BANCA D'ITALIA
EUROSISTEMA

Questioni di Economia e Finanza

(Occasional Papers)

The micro-determinants of portfolio gyrations in mutual funds: evidence from machine learning models

by Fabrizio Ferriani and Sabina Marchetti

Number 913

March 2025

BANCA D'ITALIA
EUROSISTEMA

Questioni di Economia e Finanza

(Occasional Papers)

The micro-determinants of portfolio gyrations in mutual funds: evidence from machine learning models

by Fabrizio Ferriani and Sabina Marchetti

Number 913 – March 2025

The series Occasional Papers *presents studies and documents on issues pertaining to the institutional tasks of the Bank of Italy and the Eurosystem. The* Occasional Papers *appear alongside the* Working Papers *series which are specifically aimed at providing original contributions to economic research.*

The Occasional Papers *include studies conducted within the Bank of Italy, sometimes in cooperation with the Eurosystem or other institutions. The views expressed in the studies are those of the authors and do not involve the responsibility of the institutions to which they belong.*

The series is available online at www.bancaditalia.it .

ISSN 1972-6643 (online)

Designed by the Printing and Publishing Division of the Bank of Italy

THE MICRO-DETERMINANTS OF PORTFOLIO GYRATIONS IN MUTUAL FUNDS: EVIDENCE FROM MACHINE LEARNING MODELS

by Fabrizio Ferriani* and Sabina Marchetti*

Abstract

We investigate the micro-determinants of portfolio gyrations in equity mutual funds that invest in emerging markets. Our analysis focuses on portfolio holding variations driven by asset managers' decisions, rather than by price revaluation, and matches this information with a comprehensive set of 54 stock-level characteristics. Using gradient boosting models (GBMs), we explore the non-linear relationships between stock characteristics and portfolio adjustments. Our findings show that firms' size and investment-related features, alongside equity stock attributes (e.g. market capitalization, traded volume, beta), are the most influential in explaining portfolio turnovers. Additionally, we provide evidence on how the relative importance of these characteristics shifts, based on sample partitions determined by market conditions (downturn vs recovery), investor type (institutional vs retail) and investment strategies (active vs passive).

JEL Classification: G01, G11, G14, G23.

Keywords: mutual funds, emerging markets, machine learning.

DOI: 10.32057/0.QEF.2025.913

---

* Bank of Italy, Directorate General for Economics, Statistics and Research.
Fabrizio Ferriani: fabrizio.ferriani@bancaditalia.it; Sabina Marchetti: sabina.marchetti@bancaditalia.it

1 Introduction[1]

The importance of mutual funds (MFs) in the global financial industry has significantly increased in recent years. According to data from ICI (2024), the total net assets managed by regulated open-end funds worldwide reached 68.9 trillion USD in 2023. This amount is equivalent to more than 135% of the US market capitalization at the end on 2023 and represents a growth of approximately 160% since the aftermath of the global financial crisis.

Several factors have contributed to the expansion of the MF industry. These include tighter banking regulations, which have shifted some financial intermediation to the non-banking financial sector, along with investors' appetite for innovative financial products and diversified investment strategies. Additionally, mutual funds provide relatively low-cost access to international markets, particularly in emerging markets. The growing participation of institutional, long-term-oriented investors has also played a role in sustaining this trend. However, these structural changes have not come without risks. Numerous academic studies and policymakers have highlighted potential vulnerabilities within the MF industry, such as liquidity mismatches due to daily redemptions, increasing concentration and similarities across MF portfolios, high interconnection with the broader financial market despite less intensive prudential oversight, and the absence of access to central bank liquidity facilities or other investor insurance schemes (see Jotikasthira et al., 2012, IMF, 2015, IOSCO, 2018, Fricke and Fricke, 2021, Falato et al., 2021, FSB, 2023 among many others).

The interplay between financial stability considerations and MFs is particularly relevant for emerging market economies (EMEs). These countries are generally more dependent on foreign capital, sensitive to global financial conditions - particularly monetary policy in advanced economies - and have local financial systems that are less resilient to shocks or are characterized by a less diverse local investor base (IMF, 2016; Cerutti et al., 2019; Koepke, 2019; BIS, 2021). During past episodes of financial turbulence and local shocks to emerging market economies, redemptions from MFs investing in EMEs have often been the most apparent sign of investors'rising risk aversion, exacerbating and

[1]We thank Alessio Anzuini, Raffaele Gallo, Marco Taboga and seminar participants at Banca d'Italia for their suggestions. All remaining errors are our owns

amplifying already tense market conditions (e.g., Gelos, 2011, Puy, 2016, Chari et al., 2021, Ferriani, 2021, Ferriani et al., 2023). The most recent and significant episode in this regard is the COVID-19 financial turmoil, during which portfolio outflows from EMEs exceeded 100 billion USD from January to March 2020, approximately 3.5% of the corresponding international investment positions (IMF, 2020, FSB, 2020, Eguren Martin et al., 2020, Ferriani and Natoli, 2021).

In this paper, we apply machine learning (ML) techniques to a novel dataset containing detailed information on the portfolio holdings of equity mutual funds investing in EMEs. Our aim is to identify the firms' characteristics that are most significant in explaining the variation in portfolio allocations. To achieve this, we first isolate the variation in portfolio holdings that can be attributed to asset managers' decisions, removing the component linked to market price changes. We combine this information with a comprehensive dataset of 54 variables, grouped into seven macro-categories, encompassing both high- and low-frequency firm characteristics. Our analysis spans the period from January 2019 to June 2020, covering not only the major financial turmoil caused by the Covid-19 pandemic in early 2020 and the subsequent market recovery following global policy interventions, but also a more stable period throughout 2019, when a favorable macroeconomic environment supported investor flows into EMEs. This also allows us to study whether portfolio gyrations by managers can be explained by different groups of firm characteristics across distinct market regimes.

For our study, we employ gradient boosting (GB), a machine learning approach capable of handling large datasets non-parametrically. This technique enhances model accuracy by recursively combining weaker models, such as simple decision trees. Given an information set, boosting methods assign more weight to more successful models, leading to a highly accurate final performance based on the combined ensemble of previous models in the sequence. Our findings reveal that variables related to size, investment, and equity stock attributes (e.g., market capitalization, traded volume, market beta, spread) are the most influential in explaining portfolio turnovers. The prominence of these variables may reflect, on one hand, behavioral tendencies in fund management, with managers basing their portfolio adjustments on dimensional factors that highlight well-established, large companies traditionally perceived as safer. On the other hand, the strong influence of attributes tied to trading activity and asset performance increases sensitivity to fluctuations, potentially amplifying portfolio turnover during periods of instability as managers shift their focus from fundamentals to ongoing market signals.

We conduct our analysis at both the full sample level and through several sample partitions, focusing on investment strategies (active vs. passive mutual funds), investor types (institutional vs. retail), and market conditions (downturn vs. recovery) to identify potential variations in asset managers' behavior. While variables related to size and market features consistently play a predominant role, we observe some heterogeneity across different sample partitions. Additionally, for a selected set of variables, we perform exercises based on variable permutation and local effects to demonstrate how the overall model performance responds to changes in specific variables.

In this paper, we aim to bridge two strands of literature. The first, as previously mentioned, relates to the analysis of MFs behavior in EMEs in response to monetary policy announcements or other episodes of market turmoil, with potential implications for financial stability and cross-country spillovers. We particularly focus on the Covid-19 period because of its significance in the recent history of market turmoil episodes. The validity of our research is nevertheless assured by the inclusion of adjacent time spans with different market conditions, and our approach can be generally extended to any time period. The second, more novel and rapidly growing strand of literature examines the role of machine learning (ML) and artificial intelligence (AI) techniques in identifying the most critical characteristics for portfolio asset selection, trading signal detection, asset pricing, and equity return forecasts (see, for example, Leippold et al., 2022, Li and Rossi, 2020, Kaniel et al., 2023, DeMiguel et al., 2023, Bonelli and Foucault, 2023, Zhang et al., 2023, Chen and Ren, 2022). These considerations have also recently entered the policy debate (IMF, 2024), raising concerns about potential financial stability risks stemming from the application of ML and AI in asset allocation and trading, such as increased turnover, higher asset correlations, volatility amplification, and operational and cyber risks.

To the best of our knowledge, this is the first paper to apply ML techniques to analyze the relationships between a large set of stock-level features and changes in the portfolio holdings of mutual funds.

