Birmingham City
Business School

BIRMINGHAM CITY
University

CAFÉ WORKING PAPER NO. 33

Centre for Applied Finance and Economics (CAFÉ)

The Alchemy of Multibagger Stocks: An empirical investigation of factors that drive outperformance in the stock market

Anna Yartseva

February 2025

The Alchemy of Multibagger Stocks:

An empirical investigation of factors that drive outperformance in the stock market

Anna Yartseva

Birmingham City University (United Kingdom)

Abstract

Hot growth stocks attract substantial interest from investors; however, future stock market winners are difficult to identify using traditional metrics of fundamental analysis (such as cash flow, profitability or earnings per share). This study investigates the characteristics of "multibagger stocks" – stocks that increase in value several times the original investment – and detects key drivers of their abnormal investment returns. An empirical analysis of 464 multibagger stocks listed on major American stock exchanges, each increasing in value by at least tenfold during 2009-2024, was conducted. A dynamic panel data model was developed to explain the sources of their outperformance and to predict future returns. The findings indicate that several traditional Fama-French factors, including size, value and profitability, remain significant predictors of future multibagger returns: small-cap high-value high-profitability stocks outperform. Additionally, the analysis identifies further important drivers of multibagger stock outperformance. These include fundamental, technical, and macroeconomic variables, such as high free cash flow yield, distinctive investment patterns linked to EBITDA growth, complex momentum effects with quick trend reversals that limit optimal entry points, and specific interest rate environment.

This study advances asset pricing research by developing a novel, empirically validated model to explain the multibagger phenomenon. It offers valuable practical insights for investors and asset managers and provides a robust theoretical foundation for future stock screening strategies aimed at identifying potential multibaggers and maximising capital gains.

Keywords: asset pricing; growth stocks; multibagger stocks; stock market outperformance; stock returns modelling; beating the market; Fama-French model; predictive modelling; investment strategies; financial economics.

1. Introduction

Background and context for the study

Stock market investment represents a core component of financial markets and play an important role in life of modern society. Individual investors, asset managers and academic researchers are actively striving to discover investment strategies that generate superior returns. **Despite extensive research on asset pricing and the determinants of stock performance, the systematic identification of future multibagger stocks remains one of the most sought-after yet challenging objective** in the investment analysis. The term "multibagger", popularised by Peter Lynch (1988), refers to a stock that appreciates multiple times over its initial price, usually within a relatively short period, and typically generates returns that significantly exceed market benchmarks[1]. For example, a "two-bagger" describes a stock that doubled in price while a "10-bagger" refers to a stock that increased in value tenfold, from $1 invested to $10.

While the concept of multibaggers is widely recognised and actively used in investment practice, this type of stocks has received little attention from academic researchers. There is a considerable gap in academic literature on robust methods for identifying future multibagger opportunities and understanding determinants of their exponential growth using rigorous econometric methods. Traditional asset pricing theories, whether supporting the Efficient Market Hypothesis (Fama, 1970) or the contrarian Overreaction Hypothesis (De Bondt and Thaler, 1985), tend to focus on conventional return predictors such as firm size, value, profitability, and momentum. More recent studies extend factor models by incorporating more original explanatory factors, including behavioural biases, cognitive errors, and investor emotions (Ren, 2024; Padmavathy, 2024). Despite these advances, academic literature has largely overlooked multibagger stocks as a distinct group: the asset pricing models are typically applied to broad stock markets, aiming to uncover general patterns in stock returns rather than to explains why certain stocks experience extreme capital appreciation and produce exceptional market-beating returns. As a result, the specific factors driving the abnormal multibagger returns remain largely unexplored.

Furthermore, **most insights on multibaggers** originate from investment practitioners rather than academic empirical finance research (Phelps, 1972; Oswal, 2014; Martelli, 2014). While these studies **offer useful heuristics for investment decision-**making (for instance, strong earnings growth or above-average return on capital), they **often lack necessary econometric rigour and empirical validation** of statements made. In addition, prior research has not provided a clear, actionable framework for identifying future multibaggers *ex ante*. Finally, another notable gap is that most studies on multibagger stocks **cover periods only up to 2014**, overlooking structural changes in financial markets over the past decade, such as the explosive growth of disruptive technologies, macroeconomic and geopolitical shocks, which may have altered the drivers of stock market outperformance, making previous insights potentially obsolete and less relevant to the current market environment. Filling these gaps is essential for advancing academic finance and informing practical investment strategies.

Research problem and rationale for the study

The lack of attention from the academic community, insufficient empirical scrutiny and methodological flaws of existing practitioner research limit both theoretical advancements in asset pricing and practical applications for investors seeking to identify high-growth opportunities.

The central research problem this study seeks to address is the absence of a robust, evidence-driven quantitative framework for identifying future multibagger stocks. Understanding the multibagger

[1] The term originates from baseball where it refers to "bags" or "bases" that a player reaches (e.g., single-bagger, double-bagger) that reflect the success of their play.

phenomenon is essential as it both advances financial theory and has significant practical implications. Investors need a systematic method to identify potential future winners among thousands of listed stocks while reducing exposure to speculative investments that fail to maintain returns over the longer term and create undesirable portfolio volatility. Additionally, as financial market conditions constantly evolve, it is useful to verify whether traditional explanatory variables, such as the Fama-French size and value factors, continue to predict future performance reliably.

Research aim and objectives

This study addresses numerous gaps in existing literature by conducting a comprehensive empirical analysis of multibagger stocks listed on major U.S. exchanges over 25 years (from 2000 to 2024), developing a dynamic predictive model of multibagger returns, and identifying key fundamental, technical and macroeconomic determinants of their extraordinary growth. **It aims to uncover the unique characteristics of multibagger stocks that drive their outperformance.**

To achieve this aim, **this study pursues the following research objectives:**

1. To examine the fundamental and technical characteristics that distinguish multibagger stocks that increased in value by at least tenfold during the last 15 years from non-multibaggers.
2. To evaluate the effectiveness of the traditional Fama-French five-factor model in explaining and predicting multibagger stock performance.
3. To develop an enhanced dynamic econometric model that explains multibaggers' performance and predicts their future returns by incorporating novel factors beyond those suggested by the traditional asset pricing theory.
4. To analyse the impact of macroeconomic factors and broader market conditions on multibagger stocks returns and determine whether including macroeconomic variables, such as interest rates, improves the model's predictive accuracy.
5. To generate actionable insights for investors and contribute to asset pricing literature by advancing the understanding of multibagger phenomenon.

The study attempts to answer the following research questions:

1. What unique traits differentiate multibagger stocks from other equities?
2. To what extent do conventional asset pricing factors (size, value, profitability, etc.) help to predict future returns of multibagger stocks?
3. Are there additional variables, beyond those in the Fama-French factor model, that significantly enhance the ability to predict multibagger performance?
4. How do macroeconomic conditions, such as the interest rate environments and business cycles, as well as overall stock market performance, affect multibagger stock returns?
5. What practical insights can be derived from the findings to enable investors to identify potential future multibagger opportunities?

Paper structure

The remainder of this paper is structured as follows. Section 2 reviews the existing literature, summarising key debates on the sources of market-beating stocks returns and focusing on prior research on multibagger stocks. It also highlights the limitations of existing studies on the subject and explains the author's unique contribution to academic literature. Section 3 describes the data sources, sample selection criteria and rationale for the time period chosen for analysis. Sections 4 and 5 present the basic model and the methodology for further model development employed in the study. Section 6 discusses the empirical findings in detail, comparing the performance of the traditional five-factor Fama-French model with more refined static and dynamic frameworks specifically designed to explain and predict multibagger stock performance. Section 7 concludes with key takeaways, limitations, and directions for future research.

2. Literature review

Historic development of literature on multibagger stocks

The first empirical studies on the most profitable investment strategies and stock-picking techniques dated back to the early 1930s. Wyckoff (1931) suggested a method of stock selection based on past price dynamics, volume analysis, and market psychology. His methodology focused on identifying accumulation and distribution phases within stock price cycles and entering positions before major price movements – the ideas that created a basis for modern technical trading. At the same time, Graham and Dodd (1934) laid grounds for fundamental stock market analysis, emphasising the importance of intrinsic value, margin of safety, and financial statement analysis to identify undervalued stocks overlooked by the market that are likely to deliver abnormal returns when their true value is recognised.

Since then, numerous authors attempted to identify the type of stocks that generated market-beating returns for investors, detect their unique features and formulate other methods of successful stock selection. Alternative theoretical approaches and varied empirical evidence led to contradicting conclusions. The supporters of the Efficient Market Hypothesis (originated by Fama, 1970) argued that stock market prices accurately reflect all available information about listed companies and their prospects. Therefore, stocks always trade at their fair value, and it would not be feasible to find investment opportunities with above-average risk-adjusted returns by picking "the best stocks". The advocates of the Overreaction Hypothesis point to plentiful observed cases of market inefficiencies caused by information asymmetries, market psychology, and irrational human behaviour and provide numerous examples of investment strategies that "beat the market": starting from the foundation papers by De Bondt and Thaler (1985), Chopra et al. (1992) and Jegadeesh and Titman (1993) which all demonstrate that the recent history of a stock price movement is useful in predicting future returns and identifying potential outperformers, to recent studies by Singh and Kaur (2024), Zhang and Li (2024) that convincingly show that stocks that experienced extreme recent declines exhibit significant excess returns relative to the market in future periods. As will be shown later, this study provides further empirical evidence in support of the latter contrarian idea and demonstrates how existing market inefficiencies can be exploited for substantial investment gains.

While the active discussion on sources of stock market outperformance produced numerous publications covering various geographic regions and time periods, **a particular group of highest-performing multibagger stocks that generate market-beating returns has received little attention** from the academic research community. There were limited attempts to find a formula for discovering potential future multibaggers using a bottom-up data-driven approach without assuming any theory. Only a few studies explicitly focused of identification of unique traits of the best stocks that outperformed the market for a long time period – for example, studies of 10- and 100-baggers by Phelps (1972), later books and practitioner research by Oswal (2014), Martelli (2014) and Mayer (2018).

The seminal study by Phelps' (1972) focused on 100-baggers – stocks that grow to $100 for every $1 invested. **It analysed the period from 1932 to 1971, listed 365 stocks** and attempted to uncover their common features **using anecdotal examples and case studies.** He suggested searching for small and relatively unknown companies that offer new products and new materials or exploit new production methods – things that help to solve problems and improve humans' lives – with strong earnings growth, potential for further expansion, and sound management practices and holding them for extended periods avoiding overtrading. He summarised:

> *"To make money in the stock market you must have the vision to see them, the courage to buy them and the patience to hold them. Patience is the rarest of all three."* (Phelps, 1972:8).

Phelps' research, although mainly descriptive rather than theoretical, became legendary in the investing community and laid grounds for further applied studies of multibaggers. **Mayer (2018) applied Phelps' methodology to analyse 100-baggers during the later period covering 1962-2014.** He developed Phelps' idea of long-term holding into a **"coffee-can portfolio" approach** where the best stocks are kept for at least 10 years. He also proposed focusing on stocks with the following features:

- Extended periods of earnings growth accompanied by valuation multiples (P/E, P/S etc.) expansion.
- Accelerating rather than steady growth in earnings is highly beneficial.
- High ROE[2] (exceeding 20%).
- Owner operators: talented visionary CEO, high insider ownership.
- Beaten-down and forgotten stocks after they turn around and return to profitability.
- Small cap stocks rather than mega-caps as they have higher chances of becoming multi-baggers.

The idea that **smaller companies might generate higher investment returns** compared to companies with high capitalisation due to the low base effect finds empirical support in many other studies, including well-known work by Fama and French (1993), numerous tests of their factor models using broad market data, and focused examination of the multibagger sample for size patterns by Martelli (2014).

