Trust Anomalies' Judgment: Social Ranking Theories for Portfolio Construction

Veronica Guidetti[a], Alessandra Insana[b,*]

[a] *Department of Physical, Computer and Mathematical Sciences, University of Modena and Reggio Emilia, Via G. Campi 213/b, Modena, 41126, Italy*

[b] *Department of Economics, University of Messina, Piazza Pugliatti 1, Messina, 98122, Italy*

Abstract

This paper presents a novel approach to anomaly-based portfolio construction, integrating multiple market anomalies into a unified selection framework. We adapt the Majority Judgment (MJ) method from social choice theory to financial decision-making, ranking stocks based on multiple anomalies, each treated as an independent evaluation criterion.

Using 23 years of U.S. equity data and 11 stock anomalies, we construct equal- and value-weighted decile portfolios based on both single-factor breakpoints and MJ rankings. Our empirical results show that MJ-based portfolios consistently outperform single-factor strategies across most configurations and remain robust to changes in reallocation schedules, voter count, and voting system specifications. Comparative experiments highlight that MJ is more resilient to portfolio weighting choices than conventional anomaly aggregation methods and achieves the strongest performance under annual rebalancing. These findings underscore MJ's robustness and scalability in high-dimensional financial decision-making, offering a transparent and interpretable alternative for multi-factor portfolio construction.

Keywords: Multiple criteria analysis, Social ranking functions, Majority judgment, Portfolio selection, Market anomalies.

1. Introduction

Asset pricing and portfolio management pose significant challenges revolving around the strategic selection and aggregation of stocks based on predefined criteria. The fundamental objective of investment decision-making is to maximize expected returns while prudently managing risk. Achieving this balance requires a deep understanding of asset pricing models, which are essential for analyzing price behavior and identifying predictive factors. These factors serve two key purposes: forecasting portfolio returns and elucidating the relationship between returns and firm-specific characteristics, thereby uncovering market anomalies.

Stock sorting has been a popular approach to explore and exploit market anomalies. By categorizing stocks according to a single characteristic (e.g., size or value), investors construct five or ten portfolios, where the average excess returns often exhibit a systematic upward or downward trend, providing insight into the anomaly. The identification of numerous anomalies over the past 20 years has resulted in a veritable "factor zoo" (Cochrane, 2011). Over time, these have been explored, leading to the identification of more than 300 factors (Harvey et al., 2016; Green et al., 2017; Hou et al., 2018; Feng et al., 2020). In response to these findings, investors continuously refine their strategies to harness these insights for improved portfolio performance.

In recent years, different studies aimed to combine asset anomalies for portfolio construction. Indeed, as Novy-Marx (2016) and Stambaugh and Yuan (2017) pointed out, aggregating factor anomalies can mitigate noise in mispricing measures, offering a more comprehensive and refined investment approach. Common aggregation techniques include z-score aggregation (Asness et al., 2019), the linear combination of multiple signals (Novy-Marx, 2016), aggregate rank mean

*Corresponding author
Email addresses: veronica.guidetti@unimore.it (Veronica Guidetti), ainsana@unime.it (Alessandra Insana)

(Stambaugh et al., 2015; Stambaugh and Yuan, 2017; Dong et al., 2022), and optimization-based weight aggregation (Reschenhofer, 2024). Since these methods aim to derive a single global ranking from a diverse set of pricing anomalies, the problem naturally aligns with multi-criteria decision-making (MCDM).

In financial management, MCDM models are particularly valuable for complex tasks such as stock ranking and selection, where multiple, often conflicting, criteria must be considered. Building on Markowitz's foundational work on modern portfolio theory (Markowitz, 1952), MCDM methods have been widely applied to various financial domains, including portfolio optimization, corporate and asset evaluation, and credit risk estimation (Almeida-Filho et al., 2021). These techniques range from standard MCDM approaches for portfolio selection (Pätäri et al., 2018; Marqués et al., 2020), to metaheuristic algorithms for multi-criteria portfolio optimization (Ehrgott et al., 2004; Erwin and Engelbrecht, 2023; Xidonas et al., 2011), and expert systems designed to incorporate user preferences or enhance robustness in portfolio construction (Fasanghari and Montazer, 2010; Lin et al., 2024). In recent years, there has been a shift towards more data-driven approaches to portfolio management derived from machine learning techniques (Gu et al., 2020; Mirete-Ferrer et al., 2022). A prominent example is provided by Learning to Rank (LTR) methods as in Poh et al. (2022), which have been applied in the field of stock selection and ranking.

Classical and modern MCDM approaches typically involve creating a decision matrix where each row corresponds to an alternative under consideration, and each column corresponds to a specific criterion. To compare different criteria fairly, the matrix must be normalized to eliminate units of measure. The selection (or ranking) of solutions is then based on a utility function that maps multidimensional points to a scalar value, representing the distance or similarity between candidates and ideal solutions. However, using a distance function in high-dimensional spaces is challenging due to the curse of dimensionality, where distance loses its meaning due to distance concentration effects (Aggarwal et al., 2001). Additionally, requiring the objective space to be equipped with a metric is a stringent and often unreasonable condition. Candidate ranking and numerical optimization methods in high-dimensional spaces can also become highly sensitive to model specifications, where small changes in input and output can significantly affect efficiency scores. This is the reason why most MCDM applications consider a limited number of criteria/objectives (Ehrgott et al., 2004; Fasanghari and Montazer, 2010; Xidonas et al., 2011; Pätäri et al., 2018; Lin et al., 2024). For example, Pätäri et al. (2018) showed that metric-based MCDM methods for studying pricing anomalies performed satisfactorily only in low-dimensional spaces (up to four criteria).

Nevertheless, given the increasing complexity of financial markets and the proliferation of market anomalies, the need for scalable MCDMs is getting more and more compelling. On top of the plenitude of factors to be taken into account, the high-risk profile of financial applications requires scalability to be accompanied by robustness, interpretability of the method, and transparency of results. These last requirements can hardly be provided by black-box machine learning models where the overwhelming number of parameters make the model decision-making process opaque, hindering practical deployment. This is precisely why relatively simple scoring methods for anomaly aggregation, such as z-score and mean rank, though heuristic, have become mainstays among practitioners.

To overcome the aforementioned limitations, in this work we adapt social ranking theory for multi-criteria portfolio construction, focusing specifically on the Majority Judgment (MJ) method proposed by Inada (1969) and Balinski and Laraki (2010). MJ, often used when decisions hinge on the majority's preferences or votes, stands out for its clarity and axiomatic nature. It employs a lexicographic function to aggregate individual judgments expressed in a common language of grades, with the collective preference for each alternative determined by the median of the grades received. Building on this framework, we introduce an MJ-based portfolio management strategy that combines multiple market anomalies (e.g., size, momentum, and book-to-market) to refine portfolio selection. Since MJ was originally designed for static voting contexts, we extend it to a financial setting characterized by continuous, time-varying market anomaly measures. First, we propose a universal grading criterion that converts continuous anomaly indicators into categorical grades, ensuring the method's applicability to typical portfolio management data. Next, we introduce three different methods to apply MJ on rolling window data: rank rolling, profile rolling, and vote rolling.

We evaluate our method on a comprehensive dataset of U.S. equities from January 2000 to December 2023 and eleven widely recognized stock anomalies. In our first experiment, designed to assess the effectiveness of MJ for anomaly aggregation, we adopt an annual reallocation strategy in June. Specifically, we form decile portfolios using individual anomalies or the MJ ranking as breakpoints. To account for the size effect, we employ both equal- and value-weighted strategies. The results indicate that MJ portfolios consistently outperform those constructed from single anomalies under most configurations. Next, we incorporate the rolling MJ extension and perform robustness checks by varying the reallocation schedule, rolling window length, grading systems, and the set of anomalies used. These tests confirm the reliability and consistency of the MJ method across different experimental settings, with the profile rolling approach demonstrating particular resilience. Finally, we compare our results to classical aggregation techniques commonly used in portfolio management (i.e., z-score and mean-rank). Our findings reveal that MJ is more robust to variations in portfolio weighting schemes and achieves superior performance under annual reallocation.

