

Gorzon, David; von Nitzsch, Rüdiger

Article — Published Version

Behavioral performance attribution of retail investors' portfolio returns

Financial Markets and Portfolio Management

Suggested Citation: Gorzon, David; von Nitzsch, Rüdiger (2025) : Behavioral performance attribution of retail investors' portfolio returns, Financial Markets and Portfolio Management, ISSN 2373-8529, Springer Berlin Heidelberg, Berlin/Heidelberg, Vol. 40, Iss. 2, pp. 213-244, <https://doi.org/10.1007/s11408-025-00485-6>

This Version is available at:

<https://hdl.handle.net/10419/342626>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



<http://creativecommons.org/licenses/by/4.0/>



Behavioral performance attribution of retail investors' portfolio returns

David Gorzon¹ · Rüdiger von Nitzsch¹

Accepted: 6 July 2025 / Published online: 11 August 2025
© The Author(s) 2025

Abstract

Understanding the impact of cognitive biases on retail investor performance remains a critical challenge in behavioral finance. This study introduces the Behavioral Performance Attribution framework, a novel methodology for decomposing portfolio returns based on investor biases, providing an alternative to traditional return attribution models such as the Brinson, Hood, and Beebower (BHB) model. Using a large real-world trading dataset, we apply ordinary least squares (OLS) regression to quantify the explanatory power of behavioral biases across different investor groups. The empirical results demonstrate that biases such as Action Bias and Portfolio Concentration Bias significantly correlate with returns, with the Model Explainability Ratio (MER) ranging between 43.44% and 63.54%, confirming the framework's practical applicability. Notably, we estimate separate models for different investor subgroups, revealing that bias effects vary between investors outperforming and underperforming a benchmark. While excessive trading and portfolio concentration enhance outperformance among extreme outperformers, these same behaviors amplify losses among underperforming investors. Residual analysis highlights systematic over- and underestimation patterns, suggesting the need for further refinements incorporating additional behavioral and market-driven factors. These findings contribute to the growing literature on behavioral return attribution and offer a structured approach for integrating psychological factors into investment performance analysis.

Keywords Behavioral finance · Performance attribution · Household finance · Cognitive bias · Factor models · Investment behavior · Portfolio management

✉ David Gorzon
david.gorzon@rwth-aachen.de

Rüdiger von Nitzsch
nitzsch@efi.rwth-aachen.de

¹ RWTH Aachen University, Aachen, Germany

JEL Classification G41 (Role and effects of psychological, emotional, social, and cognitive factors on decision making in financial markets) · G11 (Portfolio choice and investment decisions) · G53 (Financial literacy) · C38 (Factor models)

1 Introduction

Understanding the drivers of retail investors' portfolio returns has long been a complex challenge in behavioral finance. Retail investors, unlike institutional investors, are influenced by a range of cognitive and emotional biases, which significantly impact their decision-making and ultimately their portfolio performance. The interplay of these biases has been extensively documented in the literature, with studies highlighting their effects on overtrading, portfolio concentration, and other detrimental behaviors (e.g., Barber and Odean 2000; Goetzmann and Kumar 2008). While these biases have been identified as performance drivers, the precise extent to which they holistically contribute to individual investors' achieved returns remains underexplored.

Building on prior work (Gorzon et al. 2024; Gorzon and Nitzsch 2025), this paper introduces the Behavioral Performance Attribution framework, a novel methodology for decomposing retail investors' achieved excess portfolio returns compared to a benchmark into contributions from specific behavioral biases. Unlike traditional return attribution models such as the Brinson, Hood, and Beebower (BHB) model (Brinson et al. 1986), which assumes rational decision-making and is primarily applicable to institutional investors, the Behavioral Performance Attribution framework accounts for the broader and thus more heterogeneous group of investors. By quantifying cognitive biases and their relationship with investment performance, this approach provides a structured way to assess behavioral tendencies across different investor groups, particularly successful (those outperforming the benchmark) versus unsuccessful (those underperforming the benchmark) investors.

Using real-world trading data, we apply ordinary least squares (OLS) regression models to estimate the extent to which behavioral biases explain achieved excess returns. The Model Explainability Ratio (MER) provides a measure of how well the framework captures systematic behavioral influences, validating its effectiveness. Additionally, we analyze residual returns—the portion of performance unexplained by behavioral biases—to assess whether they differ meaningfully across investor groups and whether other market-related or learning effects contribute to investment outcomes.

This study makes several contributions. First, it offers a practical framework for return attribution that integrates behavioral factors, addressing a critical gap in the literature. Second, by demonstrating how bias attribution varies across different investor subgroups, it provides a more tailored approach to understanding investor decision-making. Third, the findings have implications for financial advisory practices, highlighting actionable strategies to mitigate costly behavioral tendencies. Finally, the study suggests directions for future research by identifying areas where behavioral attribution models could be expanded to include additional market dynamics and investor learning effects.

To date, there is no holistic behavioral attribution model that comprehensively decomposes portfolio returns while accounting for the full range of cognitive bias

factors (Table 1). This is due to the inherent complexity of investor behavior, the dynamic and interactive nature of biases, and significant data limitations. Our study presents a first pragmatic step toward such a framework, one that is implementable with transactional data alone and therefore scalable for use by brokerages, platforms, and researchers. While the model has limitations, it offers a practical foundation for quantifying behavioral tendencies and guiding further refinement of bias-aware performance analytics.

The remainder of this paper is structured as follows. Section 2 details the methodology, including the dataset, regression models, and attribution approach. Section 3 presents the results, emphasizing both group-level and individual-level findings. Section 4 discusses the broader implications of the findings, and Sect. 5 concludes by summarizing the study's contributions and outlining future research directions.

2 Methodology

2.1 Data and sample

This study utilizes a dataset of 20,000 German retail investors, spanning a six-year period from 2018 to 2023. The dataset includes monthly trading records, portfolio returns, trade volumes, asset allocations, and investor-specific cognitive bias measures. The data originates from a major German brokerage platform, ensuring a broad representation of individual investors across various demographic and behavioral profiles.

As a result of data privacy constraints, the dataset does not include demographic variables, such as income, education, or psychometric attributes. This ensures compliance with privacy regulations while focusing the model exclusively on behavioral patterns observable from trading data.

2.1.1 Performance group segmentation

To examine the differential relationships of behavioral biases and investment performance, investors are categorized into two distinct performance groups:

- **Winning Investors:** Defined as those whose annual average portfolio returns exceed the performance of the adjusted MSCI World Index (excluding dividends)¹ over the six-year period. This group constitutes approximately 25% of the dataset.

¹ The dataset has a limitation: data from the settlement account is absent. This means that dividend payments are not captured in the dataset. While this omission of dividend payments does not affect the analysis of trading behavior, it does mean that total portfolio returns may be understated, as dividends can be a significant component of investment returns. Therefore, we used an adjusted MSCI World Index excl. dividend payment as the benchmark (8.09% average return p.a.). However, as a robustness check, we also conducted the analysis for further benchmark in our prior study 1) MSCI World Index (Dist) including dividend payments (9.63% average return p.a.), and 2) the MSCI World Price Index (8.54% average return p.a.). The results were robust with regards to the MSCI World Index benchmark selection. (Gorzon and Nitzsch 2025).

Table 1 Examples of behavioral performance attribution for two winning and two losing investors. Shows individual bias contributions, total bias-attributed return, actual excess return, and resulting Model Explainability Ratio (MER). Behavioral Performance Attribution Examples using Eq. (3) $\alpha_{attrib,i} = \sum_{j=1}^N \beta_{j,i} B_{j,i}$

Behavioral Performance Attribution	Winning Investor #1	Winning Investor #2	Losing Investor #3	Losing Investor #4
Action Bias	0.19%	3.91%	-0.07%	-17.95%
Narrow Framing Bias	1.79%	2.95%	-1.98%	-0.92%
Sensation Seeking Bias	0.01%	0.00%	-0.43%	-2.42%
Extrapolation Bias	0.01%	0.29%	-0.04%	-0.07%
Familiarity Bias	-1.61%	-3.51%	-4.38%	-2.74%
Portfolio Concentration Bias	3.27%	1.95%	-1.08%	-0.49%
Sum of Behavioral Performance Attribution	3.67%	5.60%	-7.98%	-24.59%
Actual Performance of Investor	5.38%	4.06%	-5.48%	-28.93%
Model Explainability Ratio	68.16%	72.54%	68.64%	85.01%

- **Losing Investors:** Defined as those whose annual average portfolio returns underperform the adjusted MSCI World Index. This group accounts for the remaining 75% of the dataset.

This segmentation approach was validated in the study by Gorzon and Nitzsch (2025), which demonstrated significant differences in the correlations of behavioral biases and achieved returns across performance groups. By differentiating between successful and unsuccessful investors, we aim to assess whether bias-driven return attributions vary systematically between these groups. Rather than implying a strict dichotomy, this validated segmentation reflects that a one-size-fits-all model may not sufficiently capture the heterogeneity of investor behavior. Therefore, we estimate separate models for each performance group to better account for structural differences in behavioral patterns and their association with investment outcomes.

