

https://dspace.org/

Machine learning style rotation Evidence from the Johannesburg Stock Exchange

Moshe Page; David McClelland; Christiaan Auret

https://hdl.handle.net/10539/41055
IBSS

Downloaded from a DSpace repository. https://dspace.org/

Cogent Economics & Finance

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/oaef20

Machine learning style rotation – evidence from the Johannesburg Stock Exchange

Daniel Page, David McClelland & Christo Auret

To cite this article: Daniel Page, David McClelland & Christo Auret (2024) Machine learning style rotation – evidence from the Johannesburg Stock Exchange, Cogent Economics & Finance, 12:1, 2402893, DOI: 10.1080/23322039.2024.2402893

To link to this article: https://doi.org/10.1080/23322039.2024.2402893

© 2024 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group

View supplementary material

Published online: 13 Sep 2024.

Submit your article to this journal

View related articles

View Crossmark data

Full Terms & Conditions of access and use can be found at
https://www.tandfonline.com/action/journalInformation?journalCode=oaef20

COGENT ECONOMICS & FINANCE
2024, VOL. 12, NO. 1, 2402893
https://doi.org/10.1080/23322039.2024.2402893

FINANCIAL ECONOMICS | RESEARCH ARTICLE

OPEN ACCESS Check for updates

Machine learning style rotation – evidence from the Johannesburg Stock Exchange

Daniel Page[a,b], David McClelland[a,c] and Christo Auret[a]

[a]School of Economics and Finance, University of the Witwatersrand, Johannesburg, South Africa; [b]27four Investment Managers, Johannesburg, South Africa; [c]Laurium Capital, Johannesburg, South Africa

ABSTRACT

This study evaluates naïve and advanced prediction models when applied to style rotation strategies on the Johannesburg Stock Exchange ('JSE'). We apply 1- and 3-month style momentum as naïve predictors against three tree-based machine learning ('ML') algorithms (advanced predictors), namely Random Forest, XGBoost and LightGBM. Additionally, the study corrects for a shortcoming in the literature by incorporating trading costs into back-tested portfolio sorts. The results of the study are threefold. First, style rotation strategies based on advanced predictors achieve superior risk-adjusted returns when compared to naïve momentum. Of the three ML models applied, XGboost is superior, followed by LightGBM, implying that gradient boosters are superior to less advanced ensemble methods (Random Forest) which are in-turn superior to style momentum. Second, short-term momentum results in the highest share turnover across style rotation strategies, resulting in the largest negative impact associated with trading costs. Third, contrary to similar studies, the incorporation of price momentum as an independent variable in factor spanning tests renders most time-series alphas statistically insignificant.

IMPACT STATEMENT

This study considers the application of machine learning ("ML") for style rotation. The results indicate that ML based rotation signals generate excess performance relative to momentum based style rotation when applied on a cross-section of emerging market (South African) listed equities.

ARTICLE HISTORY
Received 23 March 2024
Revised 26 August 2024
Accepted 5 September 2024

KEYWORDS
Machine learning; style rotation; equity factor; emerging markets

SUBJECTS
Quantitative Finance; Statistics for Business, Finance & Economics; Machine Learning

1. Introduction

This paper explores the use of machine learning ('ML' hereafter) in timing investment styles which have been independently shown to generate performance in excess of what traditional risk-return frameworks expect. Over the past decade, literature has emerged that migrates from the investigation of singular equity styles (long-only) or factors (long-short) to their efficient combination in multi-style or multi-factor portfolios. The basis for multi-style strategies is an attempt to exploit the time-varying nature of style premia by cross-sectionally timing style exposure to generate abnormal profits. Several studies have considered the application of style and factor momentum to determine timing signals, with the majority finding that both generate significant time-series alphas (see Avramov et al., 2017; Cakici et al., 2022; Ehsani & Linnainmaa, 2022; Gupta & Kelly, 2019; Page et al., 2022; Su, 2021). Notwithstanding the simplicity of a momentum strategy, the drawbacks of naïve price momentum are largely applicable to style momentum, namely increased portfolio turnover (trading costs), idiosyncratic volatility, market risk and crash risk (see Barroso & Santa-Clara, 2015; Daniel & Moskowitz, 2016; Grobys et al., 2018). ML offers a more advanced approach to estimating style rotation signals by considering multiple explanatory variables (features) and allows for the incorporation of non-linear dynamics between features and their

CONTACT Daniel Page daniel.page@wits.ac.za School of Economics and Finance, University of the Witwatersrand, Johannesburg, South Africa

© 2024 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

dependent variables (labels), namely future style performance (see Galakis et al., 2021; Karatas & Hirsa, 2021; Ma et al., 2023; Neuhierl et al., 2023).

The application of ML to individual share price and investment style prediction is relatively nascent, with the field of study increasing dramatically over the past decade. Strader et al. (2020) and Patel et al. (2021) provide a comprehensive survey of the literature considering several supervised and unsupervised ML approaches for forecasting share prices. In reference to asset pricing, ML has been applied in several studies to disentangle the 'factor zoo', wherein ML is used to determine the data generating process of equity returns by exploring whether investment factors are 'priced' i.e. are able to explain future share or portfolio returns by minimizing out-of-sample R^2 (see Gu et al., 2020 and Leippold et al., 2022). The objective of this study differs to the abovementioned literature as its primary focus is to evaluate the predictive power of ML to the application of style rotation-based investment strategies against the well-documented application of style-level momentum as a means of style rotation. The goal of this study is therefore less about explaining return variation but rather testing the efficiency of 'advanced' predictive signals when applying a ML framework to style rotation.

This study will consider the application of both style momentum and ML to style rotation strategies on the Johannesburg Stock Exchange. The purpose of the study is to determine whether 'advanced' signal generation via three tree-based ML algorithms (Random Forest, XGBoost and LightGBM) is superior to 1- and 3-month style momentum when applied on an emerging market bourse with a notably smaller available investment universe in comparison to large developed markets. The superiority (i.e. accuracy) of signal generation is determined through back-tests where gross and net of fees portfolio performance is evaluated on nominal and risk-adjusted basis via factor spanning tests using five asset pricing models underpinned by the literature.

The results of the study are threefold. First, style rotation strategies based on advanced prediction methods achieve superior risk-adjusted returns when compared to style momentum. Of the three ML models applied, XGboost is superior, followed by LightGBM, implying that gradient boosters are superior to less advanced ensemble methods (Random Forest) which are in turn superior to style momentum. Second, style momentum results in the highest share turnover across rotation strategies, resulting in the largest negative impact associated with trading costs. Third, contrary to other style rotation strategy studies, the incorporation of price momentum as an independent variable in factor spanning tests renders most portfolio time-series alphas statistically insignificant. The study is structured as follows. Section 2 describes the data and methodological design; Section 3 presents the empirical results and Section 4 discusses the results and concludes.

2. Empirical design

2.1. Data

This study applies 30 long-only styles sorted monthly on the cross-section of shares listed on the Johannesburg Stock Exchange over the period January 2000 to August 2023. Data are sourced from Bloomberg and cover price, market capitalization, trading volume and accounting data. To mitigate the impact of survivorship bias, the dataset retains delisted counters and includes any new listings over the sample period. As an additional measure, shares are assigned a first true code (initial listing ticker) to limit the impact of structural breaks due to name changes. Returns are calculated on a total-return basis implying that prices are adjusted for dividends, special dividends, consolidations and unbundlings. To limit look-ahead bias, all accounting data are lagged three months from the reporting period date. Last, in the event of a delisting due to company failure, a share is assigned a -100% return.

2.2. Style sorts

Given the relative size and illiquidity of the JSE, the construction of long-only style portfolios follows the methodology of Page et al. (2022) where at each portfolio sort date, the cross-section of shares is limited to the top 120 counters based on market capitalization. The purpose of applying a limited universe is to ensure the 'investability' of the underlying style portfolios as well as the feasibility of their respective returns. Style proxy returns are estimated where for each style, shares are ranked monthly on their

respective style proxy z-score with the top 30 shares being assigned to the respective style portfolio. To limit the impact of share (and style) overlap across portfolios, we apply linear-weighting where shares are assigned a descending weight based on their z-score rank[1]. Portfolio returns are then calculated for the following month assuming buy and hold after which the process is repeated (see Appendix 1 for a detailed description of the underlying styles and their respective proxies).