To our knowledge, this is the first study to extend the MJ ranking method to settings with continuous outcomes and evolving characteristics, while also evaluating its performance within a MCDM framework. The consistency of our findings across different configurations and anomaly sets highlights the scalability of MJ and demonstrates its capacity to adapt to complex, dynamic financial environments. Moreover, the absence of trainable parameters and hyperparameters, in contrast to methods requiring extensive parameter tuning, makes our approach particularly practical for real-world applications. Taken together, these considerations underscore the significant potential of our method for portfolio management and financial decision-making.

This work is organized as follows: in section 2, we briefly introduce social ranking functions, focusing on the MJ methodology defined by Balinski and Laraki (2010). In section 3, we introduce our MJ-based method for portfolio construction, extending the classical methodology to deal with continuous targets and time-varying domains. The data considered in our empirical analysis are described in section 4 together with the settings used in portfolio creation. In section 5 we outline experimental results, robustness tests and comparisons with other methods in the literature. Concluding remarks are discussed in section 6.

2. Preliminaries: social ranking functions and the Majority Judgment

Social choice theory, first introduced by De Condorcet (1785) and further developed by Arrow (1951), provides a rigorous framework for aggregating individual preferences, judgments, and votes into collective decisions. It can be viewed as an MCDM framework, where voters act as evaluation criteria, assessing feasible alternatives to determine the most preferred candidate. A social ranking method gathers individual opinions on each candidate and systematically aggregates them to identify the best choice or potential ties. Let us start by introducing the basic ingredients needed to construct a social ranking function. We define a finite number of competitors or candidates as $C = \{C_1, C_2, \ldots, C_m\}$, a finite number of judges or voters $J = \{J_1, J_2, \ldots, J_n\}$, and a common language of grades $\Lambda = \{\alpha, \beta, \gamma, \dots\}$ that is a totally ordered set. Considering those inputs we can define a voting profile as an $m \times n$ matrix Φ of the grades $\Phi(i, j) = \alpha_{ij} \in \Lambda$ that all judges $J_j \in J$ assign to each of the competitors $C_i \in C$.

A *method of ranking*, $\succeq_J$, where $C_i \succeq_J C_k$ indicates that jury J prefers C_i to C_k, is defined as a complete binary relation constructed starting from May (1952) axioms. A *social-ranking function* is a method of ranking that needs to satisfy the following properties:

1. *Anonymity*. Permuting the names of voters does not change the outcome (Pareto optimality).
2. *Neutrality*. Permuting the names of candidates does not change the outcome.
3. *Transitivity*. If $C_i \succeq_J C_k$ and $C_k \succeq_J C_l$ then $C_i \succeq_J C_l$. That is, Condorcet's paradox[1] cannot occur.
4. *Independence of irrelevant alternatives*. If $C_i \succeq_J C_k$ then whatever candidates are dropped or adjoined $C_i \succeq_J C_k$. That is, Arrow's paradox[2] cannot occur.

2.1. Majority ranking

Among the social ranking methods, this work focuses on the theory of Balinski and Laraki (2010), i.e., the majority judgment, that we outline below. Majority ranking is a multi-step process that starts with candidates being graded by a jury

[1]If this condition is violated, Condorcet's paradox occurs (De Condorcet, 1785). Suppose we have three candidates, C_1, C_2, and C_3, and three voters, J_1, J_2, and J_3, who have the following preference: $C_1 \succ_{\{J_1\}} C_2 \succ_{\{J_1\}} C_3$; $C_2 \succ_{\{J_2\}} C_3 \succ_{\{J_2\}} C_1$; $C_3 \succ_{\{J_3\}} C_1 \succ_{\{J_3\}} C_2$. In this case, the transitivity property fails, and there are cyclical societal preferences $C_1 \succeq_{\{J_1,J_2,J_3\}} C_2 \succeq_{\{J_1,J_2,J_3\}} C_3 \succeq_{\{J_1,J_2,J_3\}} C_1$. Even if the preferences of individual voters are transitive, collective preferences can be non-transitive. There is no clear winner even though each pair has a majority preference.

[2]This propriety avoids Arrow's paradox (Arrow and Raynaud, 1986), ensuring that the relative ranking of two alternatives should not change when other irrelevant alternatives are introduced or removed.

based on evaluation criteria. The vote/profile matrix, Φ, where judges' preferences are stated in a universal language Λ serves as a basis for defining a method of grading. This must be an aggregation function, F, summarizing all grades assigned by judges to competitors and using the same language: $F : \Lambda^{m \times n} \to \Lambda^m$. Therefore, $F(\Phi)$ is a vector whose i-th component, F_i is the aggregated or final grade assigned to competitor $C_i \in C$ by F. To be a *social-grading aggregation function*, F needs to satisfy the following properties:

- *Anonymity*: Permuting the names of judges (voters) should not affect the final grade, i.e., $\forall i \in \{1 \ldots, m\}$, $F_i(\ldots, \alpha, \ldots, \beta, \ldots) = F_i(\ldots, \beta, \ldots, \alpha, \ldots)$.
- *Unanimity*: If all judges assign the same grade to a competitor, C_i, then the aggregated grade must be the same, i.e., $F_i(\alpha, \alpha, \ldots, \alpha) = \alpha$.
- *Monotonicity*: If every judge increases C_i's grade (or leaves it unchanged), the final aggregated grade should not decrease. The final grade should increase if all individual grades increase.

$$\begin{aligned}
\alpha_j \preceq \beta_j \ \forall j \in \{1, \ldots, n\} \quad &\Rightarrow F_i(\alpha_1, \ldots, \alpha_n) \preceq F_i(\beta_1, \ldots, \beta_n) \\
\alpha_j \prec \beta_j \ \forall j \in \{1, \ldots, n\} \quad &\Rightarrow F_i(\alpha_1, \ldots, \alpha_n) \prec F_i(\beta_1, \ldots, \beta_n)
\end{aligned}$$

By assigning a final grade to each competitor, F determines the order-of-finish of all competitors. The social grading function used by MJ is the *majority grade*, the unique middlemost aggregation function respecting consensus, i.e.,

$$F_i^{MJ} = \begin{cases} r_{(n+1)/2} & \text{if } n \text{ is odd,} \\ r_{(n+2)/2} & \text{if } n \text{ is even.} \end{cases}$$

Here, r_a denotes the a-th element of the ordered vector $r = (r_1, \ldots, r_n)$, which represents a permutation of the candidate C_i's voting profile, $(\alpha_1, \ldots, \alpha_n)$, arranged in non-increasing order: $r_1 \succeq r_2 \succeq \cdots \succeq r_n$. Therefore, a competitor's majority grade is the grade that obtains an absolute majority of the voters against any lower grade and an absolute majority or a tie against any higher grade. The MJ grading principle acknowledges consensus: if a jury is more aligned in their assessment of one candidate compared to another, the stronger agreement should be honored by awarding a final grade that is at least equal to the other option's grade.