Oswal's "Wealth creation" study (2014) attempted to build on Phelps and Mayer's qualitative insights by conducting a basic statistical analysis of multibagger stocks. Oswal focused on the Indian stock market and identified 47 stocks whose value increased 100-fold during the previous 20 years. The technological sector was identified as the largest wealth-creating sector. The 100-baggers were also found in numerous other areas from pharmaceuticals, banks and consumer retail to auto and building materials manufacturers. Oswal found that the Indian stock market itself, represented by the BSE Sensex index, was a 100-bagger too: its value increased 100-fold over 27 years between 1979 and 2006 with a CAGR of almost 19%. The average 100x period (i.e., time to achieve 100-fold returns) in India was found to be around 12 years (equivalent to 47% CAGR return) – significantly shorter than in developed markets (26 years on average according to Mayer, 2018). Oswal recommended concentrating attention on small and relatively unknown companies with sustainable high growth in earnings and quality management which were trading at low single-digit P/E, calling **his investment philosophy "QGLP" (Quality, Growth, Longevity, at reasonable Price).** According to his analysis, **the 100x phenomenon required both growth in earnings and expansion in valuation ratios.** Oswald's study concluded that to earn life-changing returns in the stock market, an investor should search for "growth in all dimensions – sales, margin and valuation" (Oswal, 2014:7).

Similar ideas were promoted by the famous investor Peter Lynch who managed the world's most profitable investment fund Fidelity Magellan. Lynch (1988, 1993) advocated the idea of "growth at reasonable prices" which allowed him to identify numerous multibaggers and grow assets under his

---

[2] List of all abbreviations used in this text is provided in the Appendix.

management from $18 million to $14 billion with an average annual return of 28.1% vs a market average of 9.1%. Using case studies and anecdotal evidence from his extensive investment practice, Lynch illustrates how these factors have historically contributed to significant outperformance of stocks in his portfolio.

The existing publications in the domain appear to agree that a multibagger is created via the alchemy of the following elements:

1. Size: company should be small and relatively unknown.

- Size is a key driver of the low base effect to enable substantial future growth in market capitalisation. Company must be small both in terms of market cap and sales volume.
- Analyst coverage and institutional holdings should be low, providing a chance to buy stocks below their intrinsic value. As a stock becomes popular, the market recognises its future growth potential and factors it into the stock price, thus, limiting future returns.
- Relatively low trading volumes that provide further mispricing opportunities.

2. High quality of business and management team.

- Proven business model required; company must be able to generate high ROE/ROCE relative to industry average.
- Wise capital allocation decisions: ability to reinvest at ROIC well above market average for exceedingly long time with the potential to compound growth over a sufficient period to create abnormally outsized returns.
- Low intensity of competition and industry tailwinds are highly desirable. Company must be growing to become a market leader (among top 3) in their respective business.
- Company must have a “moat” (as per Warren Buffett) – i.e., sustainable competitive advantage and ability to protect its competitive position from potential threats.
- Asset-light business model is advantageous, as it allows a company to avoid significant maintenance CAPEX commitments.
- A close alignment of management priorities with shareholder interests is necessary to convert company growth into share price growth.

3. Growth in all its dimensions: sales, cash flow, profit margins, valuation multiples.

- Earnings per share (EPS) growth is an absolute must and non-negotiable.
- EPS growth should preferably be combined with growth in ROE.

4. Longevity of growth:

- Company must have a large growth pathway ahead: high addressable market and low current market penetration with numerous opportunities to expand operations.
- Growth must be consistent across economic and market cycles (usually implying non-cyclical business).

5. Favourable valuation at time of purchase: future growth potential must not be fully reflected in the purchase price.

- Low P/E, PEG, and other valuation ratios at entry point.
- Outsized share price growth can be achieved via a combination of growth in two elements: earnings and valuation. This can easily be shown mathematically: as share price (P) can be decomposed into a product of earnings per share (EPS) times market value of each $ of earnings (P/E ratio):

$$Share\ price = Earnings\ per\ share\ \times \frac{Share\ price}{Earnings\ per\ share} = EPS\ \times P/E\ ratio, \tag{1}$$

hence, taking logs and differentiating with respect to time produces:

$$\frac{d}{dt}\ln P = \frac{d}{dt}\ln EPS + \frac{d}{dt}\ln P/E. \tag{1.1}$$

Therefore,

$$\hat{P} = \widehat{EPS} + \widehat{P/E}, \tag{1.2}$$

where hats $\widehat{\ }$ denote growth rates.

- Therefore, valuation multiple expansion to support earnings per share growth is also necessary to achieve outstanding investment returns (occasionally called "twin engines" of share price growth – Mayer, 2015:179).

Limitations of existing studies and author's unique contribution to literature

The consensus among authors suggests that winning stocks that deliver market-beating returns share some common features. However, the existing studies suffer from several limitations.

First, most of the features of multibaggers stocks suggested by previous publications lack rigorous empirical validation. The insights, rather than being derived from methodical quantitative analysis, frequently rely on anecdotal examples or selective case studies that may not be representative of broader market trends. As this paper will demonstrate, when the proposed characteristics are subjected to statistical testing, they frequently fail to hold (for example, most notably, the need for EPS growth which is treated as an axiom by the existing literature). The lack of empirical testing highlights the need for a more systematic data-driven approach to identifying the true determinants of multibagger stock outperformance, which this paper will implement.

Second, when attempts at empirical analysis of existing multibaggers are made, they tend to be predominantly **descriptive rather than analytical.** For instance, Mayer (2018) reports average values for P/E ratio, EPS growth rates, and total returns for multibagger stocks in his sample, while Oswal (2014) classifies firms into lists such as ten largest value creators, ten fastest or the most consistent value creators. While these studies provide retrospective snapshots of past multibagger firms' performance, they do not investigate the underlying factors that contributed to their outstanding stock appreciation. Critically, none of the existing studies in this niche area attempt to employ sophisticated econometric techniques to uncover determinants that drive abnormal stock returns or to test their statistical significance. Without a robust econometric framework, it is unclear whether previously observed characteristics of multibagger stocks are relevant for different market conditions or whether they are simply coincidental. Moreover, reliance on basic descriptive statistics without controlling for other influences makes any causal inferences unreliable, limiting practical applicability of these studies in investment practice. This lack of analytical depth represents a critical shortcoming in existing literature, which will be explicitly rectified in this study.

Third, many existing publications on the subject, both academic or practitioner, **do not provide clear actionable criteria for stock selection, as their proposed multibagger identifying methods are subjective, difficult to quantify, and problematic to implement in practice.** Many authors and investment experts provide qualitative heuristics, but their guidance is often vague. For example, the legendary Peter Lynch recommends choosing "a simple business with a boring name, doing something off-putting" (Lynch, 1989:131) – a fascinating advice, which unfortunately lacks quantifiable parameters and is open to subjective interpretation. Some studies solely rely on the analysis of the subjective statements as their key research method. For instance, Chauhan et al. (2022) attempts to identify the factors influencing the selection of multibaggers stocks and establish their hierarchy of importance based on the analysis of 15 semi-structured interviews with industry experts; however, the inherently subjective nature of this approach and the absence of empirical validation using actual stock market data limit the robustness and generalizability of their findings. Furthermore, frequently mentioned attributes such as "high quality of management" or "wise capital allocation decisions" can only be appraised retrospectively, once share price growth has already occurred, thereby reducing their practical utility for stock selection. Similarly, predicting the longevity of a company's "growth pathway ahead" is inherently challenging, particularly given the current rapid rate of technological advancements. The reliance on subjective judgments and qualitative heuristics in existing literature reveals the need for a quantifiable objective framework to guide investors in selecting potential future multibaggers – an approach this study seeks to develop.

Fourth, existing studies cover various time periods up to 2014 only, leaving a significant gap in the analysis of more recent multibagger stocks. The explosive rise of new industries, such as artificial intelligence, renewable energy, gene therapy, autonomous vehicles, blockchain and digital finance, due to disruptive technological progress during the past decade, suggests that the factors contributing to multibagger performance may have evolved. Moreover, during the recent years the global financial markets have been shaken not only by the unprecedented technological advancements but also by significant macroeconomic and political disruptions, such as the Brexit referendum, COVID-19 pandemic, consequent inflation surge and interest rate hikes, US-China trade wars, Russia-Ukraine war and Middle East conflicts, which all may have caused considerable shifts in investor behaviour. All these factors could have influenced the characteristics of multibagger stocks and drivers of their returns, raising questions whether the existing research based on outdated observations still provides relevant insights for contemporary investors. It is important to examine whether the patterns identified in earlier studies are still valid in the current market environment. This is an obvious gap which this research will address.

Next, several additional articles retrieved using the search term "multibaggers" predominantly consist of low-quality student papers and blog posts published by investment companies. These sources often employ flawed methodologies (for example, suffer from spurious regression issues – see Gunasekaran et al., 2024) and offer unverifiable investment recommendations (Alta Fox Capital, 2021; Wright Research, 2021), rendering them unsuitable for reliable academic analysis.

Furthermore, the majority of available publications, aside from the seminal works of Lynch, Phelps and Mayer, exclusively focus of Indian equities (eg., Oswal, 2014; Chauhan et al., 2022, among others). Consequently, the insights derived from these studies, based on a narrowly defined sample, may have limited applicability to developed stock markets, where regulatory frameworks, macroeconomic conditions, market dynamics, and investor behaviour differ significantly.

This paper will attempt to fill all above-mentioned research gaps. It will explicitly focus on identifying quantifiable features of multibagger stocks that drive their abnormal returns using robust panel data econometric modelling process covering more recent period of 2009-24 that was not examined in previous studies. The findings from this investigation have significant practical value as they can be converted in a practical stock screener which can be used to analyse the existing stock universe and identify companies with similar characteristics with the potential to generate similar multibagger returns in the future.

3. Data

Time period

The analysis in this paper uses data on all companies listed in major American stock exchanges (the NYSE and NASDAQ), including ADRs, sourced from the S&P Capital IQ database. The total share price returns of all listed companies were calculated for a 15-year period (from 1 January 2009 to 1 January 2024). This period was selected for analysis because it begins at the market low immediately following the end of the previous bear market caused by the global financial crisis of 2007-08 (Figure 1). This event effectively "reset" the market, initiating a new market cycle. A 15-year window is sufficiently long to allow high-growth companies to demonstrate their full potential and is commonly used in existing studies. Moreover, excluding earlier market cycles (pre-2009) reduces the impact of legacy high-performers from more traditional industries (e.g., from the dot.com era or earlier) and maintains focus on companies and their characteristics that are more relevant for success in the current market environment.