2.3. Style rotation

2.3.1. Rotation signals

Using the long-only style proxy returns described, we apply five methodologies for generating style rotation signals, namely style momentum using a 1- and 3-month estimation period (naïve) and three tree-based ML algorithms (advanced), namely Random Forest, XGBoost and LightGBM. Per Figure 1, the process of generating signals is equivalent across prediction models, where at each portfolio sort date, the respective model generates a predicted expectation (i.e. raw signal) of each styles future return performance. Raw signals are then ranked in descending order and the bottom 15 styles are discarded[2]. The top 15 rank predictions are then converted to 'relative style weights' using linear weighting where the top predicted style (rank 1) is assigned the highest weight and decrease monotonically thereafter to the lowest ranked style (rank 15)[3]. The relative style weights are then applied (via multiplication) to the respective style proxy z-scores at portfolio formation date, thereby upweighting the style proxy z-scores that are predicted to outperform and underweighting or zero-weighting style proxy z-scores that are predicted to underperform. Style weighted z-scores are then aggregated (cross-sectionally added) to determine a single composite weighted z-score for each qualifying share. Finally, shares are ranked based on the weighted z-score to determine the top 30 shares which form the respective prediction models long-only style rotation portfolio.

2.3.2. Style momentum signals (naïve)

Per Arnott et al. (2023) and Page et al. (2022), we apply 1- and 3-month momentum as both studies found evidence of 'short' estimation periods achieving the highest portfolio returns. Style momentum signals are considered naïve given the simple nature of estimating signals. Plainly, at each portfolio formation date, the 1- and 3-month (21 and 63 day) momentum is calculated for each style and applied as the raw signal which is then converted to a signal weight using the method described above.

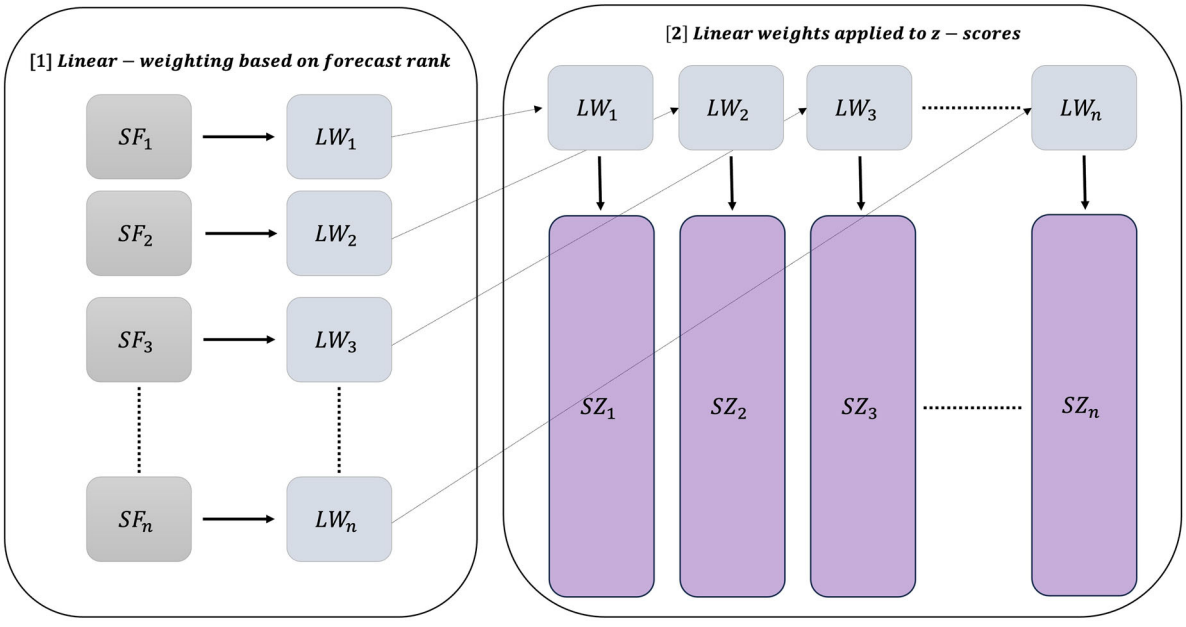


Figure 1. Signal application methodology.

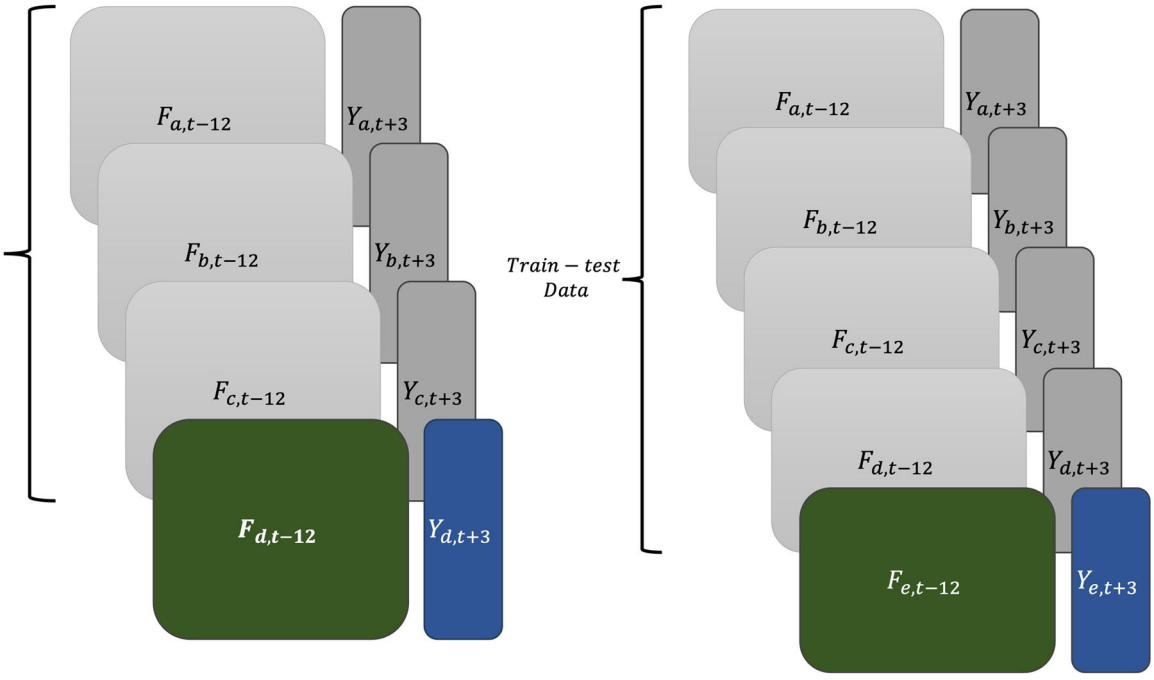


Figure 2. Machine learning process.

2.3.3. Machine learning signals (advanced)

The estimation of ML signals is more intricate. For each ML algorithm, we apply 212 (historical) nominal and risk-adjusted performance, risk and macroeconomic features (see Appendix 2 for feature descriptions) across the 30 styles, measured over the previous 1-12 months. Per Figure 2, features are mapped to their respective style's next-quarter returns in excess of the market return, where the market is proxied by the JSE ALSI total return index (J203T). As ML algorithms require a starting feature set for training, the first 11 years of the sample period (2001 – 2011) are used to initially train the model. For each ML model we assume a 70:30 train-test-split applying a randomized grid-search of possible algorithm specific parameters[4] assuming a k-fold of 3 and an iteration limit of 250[5]. Once trained, predictions are then estimated applying the most recent year (i.e. 2011) to predict the next quarters (i.e. 2012 Q1) out-of-sample style excess performance (raw signals). Once prediction signals are generated, the prior prediction feature and label data are incorporated in the training data and the process is repeated (training data now ranges from 2001 – 2012Q1 etc.). Therefore, each forecast is conducted out-of-sample on unseen data while allowing the training data to be updated with the most recent historical data post prediction.

