Risk Factors For the Indian Equity Market: Statistics, Visualization, and an Interactive Tool

Aman Foujdar, Sandeep Juneja, Atish Kumar, and Nagpurnanand Prabhala.*

June 16, 2026

Abstract

We provide a carefully constructed monthly time series of six Fama–French factors for the Indian equity market. We motivate and discuss the protocols, filters, and weights used to construct the series and in particular, what we draw from well-established practices U.S. research and what we adapt to the local market structure in India. In relatively conservative value-weighted portfolios between 2003 and 2025, long-short factor premiums range from 400 basis points to 1300 basis points per year with Sharpe ratios of up to 0.40 to 0.60. Performance metrics depend significantly on the universe, filters, and the precise protocols used for factor construction, underlining the need for caution in constructing and interpreting factor premiums. We also discuss a set of interactive tools and individual stock labels for academics and practitioners to assess and visualize the dynamics of factor premiums, cumulative returns, and portfolio performance statistics for a flexible suite of long-short strategies over multiple windows of time.

The factor library is publicly available at quantfin-scdlds.com

JEL Codes: G12, G11, G14, G17
Keywords: Factor investing, asset pricing, Fama-French Factors, India.

*Aman Foujdar and Sandeep Juneja are at Ashoka University and can be reached at aman.foujdar_ug2023@ashoka.edu.in and sandeep.juneja@ashoka.edu.in, respectively. Atish Kumar is at the Indraprastha Institute of Information Technology Delhi and can be reached at atishk@iiitd.ac.in. Juneja is Professor of Computer Science and Director, Safexpress Centre for Data, Learning, and Decision Sciences at Ashoka University. Nagpurnanand Prabhala is Francis J. Carey Professor of Busines and Professor of Fnance at The Johns Hopkins University Carey Business School. Email: prabhala@jhu.edu.
We thank CMIE for providing the financial data and several colleagues and CMIE staff members for helpful discussions.

1 Introduction

The Fama–French factors have become foundational in empirical asset pricing. The factors are a parsimonious representation of sources of systematic risk. The early work (Fama and French, 1993) supplemented the classic CAPM market factor with size and value. Subsequent work has incorporated four other factors. Carhart (1997) introduced a momentum factor, Novy-Marx (2013) considers the profitability, while Titman, Wei and Xie (2004) and Anderson and Garcia-Feijoo (2006) consider an investment factor. See also Fama and French (2016). The six risk factors have become central tools in academic research and practice, although other factors – and alternative measures of the factors - have emerged and continue to emerge resulting in what is now called the factor zoo.[1]

Factor models serve a variety of purposes. Perhaps the most important of them is that they are return benchmarks that help test whether any proposed trading strategy generates genuine alpha. Factors help decompose and attribute the style and performance of mutual fund and pension fund managers or advisors providing portfolio management services.[2] Factor premiums also serve as targets for mutual funds and ETFs that drive the large factor investing industry. In corporate finance, they serve as benchmarks to test post-announcement portfolio and calendar time returns. Factor models also improve cost of capital estimates used by managers to evaluate projects.

Understanding the factor risk premiums is also of economic interest. For example, finance research attempts to understand whether momentum in stock returns reflects factor momentum (Ehsani and Linnainmaa, 2022). The existence of a value premium and its time variation has been the focus of a large body of research. For example, Zhang (2005) attributes it to risks due to frictions in right-sizing capital stocks and countercyclical risk premiums. Chen, Petkova and Zhang (2008) and Fama and French

[1]For example, Pastor and Stambaugh (2003) consider a liquidity factor, for which Liu (2006) and Amihud (2002) propose alternative metrics. See, e.g., Harvey, Liu and Zhu (2016) and Feng, Giglio and Xiu (2020) on the factor zoo and the suspicions that they are the products of data dredging and snooping.

[2]A question that often comes up is whether factors represent dimensions of risk or not. This issue is not critical in performance attribution in practice. Using factors to calibrate returns can be interpreted as the extent to which any detected alpha is spanned by already-known factors.

(2006) examine whether the value premium is robust over long time period. Campbell, Giglio and Polk (2023) show that value booms and busts are related to news about aggregate cash flows, discount rates, and volatility.

We develop estimates of the six core Fama-French factors for the Indian equity market. India is a large developing economy with 2025 GDP of about $4 trillion, but has an extremely old and established stock market with distinct signatures in financial market development. participation, regulations, and trading. The Indian stock market was established in 1875 as The Native Share & Stock Brokers' Association, now called the BSE. Operations continued past Indian independence under strict control through the Controller of Capital Issues. A major shift occurred in 1992-1994 on the heels of stock market scandals with the establishment of the National Stock Exchange (NSE). A system of electronic screen-based trading was introduced at the time, a transition to which was largely completed around 2000. Our estimates of factor premiums start around this time.

We are not the first to develop data on factor premiums for the Indian market. A prominent four factor model that includes the three (Fama and French, 1993) factors and momentum was developed by Agarwalla, Jacob and Varma (2014) (AJV, henceforth), originally for stocks listed on the BSE and subsequently extended to NSE stocks. More recently, another prominent and ambitious effort by Jensen, Kelly and Pedersen (2023) (JKP, henceforth), constructs 153 factors across 13 "themes" for 93 countries including India. Factor construction requires the researcher to make multiple empirical choices. Both studies take different approaches, in part reflecting varying preferences on how much to customize to local markets – and what criteria to use in doing so – versus standardized choices based on U.S. cutpoints.[3]

Our approach is to lay out our own choices transparently (but as in JKP) but providing tools to users to generate, visualize, and possibly customize and assess alternative portfolios. Our preferred choices for factor estimation tilt towards more traded stocks that have more reliable return data, following the AJV approach for screens based on

[3]Here, we note the distinction between stocks used to construct factor premiums and ones used in portfolio formation and trading. Realizable returns in trading strategies need significant adjustments for transaction costs due to illiquidity (Novy-Marx and Velikov, 2015).

local market structure. We provide data for all the six Fama-French factors and provide tools to analyze the data at multiple levels, including the top 300 top 500 stocks by market capitalization and with value- and equal- weights. Besides research using factors, work under progress on the data delivery side includes auxiliary factors and the ability to implement alternative weighting schemes. We are also working on more explicit estimates of transaction costs but provide turnover computations as an interim output for users interested in the issue.

2 Data and Filters

Firm-level balance sheet, stock returns, and share active trading dates come from the CMIE Prowess DX database We have 7757 unique firms in the data with return histories from 2000 to 2025. label fiscal year t as the Given the need for historical returns for the momentum factor and based on our internal assessment of the reliability of data, we provide data starting in 2002

2.1 Price Data

The Bombay Stock Exchange (BSE) and the National Stock Exchange (NSE) are India's leading stock exchanges and it is not uncommon for firms to cross-list on both. As of 2025, the market cap of all stocks in the BSE is ₹456 trillion compared to ₹461 trillion in the NSE versus a total market capitalization of ₹464 trillion. The BSE has about 2,000 more lists but they are all small microcaps. We have price data from both exchanges. If a stock is listed on the NSE then we use NSE prices and if not, we use BSE prices. .

2.2 Accounting Data

A vast majority of Indian companies have their financial end year in March. Listed firms are typically required to release audited financials within 60 days of year end. Following Fama and French, we use a conservative 6-month gap before using the accounting data

for a year. Thus, accounting data for the year t that ends on March 31 is used only starting from October 1 of the year. For instance, the book value of a stock for the fiscal year ending on March 31, 2006 is applied to all months starting from October 1, 2006 to September 30, 2007.[4]

2.3 Filters for Indian Equity Markets

Negative Book Equity: Book equity is defined as the total shareholder equity reported in a firm's balance sheet. We identify firms with negative book value and exclude them from all factors that use book value.[5] **Figure 1** shows the yearly count of firms with negative book equity in the raw data. An important point to note here is that, we drop firms with negative book equity only for those factors which don't require book equity variable for construction.

Penny Stock: We define penny stocks as the ones whose median daily closing share price in the prior year is less than or equal to ₹10. These shares are illiquid, have noisy returns, are subject to pump-and-dump schemes, and attract less institutional interest. We briefly comment on our choice.

In the U.S., penny stock definitions come from Section 3(a)51 of the 1934 Securities and Exchange Act as amended by the 1990 Penny Stock Reform Act. We are not aware of a similar definition in India, where listing regulations specify minimum public float and market capitalization. Studies of the Indian market vary in their definitions of a penny stock. Pandey and Sehgal (2016) use a cutoff of ₹10 (roughly 10 to 20 cents) while Agarwalla, Jacob and Varma (2014) use threshold of ₹1 (1-2 cents). A cross-country study of momentum by Jegadeesh and Titman (2023) uses a much greater cutoff of $1, which corresponds to between ₹50 and ₹90 in our sample period.

[4] A four-month gap is used in AJV. Their choice results in differences in our choices for two months in a year (August and September). The AJV alternative is reasonable but With no special local market reasons (e.g., earlier deadlines for filings in India relative to the US), we choose to stay with the more conservative Fama-French choice.

[5] Return data for these firms are reported separately.

Figures 2 and 3 show that the share of firms whose share price is less than ₹10 remained relatively stable from around 2005 to 2020 at 25-30% of the total firms. The fraction of of firms with share price less than ₹30 dropped uniformly after 2020. Very few firms have prices below ₹1. **Table 1** shows the ₹10 corresponds to about the 25th percentile of the share price distribution until about 2020, after which a stock-market boom appears to have lifted the 25th percentile to around ₹27. In balance, the ₹10 cutoff appeared to be a reasonably stable and meaningful cutoff for the purposes of factor construction.

