How did the size effect and volatility influence stock returns before, during and after Covid-19 in the NYSE and NASDAQ U.S exchanges?

Bernardo Ramos da Mata

Dissertation written under the supervision of Professor Eva Schliephake

Dissertation submitted in partial fulfilment of requirements for the International MSc in Finance, at the Universidade Católica Portuguesa, September 2025.

1. Abstract

This study investigates how firm size and volatility impacted U.S. stock performance before, during, and after Covid-19 and whether volatility carries an independent price of risk once other factors have been taken into account. By using NYSE and NASDAQ common stocks, monthly double sorted portfolios are constructed based on two sizes (small and large firms) and five volatility quintiles using a 3-month rolling window of volatility. As compared with other sub-sample periods, portfolio returns exhibit a sharp decline during the Covid-19 crash period. In general, low-volatility stocks outperformed high-volatility stocks during the crash. Furthermore, this study provides evidence of a constructed VOL factor and its behavior over different sub-sample periods between 2018 and 2022. In all phases except for the crash period, the VOL factor is negative, as low-volatility stocks delivered lower average returns outside of the crash period, and higher returns during the crash period. An augmented Fama-French five factor model (FF5) is used to estimate factor exposures (betas) and a Fama-Macbeth two-pass procedure is applied to estimate the prices of risk (average λ's) and test if volatility and the FF5 is priced in the cross-section. During the pandemic, the estimated price of volatility risk turned positive ($\lambda^{VOL} > 0$), indicating that higher returns were attributed to portfolios with a positive β_{VOL}, favoring robust low-volatility companies. Overall, the findings contribute to understanding the behaviour of returns based on volatility across market cycles.

Keywords: volatility factor, size effect, fama-french five factors, fama-macbeth, covid-19, stock returns, asset pricing

Title: How did the size effect and volatility influence stock returns before, during and after Covid-19 in the NYSE and NASDAQ U.S exchanges?

Author: Bernardo Ramos da Mata

Resumo

Este estudo investiga como o tamanho das empresas e a volatilidade influenciaram as rendibilidades das ações nos EUA antes, durante e após a COVID-19 e se a volatilidade possui um preço de risco próprio após controlar pelos restantes fatores. Com ações ordinárias da NYSE e NASDAQ, constroem-se carteiras mensais através de dupla ordenação por dimensão (pequenas e grandes, ponto médio da NYSE) e por quintis de volatilidade, medida por desvio-padrão móvel de 3 meses. As rendibilidades caem acentuadamente no crash e, nesse período, as ações de baixa volatilidade superam as de elevada volatilidade. Analisa-se ainda um fator VOL construído e o seu comportamento em subperíodos de 2018–2022. Em todas as fases, exceto no crash, o fator VOL é negativo, dado que as ações de baixa volatilidade apresentam rendibilidades médias inferiores fora do crash e relativamente superiores durante o crash. Recorre-se a uma versão ampliada do modelo de cinco fatores de Fama-French (FF5) para estimar betas e aplica-se o procedimento de dois passos de Fama-MacBeth para obter preços de risco (λ médios) e testar se a volatilidade e os fatores do FF5 são remunerados na secção transversal. Durante a pandemia, o preço do risco da volatilidade torna-se positivo ($\lambda^{VOL} > 0$), indicando que rendibilidades superiores se associam a carteiras com β^{VOL} positivo, favorecendo empresas de baixa volatilidade robustas. No conjunto, os resultados ajudam a interpretar o comportamento das rendibilidades ao longo dos ciclos de mercado.

Palavras-chave: fator de volatilidade; efeito dimensão; cinco fatores de Fama-French; Fama-MacBeth; COVID-19; rendibilidades acionistas; avaliação de ativos.

Titulo: How did the size effect and volatility influence stock returns before, during and after Covid-19 in the NYSE and NASDAQ U.S exchanges?

Autor: Bernardo Ramos da Mata

2. Introduction

2.1 Big Picture Motivation

The COVID-19 pandemic caused the sharpest and most concentrated period of uncertainty in equity markets since records began. Between February 24th and March 24th, 2020, the world experienced 18 market jumps within 22 days, more than any period before with the same number of trading days (Baker et al., 2020). The frequency of jumps during this period amounted to 20 times the average pace since 1900. In response to these movements, central banks acted in a manner that changed investors' perceptions of traditional equity risk premia (Ohara & Zhou, 2021). The Covid-19 period presents a unique opportunity for the study of cross-sectional patterns of anomalies in the context of market cycle shifts and changes in risk appetite. There are two main anomalies present in this study: the size effect and the volatility effect. As Fama and French (1993) demonstrated, small-capitalization stocks have historically outperformed large-capitalization stocks and gained higher returns on average in the USA. Furthermore, the low-volatility anomaly has demonstrated that high levels of total or idiosyncratic volatility are associated with lower returns (Ang et al., 2006). The low-volatility anomaly has shown that low-risk portfolios earn higher risk-adjusted returns than high-risk portfolios, contradicting the risk-return trade-off suggested by CAPM (Frazini & Pedersen, 2014). These relationships may be affected by pandemic periods. Funding constraints and lower margins lead investors to shift from risky and volatile assets to safer, lower beta assets, resulting in a flight-to-quality effect on prices (Brunnermeier & Pedersen, 2009). Recent studies show that the premium has widened during 2020-2021 as liquidity injections and retail participation disproportionately benefitted small firms (Barber et al., 2022). The Covid-19 pandemic shortened the duration of volatility dynamics into a matter of weeks: a sharp decline, immediate government intervention and a rapid rebound, these market shifts are precisely what cause changes in the pricing of size and volatility. It is important to understand how these two characteristics behaved before, during, and after Covid-19 to be able to determine if the size premium and volatility relationship are stable risk compensations or if they depend on the market state.

2.2 What this paper does

This study focuses on a phase-by-phase assessment of how firm size and volatility influenced stock performance of all common stocks excluding financial and insurance companies on the NYSE and NASDAQ exchanges from January 2018 to December 2022, and how their influences evolved across pre-pandemic, pandemic, and post-pandemic periods. Research is conducted on a full sample (January 2018 to December 2022), pre-pandemic (Jan 2018-Dec 2019), crash (Jan 2020-Mar 2020), rebound (Apr 2020-Dec 2021) and post-pandemic (Jan 2022-Dec 2022). To answer the question, a two-dimensional portfolio sorting method is constructed: stocks are sorted by size (Small and Large) and volatility quintile (Vol-1 to Vol-5), where Vol-1 is the lowest volatility and Vol-5 is the highest volatility, resulting in ten portfolios per month. This results in 10 dynamic portfolios that are rebalanced monthly in accordance with changes in firm size and volatility. Afterwards, value-weighted (VW) and equal-weighted (EW) returns are computed for each portfolio in order to evaluate the performance of each portfolio. For each size, we construct a VOL factor, which is defined as the spread between the lowest volatility quintile (Vol-1) and the highest volatility quintile (Vol-5).

Furthermore, an investigation into if volatility has a price of risk that changes across phases is addressed, using the Fama-Macbeth method. By using a two pass Fama-Macbeth method, the Fama-French five factor model is augmented with a VOL factor in order to assess its pricing capability. In the results of the portfolio returns, the left side of the equation consists of the same 10 value-weighted size and volatility portfolios. During the first pass, rolling factor exposures (betas) are estimated for each portfolio based on 60-month windows from 2013 onwards in order to have strong historical data by January 2018 (pre-pandemic period). To avoid look-ahead bias, exposures for month t are estimated using data up to month t-1. For the second pass, a monthly cross-sectional regression of portfolio excess returns on the betas for the six factors (5FF and VOL factor) is performed, using weighted least squares to weight each portfolio's observation by how reliable it is, using the inverse of its first-pass residual variance ($w_{i,t} = \frac{1}{\sigma^2_{e,i,t}}$), so the monthly factor prices are driven by signal rather than noise. Thus, low residual variance will receive a higher weight through up-weighting and high residual variance will receive a smaller weight through down-weighting. This makes a portfolio with quieter, better explained returns a more precise observation in the cross-section, so it counts more in estimating that month's price of risk (λ_t). Then, the time-series of factor prices (λ_t) is summarized over the entire sample (2018-2022) and each sub-sample period, with Newey-West t-statistics to account for any serial correlation.

This investigation contributes to the study of whether volatility earns an independent price of risk over and above FF5, and whether the price changes over time. As a result of the construction of a VOL factor which is orthogonal to the SMB and size neutral by design, any price associated with the factor is not a spurious size effect. It is possible to trade the factors and their economic magnitudes translate directly into expected return impacts per unit of exposure, and phase specific time series provide information on risk management, factor allocation, and stress testing in different market phases, including the Covid-19 pandemic.

2.3 Preview of main findings

The main findings reveal that the size premium shows a shift during the different stages of the sample period. First, the volatility effect flips sign across time periods. In the pre-pandemic sub-sample, the VOL spread (low-Vol – high-Vol) is negative across all time windows with 3-month rolling window volatility – and remains negative with 6 and 12 months for both sizes (Large and Small) and weighted portfolios (EW and VW). During the crash and peak recession, VOL turns positive as investors move from riskier to safer stocks, for all rolling windows. During the rebound period it becomes negative again. The pattern inverts sharply and persists for 6 and 12 months rolling windows as well. In the post-pandemic phase, the effect fades showing that we return to normality and the VOL is negative during this time period.

In terms of average returns, small-cap stocks trail large cap stocks in the pre-pandemic period for both EW and VW, across all portfolios except for small-vol5 (4.80% VW) which is higher than large-vol5 (3.01% VW). During the crash, all portfolios average returns dropped significantly and they all experienced negative returns. The small portfolios suffered heavier losses than the large in VW terms. There was an average decrease in large-vol-1 returns of -4.40% and a decrease in small-vol-1 returns of -6.65%, while large-vol-5 experienced a decrease of -16.46% and small-vol5 experienced a decrease of -23.54%. The rebound period saw large portfolios outperform small portfolios by a very narrow margin in vol-1 to vol-3, and small portfolios outperform the large in vol-4 and vol-5. As a whole, large and small portfolios in the higher volatility quintiles performed better than those in the lower volatility quintiles. The results indicate that the low-volatility anomaly's reversal was primarily concentrated in the early-rebound phase: during the explosive rally from roughly April 2020 through early 2021 (driven by monetary stimulus and reopening optimism), the market for high-volatility stocks outperformed substantially, thereby eroded the advantage of low-volatility stocks. As the market became more selective and fundamentals resumed their dominance in subsequent sub-periods (post-vaccine), the low-volatility anomaly continued to reverse as high volatility stocks outperformed lower volatility stocks. This is contrary to expectations, since it was expected that stocks would begin reverting more toward low volatility and strong stocks than high volatility and risky stocks. There may be a longer period of time required for the anomaly to return to normal. This temporal insight suggests that the "flight from safety" in favor of volatile stocks was not as short-lived as expected and lasted longer than anticipated. During the recovery phase as well as post-pandemic, investors continued to favor higher volatility stocks both in large caps and small caps. By clarifying the timing, this sub-period analysis reinforces that the low volatility anomaly was a short-lived, market cycle-specific event rather than a new normal. Post-pandemic, all portfolios except small-vol 5 (5.68%) had negative returns for VW, although "large" portfolios suffered less than "small" portfolios. The average return VW for small vol-1 was -2.11%, while the average return VW for large vol-1 was -1.43%. Large vol-3 suffered -1% while small-vol3 suffered -2.71% and large-vol5 had -0.08% while Small vol-5 had 5.68%. There was a preference for large caps across all quintiles during the crash. During the rebound, the large cap advantage narrowed to only slightly better results than small caps and flipped in favour of small caps in the two highest volatility quintiles.

According to the Fama-Macbeth cross-sectional regression results incorporating the FF5 factors and the VOL factors, significant patterns in risk pricing are observed. Over the entire sample, volatility risk emerges as an economically significant and statistically significant factor. Throughout all analyses, statistical significance will be assessed at the 5% level, corresponding to a two-tailed critical value of ±1.96 for the t-statistic. The price of volatility risk (λ^{VOL}) is about -5% per month (t = -5.1), indicating that stocks with greater exposure to volatility earned substantially lower returns on average. Increasing a portfolio's volatility beta by one unit would decrease expected monthly returns by approximately 5%, emphasizing investors' aversion to volatility. A significantly negative premium was also observed for the profitability factor (RMW), suggesting that more profitable firms were not rewarded with higher returns during this period.

Examining sub-sample periods around Covid-19 revealed shifts in factor pricing that support the portfolio returns analysis. The pre-pandemic period already showed negative prices and were statistically significant for volatility risk (λ^{VOL}) and value (λ^{HML}), indicating that low volatility and growth stocks had higher valuations during this time. This shows that safety (low-volatility exposure) was prized as investors were willing to accept lower expected returns, which means these stocks traded at high valuations already. The market required a premium for bearing high-volatility risk, so investors demanded higher returns to hold it (high volatility was rewarded more in this phase), meanwhile low-volatility stocks were bid up (higher valuations), which implies lower expected returns consistent with negative λ^{VOL}. During the crash, market risk premium (MKT-RF) turned negative as high beta stocks fared worst and the value factor (HML) became even more negative as value stocks provided little protection during the crash and didn't hedge against Covid-19. The λ^{VOL} price turned positive during the crash, consistent with the portfolio returns analysis sign flip. This unusual positive λ^{VOL} suggested that exposure to the volatility risk factor was momentarily rewarded and likely reflected a flight to safety premium, as low-volatility stocks outperformed high-volatility stocks. In the rebound phase, factor pricing reversed again: the market factor earned a strong positive premium, but the volatily risk returned to more negatively priced values than the pre-pandemic. This shows that high volatility stocks severely underperformed on a risk-adjusted basis, and more stable and low-volatility stocks had better returns. Value and profitibality factors were negative – as growth firms outperformed. Finally, after the pandemic (post-pandemic), some factor prices began to reverse: the value premium turned positive after years of being negative, indicating the revival of value stocks, and the size factor turned negative, indicating that small caps lagged large caps. The volatility factor continued to carry a negative price of risk. Importantly, the volatility factor continued to carry a negative price of risk. Given that VOL = low vol – high vol, this means that low volatility defensive stocks were penalized ($\beta_{VOL} > 0$) while high volatility tilted stocks were rewarded ($\beta_{VOL} < 0$). This pattern is consistent with a risk-on trend as investors begin to take risks again following the pandemic.

Overall, the volatility risk accounted for a significant portion of cross-sectional returns, with constant negative prices except during the crash period, demonstrating how an extreme shock may temporarily reverse the usual risk-return trade-off. Outside the crash, λ^{VOL} is negative, implying that the market rewards high volatility tilts and discounts defensive tilts. As a result, investors accept lower returns for low-volatility exposure, as if paying for insurance, while they demand a premium to hold high-volatility exposures. Investors paid up for defensiveness during the crash. According to the evidence, volatility risk varies with time period and market cycle, indicating a time-varying risk appetite for investors.

2.4 Extensions/heterogeneity analyses/mechanisms

The research extends to the measurement of risk as the baseline VOL factor is constructed using a rolling 3 month window volatility; accordingly, we recompute the VOL factor using 6 and 12 months rolling windows for estimating volatility. The comparison of these time windows allows us to determine whether post-crash investment in low volatility awards is a short-term phenomenon that reflects a brief mean-reversion in very volatile stocks, or if it is a more persistent phenomenon. According to the results for increasing the rolling window from 3 to 6 and 12 months during the crash period, the VOL is positive for all time windows for VW and EW. Statistical strength decreases from t=2.26 (3 months) to t=1.59 (6 months) and t=1.19 (12 months) as longer duration windows dilute the crash months with pre-crash data. Results of the pre-pandemic study indicate that VOL is negative due to the outperformance of high-vol stocks, which are statistically significant for VW and a slightly lower significance for EW. For each rolling window, the rebound phase demonstrates a negative VOL that is highly significant. Once the crash period has ended, the inversion remains consistent and shows that investors have changed their risk behavior towards more risky investments. This could indicate that the anomaly flip was driven by short-run dynamics if the VOL factor's inversion disappeared with a longer-term volatility window. Low volatility premiums post-pandemic inversion should not be dismissed if they persist across all time frames.

Based on the NYSE exchange as the benchmark, the bottom decile of market capitalization was excluded. The purpose of this analysis is to address the criticism of micro-cap bias and determine whether these stocks are driving the results. Using these measures, the order of portfolio returns by volatility remains unchanged. When microcaps are removed, the results did not change significantly and EW levels decreased slightly. In addition, the pattern of a positive VOL factor during the crash period and a negative VOL factor during the pre-pandemic and rebound period remained intact, and small caps retained their advantage during the rebound period.

In addition, total volatility was replaced with idiosyncratic volatility (IVOL) – calculated as the rolling standard deviation from factor regressions – and it produces the same qualitative sign flip when changing periods. There is a positive mean IVOL during the crash period, a strong negative IVOL during the rebound period, and a negative IVOL between -1 and -2% during the post-pandemic period. Since the inversion persists under IVOL, it cannot be attributed solely to market swings in beta or other common factors.

There is also the possibility that momentum and short-term reversals may play a role in determining the size of the prize and the volatility of the market. Momentum refers to the likelihood that stocks that have performed well (or poorly) in recent history will perform similarly in the near future. A short-term reversal occurs when previous losers bounce back (or winners mean revert negatively) (Jegadeesh & Titman, 1993). Under 3-month rolling window volatility, stocks were divided into terciles by 6-month momentum (lag 1-month). Pre-pandemic data revealed that every tercile experienced negative mean VOL in EW and VW terms. There was a positive mean VOL for every tercile during the crash phase, as stocks with low volatility outperformed, regardless of their previous performance. Finally, the VOL of the rebound phase flipped sign again and became negative in every momentum tercile. The inversion is strongest with prior losers (momentum tercile 1), indicating that high-vol stocks outperformed low-vol stocks and is highest with stocks that had fallen the most. Even in the post-pandemic phase, negative momentum terciles persist.

Short-term reversals follow the same pattern with a higher frequency. While the crash was occurring, VOL was positive across REV terciles. This is consistent with a move into safer portfolios even among those whose values had just fallen. The rebound VOL spread is most negative among recent losers (tercile 1), negative but smaller in the middle tercile, and near zero among winners (tercile 3). In addition to the momentum explanation, the negative VOL spread in the rebound phase isn't confined to one momentum bucket, but appears across all three terciles, but is most pronounced among prior losers where bounce back pressure is greatest. As liquidity pressures and negative sentiment ease, beaten down and more volatile stocks are likely to rebound. It was evident that high volatility outperformance was widespread, but concentrated particularly in previous losers during the rebound phase.

Furthermore, the skewness tercile splits behave differently depending on the distribution of "lottery like stocks"; in particular, the effect is strongest where there is a high level of skewness. During the pre-pandemic period, mean VOL is slightly positive for low-skew stocks, but turns negative as skewness increases in the tertiles. As a result of the crash, defensive portfolio picking behaviour displays itself, and the mean VOL is positive among the low and mid skew terciles of portfolios, but it is not positive among lottery-like stocks. As a result of the rebound, the mean VOL turns negative in every skewness tercile, and it is the most negative for high-skewness stocks compared to low-skewness stocks. In 2022, the spread remains negative among high skew names, but is slightly positive for stocks with a low skew, consistent with normalization. As a result, the skewness splits indicate that the COVID-19 inversion of low volatility premium was broad based, but that it was amplified by the lottery-like segment of the market – those stocks with positively skewed payoffs. As a result, investors sought riskier assets on the rebound and chased high volatility assets the most.