2.3.4. Portfolio formation

As with the underlying style proxy portfolios, the above models are run quarterly to determine the constituents of each respective prediction-model based portfolio. Per prediction model, the top 30 shares based on the cumulative rank-weighted z-score are assigned to one of five portfolios (style momentum (1- and 3- month), Random Forest, XGBoost and LightGBM). Portfolio returns are then calculated for the following quarter assuming three weighting methodologies, namely equal, market capitalization and linear weighting. A significant drawback of both Arnott et al. (2023) and Page et al. (2022) is the failure to consider the impact of trading costs on style rotation strategies as both found that style momentum measured over short estimation and holding periods produced the highest risk-adjusted returns. To account for trading costs, we evaluate the stock-level turnover of each portfolio and apply a trading cost and security transfer tax ('STT') of 10 basis points[6] (buy and sell) and 25 basis points (buy only) respectively, based on relative weight and weight change of each share at rebalance.

2.4. Performance evaluation - factor spanning

We apply several factor spanning regressions on style rotation portfolios, specifically focusing on the unexplained performance captured by time-series alpha. Factor spanning requires the application of an

underlying asset pricing model in the form of time-series regressions that attempts to identify whether independently priced factor premiums explain the time-series variation of the underlying strategy in question. For this study, we consider a generic factor spanning model of the form:

$$R_{i,t} - rf = \alpha_i + \sum_{j=1}^{n} \beta_{i,j} X_{j,t} + \varepsilon_{i,t} \quad (1)$$

Where $R_{i,t}$ is the time-series return of the style rotation portfolio i, rf is the risk-free proxy (91-day SA Government treasury bill rate), α_i the style rotation portfolio alpha, $\beta_{i,j}$ is the vector of factor loadings of portfolio i, $X_{j,t}$ is a matrix of factor premiums and $\varepsilon_{i,t}$ is the time-series residuals (error) of portfolio i which are assumed to be normally, identically and independently distributed (*n.i.i.d*). To mitigate the impact of heteroskedasticity in the error term, all factor spanning regressions are run using Newey-West HAC consistent standard errors. In terms of the independent variables applied to the factor-spanning regressions, we consider five asset pricing models described in the literature namely, CAPM, Fama-French 3 factor ('FF3') (Fama & French, 1993), Carhart 4 factor ('CH4') (Carhart, 1997), Fama-French 5 factor ('FF5') (Fama & French, 2015) and lastly, Fama-French 6 factor model ('FF6') (Fama & French, 2018). To ensure consistency, we match the weighting methodology of the style rotation strategy to the weighting method of the independent variables used in factor spanning regressions[7].

3. Results

3.1. Nominal return analysis

Table 1 describes the average monthly returns of the five style rotation strategies in excess of the risk-free proxy. The results show that irrespective of weighting methodology or fees, the XGBoost style-rotation strategy is superior, achieving average monthly excess returns ranging from 0.99% to 1.12%, all of which are statistically significant at the 1% level. Conversely, 1- and 3-month style momentum realize the lowest average returns of between 0.75% to 0.96%. Importantly, the results show that the inclusion of trading fees negatively impacts returns of all strategies, yet style momentum experiences the greatest decline. The greater sensitivity of style momentum to trading costs is displayed in Figure 3 that follows.

Figure 3 describes the portfolio turnover analysis of 1-month style momentum against 3-month momentum and the three ML based style-rotation strategies (advanced signals). The figure shows that 1-month momentum (blue area) produces the highest portfolio turnover, followed by 3-month momentum, offering a clear explanation for negative impact of trading costs on the naïve signal strategies. As expected, style momentum, even when applied quarterly, results in higher portfolio turnover thereby offsetting the benefit of a narrow momentum estimation period as discussed in Page et al. (2022) and Arnott et al. (2023).

Figures 4–6 provide the cumulative return analysis of the five style rotation strategies against the market proxy. As shown in Table 1, across the three ML weighting methods, XGBoost appears to be the most robust strategy when compared to the benchmark. Conversely, 1- and 3-month momentum underperform the benchmark, barring the application of linear weighting. The results presented thus far provide several key insights to the dynamics of style rotation strategies on the JSE. First, naïve signals only seem to be profitable when applying a weighting methodology that signal-weights respective shares (linear weighting). Second, advanced signal-based style rotation tends to produce higher excess returns

Table 1. Monthly average excess returns.

Method	Style Momentum (21)	Style Momentum (63)	Random Forest	XGBoost	LightGBM
EW - Gross	0.847%***	0.856%***	0.971%***	1.124%***	0.992%***
EW - Net	0.782%**	0.790***	0.913%***	1.071%***	0.937%***
MCW - Gross	0.820%**	0.821%**	1.006%***	1.106%***	0.961%***
MCW - Net	0.750%**	0.751%**	0.945%***	1.053%***	0.906%***
LW - Gross	0.964%***	0.876%***	0.923%***	1.049%***	0.943%***
LW - Net	0.893%***	0.803%**	0.858%***	0.989%***	0.881%***

Monthly average excess returns of the five style rotation strategies using 1 and 3-month style momentum, Random Forest, XGBoost and LightGBM. Returns are expressed in excess of the SA Government 91-day treasury bill rate. Statistical significance of excess returns is evaluated via paired sample t-tests where ***,**,* indicates significance at the 1%, 5% and 10% level.

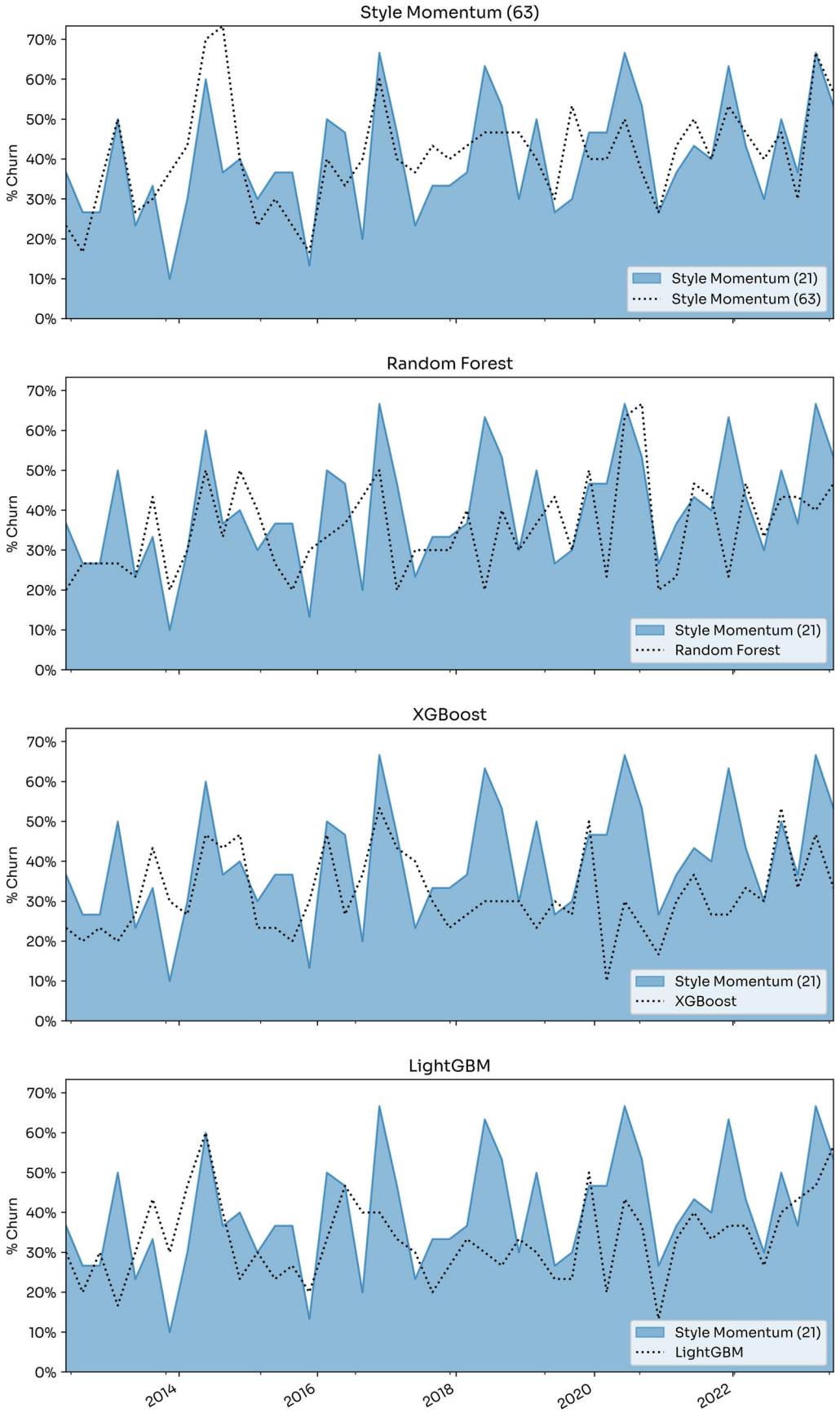


Figure 3. Portfolio turnover analysis.