Microcap: The filter drop stocks with low market capitalization, which are often hard to trade in reasonable quantities. We follow Agarwalla, Jacob and Varma (2014) and define microcaps as firms whose market capitalization as of September 30 of year t is less than 10 percent of the median market capitalization of all stocks. We apply this exclusion filter for all stocks from October 1 of year t to September 30 of year $t+1$. **Figure 4** shows the cutoff from 2021 to 2025 has been between ₹500 million and ₹700 million.

Illiquidity: We term a stock as being illiquid between October 1 of year t and September 30 of year $t+1$ if it has not traded at least once in every full trading week in the year before September 30 of yeat t. We modify the definition in Agarwalla, Jacob and Varma (2014), who require trading on any 50 days during the prior year. Our more stringent criterion of being traded in every week filters out cases with prolonged illiquidity.

Table 2 shows the holding universe by year, where for convenience, "year" is the one-year period beginning on October 1 of calendar year t to September 30 of calendar year $t+1$. **Table 3** reports the number of additional firms excluded due to negative book equity in each portfolio year.

There is a significant overlap between negative book equity firms and the remaining filters. After applying the liquidity, microcap and penny stock filter, the number firms with negative book equity has remained under 150 and with minor changes since 2016. Before 2016, the number of negative book equity firms were lower than 100 except for

one year.

3 Factor Construction

The factors constructed in this paper follow a double-sort methodology used in constricting the Fama-French factors available in the French data library. Each stock is assigned a label for each factor. Stocks are sorted, typically into above or below median categories or terciles for each label in independent sorts for each category. Intersections of sorted portfolios form factor portfolios that are used to construct long-short portfolios. The process is described below with details.

Size: We use the September end-of-month market capitalization in year y to assign the size label for the subsequent portfolio year, which is October of year y to September of year $y+1$. Firms above the 90th percentile of market capitalization are classified as Big (B) and the rest are Small (S).

As in the U.S., the Indian market has a long tail of small firms **Figure 5** shows that close to 40% of the firms have a market capitalization of lower than ₹100 crore. In the U.S. market, the remedy has been to use NYSE-AMEX market capitalizations to define size deciles that are then applied to all firms including those listed on NASDAQ. An analogous procedure is less effective given the high overlap in cross-lists at both the BSE and the NSE; moreover, the older BSE (the analog of NYSEAMEX in the U.S.) has *more* firms and more small firms. We define large firms as those belonging to the top 10th percentile of market capitalization of all firms. These firms account for 90% of the overall market capitalization in 2025. **Figure 6** shows the decile level market capitalization in the India equity markets, the mean market capitalization overlaps with the 90th percentile of market capitalization. In the U.S., the top 10% of firms by market capitalization account for about 78% of the total capitalization.

Book-to-Market (B/M): Book-to-market ratio is a signal of a value stock. High book-to-market stocks are value stocks while low book-to-market indicates growth stocks whose valuation comes less from assets-in-place than from growth opportunities. We compute it as the ratio of book equity in the March of calendar year y and market value as of September of year y, and use the label for portfolio formation purposes from October of year y through September of year $y + 1$.

Firms with B/M ratio above the 70th percentile are classified as Value (V); firms below the 30th percentile are classified as Growth (G), the remaining are Neutral (N). Firms with negative or missing book equity are excluded from value breakpoint construction and holding universe.

Operating Profitability: Operating profitability is a measure of profit divided by its lagged book equity. The numerator is sales minus cost of goods sold (COGS), selling and distribution expenses, and interest expense. The denominator is the book value of equity. Both are estimated as of year y and the label applied for all months from October of year y to September of year $y+1$. We exclude firms with negative book values. Firms above the 70th percentile of the profitability ratio are classified as Robust (R), while firms below the 30th percentile are classified as Weak (W) and the rest are Neutral (N).

Investment The investment factor captures the asset growth of a firm. Firms ranked high on this factor invest aggressively and experience high asset growth firms (A). The low investors are called conservative (C). are classified as Aggressive (A) which considered as a negative signal and while low growth firms are classified as Conservative (C). As in the original Fama and French factors, we compute investment as the change in assets between the end of March of calendar year y and $y - 1$ and divide it by the assets as of year $y - 1$, and apply the label for all months from October of year y to September of year $y + 1$.

Momentum: We define momentum as the cumulative gross return over an 11-month formation window excluding the most recent month. Thus, for example, momentum for the month of September 2013 would include cumulative 11-month returns ending on July 31, 2013. The momentum signal is the cumulative gross return over months $t-12$ through $t-2$, the geometric product of one plus the return over the time period.

A firm must have valid returns for 11 months for the signal to be computed. Firms occasionally have missing returns for some months, the by-product of gaps in the daily return sequences (that are use to iproduce monthly returns). Breakpoints are computed monthly at the 30th and 70th percentiles of the cross-sectional distribution of MOM_t. Firms above the 70th percentile are assigned to the label Winner (W) and firms below the 30th percentile are classified as Loser (L). The remaining firms are assigned the label Neutral(N). For intersections of momentum and market capitalization or size, we use the equity market value as of end of the momentum period (rather than the lagged capitalization as of March 31).

Market Factor : We define the market factor as the return on a broad market portfolio minus the risk-free rate (converted into a monthly rate by dividing by 12). We use the monthly return on the NIFTY 500 index as a proxy for the market portfolio, given its comprehensive coverage of large, mid and small cap stocks. The risk-free rate is the 91-day Treasury bill rate published on the Reserve Bank of India website.

4 Factor Cut Points and Returns

All factor portfolios except momentum are constructed based on stock decile labels as of September 30 of each year and labels are held constant until the next September. Momentum, is of course, relabeled every month. We provide labels at three levels: for the entire tradable universe, for the top 300 or 500 stocks by market cap as of September 30 of each year.

Firms are classified as big (B) or small (S) based on market size. Other factors are

typically divided into terciles and thus stock factor labels are terciles. Intersecting the two sets of labels for each stock (e.g., **L**ow, **N**eutral, **H**igh) with size labels, viz., **S**mall or **L**arge gives us six value-weighted portfolios for each factor: $\{S, B\} \times \{F_L, F_N, F_H\}$

Size Factor (SMB): The Size factor (Small Minus Big) is computed by taking the average returns of three small portfolios with three big portfolios computed within the value universe:

$$\text{SMB} = \frac{1}{3}\big[(SV - BV) + (SN - BN) + (SG - BG)\big] \tag{1}$$

where SV, SN, SG denote the Small-Value, Small-Neutral, and Small-Growth portfolios respectively, and BV, BN, BG denote their Big counterparts. With independent sorts, it is possible that not all intersections of characteristics are sufficiently populated. In India, a problem is (perhaps curiously) big value firms, a point noted by AJV, although these gaps well populated for the past decade. We do not report returs for factor portfolios with fewer than five stocks. This restriction amounts to a modification in the factor return premium calculation. For example, if BV has less than 5 firms, we drop the BV and SV portfolios. SMB is computed using only the neutral and growth pairs.

$$\text{SMB} = \frac{1}{2}\big[(SN - BN) + (SG - BG)\big] \quad \text{if } n_{BV} < 5 \tag{2}$$

Value Factor (HML): The value factor (High-Minus-Low book-to-market) is the average return of the two value portfolios minus the average return of the two growth portfolios:

$$\text{HML} = \frac{1}{2}\big[(SV - SG) + (BV - BG)\big] \tag{3}$$

The long portfolios are Small-Value and Big-Value and the short ones are Small-Growth and Big-Growth. When $n_{BV} < 5$, HML is essentially the difference between

small value and small growth:

$$\text{HML} = SV - SG \quad \text{if } n_{BV} < 5 \tag{4}$$

Momentum Factor (WML) The momentum factor (Winner-Minus-Loser) is the average return of the two winner portfolios minus the average return of the two loser portfolios:

$$\text{WML} = \frac{1}{2}\big[(SW - SL) + (BW - BL)\big] \tag{5}$$

where SW and BW denote the Small-Winner and Big-Winner portfolios and SL and BL denote the Small-Loser and Big-Loser portfolios. Momentum and size breakpoints are computed monthly for the momentum factor.

Operating Profitability Factor (RMW) The operating profitability (Robust-Minus-Weak) factor along with Investment are rely only on the balance sheet variables. It is the difference between average return of the two robust portfolios and the average return of the two weak portfolios:

$$\text{RMW} = \frac{1}{2}\big[(SR - SW) + (BR - BW)\big] \tag{6}$$

where SR and BR denote the Small-Robust and Big-Robust portfolios and SW and BW denote the Small-Weak and Big-Weak portfolios. Operating profitability labels are assigned annually using March financial year-end data with a six-month reporting lag.

Investment Factor (CMA): The investment factor (Conservative-Minus-Aggressive) is the average return of the two conservative portfolios minus the average return of the

two aggressive portfolios:

$$\text{CMA} = \frac{1}{2}\big[(SC - SA) + (BC - BA)\big] \tag{7}$$

where SC and BC denote the Small-Conservative and Big-Conservative portfolios and SA and BA denote the Small-Aggressive and Big-Aggressive portfolios. Conservative firms are those with low asset growth (below the 30th percentile) and aggressive firms are those with high asset growth (above the 70th percentile).

Table 4 & **Table 5** summarizes the long and short factor construction and respective factor breakpoints.

