Proceedings of CONF-BPS 2026 Symposium: GenAI, Labour Markets, and the Economics of Human and Financial Capital
DOI: 10.54254/2754-1169/2026.LD31217

Taming the Factor Zoo: A Comprehensive Review of Methodologies for Screening Factor Redundancy in Asset Pricing Models

Jiaxin Yu

PSB Academy, Singapore, Singapore
908169738@qq.com

Abstract. Empirical asset pricing has been changing and improving rapidly, moving from the single-factor CAPM to multiple-factor models. But on the other hand, this has also brought the so-called "Factor Zoo," that is, hundreds of potential pricing factors have been found, which leads to a big problem of dimension and redundancy. The method used in this paper is a systematic literature review and theoretical analysis, the main difficulty being to identify and remove redundant factors. The Fama-French three-factor model to the q-factor model. The paper also discusses existing screening methods for redundancies, such as spanning regression, and new techniques like machine learning. The findings of the study are that traditional spanning tests are still important, but not enough for high-dimensional data. Advanced statistics like regularization and dimension reduction give us better ways to tell apart the real pricing reasons and extra noise that doesn't matter. This paper concludes that the future of research in asset pricing is to develop screening methods based on economics. This is necessary to resolve the redundancy problem and restore simplicity to factor models.

Keywords: Factor Zoo, Factor Redundancy, Asset Pricing, Multi-factor Models

1. Introduction

The drivers of fluctuations in asset prices have constituted the defining theme of financial economics for at least several decades. It started with the Capital Asset Pricing Model (CAPM), which said that the only thing that determined the expected return on a security was the market risk (Beta). But since then came things like the size effect or the value effect as empirical anomalies showing that there is more than just one factor needed for these. This section introduces Fama and French's three-factor model [1] and their subsequent five-factor model [2], which incorporates the investment and profitability factors, as well as the q-factor model [3].

However, in spite of all this progress in theory, the academic world of finance faces a crisis known as the "Factor Zoo." Researchers have found more than 200 different anomalies and factors, and which leads to an important question about the validity and independence of the factors. The central problem is factor redundancy, with different factors often capturing the same underlying economic risk, resulting in multicollinearity and overfitting of pricing models. For example, a value

DOI: 10.54254/2754-1169/2026.LD31217

factor and a profitability factor may explain the same amount of return variance, meaning including them both would be redundant and noise for the model.

The focus of this paper is on the critical "unresolved issue" that has emerged recently in the research outlooks, i.e., the ways of screening for factor redundancy. With the proliferation of factors, the space for risks seems rich, but with no agreement on how to filter out redundant factors, the asset pricing models are not parsimonious or predictive. To review the progression of standard valuation factors, behavioral and specific market factors (e.g., CH-3 for China) for this study's goal of synthesizing the various types of redundancy screening.

The importance of this study is that it may provide a resolution to the "Factor Zoo" debate. We attempt to examine various models that incorporate these variables, including the Fama-French 5-factor model and the Hou-Xue-Zhang q-factor model, among others. This allows us to gain a rough understanding of the complexity and explanatory power of these models, enabling us to gain insight into the trade-offs between them. Additionally, as machine learning plays an increasing role in factor mining, it is important for the next generation of asset pricing theory to develop robust screening protocols.

2. The evolution of factor models and the emergence of redundancy

The basic idea of any factor model is that the expected excess return of an asset is a function of the exposure to a set of risk factors and the risk premium for each risk factor. The Capital Asset Pricing Model stated that the only risk that should be rewarded is systematic risk [4]. But, by the 1970s and 80s, evidence started to come in that CAPM couldn't explain certain anomalies. Such as why small-cap and high book-to-market stocks tend to outperform the market [1]. To solve this problem, Fama and French added a size factor and a value factor to the market factor [1]. This model has greatly improved our ability to explain cross-sectional return differences. But some new ones popped up. In 2015, Fama and French extended their model to five factors, adding profitability and investment factors derived from the dividend discount model, which associates higher expected returns with higher profitability and lower investment [2].