The Inclusion of the VOL factor as an extension to the Fama-French five factor model explored how factor pricing relationships varied during the different market phases. The VOL factor is defined monthly as low Vol minus high Vol within each size group and then averaged across sizes and placed into the Fama-French model and pricing it with Fama-Macbeth. Using 10 VW Size and volatility portfolios as test assets. The cross-section of each month is regressed on the betas to obtain six risk prices for the five factors and the VOL factor. The estimation is performed using weighted least squares (WLS) with weights from the first pass (down-weighting noisy portfolios) and Newey-West heteroskedasticity and autocorrelation (Newey-West HAC) with standard errors for the mean of each λ. This fixes the standard errors (not the coeefficients) in time series regressions when residuals can be heteroskedastic and serial correlation could be present. A potential problem with the Fama-Macbeth method, which is the next step, is autocorrelation. This is addressed. Newey-West gives robust t-stats for time-series averages such as the average price of risk per factor, which tells us whether on average, the market pays (positive λ) or penalizes (negative λ) exposure to a given factor and by how much per period. This creates an augmented FF5 model with an investable VOL factor and is evaluated based on the segmented sample periods, namely the entire sample, pre-pandemic, crash, rebound and post-pandemic.

The VOL price was a source of heterogeneity in showing how prices varied across sample periods, as it was negative on average throughout the entire sample except for the Covid-19 crash. As a result of the crash, the sign flips into a positive factor, consistent with the fact that low-volatility stocks briefly enjoyed a premium when the markets collapsed. While the SMB factor was generally weak, it became significantly negative during the pre-pandemic phase as conditions tightened. Overall, the six factor model pricing varied materially by phase. A negative λ^{VOL} means the cross section rewarded high volatility tilts (portfolios with beta VOL < 0), which is what occurred in the pre-pandemic and especially in the rebound when investor risk appetite was high. The temporary positive λ^{VOL} in the crash showed a flight to low-volatility exposure (beta VOL > 0 was favored) amidst uncertainty, the reversion in the post pandemic shows that investors again demanded higher expected returns to hold volatile stocks as policy support faded and macro risks continued in 2022. The phase by phase pricing pattern provides a mechanism consistent with market state-dependent risk appetite. The crash meant preference for stability (positive VOL), rebound showed risk apettite increase (negative VOL) and post-pandemic tightening meant a renewed aversion to volatility (negative VOL).

2.5 Literature contributions

The Covid-19 pandemic created an unprecedented stock market crash in early 2020, followed by a long recovery period. This volatile phase offers an opportunity to examine asset pricing anomalies present in the build-up, during and after the extreme conditions. The size effect means that smaller-capitalization stocks tend to outperform larger-capitalization stocks on a risk-adjusted basis (Banz, 1981). The volatility effect denotes the empirical finding that stocks with higher volatility often yield lower returns than less volatile stocks, contradicting the traditional high risk-high reward intuition (Ang et al., 2006).

Early studies by Banz, showed that smaller firms have higher-risk adjusted returns, on average, than larger firms. The size effect suggested that market capitalization contains information not captured by the Capital Asset Pricing Model (CAPM), since higher returns of small stocks couldn't be explained by market beta alone. Fama and French (1992) researched further into the size effect's importance and they evidenced that two simple variables, namely, Market Capitalization and Book-to-market equity, work together in determining the cross-sectional fluctuation in average stock returns in the United States. This showed smaller firms consistently earned higher average returns than larger firms, even after controlling for market risk, and this size premium, along with value premium, helped explain patterns in returns that CAPM could not explain. The size factor was formalized as SMB (Small Minus Big) in the Fama French three-factor model. With time, the size effect has raised questions because studies observed that the size premium weakened or disappeared in U.S markets after the 1980's, raising questions about its robustness. Asness, et al. (2018) revisit the size anomaly and argue that inconsistent performance of the size premium over time is largely attributable to the influence of "junky" small firms with poor quality characteristics. After controlling for firm quality, the size premium emerges that is much stronger and more stable over time, while the return differential between small and large stocks becomes more pronounced and persistent when quality factors are taken into account (profitability, leverage, earnings stability). Their findings discredit previous arguments against the size effect by demonstrating that many of them are the result of the volatile performance of small low-quality companies. The authors revive the size anomaly and establish it as a central empirical study with consistent findings, in addition to values and momentum, in terms of robustness. Size premiums are also discussed in relation to whether they are compensation for risk or a reflection of mispricing. There is evidence that small-cap stocks tend to provide higher returns on average than large-cap stocks, though this effect varies over time and is dependent on the fundamentals of the firm and the market environment.

In classical finance theory, higher volatility is rewarded with higher returns as a result of the risk-return tradeoff. Although this holds true under CAPM, empirical research revealed a puzzling volatility effect in the cross-section of stocks. According to Ang et al. (2006), stocks with high idiosyncratic volatility earned abnormally low returns on average. Even after controlling for size, value, and momentum effects, the highest volatility stocks underperformed the lowest volatility stocks. As a result of this counter-intuitive negative relationship, known as the idiosyncratic volatility puzzle, the notion that risk is uniformly rewarded in the market was challenged. It was confirmed by the authors that their results were robust using U.S. data, and that they were further extended internationally, strengthening the research on this anomaly. It has been shown by Frazzini and Pedersen (2014) that a strategy of betting against beta (long-low beta stocks and short high beta stocks with leverage to equalize risk) produces significant positive abnormal returns, indicating that systematic volatility is not fully compensated. Together, these results suggest that both total volatility and idiosyncratic volatility can predict lower future returns, a result that is inconsistent with traditional models and indicative of market frictions or behavioral biases. Several studies have suggested that the negative relationship between volatility and returns is not causal, but rather driven by a third factor – investors' preference for lottery-like payouts. Stocks with high idiosyncratic volatility tend to have skewed return distributions, with a low probability of experiencing extreme positive returns. According to Bali, Cakici and Whitelaw (2011), the earlier puzzle largely disappears when we take into account stocks with extremely high daily returns (max effect). They discover that there is no idiosyncratic volatility puzzle as reported by Ang et al. (2006). In a world in which poorly diversified and risk-averse investors determine stock prices, stocks with high idiosyncratic volatility are likely to generate higher future returns. Based on their analysis, Ang et al.'s results were driven by lottery-like stocks, with occasional huge upsides that investors overpaid for. When these effects are taken into account, higher volatility again corresponds to higher expected returns, as predicted by risk-based theory. Stambaugh, Yu and Yuan (2015) provide a different perspective, attributing volatility anomalies to arbitrage and mispricing limits. The study decomposes stocks based on compositive misvaluation measures into overpriced and underpriced stocks. They discover that idiosyncratic volatility is assymetric, with a strongly negative value among overpriced stocks (high IVOL stocks which are likely overvalued tend to underperform as arbitrageurs avoid shorting these risky stocks), but positive among "underpriced" stocks, where higher volatility actually predicts higher returns, as would be expected in an efficient market. The result indicates that short-sale constraints and arbitrage risks are preventing correcting the overpricing of high-volatility stocks, resulting in a deviation from the average effect (many high-volatility stocks are overpriced and earn low returns). For some, the volatility effect is explained by lottery-seeking behavior, while for others, it is explained by market frictions and arbitrage asymmetries.

Asset pricing effects don't remain constant across market cycles. In economic recessions and financial crises, out-of-sample tests can be conducted to determine whether return predictors are compensating for risk or exhibiting anomalies that could change under stress scenarios. According to Perez-Quiros and Timmerman (2000), small firms' risk and returns vary asymmetrically during economic cycles. Recessions cause small businesses' net worth to decrease, and increased credit markets will have a greater adverse effect on them than expansions. As small businesses are less likely to have collateral and external financing capacity, downturns will impact them more severely than large corporations with more stable financing. As a result, investors avoid high-risk small firms during crises and switch to larger, safer firms. When the market experiences a large contraction, the usual size premium may diminish or reverse in the short term. During the financial crisis of 2008, small cap stocks suffered greater drawdowns as a result of liquidity difficulties and risk aversion spikes. By comparison, large cap stocks (which have strong financial statements) served as relative safe havens during that time. Once the recovery period sets in, small caps can rebound significantly, sometimes outperforming as conditions normalize and risk apetite returns. This pro-cyclical behaviour of the size effect (weak in recessions, strong in recoveries) states that size-related returns are at least partly compensation for economic risk exposure. It also suggests that any analysis of the size effect should account for the market state, a factor or strategy that favors small firms may underperform in downturns and outperform over longer horizons that include recoveries.

The volatility effect is influenced by market cycles and the tendency for low volatility stocks to outperform can seem puzzling, but in turbulent periods, this anomaly can be related to investors "flight to safety." When market-wide volatility surges, high volatility stocks typically experience the sharpest drops, reflecting their higher risk and often lower quality. A portfolio with a lower volatility performs better on a relative basis, stabilizing the portfolio. Market indices were led downward by sectors and firms with uncertain futures during the COVID-19 period. Frazzini and Pedersen (2014) discovered these patterns when betting against beta factor results – in bad times, leverage on low-risk stocks is more effective than holding high-beta stocks. Additionally, extremely volatile events can exacerbate mispricing mechanisms. Overall, crises tend to accentuate investor focus on quality and safety while temporarily suppressing the usual high risk/high reward tradeoff and allowing anomalies such as size and volatility to act differently during different periods of crisis and recovery.

2.6 Structure of the Paper

This paper proceeds by delving into the data section where the retrieval procedures are described, cleaning of the data, the sample periods, description variables, VIX index graph and summary statistics for a visualisation of the sample data. The methodology section highlights the theoretical model to test the hypotheses and includes the expectations of the results. The portfolio formation is discussed showing how the size and volatility splits are performed to construct each portfolio. Additionally, the portfolio return calculations are described and excess returns alongside the VOL factor build. The augmented Fama-French five factor model and Fama-Macbeth two-pass methods are explained and the setup of each regression to estimate monthly prices of risk alphas for each factor. The robustness checks are detailed to address micro-cap noise, idiosyncratic volatility, momentum, reversal and skewness in the portfolio returns analysis. Additionally, econometric techniques are explained to test these hypotheses and to address issues such as heteroskedasticity and autocorrelation in the regressions. The results section analyses the average returns and t-stats for VW and EW portfolios for the entire sample and each sub-sample period along with the VOL factor and economic interpretation is detailed within the analysis. The regressions output for the two-pass Fama-Macbeth method is presented alongside Newey-West t-stats, R^2 and Adjusted R^2. The robustness examination discusses variations in the results of the model when addressing the earlier mentioned checks. The discussion explains the economic intuition and implications of the findings, compares these results to previous studies and critically assesses weakness of the approaches, alternative explantions and endogeneity problems. Lastly, the summary of the main findings is discussed while reiterating the paper’s contributions to the financial markets and asset pricing fields.

3. Data

This study focuses on publicly traded companies in the United States of America (USA) from 01.01.2018 until 31.12.2022. The research is further divided into sub samples: Pre-Pandemic (Jan 2018-Dec 2019), Crash (Feb 2020-Mar 2020), Rebound (Apr 2020-Dec 2021) and Post-Pandemic (Jan 2022-Dec 2022). The data gathered involves monthly data obtained from CRSP which includes: Share code, Exchange code, Price or Bid Ask Average (PRC), Returns, Shares Outstanding (SHROUT), Cumulative Factor to adjust Prices (CFA), NAICS and PERMNO. The approach to ensure reliable results includes keeping only common shares (Share codes 10 and 11), focusing on the NYSE and NASDAQ Exchanges (Exchange codes 1, 3) and removing companies within the Finance and Insurance industry (NAICS code starting in 52), which operate in fundamentally different ways compared to non-financial firms.[1] Additionally, any observation that presented an "NA" value, that observation was removed from the dataset. Price or Bid Ask Average and Shares outstanding cannot be negative values, so these were removed from the dataset if they presented negative values. PERMNO assigns a permanent identification number to all companies on the CRSP dataset, regardless of if they change their company name over the course of the sample period. It is important to track the number of entries each common stock has after data cleaning. Market Capitalization (Market Cap) was then calculated, as its not given by CRSP, the calculation used was:

$$\text{Market Cap} = PRC_t * SHROUT_t * 1000.$$

Additionally, the Fama-French 5 Factor monthly dataset is obtained from the Dartmouth Kenneth R. French website for all factors: Market Risk Premium (Mkt-Rf), Size factor (SMB), Value factor (HML), Profitability factor (RMW), Investment factor (CMA) and Risk-free rate (Rf). Consistent with literature, size breakpoints are computed each month using only NYSE firms, and then applied to the NASDAQ exchange. Specifically, we take the NYSE median market capitalization to split Small and Large. Within each size, we form volatility quintiles from vol-1 to vol-5, producing 10 Size x Vol portfolios per month (Large, Small and five quintiles per size). Value weighted (VW) portfolios are the stock's market cap divided by the portfolio's total market cap that month.

---

[1] Financial and Insurance companies have unique financial structures, generate revenue by fees, trading activities and investment returns, involve extensive regulation requirements such as the capital adequacy ratio and their assets and liabilities are mainly focused on financial instruments.

The variables used to measure the effects of company size and volatility on stock returns consist of portfolio excess returns as the dependent variable (Y) and the independent variables (X) include the constructed factor, the VOL factor and the 5 Fama-French Factors. For the dependent variable, we are using excess returns to measure actual gains above the risk free rate investment and focusing on abnormal returns. The independent variable VOL quantifies if investors receive higher returns or lower returns for holding high volatility stocks in comparison to low volatility stocks. The return difference measures the additional compensation an investor requires for the increased risk involved.

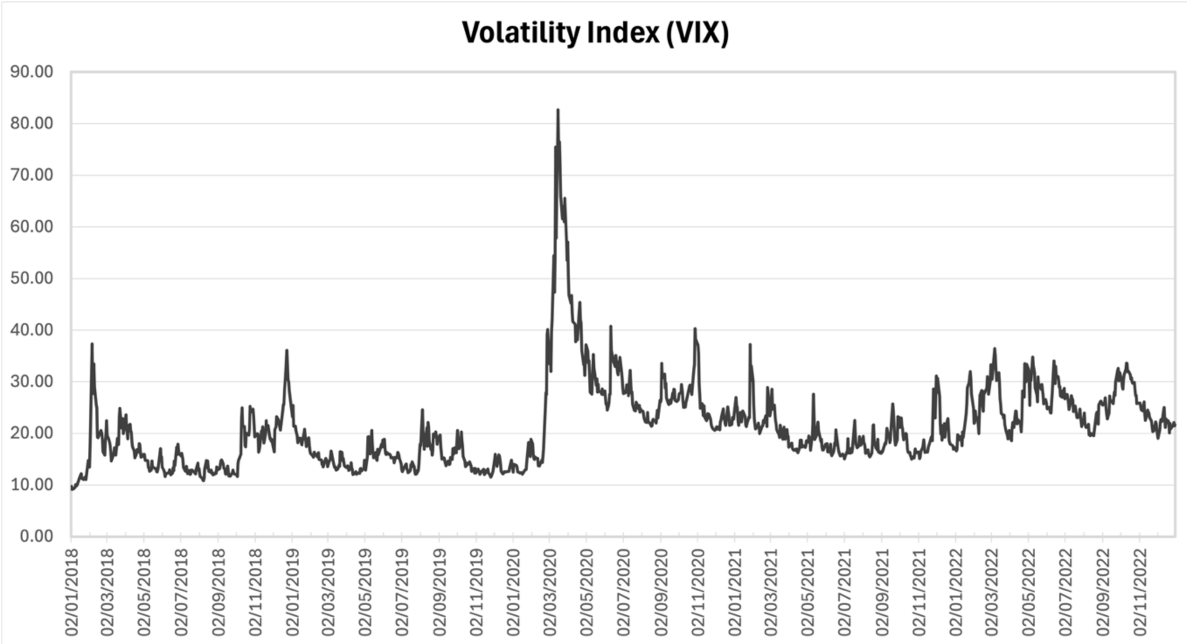
Table 1: Entire Sample and Sub-Sample Periods summary statistics

Sample Period EW	Stocks	Min	Mean	Max	Std Dev
Entire Sample	4113	-22.43%	0.62%	21.75%	7.74%
Pre-Pandemic (Jan 2018–Dec 2019)	3178	-13.47%	0.35%	14.42%	5.98%
Crash (Jan–Mar 2020)	2829	-22.43%	-10.38%	-1.25%	10.89%
Rebound (Apr 2020–Dec 2021)	3510	-6.06%	4.44%	21.75%	7.51%
Post-Pandemic (Jan–Dec 2022)	3395	-12.11%	-2.79%	9.17%	6.98%

Sample Period VW	Stocks	Min	Mean	Max	Std Dev
Entire Sample	4113	-10.44%	1.61%	15.06%	5.61%
Pre-Pandemic (Jan 2018–Dec 2019)	3178	-8.58%	1.55%	9.14%	4.38%
Crash (Jan–Mar 2020)	2829	-10.44%	-5.38%	1.31%	6.04%
Rebound (Apr 2020–Dec 2021)	3510	-4.28%	4.14%	15.06%	4.81%
Post-Pandemic (Jan–Dec 2022)	3395	-8.95%	-0.96%	10.90%	6.87%

Table 1 includes the entire sample of stocks included in the NYSE and NASDAQ stock exchanges in the United States, which includes 4,113 stocks from 01.01.2018 to 31.12.2022. This is then divided into sub-sample periods: Pre-Pandemic (Jan 2018-Dec 2019), Crash (Jan 2020-Mar 2020), Rebound (Apr 2020-Dec 2021) and Post-Pandemic (Jan 2022-Dec 2022). The entire sample equal-weighted portfolio presents a minimum return of -22.43% and a maximum of 21.75% while the value-weighted portfolio presents a minimum return of -10.44% and a maximum of 15.06%. The EW portfolio has underperformed the VW over the entire sample with an average return of 0.62% and 1.61%, respectively. The VW portfolio also outperforms the EW in every sub-sample period, except for the rebound phase. The time period which suffered the most fluctuations in the business cycle was from 01.02.2020 until 01.04.2020, labelled as the US recession period (FRED, *Federal Reserve Economic Data*).

Graph 1: Daily close of Volatility Index from January 1st, 2018 until 31st of December, 2022



The VIX at the beginning of the period exhibits values at 10 points fluctuating at low levels with two peaks at approximately 38 points before reaching the Covid recession period. The Covid recession is pronounced on the 1st of February 2020 at 17.97 and reaches its peak on the 16th of March 2020 at 82.69. According to St. Louis FED, the end of the recession period is the 1st of April 2020 where the VIX was still at 57.06 but it was falling towards level of normalcy. After the recession period, the VIX never returned to previous levels as it fluctuated between 20 and 30 points over time and reached minimums of approximately 15 points, slightly above the lowest levels experienced in 2018.

4. Methodology

After cleaning the data, the 3 month rolling window volatilities were calculated for each row. To start the data in January 2018 (first date of the entire sample), we need to include rows starting in November 2017 for 3 months rolling window and the same for 6 and 12 months (February 2017 for the robustness check of 12 month rolling window). We estimate past return volatility as rolling standard deviation of monthly returns over three look-back horizons: 3,6 and 12 months. For stock I in month t and window m:

$$\sigma_{it}^{(m)} = \sigma\,(R_{i,t-m+1},\ldots,R_{it})$$

At the end of each month, we perform the following:
1) Size splits using the NYSE median market cap breakpoints to classify stocks into Small and Large.

2) Volatility quintiles within size (vol-1 to vol-5) using cross-sectional quintiles of $\sigma_{it}^{(m)}$.

