



CAFÉ WORKING PAPER NO.29

Centre for Accountancy, Finance, and Economics (CAFÉ)

What Determines Equity Returns in Emerging Markets?

James Foye

April 2024

The views and opinions expressed in this paper are solely those of the author(s) and do not necessarily reflect those of Birmingham City University. This Working Paper presents research in progress by the author(s) and is published to elicit comments and further debate

Manuscript version: Working paper (or pre-print)

The version presented here is a Working Paper (or 'pre-print') that may be later published elsewhere.

Persistent URL: <https://www.open-access.bcu.ac.uk/>

How to cite: Please refer to the BCU open access repository item page, detailed above, for the most recent bibliographic citation information.

Copyright and reuse: The Centre for Accountancy, Finance, and Economic (CAFÉ) Working Paper Series makes this work by researchers of Birmingham City University Business School available open access under the following conditions.

Copyright © and all moral rights to the version of the paper presented here belong to the individual author(s) and/or other copyright owners.

To the extent reasonable and practicable, the material made available in CAFÉ Working Paper Series has been checked for eligibility before being made available.

Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge.

Provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

Views and opinions statement: The views and opinions expressed in this paper are solely those of the author(s) and do not necessarily reflect those of Birmingham City University.

This Working Paper presents research in progress by the author(s) and is published to elicit comments and

What Determines Equity Returns in Emerging Markets?

James Foye*

Abstract

Although factor models provide the cornerstone of financial asset pricing, their application to emerging markets has been fragmented and frequently returns inconsistent results. This paper is the first to provide a thorough evaluation of these varied findings, thereby providing a framework to guide future researchers. This not only sheds new light on returns in emerging markets but also offers broader insights into the risk versus mispricing debate in all markets.

JEL Classification: G12, G14, G15

Keywords: Anomalies, Emerging markets, Fama-French five-factor model, Carhart four-factor model, Fama-French three-factor model

* College of Accounting, Finance and Economics. Birmingham City University. james.foye@bcu.ac.uk

1. Introduction

Interest in both emerging markets and factor investing has grown rapidly in recent decades. Emerging markets have become increasingly important destinations for investors seeking high returns and diversification. Meanwhile, the Fama French (1993) three-factor model (FF93)– and subsequent extensions to it – have become the dominant methods for estimating equity returns. This vast body of research frequently returns inconsistent findings and now requires a critical synthesis.

Harvey (1995) was the first to provide a comprehensive analysis of the new equity markets created by the global trend of liberalisation. The characterisation of these markets by high returns and low correlations with developed counterparts attracted investors and motivated a range of subsequent research (see, e.g., Bekaert et al., 1998a, b; Rouwenhorst, 1999). Inevitably, this led to increasing portfolio inflows into emerging markets. Lucey et al. (2018) provide a detailed evaluation of the literature covering the effect of these inflows on integration between developed and emerging markets.

The portfolio formation process associated with testing factor models requires a significant number of stocks. This hampered early attempts to evaluate factor models in emerging markets. For example, Fama and French's (1998) early international test of the size and value effects reports detailed asset pricing tests for developed markets but only summary statistics for emerging markets. A lack of stocks would have prevented conducting a more comprehensive analysis in emerging markets. However, with time the number of stocks listed increased and Cakici et al. (2013) were able to provide the first thorough evaluation of FF93 in emerging markets. Although extensive subsequent research has been conducted, the field is highly fragmented and often reports contradictory findings.

Hence, there is now a requirement for not just a synthesis but a detailed critical analysis of this growing body of research. By cataloguing differing findings and offering possible explanations, the main contribution of this paper is to help reconcile contradicting studies, outline gaps in the literature and provide avenues for future research.

The focus of this paper is not CAPM, but rather the terms used to augment it and create multi-factor models. The primary reason for this is that while CAPM is grounded

in rigorous financial theory, other terms are based on empirical observation and their economic interpretation is unclear. Therefore, the paper begins by analysing the evidence concerning FF93's size and factor factors. This is followed by an examination of the momentum factor which has become a popular extension to FF93. The paper then turns to the investment and profitability factors of the Fama and French (2015) five-factor model (FF15). Next, country-specific considerations are examined with a particular emphasis on the implications for the size premium and the effects of national variations in earnings quality. Finally, the paper concludes with a discussion of the often-overlooked liquidity factor.

2. Three- and Four-Factor Models

2. 1. Value

2.1.1. Broad Sample

The inclusion of a value term (as measured by book-to-market equity ratio (B/M)) in the seminal FF93 model has been widely supported by extensive subsequent international research in both developed and emerging markets¹.

Fama and French (1998) provide an early test of the value premium in emerging markets. They show that there is a positive and significant value premium in portfolio sorts using 16 emerging markets. Similarly, Rouwenhorst (1999) finds that value outperforms growth in 17 out of 20 emerging markets. Although HML return could not be considered significant for each individual country, the factor is significant when formed using all 20 emerging markets combined. This prompts the question of whether factors are best formed on a national or cross-country basis, a trade-off which is discussed further in a latter section. Due to short sample periods and lack of stocks, Fama and French (1998) and Rouwenhorst (1999) do not conduct asset pricing tests for emerging markets. Later studies benefited from longer sample periods and were able to report more detailed tests.

More recently, Cakici et al. (2013), Foye (2018a) and Hanauer and Lauterbach (2019) all detect strong value premiums in each of the emerging market regions of Asia,

¹ See Cakici (2015) and Fama and French (2012, 2017) for evidence from developed markets.