5 Comparison with existing Factor libraries

As a part of constructing the factor library, we have extensively compared our methodology against the existing resources. As mentioned above, Agarwalla, Jacob and Varma (2014) and Jensen, Kelly and Pedersen (2023) publish factors for the Indian equity markets. Agarwalla, Jacob and Varma (2014) focus solely on the Indian equity markets and their methodology aims to capture the market microstructure but the factor coverage remains limited.

5.1 Comparison with Agarwalla, Jacob and Varma (2014)

Our universe construction departs from Agarwalla, Jacob and Varma (2014) on three dimensions. First, on exchange selection, since it is common for stocks to be listed on multiple exchanges, it is important to have a rule for selecting the exchange from where the prices are sourced. We follow a simple rule, choose NSE if the stock is listed both the exchanges and BSE otherwise. This essentially means that the variables are selected for PY $t+1$ by applying this rule in PY t. This rule is aimed to reflect the prominence of NSE as an exchange. On the other hand, AJV selects the exchange with higher aggregate

liquidity, though the exact liquidity variable is not specified in their website, we replicated their results with traded value, traded quantity and found that results are largely stable across the two variables. Second, for the liquidity filter, our methodology requires a firm to trade at least once in every full calendar week of the preceding portfolio year. AJV applies a 50-day trading threshold in the preceding portfolio year with no additional restrictions. This may lead to inclusion of firms which have lumpy trading activity, making it less ideal for investment.

Third, our penny stock threshold is ₹10 based on median price over the preceding portfolio year as opposed to ₹1 in AJV's case. Fourth, the microcap filter which is applied on firms below 10% of the median market capitalization on the portfolio formation date and lastly, the book equity treatment are identical across both methodologies which excludes firms with negative or missing book equity.

5.1.1 Firm Counts and Coverage

Figure 7 shows the number of firms in our universe, the Agarwalla, Jacob and Varma (2014) universe, and the full CMIE Prowess base (firms with a valid identifier) for each portfolio year from 2001 to 2025. Our filters are substantially stricter in the initial years. In 2001, we retained 858 firms against Agarwalla, Jacob and Varma (2014) 2,027 where the total base was 3,539. The gap narrows post-2010 as the market grew but our there is still a considerable difference between the two universes.

5.1.2 Impact of Exclusions

An obvious question to ask is what is the impact of impact of stricter filters as compared to AJV. This can be understood in a variety of ways such as what is the difference in market capitalization and trading volume. **Figure 8** shows the difference in the total September end market capitalization of the Agarwalla, Jacob and Varma (2014) universe, our universe, and the excluded firms for each portfolio year. Excluded firms account for 0.6–8.0% of total market capitalization, with the peak occurring in 2007. In the post-

2013 period the excluded percentage stays below 2.1%. A similar story emerges for trading volume as well, excluded firms account for under 10% of annual trading volume at their peak and typically 1–3% post-2013 **(Figure 9)**. These figures support the claim that excluded firms are economically insignificant and that their prices are likely to be unreliable for return inference.

5.1.3 Characteristics of Excluded Firms

Figure 10 reports median market capitalisation for our universe, Agarwalla, Jacob and Varma (2014), and excluded firms by portfolio year. The excluded firms' median market cap is approximately 10% of our universe throughout the sample period precisely the threshold set by the microcap filter confirming that the filter operates as intended. The median market cap of excluded firms remains below ₹300 million for most of the sample, rising only modestly in recent years as overall market levels increased.

5.2 Comparison against Jensen, Kelly and Pedersen (2023)

There is a structural difference in factor construction methodology when we compare our methodology against Jensen, Kelly and Pedersen (2023). Their Global Factor Data are constructed using CRSP, Compustat, and Compustat Global identifiers and characteristics. For each factor construction, they form monthly portfolios by dividing stocks into terciles using breakpoints computed from only non-micro stocks for each country. They define non-micro cap stocks as firms with market equity above the NYSE 20th percentile. Microcap stocks are then assigned factor label using factor breakpoints computed from the non-micro stocks.

Portfolio returns are computed using three methods, equal-weighted, value-weighted and capped value weights. For the capped value-weighted method they cap the highest market equity to be at the NYSE 80th percentile, and factor returns are defined as high-minus-low long-short portfolios. On the other hand, our factor library sources data from CMIE Prowess with Indian-market-specific sample filters. In the absence of a common

identifier between CMIE Prowess and the WRDS database, we matched firms using a multi-stage fuzzy name-matching procedure, the first level included cleaned full names after removing generic suffixes such as *Ltd.* and *Industries*. In the second level we used two and three token name prefixes, and finally on spaceless normalised names to handle spelling variants and after this procedure. After doing this, there were 113 unique firms remain unmatched.

5.2.1 Impact of exclusions

Figure 13 compares firm counts between our universe and Jensen, Kelly and Pedersen (2023) by portfolio year. Post-2012, the JKP firm universe contains close to twice as many firms as ours. We have analysed underlying reasons for such a large deviation in the universe construction and observed that the exclusions can be explained by the filters. Since 2013, over 95% of the firms which were a part of the JKP universe-only fail at least one of the filters. This shows that the differences in the tradable universe can be attributed to differences in methodology. Their factor library doesn't apply filters to account of the nuances of the Indian markets. The unexplained residual of 50-66 firms per year is stable and is due to the 113 firms that could not be matched across CMIE and WRDS identifiers.

The characteristics of JKP-only firms further validate our exclusions. As shown in Figure 13, the median price of firms belonging only to Jensen, Kelly and Pedersen (2023) is ₹12 versus ₹87 for our universe, with the distribution of their firms centred precisely at the ₹10 penny threshold. Figure 14 shows that JKP-only firms have a median of 178 trading days per year versus 248 for our universe which is significantly lower. Figure 15 reports median market capitalisation for our universe and Jensen, Kelly and Pedersen (2023)-only firms, the latter's median stays substantially below our universe.

Taken together, the evidence supports a straightforward characterisation of our exclusions: they remove firms that are microcap in size, illiquid in trading, and priced near or below the penny threshold — precisely the firms for which price-based return measure-

ment is most unreliable. These firms represent under 2% of total market capitalisation and under 3% of total trading volume in the post-2013 period.

6 Returns

6.1 Long-Only Factor Performance (June 2016 – May 2026).

Table 6 reports the performance of long-only portfolios for the Fama-French Factors and Momentum over the past ten years. Momentum delivers the largest alpha, with 22.45% annualized return along with the highest Sharpe ratio of 0.79 among all factors. **Figure 16** shows that that value and momentum have risen sharply post-2020. During this period, the long-only leg's drawdown is comparable to the NIFTY 500 benchmark which is close to -30%. Both size and value also show strong performance with an annualized return of 20% and 22.9% respectively. However, both these factors carry much higher volatility and drawdowns as compared to the NIFTY 500 benchmark. Owing to the higher volatility, the share ratio of value and size are lower than momentum and stands at 0.6 and 0.55 respectively. The profitability factor is the weakest performer of all both in terms of returns and sharpe ratio. On a long-only basis, Operating Profitability achieves an 15.54% annualized return with a sharpe ratio of Sharpe 0.46 but with the second deepest drawdown of -55.34%

6.2 Long-Short Factor Performance (June 2016 – May 2026).

Table 7 reports the performance of long-short portfolios for the Fama-French Factors and Momentum over the past ten years. Momentum remains the top performer in all metrics except volatility. In the long-short case, it generated a 15.70% annualized mean return with a sharpe ratio of 0.65. **Figure 17** shows a sharp reversal in the performance of value factor post-Covid 19 pandemic, it generated an annual mean excess returns of 2.72% which translates into a modest Sharpe ratio of 0.16, along with steepest drawdown

among all factors. The remaining factors have generated negaive annual excess returns. However the investment factors had the lowest drawdown among all the factors.

7 Conclusion

We have constructed a comprehensive factor library for Indian equity markets spanning close to 25 years. Apart from covering the popular Fama-French and Momentum factors, we have constructed additional factors which could be of interest for practitioners and academicians. We have conducted extensive analysis to compare against the existing factor libraries. Our factor library draws inspiration from the existing resources and attempts to make a methodological improvement in the Indian context. Additionally, we aim to leverage these factors to answer other interesting questions.

References

Agarwalla, Sobhesh Kumar, Joshy Jacob, and Jayanth Rama Varma. 2014. "Four factor model in Indian equities market." *Indian Institute of Management, Ahmedabad Working Paper*, , (2013-09): 05.

Amihud, Yakov. 2002. "Illiquidity and Stock Returns: Cross-Section and Time-Series Effects." *Journal of Financial Markets*, 5(1): 31–56.

Anderson, Christopher, and Luis Garcia-Feijoo. 2006. "Empirical Evidence on Capital Investment, Growth Options, and Security Returns." *Journal of Finance*, 61(1): 171–194.

Campbell, John, Stefano Giglio, and Christopher Polk. 2023. "What Drives Booms and Busts in Value." *NBER Working Paper 31589*.

Carhart, Mark M. 1997. "On persistence in mutual fund performance." *The Journal of finance*, 52(1): 57–82.

Chen, Long, Ralitsa Petkova, and Lu Zhang. 2008. "The expected value premium." *Journal of Financial Economics*, 87(2): 269–280.

Ehsani, Sina, and Juhani Linnainmaa. 2022. "Factor Momentum and the Momentum Factor." *Journal of Finance*, 77(3): 1877–1919.

Fama, Eugene, and Kenneth French. 2006. "The Value Premium and the CAPM." *Journal of Finance*, 61(5): 2163–2185.