3) Portfolio returns for each of the ten portfolios:

$$R_{pt}^{EW} = \frac{1}{N_{pt}} \sum_{i \in p} R_{it}$$

$$R_{pt}^{VW} = \sum_{i \in p} w_{it} R_{it}, \quad w_{it} = \frac{MC_{it}}{\sum_{j \in p} MC_{jt}}$$

4) Excess returns: Subtract RF

$$R_{pt}^{EW} - \mathrm{R}F_t \quad \text{and} \quad R_{pt}^{VW} - \mathrm{R}F_t$$

Afterwards, within each size and month we form the VOL factor as low-vol minus high-vol:

$$\mathrm{VOL}_t^{EW} = R_{Vol1,t}^{EW} - R_{Vol5,t}^{EW}, \quad VOL_t^{VW} = R_{Vol1,t}^{VW} - R_{Vol5,t}^{VW}$$

Then we get the equal-weight average across sizes to not contaminate VOL with the size effect:

$$\mathrm{VOL}_t = \frac{1}{2}(VOL_t^S + VOL_t^L)$$

We then use the monthly FF5 factors from 2013 until 2022 so that the first-pass rolling regressions in January 2018 have 60 months of history, but keeping the sample from 2018 to 2022. Then the two-pass Fama-Macbeth (monthly) is performed using the following method:

1. The first pass involving time series betas. For each test portfolio I, estimated rolling betas are calculated using information up until t-1:

$$r_{i,t} - r_{f,t} = \alpha_i + \beta_i^{\mathrm{MKT}}(MKT_t - RF_t) + \beta_i^{\mathrm{SMB}} SMB_t + \beta_i^{\mathrm{HML}} HML_t + \beta_i^{\mathrm{RMW}} RMW_t + \beta_i^{\mathrm{CMA}} CMA_t + \beta_i^{\mathrm{VOL}} VOL_t + \varepsilon_{i,t}$$

With a 60 month rolling window where possible, updated each month. The betas and idiosyncratic variance of the portfolio are stored, the latter of which gives efficiency weights for the second pass.

2. Second pass (cross-sectional prices of risk). For each month t we regress the cross-section of portfolio returns on their betas:

$$r_{i,t} - r_{f,t} = \lambda_t^{MKT}\widehat{\beta_{\iota,t}^{MKT}} + \lambda_t^{SMB}\widehat{\beta_{\iota,t}^{SMB}} + \lambda_t^{HML}\widehat{\beta_{\iota,t}^{HML}} + \lambda_t^{RMW}\widehat{\beta_{\iota,t}^{RMW}} + \lambda_t^{CMA}\widehat{\beta_{\iota,t}^{CMA}} + \lambda_t^{VOL}\widehat{\beta_{\iota,t}^{VOL}} + u_{i,t}$$

This cross-section is estimated using weighted least squares (WLS) with weights $w_{i,t} = \frac{1}{\widehat{\sigma_{\varepsilon,\iota,t}^2}}$, to address cross-sectional heteroskedasticity so that noisy portfolios receive smaller weight. A cross-sectional intercept is not present, which is standard when estimating prices of traded factors.

3) For each factor k, a time-series average price of risk and Newey-West t-stats were computed:

$$\overline{\lambda^k} = \frac{1}{T}\sum_{t=1}^{T}\lambda_t^k$$

$$t_{NW}\left(\overline{\lambda^k}\right) = \frac{\overline{\lambda^k}}{\widehat{sd}\left(\lambda_t^k\right)/\sqrt{T}}$$

A lag of 4 was used for the entire sample (2018-2022) and a lag of 2 was used within phases. A phase-specific mean was calculated by mapping each t to pre-pandemic, crash, rebound, and post-pandemic.

In summary, each month we are seeing how much the market requires per unit of exposure to each factor at that moment in time. The alpha t's represent the monthly prices. Taking the average of these factors, we are able to determine whether investors rewarded or punished exposure to Market, Size, Value, Profitability, Investment, and Volatility risk during 2018-2022.

As part of robustness checks, to reduce micro-cap noise, stocks below the 10th percentile of NYSE Market Cap are excluded from the data. All steps involving creating portfolios, calculating returns and VOL summaries are repeated the same way with the exclusion of the bottom decile of stocks within the portfolios (refer to appendix for results).

Another robustness check performed is the idiosyncratic volatility, where market-wide and firm specific risk are separated. Each month and for each stock I, residual (idiosyncratic) risk is estimated by fitting a rolling FF5 regression on excess returns over the past 12 months. IVOL is defined as: $\mathrm{IVOL}_{\mathrm{I,t}} = \sigma\left(\varepsilon_{i,t}\right)$, as the standard deviation of the residuals in the last 12 month window. Then the method for portfolio sorting is mirrored, split by small and large, rank stocks into IVOL quintiles (Vol1 to Vol 5) and compute EW and VW for the 10 portfolios. A traded, size neutral IVOL factor is created and summarized by phase with means and t-stats, just like total volatility. For completeness, IVOL was included but the focus was on total volatility (refer to appendix for IVOL results).

Moreover, to ensure volatility patterns aren't driven by recent performance, the VOL spread is adjusted using two predictive signals. These signals are momentum and short term reversal calculated for each stock individually and refreshed every month. The momentum is the six month cumulative return ending one month before t (skipping the most recent month):

$$\mathrm{MOM\_6}_{it} = \prod_{k=2}^{7}\left(1 + R_{i,t-k}\right) - 1$$

The short-term reversal equation is defined as:

$$\mathrm{REV1}_{\mathrm{it}} = R_{i,t-1}$$

For each month t, a rank of all stocks into terciles by momentum is done. Within each tercile, we re-form Vol-1 to Vol-5 and recompute VOL factor (Vol 1 – Vol 5) and summarize by each sub-sample period. For each month, the split is done of cross-section into terciles where 1 = weakest momentum (past losers) and 3 = strongest momentum (past winners). This allows for like-for-like results which can be directly compared with our original methodology.

Lastly, a test is performed whether the Low-High volatility premium depends on a stocks return skewness. For each stock I and month t, a 12 month rolling skewness of simple monthly returns is computed:

$$\mathrm{SK} = \frac{n}{(n-1)(n-2)}\sum_{j=1}^{n}\left(\frac{r_j - \bar{r}}{s}\right)^3, \quad n = 12$$

In this step, returns are used and no factor model is used, resulting in a time-varying, stock-level measure of asymmetry in recent returns. The cross section is sorted into skewness terciles (1 = lowest skew, 3 = highest positive skew) and the Vol-1 to Vol-5 is recalculated inside each tercile, resulting in the VOL factor (Vol1 – Vol 5) for each subsample. As a result, we are able to determine whether volatility effects are more pronounced in stocks with a high skewness than in stocks with a low skewness.

4.1 Hypotheses and expectations of results

Prior to conducting the regressions, hypotheses were developed regarding the roles of firm size and return volatility in expected stock returns. As a result of previous literature and economic intuition, size and volatility were expected to affect returns at different stages of the market cycle. It was expected that smaller firms would yield higher average returns on a risk-adjusted basis than larger firms due to the "size effect" in asset pricing since historically small-cap stocks have outperformed large-cap stocks. With respect to volatility, the hypothesis was that volatility is a factor associated with lower expected returns, rather than a source of reward. There was an expectation that the low-volatility anomaly would be strongly evident in the data, with stocks with higher volatility experiencing lower returns in the future. Even after controlling for the FF5 factors, volatility would still show a significant inverse relationship with returns.

4.2 Econometric techniques

As part of the statistical analysis used in the study, the estimation technique in the cross-sectional stage (Fama-Macbeth method) utilized weighted least squares rather than ordinary least squares in the regression. Weights were assigned to each stock based on the inverse of its first stage residual variance (the inverse of its idiosyncratic variance). This is due to the fact that stocks with noisy beta estimates (high residual variances) should be given less weight when determining factor premiums, while stocks with more precise betas (low idiosyncratic noises) should be given more weight. In this manner, errors-in-variables problems associated with estimated betas can be mitigated. By using WLS, the aim was to obtain more reliable and efficient estimates of monthly λ's, implementing a Fama-Macbeth procedure that accounts for heteroskedasticity in the cross-section. Since our sample consists of a wide range of firms and trading volumes, WLS naturally downweights the noisy residuals in the case of smaller firms. Newey-West standard errors were used in order to account for the possibility that alpha estimates may be serially correlated over time. As a result of the Newey-West adjustment, heteroskedasticity and autocorrelation consistent standard errors (HAC) were produced for each alpha value. This creates a t-statistic for each average factor premium.

5. Empirical Results

Table 2: Entire Sample Equal-Weighted and Value-Weighted 10 portfolios

Entire Sample

Portfolio	Average_EW	T_Stat_EW	Average_VW	T_Stat_VW
Large-Vol1	0.86%	2.45	1.07%	2.74
Large-Vol2	0.99%	1.85	1.47%	2.65
Large-Vol3	1.15%	1.55	1.40%	1.81
Large-Vol4	1.49%	1.49	1.54%	1.47
Large-Vol5	3.82%	2.43	4.07%	2.49
Small-Vol1	-1.28%	-2.87	0.03%	0.07
Small-Vol2	-1.51%	-2.26	0.15%	0.21
Small-Vol3	-1.63%	-1.68	0.10%	0.10
Small-Vol4	-0.78%	-0.58	1.37%	0.99
Small-Vol5	6.41%	2.76	9.04%	3.77

The table above presents the results of the entire sample portfolio, including the average returns for the EW and VW portfolios, as well as t-statistics. Results indicated that high-volatility portfolios outperformed low-volatility portfolios over the entire sample. An interesting pattern emerged, high-volatility premium was stronger in VW returns. These results indicate that volatile stocks were not only outperformed by small-cap stocks, but also by large-cap stocks. Small portfolios outperformed Large portfolios except for the Vol-5 quintile, where the Small portfolio (9.04%) performed better than the Large (4.07%). For VW's large-caps, the quintile with the highest volatility (Large-vol5) had the highest returns around 4.07% per month, compared to 1.07% for the quintile with the lowest volatility (Large-vol1). In general, the effect was more pronounced among smaller stocks, as the Small-vol5 portfolio returned on average 9.04%, while the Small-vol1 portfolio returned 0.03%.

Table 3: VOL factor in different sub-sample periods

Phase	mean_VOL_EW	t_VOL_EW	mean_VOL_VW	t_VOL_VW	Rolling Window Vol
Crash	8.41%	1.69	10.15%	2.08	3 months
Post Pandemic	-3.93%	-1.67	-4.57%	-1.72	3 months
Pre-Pandemic	-3.24%	-2.19	-4.63%	-2.84	3 months
Rebound	-10.48%	-4.22	-10.69%	-4.29	3 months

Table 3 shows the average monthly VOL factor derived by subtracting Vol-1 from Vol-5 for both Small and Large portfolios, and then calculating the average of these two factors. The VOL factor was slightly more statistically significant in VW terms during every sub-sample period. In the pre-pandemic period, the VOL factor was negative and statistically significant, which means low-volatility stocks underperformed high-volatility stocks by 4.6% per month on average. The stronger VOL factor for VW compared to EW, indicates that the pre-pandemic outperformance of high-volatility stocks was especially pronounced among larger firms in the tails. There was no micro cap issue with the high-volatility premium before the Pandemic, but rather there was a stronger effect when large companies carried more weight. During the pandemic crash, VW's VOL increased to 10.15% (t = 2.08), suggesting a flight to quality into more robust and low-volatility firms. During the rebound, the negative premium grows (-10.69 VW) compared to pre-pandemic levels, aligning with a risk-on rally following the crash which led to volatile names becoming more dominant. After the pandemic, the VOL premium remains negative (-4.57% VW), but it is not statistically significant.

Table 4: Pre-Pandemic Equal-Weighted and Value-Weighted 10 portfolios

Pre-Pandemic

Portfolio	Average_EW	T_Stat_EW	Average_VW	T_Stat_VW
Large-Vol1	1.17%	2.42	1.44%	3.07
Large-Vol2	1.23%	1.91	1.62%	2.37
Large-Vol3	0.94%	0.98	1.19%	1.21
Large-Vol4	1.36%	1.11	1.41%	1.14
Large-Vol5	2.74%	1.38	3.16%	1.55
Small-Vol1	-0.94%	-1.68	0.24%	0.40
Small-Vol2	-1.13%	-1.37	0.26%	0.29
Small-Vol3	-1.27%	-1.1	0.23%	0.19
Small-Vol4	-1.21%	-0.78	1.00%	0.58
Small-Vol5	3.97%	1.43	7.79%	2.61

In the pre-pandemic period, there was a slight positive relationship between volatility and returns, especially among small-cap stocks. There is no steady increase in returns over time as the volatility quintile increases in both portfolios. For VW, t-statistics are significant only for the Large Vol-1, Large Vol-2, and Small-Vol 5 portfolios, the rest are not significant. In terms of returns, the highest volatility quintile (Vol-5) earned 7.79% average returns with strong significance, while the lowest volatility quintile (Vol-1) earned 0.24% with low significance. The Large Vol-5 portfolio earned 3.16% returns, while the Large Vol-1 portfolio earned 1.44% returns. EW returns are slightly lower in all Large quintiles, with Large-Vol3 displaying the lowest average returns (0.94%) and being non-significant. Besides this quintile (Vol-3), the returns increase progressively as the quintile increases. For Small portfolios, the returns become more negative except for Vol-4 and Vol-5, which show better returns than the previous quintiles, while the only significant result is the Vol-5. During the period prior to the pandemic, volatility was present, but modest. Small speculative stocks outperformed, consistent with lottery-like stocks, which contributed to the higher than normal performance of the Small portfolios. The higher-than-normal performance by the highest volatility group in large and small stocks can be attributed to highly speculative attention-grabbing stock investments by investors with a purpose of quick returns who cause prices to move (Barber et al., 2022).

Table 5: Crash Period Equal-Weighted and Value-Weighted 10 portfolios

Crash

Portfolio	Average_EW	T_Stat_EW	Average_VW	T_Stat_VW
Large-Vol1	-2.29%	-0.89	-1.65%	-0.59
Large-Vol2	-5.20%	-1.77	-3.43%	-1.25
Large-Vol3	-7.39%	-1.99	-5.99%	-1.74
Large-Vol4	-10.86%	-2.16	-10.79%	-2.75
Large-Vol5	-13.51%	-1.49	-10.82%	-1.22
Small-Vol1	-5.67%	-1.68	-5.11%	-1.56
Small-Vol2	-9.70%	-2.23	-9.18%	-2.29
Small-Vol3	-13.93%	-2.05	-14.53%	-2.14
Small-Vol4	-17.12%	-1.83	-15.98%	-2.00
Small-Vol5	-11.29%	-0.96	-16.24%	-1.43

During the crash period, the volatility premium switches sign and becomes positive (VW mean 10.15% and EW mean 8.41%) per month in this period, with robust results for VW (t=2.08) and not significant for EW (t=1.69). The long-low/short-high spread loses less in this period because high-volatility stocks collapse more than low-volatility stocks. The average returns had a wide gap between the lowest and highest returns. For Large portfolios in VW, they fell from -1.65% (Vol-1) to -10.82% (Vol-5) and Small portfolios from -5.11% to -16.24%. EW portfolios saw average returns fall from -2.29% to -13.51% for Large, and for Small, it fell from -5.67% to -17.12%, with an outlier at Vol-4 (average return of -17.12%) being more negative than at Vol-5 (average return of -11.29%). Surprisingly, the EW portfolios for small-vol3 and small-vol5 earn better returns than their VW counterparts. These outliers indicate that within the small-cap high volatility portfolios vol-3 and vol-5, the relatively larger stocks fell more than the small ones. The only significant results for VW in the Large portfolios are Large-vol4, and for Small, it consists of Small-Vol2, Small-Vol3 and Small-Vol4. This indicates that the average returns that are non-significant cannot be concluded as there isn't enough evidence. All portfolios perform negatively, but the size effect is present, where large companies suffer less negative returns during the crash, providing a buffer during the market drawdown. This result is consistent with the flight-to-safety in both EW and VW as investors move to safer stocks during a crisis.

Table 6: Rebound Period Equal-Weighted and Value-Weighted 10 portfolios

Rebound

Portfolio	Average_EW	T_Stat_EW	Average_VW	T_Stat_VW
Large-Vol1	2.18%	4.94	2.46%	3.99
Large-Vol2	2.75%	3.78	3.25%	4.15
Large-Vol3	3.75%	3.57	4.06%	3.40
Large-Vol4	4.68%	3.27	4.79%	3.04
Large-Vol5	9.10%	3.5	9.61%	3.53
Small-Vol1	0.51%	0.76	1.76%	2.76
Small-Vol2	1.16%	1.18	2.80%	2.88
Small-Vol3	1.99%	1.38	3.64%	2.47
Small-Vol4	4.21%	1.93	6.02%	2.86
Small-Vol5	14.55%	3.14	15.98%	3.48

In the rebound phase, the VOL factor changed signs (-10.48% EW and -10.69% VW) with robust results (t = -4.22 and -4.29), consistent with a back to normality sign flip as all the sub-sample periods besides the crash period have negative VOL factors. The post-crisis period favored volatility as a factor. This reflects that high volatility outperformed low volatility portfolios due to monetary stimulus and a resurgence in risk appetite favoring smaller, more volatile stocks. For VW, large portfolios experienced 2.46% returns (Large-Vol1) and 9.61% (Large-Vol5) while small stocks returned 1.76% (Small-vol1) and 15.98% (Small-vol5). EW portfolios slightly underperformed VW portfolios and the returns ranged from 2.18% (Large-vol1) to 9.10% (Large-vol5) while small portfolios experienced 0.51% to 14.55% (Small-vol5). The VW portfolios outperform all the EW counterparts during this sub-period. All the portfolio in VW shows significant results while the EW portfolios show significant results for all large portfolios and only the small-vol5 for small portfolios. Stocks with high volatility were rewarded as investors piled into "lottery-like" stocks that crashed during COVID-19, many of which reside in the small cap quintile with the highest volatility (Small-vol5) as retail trading soared, amplifying volatility stocks' recovery. The rebound showed how quickly the traditional volatility-return relationship returns to normal after a market downturn.

Table 7: Post-Pandemic Period Equal-Weighted and Value-Weighted 10 portfolios

Post-Pandemic

Portfolio	Average_EW	T_Stat_EW	Average_VW	T_Stat_VW
Large-Vol1	-1.27%	-1.8	-1.43%	-1.98
Large-Vol2	-1.03%	-0.71	-0.69%	-0.43
Large-Vol3	-0.85%	-0.44	-1.00%	-0.50
Large-Vol4	-0.73%	-0.27	-0.80%	-0.27
Large-Vol5	1.09%	0.31	-0.08%	-0.02
Small-Vol1	-4.02%	-6.04	-2.11%	-2.80
Small-Vol2	-4.90%	-4.1	-2.38%	-1.60
Small-Vol3	-5.60%	-2.79	-2.71%	-1.19
Small-Vol4	-4.58%	-1.66	-1.70%	-0.53
Small-Vol5	1.47%	0.35	5.68%	1.26

In the post-pandemic period, as markets experienced inflation and tighter government policy, the volatility effect remained negative, shown in the VOL factor (-3.93% EW and -4.57% VW) with non-significant t-stats (t=-1.67 and -1.72). Unlike the crash period, low volatility stocks didn't outperform high volatility stocks, they continued to underperform. This period presents counterintuitive outcomes in both EW and VW portfolios. Additionally, there are many portfolios which show non-significant results, showing that no conclusions can be made about the results. The only significant results in VW are Large-vol1 (t-stat = -1.98) and Small-vol1 (-2.80). In EW, Small-vol1, Small-vol2 and Small-vol3. By the end of 2022, the low-volatility anomaly had not re-established itself, providing insights into the slow pace of the anomaly's return to markets following a crisis period. The continued underperformance of low-volatility stocks in the post-pandemic period reinforces that during 2018 through to 2022, low-volatility strategies consistently underachieved, except during market turmoil like COVID-19. Stocks with a high level of volatility held up relatively well in 2022, which implies that the market period shift towards favoring volatility has continued beyond the period of rebound.