Eastern Europe and Latin America. Zaremba and Czapkiewicz (2017) and Foye et al. (2013) confirm these findings for the same markets in Eastern Europe. However, Leite et al. (2018) report contradictory results and find little evidence of a value premium in any of the three regions. One possible explanation for these differing findings is that Leite et al. (2018) use a sample that begins significantly later than the other three studies – perhaps indicating the value premium has disappeared in emerging markets. This would be consistent with McLean and Pontiff (2016) and Schwert (2003) who both report that the profitability of value strategies decreased in the years following publication. However, Leite et al.'s (2018) very late sample start date, and subsequent smaller sample size, must have at least some effect on their findings. Indeed, in contrast to other broad sample emerging market studies, Leite et al. (2018) consistently report weak premiums for all factors – suggesting that the small sample size results in statistically insignificant results. While this paper proposes possible explanations for these contradictory findings, an important task for future researchers is to test these empirically.

One of the most striking findings of Fama and French (2015) is that HML is redundant in the face of investment and profitability factors. Interestingly, Foye's (2018a) factor spanning tests reveal that HML is still important for describing average returns in emerging markets, even in the presence of RMW and CMA. However, inferences from factor spanning tests can be sample specific. It is therefore unfortunate that Cakici et al. (2013), Hanauer and Lauterbach (2019) and Leite et al. (2018) do not report spanning tests as it would have made for an interesting comparison.

Leite et al. (2018) report asset pricing tests for a single factor model using HML alone. The model consistently returns lower GRS statistics and average intercepts than CAPM for Asia and Eastern Europe. The results for Latin America are not quite as conclusive, but the single factor HML model generally outperforms CAPM. Neither Cakici et al. (2013) nor Foye (2018a) report asset pricing tests for a model using HML alone. Therefore, the individual contributions of HML and SMB cannot be identified. However, Cakici et al. (2013) do report results for CAPM and show that FF93 offers an improved description of average returns. However, it is not possible to observe the individual contribution of HML.

2.1.2. Individual Countries

Although the general consensus has been supportive of the value premium in both developed and emerging markets, evidence from the Chinese A-shares market only offers mixed empirical support. Numerous studies have related this inconclusive evidence to China's idiosyncratic market imperfections.

Wang and Xu (2004) and Hu et al. (2019) both form HML factors using B/M and find no evidence of a value premium in China. Wang and Xu (2004) relate their findings to the country's institutional environment. They cite poor accounting quality as a particular concern and replace HML with an alternative factor based on free float which they believe is a proxy for company fundamentals. As this is one of the earliest tests of the determinants of Chinese stock market returns, it is possible that the institutional environment has since developed and accounting quality has improved.

Liu et al. (2019) also fail to detect a value premium. They form HML using alternative value measures and show that a factor based on earnings-to-price ratio (E/P) strongly outperforms the B/M-based factor. Like Fama and French (1992), they run a horse race and treat the choice amongst competing value measures as an entirely empirical issue. Therefore, they offer no explanation as to why E/P subsumes B/M in capturing the value effect in China.

Contrary to the above findings, numerous studies provide strong support for a B/M-based value premium in China. Surprisingly, Eun and Huang (2007) report that despite the well documented idiosyncratic institutional environment, Chinese shares are actually priced rather rationally. Unlike Wang and Xu (2004), they find a significant B/M premium despite the imperfections present during the stock market's early years. Chen (2010) finds that stock return predictability in China is generally low. However, cross-sectional regressions reveal that there is a strong positive relation between the B/M ratio and average returns. The relationship between alternative value measures – cash flow-to-price ratio (CF/P), E/P and sales growth – is nowhere near as strong.

Lin's (2017) initial summary statistics indicate that the HML premium is very low and not statistically significant in China. However, more sophisticated tests reveal that HML is important for describing average returns. HML produces a strong intercept in factor spanning tests – indicating that the factor is useful as its average return is not captured by exposure to other factors. Most importantly, asset pricing tests reveal that HML plays an important role in explaining LHS excess returns. In contrast to Fama and

French (2015), Lin (2017) reports that the HML factor is not redundant for explaining returns.

Cakici et al. (2017) examine the return predictability of alternative value measures. Using cross-sectional regressions and portfolio methods, they report that B/M, CF/P and E/P all have strong predictive power for Chinese listed stocks. Blitz et al. (2021) form HML using B/M and report a positive and significant value premium. They offer little in the way of explanation for this finding as their focus is on testing the volatility effect in China.

Hanauer et al. (2024) report that the 'standard' B/M-based HML factor offers a poor description of returns on Chinese A-shares. However, in contrast to Chen (2010), they find that specifying the value factor using E/P provides a far better description of average returns. Therefore, Hanauer et al. (2024) propose a three-factor model based on market, size, and an earnings-based value factor for use in the Chinese A-shares market.

In summary, despite broad agreement regarding the existence of Chinese market imperfections, the empirical evidence regarding the value premium is inconclusive. It is quite possible that these inconsistent results are attributable to differing sample companies, sample periods or factor construction procedures. This is definitely an area requiring further research as presently it is impossible to draw inferences regarding the impact of Chinese institutional issues on the value premium.

Outside China, there have been very few tests of FF93 published in quality peer-review journals. One exception is Foye and Valentinčič (2020) who study the Indonesian market. Consistent with the broad sample evidence of Cakici et al. (2017) and Foye (2018a), Foye and Valentinčič (2020) report a pronounced value premium in Indonesia. Spanning tests reveal that HML produces a strong intercept and cannot be explained away by other factors. They do not report results for a model with fewer than three factors. Therefore, it is not possible to observe the individual contribution of HML to the description of average returns.

2.1.3. Alternative Value Measures

The majority emerging markets studies follow Fama and French (1993) and form the value factor using B/M. Hanauer and Lauterbach (2019) test alternative definitions of

the value factor based on B/M, CF/P and E/P in 28 emerging markets. They initially report combined results for all countries in the sample and then provide details of regional tests. CF/P is more robust than the other two value measures in both portfolio sorts and Fama and MacBeth (1973) cross-sectional regressions.