Unlike previous studies (e.g., Li and Rossi, 2020), which focus on selecting high-performing mutual funds based on *average* fund-level exposures to a broad range of stock characteristics, our approach delves into the micro-determinants of portfolio rebalancing without aggregating at the portfolio level. ML techniques prove essential in this context, as they are better suited than conventional analytical tools to capture potential nonlinear relationships among variables and to handle large datasets with multiple determinants. In turn, this allows for a cleaner and more granular identification of the factors driving portfolio changes, providing deeper insights into asset managers' behavior, particularly during periods of financial market shocks. From a policy perspective, the growing adoption of ML techniques - and potentially generative AI - presents opportunities for supervisors and market surveillance to enhance risk assessment frameworks, develop vulnerability indicators, and improve performance driver analysis through multidimensional methods (IMF, 2024). Our study contributes to this debate by offering an empirical application that can be extended to other types of financial assets and integrated into supervisory monitoring frameworks, enabling a more data-driven approach to identifying emerging risks and mitigating procyclicality. This is particularly relevant for MFs investing in EMEs, where abrupt outflows or sudden shifts in fund managers' allocation patterns, driven by changes in investor risk aversion, can rapidly exacerbate market instability.

The rest of the paper is organized as follows. Section 2 introduces the dataset, Section 3 presents the machine learning technique adopted in the analysis, while Section 4 illustrates the main empirical results. Finally, Section 5 offers some concluding remarks and policy implications.

2 Data

We base our analysis on equity mutual funds investing in emerging market economies (EMEs), classified under the Morningstar category "Global Emerging Markets Equity," covering the period from January 2019 to June 2020. The use of mutual funds' portfolio holdings can significantly expand the dataset's dimension; our relatively short time span keeps it manageable while still offering a sufficiently broad and diverse sample to observe portfolio gyrations across different market conditions.

Similar to Li and Rossi (2020), we apply filters to the original sample, excluding funds with an average size smaller than 20 million USD, those with less than 80% of their portfolio invested in equities, and those with an inception date less than one year before the start of our sample. We then use Morningstar's database of historical holdings to obtain monthly data on the funds' portfolio holdings (quantities in terms of shares and relative weights). Our analysis focuses solely on the equity component, as we aim to match these holdings with firm characteristics to identify the micro-drivers of portfolio gyrations. For each fund, we limit the analysis to the first 150 positions in the monthly holdings. We believe this choice strengthens our analysis for several reasons. First, these funds typically exhibit concentrated portfolios with a relatively small set of equity stocks; in fact, the top 150 equity positions account for a median share of around 94% of total portfolio investments. Second, very small equity shares often correspond to smaller emerging market firms, making it difficult to retrieve a comprehensive list of corporate characteristics, which is crucial for this study. Lastly, focusing on very small holdings may inflate our sample with discontinuities and staggered values of share turnovers. Based on data availability, our final dataset consists of 762 unique MFs investing in 3,552 different stocks.

Each stock is uniquely identified by its ISIN, which is used to retrieve firm characteristics from LSEG. Our study is based on a comprehensive set of more than 50 stock-level characteristics; similar to Li and Rossi (2020) and Hou et al. (2020) we chose to group these characteristics into seven macro-categories: *company information, liquidity, market, profitability, solvency, size and investment, and rating*. The full list of variables is provided in Table A.1 in the Appendix and combines both high- and low-frequency characteristics (i.e., monthly, annual, time-invariant). We complement this dataset with fund-level information such as total assets under management, investment strategy (e.g., passive vs. active), and investor type (i.e., distinguishing between funds targeting either retail or professional investors). This information will be used in the empirical section to replicate the analysis on specific dataset partitions and explore potential heterogeneity in managers' behavior. Additionally, we use LSEG to obtain the average monthly stock price for each ISIN.[2]

[2]Results remain qualitatively unchanged when using the month-end closing price instead of the average.

Our final sample consists of more than 1.4 million observations, providing, for each fund and month, detailed information on the quantities and prices of each ISIN held in the fund's portfolio. We aim to study how variations in the fund's portfolio holdings can be explained by firms' characteristics. In other terms, we need to compute monthly net purchases carried out by each mutual fund and to this purpose we distinguish changes in portfolio holdings due to market price revaluation from those associated with managers' decision that represent actual financial transactions. Similar to Affinito and Santioni (2021), we measure the revaluation effect as follows:

$$rev_{j,i,t} = (p_{j,t} - p_{j,t-1}) * min(q_{i,j,t}, q_{i,j,t-1})$$

where rev refers to the revalution component of a generic share j included in the portfolio of fund i at month t, whereas p and q are respectively the average monthly price of stock j and its quantity held in the portfolio of fund i. We then measure the net purchase component by subtracting the price revaluation dimension from the overall portfolio variation associated to each asset:

$$tra_{j,i,t} = (p_{j,t} * q_{i,j,t} - p_{j,t-1} * q_{i,j,t-1}) - rev_{j,i,t}$$

where tra accounts for the part of portfolio gyrations that is associated with the fund managers' transactions and represents the main object of our investigation. In the following section, we will describe the ML techniques used to examine how variations in this component may be explained by firms' individual characteristics. The descriptive statistics for the target variable and for the other features used by our ML approach are available in Table A.2.

3 Methods

We use ensemble learning tree methods to analyze the main determinants of mutual funds' investment strategies. Ensemble learning refers to a class of ML approaches that combine multiple models to enhance predictive performance. Ensemble learning often considers regression trees as building blocks

HYPER-PARAMETER	CANDIDATE VALUES	BEST
Learning rate	Unif(0.001, 0.1)	0.009
Boosting rounds	[100, 150, ..., 500]	200
Number of leaves (max)	[5, 6, ..., 100]	52
Tree depth (max)	[3, 4, ..., 20]	18
Observations per child (min)	[1, 2, ..., 5]	1
Training instances per tree (sample ratio)	[0.3, 0.35, ..., 0.5]	0.4
Training features per tree (sample ratio)	Unif(0.1,0.3)	0.189
Training features per level (sample ratio)	Unif(0.5, 1.0)	0.848
L1 regularization parameter	Unif(0,1)	0.234
L2 regularization parameter	Unif(0,1)	0.617

Table 1: Optimal value of main hyper-parameters derived with k-fold cross-validation tuning routine, k=8. The number of leaves is key to control complexity in LightGBM. The learning strategy used by LightGBM would theoretically set the value equal to $2^{\text{maximum depth}}(= 128)$. However, practical performance often benefits from setting the value lower, to prevent overfitting and achieve better accuracy.

since they offer several advantages. Among others, they efficiently handle complex tabular datasets with multicollinearity, missing values, and outliers. Importantly, the training process of regression trees encompasses automatic feature selection, which can support the identification of the most influential factors driving investment decisions.

Ensemble learning approaches can be broadly classified into two categories: *boosting* or *bagging*. Boosting generates sequences of weak models, each learning from the residuals or errors of its predecessors to improve performance. Bagging, on the other hand, trains more complex models independently on random subsets of data and then combines their predictions. Our analysis relies on the well-established ensemble class of gradient boosting (GB) models Friedman (2001). GB methods iteratively partition the feature space by building a sequence of weak models, with each learning from the residuals of the previous one, progressively refining overall performance.[3]

GB methods excel at handling complex datasets, automatically performing feature selection, and achieving high predictive accuracy. However, they can be prone to overfitting if not carefully tuned.[4]

[3]A popular tree-based ensemble method based on bagging is the so-called Random Forest Breiman (2001), which aggregates deep uncorrelated trees. Random Forest tends to be less accurate than GB methods on complex datasets.

[4]Overfitting occurs when a model learns the training data too well, capturing noise and random fluctuations instead of underlying patterns. This leads to poor performance on new, unseen instances. The opposite of overfitting is underfitting, which occurs when the model fails to capture relevant patterns even in the training data.

Additionally, the iterative nature of GB methods can result in higher computational costs compared to other ensemble techniques. To mitigate these limitations, we employ LightGBM Ke et al. (2017), a GB framework designed to improve predictive performance and scalability for large datasets.

LightGBM leverages two main techniques during training: Gradient-Based One-Side Sampling (GOSS) and Exclusive Feature Bundling (EFB). GOSS accelerates training by focusing on data instances with larger gradients, whereas EFB efficiently handles categorical features. LightGBM differs from standard GB as it builds each tree in the sequence by iteratively selecting the leaf node with the largest potential gain at each step. This approach leads to deeper yet narrower trees compared to the traditional approach that builds the tree level by level.

While LightGBM is a powerful and efficient gradient boosting framework, other strong contenders exist in the field. Notable alternatives include XGBoost (Chen et al., 2015), that enhances performance accuracy by efficiently minimizing a regularized objective function but whose learning process relies on a level-wise splitting approach, and CatBoost (Prokhorenkova et al., 2018), which is suited to handle complex datasets with numerous categorical features without requiring extensive preprocessing.