T*able 8: Full Sample Average Factor Prices of Risk (Fama–MacBeth Estimates)*

Factor	Mean	T_NW	N
λ MKT	0.86%	1.82	94
λ SMB	-0.02%	-0.06	94
λ HML	-3.34%	-1.81	94
λ RMW	-1.25%	-3.17	94
λ CMA	-0.67%	-0.69	94
λ VOL	-5.09%	-5.10	94

The estimation of monthly prices of risk for the augmented Fama-French five factor model and the VOL factor using a two-pass Fama-Macbeth procedure, resulted in a price of volatility risk that is negative and economically large (-5.09%) per month with significant results. To interpret this, each 0.10 unit increase in exposure to VOL beta (beta to Vol 1 – Vol 5) is associated with -0.509% lower monthly excess return (0.10 * -5.09% = -0.509%), holding other betas fixed. Amongst the FF5 factors, the RMW factor is the only significant factor at -1.25% (t= -3.17). The full sample evidence points to a pronounced reward for high-volatility exposure and negative pricing of profitability. The factors and portfolios are built back to 2013, although those pre-2018 observations are used only to estimate rolling betas in the first pass. The second pass Fama-Macbeth estimations, averages and t-stats begin in 2018 and finish in 2022. Betas for the second pass FM regressions are estimated with a rolling 60-month window beginning in 2013, so N = 94 monthly cross section observations bring strong inference to the results. This setup uses lots of information for betas while keeping the evaluation horizon fixed to 2018-2022.

Table 9: Average Monthly Prices of Risk by Phase (Fama-Macbeth)

FACTOR	CRASH (JAN–MAR 2020)	POST (JAN–DEC 2022)	PRE (JAN 2018–DEC 2019)	REBOUND (APR 2020–DEC 2021)
MKT	-8.24%*** (-3.07)	-1.35% (-1.04)	+0.24% (0.33)	+3.76%*** (3.85)
SMB	-2.99% (-0.77)	-0.97%** (-2.19)	+0.56% (0.87)	+0.53% (0.52)
HML	-18.14%** (-2.33)	+7.97% (1.50)	-8.44%** (-2.18)	-5.21%*** (-3.45)
RMW	-1.35% (-0.57)	+0.69% (0.57)	-0.77% (-1.37)	-2.06%** (-2.27)
CMA	- 7.87%* (-1.85)	+5.40% (1.46)	-2.42% (-1.26)	-2.01% (-1.55)
VOL	+10.03% (1.22)	-4.63%** (-2.05)	-4.61%*** (-2.66)	-10.57%*** (-3.08)
N (MONTHS)	3	12	24	21
R^2	0.985	0.930	0.935	0.943
ADJ. R^2	0.962	0.825	0.838	0.858

Entries are average monthly prices of risk in %, with Newey–West t-statistics in parentheses. Significance: * $p<0.10$, ** $p<0.05$, *** $p<0.01$. N is the number of monthly cross-sections. R^2 and Adjusted R^2 refer to the cross-sectional second-pass fit in each sub-period.

In the pre-pandemic period, the volatility risk premium (λ^{VOL}) equalled -4.61% and was statistically significant (t = -2.66), meaning that high-volatility exposure was compensated relative to low-volatility, as investors were rewarded for taking on risk. Thus, before the shock, portfolios with higher exposure to the high volatility leg ($\beta_{VOL}< 0$) gave higher expected returns, while low volatility tilts were priced to deliver lower returns. The negative price of HML (-8.44%) with significant t-stats (-2.18) indicates that growth stocks (low book-to-market) outperformed value stocks (high book-to-market) during this period.

There is a positive turn in the volatility risk premium during the crash period (10.03%) with a non-significant t-statistic (1.22). There are only three monthly observations for the crash period (Jan-March 2020), so the inference is noisy, and the results cannot be concluded; however, the sign change is economically relevant to the investigation. Investors benefit from low-volatility exposure during the sell-off during the crash as it reduces their risk exposure. In terms of direction of sign, the results of the portfolio returns and the VOL factor are consistent with the sign flip across the crash period. Market price (MKT) of risk is sharply negative at -8.24% (t = -3.08), as is the HML factor at -18.24% (t = -2.33). Although the t-stats are significant for MKT and HML, they aren't reliable and should be treated as descriptive and not conclusive due to the short time series of λ.

With respect to the rebound period, the volatility risk premium flips and becomes significantly negative at -10.57% (t = -3.18) which indicates that portfolios with greater exposure to high volatility were rewarded, reflecting risk-taking behavior. The findings of this study are consistent with those of the portfolio returns analysis. Growth stocks dominated factor preferences as the Value and Profitability factors indicated that the cross-section discounted "quality" and paid up for growth stock winners during the rebound. It appears that the market price of risk displays significant results at 3.76% with a strong significance coefficient (t=3.85), which is consistent with the broad appetite for risk among investors and the increasing confidence in the stock market. The value factor (HML) remains significantly negative at -5.21% (t = -2.27), suggesting that the high-profitability side was penalized during the rebound. As the estimates are based on 21 monthly cross-sectional results, the t-statistics should be interpreted as indicative of the conclusions derived from the data.

In the post-pandemic phase, the volatility risk premium remains negative at -4.63% (t = -2.05) but smaller than during the rebound. Economically, a 0.10 unit increase in β_{VOL} led to a 0.46% lower expected return per month (-4.63% * 0.10). Despite macroeconomic conditions normalizing and interest rates rising, investors continued to penalize high volatility exposure and favored low-volatility characteristics in stocks. A negative SMB factor (-0.97%) and statistically significant (t = -2.19) are indicative of small-cap underperformance in 2022. This indicates a premium for large-cap exposure as liquidity tightened. A resurgence of size aversion and a persistent discount on volatility is observed in 2022, consistent with the portfolio returns analysis, where Small portfolios with high volatility underperformed, with the exception of Small-Vol5.

Across the different sub-sample periods, the adjusted R^2 indicates a range between 82.5% and 85.8%, excluding the crash period due to limited statistical power of 3 months of data. This period R^2 and adjusted R^2 should be treated with caution as results are not definitive but suggestive.

Table 10: Momentum results phase-by-phase

Phase	Tercile	mean_VOL_EW	t_VOL_EW	mean_VOL_VW	t_VOL_VW	Signal
Crash	1	0.77%	0.12	8.20%	1.56	MOM6_1
Crash	2	13.48%	1.45	15.48%	1.98	MOM6_1
Crash	3	13.27%	1.64	7.67%	1.22	MOM6_1
Post Pandemic	1	-12.57%	-3	-6.47%	-1.84	MOM6_1
Post Pandemic	2	-3.27%	-0.97	-0.89%	-0.27	MOM6_1
Post Pandemic	3	-0.54%	-0.17	-1.21%	-0.34	MOM6_1
Pre Pandemic	1	-7.86%	-2.48	-4.99%	-1.76	MOM6_1
Pre Pandemic	2	-2.15%	-0.77	-1.94%	-0.79	MOM6_1
Pre Pandemic	3	-0.69%	-0.28	-0.46%	-0.19	MOM6_1
Rebound Apr 20-Dec 21	1	-15.92%	-3.58	-8.49%	-2.98	MOM6_1
Rebound Apr 20-Dec 21	2	-10.98%	-3.48	-6.96%	-2.76	MOM6_1
Rebound Apr 20-Dec 21	3	-9.97%	-2.94	-7.69%	-2.58	MOM6_1

Table 11: Reversal results phase-by-phase

Phase	Tercile	mean_VOL_EW	t_VOL_EW	mean_VOL_VW	t_VOL_VW	Signal
Crash	1	-2.11%	-0.21	2.98%	0.36	REV1
Crash	2	8.77%	0.89	14.77%	2.31	REV1
Crash	3	14.35%	1.86	9.58%	2.13	REV1
Post Pandemic	1	-21.04%	-4.52	-14.41%	-2.96	REV1
Post Pandemic	2	-7.94%	-1.68	-5.52%	-1.44	REV1
Post Pandemic	3	8.58%	2.87	5.23%	1.43	REV1
Pre Pandemic	1	-13.95%	-5.28	-11.13%	-5.69	REV1
Pre Pandemic	2	-4.59%	-1.53	-3.83%	-1.53	REV1
Pre Pandemic	3	5.21%	2.36	3.96%	1.87	REV1
Rebound Apr 20-Dec 21	1	-26.91%	-5.63	-17.28%	-5.2	REV1
Rebound Apr 20-Dec 21	2	-15.87%	-4.12	-10.19%	-3.68	REV1
Rebound Apr 20-Dec 21	3	1.16%	0.26	-1.68%	-0.41	REV1

As part of robustness checks, the tables above show that volatility-return relationship was tested to see if the results could be explained by other known factors, such as momentum or short-term reversal. There is a possibility that, among others, the portfolios with high volatility in the rebound were disproportionately composed of recent losers who rebounded (reversal effect). To test this possibility, the data was segmented by prior 6 month momentum and 1 month reversal signals and the volatility effect was examined within those subgroups. Based on both momentum and reversal, it is observed that recent losers exhibit the strongest high volatility premium during the pre-pandemic phase. There is a focus on the 3-month rolling window and the 6-month and 12-month checks (refer to appendix). The momentum-conditioned

VOL is sharply negative in tercile 1 (prior losers) at -7.86% EW (t-stat = -2.48) and -4.99% VW (t = -1.76) but closer to zero for tercile 2 and tercile 3, although the results are inconclusive for VW due to the strength of the t-stats. The reversal results are more pronounced as tercile 1 VOL is -13.95% EW (t-stat = -5.28) and -11.13% VW (t-stat = -5.69), modestly negative in tercile 2 (-2.15% and -1.94%) with non-significant results, and positive for tercile 3, which is considered recent winners, indicating that short-run winners did not contribute to premiums. The size and significance of the VW magnitudes in tercile 1 indicate that the effect is not limited to microcaps; larger high volatility losers also mean-revert.

As a result of the crash, both momentum and reversal flipped signs, consistent with a flight-to-safety/liquidity spiral rather than momentum or reversal. With momentum, VOL turns positive in every tercile except for tercile 2 VW (15.48 % and t = 1.98), and reversal VOL is negative for tercile 1 and positive for terciles 2 and 3. Significant results can be found in VW terciles 2 and 3. Based on the results, low-volatility outperformed high-volatility for losers (tercile 1) and winners (tercile 3). In terms of reversal, the low-volatility advantage also appears regardless of whether the stock has just won or lost in the preceding month (statistically significant for both T2 and T3).

The rebound phase is characterized by the emergence of high-volatility stocks as better performers. The momentum VOL is negative across all VW terciles and just slightly positive in T3 for EW terciles but not statistically significant. The rebound wasn't a disguised momentum effect, rather it was a risk-on, high volatility outperformance, which was particularly strong for stocks that had previously sold off and remained broad after neutralizing medium-term trends. The reversal shows that high-volatility stocks outperformed low-volatility stocks last month (T1), and in T2 it remains clearly negative, indicating that the rebound in high-volatility stocks exceeded the rebound in the prior losers. Since T3 is small and non-significant, there is no statistically reliable evidence that recent winners contributed. Rebounds were led by high volatility stocks that were most damaged and past losers. In contrast with recent winners, who did not contribute meaningfully to the outperformance of high-volatility stocks, the effect was economically large and statistically significant.

As a result of the post-pandemic phase, momentum reveals that volatility is highly concentrated on recent losers. In tercile 1 (lowest momentum), the VOL spread averages -12.57% EW and -6.47 VW (t=-3, -1.84 respectively), indicating that high-volatility stocks outperformed low-volatility stocks in the losers tercile. The effect weakens towards zero for tercile 2 and 3 while being economically small and statistically insignificant. The effect of high volatility losers rebounded more sharply and investors avoided low volatility losers least, while momentum winners had little effect. Under short-term reversal, tercile 1 (largest one month losers), the VOL spread is -21.04% EW (t = -4.52) and -14.41% VW (t = - 2.96): high-volatility losers bounced more than low-volatility losers the following month. In tercile 2, the effect attenuated, although it was statistically insignificant. A noteworthy observation is that in tercile 3 (one-month winners), the effect reverses sign and low-volatility winners outperform high-volatility winners. In other words, post-pandemic markets may have rewarded high-volatility mean-reversion in recent losers while maintaining preferences for lower-volatility in recent winners.

Table 12: Skewness results phase-by-phase

Phase	Skew Tercile	mean_VOL_EW	t_VOL_EW	mean_VOL_VW	t_VOL_VW
Crash	1	23%	2.16	19%	2.94
Crash	2	14%	1.92	10%	1.86
Crash	3	-7%	-2.29	-6%	-1.82
Post 2022	1	8%	2.71	4%	1.47
Post 2022	2	-2%	-0.61	-2%	-0.60
Post 2022	3	-16%	-4.00	-9%	-2.15
Pre Pandemic	1	7%	2.26	2%	0.66
Pre Pandemic	2	-2%	-0.58	-2%	-0.61
Pre Pandemic	3	-14%	-3.62	-7%	-2.46
Rebound	1	0%	-0.14	-1%	-0.50
Rebound	2	-8%	-3.01	-8%	-3.04
Rebound	3	-21%	-4.58	-13%	-3.77

Based on the pre-pandemic period, it is evident that stocks with higher negative skews (tercile 1) and greater downside tail risks exhibit a positive VOL spread of 7% EW and 2% VW, and this spread is statistically significant for EW (t = 2.26), suggesting that low volatility outperformed high volatility. Spreads fade in tercile 2 and sharply decline in tercile 3 (positive skewed lottery-like stocks) at -14% EW and -7% VW. Low volatility premiums were mainly concentrated in crash-prone, negatively skewed stocks, while high volatility premiums dominated positively skewed stocks prior to Covid-19.

In the crash period, the VOL spread reaches 23% for EW and 19% for VW (t=2.16 and 2.94, respectively). In tercile 2, the spread reaches 14% for EW and 10% for VW, but is not statistically significant. In tercile 3, the sign flips from -7% EW to -6% VW with a high level of significance (t = -4.58 and -3.77). When the stock market was sold off, it paid off to invest in low volatility stocks in the crash-resistant segment of the market, whereas lottery-like names remained high beta leaders, with high volatility outpacing low volatility within this tercile.

The rebound results in most of the terciles flipping to negative VOL spreads, with terciles 2 and 3 showing statistically significant results. In tercile 2, the VOL factor exhibits -8% EW and -8 VW with t-statistics of -3.01 and -3.04. In tercile 3, the spread is more negative at -21% EW and -13% VW (t = -4.58, -3.77). Positively skewed, speculative companies led the rally during the return to risk period, while high volatility outperformed low volatility most strongly in the regions with the greatest lottery characteristics.

During the post-pandemic period, there are significant results for the EW portfolio in terciles 1 and 3 as well as VW exhibits in tercile 3. VOL spreads showed slightly better results in tercile 1 and slightly worse results in tercile 3. The VOL spread in tercile 1 was 8% EW and 4% VW (t=2.71 and 1.47), whereas in tercile 3, it was -16% EW and -9% VW. Despite policy tightening and volatility, low volatility again underperformed against high volatility in lottery-like names.

6. Discussion

The focus of the research involves examining the pronounced role of volatility in explaning cross-sectional returns before, during and after Covid-19. Over the full sample from 2018 to 2022, higher volatility stocks earned higher average excess returns than low volatility stocks, especially amongst small firms. As an example, the Small-vol1 portfolio earned -1.28% while Small-vol5 earned 6.41%, in EW terms. It is also evident in large-cap companies, although it is less extreme, showing a contrast to the traditional low volatility anomaly during this sample period. Ang et al. (2006) documented that U.S stocks with the highest idiosyncratic volatility underperform the lowest volatility volatility stocks by about 0.7% per month in the long run. Likewise, Frazzini and Pedersen (2014) show that because many investors face leverage constraints and cannot lever up safe assets, they instead bid up risky, high-beta stocks, causing those stocks to have low alphas (and low subsequent returns) on average. According to the pre-pandemic subsample, portfolios with high volatility had slightly higher mean returns than portfolios with low volatility, however much of that outperformance was not statistically significant, and cross-sectional regressions confirmed a negative volatility risk price. Data prior to Covid-19 indicated that investors were not rewarded for taking idiosyncratic volatility risk, following the theory that high-volatility stocks could be overvalued due to their lottery-like appeal or leverage.

A shift in the volatility-return relationship was introduced by the Covid-19 crash. The VOL factor pattern flipped in accordance with a flight-to-safety behavior in which high volatility stocks suffered more than low volatility stocks. According to the portfolio analysis, during the crash period, the highest-volatility large stocks fell 13.51% per month on average, compared to 2.29% for the lowest-volatility large stocks (a spread of over 11% per month in favor of low-volatility defensive stocks). The flight-to-quality behavior observed in previous crises (such as the financial crisis of 2008) resulted from this intuitive response in a panic market state. But the reversal was short-lived. Riskier stocks outperformed safer ones during the rebound phase, where volatility-seeking reached extreme levels. For instance, the Small-vol5 portfolio earned 14.6% EW (15.98% VW) per month in the rebound phase, compared to essentially zero for the least volatile small stocks in that segment. During the crash, this speculative surge faded and eventually reversed the low-volatility advantage. By the post-pandemic period (2022), the highest-volatility stocks were once again outperforming low-volatility stocks on average, a pattern reflected in modest positive returns from volatile stocks while stable portfolios continued to lose value. These phase-specific dynamics highlight an important finding, volatility premia are highly market phase-dependent. Investors piled into volatile, high-beta assets in calm and stimulus-fueled conditions, driving their prices and returns higher than those of strong, low-volatility stocks. Conversely, in turmoil, the market rewarded safety and punished volatility. A static low-volatility strategy actually underperformed over 2018–2022, even though theory and long-run evidence suggest it should outperform. It highlights the fact that the COVID-era market was unusual, as it contradicted the usual low-volatility anomaly during the rebound, likely as a result of speculative retail trading, policy support, and renewed investor confidence in markets. It's important to note that the sample is only from 2018-2022, while other research lasted over multi-decade periods. In accordance with historical observations of market bubbles, constrained or optimistic investors tend to buy lottery-like, high-volatility stocks (e.g. small technology firms), temporarily raising their valuations despite lower expected returns (Frazzini and Pedersen, 2014). The findings thereby both challenge and reinforce existing theories. They challenge the universality of the low-volatility anomaly – showing it can disappear or reverse in the short run under extreme conditions – yet they reinforce the notion that this anomaly exists in calmer periods and is rooted in investor behavior and risk constraints. The evidence suggests that the price of volatility risk can change from negative in normal times, where investors accept lower returns for volatile stocks as in Ang et al. (2006), to positive or zero in bull runs when speculative demand is high.

The Fama-French Five Factor (FF5) model used in the paper did not fully explain the return spreads associated with volatility-sorted portfolios during Covid-19. During the pre-pandemic period, the FF5 factors were relatively stable: size and market premia were modest, while value (HML) was already yielding negative returns. The two-pass Fama–MacBeth regressions on the pre pandemic data confirm a significantly negative price of volatility in the cross-section (λ VOL ≈ –0.046 per month, t ≈ –2.7), meaning a stock with one unit higher volatility beta had about 4.6% lower expected return per month. Even after controlling for the FF5 factors, this negative volatility price suggests that volatile stocks were overpriced prior to Covid-19 - investors were willing to pay a premium (accept lower returns) for riskier stocks. In particular, profitability factor's price (RMW) was negative before the pandemic (though not highly significant), suggesting that highly profitable firms did not earn higher returns than unprofitable firms in those years. This pattern is consistent with the fact that growth firms outperform stable firms (often less profitable, but more volatile). This displayed the problems associated with the FF5 model during the pandemic: one of its core intended premiums (value and profitability) was effectively inverted in this environment, with growth firms outperforming values, and less profitable firms outperforming profitable ones.