In addition to testing FF15, Zaremba and Czapkiewicz (2017) also examine strategies based on ten alternative measures of value in Eastern Europe. All ten measures produce positive returns and eight can be considered statistically significant at the one percent level. Despite the profitability of these value strategies, they show that FF15 can explain the value-weighted returns on long-short portfolios for all ten portfolios.

2.1.4. An Institutional Interpretation of Value

The value premium's persistence across developed and emerging market samples makes it widely regarded as a particularly robust empirical regularity. However, the economic interpretation of the premium remains unclear and the risk versus mispricing debate is one of the most contentious issues in the finance literature. The main obstacle to resolving this argument is how the joint hypothesis problem impedes disentangling competing explanations. Specially, every evaluation of an asset pricing model is a joint test of both market rationality and the model used to describe risk factors. Perhaps because this ongoing debate is so vigorous, less attention has been paid to the possibility that the correct economic interpretation of the value premium may vary according to a country's or region's institutional environment.

Some researchers have examined the issue in the post-transition emerging markets of Eastern Europe because these nations have been subject to substantial institutional upheaval over a brief period. For example, Hartwell and Malinowska (2019) characterise Poland as a volatile institution environment and report that the country's informal institutions have a greater impact on firm valuation (as measured by PE) than both formal institutions and firm-specific attributes. Foye and Mramor (2016) provide a comparative study into the differing relationship between equity returns and the decomposed elements of the B/M ratio in Eastern Europe and the U.S. Intriguingly, the direction of association between the ratio's financing component and equity returns in Eastern Europe is contrary to the U.S. findings of Penman et al. (2007). They relate their results to earlier studies which established that firms' financial leverage choices in Eastern Europe are determined by factors other than maximising shareholder

wealth. Foye and Mramor (2016) only cover a narrow region over a short period. A test of the decomposed B/M ratio in a broad sample of emerging markets would yield interesting results.

2.2. Size

2.2.1. Broad Sample

International tests only offer mixed support for the size factor in both developed² and emerging markets. This has led to the size effect often being regarded as less robust than the value premium.

Fama and French (1998) show that there is a positive but insignificant (below the traditional two-standard-error bound) value premium in portfolio sorts using 16 emerging markets over the period 1987-1995. Rouwenhorst (1999) reports a positive size premium in 12 out of 20 emerging markets. These national SMB portfolios are poorly diversified and therefore cannot be considered statistically significant at any conventional level. However, an SMB factor based on all 20 markets combined does return statistically significant results.

Similarly, Cakici et al. (2013) detect some evidence of a size premium in each of their three emerging market regions. Despite not being above the two-standard-error bound, the SMB premiums could be considered economically significant. In particular, the size premiums for Asia and Latin America are substantially higher than the premiums in either U.S. or global developed markets. Cakici et al. (2013) only report asset pricing tests for CAPM and FF93 and make no mention of spanning tests. This makes it hard to draw specific inferences about SMB's usefulness for describing average returns.

Foye (2018a) tests a more recent sample based on exactly the same markets and is less supportive of the existence of a size premium. The size premium is lower in Asia and small stocks actually marginally underperform large stocks in Eastern Europe and Latin America. Factor spanning tests reveal that the SMB factor is redundant in all three regions. Like Cakici et al. (2013), Foye's (2018a) asset pricing tests do not report whether dropping the size factor harms the model's performance. Leite et al. (2018)

² E.g. Fama and French (2017) test the five-factor model in 23 developed markets and find little evidence of a size factor.

also fail to detect a statistically significant SMB premium in any of the three economic blocks or 12 countries in their sample. They only report results for models including SMB, so it is not possible to see the effect of omitting the factor on the description of average returns.

Similarly, Hanauer and Lauterbach's (2019) summary statistics provide no evidence of a size premium. Furthermore, cross-sectional regressions indicate no clear relationship between size and returns across the sample of 28 emerging markets. Zaremba and Czapkiewicz (2017) report that although the size premium is positive in Eastern Europe's five largest stock markets, it could not be considered statistically significant at any conventional level. Similarly, Foye et al. (2013) do not detect a size premium in Eastern Europe.

2.2.2. Individual Countries

Evidence from most developed and emerging markets has returned inconsistent findings regarding the existence of a size premium. However, China has been an exception as tests of the A-shares market have been consistently supportive of the size effect.

Wang and Xu's (2004) and Eun and Huang's (2007) relatively early studies both find evidence of a size premium. Wang and Xu (2004) first apply the Fama and Macbeth (1973) procedure to individual stocks before constructing an SMB factor using Fama and French (1993) style portfolios. Both approaches show that size has significant explanatory ability. Eun and Huang (2007) eschew the portfolio approach and only report results for Fama and MacBeth (1973) style individual stock regressions. Similar to Wang and Xu (2004), they find that size is related to stocks returns. However, in another early study of the Chinese market, Chen (2010) finds that the relationship between size and returns could not be considered statistically significant at any conventional level.

More recent studies support these findings. For example, in addition to testing three broad economic blocks, Leite et al. (2018) report individual size premiums for each country in their sample. Most national size premiums are positive, China is the only country that produces an SMB factor that could be considered significantly different to zero. Results from factor spanning tests are not reported so it is unclear if the SMB

factor is redundant in the sample. A lack of stocks precludes forming national 5x5 sorts so asset pricing tests can only be conducted at a regional level.

Similarly, Lin (2017) finds that SMB returns a strong intercept in the factor spanning tests. More importantly, SMB is shown to be useful for describing average returns from the LHS portfolios. Like Eun and Huang (2007), Cakici et al. (2017) do not form an SMB factor but instead use firm-level cross-sectional regressions. They find that size exhibits strong predictive power in China. Hu et al. (2019) apply both Fama and MacBeth (1973) style cross-sectional regressions and the Fama and French (1993) portfolio approach. They report a strong size premium despite no evidence of a value premium. In fact, SMB is the strongest of the FF93 factors for explaining the cross-section of Chinese returns. Most recently, Hanauer et al. (2024) show that the Chinese SMB premium is strong.