We randomly split observations into a training set (80%) and hold out a test set for validation (20%), corresponding to 292,365 observations and 73,292 records respectively. Since our application aims to assess the main drivers of portfolio managers, we don't rely on a test set, to check out-of-sample performance but still leverage the validation set to enhance learning of externally valid representation rule of the feature set, i.e. to avoid overfitting. Our training strategy entails a tuning routine to define the optimal architecture of our GB model Hastie (2009). Tuning is carried out using k-fold cross-validation ($k = 8$) over a sparse grid of hyper-parameters.[5] We identify optimal hyper-parameter values by optimizing the quadratic loss function associated with each combination. Hyper-parameters include the number of boosting rounds, i.e. the maximum length sequence of trees, trees' characteristics, and regularization parameters (see Table 1). Given an optimal combination of hyper-parameters, we train our model. We finally validate its performance on the hold-out set.

[5] Our tuning routine randomly samples hyper-parameter values from their respective candidate set. We set the number of evaluation rounds equal to 1'000, i.e. 1'000 combinations of hyper-parameter values.

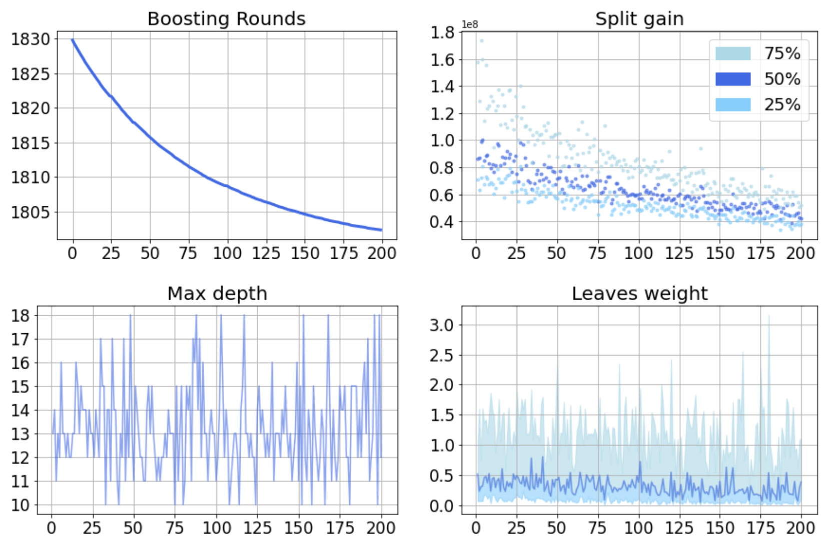


Figure 1: The figure displays basic training diagnostics for the baseline Gradient Boosting model. Top-panels show convergence of the L2 loss function optimized by learning routine (left); the gain associated with internal splits for each tree in the sequence (median, first and third quartile values; right). Bottom-panels show maximum depth of each tree in the sequence (left); the distribution of instances across leaves (median, first and third quartile percentage values; right).

We train our LGBM using GOSS boosting type for 200 boosting rounds.[6] Figure 1 shows our training strategy leads to a decreasing splitting gain trend throughout the boosting rounds, indicating effective learning and progressive refinment. Descreasing splitting gains associate with a slightly increasing pattern in tree depth, suggesting that the model was able to capture increasingly complex patterns in the data. The left-skewed distribution of leaf weights further supports the model's ability to prioritize informative splits and allocate appropriate weights to different decision paths.

4 Results

In this section, we present the main empirical results on the stock-level characteristics driving portfolio gyrations. We rely on established explainable artificial intelligence (XAI) tools to gain insights into the logical mechanism underlying the functioning of our general GB model, as well as its local specifications. XAI techniques can broadly be classified as either global or local: global methods

[6]GOSS boosting type efficiently handles imbalanced datasets by focusing on gradient information of instances with large gradients, thereby reducing computational cost and improving model performance.

provide insights into the overall behavior of a model, while local methods focus on understanding why a model made a specific decision in a given instance.

4.1 Variable importance

We begin by using global XAI techniques to examine variable importance (VI), which is a byproduct of tree-based models. VI is a quantitative measure that evaluates the relative contribution of each feature to the model's predictive performance. By scoring each feature based on its role in building each tree throughout the GB sequence, VI helps identify the variables with the most significant impact on the outcome. LightGBM provides two primary VI metrics: split *gain* and split *frequency* VIs. While both metrics can assess feature importance in ensemble tree methods, they capture different aspects. Split gain VI measures the reduction in impurity measures attributable to splitting on a particular feature across all trees in the ensemble. Split frequency VI, on the other hand, tracks how often a feature is chosen for splitting across all decision trees in the ensemble. Given our objective of identifying features most strongly associated with the target variable *tra* (quantity effect), we focus on split gain VI for our analysis.

Baseline results are displayed in Figure 2, which shows the top 20 stock-level characteristics in terms of VI. The values are normalized to sum to one to facilitate the interpretation of the relative importance of each characteristic; A complete graph of all stock-level characteristics is provided in Figure A.1 in the appendix. The top 20 most influential variables, representing less than half of the total number of covariates, account for a substantial share of the overall variable importance, over 67% Although there is some heterogeneity in terms of relative influence, no characteristic emerges as overwhelmingly more important than the others, and no covariate is entirely excluded by the GB model. The three most influential variables are reinvestment rate[7], market capitalization, and the debt to capital ratio, all with a VI of around 6.5%; the VI gradually decreases across the remaining stock characteristics reaching approximately 0.6-0.7% for the least influential ones (see Figure A.1), namely P/E ratio, total cash, and P/BV ratio. In Figure 3a we also present the breakdown of VI

[7] Reinvestment rate is computed by dividing retained earnings by common equity.

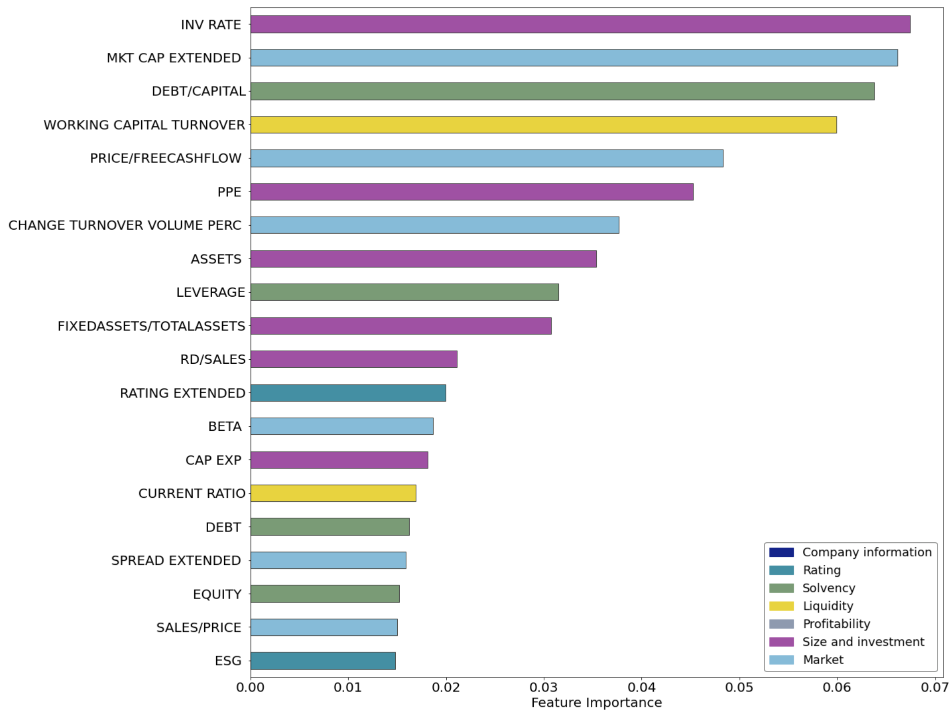
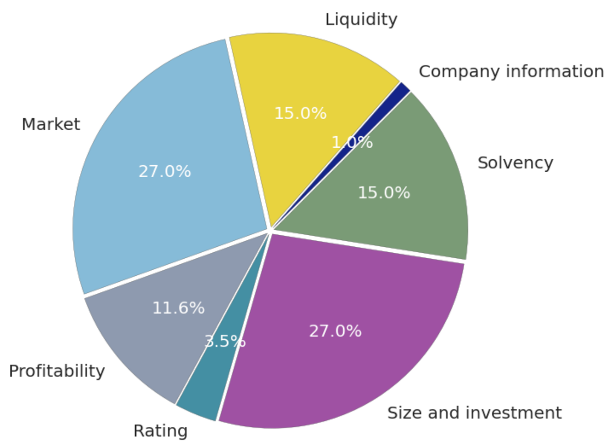


Figure 2: Figure reports the top-20 stock-level characteristics in terms of variable importance. The variable importance is normalized to sum to 1 across all 54 characteristics. Refer to Table A.1 for the full list of variables.

across the 7 categories we used to classify firms' characteristics (see Gu et al., 2020, Li and Rossi, 2020 for analogous classifications). Clearly, the VI breakdown across categories partly reflects the relative number of variables within each category (as shown in Table 3b). That said, the overall VI of variables related to size and investment, market, and solvency is more than proportional to their relative share, a result that is partly in line with the findings of Gu et al. (2020) and Leippold et al. (2022), who use machine learning techniques to identify predictive factors for asset premiums and returns. Conversely, the opposite is true for profitability and company-related information, where the overall VI is less than proportional to the share of variables in these categories.

