily return	563	0.000990	0.019894	-0.135995	-0.005807	0.001516	0.009167	0.105441
Mine daily return	563	0.000677	0.019209	-0.125590	-0.006681	0.001363	0.009784	0.071628
VT daily return	563	0.000185	0.015106	-0.123763	-0.004592	0.000886	0.005933	0.087470
VOO daily return	563	0.000443	0.015820	-0.124870	-0.003893	0.000976	0.006653	0.091087
Fedyk 5 return	559	0.004866	0.046404	-0.265341	-0.015031	0.007872	0.025282	0.193965
Mine 5 return	559	0.002979	0.043110	-0.226991	-0.013770	0.003621	0.021740	0.159707
VT 5 return	559	0.000824	0.030805	-0.214262	-0.011045	0.003956	0.014927	0.151788
VOO 5 return	559	0.002111	0.031063	-0.204425	-0.009442	0.005716	0.016506	0.162820
Fedyk 30 return	534	0.025268	0.114955	-0.431533	-0.032448	0.035874	0.072017	0.401292
Mine 30 return	534	0.012159	0.140314	-0.474366	-0.051104	0.009054	0.054639	0.431367
VT 30 return	534	0.002735	0.081324	-0.406688	-0.024941	0.017664	0.041113	0.224464
VOO 30 return	534	0.009843	0.079924	-0.401950	-0.015450	0.027341	0.046977	0.252864
Fedyk 120 return	444	0.050804	0.114293	-0.330870	-0.021276	0.046242	0.130711	0.357556
Mine 120 return	444	-0.059004	0.127178	-0.358205	-0.139490	-0.074974	-0.015135	0.439583
VT 120 return	444	-0.012628	0.085645	-0.333294	-0.070621	-0.012804	0.059034	0.186378
VOO 120 return	444	0.013786	0.079450	-0.302877	-0.037033	0.011899	0.073431	0.231377
Fedyk 563 return	563	0.168894	0.129435	-0.148334	0.100301	0.146221	0.199060	0.565657
Mine 563 return	563	-0.038567	0.143947	-0.468443	-0.114762	-0.027966	0.041386	0.380918
VT 563 return	563	0.002293	0.063753	-0.300675	-0.023998	0.014837	0.032750	0.122378
VOO 563 return	563	0.096803	0.071714	-0.166225	0.054575	0.093080	0.136742	0.253868

Table 2: Descriptive Statistics for Daily and Rolling Returns
Note: Returns accounting for investments in all types of securities. ETF returns are identical to table 1.

	count	**mean**	**std**	**min**	**25%**	**50%**	**75%**	**max**
Fedyk daily returns	563	0.000555	0.020000	-0.125369	-0.006658	0.000754	0.008755	0.097995
Mine daily returns	563	0.000450	0.018710	-0.126475	-0.006472	0.000606	0.009044	0.072808
VT daily return	563	0.000185	0.015106	-0.123763	-0.004592	0.000886	0.005933	0.087470
VOO daily return	563	0.000443	0.015820	-0.124870	-0.003893	0.000976	0.006653	0.091087
Fedyk 5 return	559	0.002684	0.047496	-0.287100	-0.017592	0.005106	0.022556	0.201768
Mine 5 return	559	0.001891	0.041707	-0.224346	-0.015152	0.002165	0.020423	0.145245
VT 5 return	559	0.000824	0.030805	-0.214262	-0.011045	0.003956	0.014927	0.151788
VOO 5 return	559	0.002111	0.031063	-0.204425	-0.009442	0.005716	0.016506	0.162820
Fedyk 30 return	534	0.012231	0.118341	-0.491574	-0.044795	0.018809	0.055543	0.391200
Mine 30 return	534	0.005924	0.134409	-0.456420	-0.056865	0.006054	0.044482	0.399350
VT 30 return	534	0.002735	0.081324	-0.406688	-0.024941	0.017664	0.041113	0.224464
VOO 30 return	534	0.009843	0.079924	-0.401950	-0.015450	0.027341	0.046977	0.252864
Fedyk 120 return	444	-0.004432	0.114878	-0.478685	-0.062876	-0.010586	0.076930	0.271451
Mine 120 return	444	-0.076568	0.124668	-0.368195	-0.159984	-0.095393	-0.037126	0.406700
VT 120 return	444	-0.012628	0.085645	-0.333294	-0.070621	-0.012804	0.059034	0.186378
VOO 120 return	444	0.013786	0.079450	-0.302877	-0.037033	0.011899	0.073431	0.231377
Fedyk 563 return	563	0.086945	0.096844	-0.361597	0.050499	0.096675	0.134919	0.345822
Mine 563 return	563	-0.103880	0.152561	-0.547158	-0.193259	-0.096858	0.008830	0.253266
VT 563 return	563	0.002293	0.063753	-0.300675	-0.023998	0.014837	0.032750	0.122378
VOO 563 return	563	0.096803	0.071714	-0.166225	0.054575	0.093080	0.136742	0.253868

Table 3: Descriptive Statistics for Daily Returns, up to February 3rd 2020.
Note: Returns accounting for investments in all types of securities.