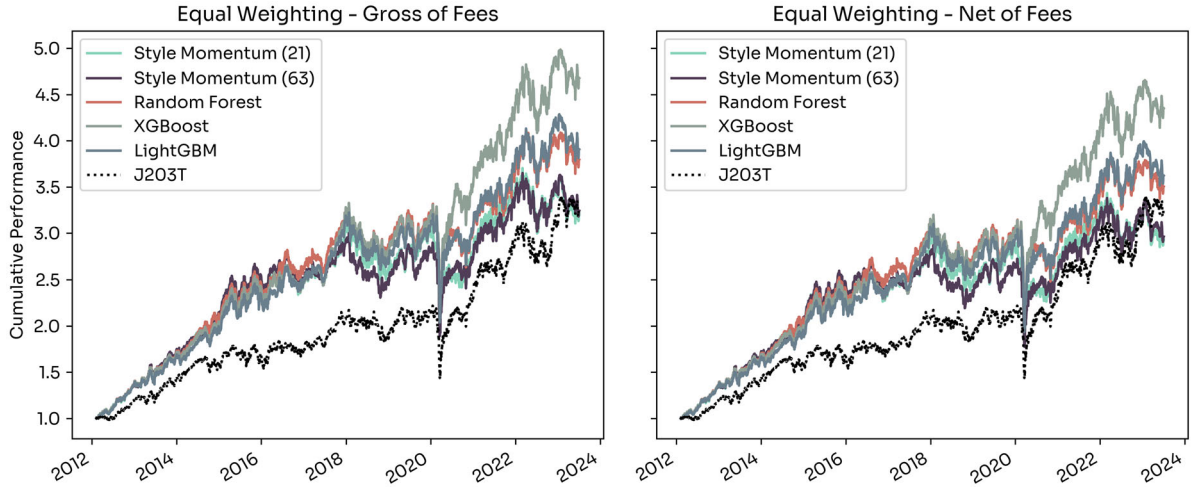


Figure 4. Equal weighted cumulative performance.

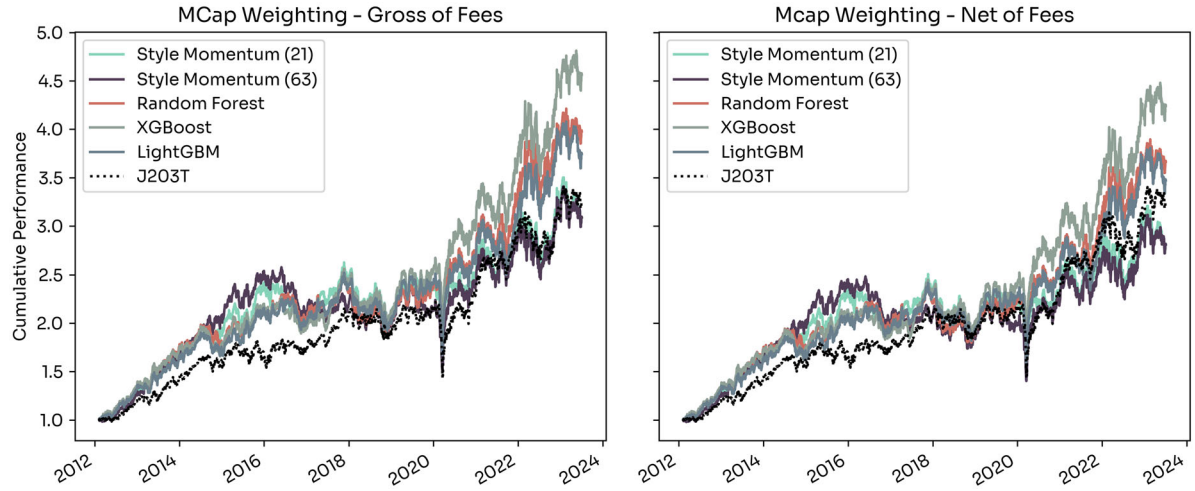


Figure 5. Market capitalisation weighted cumulative performan.

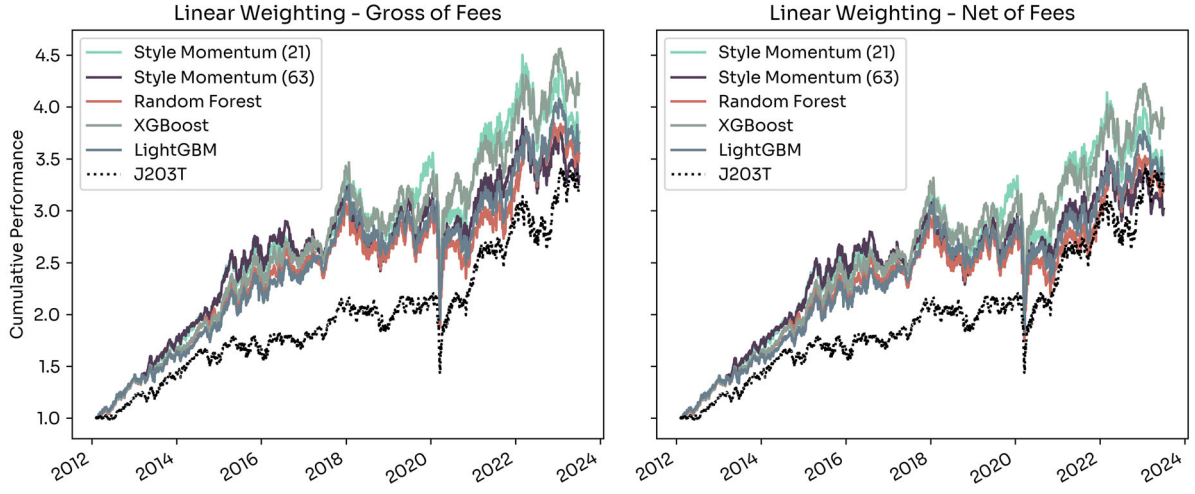


Figure 6. Linear weighted cumulative performance.

Table 2. Sharpe ratio analysis.

Method	Style Momentum (21)	Style Momentum (63)	Random Forest	XGBoost	LightGBM
EW - Gross	0.236*	0.245*	0.265**	0.297**	0.273**
EW - Net	0.219	0.227*	0.250*	0.283**	0.259**
MCW - Gross	0.212	0.220	0.267**	0.281**	0.252*
MCW - Net	0.196	0.201	0.252*	0.268**	0.238*
LW - Gross	0.251*	0.231*	0.242*	0.267**	0.248*
LW - Net	0.234*	0.213	0.226	0.252*	0.232*

Monthly Sharpe ratios of the five style rotation strategies using 1 and 3-month style momentum, Random Forest, XGBoost and LightGBM. Sharpe ratios are calculated using portfolio returns in excess of the 91-day treasury bill rate scaled by monthly standard deviation. P-values are estimated assuming that Sharpe ratios follow a Student t-distribution with a test statistic equal to $t = \sqrt{n}SR$ (where SR is the calculated Sharpe ratio and n is the sample size) and ***,**,* indicates significance at the 1%, 5% and 10% level.