To assess the feature importance of macro-categories of stock level characteristics we also rely on model agnostic XAI techniques. These methods are applicable to any type of model, providing a complementary perspective to model-specific approaches like VI. By comparing the VI rankings obtained from both model-specific and agnostic methods, we can enhance the robustness of our feature importance analysis. We base our analysis on permutation variable importance (PVI), which mea-



(a) Variable importance by category

Category	Share (%)
Company information	5.6%
Liquidity	16.7%
Market	24.1%
Profitability	18.5%
Rating	3.7%
Size and investment	20.4%
Solvency	11.1%

(b) Variable breakdown by category

Figure 3: Figure 3a displays the variable importance of the 54 stock-level characteristics classified in 7 categories: company information, liquidity, market, profitability, solvency, size and investment, and rating. Table 3b shows the share of variables included in each category.

sures the decrease in model performance when the values of a specific feature are randomly shuffled, thereby breaking the relationship between the feature and the target variable while preserving the marginal data distribution (Molnar, 2020).[8] A significant drop in the average performance indicates the relevance of a feature for model predictions. The results, shown in Figure 4, confirm that model performance is largely driven by variables in the market, and size and investment categories. At the feature level, dropping information on asset growth and change turnover volume is associated with most relevant deterioration in average performance.

4.2 Accumulated local effects

We then use accumulated local effects (ALEs) Apley and Zhu (2020) to gain insights on the extent and the direction of the impact of changes in input variables toward the output of our machine learning model. ALEs measure the marginal effect of each input variable on the model predictions,

[8]To avoid potential bias caused by specific random seeds, PVI repeatedly shuffles the values of one feature at a time and re-evaluates the model's performance over m iterations, $m = 100$ in our setup. One hot-encoded categorical features are excluded from the permutation, to avoid inconsistencies.

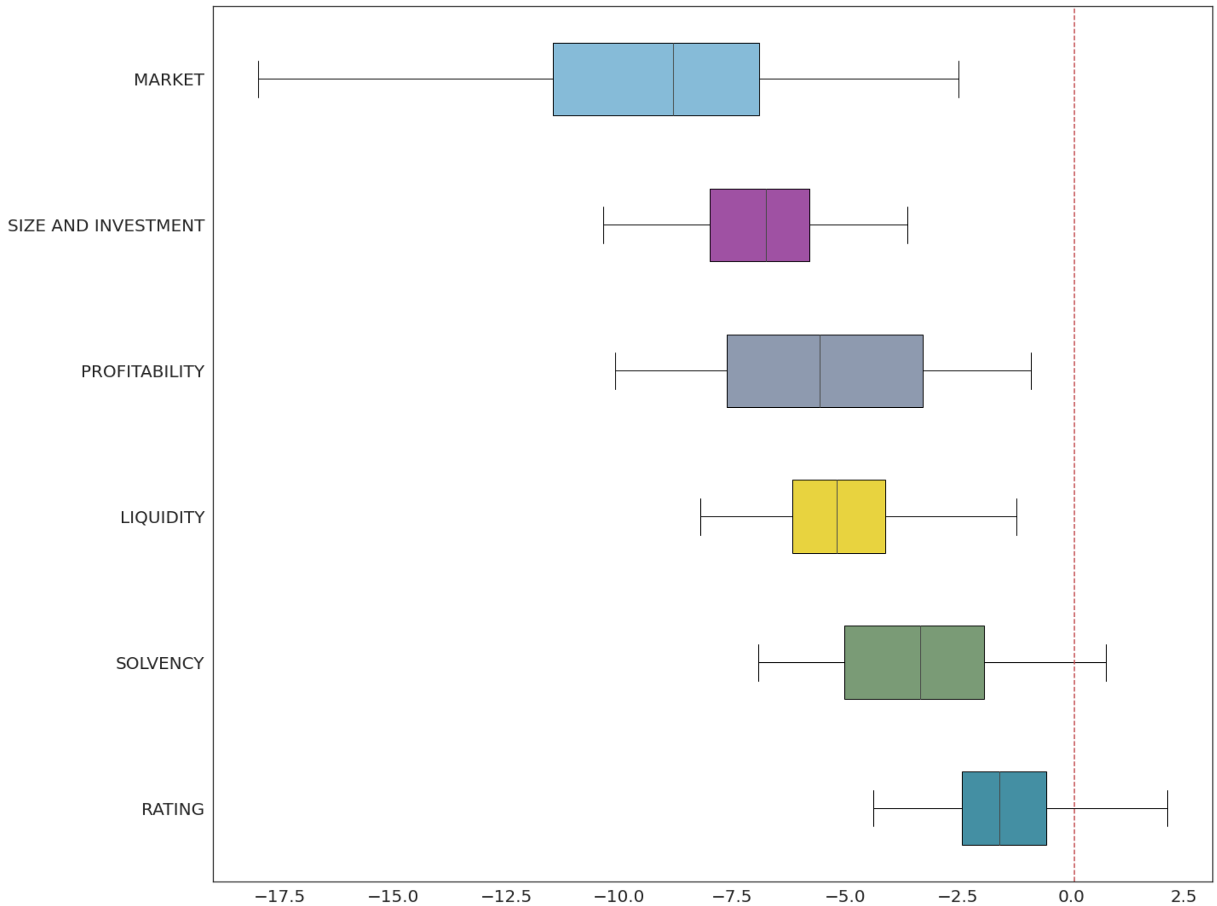


Figure 4: The figure displays the permutation variable importance (PVI) by categories of stock-level characteristics. Lower PVI values indicate a higher relevance of the variable category in explaining portfolio gyrations. Variables in the "Company Information" category are excluded from the PVI due to their categorical nature.

while holding all other variables constant. Let X_A be a continuous feature defined over domain $\mathcal{X}_A$ whose minimum observed value is $x_{A,0}$, and let $\mathbf{X}_{-A}$ be the set of all features except X_A. ALEs are derived for each value $x_A \in \mathcal{A}$ as:

$$f_{A,ALE}(x_A) = \int_{x_{A,0}}^{x_A} \mathbb{E}_{\mathbf{X}_{-A}} \left[f'(X_A, \mathbf{X}_{-A}) | X_A = x \right] dx - c$$

where f' is the local effect of x_A on f, i.e. the partial derivative in x_1 of f, the distribution mapping input features into model outputs, and c is a constant. In essence, $f_{A,ALE}(x_A)$ quantifies the deviation of the model's output from its average behavior when feature $X_A = x_A$. By accumulating

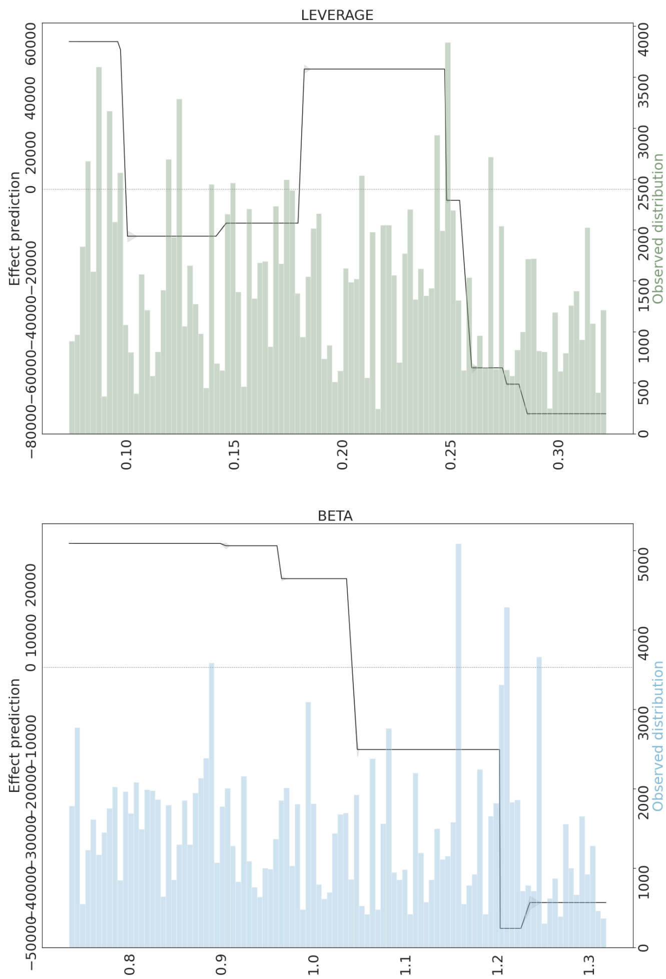


Figure 5: Figure reports the accumulated local effects (ALEs) for firm leverage (upper panel) and the stock beta (lower panel). Coloured bars represent the sample distribution of the analysed variables.