While our return metric excludes dividends, tax adjustments, and formal risk adjustment, this design choice reflects a practical modeling focus. Many common risk-adjusted models (e.g., Sharpe, CAPM, Fama–French) rely on assumptions such as rationality and diversification that may not hold for retail investors. Our aim was to maintain transparency and replicability using only transaction-level data, avoiding additional assumptions that could confound or dilute the observable impact of behavioral patterns.

2.1.2 Behavioral bias factors

The dataset includes six cognitive bias measures, identified in previous research (e.g., Barber and Odean 2000; Goetzmann and Kumar 2008) and validated in Gorzon (2024) (Table 2).

- **Action Bias:** Tendency to trade excessively, even in the absence of informational advantages.
- **Narrow Framing Bias:** Evaluating investments in isolation rather than as part of a portfolio.
- **Sensation Seeking Bias:** Engaging in high-risk trades due to preference for excitement.
- **Extrapolation Bias:** Overweighting recent market trends when making investment decisions.
- **Familiarity Bias:** Preference for investing in well-known or local securities.
- **Portfolio Concentration Bias:** Holding an undiversified portfolio, leading to increased idiosyncratic risk.
- **Disposition Effect²:** Tendency to sell winners too early and hold losers too long.

² The Disposition Effect is excluded from our analysis because the trading data does not include individual performance contributions per single trade but monthly cash flows on a portfolio level, as described in Gorzon and Nitzsch (2025) study.

Table 2 Summary of cognitive biases included in the model, with their behavioral interpretation and operational measurement used to calculate individual investor exposure scores

Bias Factor	Cognitive Bias	Operationalization	(1)	Meaning of a high value
Overtrading (OT)	Action Bias	$OT_{it} = \text{Average Portfolio Turnover} = \frac{1}{2} \cdot \frac{\text{Purchase Volume}_{i,t-1}}{\text{Portfolio Volume}_{i,t}} + \frac{1}{2} \cdot \frac{\text{Sale Volume}_{i,t-1}}{\text{Portfolio Volume}_{i,t}}$	$[0, +\infty)$	Tendency to trade frequently
Trade Clustering (TCL)	Narrow Framing Bias	$\text{Trade Clustering}_{i,t} = \frac{\text{Number of trading days}_{i,t}}{\text{Number of trades}_{i,t}}$	$[0, 1]$	Tendency to select investments individually instead of considering the broad (holistic) impact on the portfolio
Lottery Stock Preference (LSP)	Sensation Seeking Bias	$LSP_{it} = \frac{\text{Lottery Portfolio Volume}_{i,t}}{\text{Total Portfolio Volume}_{i,t}}$	$[0, 1]$	Tendency to select stocks with lottery-like features (high recent return and volatility)
Trend Chasing (TC)	Extrapolation bias	$TC_{it} = \text{Average trend before all trades } (k) = \sum_{j=1}^N r(i, j, k)$	$(-\infty, +\infty)$	Tendency to chase assets with high positive or negative recent returns
Home Bias	Familiarity Bias	$HB_{it} = \frac{\text{Domestic Portfolio Volume}_{i,t}}{\text{Total Portfolio Volume}_{i,t}}$	$[0, 1]$	Tendency to select stocks with headquarters in the investor's home country
Herfindahl–Hirschman-Index (HHI)	Portfolio Concentration Bias	$HHI_{it} = \sum_{i=1}^N \omega_i \text{weight}_i^2$	$(0, 1]$	Tendency to hold few or concentrated stock positions
Disposition Effect (DE) ^a	Loss Aversion Bias	$\text{Disposition Effect}_{i,t} = \frac{\text{Realized Gains}_{i,t}}{\text{Realized Gains}_{i,t} + \text{Unrealized Gains}_{i,t}} - \frac{\text{Realized Losses}_{i,t}}{\text{Realized Losses}_{i,t} + \text{Unrealized Losses}_{i,t}}$	$[-1, 1]$	Tendency to sell winners too early and hold losers too long

^aThe Disposition Effect is excluded from our analysis because the trading data does not include individual performance contributions per single trade but monthly cash flows on a portfolio level, as described in Gorzon and von Nitzsch (2025) study

(1) Theoretical range

These biases are quantified using measurable proxies based on observed investor behaviors, such as trading frequency, diversification, and security selection tendencies. The next sections detail the application of these biases in the behavioral performance attribution framework.

2.2 Behavioral performance attribution

This section outlines the adaptation of the Brinson, Hood, and Beebower (BHB) model for behavioral performance attribution, providing a structured approach to decompose portfolio returns based on cognitive biases. The traditional BHB model assigns return contributions to asset allocation and security selection forming the foundation for performance evaluation in institutional finance. We adapt this approach by introducing behavioral biases as explanatory factors in retail investor performance analysis.

While we refer to ‘attribution’ of returns to behavioral biases, we note that the model is correlational and does not establish causal relationships. Interpretations are therefore conditional on the associations derived from observed investor behaviors and portfolio outcomes.

2.2.1 Traditional brinson, hood, and beebower (BHB) model

The Brinson, Hood, and Beebower (BHB) model (Brinson et al. 1986) provides a structured approach to understanding portfolio performance by decomposing total returns into distinct sources of contribution. The model attributes portfolio returns to three primary components: market return, asset allocation, and security selection. The market return contribution represents the return an investor would have earned by passively tracking the benchmark index without any active decision-making. The asset allocation contribution measures the impact of strategic decisions to overweight or underweight specific asset classes relative to the benchmark allocation. Lastly, the security selection contribution captures the portion of returns generated by actively selecting securities within each asset class, independent of allocation decisions. In addition to these two effects, the model also accounts for an “Other” term, also referred to as the interaction effect or cross-product term, which represents the combined impact of asset allocation and security selection decisions interacting together (Bacon 2019).

The calculation process follows a structured approach to determine the extent to which each factor contributes to the investor’s total return. First, a policy portfolio is defined, representing the benchmark allocation of assets based on the investor’s long-term strategy. The total return of the investor’s portfolio μ_i is then compared to the return of the selected benchmark portfolio μ_{bm} to identify any over- or underperformance α_i . These deviations are decomposed into asset allocation effect $\alpha_{allocation}$, security selection effect α_{select} , and an interaction effect α_{other} (“Other”), which accounts for the simultaneous influence of both allocation and selection decisions. Mathematically, the model expresses total excess return α_i as the sum of asset allocation, security selection, and interaction effects (see Eq. 2 below).

For example, if an investor achieves a total return of 10% while the benchmark return is 6%, the model explains the 4% alpha by attributing a portion to asset allocation—such as shifting more weight into equities instead of bonds—another portion to security selection—choosing high-performing stocks within the equity segment—and a final portion to the interaction effect, which reflects how these two active decisions reinforce or offset each other. If the asset allocation effect contributes 1.5% and the security selection effect contributes 2.0%, then the interaction effect would account for the remaining 0.5% of the excess return.

Excess Return (Alpha)

$$\mu_i - \mu_{bm} = \alpha_i \text{ i.e. } 10\% - 6\% = 4\% \quad (1)$$

Performance Attribution

$$\alpha_i = \alpha_{allocation} + \alpha_{select} + \alpha_{other}, \text{ i.e. } 4\% = 1.5\% + 2.0\% + 0.5\% \quad (2)$$

Empirical studies applying the BHB framework, such as the original research on U.S. pension plans from 1974 to 1983 (Brinson et al. 1986), demonstrated that 93.6% of excess α_i return variation could be explained, highlighting its dominant role in long-term performance. The BHB model framework remains widely adopted in institutional finance for professional portfolio managers as a fundamental tool for performance attribution and investment evaluation as it is still part of the Chartered Financial Analyst program (Bacon 2019).

2.2.2 Adapting performance attribution to behavioral biases

Retail investors' returns are influenced not only by strategic investment choices but also by cognitive biases. To account for these influences, we adapt the Brinson, Hood, and Beebower framework within the context of Behavioral Finance by computing bias-attributed returns. Similar to the BHB model, which decomposes excess portfolio returns α_i into asset allocation, security selection, and interaction effects, our approach decomposes an investor's excess return α_i into an explained component $\alpha_{attrib,i}$ driven by behavioral biases and a residual component $\alpha_{resid,i}$ capturing unexplained influences conceptually similar to the BHB model's "Other" term.

While the structure of return decomposition is similar, the fundamental difference lies in attribution focus: the BHB model explains rational allocation and selection decisions relative to a chosen benchmark, whereas the Behavioral Performance Attribution framework quantifies the presence of individual investors' behavioral biases (not relative to a benchmark) and its effects on portfolio returns. Integrating bias factors into BHB would introduce double-counting and break the summation constraint, as biases affect multiple return drivers in portfolio composition, e.g., Portfolio Concentration Bias and Home Bias or stock picking, e.g., Sensation Seeking Bias and Extrapolation Bias (Gorzon et al. 2024). By treating biases as a distinct explanatory framework, we ensure a more accurate and meaningful assessment of investor performance.