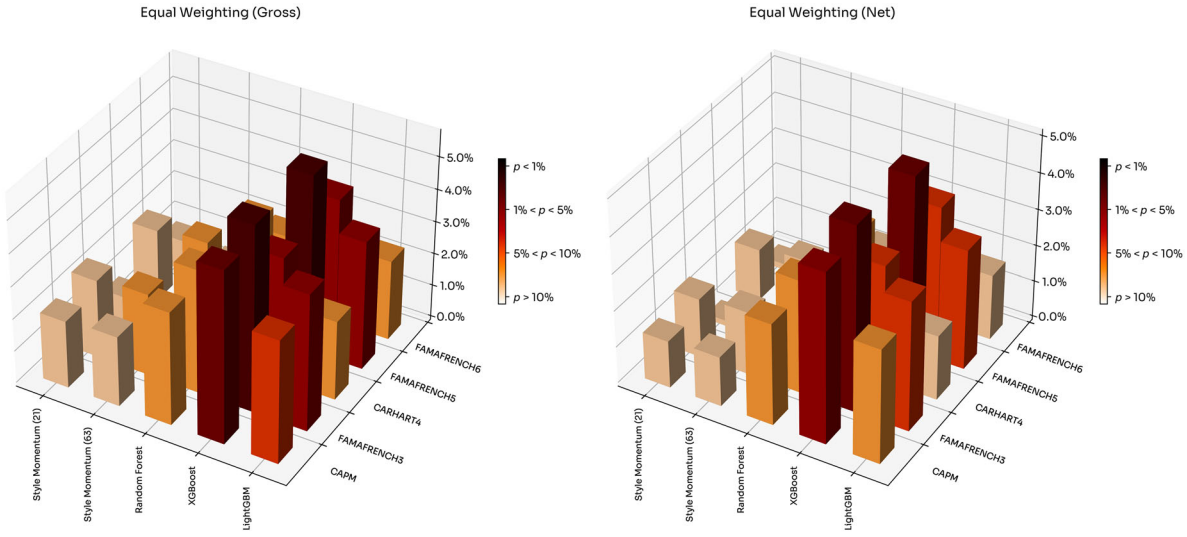


Figure 7. Equal weighted factor spanning.

on both a gross and net of fees basis. Last, of the ML models applied, XGBoost is superior gross and net of fees, with the latter performance being bolstered by lower portfolio turnover.

Table 2 describes monthly Sharpe ratios of the five style rotation strategies across the three weighting methods applied. The results show that irrespective of weighting method or the application of trading costs, XGBoost provides the highest Sharpe ratios, ranging from 0.25 to 0.3, all of which are significant at the 5% level, barring the linear weighted net of trading costs portfolio (which is still significant at the 10% level). Conversely, 1 and 3-month style momentum Sharpe ratios vary from 0.2 to 0.25, with only two of the former and three of the latter being significant at the 10% level. Comparatively, XGBoost produces aggregate Sharpe ratios that are 22% higher than those achieved by style momentum, with the difference being significant at the 5% level. To assess the impact of trading costs, an aggregate comparison of gross and net of fees Sharpe ratios is conducted. Across all strategies, the application of trading costs reduces Sharpe ratios by a statistically insignificant 6.1% on average. Conversely, and consistent with table 1 and Figure 3, 1- and 3-month style momentum Sharpe ratios are the most sensitive to trading costs, decreasing by 7.3% on average with the difference being significant at the 10% level.

3.2. Factor spanning tests

Figures 7–9 describe the time-series alpha estimates (z-axis) of the five style rotation strategies (x-axis) using the five factor spanning regressions (y-axis). Each bar represents the respective portfolios time-series annualized alpha and are color coded based on statistical significance. Focusing on Figure 7, the left-hand sub-figure shows gross-of-fee alphas and indicates that XGBoost followed by LightGBM produce the economically largest alphas, followed by Random Forest. In terms of statistical significance, a clear pattern emerges where factor spanning models that include price momentum reduce both the economic scale and statistical significance of alphas. The right-hand sub-figure emphasizes the impact of

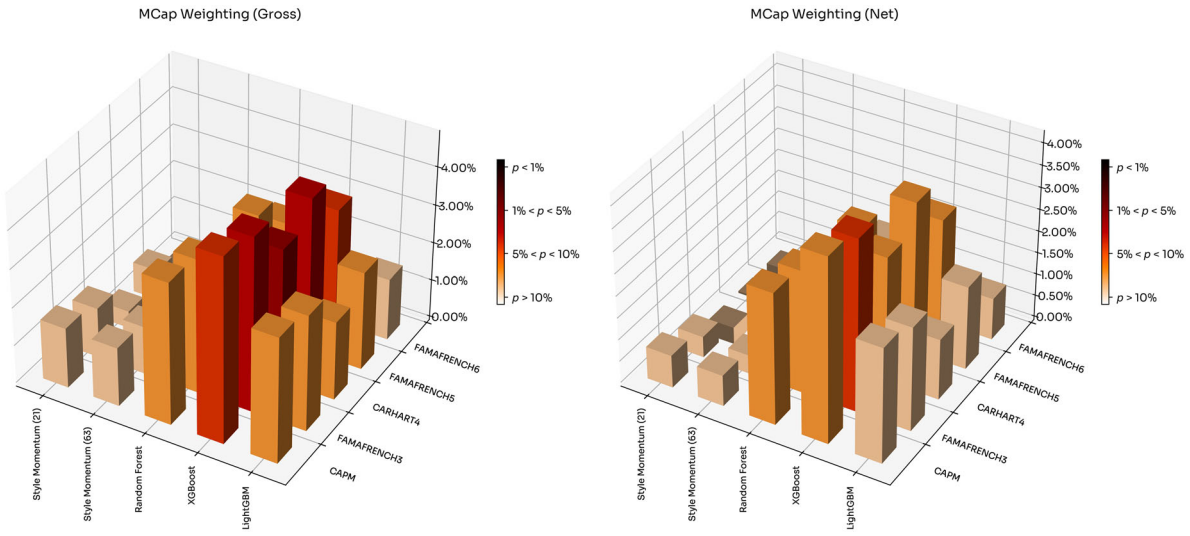


Figure 8. Market capitalisation weighted factor spanning.

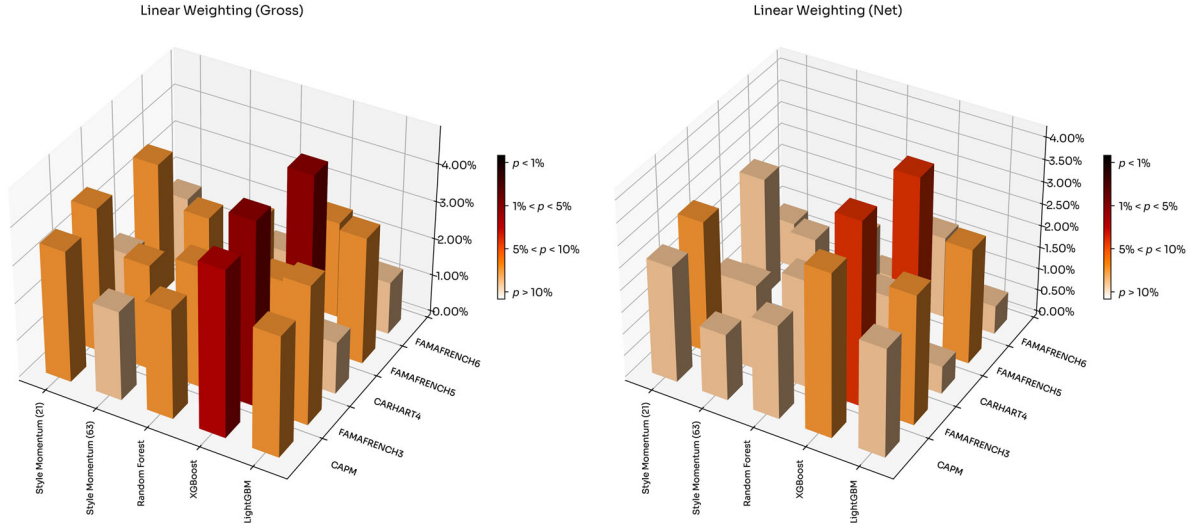


Figure 9. Linear weighted factor spanning.

trading costs on style rotation strategies. The best performing signal model, XGBoost, only produces alphas that are significantly different from zero at the 5% level when applying the market model (CAPM), FF3 and FF5 while alphas are only significant at the 10% level when applying CH4 and FF6.