these deviations across different intervals of X_A, ALEs measure how model predictions change as the feature value varies within its domain range. This cumulative perspective offers insights into the feature's impact on the model predictions, including the direction of the effect. Furthermore, it enables to identify non-linear relationships between input and output variables.[9] Figure 5 presents the ALE plots for two representative variables, the leverage (upper plot) and the stock market beta (lower plot). Both plots are indicative of how ALE tools can be useful to detect and interpret nonlinearities

[9] An alternative approach providing insights similar to ALEs' is represented by partial dependence plots (PDPs). PDPs compute the marginal effect of a feature on the predicted outcome by averaging the model's predictions over all possible combinations of the other features. While PDPs offer a straightforward way to understand feature relationships, they don't properly take into account characteristics of the joint distribution of input features. As such, they average over unrealistic combinations of feature values and fail to provide reliable insights when features are correlated.

in model features. We observe a non-linear, piecewise relationship between firm leverage and our target variable. Positive portfolio gyrations are associated with extremely low levels of firm leverage or moderately high levels, suggesting that fund managers may prefer stocks perceived as extremely safe or offering a favorable balance between moderate risk and return. Conversely, at relatively low leverage levels, the association is slightly negative, indicating that stocks with such leverage might not be sufficiently remunerative to warrant inclusion in the portfolio. For firms with high leverage, the association becomes strongly negative, reflecting a tendency for fund managers to reduce portfolio exposure to companies perceived as excessively indebted and ultimately risky. The stock market beta also exhibits a non-linear yet substantially monotonic negative relationship, with some fluctuations. Positive portfolio gyrations are associated with relatively low beta values. However, as beta exceeds 1, the expected value of portfolio gyrations gradually decreases. Higher beta values, which occur more frequently than low or moderate ones, contribute to negative portfolio gyrations. This pattern suggests that fund managers tend to avoid firms that are highly responsive to market dynamics and economic cycles.

4.3 Sample partitions

As a final exercise, we split our trained model across multiple dimensions to gain valuable insights into local dynamics and identify potential variations in the role of different variables in explaining asset managers' behavior. Similar exercises, focused on identifying predictors for stock and mutual fund returns, have revealed some heterogeneity in the results, with variations emerging when accounting for time-varying dynamics and other sample partitions (DeMiguel et al., 2023, Leippold et al., 2022, Li and Rossi, 2020). As discussed in the Introduction, the use of ML techniques in this context can serve as a practical tool for supervisors and risk managers, helping to structure tailored indicators and perform analyses that account for the diverse characteristics and features within the mutual fund industry.

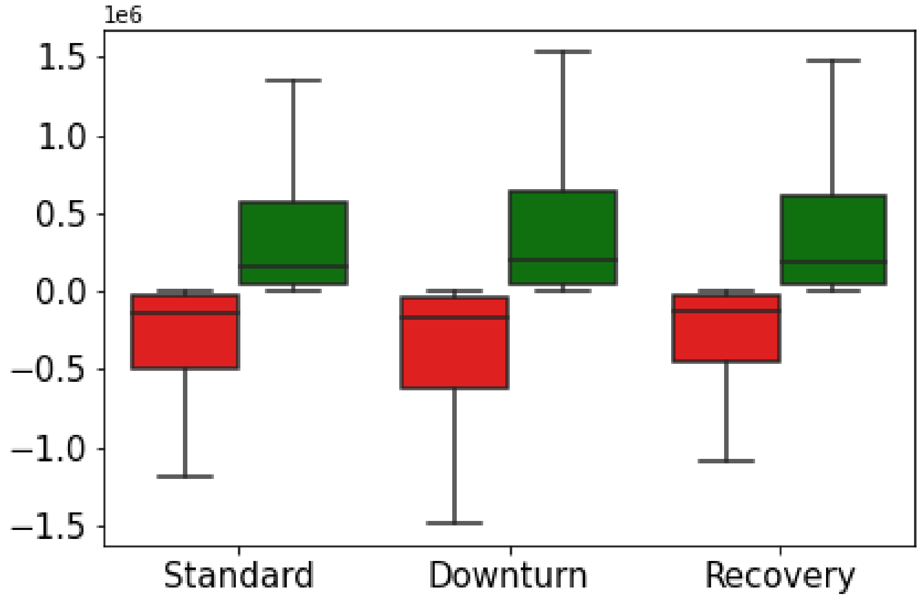


Figure 6: The figure displays the distribution of the dependent variable tra broken down by values (positive/negative) and time periods: standard, downturn, and recovery. The time periods are defined as follows: standard (Jan-2019-Jan 2020), downturn (Feb-Mar 2020), recovery (Apr-Jun 2020).

Partitions based on market regimes

The first source of heterogeneity we explore is time, as the VI of stock-level characteristics can evolve in response to changes in overall market conditions. To capture this, we partition the sample into three distinct temporal regimes and independently train the models for the standard period (Jan 2019-Jan 2020), the downturn linked to the Covid-19 crisis (Feb-Mar 2020), and the recovery phase (Apr-Jun 2020). Figure 6 shows the distribution of the dependent variable - portfolio gyrations associated with fund managers' transactions - categorized by regime and the positive/negative values of the variable of interest. Compared to the standard regime, the distribution of tra appears significantly more dispersed during the downturn period for both positive and negative values. This suggests that portfolio turnovers tend to reach more extreme levels during turbulent and volatile times. In contrast, during recovery periods, the distribution resembles that of the standard phase; however, the asymmetry between positive and negative values becomes more apparent, with positive portfolio gyrations exhibiting greater dispersion during favorable market conditions.

Figure 7 shows the VI by category across these three market regimes. Overall, our results are

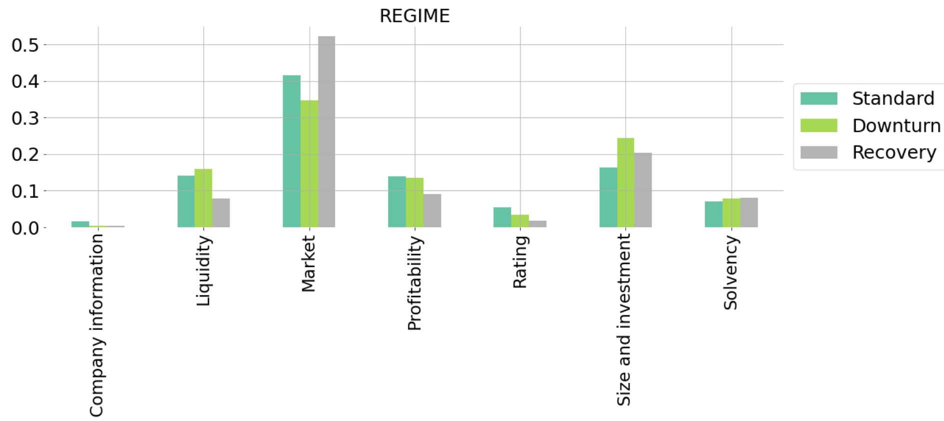


Figure 7: The figure displays the variable importance of equity stock characteristics broken down by time periods: standard, downturn, and recovery. The time periods are defined as follows: standard (Jan-2019-Jan 2020), downturn (Feb-Mar 2020), recovery (Apr-Jun 2020). Stock characteristics are grouped in macrocategories of interest, see Table A.1 for the whole list of variables.

consistent with the evidence presented in Figure 3a, though with some distinctions. Variables in the market category are most influential during tranquil market periods and the recovery phase, but their VI decreases significantly during the downturn phase. This reduction is partly offset by stock-level characteristics in the size and investment category, whose VI increases during the downturn. Variables related to firms' liquidity and profitability maintain a similar VI across both standard market phases and the downturn period, though their aggregate VI shrinks during the recovery phase, where managers may prefer portfolio adjustments driven by market conditions, following momentum and hoarding strategies. Finally, the VI assigned to ratings remains relatively small, but is generally higher during stable market conditions, when limited market developments likely lead asset managers to rely more on synthetic metrics.