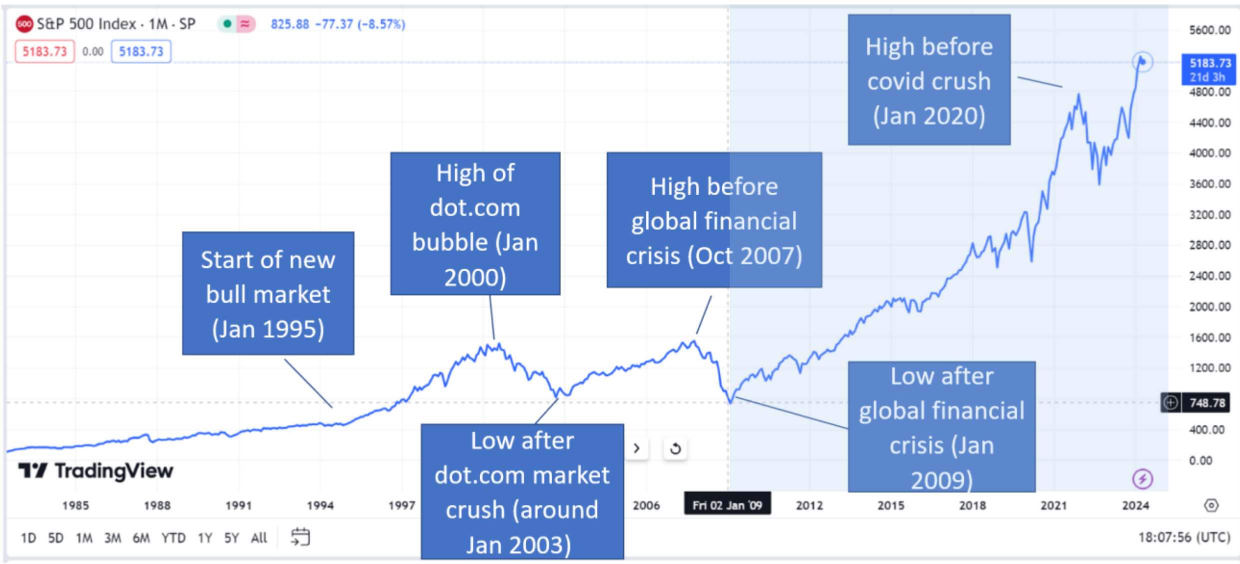

Figure 1. Choice of time period for analysis: the dynamics of S&P 500 index, 1980-2024

Source: tradingview.com

Furthermore, **the selected time period captures a broad range of market conditions and significant economic events, making it highly suitable for analysis.** During this observation period, the U.S. stock market has experienced substantial volatility and has been affected by numerous shocks, including:

- Internal political disturbances: Market-moving U.S. presidential elections (2016, 2020 and 2024).
- Global geopolitical tensions and international events: Brexit referendum (2016), U.S.-China trade war (2018), Russia-Ukraine war (ongoing since 2022), and ongoing conflicts in Israel and the Middle East.
- Commodity price shocks: Sharp oil price declines (2014-16 and 2020) and surges (2022-23); global food price crises (2010-12, 2022-23); precious metals price shocks (2011, 2020).
- Macroeconomic shocks and policy shifts: European debt crisis (2010), U.S. debt ceiling crises (2011, 2023) and a credit rating downgrade (2011), inflation surge (2021), and Federal Reserve emergency interest rates cuts (2020) and hikes (2021).
- Financial sector disruptions: Flash crush (2010), banking crisis (2023), and the approval of Bitcoin ETFs (2024).
- Other global crises: COVID-19 pandemic (2020).

The observation period covers two recessions (2009 and 2020) and consequent recoveries, periods of increasing and declining interest rates (analysed in detail in section 6.5), three bull and three bear markets with periods of S&P 500 index gains of 63-400% and declines of 25-57%, thus, providing an excellent data range for examining stock performance across diverse market conditions.

Sample selection and dataset construction

During the observation period, over five hundred enduring 10-baggers – i.e., stocks that increased in value tenfold[3] or more between 2009 and 2024 and maintained this level at the end of the observation period – were identified. Companies that temporarily achieved tenfold returns but later dropped below the 900% return threshold ("transitory" multibaggers – Oswal, 2014) were excluded. Additionally, firms with missing fundamental data were also removed from the sample.

The resulting panel dataset consists of 464 firms and includes various characteristics of these companies over a 25-year period (1 January 2000 to 1 January 2024). In other words, the dataset also examines the history of these multibaggers preceding their exceptional growth. Selected descriptive statistics for companies in the sample and other descriptive data, including sector distribution and time required to achieve tenfold share price appreciation are presented in the Appendix (Tables A2-A4).

4. Model

The starting point for the analysis is the five-factor model (Fama and French, 2015) which postulates that the expected future stock return is a function of several variables:

$$R_{it} - R_{Ft} = a_i + b_i(R_{Mt} - R_{Ft}) + s_i SMB_t + h_i HML_t + r_i RMW_t + c_i CMA_t + e_{it}, \quad (2)$$

where:

R_{it} is the return on a stock or portfolio *i* for period *t*.

R_{Ft} is the risk-free return, commonly proxied by one- or three-month Treasury bill rate.

R_{Mt} is the market return, proxied by a rate of return on a market index such as the S&P 500.

SMB_t is the size factor: 'Small Minus Big', calculated as the difference between the returns on diversified portfolios of small-cap and big-cap stocks.

HML_t is the value factor: 'High Minus Low', calculated at the difference in returns between high and low book-to-market (B/M) stocks.

RMW_t is the profitability factor: 'Robust Minus Weak', calculated as the difference in returns between stocks with robust and weak profitability.

CMA_t is the investment factor: 'Conservative Minus Aggressive', calculated as the difference in returns between stocks with low and high investment levels.

b_i measures the sensitivity of a stock's return to the overall market return, reflecting its idiosyncratic risk.

s_i, h_i, r_i, c_i measure the relevant factors exposures or payoffs.

a_i is the intercept term that is expected to be zero if the exposures to the five factors fully capture all variation in expected stock returns.

e_{it} is the zero-mean residual.

[3] The tenfold increase in share price is equivalents to 900% share price return. Dividend yield was ignored in this analysis.

In other words, **according to Fama and French**, expected stock returns depend on the performance of the broad market (or market risk premium) and a stock's exposure to size, value, profitability, and investment factors. The model suggests that **firms with smaller size, higher value, stronger profitability, and conservative investment pattern tend to outperform** in the long run.

Empirical tests of the Fama-French five-factor model typically follow a two-step process. First, portfolios are created from independent stock sorts into groups according to each factor and average returns are calculated. This analysis identifies and isolates each factor's premium in stock returns, controlling for other factors. The second step involves estimating the regression model (2), evaluating the significance of individual coefficients, assessing overall model performance, and comparing alternative specifications.

5. Method

This paper builds on the conventional analytical approach described above as a foundation and extends it by developing a more sophisticated dynamic panel model that incorporates factors unique to multibagger stocks. The sequential steps of the model development from the standard Fama-French equation (2) to the proposed model of multibagger stock returns are illustrated in the diagram below.

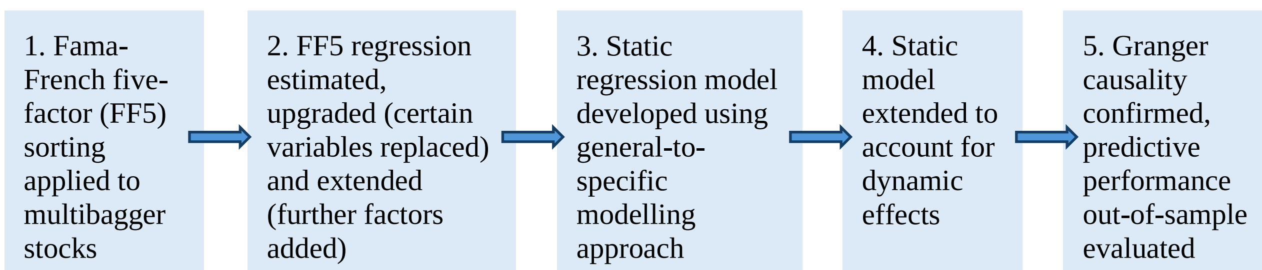

Figure 2. The modelling process: from Fama-French five-factor model to the dynamic panel model

Following the Fama-French methodology, companies in the sample were independently sorted into several quantiles (data on 1 January for the period 2000-24 were used):

- 3 groups by size (Small, Medium and Big) based on market capitalisation data.
- 3 groups by value (Low, Medium and High) using B/M ratios, calculated as total equity/market cap.
- 2 groups by profitability (Robust and Weak) measured as operating profit divided by book equity.
- 2 groups by investment (Conservative and Aggressive), proxied as the annual percentage change in total assets.

The intersection of size, value, profitability, and investment groups is used to create 36 portfolios from 3×3×2×2 individual sorts. The first letter in each portfolio name refers to the size factor, the second letter denotes value group, the third letter reflects profitability, and the fourth letter describes the investment group. For instance, the SHRA portfolio consists of stocks of small-cap companies (S), with high book-to-market value (H), robust operating profitability (R), and aggressive investment strategy (A). The number of groups and portfolios is commonly chosen based on the available sample size to ensure sufficient diversification in the resulting portfolios. Typically, 2-5 groups are created for each factor (see Fama and French, 2017; Foye, 2018, for examples). The sample for the sorting process in this paper includes 10,740 company-years, with created portfolios sizes ranging from 67 to 774 observations, providing an adequate sample size for meaningful statistical analysis.

Future (next year) actual and excess returns above the S&P 500 index are calculated for each stock, and the returns are then aggregated at the portfolio level. The results are reported in Tables 1 and 2 below. Panel A presents annual excess returns along with additional descriptive statistics for each portfolio, providing a clearer picture of the type of stocks sorted into each portfolio. Three additional panels (B, C, and D) report actual (rather than excess) annual price returns, as well as median (rather than mean) returns. A colour-coded scale is used to visualise the extent of portfolio outperformance: greener shades indicate higher returns, while redder shades represent weaker performance. The estimated regression model (equation 2) is presented in Table 3 and discussed in the section 6.2.

Panel A: Next year annual excess returns (mean values)		**Value Low**				**Value Medium**				**Value High**				**Average by size**
		Profitability Weak		Profitability Robust		Profitability Weak		Profitability Robust		Profitability Weak		Profitability Robust		
		Conserv	Aggres	Conserv	Aggres	Conserv	Aggres	Conserv	Aggres	Conserv	Aggres	Conserv	Aggres	
		1	2	3	4	5	6	7	8	9	10	11	12	13
Size Small		SLWC	SLWA	SLRC	SLRA	SNWC	SNWA	SNRC	SNRA	SHWC	SHWA	SHRC	SHRA	
	N	229	131	67	113	280	210	106	216	774	484	173	360	
	Market cap, $m	179	205	306	238	244	259	309	290	211	224	255	211	**$244m**
	BM ratio	-0.28	-0.23	0.10	0.09	0.41	0.40	0.40	0.40	1.43	1.22	2.25	1.76	
	PF, %	-37.1	-26.0	23.9	22.4	-12.9	-4.2	12.5	15.9	-6.9	-3.1	16.8	22.9	
	I growth, %	-11.7	50.8	-3.0	35.5	-7.1	39.7	1.8	49.1	-4.6	41.3	0.4	42.1	
	Annual excess return, %	4.0	48.0	19.5	37.1	-5.9	42.3	15.3	46.1	50.1	57.7	30.4	108.2	**37.7**
Size Medium		MLWC	MLWA	MLRC	MLRA	MNWC	MNWA	MNRC	MNRA	MHWC	MHWA	MHRC	MHRA	
	N	202	143	240	343	389	303	274	390	381	279	127	233	
	Market cap, $m	2,061	2,214	2,290	2,304	1,886	1,936	2,250	1,962	1,736	1,611	1,749	1,824	**$1,985m**
	BM ratio	0.11	0.15	0.14	0.15	0.40	0.40	0.39	0.38	0.88	0.84	0.87	0.85	
	PF, %	-9.9	-8.5	15.0	18.1	-3.3	0.6	12.4	14.9	-0.5	1.3	19.6	21.5	
	I growth, %	-5.0	35.7	-1.3	29.3	-1.8	34.2	0.9	30.3	-1.6	35.0	1.0	40.8	
	Annual excess return, %	-8.8	11.8	1.8	22.2	-0.9	15.2	6.4	24.7	14.4	17.4	27.3	43.1	**14.5**
Size Big		BLWC	BLWA	BLRC	BLRA	BNWC	BNWA	BNRC	BNRA	BHWC	BHWA	BHRC	BHRA	
	N	178	164	657	701	209	149	333	373	102	70	105	105	
	Market cap, $m	26,062	66,815	66,747	59,422	18,825	25,084	23,625	37,193	15,068	11,089	15,199	13,636	**$31,564m**
	BM ratio	0.12	0.13	0.11	0.13	0.39	0.37	0.36	0.35	0.82	0.87	0.78	0.84	
	PF, %	-11.6	-5.5	17.9	20.4	0.4	2.3	16.0	19.1	0.6	2.6	17.6	22.9	
	I growth, %	-3.4	34.0	0.9	25.6	-0.2	31.2	1.4	24.4	-0.1	62.4	2.0	37.1	
	Annual excess return, %	-11.5	13.9	1.8	13.4	0.6	7.7	7.1	15.6	6.3	25.0	11.5	25.2	**9.7**
	Average by I growth	**-5.4**	**24.6**	**7.7**	**24.2**	**-2.1**	**21.7**	**9.6**	**28.8**	**23.6**	**33.4**	**23.1**	**58.8**	
	Average by PF	**9.6**		**16.0**		**9.8**		**19.2**		**28.5**		**40.9**		
	Average by Value	**12.8**				**14.5**				**34.7**				