During the crash period, the factor risk premia underwent shifts. The Fama–MacBeth analysis showed the market factor’s price turned deeply negative (λ^{MKT} ≈ –8.24%, t≈–3), indicating that stocks with higher market beta suffered disproportionately – a direct reflection of the market-wide plunge. Furthermore, the value factor (HML) price also flipped sharply downward (-18.14 %), confirming that value stocks (which normally carry a premium) were particularly hard-hit during the economic downturn. The results confirm that the Covid-19 shock was asymmetric: technology and growth companies (often with high volatility and high valuations) proved to be more resilient or recovered more quickly than "cheap" value stocks. Meanwhile, the volatility factor’s price in the crash became positive (λ^{VOL} = 10.03%, t = 1.22) – an interesting result, although statistically insignificant with only three months of data. A positive λ VOL would imply investors demanded extra return for bearing volatility risk in the crash (ahigh-volatility stocks expected returns rose), which makes sense as volatility was associated with especially large losses during the panic. Essentially, the market reversed its pricing of risk: safer stocks were rewarded (negative beta and low volatility stocks provided relative shelter), whereas riskier stocks were penalized. These findings support the concept of a downside beta or crisis risk effect, in which investors pay a premium for "safer" securities in rare periods of downturn.

During the rebound, the risk pricing changed again. The market factor earned a high positive premium (λ^{MKT} = 3.76%, t= 3.85) as the market climbed up again, investors who took market exposure were rewarded in these months. Size (SMB) had a mild positive price (reflecting small-caps participating in the rally) but not very significant. Remarkably, the value factor's price remained strongly negative (λ^{HML} = –5.21%, t = –3.45) and value stocks continued to lag far behind growth, intensifying the "value problem" that had begun before Covid-19. The profitability factor (RMW) also carried a negative price in the rebound (–2%, t= –2.3), indicating that firms with strong profits were not the market's best choices in the speculative rally – rather, unprofitable or story-driven companies (biotech, meme stocks) saw superior returns. Moreover, the volatility factor's price became highly negative in the rebound ($\lambda^{VOL} \approx$ –10.57%, t ≈–3.1). Accordingly, stocks with higher volatility loadings (or simply higher idiosyncratic volatility) had higher realized returns during this phase, an inversion of the usual pricing pattern. As a measure of risk premium, a negative λ VOL indicates that volatility was not viewed as a risk to be penalized during the rebound, but rather as a source of gain. This result matches the evidence that high-volatility portfolios dramatically outperformed in the rebound. Consequently, the five-factor model fails to account for this period's cross-section. Many high-vol stocks earned high returns without corresponding factor exposure explanations, which yielded significant positive FF5 alphas. This can be interpreted as a sign of a short-term speculative bubble or sentiment-driven pricing: the traditional risk factors (which would regard volatility as undesirable risks) were temporarily overwhelmed by a flood of risk-tolerant capital seeking to acquire volatile assets.

Finally, after the pandemic, the pricing of factors began to normalize, although not entirely to pre-crisis levels. Inflation and interest rates rose in 2022, and the market suffered a general decline, which particularly affected growth stocks with high valuations. In the Fama–MacBeth results, the market factor premium turned slightly negative (λ^{MKT} = -1.35%, t= –1.04). The size factor became significantly negative ($\lambda^{SMB} \approx$ -0.97%, t = –2.2), indicating small-cap stocks underperformed large-caps in 2022 – a reversion to a more risk-averse stance. Notably, the volatility factor's price in 2022 was still negative and significant ($\lambda^{VOL} \approx$ -4.63%, t ≈ –2.05). Meanwhile, value (HML) showed a positive but mild premium in 2022 (λ^{HML} = 7.97%, t=1.5) – value stocks started to make a comeback as investors rotated out of bad growth stocks when interest rates rose (though this was not statistically robust). There was little movement in profitability and investment factors. Overall, the market was re-aligning with fundamental risk pricing by late 2021 and 2022, rewarding value and low volatility to some extent following the abnormal phase of the rebound rally.

A VOL factor provides an additional dimension of systematic risk or mispricing related to volatility. While the measure is based on total volatility, Herskovic et al. (2016) also emphasize that volatility risk matters and that a volatility factor can improve asset pricing models. The results of the study support this view - especially in volatile regimes, a volatility factor is required to explain the return patterns that are not explained by standard size, value, and profitability factors. On average, investors appeared to demand (and earn) a premium for bearing volatility risk during this period, but the premium varied greatly between sub-periods.

Several limitations are present in this research that must be addressed. The sample period is relatively short (five years) and dominated by an extreme event, which limits the statistical power. Although the goal of the study is to focus on the Covid-19 window, one needs to be cautious in extrapolating these results to other contexts, as the pandemic coincided with both an unprecedented global economic shutdown and extraordinary rapid policy responses that may not recur in the same way. In this study, the U.S. suffered greatly and may not reflect on more tranquil economies. As the Fama-Macbeth estimation window is short (2013–2022) and some sub-periods, including the three-month crash, offer limited power, Newey–West inferences are sensitive and suggestive, rather than definitive. Additionally, the portfolio sorting approach (size and volatility quintiles) simplifies what are continuous firm characteristics. It may not capture heterogeneity within quintiles. By performing double-sorts (ensuring that size and volatility effects are disentangled) and comparing value-weighted returns with equal-weighted returns, there was an attempt to minimize these effects. The results were qualitatively robust - for example, excluding the bottom decile did not alter the key patterns in volatility spreads and returns. Nevertheless, there are concerns regarding endogeneity and omitted variables. It has been suggested that volatility itself may be endogenous to expected returns. For example, firms with poor prospects or distress risk often exhibit higher volatility and lower returns (distress anomaly), which could explain part of the low-volatility phenomenon. There was an attempt made to control for known factors such as size, value, and profitability, however, there may also be other correlated characteristics that could influence the relationship (such as liquidity or credit risk). In times of crisis, government interventions (such as the Federal Reserve's actions in 2020) can break the usual link between risk and return. The study does not explicitly model these effects. As a result, we are unable to establish a direct causal relationship between volatility and its effects.

7. Conclusion

In conclusion, this study examined whether firm size and volatility can help explain the cross-section of U.S stock returns around the Covid-19 shock. Three conclusions emerge. First, volatility is priced negatively on average across 2018-2022. The full-sample Fama–MacBeth estimate for VOL's price of risk is −5.09% per month ($t = -5.10$), implying that, holding the other factors constant, assets with greater exposure to the low-minus-high volatility premium earn lower expected returns. The portfolio evidence supports this conclusion: low-volatility portfolios generally outperformed high-volatility peers, except during the period of the crash. Second, factor premia are state-dependent. Before the pandemic, VOL carried a significantly negative price (−4.61%, $t \approx -2.66$) amid a market that still rewarded higher-risk names in returns; during the crash the sign briefly flipped positive (10.03%, $t = 1.2$), consistent with the rare dynamics of panic selling; in the rebound turned sharply negative again (−10.57%, $t = -3.08$), and it remained negative in 2022 (−4.63%, $t = -2.05$). Over the same phases, HML was negative in the pre-pandemic (−8.44%, $t = -2.18$) and rebound (−5.21%, $t \approx -3.5$), while MKT swung from insignificance pre-COVID to strongly positive in the rebound (3.76%, $t \approx 3.9$) and modestly negative in 2022. Third, the results hold when (i) microcaps are excluded, (ii) momentum, reversal, and return skewness are conditioned on VOL spread, and (iii) total volatility is replaced with idiosyncratic volatility (FF5-residual IVOL). It has been shown that phase-dependence is supported by conditioning analyses: VOL rewards are highest when prior momentum is strong (rebound) and when past returns are positively skewed, and they diminish when reversal occurs, in line with shifting risk appetite and attention across regimes. Both portfolio and Fama–MacBeth evidence suggest that volatility risk is priced negatively on average, but that this pricing varies with macro-stress and risk appetite.

This study has important limitations. As the five-year sample is centered on the COVID-19 shock, statistical power is limited, and extrapolation beyond this unique phase, characterized by a global shutdown and unusually rapid policy support, should be undertaken with caution. Evidence is limited to the United States and may not be generalizable to other or calmer economies. As the Fama-MacBeth window is short and some subperiods (pandemic phase) do not provide many observations, Newey-West inferences are less certain; therefore, the results are more suggestive than definitive. By sorting portfolios into size-volatility quintiles, continuous characteristics are discretized and within-bin heterogeneity can be hidden; we mitigate this through double-sorts, equal-weighted versus value-weighted returns, and exclusion of the bottom decile, which result in qualitatively robust volatility-spread patterns and returns. However, endogeneity and the omission of variables remain concerns: volatility may be a proxy for distress or other characteristics (for example, liquidity or credit risk), and crisis-era policies may distort the usual risk-return tradeoff. Thus, there can be no definite causal identification of volatility's effect. The evidence supports a clear conclusion: volatility is a negatively priced dimension of risk in the cross-section, and its strength and sometimes its sign vary by state. The addition of a size-neutral volatility factor to the FF5 model significantly improves our understanding of returns around an unprecedented macro shock, reconciling raw performance of volatility-sorted portfolios with the formal estimate of risk-price.

8. References

Ang, A., Hodrick, R. J., Xing, Y., & Zhang, X. (2006). The cross-section of volatility and expected returns. *The Journal of Finance, 61*(1), 259–299.

Asness, C. S., Frazzini, A., Israel, R., Moskowitz, T. J., & Pedersen, L. H. (2018). Size matters, if you control your junk. *Journal of Financial Economics, 129*(3), 479–509. https://doi.org/10.1016/j.jfineco.2018.05.006

Baker, S., Bloom, N., Davis, S., Kost, K., Sammon, M., & Viratyosin, T. (2020a). The unprecedented stock market impact of COVID-19. *The Review of Asset Pricing Studies*. https://doi.org/10.3386/w26945

Bali, T. G., Cakici, N., & Whitelaw, R. F. (2011). Maxing out: Stocks as lotteries and the cross-section of expected returns. *Journal of Financial Economics, 99*(2), 427–446. https://doi.org/10.1016/j.jfineco.2010.08.014

Banz, R. W. (1981). The relationship between return and market value of common stocks. *Journal of Financial Economics, 9*(1), 3–18. https://doi.org/10.1016/0304-405X(81)90018-0

Barber, B. M., Huang, X., Odean, T., & Schwarz, C. (2022). Attention-induced trading and returns: Evidence from Robinhood users. *Journal of Finance, 77*(6), 3141–3190.

Brunnermeier, M. K., & Pedersen, L. H. (2009). Market liquidity and funding liquidity. *The Review of Financial Studies, 22*(6), 2201–2238. https://doi.org/10.1093/rfs/hhn098

Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics, 33*(1), 3–56.

Federal Reserve Economic Data. FRED. (n.d.). https://fredhelp.stlouisfed.org/fred/data/understanding-the-data/recession-bars/

Frazzini, A., & Pedersen, L. H. (2014). Betting against beta. *Journal of Financial Economics, 111*(1), 1–25.

Jegadeesh, N., & Titman, S. (1993). Returns to buying winners and selling losers: Implications for Stock Market Efficiency. *The Journal of Finance*, *48*(1), 65. https://doi.org/10.2307/2328882

O'Hara, M., & Zhou, X. (2021). Anatomy of a liquidity crisis: Corporate bonds in the COVID-19 crisis. *Journal of Financial Economics, 142*(1), 46–68.

Pérez-Quiros, G., & Timmermann, A. (2000). Firm size and cyclical variations in stock returns. *Journal of Finance, 55*(3), 1229–1262.

Stambaugh, R. F., Yu, J., & Yuan, Y. (2015). Arbitrage asymmetry and the idiosyncratic volatility puzzle. *Journal of Finance, 70*(5), 1903–1948.

9. Appendices

Appendix A: Portfolio returns 6 months rolling window robustness check

Portfolio Returns 6 months rolling window

Portfolio	Average_EW	T_Stat_EW	Average_VW	T_Stat_VW	Period
Large-Vol1	1.05%	2.92	1.30%	3.20	Entire Sample
Large-Vol2	1.06%	1.81	1.52%	2.28	Entire Sample
Large-Vol3	1.15%	1.50	1.47%	1.76	Entire Sample
Large-Vol4	1.25%	1.30	1.12%	1.12	Entire Sample
Large-Vol5	3.79%	2.49	4.04%	2.54	Entire Sample
Small-Vol1	-0.47%	-1.06	0.33%	0.68	Entire Sample
Small-Vol2	-1.05%	-1.40	0.19%	0.24	Entire Sample
Small-Vol3	-1.29%	-1.25	0.54%	0.48	Entire Sample
Small-Vol4	-0.93%	-0.69	1.74%	1.21	Entire Sample
Small-Vol5	5.10%	2.36	9.03%	3.91	Entire Sample
Large-Vol1	1.24%	2.84	1.58%	3.15	Pre-Pandemic
Large-Vol2	1.08%	1.52	1.51%	1.94	Pre-Pandemic
Large-Vol3	1.07%	1.08	1.33%	1.21	Pre-Pandemic
Large-Vol4	1.19%	0.96	1.20%	0.97	Pre-Pandemic
Large-Vol5	2.83%	1.59	3.17%	1.76	Pre-Pandemic
Small-Vol1	-0.29%	-0.53	0.40%	0.68	Pre-Pandemic
Small-Vol2	-0.72%	-0.77	0.22%	0.20	Pre-Pandemic
Small-Vol3	-1.22%	-0.96	0.33%	0.24	Pre-Pandemic
Small-Vol4	-1.10%	-0.70	1.70%	1.00	Pre-Pandemic
Small-Vol5	2.67%	1.06	7.61%	2.80	Pre-Pandemic
Large-Vol1	-2.38%	-1.30	-2.06%	-0.81	Crash
Large-Vol2	-5.44%	-1.73	-4.99%	-1.32	Crash
Large-Vol3	-8.01%	-2.09	-7.89%	-2.33	Crash
Large-Vol4	-11.54%	-2.15	-11.30%	-2.32	Crash
Large-Vol5	-13.30%	-1.41	-9.94%	-0.95	Crash
Small-Vol1	-6.10%	-2.47	-5.82%	-2.37	Crash
Small-Vol2	-10.75%	-2.29	-10.72%	-2.45	Crash
Small-Vol3	-14.66%	-2.04	-15.07%	-1.97	Crash
Small-Vol4	-17.54%	-1.75	-17.13%	-1.81	Crash
Small-Vol5	-9.54%	-0.87	-12.83%	-1.32	Crash
Large-Vol1	2.16%	4.31	2.35%	3.97	Rebound
Large-Vol2	2.96%	3.61	3.86%	4.02	Rebound
Large-Vol3	3.82%	3.82	4.12%	3.36	Rebound
Large-Vol4	4.51%	3.28	4.75%	3.13	Rebound
Large-Vol5	9.09%	3.48	9.98%	3.66	Rebound
Small-Vol1	1.36%	2.30	2.15%	3.39	Rebound
Small-Vol2	1.79%	1.74	3.10%	2.88	Rebound
Small-Vol3	2.68%	1.76	4.87%	3.07	Rebound
Small-Vol4	3.96%	1.79	6.22%	2.63	Rebound
Small-Vol5	13.07%	2.97	16.03%	3.39	Rebound
Large-Vol1	-0.41%	-0.39	-0.25%	-0.22	Post-Pandemic
Large-Vol2	-0.69%	-0.42	-0.90%	-0.49	Post-Pandemic
Large-Vol3	-1.06%	-0.50	-0.56%	-0.26	Post-Pandemic
Large-Vol4	-1.11%	-0.49	-2.29%	-0.99	Post-Pandemic
Large-Vol5	0.71%	0.22	-1.15%	-0.32	Post-Pandemic
Small-Vol1	-2.61%	-2.79	-1.50%	-1.23	Post-Pandemic
Small-Vol2	-4.23%	-2.65	-2.21%	-1.17	Post-Pandemic
Small-Vol3	-5.02%	-2.44	-2.72%	-1.14	Post-Pandemic
Small-Vol4	-4.98%	-1.96	-1.31%	-0.44	Post-Pandemic
Small-Vol5	-0.33%	-0.09	5.08%	1.25	Post-Pandemic

Appendix B: Portfolio returns 12 months rolling window robustness check

Portfolio returns for 12 month rolling window

Portfolio	Average_EW	T_Stat_EW	Average_VW	T_Stat_VW	Period
Large-Vol1	1.06%	2.48	1.36%	2.98	Entire Sample
Large-Vol2	1.10%	1.74	1.52%	2.11	Entire Sample
Large-Vol3	1.14%	1.46	1.10%	1.29	Entire Sample
Large-Vol4	1.40%	1.49	1.61%	1.62	Entire Sample
Large-Vol5	3.51%	2.45	3.99%	2.59	Entire Sample
Small-Vol1	0.09%	0.17	0.52%	0.90	Entire Sample
Small-Vol2	-0.50%	-0.61	0.34%	0.39	Entire Sample
Small-Vol3	-0.94%	-0.89	0.67%	0.57	Entire Sample
Small-Vol4	-0.76%	-0.56	2.02%	1.39	Entire Sample
Small-Vol5	4.00%	1.97	9.02%	4.12	Entire Sample
Large-Vol1	1.24%	2.46	1.57%	2.69	Pre-Pandemic
Large-Vol2	1.05%	1.32	1.40%	1.61	Pre-Pandemic
Large-Vol3	1.09%	1.13	1.03%	1.09	Pre-Pandemic
Large-Vol4	1.21%	0.96	1.58%	1.16	Pre-Pandemic
Large-Vol5	2.75%	1.67	2.91%	1.74	Pre-Pandemic
Small-Vol1	0.10%	0.15	0.52%	0.68	Pre-Pandemic
Small-Vol2	-0.46%	-0.45	0.21%	0.19	Pre-Pandemic
Small-Vol3	-0.77%	-0.62	0.65%	0.46	Pre-Pandemic
Small-Vol4	-0.90%	-0.54	1.96%	1.15	Pre-Pandemic
Small-Vol5	1.54%	0.69	7.50%	2.82	Pre-Pandemic
Large-Vol1	-2.92%	-1.02	-2.56%	-0.84	Crash
Large-Vol2	-6.00%	-1.83	-5.07%	-1.40	Crash
Large-Vol3	-8.53%	-2.11	-9.18%	-2.13	Crash
Large-Vol4	-10.59%	-2.46	-10.40%	-2.85	Crash
Large-Vol5	-12.62%	-1.34	-8.04%	-0.80	Crash
Small-Vol1	-6.94%	-2.35	-6.67%	-2.26	Crash
Small-Vol2	-11.74%	-2.44	-12.41%	-2.56	Crash
Small-Vol3	-15.15%	-1.84	-15.59%	-1.75	Crash
Small-Vol4	-17.41%	-1.95	-17.79%	-2.01	Crash
Small-Vol5	-7.79%	-0.74	-7.01%	-0.84	Crash
Large-Vol1	2.11%	3.16	2.55%	3.48	Rebound
Large-Vol2	3.13%	3.53	4.10%	3.56	Rebound
Large-Vol3	3.71%	3.48	3.70%	3.02	Rebound
Large-Vol4	4.78%	3.65	5.01%	3.46	Rebound
Large-Vol5	8.58%	3.50	10.25%	3.85	Rebound
Small-Vol1	2.00%	2.84	2.47%	3.22	Rebound
Small-Vol2	2.71%	2.47	3.70%	3.17	Rebound
Small-Vol3	3.17%	2.02	4.83%	2.91	Rebound
Small-Vol4	4.47%	2.05	7.03%	2.88	Rebound
Small-Vol5	11.35%	2.64	15.29%	3.28	Rebound
Large-Vol1	-0.14%	-0.12	-0.17%	-0.16	Post-Pandemic
Large-Vol2	-0.60%	-0.35	-1.13%	-0.64	Post-Pandemic
Large-Vol3	-0.85%	-0.40	-0.76%	-0.31	Post-Pandemic
Large-Vol4	-1.14%	-0.51	-1.27%	-0.54	Post-Pandemic
Large-Vol5	0.21%	0.07	-1.81%	-0.52	Post-Pandemic
Small-Vol1	-1.51%	-1.26	-1.07%	-0.70	Post-Pandemic
Small-Vol2	-3.38%	-1.92	-2.09%	-1.02	Post-Pandemic
Small-Vol3	-4.92%	-2.48	-2.53%	-1.04	Post-Pandemic
Small-Vol4	-5.48%	-2.13	-1.67%	-0.59	Post-Pandemic
Small-Vol5	-1.02%	-0.31	5.08%	1.43	Post-Pandemic