When aiming to rank candidate solutions, relying solely on the majority grade may be insufficient. While a candidate with a higher majority grade is naturally placed higher in the hierarchy due to the implied order of grades, there may be situations where different candidates receive the same grade. Ties between candidates are solved using the majority value, i.e., an ordered list of grades received by the candidate from the set of judges. Suppose we want to rank two candidates, C_1 and C_2, whose ordered list of grades are

$$r^{C_1} = (r_1 \succeq r_2 \succeq \cdots \succeq r_n)\,; \quad r^{C_2} = (r'_1 \succeq r'_2 \succeq \cdots \succeq r'_n)\,.$$

if their majority grades are the same, we can break the tie by removing the central entry from both vectors, re-evaluating the majority grades of the candidates, and comparing them. This procedure can be repeated until we find a mismatch in the majority grades obtained from the jury subsets. The only ties left will be those cases where $r^{C_1} \equiv r^{C_2}$ so that the two candidates are indistinguishable under the anonymity assumption. The majority value, r_v, of a candidate is the list of grades ordered to make iterative comparison straightforward:

$$r_v = \begin{cases} (r_{\tilde{m}}, r_{\tilde{m}+1}, r_{\tilde{m}-1}, r_{\tilde{m}+2}, r_{\tilde{m}-2}, \ldots r_n, r_1) & \tilde{m} = \frac{n+1}{2} \quad \text{if } n \text{ is odd,} \\ (r_{\tilde{m}}, r_{\tilde{m}-1}, r_{\tilde{m}+1}, r_{\tilde{m}-2}, r_{\tilde{m}+2}, \ldots, r_n, r_1) & \tilde{m} = \frac{n+2}{2} \quad \text{if } n \text{ is even.} \end{cases}$$

Thus, a competitor's majority value is the sequence that begins at the middle,

$r_{\tilde{m}}$, and fans out alternately from the center starting from below (above) when there is an even (odd) number of voters.

After introducing the majority value of different candidates we are able to rank different solutions by defining the *majority ranking* criterion, $\succ_{MJ}$. Given two competitors, C_1 and C_2, with majority values $r_v^{C_1}$ and $r_v^{C_2}$ then we say that $C_1 \succ_{MJ} C_2$, when $r_v^{C_1} \succ_{lexi} r_v^{C_2}$, where $\succ_{lexi}$ means lexicographically greater. Lexicographic order is a generalization of the alphabetical order of the dictionaries to sequences of ordered symbols or elements of a totally ordered set. Therefore, the majority ranking is defined by the lexicographic order among the sequences of grades that are the majority values.

In Figure 1 we depict the steps leading to the identification of majority ranking for a simple case where three judges grade four candidates using a voting system $\Lambda = \{1 \preceq 2 \preceq 3 \preceq 4\}$. These go from the initial grading of candidate solutions through the definition of majority values to the final determination of ranking.

Among the infinite ranking methods, the MJ stands out for its resistance to manipulations. The MJ permits voters to better express their opinions and always gives a transitive ranking of candidates, thus avoiding the Condorcet paradox. Moreover, the order between two candidates is not dependent on the grades received by other candidates, thus eliminating the Arrow paradox. It best combats voters' strategic manipulation, thereby inciting honest opinions. Furthermore, a candidate whose grades dominate another wins, preventing the domination paradox. Despite strategic voting and manipulation not being straightforwardly applicable to anomaly-based portfolio construction, these properties imply that MJ is robust against the presence of outliers (extreme votes) or systematic, "non-random" vote distortions (bias), rather common peculiarities of time-varying financial data.

	J_1	J_2	J_3
C_1	1	3	2
C_2	4	2	3
C_3	2	1	1
C_4	3	3	4

→

	r_1	r_2	r_3
C_1	3	2	1
C_2	4	3	2
C_3	2	1	1
C_4	4	3	3

→

	r_v
C_1	[2,1,3]
C_2	[3,2,4]
C_3	[1,1,2]
C_4	[3,3,4]

→

	r_v	**rank**
C_4	[3,3,4]	1
C_2	[3,2,4]	2
C_1	[2,1,3]	3
C_3	[1,1,2]	4

Figure 1: Transformation steps leading to majority ranking definition. The original grade profile where three judges J_i grade four candidates C_j (leftmost) is transformed into a sorted equivalent profile (center-left) through anonymity. The majority grade is identified by the central column (light gray). Subsequently, the majority value sequences are computed (center-right). The lexicographic order of the candidate majority values determines the majority ranking (rightmost).

3. Methods: Extending MJ for anomaly-based portfolio construction

This section aims to adapt the ranking framework of Balinski and Laraki (2010) to financial anomalies, treating anomalies as voters or judges and stocks as candidates to be ranked. Once a common grading language, Λ, is established, the properties outlined in section 2 naturally extend to multi-criteria portfolio construction based on pricing anomalies. Most of these properties are designed to ensure fairness and logical consistency in grade aggregation and ranking. In grade aggregation, unanimity guarantees that a stock's grade reflects the consensus among anomalies, while monotonicity ensures that stocks receiving higher individual anomaly grades are ranked accordingly. In ranking, neutrality mandates that stocks with identical anomaly grades receive the same ranking, transitivity prevents cyclical preferences that could lead to ambiguous orderings, and independence of irrelevant alternatives ensures that rankings remain stable when additional stocks are introduced or removed. Anonymity is generally a reasonable assumption, ensuring that all anomalies are treated symmetrically, aligning with standard aggregation methods such as mean-rank and z-score. For simplicity and consistency with prior methods, we assume equal anomaly weights in this work, though extending to different weightings is straightforward.[3]

[3]Following Balinski and Laraki (2010), MJ allows for integer weighting schemes without violating its core principles. This can be obtained by replicating a criterion's column in the grading matrix according to the weight assignment.

3.1. Transforming anomaly distribution into a grading system

Designing a universal grading system for pricing anomalies is challenging due to their diverse statistical properties, including varying ranges, symmetry, and tail behaviors. To ensure consistency and fairness, continuous anomaly values must be discretized into categorical grades while accounting for these distributional differences. We address this by employing *quantile binning*, a robust discretization method that allocates observations into approximately equal-sized categories. By adapting dynamically to the underlying distribution, quantile binning preserves relative rankings, enhances cross-variable comparability, and reduces sensitivity to outliers, providing a standardized and resilient grading framework across heterogeneous anomaly distributions.

Grade assignment proceeds as follows. At a given reallocation date, t_r, we identify all tradable stocks and collect their corresponding anomaly values, for that date (or within a recent window; see the next subsection). The stocks represents our competitors C while anomalies play the role of judges J. The empirical distribution of each anomaly is then constructed and discretized using quantile binning. The number of bins determines the cardinality of the grading system, i.e., the number of grades in Λ, $|\Lambda|$. For each anomaly, the grading intensity direction is determined based on empirical findings (see Table 1), where a sign of +1 (-1) indicates a direct (inverse) relationship between the anomaly and future returns. If the sign is +1, the lowest grade is assigned to the first quantile and the highest to the last; if -1, this order is reversed. Each stock is then assigned a grading vector based on the binning thresholds and intensity signs, forming the profile matrix at t_r, $\Phi(t_r, C_i, J_j) = \alpha_{i,j}(t_r) \in \Lambda$, where each anomaly $J_j \in J$ assigns a grade to each stock $C_i \in C$ tradable at t_r. This ensures a standardized and timely stock grading system across anomalies.

3.2. Rolling MJ-portfolio construction

Once the profile matrix $\Phi(t_r, C_i, J_j)$ is constructed, stocks can be straightforwardly ranked using MJ, leveraging cross-sectional anomaly grades to systematically aggregate information for decision-making. However, a common challenge in financial time series analysis is that cross-sectional estimates may overlook temporal dynamics and be highly sensitive to outliers. To mitigate these issues and evaluate the stability of MJ methods, we introduce three rolling window estimation techniques that smooth grades or rankings over time.

In each approach, C represents the set of tradable stocks on a given reallocation date t_r, with their historical data considered over the preceding w_{MJ} observations. Each anomaly in J evaluates the stocks w_{MJ} times, once at every observation within the rolling window. The quantiles used for categorical grading via quantile binning are determined based on the full distribution of factors across C over the rolling window period. This ensures that binning applies consistent thresholds across all window dates, maintaining uniform categorization and enhancing the universality of the resulting judgments.