Blitz et al. (2021) are somewhat less supportive of the size premium in China. Although they detect a positive SMB premium, it cannot be considered statistically significant at any conventional level. This is the closest any major study comes to refuting the size effect in China.

Therefore, while the empirical evidence regarding the value premium in China is ambiguous, there is a clear size effect. This provides an interesting contrast to the results from other emerging markets tests of FF93 which generally find that the size premium is less robust than the value premium.

In Indonesia, Foye and Valentinčič (2020) do not report results for a model with fewer terms than FF93. Therefore, it is not possible to observe the contribution SMB makes to describing LHS returns in Indonesia. However, summary statistics show that the size premium is weak. Furthermore, SMB's intercept from the spanning tests is below the traditional two-standard-error bound – indicating that the factor may be redundant.

2.3. Momentum

2.3.1. Broad Sample

The existence of a momentum premium suggests that past returns can predict future price movements. This strikes at the very heart of the efficient market hypothesis. It is therefore surprising that many tests of factor models do not report tabulated results for the momentum term.

For example, although Fama and French (2012) find strong momentum premiums across three developed market regions, Fama and French's (2017) latter international study ignores momentum. The situation is similar in emerging markets, Leite et al. (2018) completely omit the momentum factor in their tests. Similarly, Foye (2018a) does not report results for momentum, although he mentions in a footnote that augmenting the five-factor model with a momentum term offers no improvement.

The studies that do report tabulated results for momentum in broad sample emerging markets tests are generally supportive of the premium. Rouwenhorst's (1999) early study reports a strong momentum premium across 20 emerging markets. Cakici et al. (2013) report a strong momentum effect for the combined sample of all emerging markets. In regional tests, they find substantial premiums in Asia and Latin America, but not Eastern Europe. Similarly, Hanauer and Lauterbach (2019) find a positive and significant momentum premium across 28 emerging markets. The premium is pervasive for both equal-weighted and value-weighted sorts as well as in cross-sectional regressions. Like Cakici et al. (2013), Zaremba and Czapkiewicz (2017) fail to detect a momentum premium in European markets.

Majumder (2012) tests momentum in the form of serial correlation across the BRICs markets. Returns in China and India could be considered relatively efficient prior to the financial crisis of 2007. However, returns in both markets became inefficient in the period following the crisis. Returns in Brazil and Russia were inefficient before and after 2007.

2.3.2. Individual Countries

The studies which test momentum in China offer little support for the premium. Chen (2010) finds that momentum strategies are not profitable in the Chinese markets. Cakici et al. (2017) implement both cross-sectional and portfolio regressions to test for return predictability. Their initial univariate tests show that momentum fails to qualify as a useful cross-sectional return predictor. However, when used in conjunction with additional variables in multiple regressions, momentum does actually exhibit predictive power. When testing the volatility effect, Blitz et al. (2021) describe the momentum premium as being "completely absent" in China. In untabulated tests, Wang and Xu (2004) and Lin (2017) report that the addition of a momentum term fails to improve the

performance of F93 or FF15, respectively. Similarly, Gao et al. (2021) and Hanauer et al. (2024) both find no evidence of a momentum effect in China.

Outside China, Foye and Valentinčič (2020) find the momentum premium is very weak in Indonesia and could not be considered statistically significant at any conventional level. WML's intercept from the spanning tests is close to zero, indicating that the factor is redundant.

2.3.3. Alternative Momentum Measures

Although there are numerous definitions of momentum, the WML factor is generally formed using a Jegadeesh and Titman (1993) style relative strength approach. With the exception of Zaremba and Czapkiewicz (2017), forming a momentum factor using alternative specifications of momentum has received little attention in emerging markets. Given that Fama and French (2008, p.1653) identify momentum as the “premier anomaly”, this is a surprising oversight which needs to be addressed by future researchers.

3. Five and Six Factor Models

3.1. Profitability

Novy-Marx (2013) shows that profitable firms generate substantially higher returns than unprofitable firms and profitability has similar predictive power to B/M in predicting the cross section of returns. The paper's most interesting finding is that the performance of value strategies can be substantially improved by controlling for gross profitability (gross-profit-to-assets). Novy-Marx (2013, p.15) describes this phenomenon as “profitability represent[ing] the other side of value”. The predictive power of profitability, even after controlling for value, led Fama and French (2015) to include a profitability term in their five-factor model.

3.1.1. Broad Sample

Because the inclusion of a profitability term is a relatively recent development, only a few studies have examined the factor in a broad sample of emerging markets.

The summary statistics reported by Leite et al. (2018) provide no indication of a profitability premium in any of the 12 countries or three regions covered. Due to a lack

of stocks, asset pricing tests could only be conducted at a regional level. Unfortunately, asset pricing test results for RMW are only reported when the factor is combined with either CMA or CMA and HML. Furthermore, Leite et al. (2018) do not report details of factor spanning tests. Hence, it is difficult to make inferences about the specific contribution of RMW to the model's performance. However, a two-factor model using both RMW and CMA tends to return lower GRS statistics and average intercepts than CAPM for all three regions. Adding HML to this two-factor model fails to offer a meaningful improvement, indicating that a combination of the RMW and CMA factors leaves the HML factor redundant in all three regions.