Partions based on investment strategies and investor types

We then combine the temporal breakdown with an additional partition, splitting the sample based either on investment strategies (active vs. passive funds) or investor type (retail vs. institutional funds).[10] The VI results are displayed in Figure 8 for stock characteristics across time and investor

[10] An institutional fund is a fund that meet at least one of the following qualifications: it has the word "institutional" in its name; it has a minimum initial purchase of $100,000 or more; it states in its prospectus that it is designed for institutional investors or those purchasing on a fiduciary basis. Active funds accounts for approximately 23% of

type, and in Figure 9 for the breakdown by time and investment strategy .[11] Some key findings emerge. We observe that stock characteristics in the market category generally show a higher VI for retail mutual funds compared to institutional ones, particularly during the recovery period. Market variables also have the highest VI for institutional funds; however, portfolio turnovers for institutional funds seem to place more relative weight on features in the size and investment and solvency categories compared to retail mutual funds. Firm size generally exhibited a larger VI during the downturn period for both categories of mutual funds, indicating that the firms dimension played a major role in portfolio adjustments during the Covid-19 market collapse.

These results may be explained by the differing attitudes of the two types of investors (e.g. Ivković and Weisbenner, 2009, Evans and Fahlenbrach, 2012, Salganik-Shoshan, 2016, Goldstein et al., 2017 among others). Retail investors tend to be more sensitive to changes in market conditions, which may explain why market-related variables (e.g., market capitalization, volume, bid-ask spreads, market beta) play a more significant role in explaining portfolio gyrations for these investors. This greater reliance on market-driven variables can exacerbate volatility during downturns, amplifying market stress and contributing to financial instability in periods of market turmoil. In contrast, institutional investors typically follow long-term strategies that are less influenced by momentum trading and are more focused on long-term market fundamentals.

Regarding investment strategies, market variables explain a substantial portion of portfolio adjustments, particularly for passive funds, where these features account for around 50% of the VI, and nearly 60% during the market downturn. One possible explanation is that passive funds must replicate the composition of equity benchmarks, sometimes even synthetically through stock surrogates, as long as they comply with geographic or sector constraints. This could explain why, given these ex-ante limitations, managers' decisions are primarily driven by market conditions rather than firms' fundamentals. Conversely, for active funds, portfolio managers are more involved in stock selection, and as expected, we find that firm fundamentals-related variables generally exhibit a larger

our sample, compared to 77% for passive funds. Similarly, institutional funds make up 22% of the sample, while non-institutional funds represent the remaining 78%.

[11] Comparability with the estimates reported in Figure 3a and Figure 7 cannot be fully ensured, as the classification of funds by investor type and investment strategies is not always available.

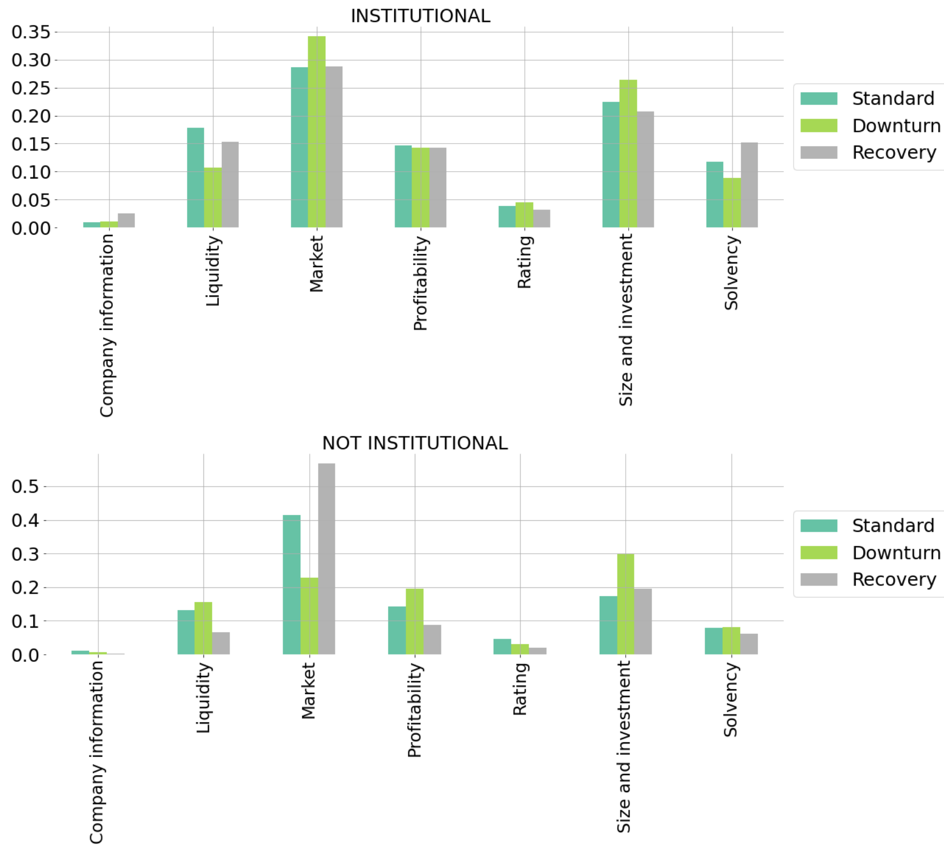


Figure 8: The figure displays the variable importance of equity stock characteristics broken down by time periods (standard, downturn, and recovery) and investor types (institutional vs retail mutual funds). The time periods are defined as follows: standard (Jan-2019-Jan 2020), downturn (Feb-Mar 2020), recovery (Apr-Jun 2020). Stock characteristics are grouped in macrocategories of interest, see Table A.1 for the whole list of variables.

VI compared to passive mutual funds.

5 Conclusions

We employ machine learning techniques to analyze the micro-determinants of portfolio gyrations in mutual funds. We focus on mutual funds investing in emerging markets, as these investments have frequently been at the epicenter of past episodes of financial instability and market turmoil. We isolate the portion of portfolio turnover directly attributable to managerial decisions and link it to a broad set of firm-level characteristics. Using gradient boosting models - a family of tree-based

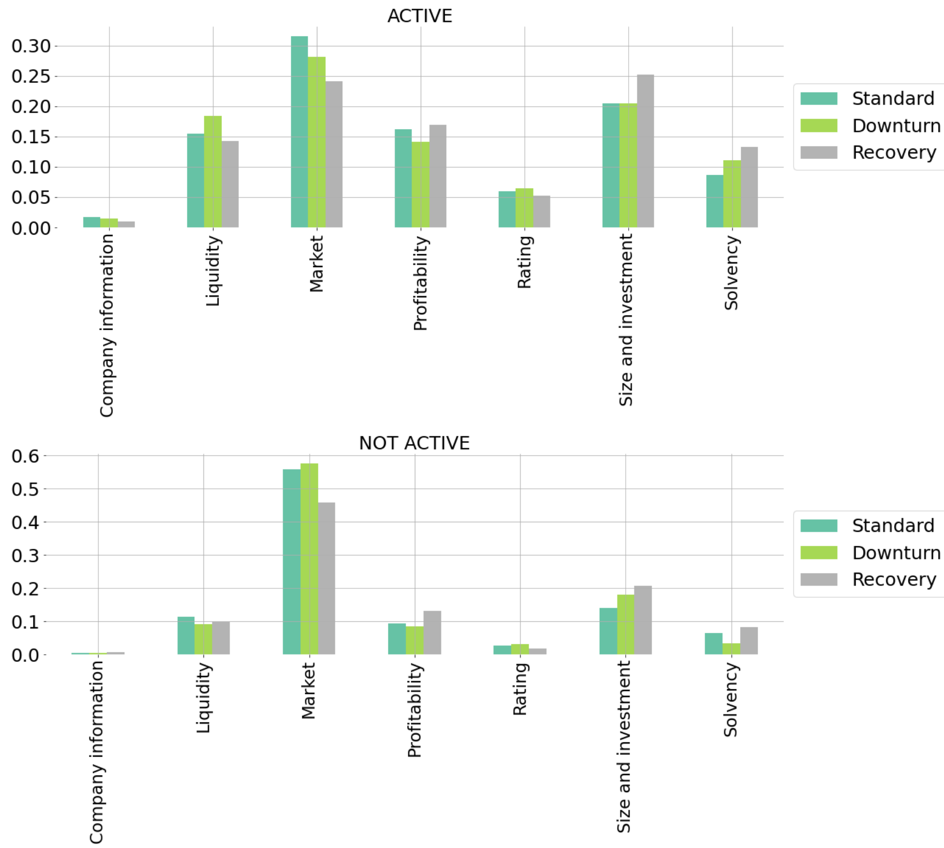


Figure 9: The figure displays the variable importance of equity stock characteristics broken down by time periods (standard, downturn, and recovery) and investment strategies (active vs passive mutual funds). The time periods are defined as follows: standard (Jan-2019-Jan 2020), downturn (Feb-Mar 2020), recovery (Apr-Jun 2020). Stock characteristics are grouped in macrocategories of interest, see Table A.1 for the whole list of variables.

methods - we shed light on the decision-making process of asset managers and identify the most influential stock-level features driving portfolio adjustments.