The adaption of the framework builds up on Gorzon et al. (2024) two requirements for a holistic evaluation of behavioral biases. First, a holistic approach implies the collective exhaustiveness of the behavioral biases considered (Minto 2002). Therefore, Gorzon et al. (2024) conducted an extensive literature review to capture all relevant (quantifiable) behavioral biases in existing literature. We include their identified (economically and statistically) significant scope of biases in our behavioral performance attribution framework as behavioral determinants of portfolio performance. Second, mutual exclusiveness means that behavioral biases do not overlap. Therefore, Gorzon and Nitzsch (2025) examined the correlation and multicollinearity among bias factors as part of their empirical analysis, reducing redundancy and enhancing the model's reliability by ensuring that each bias contributes unique explanatory power to the behavioral performance attribution. However, biases are complex as they interact and overlap complicating precise exclusiveness and exhaustiveness and hence performance decomposition. Consequently, our behavioral attribution model includes a residual (unexplained) return component $\alpha_{resid,i}$. Due to the inherent complex nature of behavioral biases, we expect this residual to be larger than in traditional performance attribution models which assume rational investor behavior.

Following the BHB model framework (Sect. 2.2.1), Eqs. 3, 4, 5 define the behavioral performance attribution framework, detailing how an investor's total excess return α_i is decomposed into a component explained by behavioral biases $\alpha_{attrib,i}$ and a residual component $\alpha_{resid,i}$ that captures influences beyond the modeled biases.

Behavioral Bias Attributable Performance

$$\alpha_{attrib,i} = \sum_{j=1}^N \beta_{j,g} B_{j,i} \quad (3)$$

Behavioral Performance Attribution

$$\alpha_i = \alpha_{attrib,i} + \alpha_{resid,i} \quad (4)$$

Residual Performance

$$\alpha_{resid,i} = \alpha_i - \alpha_{attrib,i} \quad (5)$$

where:

$\alpha_{attrib,i}$: Return attributed to N behavioral biases for investor i by the Behavioral Performance Attribution model ($N=6$ in this study).

α_i : Actual excess return relative to the benchmark for investor i .

$\alpha_{resid,i}$: Return not attributable to behavioral biases for investor i by the Behavioral Performance Attribution model.

$\beta_{j,g}$: Coefficient β of bias factor j for each performance group g is derived from OLS regression results.

$B_{j,i}$: Exposure B of bias factor j for individual investor i (normalized [0,1]).

Before we can start with the Behavioral Performance Attribution in Eq. (3), all idiosyncratic bias factor exposures $B_{j,g}$ need to be computed individually for every investor. A detailed description of the calculation procedure is provided in the paper of Gorzon and Nitzsch (2025). However, we have attached all used formulas to calculate these bias factor exposures in Table 2 in Sect. 2.1.2.

Then, from a model perspective, Adapting Performance Attribution to Behavioral Biases requires a two-step approach. First, the model estimates the importance of each bias factor $\beta_{j,g}$ using ordinary least squares (OLS) regression (Gorzon and Nitzsch 2025). The importances of each bias factor $\beta_{j,g}$, which are tailored to each performance group g (Table 3 and Table 4) indicate how strongly each bias is associated with portfolio excess performance. Second, we apply the determined importances $\beta_{j,g}$ to individual investors' idiosyncratic bias factor exposures $B_{j,i}$, which vary from person to person. By multiplying the importance (coefficient) of each bias $\beta_{j,g}$ with an investor's specific bias exposure $B_{j,i}$, we compute the performance contribution of each bias to the investor's total return. Summing these contributions across all biases provides the bias-attributed return $\alpha_{attrib,i}$ in Eq. (3). Consequently, the residual $\alpha_{resid,i}$ equals the excess return minus the explained component in Eq. (5).

The provided approach lays the groundwork for further evaluation in Sect. 2.3, where we introduce the methodology to assess the effectiveness of this behavioral performance attribution model and compare attribution outcomes across performance groups.

2.3 Explainability evaluation of behavioral bias performance attribution

2.3.1 Model over- and underestimation of individual investor's returns

To assess the effectiveness of behavioral performance attribution to explain an investor's excess returns, we introduce a Model Explainability Ratio MER in Eq. (6). This validation ensures that over- and underestimation are symmetrically evaluated,

Table 3 Estimated OLS coefficients for each bias factor among winning investors at various trimming levels

Trim in % from from both tails	Action bias	Narrow framing bias	Sensation seeking bias	Extrapolation bias	Familiarity bias	Portfolio concentra- tion bias
0,00%	1.9351***	0.0678***	0.0463	0.2699***	-0.0613***	0.2673***
1,00%	0.8664*	0.0644***	0.0588	0.1716**	-0.0384***	0.1953***
2,50%	0.8909**	0.0566***	0.0679***	0.1252**	-0.0189***	0.1036***
5,00%	1.1181***	0.0462***	0.0389**	0.0605*	-0.015***	0.0549***
10,00%	0.5696***	0.028***	0.0226**	0.0349	-0.0005	0.0144***
15,00%	0.0568***	0.0171***	0.006	0.0159**	0.0008	0.0025
20,00%	0.0306**	0.0079***	0.0052	0.0115	0.0009	0.0001

Asterisks denote significance levels: $p < 0.1^*$, $< 0.05^{**}$, $< 0.01^{***}$

Table 4 Estimated OLS coefficients for each bias factor among winning investors at various trimming levels

Trim in % from both tails	Action Bias	Narrow Framing Bias	Sensation Seeking Bias	Extrapolation Bias	Familiarity Bias	Portfolio Concentration Bias
0,00%	-2.4295	-0.0425***	-0.4211***	-0.1097**	-0.0526***	-0.0993***
1,00%	-0.937	-0.0458***	-0.3667***	-0.1683***	-0.0513***	-0.0947***
2,50%	-0.6933	-0.048***	-0.2437***	-0.1267***	-0.0489***	-0.069***
5,00%	-0.3987	-0.0517***	-0.2012***	-0.0538***	-0.0416***	-0.0394***
10,00%	-1.8812***	-0.0468***	-0.0988***	-0.0395**	-0.0378***	-0.0048
15,00%	-0.1696***	-0.0456***	-0.0513***	-0.0287**	-0.0338***	0.0112***
20,00%	-0.1838***	-0.0422**	-0.0355***	-0.0229	-0.0312***	0.0187***

Asterisks denote significance levels: $p < 0.1$ *, $p < 0.05$ **, $p < 0.01$ ***

making comparisons between deviations meaningful. We provide examples of the application of the equation in four different scenarios after its definition.

$$\text{MER}_i(\alpha_{\text{attrib},i}, \alpha_i) = \begin{cases} 100, \text{ if } (\alpha_i = \alpha_{\text{attrib},i}) \\ \left(\frac{\min(|\alpha_{\text{attrib},i}|, |\alpha_i|)}{\max(|\alpha_{\text{attrib},i}|, |\alpha_i|)} \right) * 100, \text{ if } (\alpha_i > 0 \text{ and } \alpha_{\text{attrib},i} > 0) \text{ or } (\alpha_i < 0 \text{ and } \alpha_{\text{attrib},i} < 0) \\ - \left(\frac{\min(|\alpha_{\text{attrib},i}|, |\alpha_i|)}{\max(|\alpha_{\text{attrib},i}|, |\alpha_i|)} \right) * 100, \text{ if } (\alpha_i > 0 \text{ and } \alpha_{\text{attrib},i} < 0) \text{ or } (\alpha_i < 0 \text{ and } \alpha_{\text{attrib},i} > 0) \end{cases} \quad (6)$$

where:

$\text{MER}_i(\alpha_{\text{attrib},i}, \alpha_i)$: = Model Explainability in % for investor i .

Consider four scenarios where the model-estimated return and an investor's excess return are as follows:

- Scenario A: Model estimates $\alpha_{\text{attrib},i}=4\%$ and investor excess return is $\alpha_i=5\% \rightarrow$ The model underestimates return.
- Scenario B: Model estimates $\alpha_{\text{attrib},i}=5\%$ and investor excess return is $\alpha_i=4\% \rightarrow$ The model overestimates return.
- Scenario C: Model estimates $\alpha_{\text{attrib},i}=-2\%$ and investor excess return is $\alpha_i=+5\% \rightarrow$ The model and the actual excess return indicate differences in return directions.
- Scenario D: Model estimates $\alpha_{\text{attrib},i}=5\%$, and investor excess return is $\alpha_i=5\% \rightarrow$ The model accurately estimates return.

In Scenario A, the model explains 4% out of the 5% excess return, leading to an explainability score of 80% or $0.8 \left(\frac{4\%}{5\%} \right)$. This suggests that a currently missing additional factor(s) must account for the remaining 20% of returns. In Scenario B, the model overestimates return, originally suggesting 125% explainability $(=\frac{5\%}{4\%}1.25)$, which is misleading. Therefore, we take the absolute minimum of the compared returns over the maximum of the compared returns, such that in Scenario B, we also compute $\frac{4\%}{5\%}$ leading to 80% explainability. In Scenario C, the returns have different directions and using MER, the calculation results in $-\frac{|-2\%|}{|+5\%|} * 100 = -40\%$. While a negative MER lacks intuitive interpretability, it serves a critical role in model validation by penalizing cases where the model's predictions systematically oppose actual outcomes, thereby highlighting severe attribution errors. However, a negative MER score only applies to sixteen winning investors (out of 4996) and zero losing investors (out of 15004) i.e. 0.08% of the total sample. Finally, Scenario D explains the model accurately.

