Proceedings of the 4th International Conference on Business and Policy Studies
DOI: 10.54254/2754-1169/2025.28879

Exploring the Redundancy and Novelty in Factor Discovery: A Literature Review on Correlations Between Investment Factors

Hao Yin[1*], Yuehua Wang[2†],Ziyuan Wang[3†]

[1]Westa College, Southwest University, Chongqing, China
[2]School of Business, University of Sunderland, Sunderland, UK
[3]Business School, Minnan Normal University, Zhangzhou, China
**Corresponding Author. Email:3359839235@qq.com*
[†]These authors contributed equally to this work and should be considered as co-sceond author.

Abstract. Often referred to as the "factor zoo," the fast spread of investment variables has sparked intense discussion about the true uniqueness of many recently discovered variables. With an eye toward the correlations that can compromise their uniqueness, this literature review investigates the degree to which these elements simply restate information already obtained by other elements. The consequences of these correlations for portfolio building, diversification techniques, and risk management are investigated in this review. It uses important research, including the significant work by Feng et al., in order to give a whole picture of the redundancy and value of investment variables. The results provide insights on improving factor-based investment strategies by stressing the difficulties in differentiating between really unique elements and those that repeat existing knowledge.

Keywords: Factor Zoo, Portfolio Diversification, Risk Management, Factor Correlation, Investment Strategy

1. Introduction

Driven by the search of methodical sources of returns outside the traditional bounds of asset classes, factor investing has evolved into a basic component of modern portfolio management. Extensive research on the elements of value, scale, momentum, and quality has produced investing methods including these elements. But the abundance of elements, sometimes known as the "factor zoo," begs the issue of whether many of the recently identified factors are really unique and valuable. Concerns have been raised as the number of discovered factors keeps rising that many of them are highly associated with known ones, hence providing little fresh information [1-3].

The main goal of this evaluation is to find the degree to which recently discovered elements provide special insights instead of only restating knowledge already gathered by current components. This paper especially looks at the relationships between elements and their consequences for portfolio building, diversification, and risk control [4]. The results will give

DOI: 10.54254/2754-1169/2025.28879

practitioners and researchers new perspectives on the possible duplicity in factor discovery and the difficulties it presents for sensible portfolio management.

2. Overview of factor investing

Targeting particular traits or factors, factor investing is an investment strategy whereby inclusion is intended to produce risk-adjusted returns. Investment strategies have been developed using much attention to the conventional elements of value (stocks with low price relative to fundamentals), size (small-cap stocks), momentum (stocks with strong past performance), and quality (stocks with strong fundamentals) [4,5]. The theoretical foundation of factor investing is based on asset pricing models, including the Capital Asset Pricing Model (CAPM) and the Fama-French three-factor model, which establish a link between expected returns and specific factors [3].

Advances in statistical methods and data availability have propelled factor investing's expansion by allowing academics to find an increasing range of elements supposedly explaining asset returns [1,2]. This has produced hundreds of elements across a spectrum of traits including liquidity, volatility, and more esoteric elements including corporate social responsibility and political risk [2]. But the great number of variables raises questions about data mining and the overfitting of models to historical data, which could not be indicative of future results [6].

The issue of factor redundancy, whereby newly identified factors are merely variations of established ones, presents a significant challenge for investors seeking to construct diversified portfolios [1]. Concentrated exposures and less than optimum diversification strategies might result from erroneous risk assessment resulting from element correlations [2]. Therefore, if one is to build a competent portfolio and control risk, one first has to grasp the actual informational content of components and their interrelationships [3].

3. Correlation and redundancy in factors

The study of relationships between variables is a fundamental part of determining factor originality. Based on extensive research by Feng et al [1], several of the freshly identified elements are somewhat closely linked with existing elements. This implies that the recently found elements reflect modifications of known components instead of offering new information [3]. This duplicity can mislead investors since overlapping pieces would seem to provide mistakenly positive diversity [2].

Chordia, Subrahmanyam and Tong [5] showed that when one considers their relationships with known variables, the explanatory value of novel factors is much diminished. When considering the implications of correlation, the results show that many recently discovered elements merely have a little incremental effect on asset pricing models. Moreover, the presence of redundant elements might cause overfitting in asset pricing models, therefore reducing their resilience and predictive capacity in out-of-sample research. This problem is especially relevant given the inclination of elements to show excellent performance when tested on data from the same sample (in-sample), but to underperform or vanish when tested on fresh data [1].