Fama, Eugene F, and Kenneth R French. 1993. "Common risk factors in the returns on stocks and bonds." *Journal of financial economics*, 33(1): 3–56.

Fama, Eugene F, and Kenneth R French. 2016. "Dissecting anomalies with a five-factor model." *The Review of Financial Studies*, 29(1): 69–103.

Feng, Guanhao, Stefano Giglio, and Dacheng Xiu. 2020. "Taming the Factor Zoo: A Test of New Factors." *Journal of Finance*, 75(3): 1327–1370.

Harvey, Campbell, Yan Liu, and Heqing Zhu. 2016. "... and the Cross-section of Rxpected Returns." *Review of Financial Studies*, 29: 5–68.

Jegadeesh, Narasimhan, and Sheridan Titman. 2023. "Momentum: Evidence and insights 30 years later." *Pacific-Basin Finance Journal*, 82: 102202.

Jensen, Theis Ingerslev, Bryan Kelly, and Lasse Heje Pedersen. 2023. "Is There a Replication Crisis in Finance?" *The Journal of Finance*, 78(5): 2465–2518.

Liu, Weimin. 2006. "A Liquidity-Augmented Capital Asset Pricing Model." *Journal of Financial Economics*, 82(3): 631–671.

Novy-Marx, Robert. 2013. "The other side of value: the gross profitability premium." *Journal of Financial Economics*, 108(1): 1–28.

Novy-Marx, Robert, and Mihail Velikov. 2015. "A Taxonomy of Anomalies and Their Trading Costs." *Review of Financial Studies*, 29(1): 104–147.

Pandey, Asheesh, and Sanjay Sehgal. 2016. "Explaining size effect for Indian stock market." *Asia-Pacific Financial Markets*, 23(1): 45–68.

Pastor, Lubos, and Robert Stambaugh. 2003. "Liquidity Risk and Expected Stock Returns." *Journal of Political Economy*, 111(3): 642–683.

Titman, Sheridan, K. Wei, and F. Xie. 2004. "Capital investments and stock returns." *Journal of Financial and Quantitative Analysis*, 39(4): 677–700.

Zhang, Lu. 2005. "The Value Premium." *Journal of Finance*, 60(1): 67–103.