2.3.2 Residual analysis

While the MER Ratio quantifies how much of an investor's return can be attributed to bias factors, analyzing residual deviations provides deeper insights into systematic

model tendencies of over- and underestimation. To assess these tendencies, we evaluate True Residuals and Absolute Residuals.

True Residuals (i.e., positive and negative values) indicate whether the model systematically overestimates or underestimates investor returns e.g. for winners $\alpha_{attrib,i} > \alpha_i$ or $\alpha_{attrib,i} < \alpha_i$ and for losers vice versa. Identifying directional biases in estimation can highlight missing explanatory factors which may have positive or negative effects on returns depending on the residual analysis result, e.g., if the model systematically overestimates returns for winners who outperform the benchmark, it indicates the need for an additional factor(s) that reduces the model-estimated return to more accurately decomposes actual excess returns.

Absolute Residuals (i.e., non-negative values) measure overall model accuracy by considering deviations regardless of direction. This distinction is critical when summarizing residuals using mean or median values. For instance, if the model produces residuals of +5% and -5%, a simple mean calculation will misleadingly suggest a perfect fit (0% delta), despite the actual average delta of 5%, deviating 5% in both directions.

This two-pronged residual evaluation allows for a more nuanced understanding of model performance and potential areas for refinement in behavioral bias attribution.

2.4 Trimming-based analysis

To assess the generalizability of the behavioral performance attribution framework, we implement a systematic outlier trimming approach that iteratively removes extreme-performing investors from both tails of the return distribution (Gorzon and Nitzsch 2025). While this assessment does not provide a definitive test of generalizability, it offers preliminary evidence by evaluating the model's performance (MER) across multiple subsamples. Since investor returns exhibit a normal distribution, progressively removing outliers produces refined sub-datasets that are better aligned with the average investor profile of the given sample. This analysis tests whether our framework remains applicable across different investor samples and is not disproportionately influenced by extreme observations. However, we recognize that further validation using independent trading datasets would be necessary to establish the framework's broader applicability.

By conducting new OLS regressions for each trimmed dataset, we systematically evaluate whether the attribution model continues to provide reliable and consistent results across varying subsamples (Table 5). Regression models are particularly sensitive to outliers, making robust estimation crucial in datasets with high heterogeneity, such as ours (Gorzon and Nitzsch 2025). Retail investors, unlike institutional investors, exhibit diverse trading behaviors, varying levels of financial literacy, and different investment strategies, resulting in substantial return differences. These differences increase the likelihood of extreme observations, which, if not accounted for, may distort coefficient estimates and weaken inference reliability (Gujarati and Porter 2009).

We apply symmetric trimming by progressively removing the top and bottom (from both tails) 1%, 2.5%, 5%, 10%, 15% and 20% of investors based on achieved

Table 5 Effect of outlier trimming on sample size and model fit (adjusted R^2) across trimming levels from 0 to 20% for winning and losing investors

Trim in % from both tails	% of Sample Removed	Remaining Investors	Remaining Winning Investors	Remaining Losing Investors	Adjusted R2 Winning Investors	Adjusted R2 Losing Investors
0,00%	0,00%	20,000	4996	15,004	47,10%	57,30%
1,00%	2,00%	19,600	4796	14,804	47,60%	59,00%
2,50%	5,00%	19,001	4497	14,504	52,60%	63,20%
5,00%	10,00%	18,001	3997	14,004	59,00%	66,90%
10,00%	20,00%	16,006	2999	13,007	62,80%	71,40%
15,00%	30,00%	14,004	1999	12,005	66,60%	72,70%
20,00%	40,00%	12,005	996	11,009	69,50%	74,40%

portfolio returns. This approach maintains the un-trimmed balance of the dataset while mitigating the influence of extreme outliers. For each of these six trimmed datasets, we conduct a new OLS regression analysis and evaluate the impact of trimming based on three key dimensions:

- **Model fit:** Adjusted R-squared changes to assess improvements in model performance.
- **Explainability and residual analysis:** We recompute actual excess returns, model-estimated excess returns, and absolute/true residuals to identify patterns in model attribution stability.
- **Bias attribution stability:** Changes in coefficient magnitudes and significance levels, providing insights into how robust the identified bias effects are. If removing high-performing investors significantly alters the significance of a particular bias, it suggests that the bias plays a greater role in explaining exceptional performance rather than general investment outcomes.

By incorporating this trimming-based analysis, we ensure that our findings are not driven by outliers, thereby strengthening the generalizability and applicability of the behavioral performance attribution model to retail investing.

3 Results

3.1 Behavioral performance attribution

This section presents the application of the behavioral performance attribution framework to individual investor returns. Using the methodology outlined in Sect. 2, we decompose excess returns into components attributable to cognitive biases and an unexplained residual component. The analysis is conducted separately for winning investors (those who outperform the benchmark) and losing investors (those who underperform the benchmark).

For each performance group, we first apply the behavioral performance attribution formula, computing bias-attributed returns based on the OLS regression coefficients. The calculation process is illustrated with representative examples of individual investors (Table 1 in Sect. 1). Next, we evaluate model explainability using the Model Explainability Ratio (MER), ensuring a standardized assessment of return over- and underestimation. Finally, we conduct a residual analysis, distinguishing between true residuals (directional differences between model-estimated and actual excess returns) and absolute residuals (magnitude of deviations), providing insights into the effectiveness of behavioral attribution across different investor profiles.

3.1.1 Examples of winning investors

To assess the contribution of cognitive biases to investment performance, we describe the application of the behavioral performance attribution framework to two winning investors, whose returns exceeded the benchmark ($\alpha_i > 0\%$). Using

the previously estimated OLS regression coefficients, each investor's bias-attributed return is computed by multiplying their individual bias exposures $B_{j,i}$ with the respective OLS-coefficient³ for Winning Investors $\beta_{j,wi}$ from Table 3, as outlined in Sect. 2.2.2.

We provide two examples of Winning Investors #1 and #2:

$$\begin{aligned}\beta_{j,wi} &= [\beta_{1,wi} = 1.9351, \beta_{2,wi} = 0.0678, \beta_{3,wi} = 0.0463, \beta_{4,wi} = 0.2699, \beta_{5,wi} = -0.0613, \beta_{6,wi} = 0.2673] \\ B_{j,1} &= [B_{1,1} = 0.0010, B_{2,1} = 0.2644, B_{3,1} = 0.0022, B_{4,1} = 0.0003, B_{5,1} = 0.2629, B_{6,1} = 0.1225] \\ B_{j,2} &= [B_{1,1} = 0.0202, B_{2,1} = 0.4353, B_{3,1} = 0.0000, B_{4,1} = 0.0108, B_{5,1} = 0.5720, B_{6,1} = 0.0730]\end{aligned}$$

The following section describes the step-by-step computation of Behavioral Performance Attribution for Winning Investor #1, followed by a summary of results for both investors. Finally, we extend the analysis to assess the behavioral attribution results at the performance group level for all winning investors.

First, the total return of the investor's portfolio μ_i is compared to the return of the selected benchmark portfolio μ_{bm} to identify any over- or under performance α_i .

$$\alpha_1 = \mu_1 - \mu_{bm} = 13.47\% - 8.09\% = 5.38\%$$

The attribution process follows a structured approach where the investor's total excess return will be decomposed into two components in Eq. (4), the bias-attributed return and the residual return.

$$\alpha_1 = \alpha_{attrib,1} + \alpha_{resid,1} = 5.38\%$$

To compute bias-attributed performance, we sum the contributions of each individual bias factor using Eq. (3):

$$\alpha_{attrib,1} = \sum_{j=1}^N \beta_{j,wi} B_{j,1} = 3.67\%$$

This means that 3.67% of Winning Investor #1's excess return of 5.38% can be explained by our behavioral attribution model.

The residual return, capturing the portion not explained by behavioral biases, is calculated using Eq. (5):

$$\alpha_{resid,1} = \alpha_1 - \alpha_{attrib,1} = 5.38\% - 3.67\% = 1.71\%$$

To evaluate the effectiveness of the attribution model, we compute the Model Explainability Ratio using Eq. (6).

$$\text{MER}_i(\alpha_{attrib,1}, \alpha_1) = \frac{3.67\%}{5.38\%} * 100 = 68.16\%$$

The MER indicates that 68.16% of Winning Investor #1's return is explained by behavioral performance attribution, while 31.84% remains unexplained, suggesting

³ $\beta_{1,g}$ =Action Bias, $\beta_{2,g}$ =Narrow Framing Bias, $\beta_{3,g}$ =Sensation Seeking Bias, $\beta_{4,g}$ =Extrapolation Bias, $\beta_{5,g}$ =Familiarity Bias, $\beta_{6,g}$ =Portfolio Concentration Bias.

potential influences from unmodeled factors such as market timing, macroeconomic conditions, or additional behavioral tendencies.