In addition, Hou, Xue, and Zhang also put forward the q-factor model [3]. The model is based on investment theory. It uses 4 factors, including market, size, invest, and profit. The Q-factor model contends that most anomalies can be explained by firm investment and profitability. Its supporters also said that the Fama-French value factor was no longer needed because investment and profitability had been incorporated [3]. And this revealed the problem of factor redundancy, where a "new" factor is simply an extra-noisy version of existing factors.

As research methods became easier to obtain, the number of published factors began to grow very quickly. This is known as the "Factor Zoo," and there are more than 200 factors that have been identified, including momentum and complex accounting ratios [5,6]. The key problem isn't simply a lot of people, but many lack a good economic base [5,6]. Some models just try to "explain the anomaly" with no theory, and some are really redundant. If factor A and factor B are 90% correlated, including both adds noise, not information. A host of factors exist in the current context; the crux of the issue lies in sifting through these factors to identify the genuine independent source of priced risk.

3. Methodologies for screening factor redundancy

It is a difficult statistical and theoretical problem to identify the redundant factors. According to the current literature review and development trend of factor models, this paper classifies the screening methods into three main types: spanning regressions, economic theory constraints, and machine learning/statistical reduction.

3.1. Spanning regression method

The most traditional and popular way to check for redundancy is to perform a "spanning test" with time series regression. This is the workhorse of comparing Fama-French models to the q-factor model. The logic is to test whether factor X is redundant relative to a set of benchmark factors (Market, Size, and Profitability), we regress the returns of factor X on the returns of the benchmark factors.

$$RX,t = \alpha + \beta 1 RMkt,t + \beta 2 RSize,t + \beta 3 RProf,t + \epsilon t$$

If the intercept (Alpha) of this regression is statistically indistinguishable from zero, it would mean that the benchmark factors explain the average returns of Factor X completely. In this case, factor X is considered to be redundant as it doesn't offer any additional expected return over and above the benchmark. This method is most famous for arguing that the Value factor (HML) is unnecessary in the presence of the Investment and Profitability factors. However, there are limitations to spanning tests. They are very sensitive to the benchmark model chosen. If the benchmark model is wrong, the redundancy test is useless. Furthermore, with hundreds of factors in the Zoo, running pairwise spanning tests becomes computationally and statistically infeasible due to the multiple testing problem.

3.2. Economic theory constraint method

Second is to screen by economic mechanism rather than just statistical correlation. This matches the discussion in research outlooks to verify economic principles on factors. Some factors are screened for whether they are rational risk compensation or behavioral mispricing. Take the Stambaugh Mispricing Factor Model as an example that combines anomalies with "mispricing" based on behavioral biases [7]. Redundancy screening must adapt to the local conditions of specific markets. CH-3 for China is an example [8]. It found that the standard Fama-French value factor is redundant or ineffective in China because of "shell value" in small caps. By cutting out the smallest 30% of stocks and using earnings-to-price rather than book-to-market, they constructed a simpler three-factor model that worked better [8]. This shows that "redundancy" is context-specific; a factor might be unique in the US but redundant in China.

3.3. Machine learning/statistical dimensionality reduction method

Machine learning of factors is the direction of the future for redundancy screening. Traditional OLS runs into trouble when there are lots of predictors. Principal Component Analysis transforms correlated variables into uncorrelated components. Researchers can keep the first few components that explain most of the variance and reduce 200 factors to 3 to 5 latent factors. But PCA has an issue–the results are tough to understand. They are mathematical creations rather than plain old "Value" type economic ideas. So to help make sense of it, people use other things like Lasso more. Lasso minimizes the error and penalizes large coefficients, so it sets the weak or redundant factors zero automatically. So Lasso selects the variables [6]. It aligns with the screening goals. Two things that are very correlated, then Lasso picks one and throws out the other. So it ends up with something which says I only take a main factor, and I only take one of these things or another of these.