Foye (2018a) also finds that the profitability premium in Asia is small and not statistically significant. However, in a marked contrast to the findings of Leite et al. (2018), he reports large and positive region-wide profitability premiums in Eastern Europe and Latin America. Unlike Leite et al. (2018), Foye (2018a) does report factor spanning tests and shows that the RMW factor is important in Eastern Europe and Latin America, but not Asia. Furthermore, Foye (2018a) reports more in depth asset pricing tests than Leite et al. (2018). In particular, he details results for a four-factor model that augments FF93 with RMW. This four-factor model consistently returns lower GRS statistics for Eastern Europe and Latin America, but not Asia.

Despite testing the same three regions and publishing their findings concurrently, Foye (2018a) and Leite et al. (2018) report rather different results. As discussed earlier, one reason for this is Leite et al.'s (2018) very short sample period. However, Mosoeru and Kodongo (2022) also use a very short sample period (2010 to 2015) but still find evidence of a strong profitability factor. Hence this is an issue requiring the attention of future researchers.

Foye (2018a) and Leite et al. (2018) conduct asset pricing tests and emphasize GRS statistics and average intercepts. Hanauer and Lauterbach (2019) take an alternative approach and focus on the predictive power of six alternative profitability factor definitions. Broadly similar to Foye (2018a), they report large and statistically significant profitability premiums in a combined sample of 28 emerging markets.

3.1.2. Individual Countries

The scant Chinese evidence available does not offer particularly strong support for the existence of a profitability premium.

The factor summary statistics reported by Lin (2017) indicate that RMW is not economically or statistically significant. However, spanning tests reveal that RMW returns cannot be explained away by other factors, suggesting that it may still be important for describing average returns. Most importantly, asset pricing tests reveal that augmenting FF93 with RMW consistently results in lower GRS statistics and reduced average intercepts.

Leite et al. (2018) find that the RMW premium is not economically or statistically significant in China. Blitz et al. (2021) only report rudimentary tests for profitability as their primary focus is on testing the volatility effect in China. However, the results they do report provide no evidence for a profitability premium.

Perhaps one reason for the apparent poor performance of the RMW factor in China is how profitability is measured. Hanauer et al. (2024) show that using FF15's definition of profitability to form the RMW factor yields returns close to zero. However, forming a profitability factor using ROE produces returns that are economically and statistically significant.

Outside China, Foye and Valentinčič (2020) report a very strong economically and statistically significant profitability premium in Indonesia. Spanning tests show the RMW can not be explained away by some combination of other factors. However, average intercepts and GRS statistics from asset pricing tests show that FF93+RMW is not able to offer a substantial improvement over FF93.

3.1.3. Alternative Profitability Measures

The most robust definition of the profitability factor remains a source of debate. Fama and French (2015) cite Novy-Marx (2013) as the motivation for including a profitability term in the five-factor model. However, Foye (2018b) highlights an inconsistency: Novy-Marx (2013) argues gross profitability is the best measure of economic productivity but Fama and French (2015) specify the RMW factor using operating profitability. Foye's (2018b) attempt to address this by testing FF15 using alternative specifications of RMW yields inconclusive results.

Foye (2018a) and Leite et al. (2018) only test factor models using Fama and French's (2015) definition of operating profitability. However, recognising the importance of considering different factor specifications, Hanauer and Lauterbach (2019) test six

alternative measures of profitability in 28 emerging markets. They show that GP/A (as defined by Novy-Marx (2013)) is more robust than alternative profitability definitions, including OP/BE (as used by Fama and French (2015)).

Zaremba and Czapkiewicz (2017) test 11 previously documented profitability anomalies in Eastern Europe. Five of the 11 anomalies exhibit positive returns significantly different from zero. Further tests reveal that FF15 outperforms FF93 when explaining the various profitability-based anomalies.

Hanauer and Lauterbach (2019) and Zaremba and Czapkiewicz (2017) both treat testing alternative measures of profitability as an entirely empirical exercise. Hence, they offer no explanation for the differing performance of alternative profitability specifications. However, an institutional explanation has been proposed. After failing to detect a size premium, Foye et al. (2013) respecify FF93 with an alternative factor that proxies for accounting manipulation. They relate results to the region's financial reporting environment and hypothesise that the findings may hold for emerging markets in general.

4.1. Investment

Using total asset growth as a simple measure of investment, Cooper et al. (2008) show that the top decile of investment-sorted U.S. portfolios significantly underperforms the bottom decile. International studies have also supported the asset growth effect. For example, Watanabe et al. (2013) test a broad sample of 43 developed and emerging markets and find firms with higher asset growth tend to have subsequently lower stock returns. However, negative returns associated with asset growth are weaker in emerging markets.

These findings motivated Fama and French (2015) to include an investment term based on asset growth in their five-factor model. Testing U.S. markets, they show that the inclusion of profitability and investment factors leads to the value factor becoming redundant. However, the non-U.S. evidence is less supportive of the investment factor. For example, when Fama and French (2017) test an international sample of developed markets, factor spanning tests reveal that although the investment factor is significant for North America and Asia Pacific it is trivial for Europe and Japan. In another

international study of developed capital markets, Cakici (2015) reports that the investment factor lacks explanatory power in Asia Pacific and Japan.

4.1.1. Broad Sample

Echoing their results from other factors, Leite et al. (2018) find that CMA has a t-statistic less than two for all three regions. Unfortunately, they do not conduct factor spanning tests. Furthermore, only results from asset pricing tests for models using both CMA and RMW are reported. Hence, it is difficult to make specific inferences as to whether the CMA factor alone helps offer an improved description of average returns.

Foye's (2018a) factor returns also indicate that the investment premium is not economically or statistically significant in Asia or Eastern Europe. However, he finds a strong negative CMA premium in Latin America. Factor spanning tests show that the CMA factor is redundant in all three regions. Unlike Leite et al. (2018), who only report asset pricing tests for FF93 and FF15, Foye (2018a) also tests a four-factor model based on augmenting FF93 with CMA. This four-factor model fails to produce lower GRS statistics or average returns in Asia or Eastern Europe. However, it is able to offer an improved description of average returns in Latin America.