For Winning Investor #1, the largest behavioral contributors to the outperformance $\alpha_1 = 5.38\%$ are Narrow Framing Bias (1.79%) and Portfolio Concentration Bias (3.27%), indicating that this investor's tendency to evaluate investments in isolation and to maintain a concentrated portfolio significantly contributed to the excess returns. Conversely, Winning Investor #2's ($\alpha_2 = 4.06\%$) highest positive behavioral performance attributions stem from Action Bias (3.91%) and Narrow Framing Bias (2.95%), implying that frequent trading activity, sustained over multiple days, played a dominant role in their investment outcomes. The negative contribution of Familiarity Bias (-1.61% and -3.51%) for both investors suggests that overweighting familiar (i.e., German-based) stocks negatively impacted performance, in contrast to other biases that enhanced returns.

Summing all bias contributions, Winning Investor #1 has a bias-attributed return of 3.67%, while Winning Investor #2 has 5.60%. When compared to actual achieved excess returns (5.38% and 4.06%, respectively), the Model Explainability Ratio (MER) is computed, indicating how much of the investor's performance can be attributed to the modeled biases. The results show that 68.16% of Winning Investor #1's excess return and 72.54% of Winning Investor #2's excess return can be explained by behavioral performance attribution, leaving the remainder as residual performance influenced by additional unmodeled factors.

3.1.2 Winning investors performance group results

Next, we assess the attribution's effectiveness at the group level by analyzing winning investors based on the model and investor actual absolute and true delta excess return as well as the MER. This analysis provides a broader perspective on how well cognitive biases explain investment performance at the group level and identifies potential limitations in the attribution model. The results are visualized in Fig. 1.

For consistency and robustness, we focus on median values, as they are less sensitive to extreme outliers that could distort the overall representation of investor performance. However, for completeness, mean values are also provided in Table 6 to offer an alternative perspective on the distribution of returns.

The distribution of absolute delta returns (left panel) shows that most investors exhibit relatively small deviations between actual and model-estimated returns, but with a long tail extending toward larger discrepancies. The median absolute delta return is 4.56%, while the mean is higher at 6.90%, reflecting the influence of extreme cases.

The true delta return distribution (center panel) is centered near zero but slightly skewed to the left, indicating that the model tends to overestimate outperformance for many investors. The median true delta return of -2.22% suggests that, on average, the model estimates higher excess returns than those actually achieved, reinforcing this overestimation tendency.

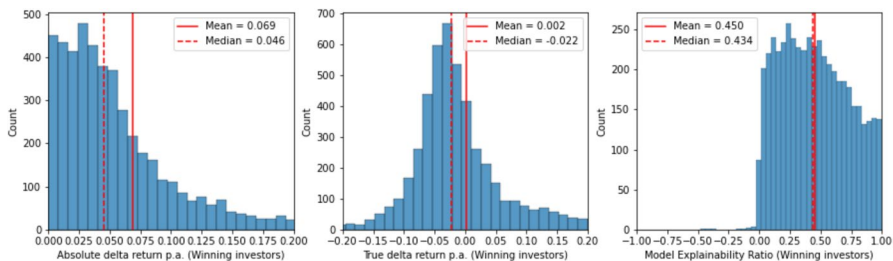


Fig. 1 Distribution of absolute and directional differences between actual and model-estimated excess returns (left and center), and distribution of Model Explainability Ratios (right) for winning investors. X-axis represents return deviation or explainability percentage; Y-axis shows investor count

To assess the proportion of performance explained by behavioral biases, we compute the Model Explainability Ratio (MER) (right panel). The MER distribution is broad and roughly uniform, with a median of 43.44%, meaning that a significant portion of excess returns remains unexplained. However, the mean MER of 50% suggests that some investors have substantially higher explainability, reinforcing the model's effectiveness for a subset of investors. The proportion of explained variance ($R^2 = 47.10\%$, Table 5) aligns closely with the MER, reinforcing the attribution's explanatory power across the performance group.

3.1.3 Examples of losing investors

To analyze the role of cognitive biases in driving underperformance, we provide two losing investors whose returns were below the benchmark ($\alpha_i < 0\%$). We proceed analogously to the Examples Winning Investors in Sect. 3.1.1, where bias-attributed returns are calculated by multiplying each investor's bias exposures $B_{j,i}$ with the respective OLS regression coefficients for Losing Investors $\beta_{j,li}$ from Table 4.

Examples of Losing Investors:

$$\begin{aligned} \beta_{j,li} &= [\beta_{1,li} = -2.4295, \beta_{2,li} = -0.0425, \beta_{3,li} = -0.4211, \beta_{4,li} = -0.1097, \beta_{5,li} = -0.0526, \beta_{6,li} = -0.0993,] \\ B_{j,3} &= [B_{1,3} = 0.0003, B_{2,3} = 0.4670, B_{3,3} = 0.0103, B_{4,3} = 0.0032, B_{5,3} = 0.8319, B_{6,3} = 0.1089] \\ B_{j,4} &= [B_{1,4} = 0.0739, B_{2,4} = 0.2162, B_{3,4} = 0.0575, B_{4,4} = 0.0062, B_{5,4} = 0.5209, B_{6,4} = 0.0489] \end{aligned}$$

For Losing Investor #3, the largest negative contributors to underperformance $\alpha_3 = -5.48\%$ stem from Narrow Framing Bias (-1.98%) and Familiarity Bias (-4.38%), suggesting that evaluating investments in isolation and over-relying on familiar stocks were primary predictors of poor performance.

Conversely, Losing Investor #4 $\alpha_4 = -28.93\%$ exhibits a particularly strong negative performance contribution from Action Bias (-17.95%), indicating that excessive trading significantly eroded returns. Sensation Seeking Bias (-2.42%) also played a role, implying a preference for high-volatility, speculative stocks that led to significant losses.

Table 6 Summary statistics for actual and model-estimated excess returns among winning investors across trimming levels. Includes absolute and true residuals, and MER

Trim in % from both tails	Actual_excess_return	Model-estimated_return	Delta_return_abs	Delta_return_true	Model_explain-ability_ratio (mer)
0,00%	4,39% (8,64%)	7,26% (8,47%)	4,56% (6,90%)	-2,22% (0,18%)	43,44% (45,03%)
1,00%	4,24% (7,66%)	6,49% (7,23%)	4,08% (5,88%)	-1,73% (0,43%)	44,02% (46,10%)
2,50%	3,86% (5,92%)	5,35% (5,57%)	3,31% (4,27%)	-1,16% (0,35%)	46,42% (47,99%)
5,00%	3,26% (4,26%)	3,99% (3,97%)	2,38% (2,83%)	-0,46% (0,29%)	49,33% (50,29%)
10,00%	2,21% (2,51%)	2,41% (2,32%)	1,38% (1,53%)	-0,07% (0,20%)	53,34% (52,84%)
15,00%	1,34% (1,45%)	1,43% (1,35%)	0,75% (0,82%)	0,02% (0,10%)	56,62% (55,49%)
20,00%	0,67% (0,66%)	0,67% (0,63%)	0,34% (0,36%)	0,03% (0,03%)	59,60% (56,29%)

The negative influence of Familiarity Bias (-4.38% and -2.74%) for both investors highlight that overweighting well-known, domestic stocks negatively impacted performance.

Summing all bias contributions, Losing Investor #3 has a bias-attributed return of $\alpha_{attrib,3} = -7.98\%$, while Losing Investor #4 has $\alpha_{attrib,4} = -24.59\%$. When compared to their actual achieved excess returns of $\alpha_3 = -5.48\%$ and $\alpha_4 = -28.93\%$, the Model Explainability Ratio (MER) is computed as 68.64% for Losing Investor #3 and 85.01% for Losing Investor #4. These results indicate that a substantial portion of investment underperformance can be attributed to behavioral biases, while the remaining variance is influenced by additional unmodeled factors such as market conditions, sentiment fluctuations, or structural inefficiencies.

3.1.4 Losing investors performance group results

Next, we assess the attribution's effectiveness at the group level by analyzing all losing investors. The overall attribution results indicate that the median actual excess return for losing investors is -6.01%, while the median model-estimated excess return is -7.49% (Table 7). The median absolute delta return ($|Actual - Model-Estimated|$) is 3.37%, whereas the median true delta return ($Actual - Model-Estimated$) is +1.06%, suggesting that the model tends to overestimate underperformance for losing investors i.e. estimating lower excess than actual returns $\alpha_{attrib,i} < \alpha_i$. The results are visualized in Fig. 2.

The distribution of absolute delta returns (left panel) shows that most investors exhibit relatively small deviations between actual and model-estimated returns, but a long tail extends toward larger discrepancies. The median absolute delta return is 3.4%, while the mean is slightly higher at 5.1%, indicating that extreme cases influence the mean more than the median.

The true delta return distribution (center panel) is symmetrically centered near zero, with a slight right skew, reflecting that while most estimates align closely with actual returns, a systematic tendency exists for the model to overestimate losses. The median true delta return of +1.06% reinforces this pattern.