Unlike the nominal return analysis presented in Table 1, Figure 8 shows that market capitalization weighting negatively impacts the economic size and statistical significance of portfolio alphas. The impact is exacerbated when including trading costs with only the XGBoost style rotation strategy producing a statistically significant alpha (p-value of less than 5%) when evaluated using the FF3 factor spanning model. Figure 9 describes the time-series alphas when applying linear weighting. Unlike equal and market capitalization weighting, linear weighting weights portfolio constituents based on the cumulative rank z-score of each share. 21-day style momentum experiences the most dramatic improvement when applying linear weighting, with alphas markedly increasing in comparison to equal and market capitalization weighting as seen in Figures 7 and 8 respectively. However, consistent with Table 2 and Figure 3, Figure 9 confirms that style rotation based on style momentum is most sensitive to trading costs. Once again, only XGBoost based style rotation produces alphas that are statistically significant (at the 10% level) limited to FF3 and FF5, both of which exclude price momentum as an independent factor premium.

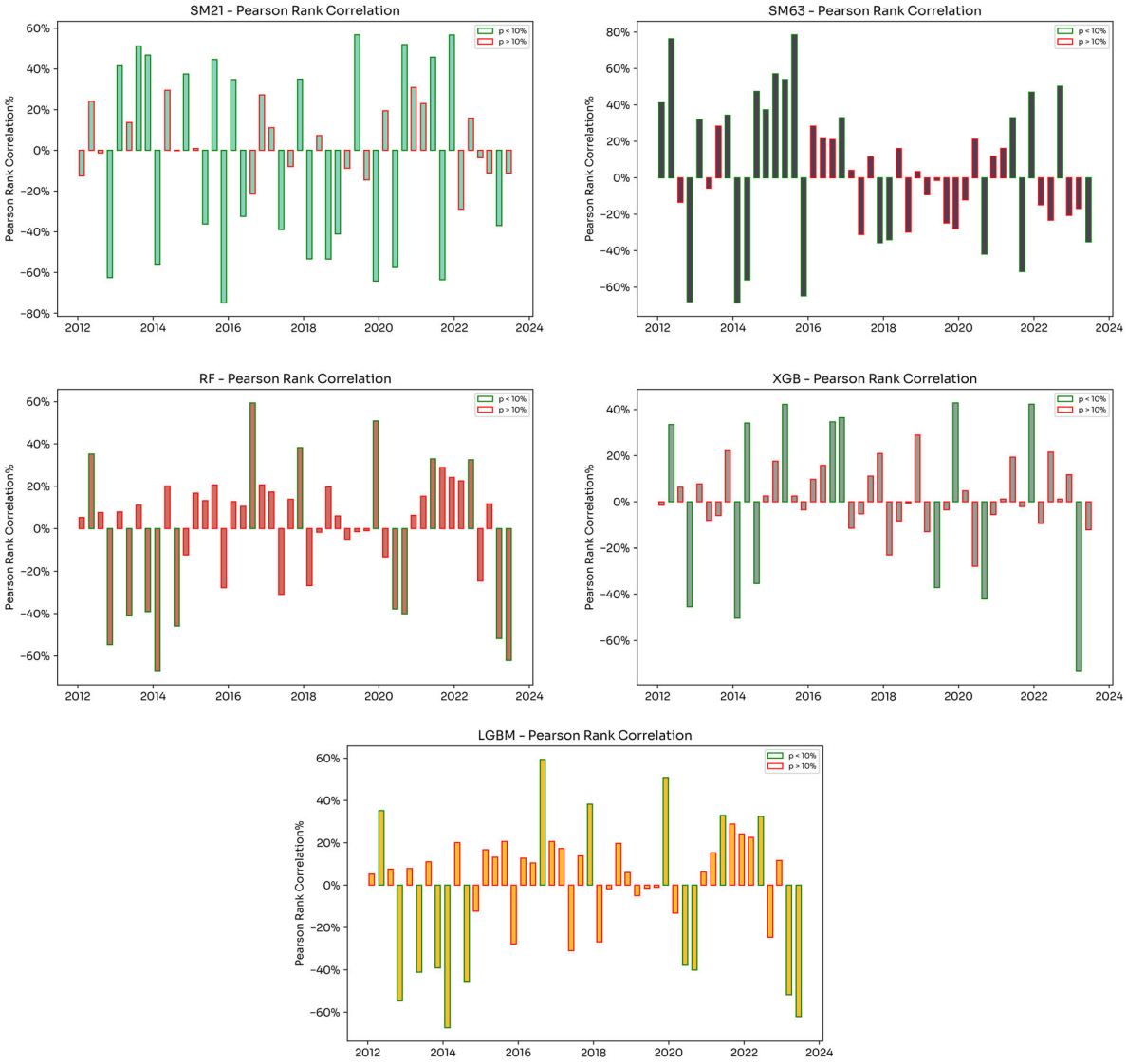


Figure 10. Spearman rank correlations.

Table 3. Spearman rank correlations.

	[1] Correlation > 0%	[2] Correlation < 0%	[3] Avg. Correlation > 0%	[4] Avg. Correlation < 0%
Style Mom. (21)	47.83%	52.17%*	32.07%	−33.01%*
Style Mom. (63)	52.17%	47.83%	33.52%	−31.21%*
Random Forest	58.70%*	41.30%	20.81%	−30.75%
XGBoost	52.17%	47.83%	19.62%	−19.29%
LightGBM	58.70%*	41.30%	20.81%	−30.75%

Summary statistics of the Spearman rank correlation results where columns [1] and [2] describe the percentage of correlations greater and less than zero. Columns [3] and [4] describe the average correlation for each model where correlations are greater and less than zero. ***,**,* indicates significance at the 1%, 5% and 10% level.

4. Robustness tests

Notwithstanding the nominal and risk-adjusted results presented, further analysis is conducted to determine the source of variation in performance of the rotation signal models. Given that the strategy applied relies on linear weighting model predictions based on rank, a natural test is to measure the accuracy of *ex ante* rank predictions against *ex post* actual ranks using Spearman rank correlation.

Figure 10 describes the rank-correlation analysis of each of the five prediction models while Table 3 summarizes the rank correlation tests. Columns [1] and [2] of Table 3 describe the percentage of times the respective model rank predictions were positively and negatively correlated with actual performance

Table 4. Intersection tests.

	[1] Predicted ∩ Actual (Top 5)	[2] Predicted ∩ Actual (Top 10)	[3] Predicted ∩ Actual (Top 15)
Style Mom. (21)	17.83%	34.13%	50.14%
Style Mom. (63)	20.43%	35.22%	51.16%
Random Forest	19.13%	35.22%	50.29%
XGBoost	21.30%**	36.30%*	51.45%
LightGBM	19.13%	35.22%	50.29%

Summary statistics of the intersection tests where table describes the average proportion of the 'top' predicted styles that intersect with the actual 'top' performing styles at each portfolio rebalance period. T-tests were applied where the mean intersection percentage is compared to the number of top styles considered scaled by the total number of styles, therefore for the top 5, 10 and 15 styles, proportion was compared to 5/30 (16.67%), 10/30 (33.34%) and 15/30 (50%). ***,**,* indicates significance at the 1%, 5% and 10% level.

ranks of the styles considered. The table indicates that 21-day style momentum achieved the lowest prediction accuracy, with only 48% of *ex ante* predicted ranks having a positive correlation with actual style *ex post* performance rank. Interestingly, Random Forest and LightGBM achieved the highest number of positive rank correlations (59%), both of which were significant at the 10% level. Columns [3] and [4] consider the average correlation values of the respective models. Interestingly, XGBoost, achieved the lowest average positive correlation, yet showed the most symmetric average negative correlation. In contrast, barring 63-day style momentum, the other prediction models all achieved worse average negative correlations. Reconciling the rank correlation analysis with performance, the superiority of XGBoost's performance could be explained by the model being less likely to produce incorrect signals as opposed to being more likely to produce correct signals.

A second accuracy (precision) test was conducted where we focused on the intersection of *ex ante* predictions in conjunction with *ex post* style performance. The basis of the test was analyzing whether the top predicted styles per model intersected with the actual style performance. Therefore, we shifted our focus from rank to the calculation of the proportion of the top 5, 10 and 15 style predictions per model at each portfolio sort date to the actual style performance over the following quarter. For example, if a model correctly predicted two of the top performing styles in the top five performing styles for the future period, a score of 2/5 (40%) was assigned. Proportions per model were then tested applying a one-sample T-test against a population mean of number of top styles scaled by total number of styles[8].