Classical method: rank rolling. In this approach, the majority ranking is computed cross-sectionally at each time within the rolling window. We call $\text{rank}^k_{C_i}$ the rank of the stock C_i at time $\tau_k = t_{r-k}$, where t_{r-k} denotes the timestamp of the k-th observation preceding t_r and $k = 0, \dots, w_{MJ} - 1$. To make rankings at different times comparable, we rescale them between 0 and 1 as:

$$z^k_{C_i} = \frac{\text{rank}^k_{C_i}}{\max\limits_{C_l \in C} \{\text{rank}^k_{C_l}\}} \in \mathbb{R}^{m \times w_{MJ}};$$

where m is the cardinality of C. The rolling ranking of each stock, rank_{C_i}, is determined by the total ordering induced by the aggregation of the rescaled ranks at different dates:

$$\text{rank}_{C_i} = rank\left(\{z(C_j)\}_{j=1,\ldots,m}, \geq\right)_i, \qquad \text{where} \qquad z(C_i) = \frac{1}{w_{MJ}} \sum_{k=0}^{w_{MJ}-1} z_{C_i}^k .$$

Longitudinal Smoothing: profile rolling. Here, the smoothed voting profile is given by the majority grade of the votes assigned to each stock by each factor over time. Let us consider the case of the judge J_j evaluating the candidate C_i on dates $(\tau_{w_{MJ}-1}, \tau_{w_{MJ}-2}, \ldots, \tau_0 = t_r)$. Aiming to find an aggregate J_j grade for C_i, we can use the majority grade of the list of $\alpha_{i,j}$ in the rolling window:

$$\phi_{i,j}^{\text{long}} = F^{MJ}(\alpha_{i,j}(\tau_{w_{MJ}-1}), \alpha_{i,j}(\tau_{w_{MJ}-2}), \ldots, \alpha_{i,j}(\tau_0))$$

that encourages judge self-consistency over time. The final ranking are obtained using the standard MJ ranking starting from the smoothed majority grade matrix: $\Phi_{w_{MJ}}^{\text{long}}(t_r, C_i, J_j) = \{\phi_{i,j}^{\text{long}}\} \in \mathbb{R}^{m \times n}$, where n is the cardinality of J.

Cross-sectional smoothing: vote rolling. This approach is orthogonal to the previous one in that the vote aggregation preceding the majority ranking is cross-sectional and not longitudinal. In this case, we obtain the final ranking from the majority grades that the stocks received by the jury on different dates. The aggregate grade of stock C_i at time τ_k is given by:

$$\phi_i^{\text{cross}}(\tau_k) = F^{MJ}(\alpha_{i,1}(\tau_k), \alpha_{i,2}(\tau_k), \ldots, \alpha_{i,n}(\tau_k));$$

where $k = 0, \ldots, w_{MJ} - 1$. The final stock ranking is obtained by applying majority ranking on the smoothed vote profile of the candidates at different dates: $\Phi_{w_{MJ}}^{\text{cross}}(t_r, C_i, \tau_k) = \{\phi_i^{\text{cross}}(\tau_k)\} \in \mathbb{R}^{m \times w_{MJ}}$. Note that this is the only proposed rolling method reducing to the majority grade instead of the standard MJ ranking in the absence of a rolling window, i.e., when $w_{MJ} = 1$.

4. Experimental setup

4.1. Data and factors

For our analysis, we examine data from two primary sources: monthly CRSP (Center for Research in Security Prices) and annual Compustat data[4]. The dataset covers the period from January 2000 to December 2023 and includes US common stocks (share codes 10 and 11) listed on the NYSE, AMEX, and NASDAQ markets. To ensure data quality, we exclude stocks with fewer than 24 observations and those with more than 90% missing observations. We adjust for delisting returns using the approach described in Bali et al. (2017). We use the French data library[5] as a benchmark for the excess market return and the risk-free rate. In our analysis, we focus on several stock characteristic anomalies: accruals (accr), market capitalization (size), momentum (mom), book-to-market ratio (bm), gross profitability (gp), asset growth (ag), and net stock issues (ns). In addition, we include four other factors to account for anomalies related to trading frictions and risk: beta (β), dollar volume of trading (dvol), illiquidity (ill), and return volatility (vol). To evaluate these factors for each stock, we follow the methodologies proposed by the authors cited in Table 1. Further details about factor evaluation can be found in the supplementary material.

4.2. Portfolio strategy construction

In what follows, we outline the methods used to evaluate the efficiency of the MJ portfolio strategy and compare it with both single-anomaly sorting and other well-established anomaly aggregation approaches.

For portfolios based on single anomalies, we sort stocks into decile portfolios on each predetermined reallocation date according to the cross-sectional distri-

[4] https://www.crsp.org/
[5] http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html

Characteristics	Acronym	Reference	Sign
Accruals	accr	Richardson et al. (2005)	-1
Asset Growth	ag	Cooper et al. (2008)	-1
Beta	β	Fama and MacBeth (1973)	1
Book-to-market ratio	bm	Fama and French (1993)	1
Dollar volume	dvol	Brennan et al. (1998), Chordia et al. (2001)	-1
Gross Profitability	gp	Novy-Marx (2013)	1
Illiquidity	ill	Amihud (2002)	1
Momentum	mom	Fama and French (1993)	1
Net stock issue	ns	Fama and French (2008), Pontiff and Woodgate (2008)	-1
Size	size	Banz (1981)	-1
Volatility	volatility	Ang et al. (2006)	-1

Table 1: Stock anomalies used in our analysis, with references and return relationships. A sign of 1 (-1) indicates higher (lower) anomaly values correspond to higher returns.

bution of each factor, which serves as the breakpoint criterion.[6] The first (tenth) portfolio contains stocks with the lowest (highest) factor values. Additionally, we construct a long-short (high-low) portfolio by taking a long position in the portfolio with the best factor values, expected to yield higher returns, and a short position in the portfolio with the worst factor values, which theoretically implies lower returns. The rightmost column of Table 1 indicates the expected correlation between returns and anomalies, helping to identify the best and worst anomaly values. The same principles apply to portfolios constructed using the MJ strategy or alternative aggregation methods: stocks with the best (worst) placements are assigned to the tenth (first) portfolio. To fairly compare single-factor with the MJ strategy, we exclude stocks having missing anomaly values at reallocation dates. This ensures a consistent stock set across all ranking methods, resulting in a reduced set of stocks for portfolio creation.

Building on this general framework, our empirical analysis examines two key

[6]Following Bali et al. (2017), we define the kth portfolio at time t as $P_{k,t} = \{i \mid B_{k-1,t} \leq F_{i,t} \leq B_{k,t}\}$. This portfolio consists of securities i whose sorting variable $F_{i,t}$ lies between the $(k-1)$th $(B_{k-1,t})$ and kth $(B_{k,t})$ breakpoints at time t. To prevent empty portfolios when breakpoints coincide, we avoid strict inequalities. Consequently, some stocks may appear in multiple portfolios, but this does not affect the analysis.

aspects of portfolio management: rebalancing frequency and the size effect. First, we explore different reallocation frequencies, i.e., monthly, semi-annual, and annual, to assess their impact on portfolio performance. Second, we examine the size effect in portfolio strategies using equal-weighted (EW) and value-weighted (VW) portfolios, which assign equal or market capitalization-based weights, respectively. Widely used in empirical asset pricing, EW portfolios emphasize raw anomaly effects, often driven by small, illiquid stocks with higher trading costs, while VW portfolios better represent real-world capital allocations. Further analyses of the size effect in the supplementary material exclude stocks with a market capitalization lower than the 5th NYSE percentile, providing a stress test for our anomaly-driven strategy.

The MJ-based strategy configuration. In all experiments, we fix the grade cardinality at $|\Lambda| = 6$. However, in subsection 5.3, we assess the sensitivity of our results to different grading systems, demonstrating the robustness of the MJ approach to variations in the grade scale. Moreover, while partitioning stocks into equally sized subsets is straightforward for single-factor portfolio sorting, the MJ-based strategy may occasionally encounter ties. To address this, we redefine grades using percentile binning and compute majority grades at each reallocation date. Final ranks are then refined by applying percentile-based sorting to tied elements. Finally, for rolling MJ strategies, handling stocks with missing factor values on dates other than reallocation dates is crucial. To mitigate this issue, we impute missing grades using the median.[7]

[7]This approach, known as *general majority-ranking*, helps prevent bias when comparing candidates evaluated by different numbers of judges (Balinski and Laraki, 2010).