DOI: 10.54254/2754-1169/2026.LD31217

4. Discussion

4.1. Argument of the redundancy of factors

There's an argument as to the redundancy of factors. The argument is not simply a matter of statistics. Moreover, a question of philosophy. A common analogy in the literature is to think of factors as "nutrients" and assets as "food." As nutrition science seeks to find which vitamins are important, so asset pricing seeks to find which risk premiums are important. In other words, the objective is not to use all the factors that were ever discovered, but to find those few drivers that really matter for explaining asset returns. The endgame is to build a simpler, clearer, and stronger pricing model.

But the problem is that "statistical redundancy" doesn't mean "economic redundancy." You can have two factors that are statistically correlated with each other, which means that the returns of the factors tend to go up and down at the same time. But the two factors could be coming from totally different economic or risk sources. For example, a factor based on firm investment and a factor based on accounting profitability will be highly correlated in the data. But the first is a sign of rational, firm decisions about future growth, while the second may be more of a short-term market reaction or mispricing of earnings. We end up removing things that are really having some unique economic signal, but just happen not to be very different from other factors (which the statistical test will highlight). So, we need to be able to dig beneath the figures and so on. Hence, a good screening framework cannot just rely on the numbers, it also includes the economic logic/theory for each factor.

Additionally, the "Factor Zoo" problem is exacerbated by data mining bias. Many of the factors that seem important are just important because the researchers have looked at thousands and thousands of different combinations of variables or models. This is the classic "multiple testing" issue-if you run enough statistical tests, some are bound to be statistically significant just by chance. Thus, more careful screening of factors becomes urgent. We need methods to separate the wheat from the chaff–real, persistent effects versus false discoveries. Other successful examples, such as Stambaugh's mispricing factor, and China's own CH-3 factor, also show that successful screening often has a combination of top-down theory and bottom-up stats. Theories give us the place to start and the reason why we include something, but the data testing tells us if it really works or not.

4.2. Methods of screening redundancy

The main methods of screening factor redundancy can be classified into three types. They look at different things, but it doesn't have to be one or the other. The first method is spanning regression. This approach is focused on Statistical Testing. To know whether returns of new factors can be explained completely by the benchmark factors. If it can, then the new factor is redundant. This method's strength is that the logic is quite straightforward, which gives a direct statistical basis. But it's not all good. This result is dependent on the benchmark model we select. If the benchmark is off, the conclusion is off. Additionally, running the pairwise spanning tests would become quite difficult with so many factors. It creates a big computational task and makes mistakes more likely because there are many hypothesis tests.

Second, Economic Theory Constraint. In this method, the meaning of the factors is economic in nature. It should have a strong theoretical background or reasonable financial logic in economics before it is considered. For example, the investment factor according to the q-theory, or the profitability factor according to the dividend discount model. This method helps to avoid pure data

DOI: 10.54254/2754-1169/2026.LD31217

mining. It makes sure that the chosen factors can be explained economically and could do better out of sample. But the challenge is that economic theory itself can be debatable or slow to develop. It could not pick up the quickly changing market situations sometimes. This can exclude some empirically useful factors.

The third method is Machine Learning and Statistical Dimensionality Reduction. These techniques are designed for high-dimensional data. Take principal component analysis (PCA) as an example, which can transform many related factors into a few unrelated components. Methods such as LASSO add a penalty to the regression that forces the coefficient on weak or redundant factors towards zero, thus selecting variables. And these methods are good at dealing with the scale of the "Factor Zoo." But they have problems too. The components from PCA tend to be linear combinations of the original factors and are difficult to interpret economically. Machine learning models can choose factors, but these are "black-box," and the selection process is not obvious. It's hard to explain the choices using regular economics talk.