Panel B: Next year annual excess returns (median values)	**Value Low**				**Value Medium**				**Value High**			
	Profitability Weak		Profitability Robust		Profitability Weak		Profitability Robust		Profitability Weak		Profitability Robust	
	Conserv	Aggres	Conserv	Aggres	Conserv	Aggres	Conserv	Aggres	Conserv	Aggres	Conserv	Aggres
	1	2	3	4	5	6	7	8	9	10	11	12
Size Small	SLWC	SLWA	SLRC	SLRA	SNWC	SNWA	SNRC	SNRA	SHWC	SHWA	SHRC	SHRA
	-20.9	-1.4	-1.5	15.6	-11.3	12.1	7.3	28.6	11.8	28.9	10.7	41.8
Size Medium	MLWC	MLWA	MLRC	MLRA	MNWC	MNWA	MNRC	MNRA	MHWC	MHWA	MHRC	MHRA
	-13.3	2.6	0.7	13.7	-4.6	8.4	3.0	13.6	4.0	9.6	12.5	25.5
Size Big	BLWC	BLWA	BLRC	BLRA	BNWC	BNWA	BNRC	BNRA	BHWC	BHWA	BHRC	BHRA
	-12.9	10.8	0.1	8.1	-0.5	2.3	4.2	11.8	0.9	9.3	9.7	16.6

Table 1. Fama-French five-factor model applied to multibagger stocks (annual excess returns for portfolios generated from 3×3×2×2 sorts)

Panel C: Next year annual price returns (mean values)	Value Low				Value Medium				Value High			
	Profitability Weak		Profitability Robust		Profitability Weak		Profitability Robust		Profitability Weak		Profitability Robust	
	Conserv	Aggres	Conserv	Aggres	Conserv	Aggres	Conserv	Aggres	Conserv	Aggres	Conserv	Aggres
Size Small	SLWC	SLWA	SLRC	SLRA	SNWC	SNWA	SNRC	SNRA	SHWC	SHWA	SHRC	SHRA
	7.3	55.5	24.5	46.9	-5.4	48.2	20.2	54.7	56.4	67.9	40.7	118.6
Size Medium	MLWC	MLWA	MLRC	MLRA	MNWC	MNWA	MNRC	MNRA	MHWC	MHWA	MHRC	MHRA
	-7.7	18.3	7.2	30.6	2.4	24.0	13.4	34.1	22.4	28.1	37.6	56.9
Size Big	BLWC	BLWA	BLRC	BLRA	BNWC	BNWA	BNRC	BNRA	BHWC	BHWA	BHRC	BHRA
	-9.5	23.1	7.9	24.8	3.8	16.4	16.5	28.0	17.8	36.0	22.2	39.5

Panel D: Next year annual price returns (median values)	Value Low				Value Medium				Value High			
	Profitability Weak		Profitability Robust		Profitability Weak		Profitability Robust		Profitability Weak		Profitability Robust	
	Conserv	Aggres	Conserv	Aggres	Conserv	Aggres	Conserv	Aggres	Conserv	Aggres	Conserv	Aggres
Size Small	SLWC	SLWA	SLRC	SLRA	SNWC	SNWA	SNRC	SNRA	SHWC	SHWA	SHRC	SHRA
	-18.1	7.7	7.0	25.7	-4.5	17.0	12.2	40.2	18.9	37.4	18.8	55.4
Size Medium	MLWC	MLWA	MLRC	MLRA	MNWC	MNWA	MNRC	MNRA	MHWC	MHWA	MHRC	MHRA
	-9.4	13.6	8.3	24.1	0.8	18.4	9.8	26.8	12.2	24.0	24.1	41.5
Size Big	BLWC	BLWA	BLRC	BLRA	BNWC	BNWA	BNRC	BNRA	BHWC	BHWA	BHRC	BHRA
	-7.6	18.2	7.8	21.3	5.4	13.1	15.4	24.6	11.8	22.0	22.4	30.6

Table 2. Fama-French five-factor model applied to multibagger stocks (annual observed returns for portfolios generated from 3×3×2×2 sorts)

6. Results and discussion of findings

6.1. Analysis of 3×3×2×2 sorts

Size effect

When controlling for value, profitability and investment factors, the size effect is evident**: small-cap stocks outperform medium and large companies in 11 out of 12 cases,** except for the SNWC, MNWC, and BNWC portfolios in column 5 (Table 1). The average values in column 13 demonstrate this pattern clearly: large firms (with an average market capitalisation of approximately $32 billion) outperform the market by 9.7% annually, mid-sized firms (with an average capitalisation of $2 billion) by 14.5%, while small companies (with a market cap below $250 million) achieve an average excess return of 37.7% per year.

However, when median values are considered instead of means, the results become less conclusive. Panel B (reported in Appendix) shows that four small-cap portfolios (SLWC, SLWA, SLRC, and SNWC) and two medium-sized portfolios (MLWC and MNWC) exhibit negative median excess returns. Some of these portfolios not only trail the market but also experience an actual decline in share prices, generating losses for investors. This suggests that **small-cap classification alone is not a sufficient condition for outperformance**, as other factors have a significant impact on stock returns. In all instances of underperformance mentioned above, the most apparent explanatory factor is the value effect.

Value effect

Companies were sorted into value groups according to their book-to-market ratios, calculated as total equity divided by market capitalisation. **A low book-to-market value (B/M < 1), i.e., low equity and relatively high market cap, implies that investors are paying more for a company than its net assets are worth.** Notably, two portfolios within the low-value group (small-cap SLWC and SLWA) include companies with negative equity, meaning that their total liabilities exceed company assets. The average B/M ratio for the low-value group is only 0.06, implying that the intrinsic value of these companies is approximately 6% of what investors are paying for them, confirming their extreme overvaluation.

On the contrary, **a high book-to-market ratio (B/M > 1), i.e., high equity and relatively low market cap, indicates that the book value of a company exceeds the market price of its shares.** The average B/M ratio for portfolios in the high-value group is 1.10, suggesting that these companies are undervalued by the market and their shares trade at a 10% discount relative to their intrinsic value.

For a rational investor, it appears logical to invest in stocks offering strong fundamental value and avoid or sell overvalued or negative-equity stocks. The empirical data on excess returns confirm that **the value effect is present among multibagger stocks: within each size group, high-value companies consistently generate superior returns.** As the lowest row of panel A indicates, on average, low-value multibagger stocks outperform the S&P 500 by 12.8% annually, medium-value stocks by 14.5%, while high value portfolios generate 34.7% excess price return annually, demonstrating an obvious positive relationship between B/M value and stock performance.

When controlling for size, profitability, and investment factors, the value effect remains consistently present across all portfolio sorts. For instance, comparing the average excess return for companies with weak profitability and a conservative investment policy (columns 1, 5, and 9) demonstrates a clear trend (the relevant section of panel A is reproduced below for clarity). One can see that low book-to-market firms generate an annual excess return of -5.4%, medium-value firms -2.1%, while high-value firms achieve 23.6%. Companies with robust profitability demonstrate a similar pattern: 7.7%, 9.6% and 23.1% (columns 3, 7, and 11 respectively), demonstrating that higher B/M ratios are associated with greater excess return. The same trend is observed among companies with aggressive investment policies.

	Value Low				Value Medium				Value High			
	Profitability Weak		Profitability Robust		Profitability Weak		Profitability Robust		Profitability Weak		Profitability Robust	
	Conserv	Aggres	Conserv	Aggres	Conserv	Aggres	Conserv	Aggres	Conserv	Aggres	Conserv	Aggres
	1	2	3	4	5	6	7	8	9	10	11	12
Average by I growth	-5.4	24.6	7.7	24.2	-2.1	21.7	9.6	28.8	23.6	33.4	23.1	58.8
Average by PF	9.6		16.0		9.8		19.2		28.5		40.9	
Average by Value	12.8				14.5				34.7			

Although the pattern is not perfectly linear in mean values (e.g., the medium-value SNWC portfolio underperforms the low-value SLWC portfolio), it becomes perfectly consistent when median values are considered instead (see Table 1 Panel B). In other words, when **controlling for other factors, all high-value portfolios consistently outperform medium-value portfolios, which, in turn, outperform low-value company groups**.

It should also be noted that all portfolios which generated negative annual price returns (as shown in columns 1 and 5 of panels C and D) belong to low or medium value groups and suffer from low profitability and lack of investment. According to the descriptive statistics in panel A, a combination of a book-to-market value above 0.40 and positive operating profitability classifies a company into a portfolio with significantly higher chances of positive excess returns and a reduced likelihood of losses for investors. This insight has practical implications, as it can be used to develop effective screens for stock picking.

The sorting process not only identifies which types of stocks have the potential to outperform (characterised by strong size and value factors) but also highlights which stocks to avoid or sell short. The analysis of median values, which prove more insightful than means, reveals that having a low book-to-market value and weak profitability poses a greater risk to small-cap stocks (as shown in column 1 of panel D). For example, portfolios with weak profitability not only underperform the S&P 500, generating negative excess returns, but also experience share price declines of 18.1%, 9.4%, and 7.6% annually for small-, medium-, and large-cap stocks respectively, indicating that the smaller the company, the more severe the losses for investors. Therefore, risk-averse investors should consider avoiding these types of stocks. Alternatively, given the extent of their underperformance, stocks with these characteristics (negative equity with B/M ≤ 0, negative operating profitability, and a small market cap below $200 million) might be considered for short strategies.

Profitability effect

The profitability effect is also present in the sample of multibagger stocks: controlling for other factors, **portfolios with weak profitability generate lower excess returns compared to portfolios with robust profitability** (for example, the averages are 9.6% vs. 16.0% for low B/M groups, 9.8% vs. 19.2% for medium value groups, and 28.5% vs. 40.9% for high-value companies). The same pattern can be seen when comparing other portfolio pairs. For example, portfolios with weak profitability and conservative investment policies generate -5.4%, -2.1%, and 23.6% returns (columns 1, 5, and 9 in panel A, reproduced below) and significantly higher returns of 7.7%, 9.6%, and 23.1% when they exhibit robust profitability (columns 3, 7, and 11). This tendency is observed in 22 out of 27 individual comparisons in panel A (82% of cases).