5. Results

We now present the results of our empirical analysis. First, we assess whether MJ effectively aggregates single anomaly signals by comparing its performance (without rolling window smoothing) to strategies based on individual anomalies with annual rebalancing. Afterward, we consider rolling methodologies to smooth MJ's evaluations and we perform robustness checks to validate their stability. Furthermore, we assess the resilience of the MJ strategy against varying number of voters and cardinality of the grading system. Finally, we compare MJ-based results with standard anomaly aggregation methods for portfolio construction.

5.1. The MJ on annual reallocation

Our first experiment examines portfolio aggregation with annual rebalancing. Following the classical approach of Fama and French (e.g., Fama and French (2008)), we rebalance portfolios at the end of each June, starting in 2000. Across 23 rebalancing dates, we construct ten portfolios as described in the previous section. On average, each portfolio contains 306 stocks. Annual rebalancing helps limit transaction costs (though not explicitly accounted for) by reducing trading frequency. We also analyze turnover rates to quantify the percentage of stocks reallocated at each rebalancing. For detailed formula derivations, see the supplementary material.

In Table 2, we present the theoretical top-performing (Best) and long-short portfolios of the MJ strategy, alongside eleven single-factor sorting strategies. For a comprehensive view of all portfolios, please refer to the supplementary material. The results demonstrate that MJ effectively aggregates anomaly signals, yielding significantly high average excess returns in both EW and VW for the best and long-short strategies (see also the cumulative excess returns of the portfolios in Figure 2). This occurs despite only a few factors, such as dvol, gp, and ns, demonstrating statistical significance in both EW and VW long-short strategies.

EW Portfolios													
		acc	ag	beta	bm	mom	dvol	gp	ill	ns	size	volatility	mj
Best	returns	0.0995	0.1190	0.0748	**0.1839**	0.0579	**0.2048**	**0.1433**	**0.1824**	**0.1494**	**0.1521**	**0.1068**	**0.2149**
	t-stat	(1.718)	(1.754)	(0.929)	(2.988)	(1.088)	(4.242)	(2.991)	(3.299)	(3.580)	(2.391)	(4.211)	(4.167)
	Volatility	0.2808	0.3289	0.3901	0.2983	0.2579	0.2341	0.2322	0.2680	0.2022	0.3083	0.1230	0.2500
	Sharpe ratio	0.3544	0.3618	0.1917	0.6163	0.2243	0.8750	0.6170	0.6805	0.7385	0.4932	0.8687	0.8596
Long-Short	returns	0.0283	**0.1122**	0.0029	**0.1312**	-0.0315	**0.1294**	**0.1074**	**0.1008**	**0.1881**	0.0665	0.0744	**0.1986**
	t-stat	(1.916)	(3.677)	(0.053)	(4.363)	(-0.616)	(3.623)	(2.770)	(2.419)	(4.388)	(1.342)	(1.193)	(5.787)
	Volatility	0.0715	0.1479	0.2593	0.1458	0.2476	0.1732	0.1879	0.2020	0.2078	0.2402	0.3025	0.1664
	Sharpe ratio	0.3953	0.7585	0.0110	0.8999	-0.1271	0.7473	0.5715	0.4989	0.9051	0.2767	0.2460	1.1938
VW Portfolios													
Best	returns	0.0623	0.0887	0.0887	**0.1145**	0.0692	**0.1292**	**0.0826**	**0.0932**	**0.1181**	0.0390	**0.0867**	**0.1542**
	t-stat	(1.334)	(1.872)	(1.225)	(2.306)	(1.292)	(3.726)	(2.501)	(2.267)	(3.694)	(0.657)	(3.586)	(3.675)
	Volatility	0.2264	0.2296	0.3511	0.2408	0.2597	0.1681	0.1602	0.1992	0.1549	0.2874	0.1172	0.2034
	Sharpe ratio	0.2753	0.3862	0.2528	0.4757	0.2665	0.7686	0.5159	0.4676	0.7621	0.1356	0.7397	0.7580
Long-Short	returns	0.0106	0.0665	0.0158	0.0605	0.0040	**0.0594**	**0.0971**	0.0224	**0.1181**	-0.0331	0.0667	**0.1033**
	t-stat	(0.408)	(1.873)	(0.238)	(1.588)	(0.065)	(2.430)	(2.231)	(0.753)	(4.106)	(-0.701)	(1.027)	(3.337)
	Volatility	0.1261	0.1721	0.3217	0.1846	0.2973	0.1186	0.2110	0.1444	0.1395	0.2291	0.3148	0.1501
	Sharpe ratio	0.0842	0.3865	0.0492	0.3276	0.0134	0.5013	0.4601	0.1553	0.8469	-0.1445	0.2119	0.6883

Table 2: Key statistics of the EW/VW best and long-short portfolios for single anomaly strategies and MJ marked (mj). T-statistics for excess returns are provided, with 5% significance levels in bold. Excess return, volatility, and Sharpe ratios are annualized.

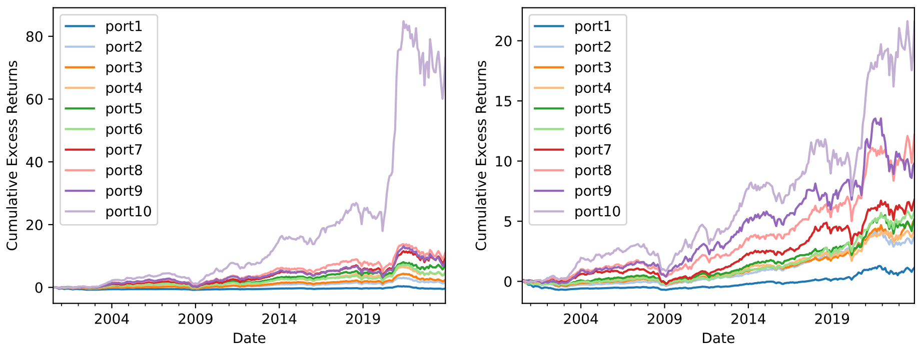

Figure 2: Cumulative excess returns of the MJ decile portfolios under EW (left) and VW (right).

Notably, contrary to prior empirical evidence, the EW momentum-based (mom) and the VW size-based (size) portfolios exhibit negative returns in the long-short strategy.[8] The best (long-short) EW MJ portfolio achieves an impressive and statistically significant yearly excess return of 21.5% (19.9%) with a Sharpe ratio of 0.86 (1.20), surpassing all other conventional strategies. In comparison, the best (long-short) portfolio of the dvol (ns) strategy yields slightly lower but still significant yearly excess returns of 20.5% (18.8%) with a Sharpe ratio of 0.87 (0.90). For the VW portfolios, while returns are generally lower, our aggregation

[8]As suggested by Aghassi et al. (2023), this inconsistency could be attributed to the underperformance of certain anomalies during the 2018 to 2020 period.

method still delivers the highest return for the best portfolio, achieving a 15% yearly excess return with a Sharpe ratio of 0.76, and the second-best result for the long-short portfolio showing a 10.3% excess return with a Sharpe ratio of 0.69. The most profitable long-short strategy in the VW setup is the ns strategy, which provides an excess return of 11.8% and a Sharpe ratio of 0.85. In terms of reallocation rate, our MJ-based strategy is comparable to the top-performing strategies derived from single factor sorting (see supplementary material).

In the supplementary material, we conduct further analyses to investigate the drivers of excess returns in the ten MJ-based portfolios by regressing their returns on classical risk factors, i.e., excess market return, size, value, and momentum. The findings for portfolio 10 and the long-short strategy indicate that, despite some sizable factor exposure, the MJ-based ranking identifies a distinct return source beyond traditional risk factors.