	count	**mean**	**std**	**min**	**25%**	**50%**	**75%**	**max**
Mine	429	-0.000328	0.013581	-0.049700	-0.006121	-0.000106	0.007070	0.065745
Fedyk	429	0.000246	0.012218	-0.047821	-0.005735	0.000549	0.007639	0.052644
VT	429	0.000200	0.008370	-0.031068	-0.003812	0.000766	0.004879	0.036545
VOO	429	0.000492	0.008938	-0.032828	-0.003072	0.000783	0.005103	0.049350

Table 4: Results for One-Way ANOVA Tests on Daily Returns
This results cover the whole period.

subset	**F Statistic**	**p-value**
(Fedyk, VT)	0.123106	0.725755
(VOO, VT)	0.078475	0.779426
(Mine, VT)	0.068498	0.793584
(Fedyk, VOO)	0.010935	0.916736
(Fedyk, Mine)	0.008335	0.927275
(Fedyk, VOO, VT)	0.069435	0.932923
(Fedyk, Mine, VT)	0.062966	0.938977
(Mine, VOO, VT)	0.046579	0.954490
(Fedyk, Mine, VOO, VT)	0.045550	0.987099
(Fedyk, Mine, VOO)	0.006693	0.993330
(Mine, VOO)	0.000046	0.994589

Table 5: Results for One-Way ANOVA Tests on Daily Returns, up to February 3rd 2020

subset	**F Statistic**	**p-value**
(Mine, VOO)	0.861444	0.353597
(Fedyk, Mine)	0.398907	0.527823
(Mine, VT)	0.321035	0.571135
(Mine, VOO, VT)	0.517251	0.596281
(VOO, VT)	0.243873	0.621550
(Fedyk, Mine, VOO)	0.451564	0.636733
(Fedyk, Mine, VT)	0.273712	0.760596
(Fedyk, Mine, VOO, VT)	0.343216	0.794081
(Fedyk, VOO)	0.055235	0.814249
(Fedyk, VT)	0.028370	0.866282
(Fedyk, VOO, VT)	0.092718	0.911456

Table 6: Results for Fligner Tests on Daily Returns
Results cover the whole period.

subset	F Statistic	p-value
(Fedyk, Mine, VOO, VT)	73.410013	0.000000
(Mine, VOO, VT)	52.796594	0.000000
(Fedyk, VOO, VT)	49.720452	0.000000
(Fedyk, Mine, VT)	47.695216	0.000000
(Fedyk, Mine, VOO)	46.156948	0.000000
(Mine, VT)	38.480331	0.000000
(Mine, VOO)	37.127424	0.000000
(Fedyk, VT)	36.387775	0.000000
(Fedyk, VOO)	34.874042	0.000000
(Fedyk, Mine)	0.031996	0.858037
(VOO, VT)	0.003157	0.955192

Table 7: Results for Fligner Tests on Daily Returns, up to February 3rd 2020

subset	F Statistic	p-value
(Fedyk, Mine, VOO, VT)	83.536860	0.000000
(Mine, VOO, VT)	64.187897	0.000000
(Fedyk, VOO, VT)	54.358126	0.000000
(Fedyk, Mine, VT)	53.377374	0.000000
(Fedyk, Mine, VOO)	51.856316	0.000000
(Mine, VT)	45.431002	0.000000
(Mine, VOO)	43.773368	0.000000
(Fedyk, VT)	39.157744	0.000000
(Fedyk, VOO)	37.465317	0.000000
(Fedyk, Mine)	0.576961	0.447506
(VOO, VT)	0.002907	0.957003

Table 8: Confidence Interval and Point Estimates for $\hat{\gamma}$ covering the whole period

Horizon	$\hat{\gamma}$ **(Mine)** [95% CI]	$\hat{\gamma}$ **(Fedyk)** [95% CI]
Daily	2.875 [0.424, 20.026]	3.071 [0.896, 20.745]
Weekly	2.908 [0.703, 24.608]	3.101 [0.651, 26.497]
Monthly	2.51 [-0.518, 29.888]	3.693 [-0.763, 33.333]

Appendix B Handling Missing Data

The original Robinhood dataset contains missing values for 3,331 securities, primarily in the earlier periods. This means that these securities don't have information for a certain date.

To ensure consistency we adopt a similar method as [Fedyk, 2024]. Their Robinhood portfolio is constructed using the available securities on a daily basis, hence securities with missing values are simply not taken into account for the day. Moreover we drop all securities that they have defined as problematic in the appendix.

Since our CRSP dataset is also a bit different from the one they use, we drop entirely securities that have more than one entry per day.

References

[Barber and Odean, 2000] Barber, B. M. and Odean, T. (2000). Trading is hazardous to your wealth: The common stock investment performance of individual investors. *The Journal of Finance*, 55(2):773–806.

[Cochrane, 2005] Cochrane, J. H. (2005). *Asset Pricing*. Princeton University Press, Princeton, NJ, revised edition edition.

[Fedyk, 2024] Fedyk, V. (2024). This time is different: Investing in the age of robinhood. Working paper, Arizona State University. Version dated November 2024. Check author's website or SSRN for latest version.

[Hansen and Singleton, 1982] Hansen, L. P. and Singleton, K. J. (1982). Generalized instrumental variables estimation of nonlinear rational expectations models. *Econometrica*, 50(5):1269–1286.

[Welch, 2022] Welch, I. (2022). The wisdom of the robinhood crowd. *Journal of Finance*, 77(3):1489–1527.