	Value Low				Value Medium				Value High			
	Profitability Weak		Profitability Robust		Profitability Weak		Profitability Robust		Profitability Weak		Profitability Robust	
	Conserv	Aggres	Conserv	Aggres	Conserv	Aggres	Conserv	Aggres	Conserv	Aggres	Conserv	Aggres
	1	2	3	4	5	6	7	8	9	10	11	12
Average by I growth	-5.4	24.6	7.7	24.2	-2.1	21.7	9.6	28.8	23.6	33.4	23.1	58.8
Average by PF	9.6		16.0		9.8		19.2		28.5		40.9	
Average by Value	12.8				14.5				34.7			

The analysis of median values and actual annual price returns reiterates the previous conclusion. In 7 out of 8 cases of negative excess returns (panel B) and in all 4 out of 4 cases of actual negative price returns (panel D), the companies exhibited weak operating profitability (columns 1, 2, and 5). According to the descriptive statistics data, the mean operating profitability for these eight portfolios amounts to -9.6%, emphasising the importance of avoiding loss-making companies for investors who aim to outperform the market.

Investment effect

Additionally, it is worth noting that **all loss-making portfolios** with negative price returns mentioned above (SLWC, MLWC, BLWC, and SNWC in panels C and D) share another common feature beyond weak profitability – they all **adhere to a conservative investment approach.** Their average year-on-year growth of assets is negative (-6.8% compared to +40.0% for similar companies in the higher investment quantile). In other words, their total assets are shrinking; these companies are not investing enough even to maintain their existing production capabilities, let alone expand their assets and create the foundation for future growth. This underinvestment, likely caused by weak profitability (-17.9% on average), serves as a red flag for stock investors, pinpointing companies to avoid or potentially short sell. This observation highlights the **importance of robust investment in company assets to remain competitive** and potentially deliver high future share price returns, becoming multibaggers, – the finding which contradicts the propositions of the five-factor model.

According to Fama and French (2015), the investment factor coefficient should be negative; that is, a more aggressive investment rate (year-on-year growth of total assets) leads to lower future share price returns. Using data on stocks traded on the NYSE from 1963 to 2013, Fama and French found empirical evidence that the investment factor is statistically significant and particularly important for small-cap stocks, as portfolios of small-cap firms that invest aggressively despite weak profitability, tend to underperform the most. However, in the sample of multibagger stocks analysed in this paper, the pattern is strikingly different.

Controlling for other factors, pairwise comparisons of conservative and aggressive investment portfolios show **higher average returns for companies with higher asset growth.** For example, in panel A, the mean excess returns for portfolios with weak profitability and conservative investment are -5.4%, -2.1%, and 23.6%, compared to 24.6%, 21.7%, and 33.4% respectively for companies with similar profitability and other characteristics but a high investment rate. This pattern is observed across all 24 possible pairwise comparisons of portfolios with varying sizes, values, and profitability levels within the table.

	Value Low				Value Medium				Value High			
	Profitability Weak		Profitability Robust		Profitability Weak		Profitability Robust		Profitability Weak		Profitability Robust	
	Conserv	Aggres	Conserv	Aggres	Conserv	Aggres	Conserv	Aggres	Conserv	Aggres	Conserv	Aggres
	1	2	3	4	5	6	7	8	9	10	11	12
Average by I growth	-5.4	24.6	7.7	24.2	-2.1	21.7	9.6	28.8	23.6	33.4	23.1	58.8
Average by PF	9.6		16.0		9.8		19.2		28.5		40.9	
Average by Value	12.8				14.5				34.7			

This is a distinctive feature of multibagger stocks, not observed in other empirical studies based on less restricted samples of stocks. The persistence of the investment effect in all 100% of cases suggests that **to outperform the market and potentially become a multibagger, companies need to aggressively invest in future growth.**

Summary of key findings from Fama-French sorts

The stock sorting exercise demonstrates that all conventional Fama-French variables considered to be the main drivers of stock returns in existing asset pricing literature – size, valuation, profitability, and investment – play an important role in driving the returns of multibagger stocks.

- **Size effect:** small-cap stocks outperform medium and large companies.
- **Value effect:** companies with a high book-to-market ratio outperform.
- **Profitability effect:** companies with robust profitability outperform.
- **Investment effect:** companies with aggressive investment strategies outperform.

The relative importance of these factors, along with their statistical significance and predictive power, will be tested more formally within the panel regression framework in the next section.

6.2. Regression analysis: original and upgraded five-factor models

Standard Fama-French regression estimation and results

In the next stage, annual data on 464 multibagger companies for the period 2000-2024 (11,600 company-year observations) were used to estimate the pooled regression with panel-corrected errors for the following model:

$$R_{it} = \alpha + \gamma_1 R_{Mt} + \gamma_2 R_{Ft} + \gamma_3 Size_{i.t-1} + \gamma_4 Value_{i,t-1} + \gamma_5 Profitability_{i,t-1} + \gamma_6 Investment_{i,t-1} + \varepsilon_{it}, \quad (3)$$

where:

R_{it} is the annual price return on a stock *i* for period *t* (dividend yield is ignored).

R_{Mt} is market return on the S&P 500 index, and R_{Ft} is risk-free return on the three-month T-bill, both common across companies and varying over time.

$Size, Value, Profitability,$ and $Invesment$ are factors proxied by the log of market cap, B/M value, operating profitability, and year-on-year assets growth, varying across companies and over time.

The actual values of these variables were used rather than differences between the top and bottom quantiles as in the original Fama and French paper (2015) to improve clarity of interpretation. To mitigate potential endogeneity within the set of independent variables, lagged values were used as predictors for future stock returns. Generalized Least Squares (GLS) rather than Ordinary Least Squares (OLS) estimator was used to account for heteroscedasticity in the data. The estimation results are reported in Table 3 below.

All coefficients in the estimated pooled regression are statistically significant and have the expected signs, as discovered during the sorting stage. However, the operating profitability coefficient is close to zero (0.001), implying a minimal impact on future stock returns. The market return coefficient equals 1.82, signalling that multibagger stocks have high CAPM betas, while the risk-free return has a negative coefficient of -2.91, which is expected (the higher the return on risk-free assets, the lower the incentive to take on additional risk).

```
Cross-sectional time-series FGLS regression

Coefficients:  generalized least squares
Panels:        homoskedastic
Correlation:   no autocorrelation

Estimated covariances      =         1          Number of obs     =      9,757
Estimated autocorrelations =         0          Number of groups  =        461
Estimated coefficients     =         7          Obs per group:
                                                              min =          2
                                                              avg =   21.16486
                                                              max =         23
                                                Wald chi2(6)      =    1950.98
Log likelihood             = -58166.49          Prob > chi2       =     0.0000
```

Price_return	Coefficient	Std. err.	z	P>|z|	[95% conf.	interval]
return_SP500	1.825973	.0561582	32.51	0.000	1.715905	1.936041
risk_free	-2.910148	.571251	-5.09	0.000	-4.02978	-1.790517
ln_market_cap						
L1.	-9.593112	.4591462	-20.89	0.000	-10.49302	-8.693202
BM						
L1.	3.612937	.3952995	9.14	0.000	2.838165	4.38771
Oper_Profit						
L1.	.0014991	.0002383	6.29	0.000	.0010321	.0019662
growth_assets						
L1.	.1948994	.0158082	12.33	0.000	.1639158	.225883
_cons	83.93175	3.507912	23.93	0.000	77.05637	90.80713

Table 3. Fama-French five-factor model: GLS pooled regression estimation

The issue with this model lies in the intercept term size. According to Fama and French (2015), if an asset pricing model completely captures expected returns, the estimated intercept term should be indistinguishable from zero. They reject this hypothesis in their own paper, however. In our sample of multibagger stocks, the intercept term is extremely high and strongly statistically significant (83.9 with a p-value of 0.000). Thus, **according to Fama-French criteria, the five-factor model fails to fully capture the expected returns when applied to multibagger stocks.** A significant proportion of share price growth remains unexplained by the proposed factors, suggesting the existence of additional variables that drive the stock returns of these companies. **This is why the conventional five-factor model was modified and extended.** The next section will describe how the inclusion of additional explanatory variables improves model performance.

Upgraded Fama-French factor model estimation and results

In order to improve model fit, **alternative metrics were tested as proxies** for size (market cap, total enterprise value, total assets, total equity, total capital, total sales, and size classification dummy), valuation (book-to-market, price-to-earnings, and price-to-sales ratios), profitability (operating profit margin, net profit margin, EBITDA margin, return on capital, return on equity), and investment (in addition to asset growth, new dummy variables were created by comparing the company's asset growth with EBITDA and free cash flow growth). The models were evaluated based on the individual and joint significance of coefficients, as well as Akaike Information Criterion (AIC) and Schwarz Bayesian Criterion (SBC) (Akaike, 1974; Schwarz, 1978). Both criteria provide a systematic way to compare alternative specifications and select the best-fitting model, accounting for model complexity and the number of parameters. The model with the lowest AIC and SBC was selected as an 'upgraded' version of the traditional Fama-French model for multibagger stocks, as reported in Table 4 below.

Dependent variable: Annual price return	Original FF5 model		Upgraded FF5 model	
Const	190.277	***	127.082	***
Return SP500	1.936	***	1.452	***
Risk-free return	-0.917	***	-0.867	***
ln Market cap(t-1)	-24.527	***		
ln TEV(t-1)			-14.613	***
BM(t-1)	2.695			
PE(t-1)			-0.178	***
Operating profit margin(t-1)	0.002	***		
EBITDA margin(t-1)			0.709	***
Assets growth(t-1)	0.163	***	0.389	***
Inv dummy(t-1)			-22.789	***
Number of observations	9757		7596	
Number of firms	461		437	
F-test (prob>F)	126.980	***	114.940	***
R2 overall	0.132		0.161	
AIC	115,489		84,382	
SBC	115,532		84,430	

*** significant at 1% level

Table 4. Comparison of regression results for the standard Fama-French five-factor model and its upgraded version (fixed effect panel models)

Apart from changes in the set of explanatory variables, alternative functional forms were evaluated for an upgraded panel model (pooled vs. fixed effects vs. random effects). **The optimal choice of functional form was determined using conventional tests.** According to the Breusch and Pagan Lagrangian Multiplier test (prob = 0.186), the null hypothesis of zero variance in the individual error term cannot be rejected, indicating that the random effect model is inappropriate, while pooled regression might be adequate. However, the F-test, which assesses whether all individual effects are zero (prob = 0.000), indicates that individual company dummies are jointly significant, rejecting pooled OLS in favour of the fixed effects model. The Hausman test, with a prob=0.000, also confirms that **the fixed effect model is consistent and preferred** over the random effects model.

As can be seen, **three variables in the new model were replaced with alternative metrics:** total enterprise value (TEV) was used instead of market cap as a measure of size, price-to-earnings ratio (P/E) was used instead of book-to-market as a valuation measure, and EBITDA margin replaced operating profitability. **All are significant at the 1% level, have the expected signs, and reflect the factor effects explained previously.** Controlling for other variables, larger company size reduces future expected returns, while higher profitability and strong asset growth increase expected returns. A high P/E ratio implies that the company is overvalued (investors pay more for company earnings), therefore, it is equivalent to a low B/M value, leading to lower future stock returns (hence, the P/E coefficient is expected to have a negative sign).

Apart from these replacements, **a new *Inv* dummy variable was introduced** (=1 if the asset growth rate exceeds the EBITDA growth rate), which turned out to be highly significant. The estimated coefficient is -22.789, implying that when a company expands its assets at a rate exceeding its EBITDA growth, the stock price return for the following year tends to be 22.8 percentage points lower. In other words, **multibagger stocks exhibit a unique investment pattern that distinguishes them from other stocks**: they must invest aggressively but also require sufficient EBITDA growth to make the investment affordable and sustainable.