5.2. Rolling MJ and lag independency

Here, we present the results obtained using different rolling window extension and aggregation methods, considering monthly, semi-annual, and annual reallocation rates. To ensure a fair performance estimation across all configurations, we accounted for all distinguishable reallocation dates (lags) to estimate lag-independent excess returns. This approach mitigates potential biases arising from specific choices of reallocation months, which could otherwise restrict the analysis of portfolio strategies' impact. Specifically, we evaluated all possible strategies under a fixed reallocation rate, MJ-window size, and rolling window prescription method. For monthly reallocation frequencies, the concept of lag becomes irrelevant, as reallocation occurs at all available times.

Figure 3 compares different MJ approaches with strategies based on single anomaly values, displaying the lag-independent average excess returns and 95% confidence intervals for both EW (top) and VW (bottom) portfolio constructions.

The plot highlights the stability of our strategy across varying lags, as the lag-independent results closely align with those obtained in the previous section for annual reallocation in June. Our findings indicate that the vote-rolling method underperforms other MJ approaches, particularly in VW cases and for shorter windows. This under-performance likely stems from the temporal stability of majority grades, which reduces the informational content of the final ranking. The issue is exacerbated with a one-month window, where vote-rolling effectively collapses into majority grading (not majority ranking), further degrading performance. In contrast, profile and rank MJ rolling methods remain stable across different configurations and yield excellent performance. This is surprising, especially given that certain factors (beta, accruals, and cumulative returns) do not exhibit significant positive returns in all setups and show significant variations when changing portfolio weights. Rank and profile rolling MJ methods (including the baseline MJ strategy) consistently prove superior or at least comparable to baseline anomalies in terms of excess return performance. This consistency is not due to increased volatility, making MJ strategies preferable to single-sorting strategies.

The excess returns of rank and profile rolling MJ strategies exhibit minimal sensitivity to rolling window width but decline with lower reallocation frequencies in VW settings. As demonstrated in the supplementary material, the benefits of rolling windows are observed only in VW strategies when nano stocks, recognized as the primary carriers of pricing anomalies (Hou et al., 2018), are excluded. This phenomenon arises because mispricing signals in nano stocks are highly pronounced, easily identifiable, and thus rapidly arbitraged away. Consequently, applying a rolling MJ strategy in their presence may be counterproductive, as it could select stocks for which the arbitrage opportunity has already dissipated. In contrast, when only larger stocks, less susceptible to anomalies, are available, the rolling window approach helps stabilize votes or ranks, mitigating fluctuations

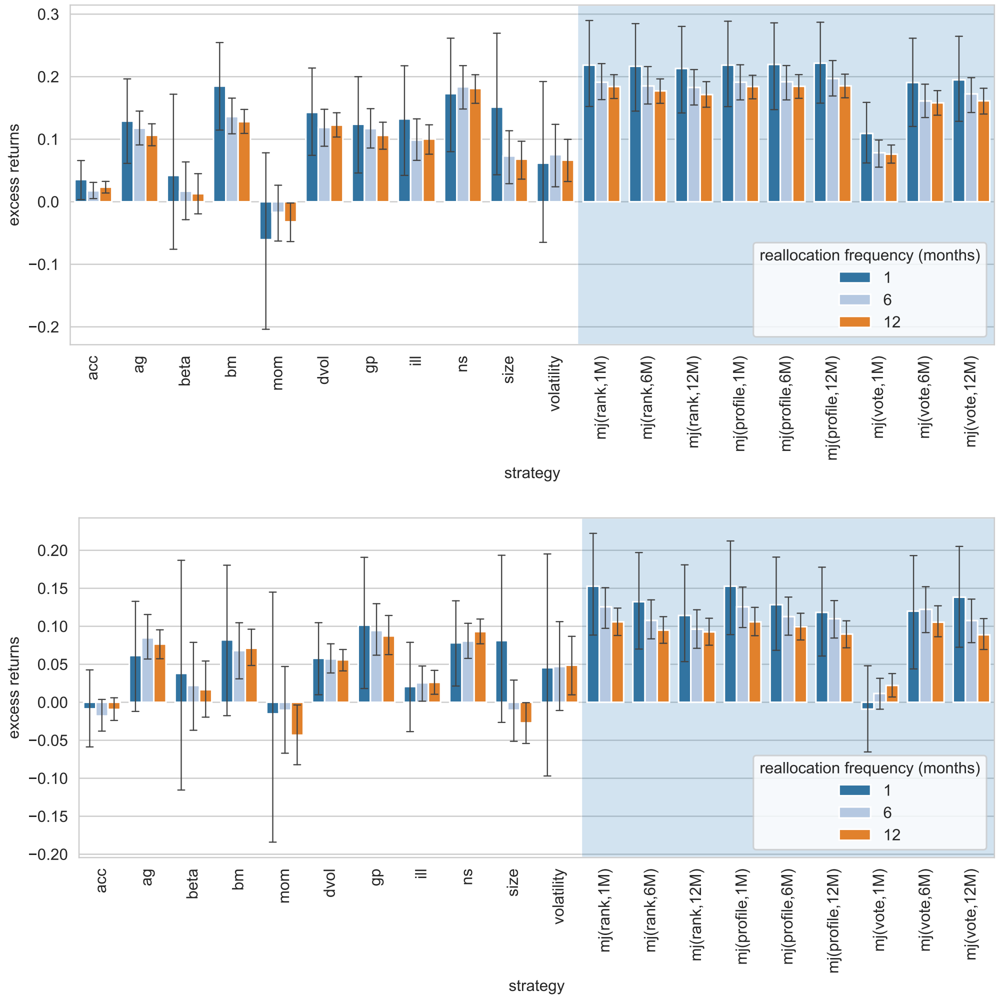

Figure 3: Lag-independent average excess returns of the EW (top) and VW (bottom) long-short portfolios based on the whole universe of US stocks.

and enhancing overall robustness.

The supplementary material contains the Sharpe ratio estimates and confidence intervals of the aforementioned strategies and the results obtained on the regularized subset of stocks with a market capitalization higher than the 5th percentile of the NYSE.

5.3. Robustness tests

We begin by examining the sensitivity of our method to different grading cardinalities. This analysis is crucial, as selecting a small number of grades in MJ can reduce differentiation between stocks, increasing the prevalence of ties, while a larger number of grades makes the grading outcome more susceptible to vote fluctuations and minor variations in voter opinions. To explore this effect, we vary the k-quantiles that define the grading system. The lag-independent excess returns for the long-short MJ strategy, without a rolling window and with annual reallocation, are presented in Figure 4. The EW strategy exhibits an initial increase followed by a slight decline as the number of grades increases, suggesting that at least five grades should be used. This aligns with common practice in MJ-based approaches, as noted in Balinski and Laraki (2010). Conversely, the VW setting does not display a clear pattern, likely due to the limited impact of pricing anomalies in large stocks, which diminishes the effectiveness of cross-sectional grade assignment.

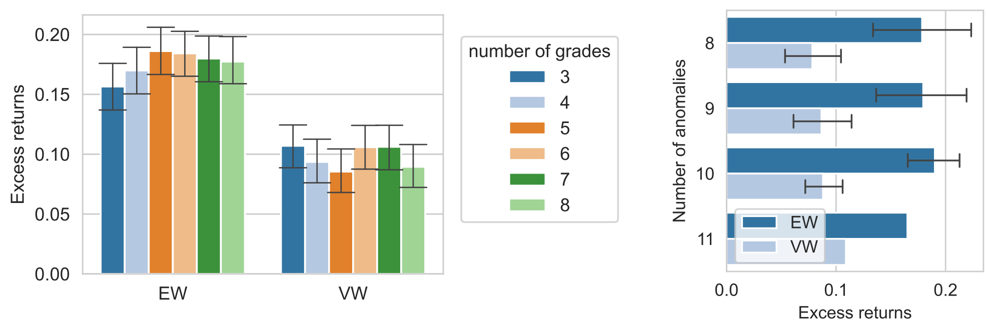

Figure 4: Lag independent average excess returns and confidence intervals of the EW/VW long-short portfolio using varying grade cardinality (left) and numbers of voters (right).