The intersection tests per Table 4 are less stringent than Spearman correlation as they focus on whether a subset of predictions is correct as opposed to the preciseness of the overall prediction rank. As shown in column [1], the average proportion of XGBoost correctly predicting at least one of the top five styles is 21.3% and is significant at the 5% level. This implies that XGBoost correctly predicts the performance of at least one style in the top five preforming styles at each portfolio formation date. Column [2] is consistent with column [1] as it shows that when extending the test to the top 10 styles in each period, XGBoost's predictions are 36.3% accurate (and is significant at the 10% level). Lastly, column [3] shows that none of the average intersection scores are significantly different from zero, however, XGBoost achieves the highest average intersection score of 51.5%.

5. Discussion and conclusion

The application of style rotation to capture time-varying style premiums has garnered significant attention across academics and practitioners alike. Most style rotation literature has focused on style or factor momentum as the key predictor of future style returns, emphasizing that a simplistic or naïve signal can generate investment returns that are not explained by well-established asset pricing models. This study expands the current body of literature on several fronts. First, it builds on the existing body of emerging market knowledge by extending the signal generation models to include ML algorithms, namely Random Forest, XGBoost and LightGBM. Building on the work of Page et al. (2022), we find that 21-day style momentum is the worst performing signal model on both a gross and net of fees basis. Moreover, when considering trading costs, 21- and 63-day style momentum suffers the worst decline due to its high turnover. Second, consistent with the Cakici et al. (2022) we find that the inclusion of price momentum in factor spanning regressions dramatically reduces style rotation alphas, however, the impact is sensitive to the weighting method applied. Third, of the ML models considered, XGBoost is

the most robust in generating style rotation signals, followed by LightGBM. In robustness tests, we show that XGBoost's superior performance is possibly attributable to it being less likely to incorrectly predict style performance ranks using Spearman rank correlations. Additionally, we show that XGBoost achieves the highest proportional precision when considering the intersection between *ex ante* predicted and *ex post* performance across the top 5, 10 and 15 styles in each quarter.

The result points to the benefit of gradient boosters being superior signal generators, which may be directly attributable to such models being more adept to identify and exploit non-linear relationships between historical performance-based style features and their future returns. An important delimitation of this study is the application of only three ML algorithms, all of which are tree-based learners and therefore cannot be deemed to be an exhaustive study of the application of ML to style rotation. A major outcome of this study is that ML driven style rotation strategies on an emerging market bourse, that suffers restraints in terms of liquidity and investable universe of shares, are feasible and profitable. More importantly, the study answers the question posed by Page et al. (2022) in their conclusion regarding the plausibility of 'advanced forecasting techniques' for style rotation. The evidence presented clearly shows that 'advanced' signal models are superior in terms of forecast accuracy and less prone to portfolio churn when compared to style momentum. Areas of future research include broadening the dataset to include other emerging market bourses, considering an exhaustive list of ML models as well as the application of Shapley Additive Explanations ('SHAP') (per Lundberg & Lee, 2017) tests to determine the relative contribution of each feature across ML models, providing insight to model performance based on the standardized set of features applied.

Notes

1. The result is a style portfolio that assigns the highest initial weight to the highest z-score share, with weights decreasing monotonically to the lowest weight for the 30th ranked share. Weights are bound between a maximum of 10% and minimum of 1% and standardized to sum to unity.
2. We limit the number of styles applied in style rotation strategies for robustness purposes, allowing both the naïve and advanced signals to drop styles that are forecast to underperform.
3. As with long-only style portfolios, maximum weight is set to 10% and minimum weight to 1% and normalized to equal unity.
4. The hyperparameter set applied to each of the tree-based learners is constructed to ensure uniformity across the models. This implies that we apply the same set of hyperparameters for the *max depth*, *subsample*, *learning rate*, *n-estimators* and *column sample by tree*. In terms of the hyperparameters chosen, we were guided by the respective model library documentation on the recommended hyperparameter ranges (see https://scikit-learn.org, https://xgboost.readthedocs.html and https://lightgbm.readthedocs.html). Jupyter notebooks and Python code are available on request from the authors.
5. All machine learning models were run using an Apple Macbook Air (M2) with the average time for model fitting and prediction being ±25 mins.
6. The trading costs considered are consistent with South African average brokerage costs for large institutional investors which range from 0.05%-0.18%. We have applied 0.1% for the purpose of brevity but trading costs can increase dramatically for individuals (between 0.25% to 1%).
7. CAPM model applies the market risk-premium, Fama-French 3 factor model includes the market risk premium as well as *SMB* (small minus big) and *VMG* (value minus growth), Carhart includes *WML* (winner minus loser) as well as the Fama-French 3 factors. The Fama-French 5 factor model builds on the Fama-French 3 factor model by including *CMA* (conservative minus aggressive asset growth) and *RMW* (robust minus weak profitability) while the Fama-French 6 factor considers the Fama-French 5 factors but includes *UMD* (up minus down or winner minus loser).
8. In the case of the top 5 styles, average proportions were tested assuming a population mean of 5 (number of top styles) divided by 30 (number of styles) or 16.67%. The population mean assumption was obviously increased for top 10 (10/30 or 33.34%) and 15 (15/30 or 50%).

Authors' contributions

Daniel Page: Conceptualization, empirical design and execution, analysis and interpretation, drafting, revision. David McClelland: Conceptualization, empirical design and execution, analysis and interpretation, review and final approval. Christo Auret: Conceptualization, empirical design, review and final approval. All authors have reviewed the study and provided final approval for the submitted version of this paper.

Disclosure statement

The authors have no interests to declare.

Funding

No funding was received.

About the authors

Daniel Page, PhD - Associate Professor, University of the Witwatersrand, School of Economics and Finance

David McClelland (MCom) - Lecturer, University of the Witwatersrand, School of Economics and Finance

Christo Auret, DPhil - Professor, University of the Witwatersrand, School of Economics and Finance

ORCID

Daniel Page http://orcid.org/0000-0003-2873-4973
Christo Auret http://orcid.org/0000-0002-7357-852X

Data availability statement

All data, code and other supporting materials are available on reasonable request from the corresponding author (daniel.page@wits.ac.za).

References

Arnott, R. D., Kalesnik, V., & Linnainmaa, J. T., (2023). *Factor momentum*. *The Review of Financial Studies*, *36*(8), 3034–3070. https://doi.org/10.1093/rfs/hhad006

Avramov, D., Cheng, S., Schreiber, A., & Shemer, K. (2017). Scaling up market anomalies. *The Journal of Investing*, *26*(3), 89–105. https://doi.org/10.3905/joi.2017.26.3.089

Barroso, P., & Santa-Clara, P. (2015). Momentum has its moments. *Journal of Financial Economics*, *116*(1), 111–120. https://doi.org/10.1016/j.jfineco.2014.11.010

Bouzida, F., & Pn, D. (2020). Is factors timing overrated. Available at SSRN 3697314.

Cakici, N., Fieberg, C., Metko, D., & Zaremba, A. (2022). Factor momentum versus stock price momentum: A revisit. Available at SSRN 4454976.

Carhart, M. (1997). On persistence in mutual find performance. *The Journal of Finance*, *52*(1), 57–82. https://doi.org/10.1111/j.1540-6261.1997.tb03808.x

Daniel, K., & Moskowitz, T. (2016). Momentum crashes. *Journal of Financial Economics*, *122*(2), 221–247. https://doi.org/10.1016/j.jfineco.2015.12.002

Ehsani, S., & Linnainmaa, J. T. (2022). Factor momentum and the momentum factor. *The Journal of Finance*, *77*(3), 1877–1919. https://doi.org/10.1111/jofi.13131

Fama, E., & French, K. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, *33*(1), 3–56. https://doi.org/10.1016/0304-405X(93)90023-5

Fama, E., & French, K. (2015). A five-factor asset pricing model. *Journal of Financial Economics*, *116*(1), 1–22. https://doi.org/10.1016/j.jfineco.2014.10.010

Fama, E., & French, K. (2018). Choosing factors. *Journal of Financial Economics*, *128*(2), 234–252. https://doi.org/10.1016/j.jfineco.2018.02.012

Galakis, J., Vrontos, I., & Vrontos, S. (2021). Style rotation revisited. *The Journal of Financial Data Science*, *3*(2), 110–133. https://doi.org/10.3905/jfds.2021.1.059

Grobys, K., Ruotsalainen, J., & Äijö, J. (2018). Risk-managed industry momentum and momentum crashes. *Quantitative Finance*, *18*(10), 1715–1733. https://doi.org/10.1080/14697688.2017.1420211

Gu, S., Kelly, B., & Xiu, D. (2020). Empirical asset pricing via machine learning. *The Review of Financial Studies*, *33*(5), 2223–2273. https://doi.org/10.1093/rfs/hhaa009

Gupta, T., & Kelly, B. (2019). Factor momentum everywhere. *The Journal of Portfolio Management*, *45*(3), 13–36. https://doi.org/10.3905/jpm.2019.45.3.013

Karatas, T., & Hirsa, A. (2021). Two-stage sector rotation methodology using machine learning and deep learning techniques. *arXiv Preprint arXiv:2108.02838*. 1–38.