8 Figures and Tables

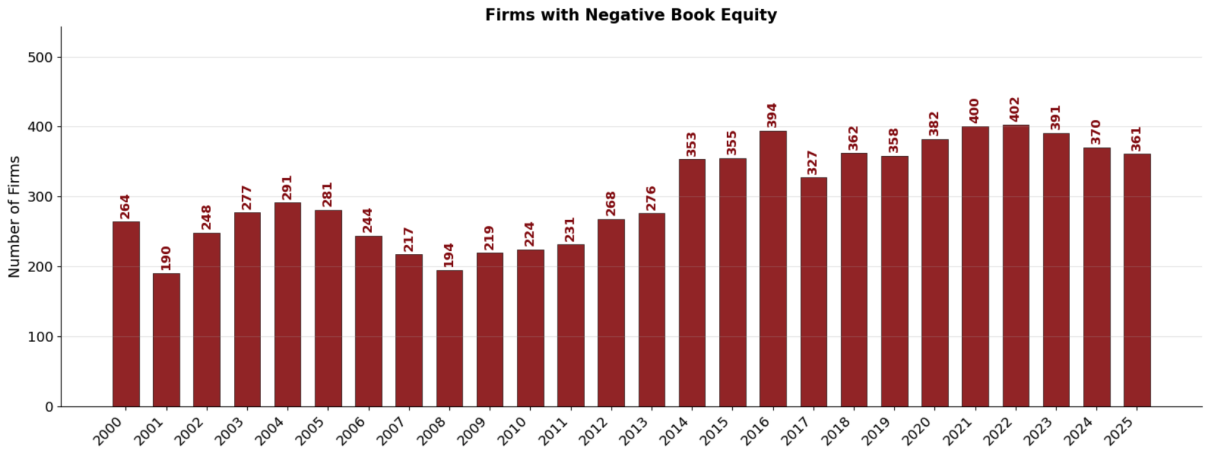

Figure 1: Firms with negative Book Equity as of Portfolio Formation date

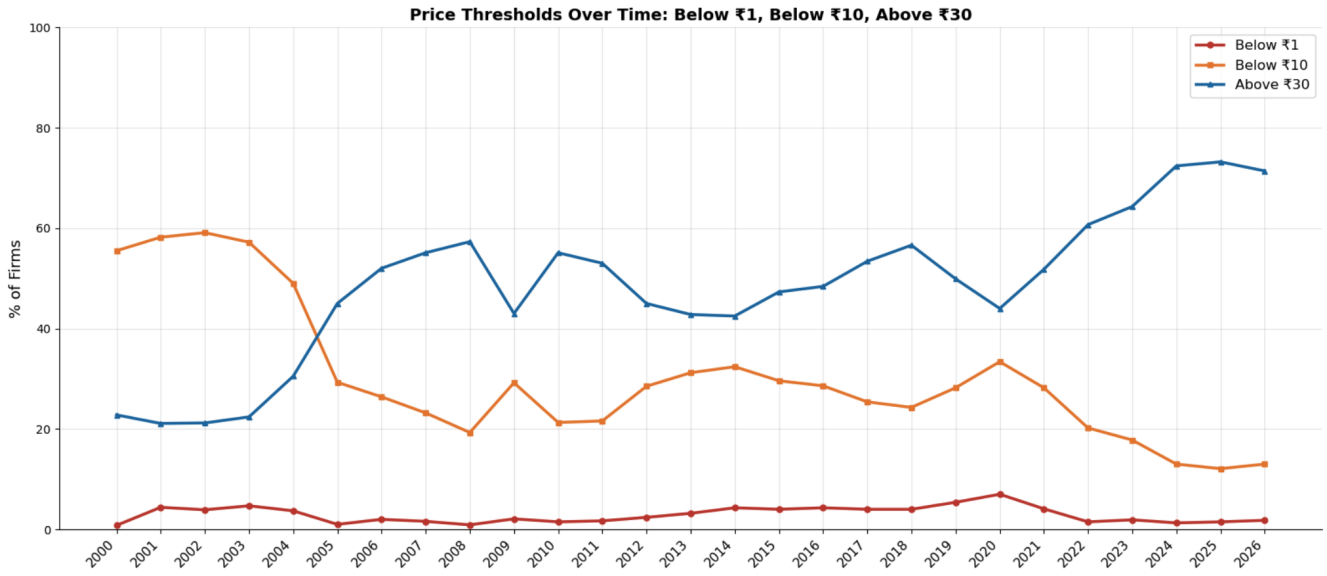

Figure 2: Percentage of firms in each price bucket, PY2000-PY2025

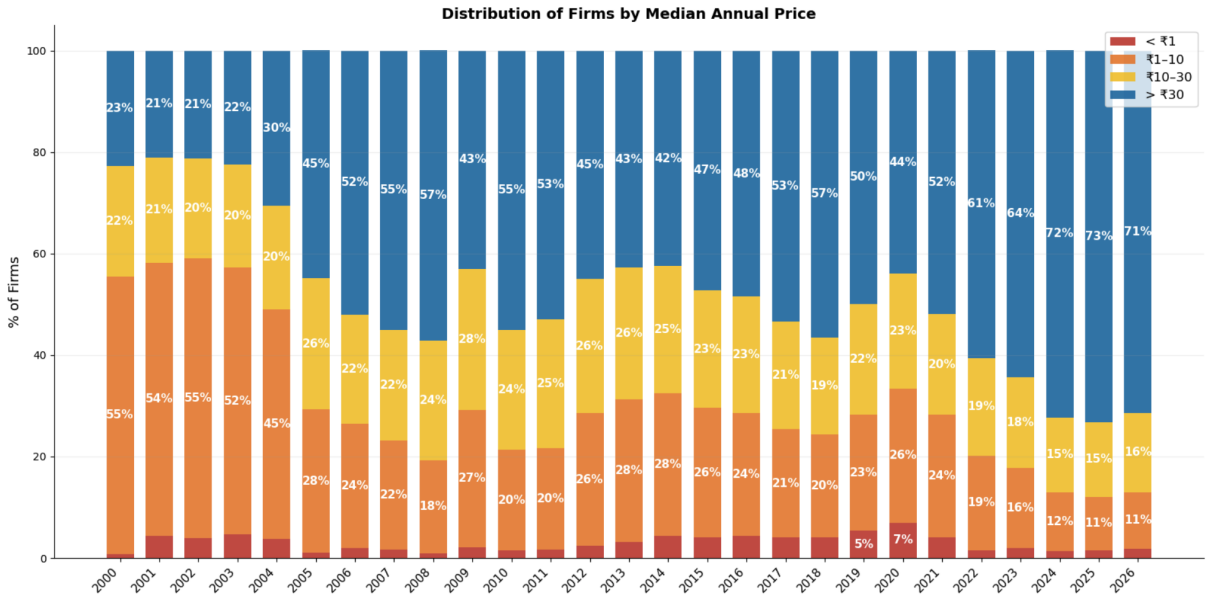

Figure 3: Distribution of firms in each price bucket, PY2000-PY2025

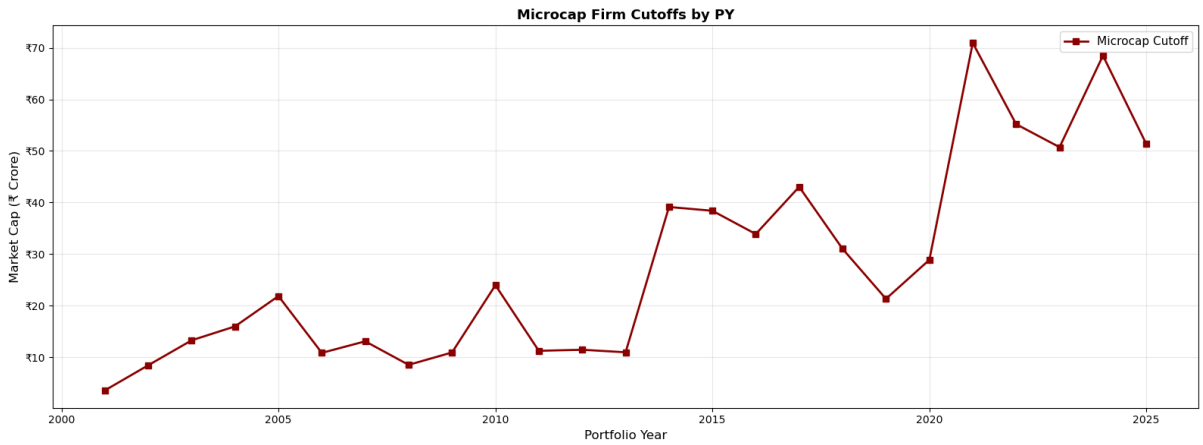

Figure 4: Microcap Cutoff in Crores

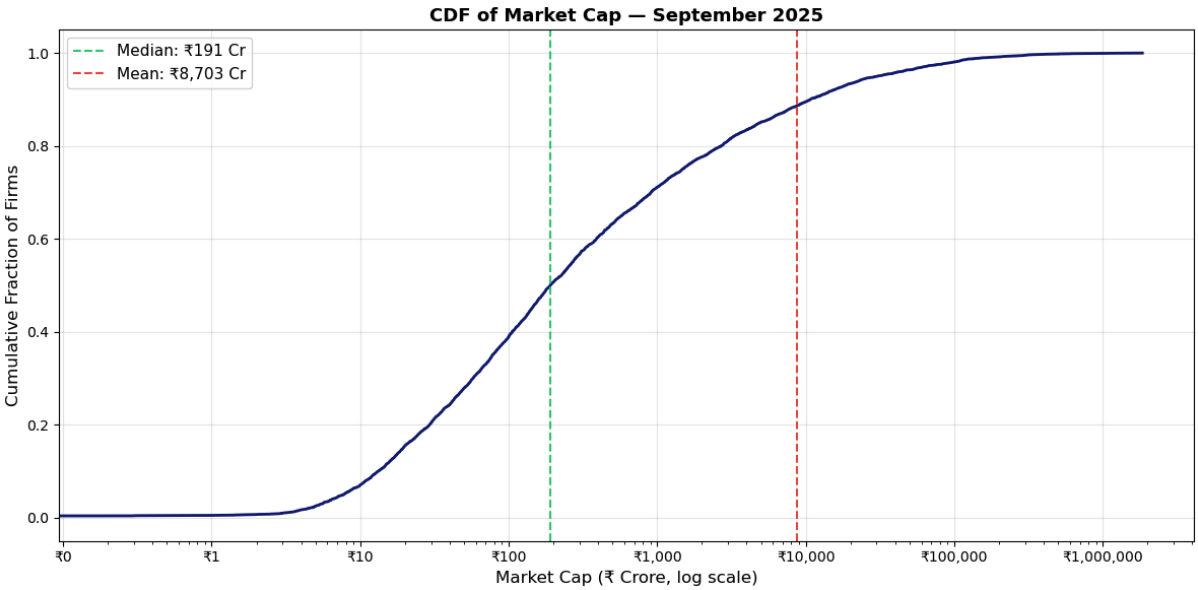

Figure 5: Firms with negative Book Equity as of Portfolio Formation date

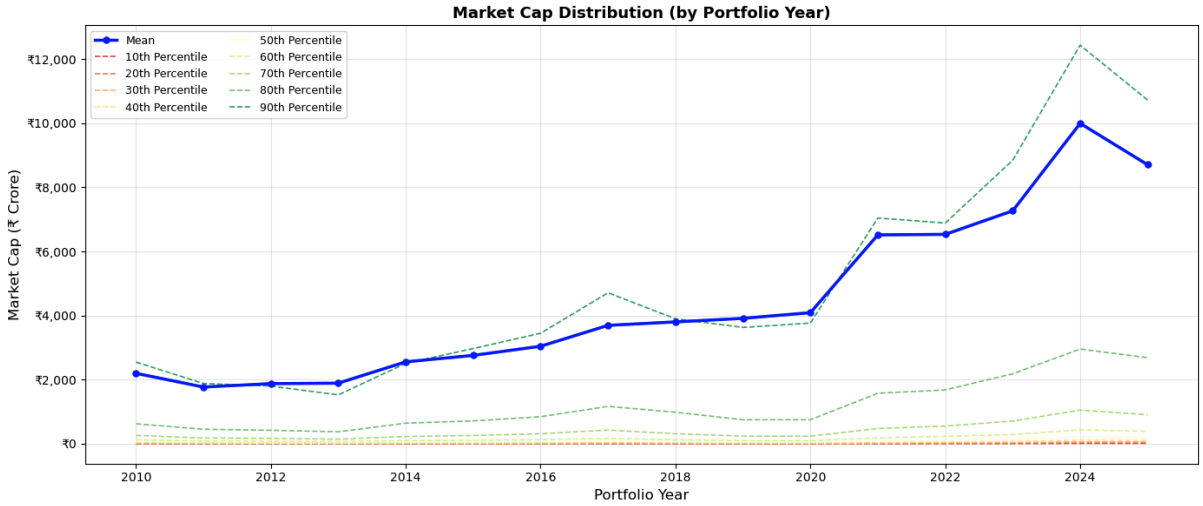

Figure 6: Firms with negative Book Equity as of Portfolio Formation date

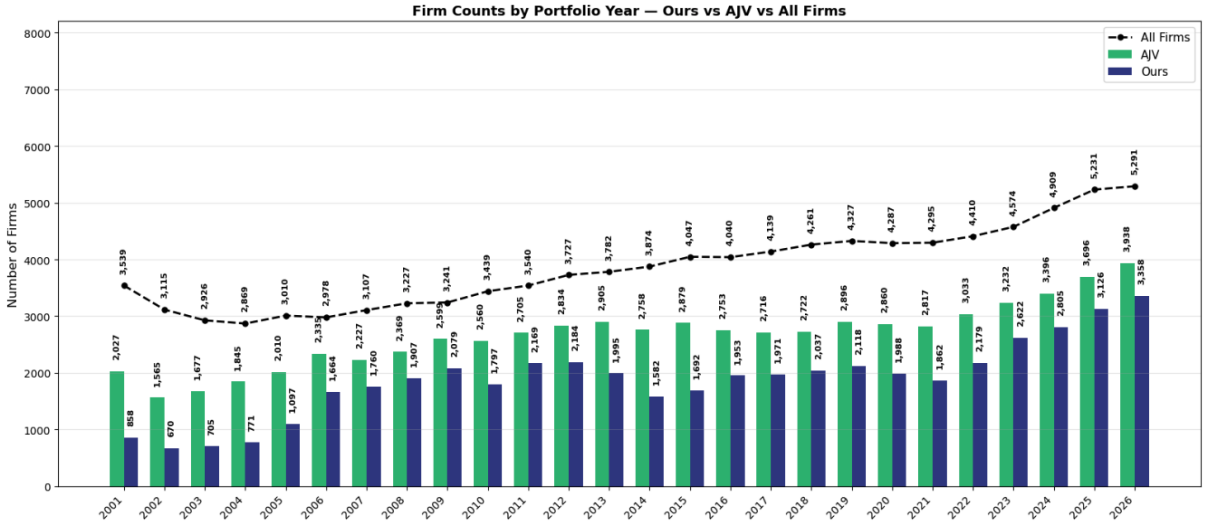

Figure 7: Firm counts by portfolio year. Black dotted line represents the total CMIE prowess base, the maroon and blue bars represent Our and AJV (Agarwalla, Jacob and Varma (2014)) tradable universe each portfolio year.

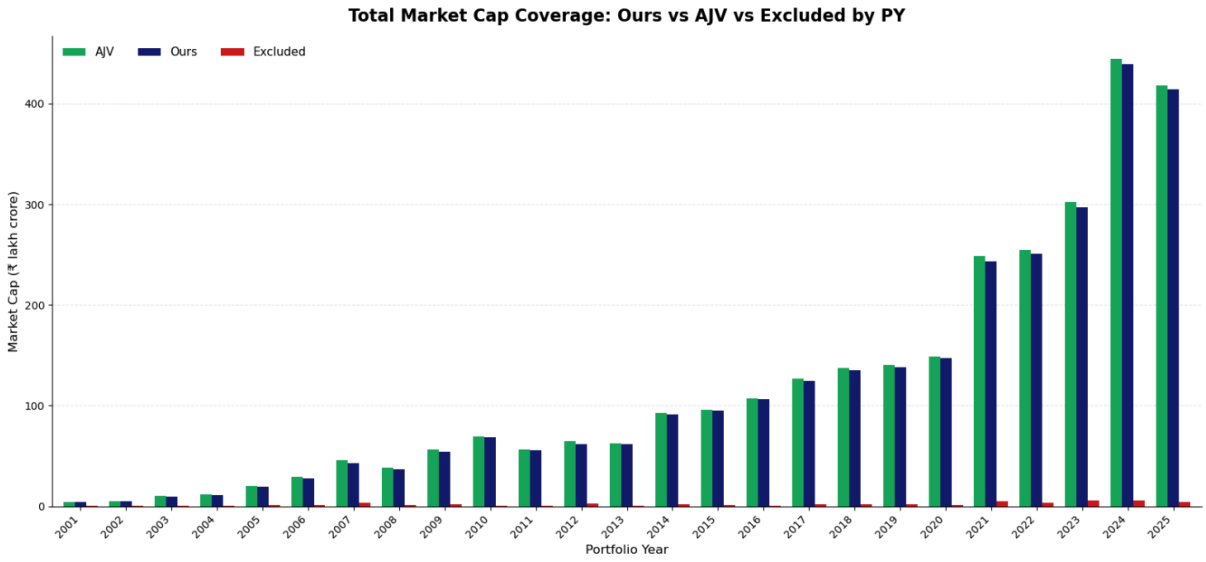

Figure 8: PY denotes portfolio year, defined as October of year t to September of year $t+1$.