We now examine MJ's sensitivity to the number of jury components, which can influence ranking outcomes. While our analysis focuses on the most common pricing anomalies, it is crucial to assess whether specific factor choices significantly impact results and whether the inclusion or exclusion of certain factors alters performance. To this end, we evaluate all possible juries formed from random combinations of 8, 9, 10, or 11 factors, yielding 165, 55, 11, and 1 unique scenarios, respectively. Figure 4 illustrates the mean annualized excess returns and their 95% confidence intervals for the MJ long-short portfolio strategy without a rolling window. Our results indicate that average excess returns remain largely stable across different jury sizes, highlighting the robustness of the MJ method across a broad range of decision-making scenarios. This stability underscores MJ's scalability to higher-dimensional decision-making frameworks.

Other experiments concerning the exploration of grade aggregation systems other than MJ (building the ranking system starting from the 75% of the grade profile instead of the median to increase consensus) or potential refinements of long-short portfolio strategy creation methods can be found in the supplementary material. For the sake of brevity, here we only mention that MJ appears to be the most reliable and robust aggregation strategy and it does not require further extension to construct long-short portfolios.

5.4. Comparison with other standard aggregation methods

In this paragraph we compare the MJ method with other common anomaly aggregation strategies widely used by practitioners, i.e., mean rank (mr) and zscore (zs), which are the building blocks of current approaches to anomaly-based portfolio construction (Stambaugh et al., 2015; Asness et al., 2019).

Here we focus on long-short strategies and, to provide an in-depth comparison, we extend the study of the size effect by introducing a parametric family of weighting schemes wherein each stock is weighted by $\mathrm{ME}(s_i)^\gamma$, $\gamma \in [0, 1]$, where $\mathrm{ME}(s_i)$ is the market equity of the stock s_i. The extremes $\gamma = 0$ (EW) and $\gamma = 1$ (VW) define the boundaries of our analysis. Figure 5 reports the excess returns of MJ, mean-rank, and z-score approaches under monthly (left) and annual reallocation (right), respectively, for $\gamma \in \{0, 0.2, 0.4, 0.6, 0.8, 1\}$. Under monthly

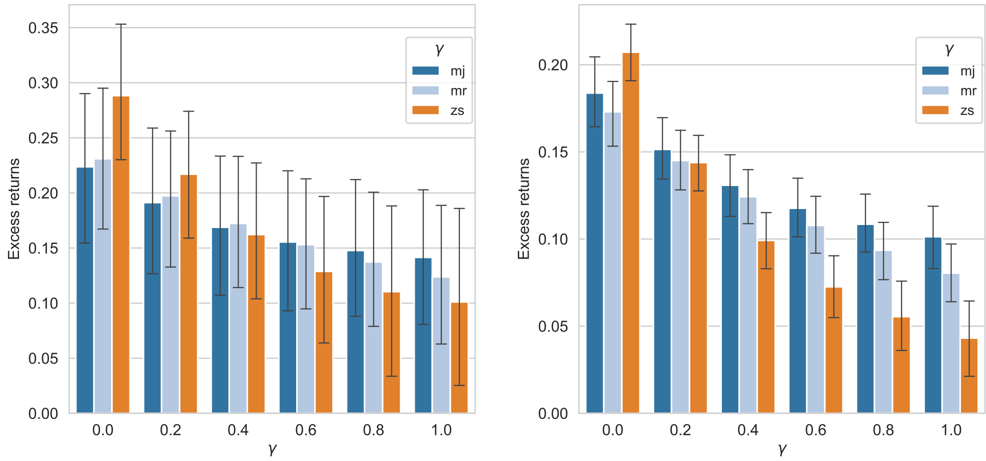

Figure 5: Lag-independent average excess returns and 95% confidence intervals of the long-short portfolio using monthly (left) and annual (right) reallocation. It considers different weighting schemes, parametrized by γ, and aggregation methods.

reallocation, MJ underperforms under strict EW but exhibits a more gradual decline in performance as γ increases; it ultimately achieves the best results in the VW setting. Under annual reallocation, MJ consistently outperforms mean-rank and displays greater robustness than z-score to deviations in γ: even a modest departure from $\gamma = 0$ confers an advantage to MJ over z-score. We attribute the z-score method's pronounced sensitivity to weighting schemes to its reliance on mean and standard deviation estimates in potentially skewed and heavy-tailed distributions (typical of several financial anomalies), making it more vulnerable to outliers. The mean-rank method, although mitigating scaling issues by relying on ordinal information, can overweight stocks that exhibit extreme ranks on only a subset of anomalies, thus failing to capture their overall performance profile. By contrast, MJ mitigates these pitfalls by effectively synthesizing the signals from multiple anomalies without disproportionately relying on small stocks or being unduly influenced by extreme values.

6. Conclusion

In this paper, we introduce a methodology that adapts the MJ method by Balinski and Laraki (2010), traditionally used in social choice contexts, for anomaly-based portfolio aggregation. In our setting, stocks represent competitors that are judged by market anomalies, here used as multiple evaluation criteria.

We assess the performance of our strategy on US stocks, including those on the NYSE, AMEX, and Nasdaq markets, from January 2000 to December 2023. We consider 11 market anomalies that are widely used and showed significant potential in cross-sectional studies. Following the traditional approach to anomaly exploitation, we create ten single-factor and MJ-based portfolios, that are either EW or VW. We show that our method successfully aggregates anomalies, outperforming single-sorted strategies in most settings.

We conduct several tests underscoring the robustness of our method across different reallocation frequencies, number of voters, and cardinality of votes used in the judgment system. Additionally, we examine the potential of rolling window methods for managing fluctuations in anomalies over time. The EW results indicate that performance remains stable across various rolling window choices. However, this approach enhances performance only in the VW setting and when excluding nano stocks, suggesting that rolling windows are particularly useful for distinguishing large-cap stocks.

By comparing the MJ-based with other common anomaly aggregation strategies, we find that it better mitigates the small-cap bias typical of many anomalies, retaining its efficacy across a broad range of weighting schemes. In particular, MJ outperforms competitors on annual reallocation, ensuring lower transaction costs and thus strengthening its appeal for both academic research and practical, large-scale portfolio management.

These results, along with the method's stability across a broad range of jury members, highlight the reliability and scalability of MJ in higher-dimensional decision-making spaces, making it potentially valuable for various application settings. Moreover, we believe our study enhances decision support for financial investors by introducing a novel portfolio construction strategy that is parsimonious, requires minimal hyperparameter tuning, and is firmly grounded in robust theoretical principles.

Acknowledgments. We would like to express our gratitude to M. Milasi, S. Moretti, and R. Lucchetti who engaged in valuable discussions regarding the theoretical methodology and its applicability to financial contexts. We also warmly thank W. Distaso who assisted in revising the financial implications of our findings.

Declaration of competing interest. None.

Funding. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

References

Aggarwal, C.C., Hinneburg, A., Keim, D.A., 2001. On the surprising behavior of distance metrics in high dimensional space, in: Database Theory—ICDT 2001: 8th International Conference London, UK, January 4–6, 2001 Proceedings 8, Springer. pp. 420–434. doi:10.1007/3-540-44503-X_27.

Aghassi, M., Asness, C., Fattouche, C., Moskowitz, T., 2023. Fact, fiction, and factor investing. The Journal of Portfolio Management 49, 57–94. doi:10.3905/jpm.2022.1.453.