Our findings indicate that characteristics related to size and investment, along with equity stock attributes (e.g., market capitalization, traded volume, market beta, spread), exert the strongest influence on portfolio turnovers. This pattern suggests that portfolio allocation decisions are are significantly driven by proxies of corporate soundness (e.g., firm size) while also highlighting that changes in allocation are highly sensitive to trading activity and asset performance, potentially amplifying turnover volatility during periods of market turmoil. We also identify two key ways in which the relative importance of firm characteristics varies. First, we show that non-linear relationships exist between portfolio turnovers and stock-level variables, emphasizing the advantages of flexible, data-driven methods such as ML in capturing these nonlinearities. Second, our results reveal that the significance of specific stock-level variables is not fixed; rather, it shifts in response to market regimes and other sample partitions, including funds' investment strategies and targeted investor types.

This study offers a first application of machine learning techniques to analyze portfolio gyrations and can be extended to comparable investments (e.g., funds investing in advanced economies or bond funds). In this regard, machine learning techniques can provide insights into fund manager behavior, demonstrating their potential for practical use in risk monitoring and supervision (IMF, 2024). Mutual funds, given their wide heterogeneity in terms of asset portfolios, investor composition, and redemption schemes, offer an ideal field for such applications. The ability of machine learning models to handle large, complex datasets and provide granular insights into asset manager behavior presents a valuable opportunity for data-driven policymaking and the development of tailored risk monitoring frameworks. These frameworks can dynamically capture the non-linear relationships between a broad set of holding characteristics and changes in mutual fund portfolios. Supervisors could then apply machine learning-based tools to devise targeted indicators and detect potential imbalances in portfolio characteristics. This, in turn, could enhance supervisory frameworks and help anticipate shifts in asset manager behavior, enabling more proactive interventions.

References

Affinito, M. and Santioni, R. (2021), “When the panic broke out: COVID-19 and investment funds’ portfolio rebalancing around the world,” *Bank of Italy Working Paper*, 1342.

Apley, D. W. and Zhu, J. (2020), “Visualizing the effects of predictor variables in black box supervised learning models,” *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 82, 1059–1086.

BIS (2021), “Changing patterns of capital flows,” *CGFS report*.

Bonelli, M. and Foucault, T. (2023), “Displaced by Big Data: Evidence from Active Fund Managers,” *Available at SSRN 4527672*.

Breiman, L. (2001), “Random forests,” *Machine learning*, 45, 5–32.

Cerutti, E., Claessens, S., and Puy, D. (2019), “Push factors and capital flows to emerging markets: why knowing your lender matters more than fundamentals,” *Journal of international economics*, 119, 133–149.

Chari, A., Dilts Stedman, K., and Lundblad, C. (2021), “Taper tantrums: Quantitative easing, its aftermath, and emerging market capital flows,” *The Review of Financial Studies*, 34, 1445–1508.

Chen, R. and Ren, J. (2022), “Do AI-powered mutual funds perform better?” *Finance Research Letters*, 47, 102616.

Chen, T., He, T., Benesty, M., Khotilovich, V., Tang, Y., Cho, H., Chen, K., Mitchell, R., Cano, I., Zhou, T., et al. (2015), “Xgboost: extreme gradient boosting,” *R package version 0.4-2*, 1, 1–4.

DeMiguel, V., Gil-Bazo, J., Nogales, F. J., and Santos, A. A. (2023), “Machine learning and fund characteristics help to select mutual funds with positive alpha,” *Journal of Financial Economics*, 150, 103737.

Eguren Martin, F., Joy, M., Maurini, C., Nispi Landi, V., Schiavone, A., van Hombeeck, C., and Moro, A. (2020), "Capital flows during the pandemic: lessons for a more resilient international financial architecture," *Bank of Italy Occasional Paper.*

Evans, R. B. and Fahlenbrach, R. (2012), "Institutional investors and mutual fund governance: Evidence from retail–institutional fund twins," *The Review of Financial Studies*, 25, 3530–3571.

Falato, A., Goldstein, I., and Hortaçsu, A. (2021), "Financial fragility in the COVID-19 crisis: The case of investment funds in corporate bond markets," *Journal of Monetary Economics*, 123, 35–52.

Ferriani, F. (2021), "From taper tantrum to Covid-19: Portfolio flows to emerging markets in periods of stress," *Journal of international financial markets, institutions and money*, 74, 101391.

Ferriani, F. and Natoli, F. (2021), "ESG risks in times of Covid-19," *Applied Economics Letters*, 28, 1537–1541.

Ferriani, F., Gazzani, A., and Natoli, F. (2023), "Flight to climatic safety: local natural disasters and global portfolio flows," *Bank of Italy Temi di Discussione (Working Paper) No*, 1420.

Fricke, C. and Fricke, D. (2021), "Vulnerable asset management? The case of mutual funds," *Journal of Financial Stability*, 52, 100800.

Friedman, J. H. (2001), "Greedy function approximation: a gradient boosting machine," *Annals of statistics*, pp. 1189–1232.

FSB (2023), "Revised Policy Recommendations to Address Structural Vulnerabilities from Liquidity Mismatch in Open-Ended Funds," .

FSB, F. S. B. (2020), "Holistic Review of the March Market Turmoil," .

Gelos, M. R. (2011), "International mutual funds, capital flow volatility, and contagion–a survey," .

Goldstein, I., Jiang, H., and Ng, D. T. (2017), "Investor flows and fragility in corporate bond funds," *Journal of Financial Economics*, 126, 592–613.

Gu, S., Kelly, B., and Xiu, D. (2020), "Empirical asset pricing via machine learning," *The Review of Financial Studies*, 33, 2223–2273.

Hastie, T. (2009), "The elements of statistical learning: data mining, inference, and prediction," .

Hou, K., Xue, C., and Zhang, L. (2020), "Replicating anomalies," *The Review of financial studies*, 33, 2019–2133.

ICI (2024), "Investment Company Fact Book," .

IMF (2015), "The asset management industry and financial stability," *Global Financial Stability Report*.

IMF (2016), "The growing importance of financial spillovers from emerging market economies," *Global Financial Stability Report*.

IMF (2020), "managing volatile portfolio flows," *Global Financial Stability Report*.

IMF (2024), "Advances in artificial intelligence: implications for capital market activities," *Global Financial Stability Report*.

IOSCO (2018), "Recommendations for Liquidity Risk Management for Collective Investment Schemes," .

Ivković, Z. and Weisbenner, S. (2009), "Individual investor mutual fund flows," *Journal of Financial Economics*, 92, 223–237.

Jotikasthira, C., Lundblad, C., and Ramadorai, T. (2012), "Asset fire sales and purchases and the international transmission of funding shocks," *The Journal of Finance*, 67, 2015–2050.

Kaniel, R., Lin, Z., Pelger, M., and Van Nieuwerburgh, S. (2023), "Machine-learning the skill of mutual fund managers," *Journal of Financial Economics*, 150, 94–138.

Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., and Liu, T.-Y. (2017), "Lightgbm: A highly efficient gradient boosting decision tree," *Advances in neural information processing systems*, 30.

Koepke, R. (2019), "What drives capital flows to emerging markets? A survey of the empirical literature," *Journal of Economic Surveys*, 33, 516–540.

Leippold, M., Wang, Q., and Zhou, W. (2022), "Machine learning in the Chinese stock market," *Journal of Financial Economics*, 145, 64–82.

Li, B. and Rossi, A. G. (2020), "Selecting mutual funds from the stocks they hold: A machine learning approach," *Available at SSRN 3737667.*

Molnar, C. (2020), *Interpretable machine learning*, Lulu. com.

Prokhorenkova, L., Gusev, G., Vorobev, A., Dorogush, A. V., and Gulin, A. (2018), "CatBoost: unbiased boosting with categorical features," *Advances in neural information processing systems*, 31.

Puy, D. (2016), "Mutual funds flows and the geography of contagion," *Journal of International Money and Finance*, 60, 73–93.