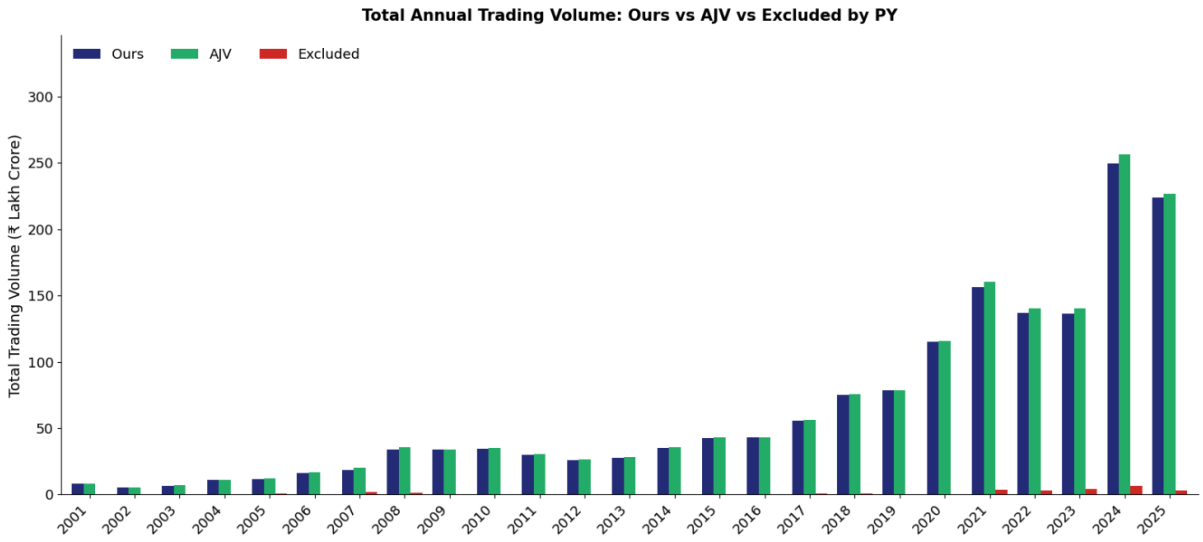

Figure 9: PY denotes portfolio year, defined as October of year t to September of year $t+1$.

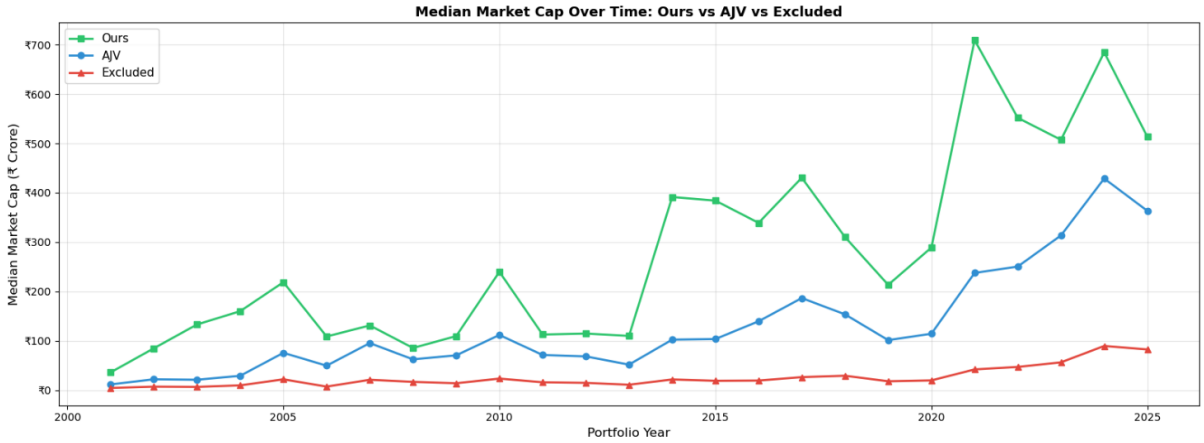

Figure 10: PY denotes portfolio year, defined as October of year t to September of year $t+1$.

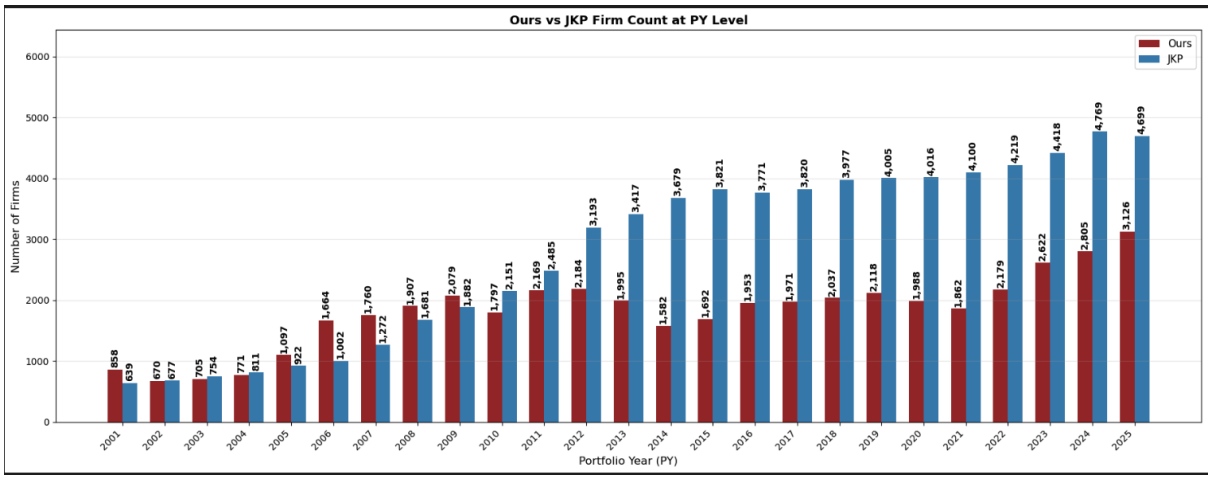

Figure 11: The red and blue bars represent the number of firms in the holding universe using Our and JKP's methodology respectively

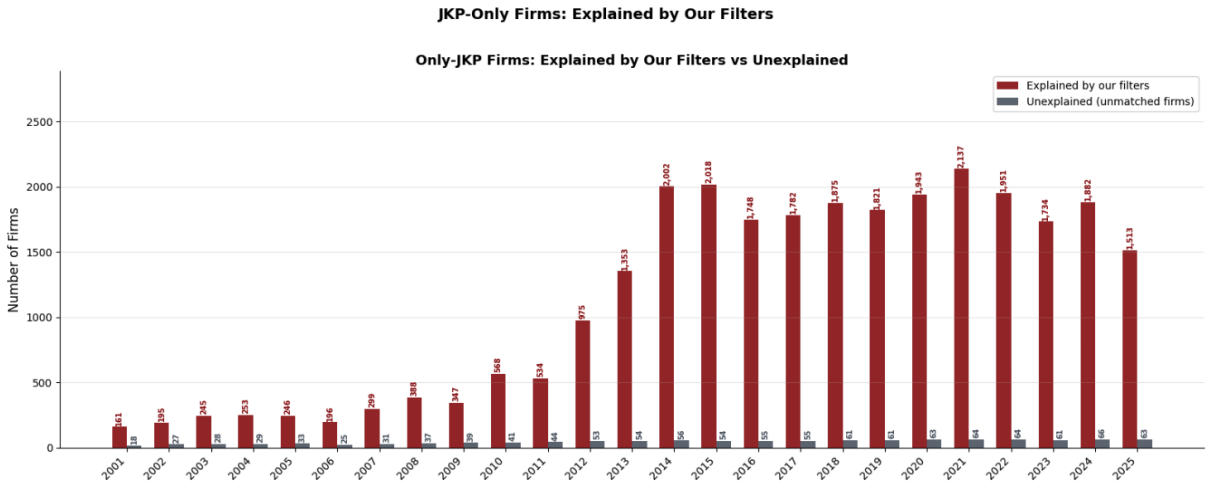

Figure 12: The red bars represent the numbers of firms are filtered from our exclusions criteria but included in JKP universe, the black bars are the unmatched/unexplained firms