These changes in model specification led to noticeable improvements. While R^2 is not directly interpretable as a measure of goodness of fit in panel models, both AIC and SBC information criteria are significantly lower in the upgraded version, reflecting an improved model fit. The coefficient for the profitability factor, which was very close to zero (at 0.002) in the standard model, is now drastically higher (at 0.709) in the revised model, where operating profit margin has been replaced with EBITDA profitability as a regressor. This change implies a more meaningful impact on future stock returns. Additionally, the intercept term is significantly lower, indicating that **the explanatory factors in the upgraded model more effectively capture variations in future stock returns compared to the original version** of the FF5 model.

6.3. Static and dynamic models of multibagger returns: estimation and analysis of results

General-to-specific modelling process

At the next stage, to estimate a more comprehensive model that explains the future returns of multibagger stocks with additional explanatory variables added to the Fama-French set of factors, **Hendry's general-to-specific modelling methodology** was employed. This approach, which has proven to be highly popular in empirical studies, **aims to uncover the optimal dynamic structure of the data without imposing any restrictive assumptions** on what the true model specification might be. (For a detailed theoretical description see Hendry's seminal work (1995) or the later review of available literature by Campos (2005)). This approach offers significant advantages in empirical research, ensuring that the resulting model is parsimonious, data-driven, and statistically robust. This is particularly relevant and effective for a study aimed at identifying the most influential factors driving multibagger stock returns, as **it provides the opportunity to evaluate the effects of the widest range of potential explanatory variables.**

The general-to-specific modelling methodology involves starting with a general model that captures the underlying data generation process and progressively simplifying it to a more specific, parsimonious form without losing essential information. Initially, the regression model is over-parameterised by introducing a generous number of explanatory variables and lags for both dependent and explanatory variables on the right-hand side of the equation. Simplifications of the general model are then conducted through a series of reductions in lag lengths and the exclusion of insignificant variables one by one. In the first stage, the least statistically significant coefficient with the highest p-value is eliminated, and the model is re-estimated. In the second stage, the next variable with the least statistically significant coefficient is removed, and the model is re-estimated again. This process is repeated until a parsimonious model, which contains only a set of statistically significant regressors, exhibits good statistical properties, and remains reasonably stable over time, is obtained. After the elimination process is complete, the variable deletion F-test is implemented to evaluate the overall significance of the excluded variables to ensure against imposing invalid restrictions. The resulting parsimonious model is then used for further analysis, forecasting, hypothesis testing, simulations, and other research purposes.

The variables included in the over-parameterised regression were chosen based on several considerations: theoretical suggestions from existing literature on factors that might drive multibagger stock returns (Phelps, 1972; Mayer, 2018), empirical testing of their *ex-ante* predictive power (Tortoriello, 2008), exploratory analysis of the multibagger dataset, and the strength of the calculated correlation with future stock returns. These considerations led to the selection of the following groups of potential explanatory variables:

- **Earnings growth**: Analysed growth in revenue; gross, operating, net profit, and EBITDA; growth of free cash flow, earnings per share, and similar metrics. Also examined growth in assets, equity, capital, and tangible book value. Both year-on-year short-term growth rates, longer term cumulative growth, and 5-year CAGR rates were considered.
- **Valuation:** Various ratios such as B/M, P/E, P/S, P/B, FCF/P, EV/EBITDA, EV/FCF, EV/sales.

- **Profitability:** ROC, ROE, ROA; gross, net, operating profit margins, EBITDA margin, levered and unlevered FCF margin; and other less conventional metrics.
- **Quality:** Earnings quality (measured as levered FCF / operating income ratio), cash ROIC (FCF / capital ratio) – both reflecting the firm's ability to convert profit recognised in financial statements into actual disposable funds – and the firm's profitability compared to industry averages.
- **Capital allocation:** Dividend yield, debt increase and reduction, new share issuance and buybacks.
- **Indebtedness, liquidity, solvency ('red flags')**: Long-term and total debt to equity ratios, debt / capital, debt cover, EBITDA/interest expense ratio, current and quick ratios, and Altman score.
- **Technical factors:** Momentum (1, 3, 6, 9, 12, 24, and 36-month) and 12-month price range.
- **Other variables:** Apart from the conventional risk-free 3-month T-bill rate and market return on the S&P 500 index, analysed the impact of the interest rate environment (Fed rate and its changes), business cycle stages (reflected by dummies), firms' R&D and marketing 'propensity' (measured as the proportion of profit spent on R&D or marketing accordingly), various investment dummies (e.g., those indicating whether asset growth exceeds cash flow or EBITDA growth), analyst coverage (to test the common belief that a multibagger company must be relatively unknown), various comparisons with prime industry metrics, and time effects.

Overall, the impact of more than 150 variables and their lags on multibagger returns was examined. The dependent variable in all models is the annual risk-adjusted stock price return (measured here as stock price return minus risk-free return). All models were estimated using data from 2000 to 2022. Two years of observations (2023-2024) were reserved to evaluate the models' out-of-sample predictive power, using models' root mean squared error (RMSE), mean absolute error (MAE), and Chow's second predictive failure test (Chow, 1960).

Preliminary diagnostic tests indicated the presence of heteroscedasticity and first-order autocorrelation in the data. The modified Wald test for groupwise heteroscedasticity in fixed effect regression resulted in prob=0.000, hence the null hypothesis (H0) of homoscedasticity was rejected at 1%; the Wooldridge test for autocorrelation in panel data produced a prob=0.015, hence H0 of no autocorrelation was also rejected at the 5% level. Consequently, the *cluster()* option was employed in Stata code to control for both heteroscedasticity and autocorrelation, ensuring accurate error estimates in both static and dynamic models. The resulting best parsimonious models, which are theoretically sound, have passed diagnostic tests, and exhibit excellent predictive power, are employed for further analysis.

The next subsections explain the differences between static and dynamic specifications and provide a discussion on the most appropriate estimating techniques for the panel dataset utilised in this study.

Static models of future stock returns (panel regressions with fixed effects)

The basic form of the static panel regression can be written as:

$$Y_{it} = \beta_1 X_{1,it-1} + \beta_2 X_{2,it-1} + \dots + \beta_k X_{k,it-1} + \mu_i + \epsilon_{it}, \quad (4)$$

where:

Y_{it} is the annual return on firm *i* stock in year *t*,

$X_{1\dots k,it-1}$ represent exposures to factors that drive stock returns (such as size, value, profitability etc.) for firm *i* at time *t*-1,

$\beta_{1\dots k}$ are the regression coefficients or payoffs to a relevant factor,

μ_i is the unobserved firm-specific fixed effect that captures time-invariant characteristics specific to each company (e.g., corporate culture, visionary CEO, efficient decision-making and so on),

ϵ_{it} is the idiosyncratic error term that captures firm- and time-specific variation not explained by the model's predictors.

Dynamic models of future stock returns: alternative functional forms for large N small T panels

A dynamic model is a model in which the current value of the dependent variable Y_t is a function of a set of independent variables $X_{1\ldots k}$ and its own past values. In other words, an additional lagged dependent variable $Y_{i,t-1}$ that captures the dynamic relationship in introduced on the RHS:

$$Y_{it} = \theta Y_{i,t-1} + \beta_1 X_{1,it-1} + \beta_2 X_{2,it-1} + \ldots + \beta_k X_{k,it-1} + \mu_i + \epsilon_{it}, \tag{5}$$

The inclusion of lags of Y on the right-hand side is appropriate in situations where the time series exhibit inertia as it allows to better capture the dynamics of the adjustment process. This might be particularly relevant for modelling stock returns where the momentum effect is well-documented – see, for example, the seminal paper by Jegadeesh and Titman (1993) or application to mutual funds by Carhart (1997) and other asset classes by Asness et al. (2013). Apart from improving model specification, the dynamic modelling framework enables Granger causality testing to determine whether, after controlling for past values of Y, past values of X help to forecast Y, indicating that X Granger-causes Y (Wooldridge, 2022). This feature is essential for identifying factors that actively drive future stock returns rather than merely correlating with them.

In empirical modelling the equation (5) is transformed to eliminate the unobserved fixed effect μ_i using within transformation or first differencing (Baltagi, 2021). Both transformations have their merits and can be appropriate in different scenarios (depending on a particular panel structure, number of entities within a panel N and time periods T). Two approaches result in a slightly different model structure and require different estimation techniques.

Within (or fixed effects) transformation involves demeaning the variables across time by subtracting the firm-specific means over time for each variable (Wooldridge, 2010). Since term μ_i is constant over time, the difference between μ_i and its mean over observation period $\bar{\mu}_i$ is zero which effectively eliminates this term from the equation:

$$Y_{it} - \bar{Y}_i = \dot{\theta}(Y_{i,t-1} - \bar{Y}_i) + \dot{\beta}_1(X_{1,it-1} - \bar{X}_{1,i}) + \dot{\beta}_2(X_{2,it-1} - \bar{X}_{2,i}) + \ldots + \dot{\beta}_k(X_{k,it-1} - \bar{X}_{k,i}) + (\epsilon_{it} - \bar{\epsilon}_i) \text{ or} \tag{6}$$

$$\tilde{Y}_{it} = \dot{\theta}\tilde{Y}_{i,t-1} + \dot{\beta}_1\tilde{X}_{1,it-1} + \dot{\beta}_2\tilde{X}_{2,it-1} + \ldots + \dot{\beta}_k\tilde{X}_{k,it-1} + \tilde{\epsilon}_{it}, \tag{6.1}$$

where terms $\bar{Y}_i$, $\bar{X}_i$ and $\bar{\epsilon}_i$ denote means of relevant variables for firm i over all time periods. This transformation is preferred when the unobserved individual heterogeneity is assumed to be correlated with the regressors. However, the current consensus is that the within estimator produces biased and inconsistent results for panels with small T (Baltagi, 2022).

The alternative **first differencing (FD)** approach eliminates the unobserved time-invariant effects μ_i by subtracting the previous period's values from the current period's values for each variable (Anderson and Hsiao, 1982), and yields the following equation:

$$Y_{it} - Y_{i,t-1} = \ddot{\theta}(Y_{i,t-1} - Y_{i,t-2}) + \ddot{\beta}_1(X_{1,t-1} - X_{1,t-2}) + \ddot{\beta}_2(X_{2,t-1} - X_{2,t-2}) + \ldots + \ddot{\beta}_k(X_{k,t-1} - X_{k,t-2}) + (\epsilon_{it} - \epsilon_{i,t-1}) \text{ or} \tag{7}$$

$$\Delta Y_{it} = \ddot{\theta}\Delta Y_{i,t-1} + \ddot{\beta}_1\Delta X_{1,it-1} + \ddot{\beta}_2\Delta X_{2,it-1} + \ldots + \ddot{\beta}_k\Delta X_{k,it-1} + \Delta\epsilon_{it}, \tag{7.1}$$

This transformation removes the time-invariant fixed effect μ_i and produces consistent estimates, however, the differenced lag $\Delta Y_{i,t-1}$ on the RHS in model (7.1) introduces potential endogeneity as it is corelated with the error term $\Delta\epsilon_{it}$. The endogeneity problem necessitates the use of instrumental variables estimators, which use deeper lags in differenced form $\Delta Y_{i,t-2} = Y_{i,t-2} - Y_{i,t-3}$ or simply levels $Y_{i,t-2}$ as instruments for $\Delta Y_{i,t-1}$ (as they are uncorrelated with the error term $\Delta\epsilon_{it}$).