Another important study by Pukthuanthong and Roll [7]looked at the redundancy of global risk factors and came to the conclusion that numerous often regarded as unique sources of hazard are actually rather closely linked with traditional risk variables. This repetition makes it more difficult to find really creative ideas that could lower portfolio risk and boost diversification [2]. Moreover, Moskowitz [8] suggested that the multiplication of related variables causes a behavior he dubbed

DOI: 10.54254/2754-1169/2025.28879

"factor crowding". This event, which could be more clear-cut during market downturns, drives investors to unwittingly focus their risk exposure on a small range of underlying risks [6].

3.1. Statistical challenges in identifying redundant factors

The presence of statistical problems complicates the process of spotting repeated elements even farther. Many factor discovery methods are sensitive to false positives coming from the use of different testing approaches, the effect of selection bias, and the intrinsic noisy quality of financial data [2,5]. Academic research, for instance, routinely tests hundreds of possible variables without sufficiently correcting for the likelihood of producing statistically significant results by chance [3]. The phenomena of data snooping biases, whereby researchers, either deliberately or inadvertently, choose elements that match historical data but lack a real economic justification, aggravates this problem even more [5].

Researchers have proposed several approaches meant to manage for data mining and redundancy in order to handle these statistical issues [1]. Using strict out-of-sample testing and cross-valuation methods helps one confirm the validity of recently identified components [2]. Moreover, machine learning techniques have been applied to examine large amounts of data and pinpoint elements that really differ from those that have been found before [9]. Even these advanced methods, however, are not immune to the problem of repetition since many elements show notable interaction in their informational content [2].

4. Implications for portfolio management

Factor correlations have important consequences for portfolio building, diversification techniques, risk management [4]. High correlation between factors can cancel out the advantages of diversity as, especially in times of market stress, similar factors usually show similar performance throughout a spectrum of market situations [7]. This phenomena, known as "factor crowding," shows when several components exhibit similar performance patterns, hence magnifying drawdowns during crisis [10].

Empirical data shows that factor correlations rise in times of market stress, therefore lessening the advantages of diversification at exactly the moment most needed [5]. Asness, Moskowitz and Pedersen [4], for instance, showed that elements like value and momentum often become more linked during market downturns, therefore offsetting the supposed diversification advantages linked with these elements [7]. As several risk exposures arise concurrently, this convergence of factor performance might cause a clear decrease in portfolio value [6].

Moreover, good risk control depends on a knowledge of factor correlations [3].Redundant features may seem to provide new dimensions of risk exposure but, in fact, they just reflect present ones, hence disguise the actual risk profile of a portfolio [2]. This could cause misallocation of money and an underestimating of risk levels, which at last generates fewer than ideal investment options [11-12]. Moreover, the way diversification shown by related events is misconstrued could lead investors to ignore underlying dangers, especially in volatile markets [5].

4.1. The impact of factor correlations on asset allocation

Decisions on asset allocation clearly depend on factor correlations, which also demand more research [5]. High factor correlation reduces the incremental effect of every extra component on portfolio risk and return [3]. This can lead to a concentration of risk exposures not very clear from a

DOI: 10.54254/2754-1169/2025.28879

naive diversification standpoint [5]. Jacobs and Levy [13], for instance, argue that portfolios built with many related parts typically show more volatility and worse risk-adjusted returns than those produced of highly dissimilar constituents [3].

Asset managers have a responsibility to closely evaluate the interactions between factors and take into account such relationships for asset allocation in order to reduce these risks [5]. Principal component analysis (PCA) and other methods can be used to determine the basic determinants of factor performance and hence simplify factor models by removing duplicated factors [3]. Furthermore helping managers to find any shortcomings in factor-based portfolios is stress testing and scenario analysis, which helps them to modify their allocation as necessary [4].

4.2. Implications for diversification and drawdown risk

Redundant components directly influence the variety of portfolios and drawdown risk [3]. Although in actuality one is focusing on same risk factors, the existence of duplicate elements may give the impression of diversity [5]. During times of market stress, this can especially be troublesome since correlated elements often perform poorly together, hence aggravating drawdowns [8]. Factor correlations tend to rise during market downturns, according to Ang, Hodrick, Xing, and Zhang [11], therefore greatly reducing the diversification advantages that factors are meant to offer [7].