Leippold, M., Wang, Q., & Zhou, W. (2022). Machine-learning in the Chinese factor zoo. *Journal of Financial Economics*, *145*(2), 64–82. https://doi.org/10.1016/j.jfineco.2021.08.017

Lundberg, S. M., & Lee, S. I. (2017). A unified approach to interpreting model predictions. *Advances in Neural Processing Systems*, *30*, 1–10.
Ma, T., Liao, C., & Jiang, F. (2023). Timing the factor zoo via deep learning: Evidence from China. *Accounting & Finance*, *63*(1), 485–505. https://doi.org/10.1111/acfi.13033
Neuhierl, A., Randl, O., Reschenhofer, C., & Zechner, J. (2023). Timing the factor zoo. Available at SSRN.
Page, D., McClelland, D., & Auret, C. (2022). Style rotation on the JSE. *Finance Research Letters*, *46*, 102504. https://doi.org/10.1016/j.frl.2021.102504
Patel, R., Choudhary, V., Saxena, D., & Singh, A. K. (2021). *Review of stock prediction using machine learning techniques* [Paper presentation]. 2021 5th International Conference on Trends in Electronics and Informatics, 840–846.
Strader, T. J., Rozycki, J. J., Root, T. H., & Huang, Y. H. (2020). Machine learning stock market prediction studies: review and research directions. *Journal of International Technology and Information Management*, *28*(4), 63–83. https://doi.org/10.58729/1941-6679.1435
Su, C. (2021). A comprehensive investogation into style momentum strategies in China. *Financial Markets and Portfolio Management*, *35*(1), 101–144. https://doi.org/10.1007/s11408-020-00375-z

Appendix 1. Long-only style portfolio proxies

Style	Style Group	Proxy
Price momentum	Momentum	Cumulative price return measured over the previous 252 day trading window, skipping the most recent 21 days
Price momentum (no skip)	Momentum	Cumulative price return measured over previous 252 day trading window
Earnings momentum	Momentum	Change in reported earnings per share adjusted for look-ahead bias
Idiosyncratic momentum	Momentum	Cumulative return measured using returns orthogonalised on the benchmark
200-day moving average	Momentum	Current price level scaled by 200-day price moving average
Log-price second derivative	Momentum	Regression run on log prices over most recent 21-days against time and time squared where coefficient of time squared dictates increasing or decreasing first derivative
Rachev ratio	Momentum	95% conditional VaR scaled by 5% conditional VaR
Probablistic momentum	Momentum	t-statistic calculated using excess returns measured over the most recent 252 days
Moving Avg & Second Derivative	Momentum	Combination of 200 day moving average and log-price second derivative
Value (Book-to-market)	Value & Growth	Latest book-to-market ratio adjusted for look-ahead bias
Growth (Book-to-market)	Value & Growth	Latest book-to-market ratio adjusted for look-ahead bias
Value (Earnings yield)	Value & Growth	Latest earnings yield (earnings per share scaled by price per share) adjusted for look-ahead bias
Growth (Earnings yield)	Value & Growth	Latest earnings yield (earnings per share scaled by price per share) adjusted for look-ahead bias
Value (Cashflow-to-price)	Value & Growth	Latest cashflow per share scaled by price adjusted for look-ahead bias
Growth (Cashflow-to-price)	Value & Growth	Latest cashflow per share scaled by price adjusted for look-ahead bias
Value (Dividend yield)	Value & Growth	Trailing 12 month dividends scaled by price
Growth (Dividend yield)	Value & Growth	Trailing 12 month dividends scaled by price
Market Beta	Low Volatility/Risk	252-day market beta estimated using regressing share returns on the market benchmark
Idiosyncratic risk	Low Volatility/Risk	Standard deviation calculated using residuals orthogonalised on the market benchmark
Currency risk	Low Volatility/Risk	252-day currency beta estimated by regressing share returns on changes in the USDZAR exchnage rate
Debt-to-assets	Quality	Latest debt-to-assets adjusted for look-ahead bias
Debt-to-equity	Quality	Latest debt-to-equity adjusted for look-ahead bias
NOPAT-to-assets	Quality	Latest net operating profit after tax scaled by assets adjusted for look-ahead bias
Return on equity	Quality	Return on equity (Net profit scaled by shareholder equity) adjusted for look-ahead bias
(Strader et al., 2020)*Return on assets*	Quality	Return on assets (Net profit scaled by total assets) adjusted for look-ahead bias
Return on invested capital	Quality	Net operating profit scaled by net operating assets
Asset growth	Quality	Change in reported assets where assets are adjusted for look-ahead bias
Sales growth	Quality	Change in reported sales/revenues where assets are adjusted for look-ahead bias
Book-value equity growth	Quality	Change in reported book-value per share adjusted for look-ahead bias
Net profit growth	Quality	Change in reported net profit adjusted for look-ahead bias

Appendix 2. Machine learning features

Feature	Detail
Excess style momentum	Cumulative style returns in excess of the market proxy measured over the previous 1-12 months (12 features)
Excess style volatility	Standard deviation of excess style returns measured over the previous 1-12 months (12 features)
Excess style kurtosis	Kurtosis of excess style returns measured over the previous 1-12 months (12 features)
Excess style skewness	Skewness of excess style returns measured over the previous 1-12 months (12 features)
Information ratio	Historical excess style return scaled by standard deviation measured over the previous 1-12 months (12 features)
Omega ratio	Cumulative positive excess style returns scaled by cumulative negative excess style returns measured over the previous 1-12 months (12 features)
VIX correlation	Correlation between styles and changes in the VIX index measured over the previous 1-12 months (12 features)
Short-technical indicator	Historical 21-day moving average of cumulative excess style returns scaled by 252-day moving average (11 feature)
Medium technical indicator	Historical 63-day moving average of cumulative excess style returns scaled by 252-day moving average (9 feature)
Batting average	Excess style performance consistency (number of positive days scaled by total days) measured over 1-12 months (12 features)
Time-weighted batting average	Excess style performance consistency (number of positive days scaled by total days) measured over 1-12 months applying a degradation factor that upweights more recent periods (12 features)
Probabilistic momentum	t-statistic/p-value of excess style momentum measured over the past 1-12 months (12 features)
Rachev ratio	Rachev ratio measured on excess style returns over the past 1-12 months (12 features)
Relative volatility	Style kurtosis scaled by market volatility over the past 1-12 months (12 features)
Relative kurtosis	Style kurtosis scaled by market kurtosis over the past 1-12 months (12 features)
Relative skewness	Style skewness scaled by market skewness over the past 1-12 months (12 features)
Term-structure correlation	Correlation between excess style returns and South African term structure (proxied by the generic 10-year SA Government bond in excess of the 91-day treasury bill) measured over the past 1-12 months (12 features)
USDZAR exchange rate correlation	Correlation between excess style returns and changes in the USDZAR exchange rate measured over the past 1-12 months (12 features)

The features considered are all measured in excess of the benchmark or relative to a specific macroeconomic proxy. Each style is estimated across three themes namely excess performance, excess risk and excess risk-adjusted performance. We include three additional macroeconomic proxies, specifically correlation with the VIX (implied volatility of S&P500 options), SA bond yield spread and changes in the South Africa Rand/US Dollar exchange rate. The purpose of the macroeconomic based features is to introduce style sensitivity to global and local economic proxies that represent global risk-sentiment (VIX) and local risk-sentiment (yield spread and currency performance). Importantly, the total number of individual features totals 212 as each of the 13 feature groups consider 9-12 periods.