1. Data Preprocessing

Raw data available included prices, dividend yield, outstanding shares and a top 1100 market cap mapping. These time series were available at different frequencies including weekly, monthly and quarterly. Since the construction methodology of the conservative portfolio requires rebalancing every quarter, we need the data points only every quarter, as that is our decision point. So, the first step was to filter and align all time series to make them quarterly.

Buyback yield was computed by calculating the decrease in number of market shares as a percentage of their rolling 12 month average. Buyback yield, is added to the dividend yield to compute the net payout yield (quarterly) for each stock.

Using the weekly price data, we compute the weekly returns, which are then used to compute rolling 36-month realized volatility on a weekly frequency. The last week of every quarter is sampled as the realized volatility which we consider at our decision point (quarterly rebalancing of portfolio).

The 12-1 month momentum is computed as the ratio of the stock price 4 weeks ago to the price of the stock 52 weeks ago. This is again sampled as the last week of each quarter to make it a quarterly series.

For analyses requiring ordering based on market cap, we also needed market capitalization data for all firms, which is simply computed as the product of the price and outstanding shares of each stock at each point in time. This is again sampled to take the last week of every quarter.

Data for the BSE-100 index was sourced from investing.com to use as a benchmark for portfolio performance.

The time series are trimmed to accommodate only the period for which all data is available for most stocks for uniformity purposes. So our analysis spans from Q1 2005 to Q1 2024.

For reference, all data preprocessing can be found in the 'preprocessing.ipynb' notebook. All the final processed and trimmed data is in the 'clean data' folder.

2. Analysis and Results

After all preprocessing was done, all required metrics for portfolio construction were available in neatly aligned quarterly time series.

We work with the understanding that portfolio rebalancing happens at the close of every quarter. So, our portfolio starts at the end of Q1 2005 and all positions are squared of at the end of Q1 2024.

2.1. Simulation of conservative and speculative portfolios with equal and market weighting

Portfolio construction follows exactly the methodology as described in the core paper, with the small difference that our pool of stocks is the top 1100 stocks by market cap at each quarter, instead of the top 1000. For market weighting, we use the market cap which is computed as the product of price and outstanding shares for a given quarter. All portfolios are simulated with a starting cash value of 1 crore to be invested. portfolios are evaluated based on growth, quarterly returns, volatility, and rolling 36-month Sharpe ratio.

Both conservative and speculative portfolios are compared with the the BSE 100 index as a benchmark.

The results of our analysis align with those in the core paper, with the conservative portfolio outperforming the BSE with both weighting methods, and the speculative portfolio falling short of the same, with both weighting methods.

The conservative portfolio also seems to consistently maintain a higher rolling sharpe ratio than the BSE and speculative portfolios.

These results reinforce the utility of the conservative investing formula, which investors can leverage to make significantly higher returns at much lower risk than a speculative portfolio.

The code for this portfolio simulation is in the 'Conservative Formula.ipynb' notebook.

2.2. Testing for the low-risk anomaly with decile portfolios

Decile portfolios are constructed by grouping stocks in increasing order of realized 36 month volatility out of the top 1100 market cap stocks at each quarter. With decile 1 being the lowest and decile 10 being the highest volatility.

Our analysis reveals that not only do returns not monotonically increase with volatility, but even show a slight decrease instead. This is slightly different from the results described in the core paper, where the relationship is initially flat and then shows a slight downward slope at higher risk.

The code and plots for this can be found in the 'deciles.ipynb' notebook.

2.3. Testing for the low-risk anomaly with across market cap sized with 5x5 size-volatility portfolios

We run two simulations, testing 5x5 portfolios with both equal weights and Fama French size breakpoints.

The low volatility anomaly appears across all 5 size groups with equal buckets, but with fama-french size breakpoints, the result does not hold with the size-1 bucket, indicating that the stocks in the 3rd percentile by market cap do not follow this trend. This shows that for small-cap stocks, higher volatility is indeed rewarded with higher returns.

However, upon closer examination, the returns and volatilites of the size-1 portfolios are extremely inflated, way beyond all other size buckets, possibly indicating the presence of outliers or extremely noisy data. So, we can conclude that the Indian market largely does follow the low risk anomaly, across sizes.

The code and plots for this can be found in the '5x5 size-volatility.ipynb' notebook.