4.1.2. Individual Countries

The investment factor has been shown to be very weak in China. Contrary to the findings of Cooper et al. (2008) and Watanabe et al. (2013), Chen (2010) reports a positive relationship between asset growth and stock returns in China. However, this finding only holds at the 10 percent significance level.

Although Lin (2017) finds that FF15 offers an improved description of average returns compared to FF93, spanning tests reveal CMA can be entirely explained by a combination of other factors. Hence, the CMA factor appears to be redundant. Furthermore, when Lin (2017) tests a four-factor model – augmenting FF93 with RMW but not CMA – it returns very similar GRS statistics and average intercepts to FF15. Therefore, all Lin's (2017) findings point to CMA offering little contribution to the description of average returns in China. Hanauer et al. (2024) also find that the investment factor is weak in China. Similarly, Blitz et al.'s (2021) rudimentary tests of CMA provide no evidence of a Chinese investment premium.

Outside China, Foye and Valentinčič (2020) report a strong negative CMA premium in Indonesia. Furthermore, the large intercept returned by the spanning tests shows that the investment premium cannot be explained away by other factors. However, asset pricing tests show augmenting FF93 with CMA only offers a small improvement in terms of reduced average intercepts and GRS statistics.

Chen et al. (2008) test nine Pacific-Basin markets³ between 1981 and 2004 and find a pervasive asset growth effect in every market except Taiwan. Their results are robust to controls for the market, size and value factors. This makes for an interesting contrast with more recent broad sample studies of Asian emerging markets – e.g. Foye (2018) and Leite et al. (2018) – which find no evidence of an investment premium.

Hence, although broad sample emerging market tests offer little support for the CMA factor, there is evidence of local premiums.

4.1.3. Alternative Investment Measures

The lack of support for an investment term in broad sample emerging market tests may be the result of the factor being incorrectly defined. Hanauer and Lauterbach (2019) test alternative specifications of investment. Although all four investment premiums have t-values above two in the portfolio sort tests, Fama and MacBeth (1973) regressions reveal that asset growth (as used by Cooper (2008) and Fama and French (2015)) is insignificant when combined with net stock issues (Pontiff and Woodgate (2008)) or composite equity issuance (Daniel and Titman (2006)).

Zaremba and Czapkiewicz (2017) use a sample of five Eastern European markets to test a range of previously established investment anomalies. They show that out of the seven investment measures tested, only one could be considered statistically significant at any conventional level. The study treats this as an entirely empirical exercise and does not consider if the findings are the result of the local institutional environment.

Conducting asset pricing tests using alternative specifications of the CMA factor offers an interesting avenue for future empirical research.

³ China, Hong Kong, Indonesia, Japan, Malaysia, Singapore, South Korea, Taiwan and Thailand.

5. Forming Portfolios: National Vs International

The decision of whether to form factors on a national, regional or global basis significantly affects results. Griffin's (2002) seminal study clearly established that national factors offer a better description of average returns compared to global equivalents. Factors formed on a global or regional basis may be too broad to capture local effects. However, estimating factors on a national basis can lead to poorly diversified portfolios from which robust inferences cannot be made. This has led many international tests to use factors formed on a cross-country basis.

5.1. Forming Size Factors

Market capitalisation is reported as a monetary value, rather than as a ratio or percentage. This has a particular impact on the decision to form size portfolios on a national or cross-country basis – a stock considered large in one country may be categorised as small in another country.

Rouwenhorst (1999) forms individual SMB factors for 20 emerging markets excluding China. Initial tests show that factor returns have a strong local character as the correlations across markets are low, even within the same region. Despite this, national size premiums are only significant for a small number of markets. However, when a single SMB factor is constructed as an equally-weighted average of the national factors, it is positive and above the traditional two-standard-error bound. The most probable reason for these apparently contradictory findings is that, despite the local character of returns, the national SMB portfolios are too poorly diversified to draw meaningful inferences from.

5.2. Country-Specific Institutional Issues Regarding Earnings Quality

Given the accounting-based nature of CMA, HML and RMW, the quality of a nation's financial reporting is surely going to impact the usefulness of factor models. However, many international studies treat testing these factors as an entirely empirical exercise and do not consider the impact of local institutional issues. As well as overlooking a source of potential additional insight, this further complicates the issue of whether to form factors on a local or cross-country basis.

Foye et al. (2013) address accounting quality concerns for companies in Eastern Europe. They propose dropping the redundant size factor and replacing it with a term

they believe acts as a proxy for earnings management. They find that their respecified three-factor model offers an improved description of returns. However, these findings only cover a narrow region over a very short sample period (2005-2012). There has been no subsequent research into testing an earning management factor in other emerging markets.

Foye and Valentinčič (2020) choose to test FF15 in Indonesia as previous research has characterised the nation as having a particularly idiosyncratic financial reporting environment. They show that the performance of FF15 in Indonesia is very different to the results reported in broad sample tests of East Asian emerging markets. Furthermore, they posit that the nation's poor accounting quality explains the inability of FF15 to offer an improved description of average returns.

6. Liquidity Factor

Emerging markets are generally characterized by higher liquidity costs than their developed counterparts. This has motivated a substantial body of research⁴. Despite this, there has been a surprising lack of studies testing a liquidity factor in emerging markets.

Liu (2006) develops a liquidity factor that cannot be explained by CAPM or FF93 for stocks listed on NYSE. However, no study has replicated this model in a broad sample of emerging markets. This is a glaring omission from the literature. Bekaert et al. (2007) test a two-factor model incorporating the market portfolio and a liquidity term. They show that local market liquidity is an important driver of expected returns across a broad range of 18 emerging markets. No analysis of the relationship between liquidity and FF93 or FF15 is provided.