Almeida-Filho, A.T.d., de Lima Silva, D.F., Ferreira, L., 2021. Financial modelling with multiple criteria decision making: A systematic literature review. Journal of the Operational Research Society 72, 2161–2179. doi:10.1080/01605682.2020.1772021.

Amihud, Y., 2002. Illiquidity and stock returns: cross-section and time-series effects. Journal of Financial Markets 5, 31–56. doi:10.1016/S1386-4181(01)00024-6.

Ang, A., Hodrick, R.J., Xing, Y., Zhang, X., 2006. The cross-section of volatility and expected returns. The Journal of Finance 61, 259–299. doi:10.1111/j.1540-6261.2006.00836.x.

Arrow, K.J., 1951. Social choice and individual values. volume 12. Yale University Press.

Arrow, K.J., Raynaud, H., 1986. Social choice and multicriterion decision-making. MIT Press Books 1.

Asness, C.S., Frazzini, A., Pedersen, L.H., 2019. Quality minus junk. Review of Accounting Studies 24, 34–112. doi:10.1007/s11142-018-9470-2.

Bali, T., Engle, R., Murray, S., 2017. Empirical Asset Pricing: The Cross Section of Stock Returns: An Overview. John Wiley & Sons, Ltd. pp. 1–8. doi:10.1002/9781118445112.stat07954.

Balinski, M., Laraki, R., 2010. Majority Judgment: Measuring, Ranking, and Electing. The MIT Press. doi:10.7551/mitpress/9780262015134.001.0001.

Banz, R.W., 1981. The relationship between return and market value of common stocks. Journal of Financial Economics 9, 3–18. doi:10.1016/0304-405X(81)90018-0.

Brennan, M.J., Chordia, T., Subrahmanyam, A., 1998. Alternative factor specifications, security characteristics, and the cross-section of expected stock returns. Journal of Financial Economics 49, 345–373. doi:10.1016/S0304-405X(98)00028-2.

Chordia, T., Subrahmanyam, A., Anshuman, V., 2001. Trading activity and expected stock returns. Journal of Financial Economics 59, 3–32. doi:10.1016/S0304-405X(00)00080-5.

Cochrane, J.H., 2011. Presidential address: Discount rates. The Journal of Finance 66, 1047–1108. doi:10.1111/j.1540-6261.2011.01671.x.

Cooper, M.J., Gulen, H., Schill, M.J., 2008. Asset growth and the cross-section of stock returns. The Journal of Finance 63, 1609–1651. doi:10.1111/j.1540-6261.2008.01370.x.

De Condorcet, N., 1785. Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix. Cambridge University Press.

Dong, X., Li, Y., Rapach, D.E., Zhou, G., 2022. Anomalies and the expected market return. The Journal of Finance 77, 639–681. doi:10.1111/jofi.13099.

Ehrgott, M., Klamroth, K., Schwehm, C., 2004. An MCDM approach to portfolio optimization. European Journal of Operational Research 155, 752–770. doi:10.1016/S0377-2217(02)00881-0.

Erwin, K., Engelbrecht, A., 2023. Meta-heuristics for portfolio optimization. Soft Computing 27, 19045–19073. doi:10.1007/s00500-023-08177-x.

Fama, E.F., French, K.R., 1993. Common risk factors in the returns on stocks and bonds. Journal of Financial Economics 33, 3–56. doi:10.1016/0304-405X(93)90023-5.

Fama, E.F., French, K.R., 2008. Dissecting anomalies. The Journal of Finance 63, 1653–1678. doi:10.1111/j.1540-6261.2008.01371.x.

Fama, E.F., MacBeth, J.D., 1973. Risk, return, and equilibrium: Empirical tests. Journal of

Political Economy 81, 607–636. doi:10.1086/260061.

Fasanghari, M., Montazer, G.A., 2010. Design and implementation of fuzzy expert system for tehran stock exchange portfolio recommendation. Expert Systems with Applications 37, 6138–6147. doi:10.1016/j.eswa.2010.02.114.

Feng, G., Giglio, S., Xiu, D., 2020. Taming the factor zoo: A test of new factors. The Journal of Finance 75, 1327–1370. doi:10.1111/jofi.12883.

Green, J., Hand, J.R.M., Zhang, X.F., 2017. The characteristics that provide independent information about average U.S. monthly stock returns. The Review of Financial Studies 30, 4389–4436. doi:10.1093/rfs/hhx019.

Gu, S., Kelly, B., Xiu, D., 2020. Empirical asset pricing via machine learning. The Review of Financial Studies 33, 2223–2273. doi:10.1093/rfs/hhaa009.

Harvey, C.R., Liu, Y., Zhu, H., 2016. ... and the cross-section of expected returns. The Review of Financial Studies 29, 5–68. doi:10.1093/rfs/hhv059.

Hou, K., Xue, C., Zhang, L., 2018. Replicating anomalies. The Review of Financial Studies 33, 2019–2133. doi:10.1093/rfs/hhy131.

Inada, K.i., 1969. The simple majority decision rule. Econometrica 37, 490–506. doi:10.2307/1912796.

Lin, H., Zhang, Y., Yang, X., 2024. Online portfolio selection of integrating expert strategies based on mean reversion and trading volume. Expert Systems with Applications 238, 121472. doi:10.1016/j.eswa.2023.121472.

Markowitz, H., 1952. Portfolio selection. The Journal of Finance 7, 77–91. doi:10.2307/2975974.

Marqués, A.I., García, V., Sánchez, J.S., 2020. Ranking-based MCDM models in financial management applications: analysis and emerging challenges. Progress in Artificial Intelligence 9, 171–193. doi:10.1007/s13748-020-00207-1.

May, K.O., 1952. A set of independent necessary and sufficient conditions for simple majority decision. Econometrica 20, 680–684. doi:10.2307/1907651.

Mirete-Ferrer, P.M., Garcia-Garcia, A., Baixauli-Soler, J.S., Prats, M.A., 2022. A review on machine learning for asset management. Risks 10. doi:10.3390/risks10040084.

Novy-Marx, R., 2013. The other side of value: The gross profitability premium. Journal of Financial Economics 108, 1–28. doi:10.1016/j.jfineco.2013.01.003.

Novy-Marx, R., 2016. Testing strategies based on multiple signals. NBER Working Paper No. w21329 .

Pätäri, E., Karell, V., Luukka, P., Yeomans, J.S., 2018. Comparison of the multicriteria decision-making methods for equity portfolio selection: The US evidence. European Journal of Operational Research 265, 655–672. doi:10.1016/j.ejor.2017.08.001.

Poh, D., Lim, B., Zohren, S., Roberts, S., 2022. Enhancing cross-sectional currency strategies by context-aware learning to rank with self-attention. The Journal of Financial Data Science doi:10.3905/jfds.2022.1.099.

Pontiff, J., Woodgate, A., 2008. Share issuance and cross-sectional returns. The Journal of Finance 63, 921–945. doi:10.1111/j.1540-6261.2008.01335.x.

Reschenhofer, C., 2024. Combining factors. Economics Letters 235, 111510. doi:10.1016/j.econlet.2023.111510.

Richardson, S.A., Sloan, R.G., Soliman, M.T., İrem Tuna, 2005. Accrual reliability, earnings persistence and stock prices. Journal of Accounting and Economics 39, 437–485. doi:10.1016/j.jacceco.2005.04.005.

Stambaugh, R.F., Yu, J., Yuan, Y., 2015. Arbitrage asymmetry and the idiosyncratic volatility puzzle. The Journal of Finance 70, 1903–1948. doi:10.1111/jofi.12286.

Stambaugh, R.F., Yuan, Y., 2017. Mispricing factors. The Review of Financial Studies 30, 1270–1315. doi:10.1093/rfs/hhw107.

Xidonas, P., Mavrotas, G., Zopounidis, C., Psarras, J., 2011. IPSSIS: An integrated multicriteria decision support system for equity portfolio construction and selection. European Journal of Operational Research 210, 398–409. doi:10.1016/j.ejor.2010.08.028.