Salganik-Shoshan, G. (2016), "Investment flows: Retail versus institutional mutual funds," *Journal of Asset Management*, 17, 34–44.

Zhang, J., Peng, Z., Zeng, Y., and Yang, H. (2023), "Do big data mutual funds outperform?" *Journal of International Financial Markets, Institutions and Money*, 88, 101842.

A List of variables

1	Acid test	Liquidity	28	Operating margin	Profitability
2	Asset growth	Size & investment	29	P/E ratio	Market
3	Total assets	Size & investment	30	PPE	Size & investment
4	Beta	Market	31	Price/Cash flow	Market
5	Capital expenditure	Size & investment	32	Price/Free cash flow	Market
6	Capital expenditure/Revenues	Size & investment	33	Price to book ratio	Market
7	Cash flow/Operating income	Profitability	34	Credit rating	Rating
8	Cash	Liquidity	35	R&D/Market value	Size & investment
9	Cash ratio	Liquidity	36	R&D/Sales	Size & investment
10	Cash flow/Total debt	Solvency	37	Analysts recommendation	Market
11	Change in turnover	Market	38	Revenues	Profitability
12	Current ratio	Liquidity	39	ROA	Profitability
13	Total debt	Solvency	40	ROCE	Profitability
14	Debt/capital	Solvency	41	ROE	Profitability
15	Dividend yield	Profitability	42	ROIC	Profitability
16	EBIT/Total assets	Profitability	43	Revenues/Total assets	Size & investment
17	Total equity	Solvency	44	Revenues/Cash	Liquidity
18	ESG score	Rating	45	Sales/Price	Market
19	Fixed assets turnover	Size & investment	46	Revenues/Receivables	Liquidity
20	Fixed assets/Total assets	Size & investment	47	Bid-Ask spread	Market
21	Interest coverage ratio (ICR)	Solvency	48	Industrial sector	Company information
22	Investment rate	Size & investment	49	State owned enterprise	Company information
23	Leverage	Solvency	50	Country of headquarters	Company information
24	Market value	Market	51	Turnover	Market
25	Momentum	Market	52	Price volatility	Market
26	Operating income	Profitability	53	Working capital turnover	Liquidity
27	Operating cash flow ratio	Liquidity	54	Working capital/Total assets	Liquidity

Table A.1: The table presents the list of variables

Variable	Mean	Standard Deviation
Portfolio gyrations by FMs' transactions	69'209.86	995'967.46
Acid test	0.76	1.59
Asset growth	0.03	5.25
Total assets (USD billion)	36.29	64.71
Beta	1.02	0.48
Capital expenditure (USD billion)	-2.39	5.56
Capital expenditure/Revenues	-0.13	1.08
Cash flow/Operating income	1.39	3.40
Cash (USD billion)	1.50	3.83
Cash ratio	0.25	0.68
Cash flow/Total debt	81.88	2'299.46
Change in turnover	0.19	2.25
Current ratio	1.87	1.91
Total debt (USD billion)	7.30	14.43
Debt/capital	0.67	2.56
Dividend yield	3.39	5.33
EBIT/Total assets	0.11	0.09
Total equity (USD billion)	17.63	37.19
ESG score	53.03	18.81
Fixed assets turnover	48.25	3'128.48
Fixed assets/Total assets	0.30	0.22
Interest coverage ratio (ICR)	1'062.40	47'127.11
Investment rate	6.68	108.02
Leverage	0.20	0.16
Market value (USD billion)	-2.39	5.56
Momentum	0.77	15.81
Operating income (USD billion)	3.41	7.87
Operating cash flow ratio	0.65	3.10
Operating margin	-0.04	13.86
P/E ratio	35.20	371.74
PPE (USD billion)	13.18	32.71
Price/Cash flow	1.78 1e-5	12.90 1e-5
Price/Free cash flow	1.04 1e-5	6.86 1e-5
Price to book ratio	4.06	12.43
R&D/Market value	41'523.55	150'236.66
R&D/Sales	0.31	12.84
Analysts recommendation	66.12	25.68
Revenues (USD billion)	22.53	48.32
ROA	8.95	9.87
ROCE	0.16	0.15
ROE	16.44	156.44
ROIC	13.16	14.90
Revenues/Total assets	0.76	0.56
Revenues/Cash	135'833.50	3'396'406.91
Sales/Price	2.28 1e6	202.86 1e6
Revenues/Receivables	12.87	50.64
Bid-Ask spread	0.49	3.38
Turnover	48.23	3'127.67
Price volatility	0.33	0.15
Working capital turnover	4.49	293.62
Working capital/Total assets	0.16	0.18

Table A.2: The table presents the descriptive statistics for our dataset.

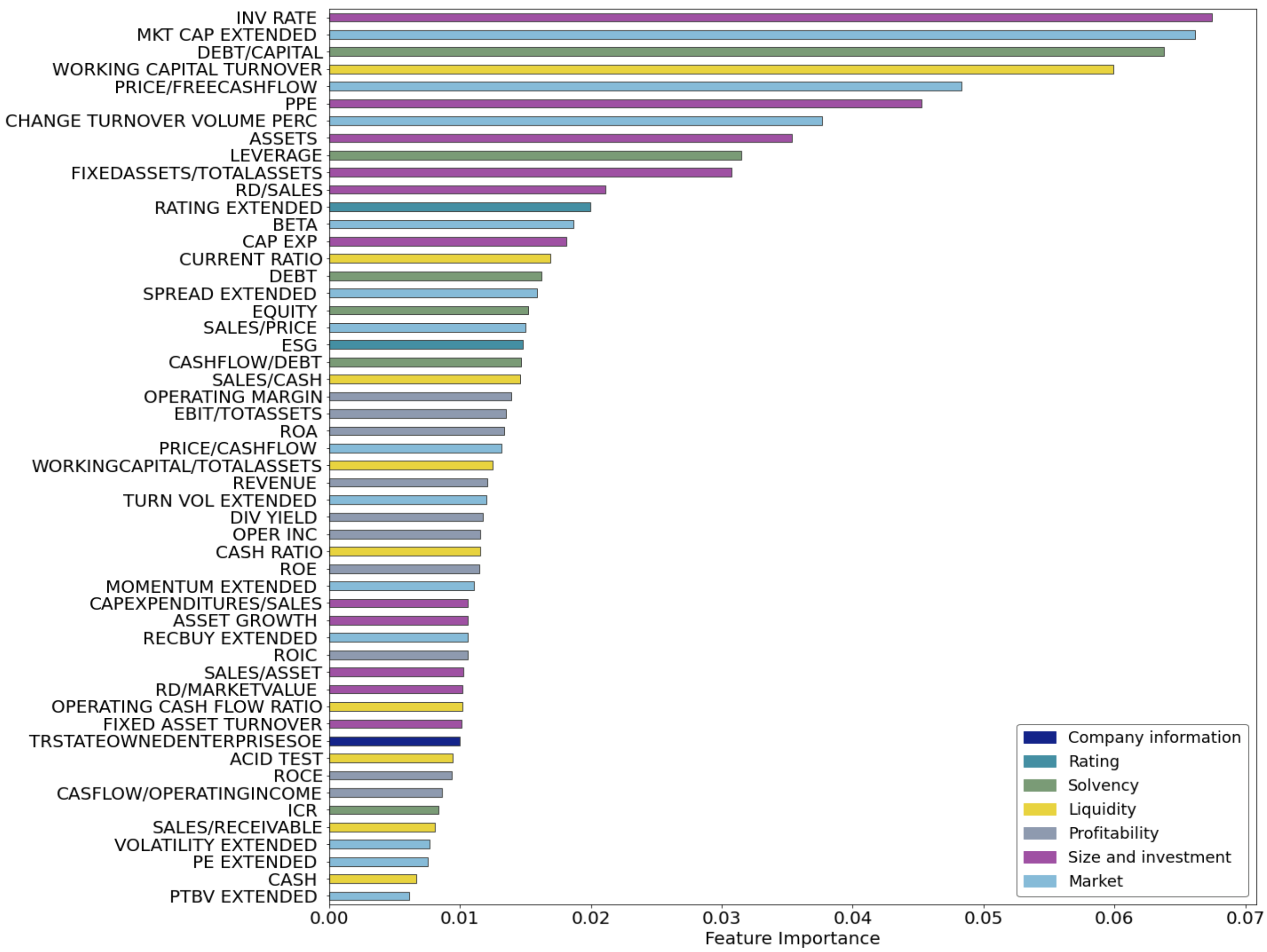


Figure A.1: Figure ranks stock-level characteristics in terms of variable importance. The variable importance is normalized to sum to 1 across all 54 characteristics. Refer to Table A.1 for the full list of variables.