Appendix C: Portfolio returns 3 months rolling window excluding bottom decile

Portfolio returns for 3 months rolling window excluding bottom decile

Portfolio	Average_EW	T_Stat_EW	Average_VW	T_Stat_VW	Period
Large-Vol1	0.89%	2.53	1.06%	2.71	Entire Sample
Large-Vol2	1.01%	1.94	1.52%	2.75	Entire Sample
Large-Vol3	1.22%	1.65	1.24%	1.65	Entire Sample
Large-Vol4	1.43%	1.45	1.61%	1.54	Entire Sample
Large-Vol5	3.68%	2.40	3.87%	2.45	Entire Sample
Small-Vol1	-0.03%	-0.08	0.25%	0.58	Entire Sample
Small-Vol2	0.00%	-0.01	0.38%	0.57	Entire Sample
Small-Vol3	0.13%	0.14	0.47%	0.51	Entire Sample
Small-Vol4	0.54%	0.44	1.00%	0.80	Entire Sample
Small-Vol5	5.71%	2.65	6.54%	3.05	Entire Sample
Large-Vol1	1.18%	2.45	1.40%	3.01	Pre-Pandemic
Large-Vol2	1.21%	1.95	1.73%	2.58	Pre-Pandemic
Large-Vol3	0.96%	1.04	1.06%	1.15	Pre-Pandemic
Large-Vol4	1.32%	1.09	1.55%	1.25	Pre-Pandemic
Large-Vol5	2.72%	1.41	3.02%	1.54	Pre-Pandemic
Small-Vol1	0.18%	0.31	0.48%	0.82	Pre-Pandemic
Small-Vol2	0.17%	0.20	0.53%	0.61	Pre-Pandemic
Small-Vol3	0.21%	0.19	0.44%	0.40	Pre-Pandemic
Small-Vol4	0.50%	0.33	0.93%	0.59	Pre-Pandemic
Small-Vol5	4.48%	1.80	4.99%	1.94	Pre-Pandemic
Large-Vol1	-2.31%	-0.92	-1.65%	-0.58	Crash
Large-Vol2	-4.99%	-1.74	-3.39%	-1.24	Crash
Large-Vol3	-7.10%	-1.94	-5.77%	-1.75	Crash
Large-Vol4	-10.74%	-2.16	-10.80%	-2.79	Crash
Large-Vol5	-13.74%	-1.53	-10.75%	-1.27	Crash
Small-Vol1	-4.61%	-1.54	-4.23%	-1.42	Crash
Small-Vol2	-8.06%	-2.29	-8.20%	-2.28	Crash
Small-Vol3	-12.27%	-2.15	-12.06%	-2.19	Crash
Small-Vol4	-16.00%	-2.01	-15.11%	-1.99	Crash
Small-Vol5	-19.84%	-1.80	-17.37%	-1.57	Crash
Large-Vol1	2.20%	4.98	2.46%	3.95	Rebound
Large-Vol2	2.80%	3.93	3.29%	4.22	Rebound
Large-Vol3	3.81%	3.63	3.72%	3.13	Rebound
Large-Vol4	4.54%	3.17	4.79%	3.03	Rebound
Large-Vol5	8.82%	3.52	9.15%	3.52	Rebound
Small-Vol1	1.60%	2.64	1.80%	3.12	Rebound
Small-Vol2	2.46%	2.68	2.86%	3.18	Rebound
Small-Vol3	3.29%	2.59	3.69%	2.84	Rebound
Small-Vol4	4.67%	2.56	5.08%	2.76	Rebound
Small-Vol5	11.98%	2.95	12.88%	3.22	Rebound
Large-Vol1	-1.19%	-1.63	-1.41%	-1.92	Post-Pandemic
Large-Vol2	-1.03%	-0.73	-0.76%	-0.48	Post-Pandemic
Large-Vol3	-0.74%	-0.38	-1.02%	-0.52	Post-Pandemic
Large-Vol4	-0.74%	-0.28	-0.76%	-0.26	Post-Pandemic
Large-Vol5	0.97%	0.28	-0.02%	-0.01	Post-Pandemic
Small-Vol1	-2.17%	-2.67	-1.83%	-2.34	Post-Pandemic
Small-Vol2	-2.65%	-1.79	-2.12%	-1.38	Post-Pandemic
Small-Vol3	-2.46%	-1.17	-1.98%	-0.90	Post-Pandemic
Small-Vol4	-2.49%	-0.88	-1.99%	-0.67	Post-Pandemic
Small-Vol5	3.59%	0.84	4.50%	1.04	Post-Pandemic

Appendix D: VOL factor by phase and VOL factor by phase excluding bottom decile

VOL Factor by phase using different rolling windows

Phase	mean_VOL_EW	t_VOL_EW	mean_VOL_VW	t_VOL_VW	Window
Crash	8.41%	1.69	10.15%	2.08	3m
Post-Pandemic	-3.93%	-1.67	-4.57%	-1.72	3m
Pre-Pandemic	-3.23%	-2.18	-4.62%	-2.83	3m
Rebound	-10.47%	-4.23	-10.69%	-4.29	3m
Crash	7.18%	1.26	7.48%	1.32	6m
Post-Pandemic	-1.70%	-0.89	-2.83%	-1.28	6m
Pre-Pandemic	-2.27%	-1.78	-4.39%	-3.13	6m
Rebound	-9.37%	-3.90	-10.77%	-4.24	6m
Crash	5.27%	0.96	2.91%	0.55	12m
Post-Pandemic	-0.42%	-0.25	-2.26%	-1.10	12m
Pre-Pandemic	-1.48%	-1.37	-4.16%	-3.18	12m
Rebound	-7.91%	-3.44	-10.26%	-4.15	12m

VOL Factor by phase using different rolling windows excluding bottom decile

Phase	mean_VOL_EW	t_VOL_EW	mean_VOL_VW	t_VOL_VW	Window
Crash	13.33%	2.75	11.13%	2.32	3m_excl10
Post-Pandemic	-3.96%	-1.70	-3.86%	-1.54	3m_excl10
Pre-Pandemic	-2.92%	-2.16	-3.07%	-2.16	3m_excl10
Rebound	-8.50%	-3.82	-8.88%	-4.04	3m_excl10
Crash	12.23%	2.18	9.73%	1.71	6m_excl10
Post-Pandemic	-2.16%	-1.14	-1.77%	-0.87	6m_excl10
Pre-Pandemic	-2.66%	-2.38	-2.99%	-2.51	6m_excl10
Rebound	-7.93%	-3.87	-8.69%	-4.15	6m_excl10
Crash	10.26%	2.03	7.35%	1.38	12m_excl10
Post-Pandemic	-1.28%	-0.79	-1.00%	-0.54	12m_excl10
Pre-Pandemic	-1.99%	-2.07	-2.46%	-2.43	12m_excl10
Rebound	-7.16%	-3.67	-8.28%	-4.19	12m_excl10

Appendix E: IVOL by phase

IVOL by phase

Phase	mean_VOL_EW	t_VOL_EW	mean_VOL_VW	t_VOL_VW
Crash	-3.46%	-1.44	-4.44%	-1.42
Post-Pandemic	-0.58%	-0.48	-2.96%	-1.95
Pre-Pandemic	-1.37%	-1.70	-3.63%	-3.24
Rebound	-6.41%	-3.59	-8.60%	-4.39

Appendix F: Momentum by phase robustness checks

Phase	Tercile	mean_VOL_EW	t_VOL_EW	mean_VOL_VW	t_VOL_VW	Signal	Window
Crash	1	-0.42%	-0.05	5.05%	0.47	MOM6_1	6m
Crash	2	11.60%	1.06	15.83%	1.44	MOM6_1	6m
Crash	3	10.07%	1.27	4.17%	0.49	MOM6_1	6m
Post Pandemic	1	-9.41%	-2.73	-7.05%	-2.57	MOM6_1	6m
Post Pandemic	2	-2.33%	-0.98	2.53%	1	MOM6_1	6m
Post Pandemic	3	1.60%	0.61	0.36%	0.13	MOM6_1	6m
Pre Pandemic	1	-6.61%	-2.16	-4.85%	-2.27	MOM6_1	6m
Pre Pandemic	2	-1.04%	-0.45	-1.89%	-0.88	MOM6_1	6m
Pre Pandemic	3	0.64%	0.31	-0.56%	-0.28	MOM6_1	6m
Rebound Apr 20-Dec 21	1	-14.31%	-3.16	-8.27%	-2.62	MOM6_1	6m
Rebound Apr 20-Dec 21	2	-11.19%	-3.39	-8.16%	-3.23	MOM6_1	6m
Rebound Apr 20-Dec 21	3	-6.96%	-2.2	-7.48%	-2.34	MOM6_1	6m
Crash	1	-1.73%	-0.23	2.64%	0.25	MOM6_1	12m
Crash	2	8.38%	0.94	14.46%	1.39	MOM6_1	12m
Crash	3	6.57%	0.93	1.59%	0.24	MOM6_1	12m
Post Pandemic	1	-5.28%	-1.72	-2.15%	-0.7	MOM6_1	12m
Post Pandemic	2	-1.59%	-0.75	2.78%	1.1	MOM6_1	12m
Post Pandemic	3	0.33%	0.15	0.29%	0.1	MOM6_1	12m
Pre Pandemic	1	-3.79%	-1.54	-4.50%	-2.69	MOM6_1	12m
Pre Pandemic	2	-0.41%	-0.24	-1.33%	-0.78	MOM6_1	12m
Pre Pandemic	3	0.68%	0.4	-0.62%	-0.37	MOM6_1	12m
Rebound Apr 20-Dec 21	1	-11.03%	-2.45	-6.30%	-2.08	MOM6_1	12m
Rebound Apr 20-Dec 21	2	-9.55%	-2.88	-8.33%	-3.53	MOM6_1	12m
Rebound Apr 20-Dec 21	3	-6.44%	-2.18	-8.20%	-2.83	MOM6_1	12m

Appendix G: Reversal by phase robustness checks

Reversal by Phase

Phase	Tercile	mean_VOL_EW	t_VOL_EW	mean_VOL_VW	t_VOL_VW	Window	Signal
Crash	1	2.48%	0.21	5.69%	0.47	6m	REV1
Crash	2	6.43%	0.71	14.10%	1.38	6m	REV1
Crash	3	6.24%	0.81	3.82%	0.58	6m	REV1
Post Pandemic	1	-9.29%	-2.56	-2.83%	-0.72	6m	REV1
Post Pandemic	2	-3.52%	-1.20	-2.41%	-0.77	6m	REV1
Post Pandemic	3	2.44%	0.98	2.81%	1.07	6m	REV1
Pre Pandemic	1	-6.15%	-2.63	-5.42%	-3.41	6m	REV1
Pre Pandemic	2	-2.35%	-1.16	-1.45%	-0.80	6m	REV1
Pre Pandemic	3	0.35%	0.22	0.61%	0.35	6m	REV1
Rebound Apr 20-Dec 21	1	-16.78%	-3.69	-9.97%	-3.24	6m	REV1
Rebound Apr 20-Dec 21	2	-11.33%	-3.49	-8.04%	-3.10	6m	REV1
Rebound Apr 20-Dec 21	3	-5.99%	-1.70	-7.06%	-2.34	6m	REV1
Crash	1	3.16%	0.33	5.49%	0.47	12m	REV1
Crash	2	6.28%	0.77	9.25%	0.89	12m	REV1
Crash	3	1.47%	0.19	0.71%	0.13	12m	REV1
Post Pandemic	1	-3.17%	-0.97	0.65%	0.17	12m	REV1
Post Pandemic	2	-1.65%	-0.70	-2.03%	-0.63	12m	REV1
Post Pandemic	3	1.07%	0.47	1.70%	0.61	12m	REV1
Pre Pandemic	1	-1.94%	-1.15	-2.36%	-1.92	12m	REV1
Pre Pandemic	2	-0.81%	-0.53	-1.27%	-0.79	12m	REV1
Pre Pandemic	3	-0.62%	-0.43	-0.92%	-0.66	12m	REV1
Rebound Apr 20-Dec 21	1	-11.72%	-2.76	-7.61%	-2.66	12m	REV1
Rebound Apr 20-Dec 21	2	-8.60%	-2.89	-7.58%	-2.68	12m	REV1
Rebound Apr 20-Dec 21	3	-6.13%	-2.03	-7.98%	-2.78	12m	REV1

Appendix H: Skewness by phase robustness checks

Skewness by phase

Crash	1	18.45%	1.47	15.47%	1.47	6m
Crash	2	14.89%	1.53	10.34%	1.16	6m
Crash	3	-4.71%	-0.73	-4.18%	-0.86	6m
Post Pandemic	1	6.19%	2.53	1.79%	0.62	6m
Post Pandemic	2	0.28%	0.11	-0.70%	-0.23	6m
Post Pandemic	3	-9.05%	-2.79	-0.19%	-0.06	6m
Pre Pandemic	1	4.72%	1.86	-0.18%	-0.08	6m
Pre Pandemic	2	-1.47%	-0.55	-2.35%	-1.00	6m
Pre Pandemic	3	-10.05%	-2.73	-7.16%	-3.00	6m
Rebound Apr 20-Dec 21	1	-2.27%	-0.91	-2.89%	-1.30	6m
Rebound Apr 20-Dec 21	2	-7.16%	-2.63	-8.29%	-2.67	6m
Rebound Apr 20-Dec 21	3	-15.64%	-3.67	-11.32%	-3.42	6m
Crash	1	12.32%	1.13	10.59%	0.89	12m
Crash	2	12.12%	1.38	9.00%	1.20	12m
Crash	3	-1.43%	-0.19	-2.96%	-0.62	12m
Post Pandemic	1	4.24%	1.86	1.95%	0.64	12m
Post Pandemic	2	1.33%	0.58	0.77%	0.27	12m
Post Pandemic	3	-4.05%	-1.48	0.83%	0.24	12m
Pre Pandemic	1	3.36%	1.61	-1.13%	-0.63	12m
Pre Pandemic	2	-0.13%	-0.06	-1.83%	-0.89	12m
Pre Pandemic	3	-5.43%	-1.71	-4.91%	-2.22	12m
Rebound Apr 20-Dec 21	1	-3.12%	-1.49	-3.86%	-1.91	12m
Rebound Apr 20-Dec 21	2	-5.67%	-2.24	-7.75%	-2.62	12m
Rebound Apr 20-Dec 21	3	-11.27%	-2.61	-12.74%	-2.54	12m