Arellano and Bond (1991) proposed a more efficient generalized method of moments procedure **(difference GMM)** than the earlier Anderson and Hsiao (1982) estimator, which has since become highly popular in empirical modelling of dynamic panel data. The Arellano-Bond estimator (the *xtabond* command in Stata) uses lagged levels of the dependent variable as instruments for the first-differenced lagged dependent variable. It is designed for datasets with many panels and few time periods (that is particularly suitable for the multibagger sample under consideration with N=464 and T=25). However, it relies on the assumption of no autocorrelation in the idiosyncratic errors, requiring separate verification. One- and two-step versions of this estimator exists: one-step estimator assumes homoscedasticity, while two-step procedure accounts for heteroskedasticity and autocorrelation making it asymptotically more efficient but requires a larger sample.

The newer **system GMM estimator**, proposed by Arellano and Bover (1995) and fully developed by Blundell and Bond (1998), extends the difference GMM method by combining equations in levels (5) and differences (7) in the system of simultaneous equations. System GMM uses both lagged levels as instruments for the differenced equation as in the Arellano-Bond estimator, and lagged differences as instruments for the level equation (the *xtdpdsys* Stata command). This approach increases the number of moment conditions reducing bias and improving the efficiency of the estimator. It is particularly advantageous when the dependent variable is highly persistent. In other words, when Y has a strong autoregressive component and changes slowly over time, system GMM mitigates the weak instrument problem that affects difference GMM in such cases. Like the original Arellano-Bond estimator, system GMM is recommended for datasets with large number of panel units N and relatively small number of time periods T. This estimator requires an assumption that there is no autocorrelation in the idiosyncratic errors.

Roodman (2009) proposes an alternative version of the system GMM estimator that is also suitable for the multibagger sample under consideration (*xtabond2* Stata command). This approach is designed to deal with cases where idiosyncratic errors are heteroskedastic and correlated within (i.e., over time for each individual firm) but not across panel units. It provides more flexibility in manually specifying range of lags, collapsing them and explicitly classifying variables as strictly exogenous, endogenous or predetermined. The two-step version further improves the estimation efficiency in large samples but added complexity might be not justified for smaller panels.

The choice between these estimators depends on the structure of the data and the assumptions made regarding the error term, which can be rather subjective and not always fully testable. Therefore, all specifications described above were estimated and then compared to identify common inferences. The results are reported in Table 4 below. All models have been tested for the validity of instruments used in the estimation process (using Arellano-Bond autocorrelation test, Sargan / Hansen test of overidentifying restrictions and Difference-in-Hansen test). Postestimation diagnostic tests which can be calculated for dynamic panel models do not provide clear guidance on which specification is “true” or “better” as was the case with more straightforward static models. As information criteria used for model selection cannot be calculated for dynamic IV models, and the R^2 is not interpretable in this context, they are not reported.

Dependent variable: Annual stock price return - rsik-free return	Static models				Dynamic models									
	FE model (levels)		FE model (first differences)		IV FE model (Within transformation, Wooldridge)		IV FD model (First differences transformation, Anderson-Hsiao)		Difference GMM 2-step (Arellano-Bond)		System GMM 2-step (Roodman)		System GMM 1-step (Blundell-Bond)	
Model number	1		2		3		4		5		6		7	
Const	95.609	***	5.710	***	98.708	***	7.090	***	156.896	***	96.592	***	141.395	***
SP500_return_adj	0.933	***	0.696	***	0.924	***	0.537	***	0.753	***	0.903	***	0.833	***
L.Interest_envir	-11.048	***	-1.557						-7.926	***	-12.067	***	-10.335	***
L.ln_TEV	-7.495	***	-41.501	***	-9.012	***	-49.037	***	-17.372	***	-5.275	***	-15.228	***
L.Margin_EBITDA	0.290	**	0.844	**										
L.ROA					0.432	**	1.882	***	1.732	***			1.828	***
L.Earnings_quality					-0.074	**								
L.Assets_growth	0.191	***	0.039		0.082	***					0.242	***		
L.Inv	-11.437	***	-10.408	***	-9.034	***	-5.668	***	-4.738	**	-11.033	***	-4.115	**
L.B/M					31.239	***	40.091	***	41.326	***	7.211	***	42.427	***
L.FCF/P	75.531	***	81.999	***	46.444	***	51.805	***	51.874	***	47.782	***	51.831	***
L.EV/sales					-0.018	**								
L.EV/EBITDA					-0.005	**								
L.Dividend_yield	3.595	*	4.478	*										
L.Price_range	-0.703	***	-0.672	***	-0.921	***	-0.858	***	-0.862	***	-0.896	***	-0.879	***
L.Momentum_1m											0.266	**		
L.Momentum_3m											-0.240	***		
L.Momentum_6m	-0.143	***	-0.288	***	-0.811	***	-0.489	***	-0.454	***	-0.089	**	-0.443	***
L1.Price_return_adj					0.485	***	0.156	***	0.127	***			0.119	***
L2.Price_return_adj									-0.024	***			-0.029	**
Observations	4109		4109		8084		7611		7611		8357		8121	
AIC	42660		41084											
SBC	42724		41146											
Number of instruments < N?					1<443	pass	1<443	pass	240<443	pass	33<443	pass	261<443	pass
Arellano-Bond autocorrelation test: p-values for AR(1)									0.000	pass	0.000	pass	0.000	pass
Arellano-Bond autocorrelation test: p-values for AR(2)									0.837	pass	0.235	pass	0.598	pass
Hansen / Sargan test for overidentifying restrictions											0.004	fail		
Difference-in-Hansen Test (for GMM models)											0.000/0.803	unclear		

Table 5. Comparison of static and dynamic models of future stock returns

* Significant at 10% level

** Significant at 5% level

*** Significant at 1% level

The estimation results include both expected findings and unique insights.

As can be seen from table 5, **all estimated coefficients have expected signs** apart from a single variable (earnings quality) in Model 3. This model (FE Within estimator) contains three additional regressors that do not appear in other specifications: earnings quality, EV/sales and EV/EBITDA valuation ratios. All these additional explanatory variables, although statistically significant, have low estimated coefficients, implying that their role in driving future stock returns is relatively small. These variables were removed from the set of regressors in other specifications as a part of the general-to-specific elimination process as they became statistically insignificant when alternative IV estimation methods were used. The same is true for 1-month and 3-month momentum variables – they turned out to be significant in only one of the specifications (model 6).

The current market return on S&P 500 is significant in all models and has a positive sign, implying that the portfolio of multibagger stocks moves in line with the rest of the market. The estimated coefficient varies from 0.54 to 0.93, which is consistent with the conventional asset pricing theory.

As in previously discussed Fama-French type models, **the size factor proxied by logged TEV is strongly significant**. It appears in all regressions and has a negative coefficient as suggested by Fama-French theory, suggesting that the bigger the size of the company, the lower future stock price growth tends to be (*ceteris paribus*). The coefficient size varies significantly across specifications though suggesting that the extent of its influence on stock returns is less certain.

The profitability factor is also significant. When the dynamic processes are explicitly accounted for, the EBITDA profit margin (that was the preferred profitability metric in the upgraded five-factor model and static models 1-2) becomes statistically insignificant and is replaced by ROA in dynamic models 3-7. None of other variables that could potentially be used as a proxy for profitability or efficiency (such as gross/net/operating profit margin, ROE, ROC, or cash ROIC) were statistically significant in any of the dynamic models. The estimated coefficient is consistent with the theoretical predictions: it suggests that, controlling for other factors, more **profitable companies with higher ROA deliver higher future stock returns**, however, the size of this coefficient is rather small (between 0.4 and 1.9 only). Notably, one of the main explanatory variables with the highest impact on future returns which is found to be strongly significant (FCF/P) can also be interpreted as a measure of profitability.

Growth variables that were tested turned out to be insignificant in the dynamic modelling process. This includes both past-year growth rates and longer-term 5-year CAGR rates. **Growth of EBITDA, EPS, and FCF per share variables are insignificant** and were not included in the final parsimonious models. Thus, the suggestion from the popular literature that to deliver high share price growth, the company must demonstrate significant growth of earnings for extended period, is not supported by the empirical evidence, which is surprising. Growth of assets rate (representing the Fama-French investment factor) is statistically significant in 3 out of 7 specifications, but the estimated coefficient is not high (0.08-0.24) suggesting a limited impact on future returns.

The investment dummy[4], however, is negative and strongly significant in both static and dynamic frameworks. It reveals a **specific investment pattern for multibagger stocks: if the growth of assets exceeds the growth of EBITDA in a particular year, stock returns above risk-free rates next year tend to be 4-11 percentage points lower** (controlling for other factors). In other words, firms must invest and grow their assets; however, the investment must remain affordable and covered by growing EBITDA. This influence appears important for high-performing stocks – this unique finding of this study.

The interest environment dummy[5] turned out to be insignificant in the conventional IV models (models 3 and 4) and was eliminated during general-to-specific modelling process; however, it remained strongly

[4] Inv dummy =1 if year-on-year growth of assets exceeds year-on-year growth of EBITDA, = 0 otherwise.
[5] Interest environment dummy = 1 if the Fed rate is growing in a particular year, =0 otherwise.

significant when more advanced GMM estimators were used. This variable suggests that controlling for other factors, **when interest rates are increasing, this macroeconomic environment depresses the return of multibagger stocks above risk-free rate next year by approximately 8-12 percentage points.**

The negative impact of rising interest rates on growth stocks is well-documented (see, for example, Bernanke and Kuttner, 2005) and straightforward to interpret. The market value of a listed company depends on the present value of its future cash flows. Spikes in interest rates not only increase the cost of capital for firms but also raise the discount rate used in present value calculations, thereby depressing company valuations. This effect is more pronounced in growth ('hot' or 'glamour') stocks, which often rely on promised earnings projected into the distant future, compared to 'value' companies. Consequently, growth stocks are more adversely affected by increases in discount rates. As changes in the interest environment dummy variable affects all stocks in the sample and is not company-specific, it would not be useful for stock selection purposes but can still enhance returns forecasts for high-growth stocks.

The value factor appears to be playing the biggest role in explaining stock returns both in static and dynamic specifications. The value factor is represented by two main variables: book-to-market (B/M) as in Fama-French models and a new FCF-to-price (FCF/P) ratio[6]. These variables have the highest coefficients in absolute terms, positive sign as expected, and strongly statistically significant. They deliver a very clear message that the future price return is strongly related to company valuation. An increase in the company's B/M and FCF/P ratios of 1% is associated with a 7-52% increase in future share price return. Interestingly, this implies that the growth vs value debate in the investment industry might be meaningless: as high-growth stocks must also be value stocks to demonstrate their superiority!

As mentioned earlier, two further valuation variables (EV/sales and EV/EBITDA) turned out to be significant in one of the models (the within estimator) but not in more intricate modelling frameworks. The valuation ratio P/E which is most commonly used in the industry to describe a valuation of the stock, turned out to be not useful in the quantitative empirical analysis and not predictive of future returns. Not only this variable was statistically insignificant when included among regressors, it also tended to skew other coefficients dramatically. **The P/E ratio is problematic for the modelling purposes** for two reasons. First, the company might have negative earnings (i.e., loss-making), making the P/E ratio not interpretable for this period, reducing the number of data points available for analysis. This reduces the sample to profit-making companies only. Secondly, when company earnings are very small, the denominator (E in the ratio) can be close to zero, forcing the ratio itself tend to infinity. These extreme values of P/E cannot be considered outliers as they are valid observations, but their presence make running regression problematic. That is why P/E as a measure of value was avoided in this study.