To assess the proportion of underperformance explained by cognitive biases, we compute the Model Explainability Ratio (MER) (right panel). The MER distribution is relatively uniform, with a median of 59.17%, meaning that cognitive biases account for the majority of the individual actual excess return, leaving 40.83% as unexplained residual return. The mean MER of 57.4% confirms the robustness of this estimate, as most investors exhibit similar explainability levels. The proportion of explained variance ($R^2 = 57.30\%$, Table 5) aligns closely with the MER, reinforcing the model's effectiveness in explaining behavior-driven underperformance.

Table 7 Summary statistics for actual and model-estimated excess returns among losing investors across trimming levels. Includes absolute and true residuals, and MER

Trim in % from both tails	Actual_excess_return	Model-estimated_return	Delta_return_abs	Delta_return_true	Model_explain-ability_ratio (mer)
0,00%	-6,01% (-8,58%)	-7,49% (-8,29%)	3,37% (5,09%)	1,06% (-0,29%)	59,17% (57,40%)
1,00%	-5,98% (-8,20%)	-7,49% (-7,93%)	3,36% (4,86%)	1,02% (-0,27%)	58,72% (57,16%)
2,50%	-5,84% (-7,51%)	-6,97% (-7,25%)	3,19% (4,27%)	0,77% (-0,26%)	59,36% (57,69%)
5,00%	-5,64% (-6,80%)	-6,44% (-6,54%)	2,95% (3,71%)	0,53% (-0,26%)	59,94% (58,40%)
10,00%	-5,26% (-5,86%)	-5,60% (-5,63%)	2,51% (2,99%)	0,17% (-0,23%)	61,64% (59,86%)
15,00%	-4,84% (-5,19%)	-5,06% (-4,95%)	2,23% (2,54%)	-0,03% (-0,24%)	62,52% (60,47%)
20,00%	-4,48% (-4,65%)	-4,51% (-4,45%)	1,96% (2,22%)	-0,11% (-0,20%)	63,54% (61,02%)

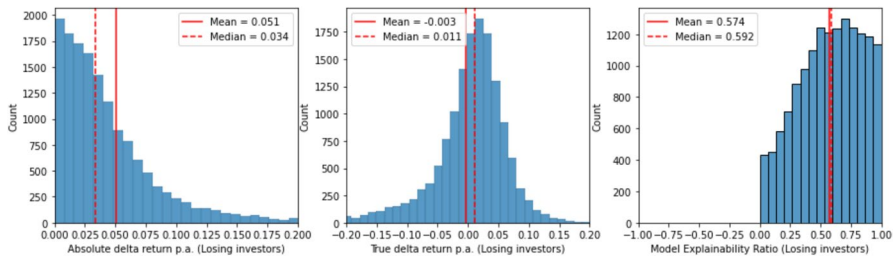


Fig. 2 Distribution of absolute and directional differences between actual and model-estimated excess returns (left and center), and distribution of Model Explainability Ratios (right) for losing investors. X-axis represents return deviation or explainability percentage; Y-axis shows investor count

3.2 Trimming analysis

To assess the generalizability and robustness of the behavioral performance attribution model, we conduct a trimming-based analysis by iteratively removing extreme-performing investors from both tails of the return distribution. This process systematically evaluates how outliers influence model fit, explainability metrics and bias attribution stability. By progressively removing the top and bottom 1%, 2.5%, 5%, 10%, 15%, and 20% of investors based on achieved portfolio returns, we assess whether extreme observations distort regression estimates and whether their exclusion improves the reliability and consistency of our findings.

3.2.1 Sample reduction and model fit (R2 scores)

Initially, the full sample consists of 20,000 investors, including 4,996 winning investors and 15,004 losing investors. As trimming increases, the sample shrinks, with the most extreme case (20% trimming from both tails) reducing the dataset to 12,005 investors, comprising 996 winning investors and 11,009 losing investors (Table 5). This reduction removes the highest and lowest performers progressively, aiming to improve the reliability of bias attribution by minimizing distortions introduced by outliers.

Trimming has a notable impact on model fit, measured by the adjusted R^2 score, which reflects the proportion of variance in investor returns explained by the behavioral attribution model. In the full sample (0% trimming), adjusted R^2 for winning investors is 47.10% and for losing investors is 57.30%. As trimming progresses, adjusted R^2 steadily increases for both groups, reaching 69.50% for winning investors and 74.40% for losing investors at the 20% trim level (Table 5). This trend indicates that extreme-performing investors introduce noise into the model, reducing its explanatory power. Their exclusion results in a more stable regression fit, improving the alignment between bias exposure and return attribution. The larger increase in adjusted R^2 for losing investors compared to winning investors suggests that extreme losses contribute more to model instability than extreme gains.

Overall, these results confirm that the behavioral performance attribution model performs more reliably when extreme investors are excluded, supporting the robustness of the framework across different subsamples.

3.2.2 Explainability (MER) & residual analysis

Trimming extreme-performing investors also significantly impacts model explainability, as measured by the Model Explainability Ratio (MER), and the accuracy of return predictions, assessed through residual analysis (Table 6 and Table 7).

For winning investors, MER initially is at 43.44%, indicating that cognitive biases explain a moderate share of their excess returns. As extreme performers are trimmed, MER progressively increases to 59.60% at the 20% trim level, aligning with the improvement in adjusted R^2 (Table 6 and Table 7). The increased MER confirms the improved model fit (Sect. 3.2.1) and suggests that bias attributions become more accurate when extreme winners are removed, reducing unexplained return components.

For losing investors, MER starts at 59.17% and rises to 63.54% at the 20% trim level. While this increase is relatively smaller than for winning investors, it confirms that trimming enhances attribution accuracy for underperforming investors as well (Table 7). The stronger initial MER for losing investors suggests that biases already explain a greater portion of underperformance, whereas for winners, other unmodeled factors—such as skill-based trading advantages—may play a more significant role.

Residual analysis further supports these findings. Absolute residuals, which capture total deviation between actual and model-estimated excess returns, decrease as trimming increases, indicating that predictions become more precise. True residuals, which measure directional errors, reveal that:

- Winning investors' returns are initially overestimated (median true residual = -2.22%). As trimming increases, this bias declines to 0.03%, suggesting that extreme winners contribute disproportionately to model overestimation.
- Losing investors' returns are initially overestimated (median true residual = +1.06%). With trimming, this bias moderates to -0.11%, reducing the degree of overprediction.

These results highlight that extreme investors introduce distortions into bias attributions, and their exclusion leads to a more precise and stable behavioral performance attribution model. The alignment between improvements in MER, reductions in residuals, and increases in adjusted R^2 confirms that trimming enhances model reliability by refining bias attributions and improving prediction accuracy.

In addition, from a generalization perspective, the results affirm that the Behavioral Performance Attribution framework remains robust across progressively refined investor samples. The fact that certain biases persist across different trim levels suggests that the framework is not solely driven by extreme cases but can reliably capture behavioral tendencies among the broader retail investor population.

3.2.3 Bias attribution stability (coefficients and significance levels)

To assess the stability of bias attributions across different investor samples, we analyze how the estimated OLS coefficients evolve as extreme performers are progressively removed from both tails of the return distribution. This analysis provides insights into whether specific biases are consistently significant across the full sample or if their explanatory power is driven primarily by extreme investors.

3.2.4 Winning investors (Outperforming the chosen benchmark)

For winning investors, the initial OLS regression (0% trimming) reveals that Action Bias (1.9351) and Portfolio Concentration Bias (0.2673) have the largest magnitudes among the estimated coefficients (Table 3). However, as trimming increases, these values systematically decline, with Action Bias falling to 0.0306 and Portfolio Concentration Bias to 0.0001 at the 20% trimming level. This substantial reduction suggests that these biases play a more prominent role in the returns of extreme-performing investors rather than the broader sample. Similarly, the statistical significance of Action Bias and Portfolio Concentration Bias diminishes as more extreme observations are removed, indicating that these behavioral effects are more relevant for outliers than for the general investor population.

Other biases, such as Narrow Framing Bias (0.0678) and Extrapolation Bias (0.2699), also experience a decrease in magnitude but remain statistically significant at various trimming levels. Narrow Framing Bias, in particular, retains significance across most trim levels, suggesting that its impact on investment decisions persists across both extreme and average investors. Conversely, Familiarity Bias (-0.0613) loses significance early in the trimming process, indicating that its effect on investment returns is largely concentrated among outlier investors.

3.2.5 Losing investors (Underperforming the chosen benchmark)

For losing investors, the initial OLS regression results (0% trimming) show that Sensation Seeking Bias (-0.4211) and Action Bias (-0.4295) are the strongest negative contributors to portfolio returns (Table 4). As trimming progresses, Sensation Seeking Bias weakens in magnitude, declining to -0.0355 at 20% trimming, while Action Bias fluctuates significantly before stabilizing at -0.1838. This pattern suggests that extreme underperformance is often linked to heightened speculative trading behaviors, which lose explanatory power when extreme investors are removed.

Further, Narrow Framing Bias (-0.0425) and Portfolio Concentration Bias (-0.0993) exhibit greater stability compared to the other biases. Their coefficients remain relatively consistent throughout the trimming process, and their statistical significance is preserved across most levels. This suggests that these biases exert a more uniform effect across different investor segments, rather than being driven primarily by extreme winners or losers.