Appendix I: Final code

```
# ==========================
# Thesis  + robustness (3m, 6m, 12m)
# Non-overlapping phases; with micro-cap exclusion, IVOL (FF5),
# momentum/reversal conditioning, and skewness conditioning (fixed).
# All outputs have "20aug_".
# ==========================
rm(list = ls())

# ---- One-time auto-installer for needed packages ----
local({
  needed <- c("sandwich", "lmtest")
  have   <- rownames(installed.packages())
  miss   <- setdiff(needed, have)
  if (!length(getOption("repos")) || getOption("repos")["CRAN"] %in% c("", "@CRAN@")) {
    options(repos = c(CRAN = "https://cloud.r-project.org"))
  }
  options(download.file.method = "libcurl")
  if (length(miss)) {
    message("Installing missing packages: ", paste(miss, collapse = ", "))
    install.packages(miss, dependencies = TRUE)
  }
})

# --- Libraries ---
suppressPackageStartupMessages({
  library(readxl)
  library(dplyr)
  library(zoo)
  library(tidyr)
  library(ggplot2)
```

```
 library(xtable)
 library(e1071)   # for skewness (type=3)
 library(openxlsx) # for Excel output
 # NEW for FM/FF5
 library(sandwich) # HAC SEs
 library(lmtest)  # coeftest / convenience with HAC
})

# ----------------------------
# 1) Load & clean CRSP panel
# ----------------------------
crsp_path <- "/Users/bernardomata/Desktop/Thesis/Data Final/FINAL stock data thesis.xlsx"
ff5_path  <- "/Users/bernardomata/Desktop/Thesis/Data Final/FINAL fama french 5 monthly.csv"

data <- read_excel(crsp_path)
colnames(data) <- c(
 "PERMNO","Names_Date","Share_Code","Exchange_Code","Ticker_Symbol",
 "Company_Name","NAICS","Price_or_BidAsk_Average","Returns",
 "Shares_Outstanding","Cumulative_Factor_to_Adjust_Prices"
)
data$Names_Date <- as.Date(data$Names_Date)

# keep NYSE/NASDAQ common stocks; drop financials; drop NA; require non-negatives
data <- data %>%
 filter(Exchange_Code %in% c(1,3),
     Share_Code %in% c(10,11)) %>%
 filter(!grepl("^52", as.character(NAICS)))

req_cols <- c("PERMNO","Names_Date","Share_Code","Exchange_Code","Ticker_Symbol",
       "Company_Name","NAICS","Price_or_BidAsk_Average","Returns",
       "Shares_Outstanding","Cumulative_Factor_to_Adjust_Prices")
data <- data[complete.cases(data[, req_cols]), ]

nonneg <- c("Price_or_BidAsk_Average","Shares_Outstanding","Cumulative_Factor_to_Adjust_Prices")
data <- data %>% filter(if_all(all_of(nonneg), ~ . >= 0))

# Market cap
data <- data %>%
 mutate(Market_Cap = as.numeric(Price_or_BidAsk_Average) * as.numeric(Shares_Outstanding) * 1000) %>%
 filter(Market_Cap > 0, Cumulative_Factor_to_Adjust_Prices > 0)

data$PERMNO <- as.character(data$PERMNO)

# ----------------------------
# 2) Periods & helpers  (NON-OVERLAPPING PHASES)
# ----------------------------
periods <- list(
 entire_sample = c("2018-01-01","2022-12-31"),
 pre_pandemic  = c("2018-01-01","2019-12-31"),
 crash       = c("2020-01-01","2020-03-31"),
 rebound      = c("2020-04-01","2021-12-31"),
 post_pandemic = c("2022-01-01","2022-12-31")
)

label_map <- c(
 entire_sample = "Entire Sample",
 pre_pandemic  = "Pre-Pandemic (Jan 2018–Dec 2019)",
 crash       = "Crash (Jan–Mar 2020)",
 rebound      = "Rebound (Apr 2020–Dec 2021)",
 post_pandemic = "Post-Pandemic (Jan–Dec 2022)"
)

# helper to label phases for VOL summaries (no "Other")
```

```r
phase_label <- function(x) {
  d <- as.Date(x)
  dplyr::case_when(
    d >= as.Date("2018-01-01") & d <= as.Date("2019-12-31") ~ "Pre_2018–Dec19",
    d >= as.Date("2020-01-01") & d <= as.Date("2020-03-31") ~ "Crash_Jan–Mar20",
    d >= as.Date("2020-04-01") & d <= as.Date("2021-12-31") ~ "Rebound_Apr20–Dec21",
    d >= as.Date("2022-01-01") & d <= as.Date("2022-12-31") ~ "Post_2022",
    TRUE ~ NA_character_
  )
}

# -----------------------------
# 3) Core functions
# -----------------------------
# Build the double sort & VOL factor for any volatility window (fixed to 2018–2022 for baseline)
build_sort_for_window <- function(data_in, vol_months = 3) {
  d <- data_in %>%
    arrange(PERMNO, Names_Date) %>%
    group_by(PERMNO) %>%
    mutate(Volatility = rollapply(Returns, width = vol_months, FUN = sd,
                                  align = "right", fill = NA)) %>%
    ungroup() %>%
    filter(!is.na(Volatility),
           Names_Date >= as.Date("2018-01-01"),
           Names_Date <= as.Date("2022-12-31"))

  # NYSE median breakpoints for Size
  nyse_bp <- d %>% filter(Exchange_Code == 1) %>%
    group_by(Names_Date) %>%
    summarise(NYSE_Median_MC = median(Market_Cap, na.rm = TRUE), .groups = "drop")

  d <- d %>%
    left_join(nyse_bp, by = "Names_Date") %>%
    mutate(Size_Portfolio = ifelse(Market_Cap < NYSE_Median_MC, "Small","Large")) %>%
    group_by(Names_Date, Size_Portfolio) %>%
    mutate(Volatility_Quintile = ntile(Volatility, 5)) %>%
    ungroup() %>%
    mutate(Portfolio  = paste(Size_Portfolio, paste0("Vol", Volatility_Quintile), sep = "-"),
           Size_Group = ifelse(grepl("^Small", Portfolio), "Small", "Large"),
           Vol_Q      = as.integer(sub(".*Vol","", Portfolio)))

  # EW/VW portfolio returns (monthly)
  port_ret <- d %>%
    group_by(Names_Date, Portfolio) %>%
    summarise(
      EW_Return = mean(Returns, na.rm = TRUE),
      Total_MC  = sum(Market_Cap, na.rm = TRUE),
      VW_Return = sum(Returns * (Market_Cap/Total_MC), na.rm = TRUE),
      .groups = "drop"
    )

  # VOL (Q1 - Q5) inside each size group
  vol_factor <- d %>%
    group_by(Names_Date, Size_Group, Vol_Q) %>%
    summarise(
      EW = mean(Returns, na.rm = TRUE),
      VW = sum(Returns * (Market_Cap/sum(Market_Cap, na.rm = TRUE)), na.rm = TRUE),
      .groups = "drop"
    ) %>%
    pivot_wider(names_from = Vol_Q, values_from = c(EW, VW), names_prefix = "Q") %>%
    mutate(VOL_EW = EW_Q1 - EW_Q5,
           VOL_VW = VW_Q1 - VW_Q5) %>%
    select(Names_Date, Size_Group, VOL_EW, VOL_VW)
```

```r
  list(panel = d, port_ret = port_ret, vol_factor = vol_factor)
}

# --- identical builder with a flexible date range (for FM/FF5 only) ---
build_sort_for_window_range <- function(data_in, vol_months = 3,
                                        start_date = "2013-01-01",
                                        end_date   = "2022-12-31") {
  d <- data_in %>%
    arrange(PERMNO, Names_Date) %>%
    group_by(PERMNO) %>%
    mutate(Volatility = rollapply(Returns, width = vol_months, FUN = sd,
                                  align = "right", fill = NA)) %>%
    ungroup() %>%
    filter(!is.na(Volatility),
           Names_Date >= as.Date(start_date),
           Names_Date <= as.Date(end_date))

  nyse_bp <- d %>% filter(Exchange_Code == 1) %>%
    group_by(Names_Date) %>%
    summarise(NYSE_Median_MC = median(Market_Cap, na.rm = TRUE), .groups = "drop")

  d <- d %>%
    left_join(nyse_bp, by = "Names_Date") %>%
    mutate(Size_Portfolio = ifelse(Market_Cap < NYSE_Median_MC, "Small","Large")) %>%
    group_by(Names_Date, Size_Portfolio) %>%
    mutate(Volatility_Quintile = ntile(Volatility, 5)) %>%
    ungroup() %>%
    mutate(Portfolio  = paste(Size_Portfolio, paste0("Vol", Volatility_Quintile), sep = "-"),
           Size_Group = ifelse(grepl("^Small", Portfolio), "Small", "Large"),
           Vol_Q      = as.integer(sub(".*Vol","", Portfolio)))

  port_ret <- d %>%
    group_by(Names_Date, Portfolio) %>%
    summarise(
      EW_Return = mean(Returns, na.rm = TRUE),
      Total_MC  = sum(Market_Cap, na.rm = TRUE),
      VW_Return = sum(Returns * (Market_Cap/Total_MC), na.rm = TRUE),
      .groups = "drop"
    )

  vol_factor <- d %>%
    group_by(Names_Date, Size_Group, Vol_Q) %>%
    summarise(
      EW = mean(Returns, na.rm = TRUE),
      VW = sum(Returns * (Market_Cap/sum(Market_Cap, na.rm = TRUE)), na.rm = TRUE),
      .groups = "drop"
    ) %>%
    pivot_wider(names_from = Vol_Q, values_from = c(EW, VW), names_prefix = "Q") %>%
    mutate(VOL_EW = EW_Q1 - EW_Q5,
           VOL_VW = VW_Q1 - VW_Q5) %>%
    select(Names_Date, Size_Group, VOL_EW, VOL_VW)

  list(panel = d, port_ret = port_ret, vol_factor = vol_factor)
}

# Per-period portfolio stats (EW/VW, t-stat, cumulative, Sharpe)
calculate_portfolio_returns <- function(data, start_date, end_date, period_label) {
  filtered_data <- data %>%
    filter(Names_Date >= as.Date(start_date) & Names_Date <= as.Date(end_date))
  portfolio_stats <- filtered_data %>%
    group_by(Portfolio) %>%
    summarise(
```

```
   Average_EW   = mean(EW_Return, na.rm = TRUE),
   T_Stat_EW    = Average_EW / (sd(EW_Return, na.rm = TRUE) / sqrt(n())),
   Cumulative_EW = prod(1 + EW_Return, na.rm = TRUE) - 1,
   Average_VW   = mean(VW_Return, na.rm = TRUE),
   T_Stat_VW    = Average_VW / (sd(VW_Return, na.rm = TRUE) / sqrt(n())),
   Cumulative_VW = prod(1 + VW_Return, na.rm = TRUE) - 1,
   Sharpe_EW    = mean(Excess_EW_Return, na.rm = TRUE) / sd(Excess_EW_Return, na.rm = TRUE),
   Sharpe_VW    = mean(Excess_VW_Return, na.rm = TRUE) / sd(Excess_VW_Return, na.rm = TRUE),
   .groups = "drop"
  ) %>%
  mutate(Period = period_label)
 return(portfolio_stats)
}

# -----------------------------
# 4) FF5 factors (RF-only for portfolio excess; full FF5 for IVOL)
# -----------------------------
ff5_all <- read.csv(ff5_path, skip = 3)
colnames(ff5_all) <- c("Date","Mkt_RF","SMB","HML","RMW","CMA","RF")
ff5_all <- ff5_all %>%
 mutate(
  Date = as.Date(paste0(Date, "01"), format = "%Y%m%d"),
  across(c(Mkt_RF, SMB, HML, RMW, CMA, RF), ~ as.numeric(.)/100),
  Year_Month = format(Date, "%Y-%m")
 )

ff5_RF <- ff5_all %>% select(Year_Month, RF)   # for portfolio excess returns

# -----------------------------
# 5) Pretty slugs for filenames + LaTeX exporters
# -----------------------------
safe_slug <- function(x) gsub("[^A-Za-z0-9]+","_", x)

format_period_tables <- function(all_returns, period_label) {
 dfp <- all_returns %>%
  filter(Period == period_label) %>%
  mutate(Portfolio = factor(Portfolio,
             levels = c(paste0("Large-Vol",1:5), paste0("Small-Vol",1:5)))) %>%
  arrange(Portfolio)
 ew_tbl <- dfp %>%
  transmute(
   Portfolio,
   `Average_EW (T_Stat_EW)` = sprintf("%.2f\\%%\n(%.2f)", Average_EW*100, T_Stat_EW),
   Cumulative_EW = sprintf("%.2f\\%%", Cumulative_EW*100),
   Sharpe_EW    = sprintf("%.2f",   Sharpe_EW)
  )
 vw_tbl <- dfp %>%
  transmute(
   Portfolio,
   `Average_VW (T_Stat_VW)` = sprintf("%.2f\\%%\n(%.2f)", Average_VW*100, T_Stat_VW),
   Cumulative_VW = sprintf("%.2f\\%%", Cumulative_VW*100),
   Sharpe_VW    = sprintf("%.2f",   Sharpe_VW)
  )
 list(ew = ew_tbl, vw = vw_tbl)
}

export_all_period_tables <- function(all_returns, window_tag = "3m") {
 for (nm in names(label_map)) {
  lab  <- label_map[[nm]]
  tbls <- format_period_tables(all_returns, lab)
  slug <- paste0(window_tag, "_", safe_slug(lab))
  print(xtable(tbls$ew,
         caption = paste0("Equal-Weighted Portfolio Returns, ", lab, " (", window_tag, ")"),
```

```r
            label   = paste0("tab:EW_", slug)),
        file = paste0("20aug_EW_", slug, ".tex"),
        include.rownames = FALSE,
        sanitize.text.function = identity)
    print(xtable(tbls$vw,
            caption = paste0("Value-Weighted Portfolio Returns, ", lab, " (", window_tag, ")"),
            label   = paste0("tab:VW_", slug)),
        file = paste0("20aug_VW_", slug, ".tex"),
        include.rownames = FALSE,
        sanitize.text.function = identity)
  }
}

# -----------------------------
# 6) Core runner for a window (baseline, 2018–2022 only)
# -----------------------------
run_pipeline_for_window <- function(res, tag_label = "3m") {
  # Merge RF to compute excess returns
  pr <- res$port_ret %>%
    mutate(Year_Month = format(Names_Date, "%Y-%m")) %>%
    left_join(ff5_RF, by = "Year_Month") %>%
    filter(!is.na(RF)) %>%
    mutate(
      Excess_EW_Return = EW_Return - RF,
      Excess_VW_Return = VW_Return - RF
    )

  # Per-period portfolio tables
  out_list <- lapply(names(periods), function(nm) {
    interval <- periods[[nm]]
    calculate_portfolio_returns(pr, interval[1], interval[2], label_map[[nm]])
  })
  all_returns <- bind_rows(out_list)

  # Export per-period tables
  export_all_period_tables(all_returns, window_tag = tag_label)

  # VOL factor summary by non-overlapping phases
  vol_period <- res$vol_factor %>%
    mutate(Phase = phase_label(Names_Date)) %>%
    filter(!is.na(Phase)) %>%
    group_by(Phase) %>%
    summarise(
      mean_VOL_EW = mean(VOL_EW, na.rm = TRUE),
      t_VOL_EW    = mean_VOL_EW/(sd(VOL_EW, na.rm = TRUE)/sqrt(sum(!is.na(VOL_EW)))),
      mean_VOL_VW = mean(VOL_VW, na.rm = TRUE),
      t_VOL_VW    = mean_VOL_VW/(sd(VOL_VW, na.rm = TRUE)/sqrt(sum(!is.na(VOL_VW)))),
      .groups = "drop"
    ) %>% mutate(Window = tag_label)

  # Save CSVs for record
  write.csv(all_returns, paste0("20aug_returns_", tag_label, ".csv"), row.names = FALSE)
  write.csv(vol_period,  paste0("20aug_vol_period_summary_", tag_label, ".csv"), row.names = FALSE)

  list(all_returns = all_returns,
       portfolio_returns = pr,
       vol_period = vol_period,
       vol_ts = res$vol_factor,
       panel = res$panel)
}

# -----------------------------
# 7) Build 3/6/12 and run (baseline, 2018–2022 only)
```

```r
# ----------------------------
res3  <- build_sort_for_window(data, vol_months = 3)
res6  <- build_sort_for_window(data, vol_months = 6)
res12 <- build_sort_for_window(data, vol_months = 12)

out3  <- run_pipeline_for_window(res3,  "3m")
out6  <- run_pipeline_for_window(res6,  "6m")
out12 <- run_pipeline_for_window(res12, "12m")

# Combined VOL comparison across windows (non-overlapping phases)
vol_compare <- bind_rows(out3$vol_period, out6$vol_period, out12$vol_period)
write.csv(vol_compare, "20aug_vol_period_summary_all_windows.csv", row.names = FALSE)

# Baseline objects (3m) if you need them later for plots/EDA
data_sorted_3m      <- out3$panel
portfolio_returns_3m <- out3$portfolio_returns
vol_factor_3m       <- out3$vol_ts

# ============================================================
# ROBUSTNESS 1) MICRO-CAP EXCLUSION (bottom 10% by NYSE MC)
# ============================================================
nyse_p10 <- data %>%
 filter(Exchange_Code == 1) %>%
 group_by(Names_Date) %>%
 summarise(NYSE_P10_MC = quantile(Market_Cap, 0.10, na.rm = TRUE), .groups = "drop")

apply_microcap_filter <- function(df) {
 df %>% left_join(nyse_p10, by = "Names_Date") %>%
  filter(!is.na(NYSE_P10_MC), Market_Cap >= NYSE_P10_MC) %>%
  select(-NYSE_P10_MC)
}

data_excl10 <- apply_microcap_filter(data)

res3_ex  <- build_sort_for_window(data_excl10, vol_months = 3)
res6_ex  <- build_sort_for_window(data_excl10, vol_months = 6)
res12_ex <- build_sort_for_window(data_excl10, vol_months = 12)

compute_all_returns_only <- function(port_ret) {
 pr <- port_ret %>%
  mutate(Year_Month = format(Names_Date, "%Y-%m")) %>%
  left_join(ff5_RF, by = "Year_Month") %>%
  filter(!is.na(RF)) %>%
  mutate(
   Excess_EW_Return = EW_Return - RF,
   Excess_VW_Return = VW_Return - RF
  )
 out_list <- lapply(names(periods), function(nm) {
  interval <- periods[[nm]]
  calculate_portfolio_returns(pr, interval[1], interval[2], label_map[[nm]])
 })
 bind_rows(out_list)
}

compute_vol_period_only <- function(vol_ts, tag) {
 vol_ts %>%
  mutate(Phase = phase_label(Names_Date)) %>%
  filter(!is.na(Phase)) %>%
  group_by(Phase) %>%
  summarise(
   mean_VOL_EW = mean(VOL_EW, na.rm = TRUE),
   t_VOL_EW    = mean_VOL_EW/(sd(VOL_EW, na.rm = TRUE)/sqrt(sum(!is.na(VOL_EW)))),
   mean_VOL_VW = mean(VOL_VW, na.rm = TRUE),
```

```
    t_VOL_VW    = mean_VOL_VW/(sd(VOL_VW, na.rm = TRUE)/sqrt(sum(!is.na(VOL_VW)))),
    .groups = "drop"
  ) %>% mutate(Window = tag)
}

# Returns + VOL summaries (excl. bottom decile)
all_returns_3m_ex  <- compute_all_returns_only(res3_ex$port_ret)
all_returns_6m_ex  <- compute_all_returns_only(res6_ex$port_ret)
all_returns_12m_ex <- compute_all_returns_only(res12_ex$port_ret)

write.csv(all_returns_3m_ex,  "20aug_returns_3m_excl10.csv",  row.names = FALSE)
write.csv(all_returns_6m_ex,  "20aug_returns_6m_excl10.csv",  row.names = FALSE)
write.csv(all_returns_12m_ex, "20aug_returns_12m_excl10.csv", row.names = FALSE)

vol_ex_3m  <- compute_vol_period_only(res3_ex$vol_factor,  "3m_excl10")
vol_ex_6m  <- compute_vol_period_only(res6_ex$vol_factor,  "6m_excl10")
vol_ex_12m <- compute_vol_period_only(res12_ex$vol_factor, "12m_excl10")

write.csv(vol_ex_3m,  "20aug_vol_period_summary_3m_excl10.csv",  row.names = FALSE)
write.csv(vol_ex_6m,  "20aug_vol_period_summary_6m_excl10.csv",  row.names = FALSE)
write.csv(vol_ex_12m, "20aug_vol_period_summary_12m_excl10.csv", row.names = FALSE)

vol_ex_all <- bind_rows(vol_ex_3m, vol_ex_6m, vol_ex_12m)
write.csv(vol_ex_all, "20aug_vol_period_summary_all_windows_excl10.csv", row.names = FALSE)

# ===============================================================
# ROBUSTNESS 2) IDIOSYNCRATIC VOLATILITY (FF5 residual IVOL)
# ===============================================================
stock_ff5 <- data %>%