Testing Liu's (2006) model in an emerging market sample and reexamining Bekaert et al.'s (2007) model in conjunction with FF93 and FF15 are important avenues for future research.

⁴ E.g. Batten et al. (2014), Fong et al. (2017) and Lesmond (2005).

7. Conclusion

A substantial body of literature empirically tests factor models in emerging markets. Although this research has produced useful insight into the behaviour of asset prices in these nations, the findings are frequently contradictory. Furthermore, most studies apply the same methodology used in the U.S. with little regard to local institutional issues.

This paper lays the foundations for a resolution to these issues. Firstly, I catalogue the various contradictions in previous research and propose possible explanations. Thereby providing a guide for future researchers aiming to disentangle these conflicting findings. Secondly, I highlight the failure to consider institutional issues when testing factor models as a major limitation of most current research. In particular, I identify earnings quality and liquidity as two important but neglected concerns in emerging markets and show how these could be incorporated into factor models. This is clearly required if future empirical research is ever to determine whether emerging market investors demand the same risk premiums as developed markets. A definitive answer to this question would remove a major obstacle to disentangling the risk versus mispricing debate.

In addition to proposing resolutions to inconsistent findings and urging the incorporation of factors to capture institutional concerns, I also highlight important omissions from the extant literature. From surveying earlier research, it quickly becomes apparent that there is a clear need for future researchers to make their data publicly available. This would allow the quick resolution of issues such as determining whether contradictory findings are merely the result of differing sample periods or are actually the product of a structural break. Furthermore, there needs to be a more coordinated approach to future empirical research. For example, the formation of breakpoints is currently largely arbitrary – this makes meaningful comparisons across studies impossible. This is a particular problem when studying the size effect. Finally, although the Chinese market is well-studied, there has been a lack of quality country-specific research into other emerging markets. The exclusion of the large and relatively liquidity capital markets of India is particularly surprising. This is perhaps the most obvious omission from the literature testing factor models in emerging markets.

References

- Batten, J.A. and Vo, X.V., 2014. Liquidity and return relationships in an emerging market. *Emerging Markets Finance and Trade*, 50(1), pp.5-21.
- Bekaert, G., Erb, C.B., Harvey, C.R. and Viskanta, T.E., 1998a. The behavior of emerging market returns. In *Emerging Market Capital Flows: Proceedings of a Conference held at the Stern School of Business, New York University on May 23–24, 1996* (pp. 107-173). Boston, MA: Springer US.
- Bekaert, G., Erb, C.B., Harvey, C.R. and Viskanta, T.E., 1998b. Distributional characteristics of emerging market returns and asset allocation. *Journal of Portfolio Management*, 24(2), pp.102-116.
- Bekaert, G., Harvey, C.R. and Lundblad, C., 2007. Liquidity and expected returns: Lessons from emerging markets. *The review of financial studies*, 20(6), pp.1783-1831.
- Blitz, D., Hanauer, M.X. and van Vliet, P., 2021. The Volatility Effect in China. *Journal of Asset Management*, pp.1-12.
- Cakici, N., Chan, K. and Topyan, K., 2017. Cross-sectional stock return predictability in China. *The European Journal of Finance*, 23(7-9), pp.581-605.
- Cakici, N., Fabozzi, F.J. and Tan, S., 2013. Size, value, and momentum in emerging market stock returns. *Emerging Markets Review*, 16, pp.46-65.
- Cakici, N., 2015. The five-factor Fama-French model: International evidence. Available at SSRN 2601662.
- Chen, S., Yao, T., Yu, T. and Zhang, T., 2008. Asset growth and stock returns: evidence from the Pacific-Basin markets. *Pacific-Basin Finance Journal*, 19(1), pp.115-139.
- Chen, X., Kim, K.A., Yao, T. and Yu, T., 2010. On the predictability of Chinese stock returns. *Pacific-Basin Finance Journal*, 18(4), pp.403-425.
- Cooper, M.J., Gulen, H. and Schill, M.J., 2008. Asset growth and the cross-section of stock returns. *Journal of Finance*, 63(4), pp.1609-1651.
- Daniel, K. and Titman, S., 2006. Market reactions to tangible and intangible information. *The Journal of Finance*, 61(4), pp.1605-1643.

- Eun, C.S. and Huang, W., 2007. Asset pricing in China's domestic stock markets: Is there a logic?. *Pacific-Basin Finance Journal*, 15(5), pp.452-480.
- Fama, E.F. and French, K.R., 1992. The cross-section of expected stock returns. *The Journal of Finance*, 47(2), pp.427-465.
- Fama, E.F. and French, K.R., 1993. Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, 123(3), pp. 441-463.
- Fama, E.F. and French, K.R., 1998. Value versus growth: The international evidence. *The Journal of Finance*, 53(6), pp.1975-1999.
- Fama, E.F. and French, K.R., 2008. Dissecting anomalies. *The Journal of Finance*, 63(4), pp.1653-1678.
- Fama, E.F. and French, K.R., 2012. Size, value, and momentum in international stock returns. *Journal of financial economics*, 105(3), pp.457-472.
- Fama, E.F. and French, K.R., 2015. A five-factor asset pricing model. *Journal of financial economics*, 116(1), pp.1-22.
- Fama, E.F. and French, K.R., 2017. International tests of a five-factor asset pricing model. *Journal of Financial Economics*, 123(3), pp.441-463.
- Fama, E.F. and MacBeth, J., 1973. Risk, return, and equilibrium: Empirical tests. *Journal of Political Economy*, 81(3), pp.607-636.
- Fong, K.Y., Holden, C.W. and Trzcinka, C.A., 2017. What are the best liquidity proxies for global research?. *Review of Finance*, 21(4), pp.1355-1401.
- Foye, J., Mramor, D. and Pahor, M., 2013. A Respecified Fama French Three-Factor Model for the New European Union Member States. *Journal of International Financial Management & Accounting*, 24(1), pp.3-25.
- Foye, J. and Mramor, D., 2016. A new perspective on the international evidence concerning the book-price effect. *Emerging Markets Finance and Trade*, 52(10), pp.2348-2363.
- Foye, J., 2018a. A comprehensive test of the Fama-French five-factor model in emerging markets. *Emerging Markets Review*, 37, pp.199-222.
- Foye, J., 2018b. Testing alternative versions of the Fama–French five-factor model in the UK. *Risk Management*, 20(2), pp.167-183.