These three methods can be used independently. They can, and they should work together in a combo. A possible step-by-step procedure would be to start with economic theory to define the initial set of candidate factors. Step 1: Exclude factors that are obviously without a theoretical basis. Step 2: apply spanning regression tests using core models that have been well established (e.g., the 5-factor model) as a first screening. This step removes factors that can be fully explained by the benchmarks already present. Third, run machine learning or dimensionality reduction techniques on the large set of factors remaining. Step 3: refine and optimize the selection of factors, arriving at a final small group of factors which is statistically effective and economically meaningful. This multi-stage approach takes advantage of the best parts of each way, like using several tools instead of just one, and that can make it better.

4.3. Challenges

However, the current method is still faced with the core problem of how to perfectly match the definition of "statistical redundancy" and "economic redundancy." Each method has its limitations. Statistical methods may mistakenly remove some useful factors due to the specification of models or data issues. Theory-based methods may hold on to the ineffective elements because of the limits of the theory. Machine learning may give up on interpretability in favor of a better fit. These shortcomings highlight important areas for further research. We will have to look for a framework or methodology which introduces this logic directly and is more flexible. For example, the development of regularization algorithms that support theoretical priors, or to create comprehensive metrics measuring the statistical significance and economic plausibility. And the relationship between them is not fixed. They can shift as the market environment, regulations, and investor behavior change over time. Therefore, it should also take a dynamic approach in the future screening framework. They need to be able to detect and adapt to such changes. And if we keep chipping away at those problems we'll get closer to harnessing the zoo of factors. That would allow for building of simpler and more robust asset pricing models built in an economic framework. This progress is significant both academically and for real-world investment.

5. Conclusion

This study found that although growth of factors can give you more ways to look for alphas, you need strict screening. The traditional spanning tests, the theory-based adaptations of them, and the new machine learning techniques. As for the three types of methods, they are not contradictory but

DOI: 10.54254/2754-1169/2026.LD31217

complementary: Spanning Regression gives us a basic statistical benchmark; Economic Theory Constraint gives us the main logic; Machine Learning/Statistical Reduction gives us the efficient handling for high-dimensional data. But they have not completely solved the problem of matching statistics with economic redundancy. This creates room for further research to create a fuller screening framework.

This research adds to the work that puts together various ways of looking for problems in one place. It indicates that redundancy is not an all-or-nothing concept, and it is model-dependent. For the practitioner and portfolio manager, this study matters for quantitative investing. Knowing how to screen out redundant factors helps build better, cheaper trading strategies.

One shortcoming of the study is that it only reviews current models, and does not perform a new test for all 200+ factors. Furthermore, machine learning requires high-quality datasets and a history of that data which will limit its usage in markets which have very little historical data.

Future research could focus on the "Unresolved Issues," like trying to build a single, Lasso-type framework for the economy. It would also be interesting to study the way that factor redundancy changes over time.

In short, this paper is about giving some new ideas about the zoo of factors and how much of a big deal it is to have a simple model. By describing the redundancy screening mechanism, this paper can contribute to more solid, interpretable, and theory-supported models of asset pricing in the future.

References

[1] 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.

[2] Fama, E. F., & French, K. R. (2015). A five-factor asset pricing model. Journal of Financial Economics, 116(1), 1-22.

[3] Hou, K., Xue, C., & Zhang, L. (2015). Digesting anomalies: An investment approach. The Review of Financial Studies, 28(3), 650-705.

[4] Fan, J., Li, R., Zhang, C.-H., & Zou, H. (2020). Statistical Foundations of Data Science (1st ed.). Chapman and Hall/CRC. https: //doi.org/10.1201/9780429096280

[5] Cochrane, J. H. (2011). Presidential address: Discount rates. The Journal of Finance, 66(4), 1047-1108.

[6] Harvey, C. R., Liu, Y., & Zhu, H. (2016). … and the cross-section of expected returns. The Review of Financial Studies, 29(1), 5-68.

[7] Stambaugh, R. F., & Yuan, Y. (2017). Mispricing factors. The Review of Financial Studies, 30(4), 1270-1315.

[8] Liu, J., Stambaugh, R. F., & Yuan, Y. (2019). Size and value in China. Journal of Financial Economics, 134(1), 48-69.