 mutate(Year_Month = format(Names_Date, "%Y-%m")) %>%
 left_join(ff5_all, by = "Year_Month") %>%
 filter(!is.na(RF)) %>%
 arrange(PERMNO, Names_Date) %>%
 mutate(Excess_R = Returns - RF)

compute_ivol_ff5 <- function(df, width = 12) {
 df <- df %>% arrange(Names_Date)
 n  <- nrow(df)
 iv <- rep(NA_real_, n)
 if (n >= width) {
  for (i in width:n) {
   sub <- df[(i-width+1):i, c("Excess_R","Mkt_RF","SMB","HML","RMW","CMA")]
   if (any(!complete.cases(sub))) next
   fit <- try(lm(Excess_R ~ Mkt_RF + SMB + HML + RMW + CMA, data = sub), silent = TRUE)
   if (inherits(fit,"try-error")) next
   iv[i] <- sd(residuals(fit))
  }
 }
 df$IVOL <- iv
 df
}

stock_ivol <- stock_ff5 %>%
 group_by(PERMNO) %>%
 group_modify(~compute_ivol_ff5(.x, width = 12)) %>%
 ungroup()

nyse_bp2 <- stock_ivol %>%
 filter(Exchange_Code == 1) %>%
 group_by(Names_Date) %>%
 summarise(NYSE_Median_MC = median(Market_Cap, na.rm = TRUE), .groups = "drop")

vol_ivol_phase <- stock_ivol %>%
```

```r
  filter(!is.na(IVOL)) %>%
  left_join(nyse_bp2, by = "Names_Date") %>%
  mutate(Size_Group = ifelse(Market_Cap < NYSE_Median_MC, "Small", "Large")) %>%
  group_by(Names_Date, Size_Group) %>%
  mutate(IV_Q = ntile(IVOL, 5)) %>%
  ungroup() %>%
  group_by(Names_Date, Size_Group, IV_Q) %>%
  summarise(
   EW = mean(Returns, na.rm = TRUE),
   VW = sum(Returns * (Market_Cap/sum(Market_Cap, na.rm = TRUE)), na.rm = TRUE),
   .groups = "drop"
  ) %>%
  pivot_wider(names_from = IV_Q, values_from = c(EW, VW), names_prefix = "Q") %>%
  mutate(VOL_EW = EW_Q1 - EW_Q5, VOL_VW = VW_Q1 - VW_Q5) %>%
  mutate(Phase = phase_label(Names_Date)) %>%
  filter(!is.na(Phase)) %>%
  group_by(Phase) %>%
  summarise(
   mean_VOL_EW = mean(VOL_EW, na.rm = TRUE),
   t_VOL_EW    = mean_VOL_EW/(sd(VOL_EW, na.rm = TRUE)/sqrt(sum(!is.na(VOL_EW)))),
   mean_VOL_VW = mean(VOL_VW, na.rm = TRUE),
   t_VOL_VW    = mean_VOL_VW/(sd(VOL_VW, na.rm = TRUE)/sqrt(sum(!is.na(VOL_VW)))),
   .groups = "drop"
  )
write.csv(vol_ivol_phase, "20aug_vol_by_ivol_all_windows.csv", row.names = FALSE)

# ================================================================
# ROBUSTNESS 3) MOMENTUM (MOM6_1) & SHORT-TERM REVERSAL (REV1)
# ================================================================
vol_by_signal_terciles <- function(panel, signal_col, tag_label) {
  df <- panel %>% arrange(PERMNO, Names_Date) %>% group_by(PERMNO) %>% ungroup()
  df <- df %>%
   filter(!is.na(.data[[signal_col]])) %>%
   group_by(Names_Date) %>%
   mutate(Tercile = ntile(.data[[signal_col]], 3)) %>%
   ungroup()

  vol_ts <- df %>%
   mutate(Vol_Q = as.integer(sub(".*Vol","", Portfolio))) %>%
   filter(Vol_Q %in% c(1,5)) %>%
   group_by(Names_Date, Tercile, Vol_Q) %>%
   summarise(
    EW = mean(Returns, na.rm = TRUE),
    VW = sum(Returns * (Market_Cap/sum(Market_Cap, na.rm = TRUE)), na.rm = TRUE),
    .groups = "drop"
   ) %>%
   pivot_wider(names_from = Vol_Q, values_from = c(EW, VW), names_prefix = "Q") %>%
   mutate(VOL_EW = EW_Q1 - EW_Q5,
       VOL_VW = VW_Q1 - VW_Q5)

  vol_phase <- vol_ts %>%
   mutate(Phase = phase_label(Names_Date)) %>%
   filter(!is.na(Phase)) %>%
   group_by(Phase, Tercile) %>%
   summarise(
    mean_VOL_EW = mean(VOL_EW, na.rm = TRUE),
    t_VOL_EW    = mean_VOL_EW/(sd(VOL_EW, na.rm = TRUE)/sqrt(sum(!is.na(VOL_EW)))),
    mean_VOL_VW = mean(VOL_VW, na.rm = TRUE),
    t_VOL_VW    = mean_VOL_VW/(sd(VOL_VW, na.rm = TRUE)/sqrt(sum(!is.na(VOL_VW)))),
    .groups = "drop"
   ) %>% mutate(Window = tag_label, Signal = signal_col)

  vol_phase
```

```
}

add_mom_rev <- function(panel) {
  panel %>%
    arrange(PERMNO, Names_Date) %>%
    group_by(PERMNO) %>%
    mutate(
      # MOM6_1: product of (1+return) from t-7..t-2  => roll on lag(Returns, 2)
      MOM6_1 = rollapply(1 + dplyr::lag(Returns, 2), width = 6,
                         FUN = function(x) { if(any(!is.finite(x))) return(NA_real_); prod(x) - 1 },
                         align = "right", fill = NA),
      # REV1: last month's return
      REV1   = dplyr::lag(Returns, 1)
    ) %>%
    ungroup()
}

res3$panel  <- add_mom_rev(res3$panel)
res6$panel  <- add_mom_rev(res6$panel)
res12$panel <- add_mom_rev(res12$panel)

mom_3  <- vol_by_signal_terciles(res3$panel,  "MOM6_1", "3m")
mom_6  <- vol_by_signal_terciles(res6$panel,  "MOM6_1", "6m")
mom_12 <- vol_by_signal_terciles(res12$panel, "MOM6_1", "12m")

rev_3  <- vol_by_signal_terciles(res3$panel,  "REV1", "3m")
rev_6  <- vol_by_signal_terciles(res6$panel,  "REV1", "6m")
rev_12 <- vol_by_signal_terciles(res12$panel, "REV1", "12m")

write.csv(mom_3,  "20aug_vol_by_momentum_terciles_3m.csv",  row.names = FALSE)
write.csv(mom_6,  "20aug_vol_by_momentum_terciles_6m.csv",  row.names = FALSE)
write.csv(mom_12, "20aug_vol_by_momentum_terciles_12m.csv", row.names = FALSE)

write.csv(rev_3,  "20aug_vol_by_reversal_terciles_3m.csv",  row.names = FALSE)
write.csv(rev_6,  "20aug_vol_by_reversal_terciles_6m.csv",  row.names = FALSE)
write.csv(rev_12, "20aug_vol_by_reversal_terciles_12m.csv", row.names = FALSE)

mom_all <- bind_rows(mom_3, mom_6, mom_12)
rev_all <- bind_rows(rev_3, rev_6, rev_12)
write.csv(mom_all, "20aug_vol_by_momentum_all_windows.csv", row.names = FALSE)
write.csv(rev_all, "20aug_vol_by_reversal_all_windows.csv", row.names = FALSE)

# ================================================================
# ROBUSTNESS 4) SKEWNESS conditioning (fixed)
# ================================================================
sk_fun <- function(x) {
  x <- x[is.finite(x)]
  if (length(x) < 3) return(NA_real_)
  e1071::skewness(x, type = 3)
}

vol_by_skewness <- function(panel, tag_label) {
  df <- panel %>%
    arrange(PERMNO, Names_Date) %>%
    group_by(PERMNO) %>%
    mutate(SK_12m = rollapply(Returns, width = 12,
                              FUN = sk_fun, align = "right", fill = NA)) %>%
    ungroup() %>%
    filter(!is.na(SK_12m)) %>%
    group_by(Names_Date) %>%
    mutate(Sk_Tercile = ntile(SK_12m, 3)) %>%
    ungroup()
```

```
 vol_ts <- df %>%
  mutate(Vol_Q = as.integer(sub(".*Vol","", Portfolio))) %>%
  filter(Vol_Q %in% c(1,5)) %>%
  group_by(Names_Date, Sk_Tercile, Vol_Q) %>%
  summarise(
   EW = mean(Returns, na.rm = TRUE),
   VW = sum(Returns * (Market_Cap/sum(Market_Cap, na.rm = TRUE)), na.rm = TRUE),
   .groups = "drop"
  ) %>%
  pivot_wider(names_from = Vol_Q, values_from = c(EW, VW), names_prefix = "Q") %>%
  mutate(VOL_EW = EW_Q1 - EW_Q5,
      VOL_VW = VW_Q1 - VW_Q5)

 vol_phase <- vol_ts %>%
  mutate(Phase = phase_label(Names_Date)) %>%
  filter(!is.na(Phase)) %>%
  group_by(Phase, Sk_Tercile) %>%
  summarise(
   mean_VOL_EW = mean(VOL_EW, na.rm = TRUE),
   t_VOL_EW   = mean_VOL_EW/(sd(VOL_EW, na.rm = TRUE)/sqrt(sum(!is.na(VOL_EW)))),
   mean_VOL_VW = mean(VOL_VW, na.rm = TRUE),
   t_VOL_VW   = mean_VOL_VW/(sd(VOL_VW, na.rm = TRUE)/sqrt(sum(!is.na(VOL_VW)))),
   .groups = "drop"
  ) %>% mutate(Window = tag_label)

 vol_phase
}

sk3  <- vol_by_skewness(res3$panel,  "3m")
sk6  <- vol_by_skewness(res6$panel,  "6m")
sk12 <- vol_by_skewness(res12$panel, "12m")

write.csv(sk3,  "20aug_vol_by_skewness_terciles_3m.csv",  row.names = FALSE)
write.csv(sk6,  "20aug_vol_by_skewness_terciles_6m.csv",  row.names = FALSE)
write.csv(sk12, "20aug_vol_by_skewness_terciles_12m.csv", row.names = FALSE)

sk_all <- bind_rows(sk3, sk6, sk12)
write.csv(sk_all, "20aug_vol_by_skewness_all_windows.csv", row.names = FALSE)

# ===================================================================
# SUMMARY STATISTICS TABLES (EW & VW) FOR ALL SAMPLE PERIODS
# ===================================================================
summarise_period_stats <- function(start_date, end_date, label_text) {
 dfp <- data %>%
  filter(Names_Date >= as.Date(start_date),
      Names_Date <= as.Date(end_date))

 monthly <- dfp %>%
  group_by(Names_Date) %>%
  summarise(
   EW = mean(Returns, na.rm = TRUE),
   VW = {
    w <- Market_Cap / sum(Market_Cap, na.rm = TRUE)
    sum(Returns * w, na.rm = TRUE)
   },
   .groups = "drop"
  )

 list(
  EW = tibble(
   `Sample Period` = label_text,
   Stocks       = n_distinct(dfp$PERMNO),
   Min         = min(monthly$EW, na.rm = TRUE),
```

```
    Mean         = mean(monthly$EW, na.rm = TRUE),
    Max          = max(monthly$EW, na.rm = TRUE),
    `Std Dev`      = sd(monthly$EW, na.rm = TRUE)
   ),
   VW = tibble(
    `Sample Period` = label_text,
    Stocks       = n_distinct(dfp$PERMNO),
    Min          = min(monthly$VW, na.rm = TRUE),
    Mean         = mean(monthly$VW, na.rm = TRUE),
    Max          = max(monthly$VW, na.rm = TRUE),
    `Std Dev`      = sd(monthly$VW, na.rm = TRUE)
   )
  )
}

stats_list <- lapply(names(periods), function(nm) {
  rng <- periods[[nm]]
  lab <- label_map[[nm]]
  summarise_period_stats(rng[1], rng[2], lab)
})

ew_summary <- bind_rows(lapply(stats_list, `[[`, "EW"))
vw_summary <- bind_rows(lapply(stats_list, `[[`, "VW"))

wb <- createWorkbook()
addWorksheet(wb, "EW")
addWorksheet(wb, "VW")
writeData(wb, "EW", ew_summary)
writeData(wb, "VW", vw_summary)
saveWorkbook(wb, "20aug_summary_stats.xlsx", overwrite = TRUE)

fmt_percent <- function(x) sprintf("%.2f\\%%", x * 100)

ew_tex <- ew_summary %>%
  mutate(
   Min   = fmt_percent(Min),
   Mean  = fmt_percent(Mean),
   Max   = fmt_percent(Max),
   `Std Dev` = fmt_percent(`Std Dev`)
  )

vw_tex <- vw_summary %>%
  mutate(
   Min   = fmt_percent(Min),
   Mean  = fmt_percent(Mean),
   Max   = fmt_percent(Max),
   `Std Dev` = fmt_percent(`Std Dev`)
  )

print(xtable(ew_tex,
         caption = "Summary Statistics of Monthly Universe Returns (Equal-Weighted)",
         label   = "tab:summary_stats_EW"),
    file = "20aug_summary_stats_EW.tex",
    include.rownames = FALSE,
    sanitize.text.function = identity)

print(xtable(vw_tex,
         caption = "Summary Statistics of Monthly Universe Returns (Value-Weighted)",
         label   = "tab:summary_stats_VW"),
    file = "20aug_summary_stats_VW.tex",
    include.rownames = FALSE,
    sanitize.text.function = identity)
```

```
# ==========================
# FM/FF5 SECTION (uses 2013–2022 ONLY FOR ESTIMATION)
# Keeps all 2018–2022 reporting intact above.
# ==========================

# Helper: traded VOL factor from a VOL timeseries (use VW), equal-weight Large & Small
make_traded_VOL <- function(vol_ts) {
  vol_ts %>%
    select(Names_Date, Size_Group, VOL_VW) %>%
    tidyr::pivot_wider(names_from = Size_Group, values_from = VOL_VW) %>%
    transmute(Names_Date, VOL = rowMeans(across(c(Large, Small)), na.rm = TRUE))
}

# Build 3m Size×Vol portfolios over 2013–2022 for FM
res3_FM <- build_sort_for_window_range(data, vol_months = 3,
                                       start_date = "2013-01-01",
                                       end_date   = "2022-12-31")

# Traded VOL factor (3m, 2013–2022)
VOL_3m_FM <- make_traded_VOL(res3_FM$vol_factor)

# Factors matrix (FF5 + VOL), decimals
fact_6 <- ff5_all %>%
  transmute(Year_Month = format(Date, "%Y-%m"),
            MKT = Mkt_RF, SMB = SMB, HML = HML, RMW = RMW, CMA = CMA)

fact_6 <- VOL_3m_FM %>%
  mutate(Year_Month = format(Names_Date, "%Y-%m")) %>%
  select(Year_Month, VOL) %>%
  right_join(fact_6, by = "Year_Month") %>%
  relocate(Year_Month)

# Test assets = 10 VW Size×Vol portfolios (3m) over 2013–2022, in excess returns
portfolio_returns_3m_FM <- res3_FM$port_ret %>%
  mutate(Year_Month = format(Names_Date, "%Y-%m")) %>%
  left_join(ff5_RF, by = "Year_Month") %>%
  filter(!is.na(RF)) %>%
  mutate(Excess_VW_Return = VW_Return - RF)

lhs_wide_FM <- portfolio_returns_3m_FM %>%
  select(Year_Month, Portfolio, Excess_VW_Return) %>%
  tidyr::pivot_wider(names_from = Portfolio, values_from = Excess_VW_Return)

# Merge and keep common months; add Phase (only 2018+ will have phase labels)
fm_panel <- lhs_wide_FM %>%
  inner_join(fact_6, by = "Year_Month") %>%
  mutate(Names_Date = as.Date(paste0(Year_Month, "-01")),
         Phase = phase_label(Names_Date)) %>%
  arrange(Names_Date)

# First pass: rolling betas on (MKT, SMB, HML, RMW, CMA, VOL)
estimate_rolling_betas <- function(df, port_name, window = 60, min_window = 24) {
  y <- df[[port_name]]
  X <- as.matrix(df[, c("MKT","SMB","HML","RMW","CMA","VOL")])
  n <- nrow(df)
  out <- vector("list", n)
  coln <- c("beta_MKT","beta_SMB","beta_HML","beta_RMW","beta_CMA","beta_VOL","resid_var")
  for (t in seq_len(n)) {
    i2 <- t - 1
    i1 <- max(1, i2 - window + 1)
    if (i2 - i1 + 1 >= min_window) {
      y_win <- y[i1:i2]
      X_win <- X[i1:i2, , drop = FALSE]
```

```
   if (all(is.finite(y_win)) && all(is.finite(X_win))) {
    fit <- try(lm(y_win ~ X_win - 1), silent = TRUE)
    if (!inherits(fit, "try-error")) {
     b <- coef(fit)
     if (length(b) == 6 && all(is.finite(b))) {
      e <- residuals(fit)
      out[[t]] <- c(b, var(e, na.rm = TRUE))
     }
    }
   }
  }
  if (is.null(out[[t]])) out[[t]] <- rep(NA_real_, 7)
 }
 bet <- as.data.frame(do.call(rbind, out))
 names(bet) <- coln
 bet$Portfolio  <- port_name
 bet$Year_Month <- df$Year_Month
 bet
}

port_names <- setdiff(names(lhs_wide_FM), "Year_Month")
beta_list <- lapply(port_names, function(p) estimate_rolling_betas(fm_panel, p, 60, 24))
betas_long <- bind_rows(beta_list)

# Second pass: monthly cross-sectional regression y_it on betas_i,t (WLS by 1/resid_var)
# Captures per-month R2 and AdjR2, then attaches phase labels.
estimate_lambdas_each_month <- function(panel_df, betas_df) {
 months <- panel_df$Year_Month
 K <- 6L
 lam_ts <- vector("list", length(months))
 names(lam_ts) <- months

 for (m in seq_along(months)) {
  ym <- months[m]
  y_row <- panel_df[m, port_names]
  y <- as.numeric(y_row[1, ])
  Bm <- betas_df %>% filter(Year_Month == ym) %>%
   arrange(match(Portfolio, port_names))

  keep <- complete.cases(Bm[, paste0("beta_", c("MKT","SMB","HML","RMW","CMA","VOL"))],
              Bm$resid_var, y)
  y_sub <- y[keep]

  if (sum(keep) >= K + 1) {
   X_sub <- as.matrix(Bm[keep, paste0("beta_", c("MKT","SMB","HML","RMW","CMA","VOL"))])
   w <- 1 / Bm$resid_var[keep]
   w[!is.finite(w) | w <= 0] <- median(w[is.finite(w) & w > 0], na.rm = TRUE)
   fit <- lm(y_sub ~ X_sub - 1, weights = w)

   smry <- summary(fit)
   r2    <- unname(smry$r.squared)
   ar2   <- unname(smry$adj.r.squared)
   n_as  <- length(y_sub)

   lam_ts[[m]] <- c(coef(fit), r2, ar2, n_as)
  } else {
   lam_ts[[m]] <- rep(NA_real_, K + 3)
  }
 }

 Lam <- as.data.frame(do.call(rbind, lam_ts))
 names(Lam) <- c("lambda_MKT","lambda_SMB","lambda_HML",
         "lambda_RMW","lambda_CMA","lambda_VOL",
```

```
              "R2","AdjR2","N_assets")
  Lam$Year_Month <- months
  Lam$Names_Date <- as.Date(paste0(months, "-01"))
  Lam$Phase      <- phase_label(Lam$Names_Date)
  Lam
}

lambda_ts <- estimate_lambdas_each_month(fm_panel, betas_long)

# HAC (Newey–West) t-stat for the mean of each lambda series
nw_t_of_mean <- function(x, lag = 4) {
  x <- x[is.finite(x)]
  if (length(x) < 6) {
    return(mean(x, na.rm = TRUE) / (sd(x, na.rm = TRUE) / sqrt(sum(is.finite(x)))))
  }
  fit <- lm(x ~ 1)
  se  <- sqrt(sandwich::NeweyWest(fit, lag = lag, prewhite = FALSE)[1,1])
  as.numeric(coef(fit)[1] / se)
}

summarise_lambda <- function(Lam, col) {
  tibble(
    Factor = col,
    Mean   = mean(Lam[[col]], na.rm = TRUE),
    T_NW   = nw_t_of_mean(Lam[[col]], lag = 4),
    N      = sum(is.finite(Lam[[col]]))
  )
}

overall_tbl <- bind_rows(lapply(c("lambda_MKT","lambda_SMB","lambda_HML",
                                  "lambda_RMW","lambda_CMA","lambda_VOL"),
                                function(cn) summarise_lambda(lambda_ts, cn)))

# by phase (Crash has T=3 → treat with caution)
by_phase_tbl <- lambda_ts %>%
  tidyr::pivot_longer(cols = starts_with("lambda_"),
                      names_to = "Factor", values_to = "Value") %>%
  filter(!is.na(Phase)) %>%
  group_by(Phase, Factor) %>%
  summarise(
    Mean = mean(Value, na.rm = TRUE),
    T_NW = nw_t_of_mean(Value, lag = 2),
    N    = sum(is.finite(Value)),
    .groups = "drop"
  )

# Phase-level average R2 and AdjR2 (from monthly cross-sections)
phase_r2 <- lambda_ts %>%
  filter(!is.na(Phase)) %>%
  group_by(Phase) %>%
  summarise(
    R2    = mean(R2,    na.rm = TRUE),
    AdjR2 = mean(AdjR2, na.rm = TRUE),
    .groups = "drop"
  )

# Attach R2s to the by-phase table
by_phase_tbl <- by_phase_tbl %>%
  left_join(phase_r2, by = "Phase") %>%
  relocate(R2, AdjR2, .after = N)

# Save FM outputs (these do not overwrite earlier 2018–2022 portfolio files)
write.csv(lambda_ts,   "20aug_fm_lambda_timeseries_3m_vw.csv", row.names = FALSE)
```

```
write.csv(overall_tbl, "20aug_fm_lambda_overall_3m_vw.csv",   row.names = FALSE)
write.csv(by_phase_tbl,"20aug_fm_lambda_byphase_3m_vw.csv",    row.names = FALSE)

# ===========================
# End of script
# ===========================
```