Moreover, the presence of redundant elements causes difficulties with respect to risk attribution and performance assessment [5]. Investors may find it difficult to separate elements that are actually causing returns from those that are only helping other linked aspects perform [2]. This can lead to poor investing choices since the overlap in factor exposures causes the apparent risk-adjusted performance of a portfolio to be exaggerated [3].

5. Behavioral vs. rational risk factors

The argument on whether investment elements reflect rational or behavioural risks adds still another level of complication to the analysis of factor redundancy [2]. Those elements judged to be logical are those that reflect methodical risks for which investors are rewarded [1]. Among these elements are those of market, size, and value aspects [4]. In contrast, behavioural factors reflect predictable patterns in investor behaviour, such as overreaction, herding, and sentiment, which can result in exploitable mispricing [12].

It is difficult to separate between elements that are truly unique and those that only reflect fluctuations of existing hazards since the possibility of both rational and behavioral elements to demonstrate strong connections. Value and quality criteria, for example, usually show a positive link since they both cover aspects of a company's basic characteristics, such stability and profitability [2]. Likewise, given the impact of investor psychology on stock price movements, momentum and sentiment elements could coincide [3].

Understanding the nature of these correlations is absolutely crucial for factor investors since it directly affects the durability of factor returns [5]. Over the long run, rational considerations are expected to offer constant risk-adjusted returns [4]. On the other hand, behavioural elements could present fleeting chances that fade when market players learn of them [9]. To build strong factor-based portfolios, then, it is imperative to separate logical from behavioural elements and consider their relationships [3].

DOI: 10.54254/2754-1169/2025.28879

6. Conclusion

Examining the body of research on investment variables exposes a rising worry about the duplicity of recently discovered variables and how they affect portfolio management [1]. Often, correlations between elements show that many so-called new components are really variants of existing ones, with only little additional value [3]. The design of an ideal portfolio can be complicated by the presence of several elements with comparable properties [5]. Overlapping elements could skew risk evaluations and lower the benefits of diversity a portfolio offers [2].

Future studies should focus on the development of more strong techniques for the identification of truly unique factors with an eye toward separating their unique contributions from those of current factors [1]. This covers the development of out-of-sample testing methods, statistical methods to adjust for data mining, and machine learning tools to find discrete factors [2,11]. Moreover, the factor investing scene would be much improved by empirical studies that carefully analyze the correlations and redundancies of components under various market environments [7].

Finally, the capacity to separate truly distinctive elements from duplicate restatements would improve the efficacy of factor-based investing methods, therefore helping investors in the management of risk and the realization of outstanding portfolio results [3].

References

[1] Feng, G., Giglio, S., & Xiu, D. (2020). Taming the factor zoo: A test of new factors. Journal of Finance, 75(3), 1327-1370
[2] Harvey, C. R., Liu, Y., & Zhu, H. (2016). …and the cross-section of expected returns. Review of Financial Studies, 29(1), 5-68.
[3] Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. Journal of Financial Economics, 33(1), 3-56.
[4] Carhart, M. M. (1997). On persistence in mutual fund performance. Journal of Finance, 52(1), 57-82.
[5] Chordia, T., Subrahmanyam, A., & Tong, Q. (2014). Have capital market anomalies attenuated in the recent era of high liquidity and trading activity? Journal of Accounting and Economics, 58(1), 41-58.
[6] McLean, R. D., & Pontiff, J. (2016). Does academic research destroy stock return predictability? Journal of Finance, 71(1), 5-32.
[7] Pukthuanthong, K., & Roll, R. (2009). Global market integration: An alternative measure and its application. Journal of Financial Economics, 94(2), 214-232.
[8] Moskowitz, T. J. (2010). The market risk of hedge funds. Journal of Finance, 65(4), 1247-1288.
[9] Gu, S., Kelly, B., & Xiu, D. (2020). Empirical asset pricing via machine learning. Review of Financial Studies, 33(5), 2223-2273.
[10] Asness, C. S., Moskowitz, T. J., & Pedersen, L. H. (2013). Value and momentum everywhere. Journal of Finance, 68(3), 929-985.
[11] Ang, A., Hodrick, R. J., Xing, Y., & Zhang, X. (2006). The cross-section of volatility and expected returns. Journal of Finance, 61(1), 259-299.
[12] Barberis, N., Shleifer, A., & Vishny, R. (1998). A model of investor sentiment. Journal of Financial Economics, 49(3), 307-343.
[13] Jacobs, B. I., & Levy, K. N. (2019). Enhanced active equity strategies. Financial Analysts Journal, 75(3), 56-71.