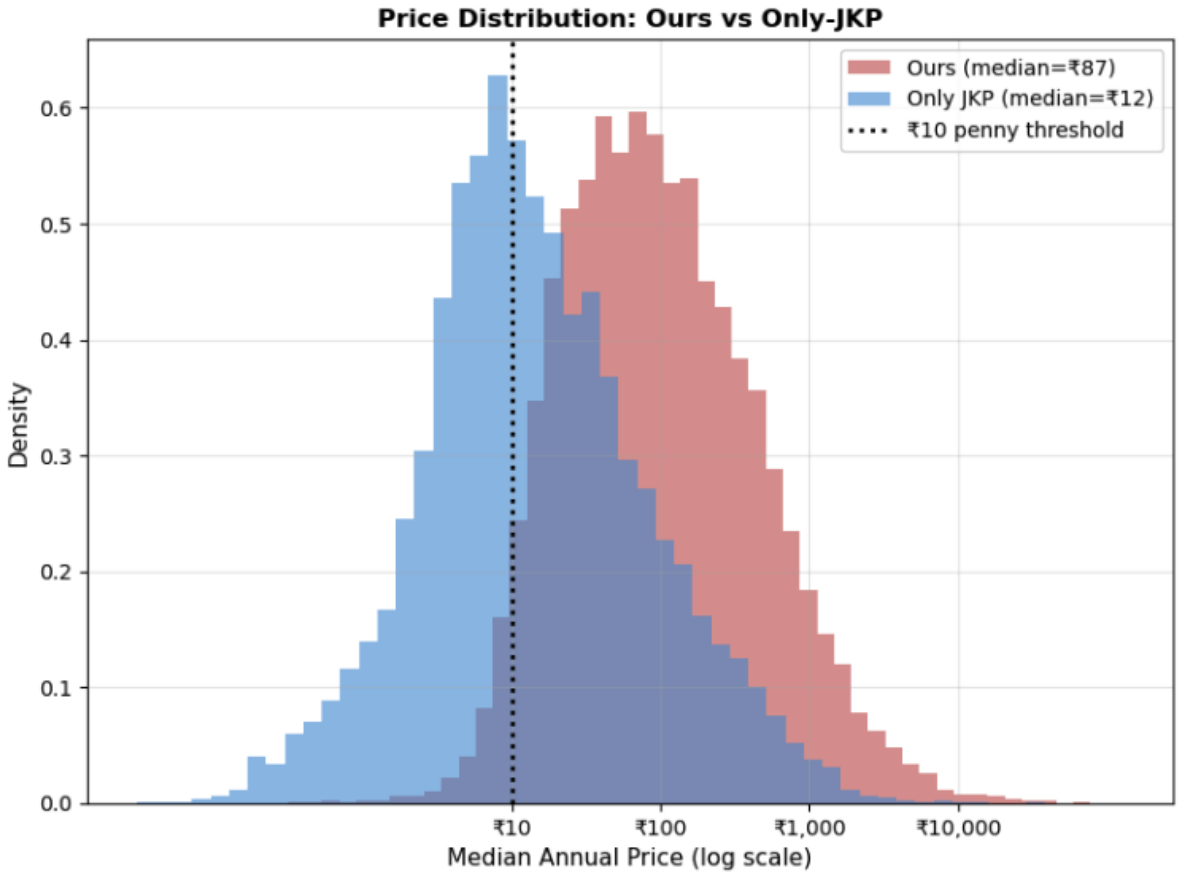

Figure 13: The red and blue bars show the distribution of median prices for Our and JKP universe

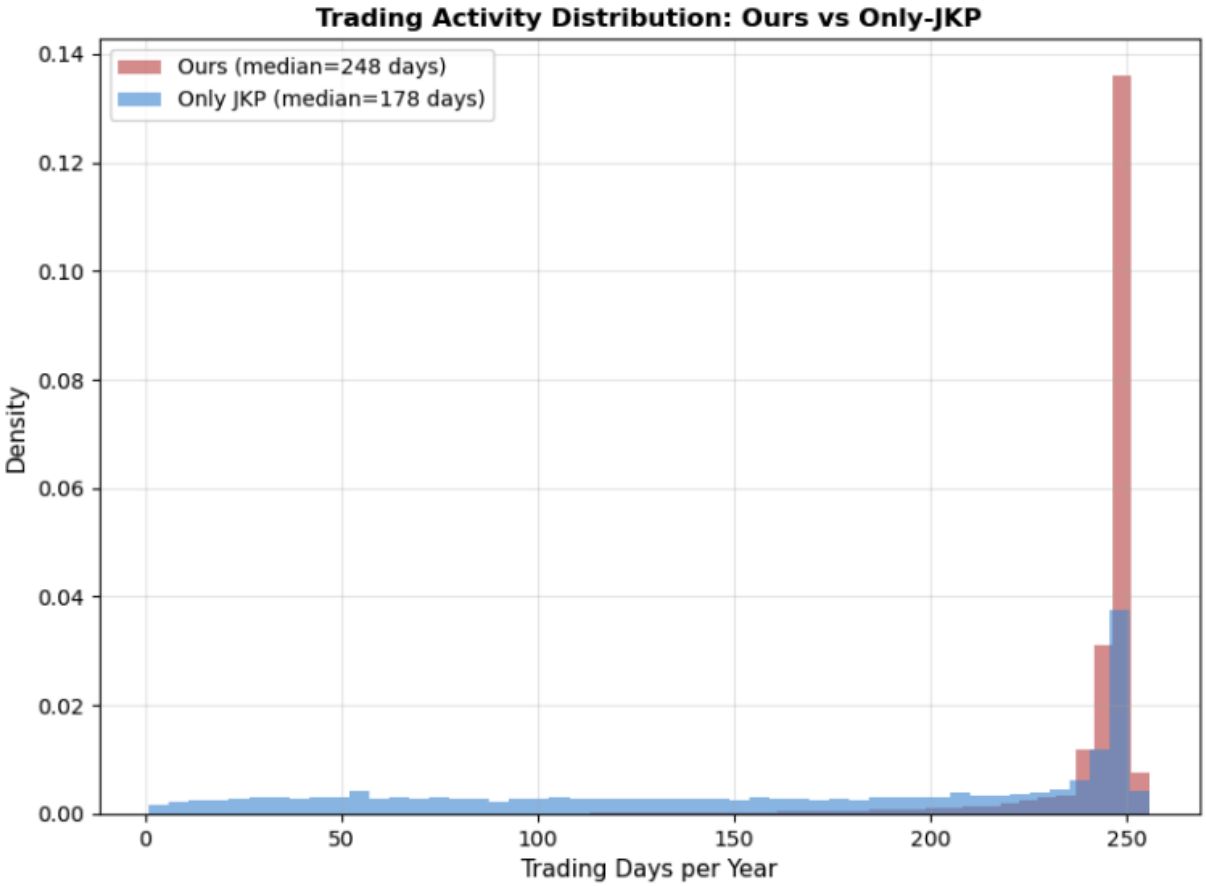

Figure 14: The red and blue bars show the distribution of trading days for Our and JKP universe

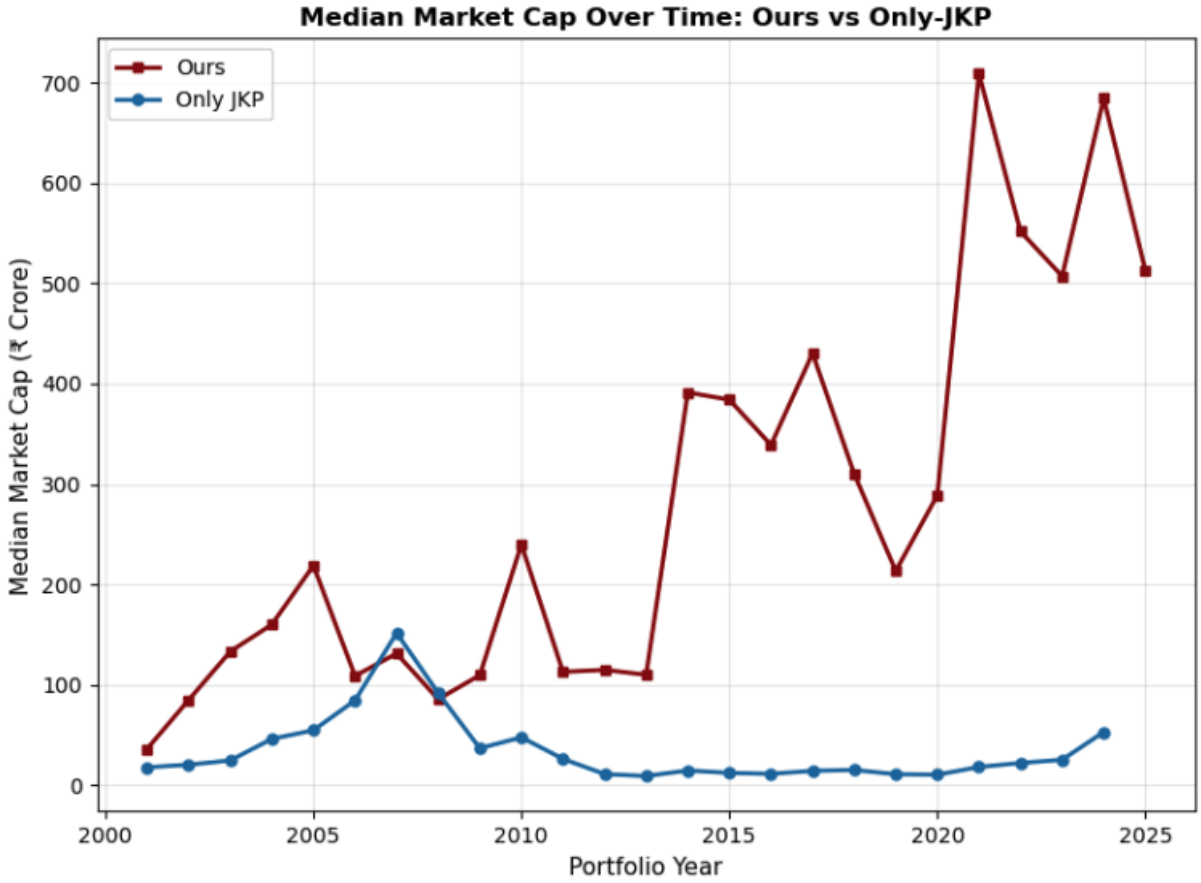

Figure 15: The red and blue lines represent the median market capitalization of firms part of Our and JKP's holding universe