Overall, these findings highlight that some behavioral biases—particularly those related to trading frequency and portfolio concentration—are more

pronounced among outliers, while others, such as Narrow Framing Bias, influence a broader range of investors. This has important implications for behavioral performance attribution, emphasizing the need to account for varying bias effects across different investor segments.

4 Discussion

4.1 Comparing behavioral performance attribution to the BHB model

The Behavioral Performance Attribution framework introduced in this study represents a fundamental shift from traditional return attribution models such as the Brinson, Hood, and Beebower (BHB) model (Brinson et al. 1986).

The BHB model assumes that investment decisions are more rational and driven by portfolio optimization (compared to retail investors). Hence the model is designed to evaluate professional and therefore financial literate portfolio managers' decisions at an institutional level, where allocation strategy plays a dominant role (Brinson et al. 1986). However, institutional portfolio managers tend to be a more homogeneous group, as they operate under structured investment mandates, risk management frameworks, and regulatory constraints, leading to more disciplined and systematic decision-making. In contrast, retail investors are heterogeneous, exhibiting a wide range of behaviors driven by varying levels of financial literacy, risk tolerance, heuristics and emotional biases which result in irrational decisions (Barber and Odean 2013; Gorzon et al. 2024). This distinction suggests that behavioral biases are more pronounced and diverse among retail investors which was confirmed in prior studies (e.g. overview of prior studies in Gorzon et al. 2024 and recent empirical results in Gorzon and Nitzsch 2025), whereas institutional investors, despite being subject to biases, may display more consistent and rational investment patterns. Therefore, one of the primary advantages of Behavioral Performance Attribution is its ability to quantify behavioral tendencies (biases) especially for the heterogeneous group of retail investors and raise awareness of their (costly) present. Consequently, the Behavioral Performance Attribution model decomposes achieved portfolio returns based on behavioral biases, rather than solely focusing on strategic allocation effects.

Our empirical results confirm that Behavioral Performance Attribution effectively explains a significant portion of investment returns, as indicated by the Model Explainability Ratio (MER). The MERs for winning investors (between 43.44% and 59.60%) and for losing investors (between 59.17% and 63.54%) suggest that behavioral biases provide valuable insights to better understand positive as well as negative investment performance of retail investors.

Overall, Behavioral Performance Attribution provides a complementary alternative to the BHB model, offering an intuitive approach for individual investors. While strategic asset allocation remains a critical determinant of returns particularly in

institutional investing, understanding behavioral tendencies provides an additional explanatory layer that can help investors optimize their decision-making process and thus portfolio performance.

4.2 Residual analysis: overestimation and underestimation patterns in behavioral attribution

The residual analysis provides key insights into the model's accuracy and the extent to which behavioral biases explain individual investment performance. While the Behavioral Performance Attribution framework effectively captures cognitive biases driving investment decisions, the residuals indicate systematic overestimation for both winning and losing investors. This suggests that additional behavioral and market-related factors may be necessary to fully explain investor returns.

One potential explanation for these residual patterns is the correlation between biases, which suggests that their effects may not be purely additive but interactive, similar to the BHB model 'interaction' or 'other' term. Certain biases, such as Sensation Seeking Bias and Extrapolation Bias, tend to co-occur, amplifying momentum and risk-seeking trading behavior, while others, such as Portfolio Concentration Bias and Familiarity Bias interact in shaping portfolio composition (Gorzon and Nitzsch 2025). Given these dependencies, a simple summation of bias effects may overstate their individual contributions to returns, as overlapping effects are not accounted for.

A possible refinement of the framework could involve interaction terms between biases, capturing how specific behavioral tendencies reinforce or counteract each other. This adjustment would help correct overestimation by subtracting overlapping effects, leading to a more accurate decomposition of bias-attributed returns. Additionally, it is possible that performance itself influences behavioral tendencies, e.g., prior success may lead to increased portfolio concentration or more frequent trading, rather than these behaviors causing outperformance. Our model does not rule out such reverse causality, and future research could explore both interaction effects and bidirectional relationships to improve the explanatory power and robustness of behavioral performance attribution. Incorporating nonlinear structures or dynamic bias profiles over time may further reduce residual patterns and strengthen attribution accuracy.

4.2.1 Overestimation of outperformance for winning investors

For winning investors, the model overestimates performance, as indicated by a median true residual of -2.22% (i.e., model-estimated returns exceed actual returns). This implies that while biases such as Action Bias and Portfolio Concentration Bias contribute to outperformance, additional return-reducing factors are missing from the framework.

A key limitation in this study is the exclusion of the Disposition Effect, because the trading data set did not include profit and loss positions on an individual trade level, a well-documented behavioral bias where investors sell winning stocks too

early and hold onto losing stocks too long (Shefrin and Statman 1985). However, prior research suggests that the Disposition Effect can lower realized returns by reducing exposure to long-term winners and increasing exposure to declining assets (Odean 1998). The omission of the Disposition Effect may partially explain the model's overestimation of winning investors' performance.

Beyond the Disposition Effect, one additional behavioral factor may include Market Sentiment Effects. Winners may be partially influenced by momentum trading and trend-following behavior, which is driven by collective investor sentiment rather than purely individual biases. The model currently does not account for broader market sentiment effects, e.g. a factor which quantifies if an investor's trades are related to broader market trading volumes (trends), that may influence retail investor performance (Baker and Wurgler 2006).

4.2.2 Overestimation of underperformance for losing investors

For losing investors, the model overestimates losses, as evidenced by a median true residual of +1.06% (i.e., model-estimated losses exceed actual losses). While the framework effectively captures well-documented biases such as Action Bias (excessive trading) and Sensation Seeking Bias (speculative investing)—which are strongly associated with underperformance (e.g. Barber and Odean 2000; Bali et al. 2017)—it may omit certain behavioral tendencies that help mitigate losses, leading to an overestimation of underperformance.

One potential missing factor is risk-aversion-induced caution, where some investors, after experiencing losses, instinctively shift to lower-risk strategies, such as reducing trading activity or reallocating capital to more stable assets. While biases like Action Bias contribute to excessive trading, not all losing investors continue to trade aggressively—some may react by scaling back risky behaviors, which softens their underperformance and is not currently accounted for in the model.

Additionally, bias interactions could contribute to the overestimation. Certain biases may counterbalance each other—for example, an investor with strong Sensation Seeking Bias may still exhibit elements of risk aversion in highly volatile conditions, leading to inconsistent trading behavior that the model does not fully capture.

A potential refinement to the model could involve introducing risk-reduction behaviors as an additional explanatory factor or incorporating interaction effects to adjust for the counterbalancing tendencies of certain biases.

4.3 Generalization of the behavioral attribution framework

The increase in adjusted R^2 and the Model Explainability Ratio (MER) after trimming provides clear evidence that the behavioral attribution model improves in predictive accuracy when extreme investors are excluded (Sect. 3.2) and therefore confirms the successful dynamic adaption of the framework to varying sub-datasets.

4.3.1 Bias attribution stability and tailored factor weights in subsamples

One of the most notable findings is that as trimming progresses, both the magnitudes and significance levels of bias coefficients change. This dynamic adjustment of the model suggests that certain biases play a more dominant role among extreme investors, while others are more relevant for average investors. Instead of applying a static attribution across all winners and all losers, the model adapts its factor weights (betas) to different investor subsets, making it more precise (tailored) for individual investor profiles and further confirming the general applicability of the behavioral attribution framework.

For winning investors, Action Bias and Portfolio Concentration Bias initially have strong positive coefficients (1.9351 and 0.2673, respectively), suggesting that high trading frequency and concentrated portfolios contribute to outperformance (Table 3). However, as trimming removes extreme performers, these effects weaken significantly, with Action Bias decreasing to 0.0306 at 20% trimming and Portfolio Concentration Bias nearly disappearing. This suggests that while heavy trading and portfolio concentration benefit some extreme winners, they are not sufficient predictors of success for the broader performance group. Conversely, Narrow Framing Bias remains significant throughout trimming, indicating that actively monitoring and adjusting portfolios is a consistent trait of successful investors across the entire performance group. Interestingly, Familiarity Bias—which is initially negative for the performance group of winning investors including extreme winners—loses its statistical significance and magnitude (coefficient) as trimming progresses. This suggests that over-reliance on domestic stocks especially hurts extreme winners in our defined performance group, while it is less relevant for low or moderate winners, e.g., for the sub sample excluding 1% extreme winners.

For losing investors, the results differ. Unlike winners—losing investors display more consistent and severe bias effects on portfolio returns across trimming levels (Table 4). This is consistent with prior literature, which finds that irrational behavioral biases tend to be more pronounced among unsuccessful investors (Barber and Odean 2000). Conversely to winners, being active in trading (Action Bias) and adjusting portfolios (Narrow Framing Bias) are detrimental behavioral tendencies for achieved portfolio performance. Instead, buy-and-hold strategies may provide a more effective approach for these investors e.g. studies have shown that long-term holding reduces transaction costs and mitigates the negative impact of emotional decision-making (Barber and Odean 2000) and that long-term stockholders benefit from compound returns, as the majority of market gains come from a small subset of outperforming stocks (Bessembinder 2018). By reducing exposure to biases such as Action Bias, Narrow Framing Bias and Sensation Seeking Bias, buy-and-hold strategies can serve as a countermeasure to the detrimental effects of excessive trading.