- Foye, J. and Valentinčič, A., 2020. Testing factor models in Indonesia. *Emerging Markets Review*, 42, pp.100628.
- Gao, Y., Han, X., Li, Y. and Xiong, X., 2021. Investor heterogeneity and momentum-based trading strategies in China. *International Review of Financial Analysis*, 74, pp.101654.
- Griffin, J.M., 2002. Are the Fama and French factors global or country specific?. *The Review of Financial Studies*, 15(3), pp.783-803.
- Hanauer, M.X., Jansen, M., Swinkels, L. and Zhou, W., 2024. Factor models for Chinese A-shares. *International Review of Financial Analysis*, 91, pp.102975.
- Hanauer, M.X. and Lauterbach, J.G., 2019. The cross-section of emerging market stock returns. *Emerging Markets Review*, 38, pp.265-286.
- Hartwell, C.A. and Malinowska, A.P., 2019. Informal institutions and firm valuation. *Emerging Markets Review*, 40, pp.100603.
- Harvey, C.R., 1995. Predictable risk and returns in emerging markets. *The Review of Financial Studies*, 8(3), pp.773-816.
- Hu, G.X., Chen, C., Shao, Y. and Wang, J., 2019. Fama–French in China: size and value factors in Chinese stock returns. *International Review of Finance*, 19(1), pp.3-44.
- Jegadeesh, N. and Titman, S., 1993. Returns to buying winners and selling losers: Implications for stock market efficiency. *The Journal of Finance*, 48(1), pp.65-91.
- Leite, A.L., Klotzle, M.C., Pinto, A.C.F. and da Silva, A.F., 2018. Size, value, profitability, and investment: Evidence from emerging markets. *Emerging Markets Review*, 36, pp.45-59.
- Lesmond, D.A., 2005. Liquidity of emerging markets. *Journal of Financial Economics*, 77(2), pp.411-452.
- Lin, Q., 2017. Noisy prices and the Fama–French five-factor asset pricing model in China. *Emerging Markets Review*, 31, pp.141-163.
- Liu, W., 2006. A liquidity-augmented capital asset pricing model. *Journal of financial Economics*, 82(3), pp.631-671.

- Liu, J., Stambaugh, R.F. and Yuan, Y., 2019. Size and value in China. *Journal of Financial Economics*, 134(1), pp.48-69.
- Lischewski, J. and Voronkova, S., 2012. Size, value and liquidity. Do they really matter on an emerging stock market?. *Emerging Markets Review*, 13(1), pp.8-25.
- Lucey, B.M., Vigne, S.A., Ballester, L., Barbopoulos, L., Brzeszczynski, J., Carchano, O., Dimic, N., Fernandez, V., Gogolin, F., González-Urteaga, A., Goodell, J.W., Helbing, P., Ichev, R., Kearney, F., Laing, E., Larkina, C.J., Lindblad, A., Lončarski, I., Ly, K.C., Marinč, M., McGee, R.J., McGroarty, F., Neville, C., O'Hagan-Luff, M., Piljak, V., Sevcica, A., Sheng, X., Stafylas, D., Urquhart, A., Versteeg, R., Vu, A.N., Wolfe, S., Yarovayar, L. and Zaghini, A., 2018. Future directions in international financial integration research-A crowdsourced perspective. *International Review of Financial Analysis*, 55, pp.35-49.
- McLean, R.D. and Pontiff, J., 2016. Does academic research destroy stock return predictability?. *The Journal of Finance*, 71(1), pp.5-32.
- Majumder, D., 2012. When the market becomes inefficient: Comparing BRIC markets with markets in the USA. *International Review of Financial Analysis*, 24, pp.84-92.
- Mosoeu, S. and Kodongo, O., 2022. The Fama-French five-factor model and emerging market equity returns. *The Quarterly Review of Economics and Finance*, 85, pp.55-76.
- Novy-Marx, R., 2013. The other side of value: The gross profitability premium. *Journal of Financial Economics*, 108(1), pp.1-28.
- Penman, S.H., Richardson, S.A. and Tuna, I., 2007. The book-to-price effect in stock returns: accounting for leverage. *Journal of accounting research*, 45(2), pp.427-467.
- Pontiff, J. and Woodgate, A., 2008. Share issuance and cross-sectional returns. *The Journal of Finance*, 63(2), pp.921-945.
- Rouwenhorst, K.G., 1999. Local return factors and turnover in emerging stock markets. *The journal of finance*, 54(4), pp.1439-1464.
- Schwert, G.W., 2003. Anomalies and market efficiency. *Handbook of the Economics of Finance*, 1, pp.939-974.

Wang, F. and Xu, Y., 2004. What determines Chinese stock returns? *Financial Analysts Journal*, 60(6), pp.65-77.

Watanabe, A., Xu, Y., Yao, T. and Yu, T., 2013. The asset growth effect: Insights from international equity markets. *Journal of Financial Economics*, 108(2), pp.529-563.

Zaremba, A. and Czapkiewicz, A., 2017. Digesting anomalies in emerging European markets: A comparison of factor pricing models. *Emerging Markets Review*, 31, pp.1-15.