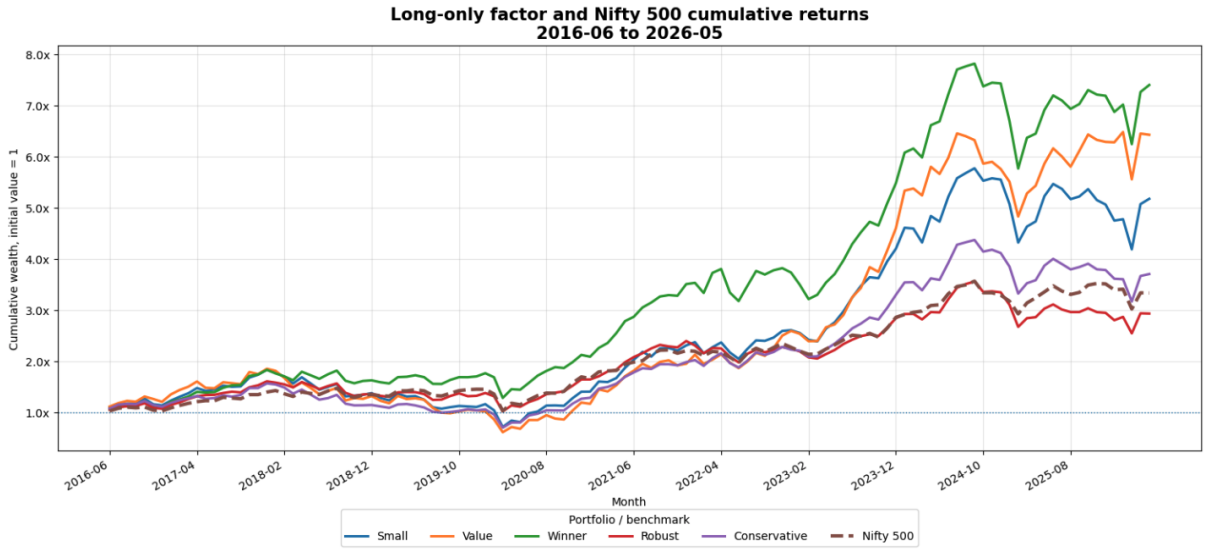

Figure 16: Each line represents long only cumulative wealth starting from ₹1 using value weighted methodology

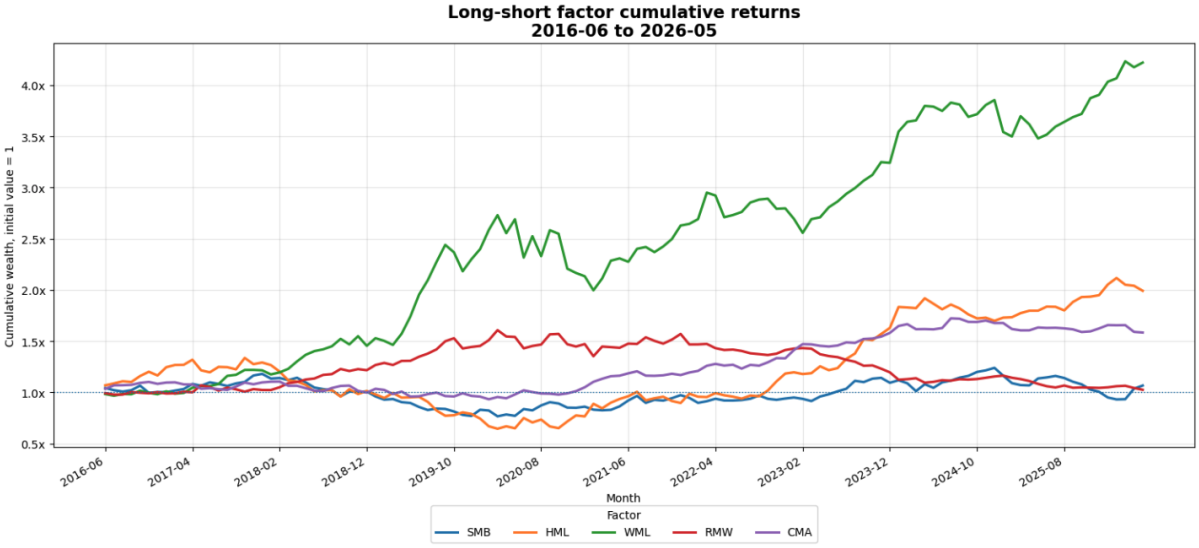

Figure 17: Each line represents long-short cumulative wealth starting from ₹1 using value weighted methodology

TABLE 1: Cross-sectional distribution of median annual stock price by portfolio year. Each cell reports the cross-sectional percentile of firm-level median daily prices within the portfolio year universe.

PY	**P10**	**P25**	**P50**	**P90**
2005	4.08	8.50	24.58	247.82
2010	4.90	12.50	36.95	323.70
2015	2.78	7.99	26.31	410.25
2020	1.70	6.30	22.75	349.38
2025	8.06	27.01	96.77	962.82

TABLE 2: Firm Counts by Portfolio Year: All Firms and Holding Firms

Portfolio Year	**All Firms**	**Holding Firms**
2001	3,539	858
2002	3,115	670
2003	2,926	705
2004	2,869	771
2005	3,010	1,097
2006	2,978	1,664
2007	3,107	1,772
2008	3,227	1,907
2009	3,241	2,079
2010	3,439	1,797
2011	3,540	2,162
2012	3,727	2,184
2013	3,782	1,995
2014	3,874	1,582
2015	4,047	1,692
2016	4,040	1,953
2017	4,139	1,971
2018	4,261	2,037
2019	4,327	2,118
2020	4,287	1,988
2021	4,295	1,862
2022	4,410	2,179
2023	4,574	2,622
2024	4,909	2,803
2025	5,231	3,126
2026	5,291	3,558

TABLE 3: Number of Firms Excluded Due to Negative Book Equity by Portfolio Year

PY	Negative BE Firms
2001	33
2002	10
2003	24
2004	27
2005	45
2006	113
2007	79
2008	80
2009	89
2010	49
2011	93
2012	92
2013	81
2014	60
2015	90
2016	128
2017	122
2018	116
2019	129
2020	99
2021	71
2022	98
2023	134
2024	117
2025	146

TABLE 4: Long and short legs for each factor return.

Factor	**Long leg**	**Short leg**	**Label update**
SMB	$SV + SN + SG$	$BV + BN + BG$	Annual (September)
HML	$SV + BV$	$SG + BG$	Annual (September)
WML	$SW + BW$	$SL + BL$	Monthly (both size and momentum)
RMW	$SR + BR$	$SW + BW$	Annual (September)
CMA	$SC + BC$	$SA + BA$	Annual (September)

Factor	Metric	Breakpoints	Labels
Op. Profitability	OpProf	30/70 pctile	W (Weak), N (Neutral), R (Robust)
Investment	Asset Growth	30/70 pctile	C (Conservative), N (Neutral), A (Aggressive)
Value	Book-to-Market	30/70 pctile	G (Growth), N (Neutral), V (Value)
Momentum	12-1 Mo Ret	30/70 pctile	L (Loser), N (Neutral), W (Winner)
Size	Market Cap	90th pctile	S (Small), B (Big)

TABLE 5: Factor Breakpoints and Labels

TABLE 6: 10-Year Long-Only Factor Performance: June'2016 - May'2026

Factor	Description	Cumulative	Ann. Returns	Ann. Excess	Volatility	Sharpe	Max Drawdown
Small	Size	4.2x	19.99%	14.37%	25.99%	0.55	-61.04%
Value	Value	5.4x	22.91%	17.30%	28.96%	0.60	-67.01%
Winner	Momentum	6.4x	22.45%	16.84%	21.19%	0.79	-29.95%
Robust	Profitability	1.9x	12.51%	6.90%	18.34%	0.38	-36.61%
Conservative	Investment	2.7x	15.54%	9.93%	21.63%	0.46	-55.34%
Nifty 500	Benchmark	2.3x	13.57%	7.96%	17.00%	0.47	-29.98%

TABLE 7: 10-Year Long-Short Factor Performance:: June'2016 - May'2026

Factor	Description	Cumulative	Ann. Mean	Ann. Mean Excess	Volatility	Sharpe	Max Drawdown
SMB	Size	0.07x	1.42%	-4.18%	12.54%	-0.33	-35.09%
HML	Value	0.99x	8.32%	2.72%	17.02%	0.16	-51.76%
WML	Momentum	3.22x	15.70%	10.10%	15.63%	0.65	-26.89%
RMW	Operating Profitability	0.03x	0.70%	-4.90%	9.49%	-0.52	-36.22%
CMA	Investment	0.58x	4.91%	-0.69%	7.77%	-0.09	-15.60%