These findings underscore the adaptability of the Behavioral Performance Attribution framework, demonstrating that bias effects are not uniform across investors and that a dynamic, tailored approach enhances its explanatory power, making it a robust tool for analyzing retail investor behavior across diverse performance groups.

5 Conclusion

5.1 4.1. Summary and key findings

The findings of this study reinforce the significance of behavioral biases in shaping retail investor performance. The Behavioral Performance Attribution framework offers a novel approach to decomposing investment returns by quantifying the impact of cognitive biases, providing deeper insights beyond traditional allocation models like BHB. Our results confirm that behavioral biases are statistically significant for both winning and losing investors, with Action Bias and Portfolio Concentration Bias being significantly associated with performance outcomes. The model effectively captures these influences, as evidenced by its strong Model Explainability Ratio (MER). Furthermore, by applying trimming-based robustness checks and estimating separate models for different investor groups, we improved attribution precision and accounted for structural heterogeneity in behavior. Notably, losing investors display more persistent behavioral distortions, and buy-and-hold strategies may offer a more suitable alternative to mitigate the negative effects of excessive trading. Overall, the study highlights the importance of integrating behavioral factors into return attribution models to enhance (retail) investor awareness and decision-making.

5.2 Limitations

Despite its contributions, the study has several limitations. Methodologically, the absence of trade-level profit and loss data restricts the model from incorporating biases such as the Disposition Effect, which could further refine its explanatory power. Additionally, residual overestimation for both winning and losing investors suggests that certain behavioral and market-driven effects, such as market sentiment, and emotional trading under stress remain unaccounted for. Another route may be to further integrate interaction terms (between biases) into the model, similar to the BHB model. However, while implementing interaction terms would increase explainability, the questions remains whether it would still provide intuitive practical insights for investors. Similarly, the model does not include demographic or psychological controls, such as income, education, or risk preferences. This is a deliberate design choice aimed at maximizing applicability across real-world data environments where such variables are often unavailable (like in our data set). Future research with access to more granular investor profiles could test whether these factors add explanatory power beyond transaction-based behavioral proxies. Conceptually, while the framework is well-suited for retail investors, its applicability to institutional portfolio managers remains uncertain, given their more structured investment processes and regulatory constraints. Furthermore, the model assumes relatively stable bias exposures, whereas investor behavior may evolve dynamically over time. These limitations highlight the need for continued refinement and validation of the framework.

5.3 Implications for research

Future research should expand the application of the Behavioral Performance Attribution framework to diverse trading datasets to further validate its generalizability across different investor populations and market conditions. Incorporating additional behavioral biases, such as the Disposition Effect and overconfidence in recovery strategies, may improve the model's explanatory power. Moreover, integrating broader market factors, including sentiment analysis and macroeconomic influences, could enhance the precision of performance attribution. The potential for hybrid models that combine behavioral attribution with traditional return attribution frameworks, such as BHB, warrants further exploration to create a more comprehensive analytical approach. Finally, longitudinal studies assessing how investors' behavioral tendencies evolve over time could provide deeper insights into the stability and adaptability of bias effects.

5.4 Implications for practice

For retail investors, the findings emphasize the importance of recognizing and mitigating behavioral biases that negatively impact portfolio performance. Given the strong link between excessive (Action Bias), continuous trading (Narrow Framing Bias) and underperformance, adopting a disciplined buy-and-hold strategy may help investors reduce transaction costs and improve long-term returns. Additionally, increased self-awareness of biases like Sensation Seeking Bias and Action Bias can enable investors to implement structured decision-making processes, such as predefined rebalancing rules, to minimize impulsive trading.

For financial advisors, the Behavioral Performance Attribution model provides a valuable tool for identifying client biases and integrating behavioral insights into portfolio reporting. By incorporating bias-aware metrics into performance reviews, advisors can help clients recognize the financial consequences of irrational decision-making and implement tailored investment strategies that mitigate harmful tendencies. For example, clients with strong Familiarity Bias may benefit from more diversified investment strategies or automated trading systems that reduce emotional decision-making and excessive trading.

For brokerage houses serving self-directed investors, the Behavioral Performance Attribution model offers an innovative way to analyze trading behavior beyond standard portfolio metrics. By providing deeper insights into how cognitive biases impact performance, brokerage platforms can enhance investor reporting and decision-making support. Additionally, this approach can be extended into a bias early warning system, alerting investors when their trading patterns indicate a high likelihood of underperformance. Integrating real-time behavioral analytics into trading platforms enables brokers to offer personalized nudges or risk alerts, fostering greater investor awareness and more disciplined, long-term trading behavior.

For regulators and policymakers, the study highlights the need to enhance investor protection through behavioral finance-based interventions. Regulatory measures,

such as transparency requirements for trading costs and behavioral nudges that discourage exposure to detrimental behavioral tendencies, could help avoiding costly investment mistakes. Furthermore, financial literacy programs focusing on the impact of biases on decision-making may empower investors to make more rational and informed choices.

By integrating behavioral insights into investment strategies, advisory practices, and regulatory frameworks, the findings of this study provide actionable steps toward improving investor decision-making and financial well-being.

Author contribution D.G. wrote the main manuscript text and prepared all figures and exhibits. RvN reviewed the main manuscript text as well as all figures and exhibits.

Funding Open Access funding enabled and organized by Projekt DEAL.

Data availability No datasets were generated or analysed during the current study.

Declarations

Conflict of interest The authors have no conflict of interests to declare that are relevant to the content of this article.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Bacon, C.R.: Performance attribution: history and progress. *SSRN Electron. J.* (2019). <https://doi.org/10.2139/ssrn.3668966>
- Baker, M., Wurgler, J.: Investor sentiment and the cross-section of stock returns. *J. Finance* **61**, 1645–1680 (2006). <https://doi.org/10.1111/j.1540-6261.2006.00885.x>
- Bali, T.G., Brown, S.J., Murray, S., Tang, Y.: A lottery-demand-based explanation of the beta anomaly. *J. Financ. Quant. Anal.* **52**, 2369–2397 (2017). <https://doi.org/10.1017/S0022109017000928>
- Barber, B.M., Odean, T.: Trading is hazardous to your wealth: the common stock investment performance of individual investors. *J. Finance* **55**, 773–806 (2000). <https://doi.org/10.1111/0022-1082.00226>
- Barber, B.M., Odean, T.: The behavior of individual investors. In: Constantinides, G.M., Harris, M., Stulz, R.M. (eds.) *Handbook of the Economics of Finance*, pp. 1533–1570. Elsevier (2013). <https://doi.org/10.1016/B978-0-44-459406-8.00022-6>
- Bessembinder, H.: Do stocks outperform Treasury bills? *J. Financ. Econ.* **129**, 440–457 (2018). <https://doi.org/10.1016/j.jfineco.2018.06.004>
- Brinson, G.P., Hood, L.R., Beebower, G.L.: Determinants of Portfolio Performance. *Financ. Anal. J.* **42**, 39–44 (1986). <https://doi.org/10.2469/faj.v42.n4.39>
- Goetzmann, W.N., Kumar, A.: Equity portfolio diversification*. *Rev. Finance* **12**, 433–463 (2008). <https://doi.org/10.1093/rof/rfn005>
- Gorzon, D., Bormann, M., Von Nitzsch, R.: Measuring costly behavioral bias factors in portfolio management: a review. *Financ. Mark. Portf. Manag.* **38**, 265–295 (2024). <https://doi.org/10.1007/s11408-024-00444-7>

- Gorzon, D., von Nitzsch, R.: Measuring the effects of cognitive biases on retail investors' portfolio returns. *J Invest.* (2025). <https://doi.org/10.3905/joi.2025.1.357>
- Gujarati, D.N., Porter, D.C.: *Basic Econometrics*. 5th Edition, McGraw-Hill Irwin, Boston (2009). ISBN-10:9780071333450. ISBN-13:978-0071333450
- Minto, B.: *The pyramid principle: logic in writing and thinking*, 3rd edn. Financial Times/Prentice Hall, London (2002)
- Odean, T.: Are investors reluctant to realize their losses? *J. Finance* **53**, 1775–1798 (1998). <https://doi.org/10.1111/0022-1082.00072>
- Shefrin, H., Statman, M.: The disposition to sell winners too early and ride losers too long: theory and evidence. *J. Finance* **40**, 777–790 (1985). <https://doi.org/10.1111/j.1540-6261.1985.tb05002.x>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

David Gorzon holds a Ph.D. in Behavioral Finance from the Decision Theory and Financial Services Group at RWTH Aachen University. His research focuses on measuring and explaining the effects of cognitive biases on retail investors' portfolio performance, including the development of the Behavioral Performance Attribution (BPA) model.

Ruediger von Nitzsch is chair of the Decision Theory and Financial Services Group at RWTH Aachen University. His current research focuses on financial issues as well as supporting decision-making behavior by the use of professional online decision support tools.