Technical factors and momentum play a noticeable role in explaining future stock returns. The impact is complex and highly dynamic, implying that multibagger stocks have a *term structure* of expected returns, and the pattern is not as beneficial as commonly assumed by the industry. According to the Momentum Investing idea, share prices exhibit strong persistence over time and tend to follow a trend: the stocks that grew in the recent past will continue to grow in the near future, and similarly, declining stocks tend to underperform in the future (Jegadeesh and Titman, 1993; Asness et al., 2013). The analysis of multibagger shows that this momentum effect (if present) is only short-lived: 1-month momentum is the only variable that has a positive coefficient, and it is only significant in a single model. All other momentum regressors (3- and 6-month) are negative, suggesting a **quick trend reversal** process for multibagger stocks. In other words, if a stock had was growing in the preceding 3-6 months, it is more likely to decline in the next year. **The price range variable[7] is also negative** and strongly statistically significant, indicating that

[6] The FCF/P ratio (also called free cash flow yield) is a valuation, and a profitability metric used to assess the attractiveness of a stock based on its cash-generating ability relative to its share price.

[7] Price range shows how close the current stock price is to its 12-month high and is calculated as (current price – 12-month low) / (12-month high – 12-month low) x 100%. The variable varies from 0 (if the current price is the 12-month lowest) to 100% (if the current price is equal to 12-month high).

the closer the current stock price to its 12-month high, the lower next year's price return tends to be. These findings align with the Overreaction Hypothesis (De Bondt and Thaler, 1985; Zhang and Li, 2024; Singh and Kaur, 2024).

Numerous other variables were tested as a part of the general-to-specific modelling approach but were found to be insignificant. Specifically, **the indebtedness and soundness of financial position** (debt-to-capital ratio, debt cover), **solvency** (Altman score), and **capital allocation decisions** (debt increase, share buybacks, dividend repayments) **were found to be insignificant for future stock returns.** Dividend yield turned out to be significant in static models but not in dynamic framework. This is an interesting finding as it implies that multibagger stocks tend to provide both abnormal capital appreciation and dividend income to investors. In fact, at the beginning of the observation period, 58% of multibagger companies paid dividends, growing to 78% of the sample by January 2024. The author's own **'R&D propensity' variable[8] was also tested:** the hypothesis was that companies that invest a large proportion of available funds into developing new innovative products and ideas should demonstrate higher share price growth; however, this idea was not supported by the empirical evidence.

Granger causality in stock returns

As lags of independent variables $X_{1\ldots k,it-1}$ discussed above turned out to be statistically significant in the explanatory regressions of stock returns Y_{it}, one can **conclude that variables $\boldsymbol{X_{1\ldots k}}$ Granger-cause $\boldsymbol{Y}$**. In other words, the factors identified in this study *drive* future returns of multibagger stocks. They are not simply associated or correlated with increasing returns, they are predictive of future stock performance. The forecasting power of estimated models is discussed in more detail in the next section.

6.4. Predictive power and out-of-sample forecasting

All regressions mentioned above were estimated using the data for 2000-22, with 2023-24 observations reserved for out-of-sample forecasting. The estimated parameters from the training period were then used to predict future stock returns both in-sample and out-of-sample. Forecasting performance was evaluated separately during bull and bear markets, and in different interest rate environments (when interest rates are increasing – the environment that tends to depress growth stocks – and when interest rates are stable or declining). The means of the forecasts are shown in Figure 3 (the out-of-sample prediction period is shaded), and some further forecasting statistics is available in Appendix Table A1.

As can be seen from the graphs of observed and predicted values, **the estimated models trace the direction of market change out-of-sample very well**: both the portfolio share prices decline in 2023 and the consequent growth in 2024 were forecasted with remarkable accuracy. **In-sample forecasting power is also notable**: at the very beginning of the observation period in 2002 when not enough training data were available yet, the model could not pick up the direction of portfolio returns but this quickly improved starting from 2003 onwards.

Interestingly, **the models are overly pessimistic both in times of bear and bull markets**, overstating the extent of predicted portfolio decline (e.g., low in 2023) and underestimating the extent of predicted portfolio rise (for example, peaks in stock performance in 2004 and 2010). All of the models predicted a decline in stock performance in 2021 (which is interesting given that the market turmoil during this period was caused by the pandemic - the "black swan" event which is completely unpredictable by definition. All models were overly pessimistic forecasting between 6.4% to 15.6% mean portfolio returns while the actual return

[8] Measured as a percentage of available cash allocated to R&D expenditure (=R&D expense / Levered FCF x 100%).

amounted to 41.6%. The only year in which the model significantly overstated the portfolio performance is 2013 (the models predicted between 52.5% to 63.9% share price growth while the realised return was 38.4% only). However, this local portfolio high was indeed achieved a year later: in other words, the models were able to foresee the increased returns but the portfolio took one year longer than expected to achieve these. Throughout the whole forecasting exercise, there was not a single year when the models predicted an increase in stock prices while the stocks would fall in reality – which is reassuring for investors.

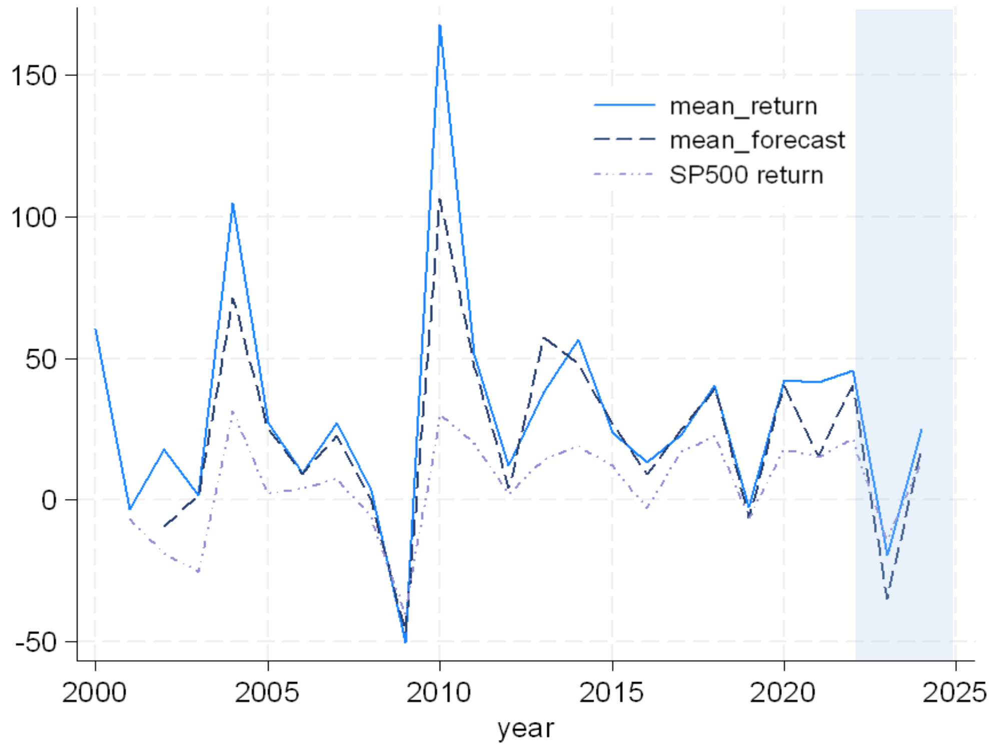

Figure 3. Mean returns of multibagger stock portfolio vs predicted values vs S&P 500 returns

It is also interesting to see how the models' predictive power is affected by the changing macroeconomic interest rate environment. While the estimated models tend to be overly pessimistic (the average forecasting error across all models over all forecasted periods is negative at -6.63%), they are noticeably more pessimistic in a stable or declining interest rates environment (the average forecasting error for these periods increases to -9.92% – see Table A1 in Appendix). When the Fed increases its rate, the models pick the negative effects of higher discount rates on growth stocks very well, and the average forecasting error drops to mere a 1.68% in absolute terms.

To summarise, **the estimated models systematically underestimate the extent of future portfolio performance.** The predictive performance of models is biased, however, the direction of this bias is consistent across all models and all forecasting periods. It still provides valuable information and, in fact, **the direction of this forecasting bias is favourable for investors** who might attempt to use this model in investment decisions. In all cases, the estimated models tend to err on the side of caution, especially during periods of extreme volatility in the markets (periods of extreme highs or extreme lows) predicting lower risk-adjusted returns than actual realised returns, which is arguably a good thing for an investor as it **provides some built-in margin of safety.**

7. Implications and conclusion

Summary of key findings

This paper focuses on a comprehensive analysis of a specific type of stock – multibagger stocks listed on major U.S. stock exchanges that increased in value by at least tenfold from 2009 to 2024. The panel data analysis of 464 multibaggers identified during this period pinpointed several significant factors that explain the sources of their outperformance relative to market averages. The findings indicate that several traditional Fama-French factors, including size, value, and profitability, remain significant determinants of future multibagger returns. Additionally, the analysis identifies other important drivers of stock outperformance. These include fundamental, technical, and macroeconomic variables, such as high free cash flow yield, distinctive investment patterns, complex momentum effects with quick trend reversals, and a specific interest rate environment, which are essential for growth stocks to demonstrate their full potential. A summary of the key findings is provided below:

- **Many common beliefs related to multibagger stocks are not supported by empirical evidence** (for instance, the assumption that strong earnings growth is necessary for significant stock appreciation).
- **Small-cap, high-value, and high-profitability stocks tend to outperform**, supporting the Fama and French (2015) factor investing principles and their applicability to high-growth investment strategies.
- **Aggressive investment is beneficial for stock growth; however, it must be supported by corresponding increases in EBITDA.** An aggressive investment strategy only reduces future returns when a firm expands its asset base at a rate exceeding its earnings growth, indicating a more complex interaction between investment spending and future stock performance than is typically postulated by traditional factor models.
- **Robust cash flow yield is the most important driver of multibagger stock outperformance.**
- Macroeconomic factors, such as interest rates, significantly influence returns; for example, **a rising Fed interest rate depresses the next-year stock returns by 10.1%.**
- **Momentum effects are important**; however, the share price dynamics of multibagger stocks are complex, characterised by **rapid trend reversals.** The closer the current stock price is to its 12-month high, the lower the next-year price return tends to be.
- **The entry point is critical for future returns. Specifically, the stock should be close to its 12-month low at the time of purchase** and, ideally, have fallen in price considerably in the preceding six months.
- **The observed fundamental features of listed companies and their recent performance reliably predict future stock returns**, challenging the Efficient Market Hypothesis.

Contribution to academic literature

The empirical investigation reveals several unique findings that have not been previously published, such as the effects of cash flow yield and the distinctive investment patterns of multibagger companies. This research makes substantial contributions to both academic literature and investment practice. **It advances existing theories is financial economics by focusing on a niche subset of stocks and testing established asset pricing models with novel empirical evidence**, thereby offering a more comprehensive understanding of the factors driving exceptional stock returns. Additionally, **it proposes data-driven ideas for enhancing factor-based investment strategies**.

Refinement of existing asset pricing models and novel return factors: This research confirms that the traditional Fama-French factors – size, value, and profitability – remain significant determinants of future multibagger returns, demonstrating that smaller, undervalued companies with higher profitability metrics typically outperform. However, it goes beyond these traditional models by identifying additional unique explanatory variables: high free cash flow yield, aggressive investment linked to EBITDA growth, stable interest rate environments, and entry stock price close to its 12-month low.

These novel factors are crucial for predicting stock outperformance and systematic stock selection for high-growth portfolios. This integration challenges the conventional wisdom within asset pricing literature, which primarily focuses on surface-level financial metrics. The findings show that **while traditional fundamental factors form a necessary foundation, they alone are insufficient for identifying the highest performing stocks** within the stock universe.

Empirical testing of previous assumptions treated as axioms: Contrary to many practitioner-oriented publications that lack rigorous statistical analysis (such as Phelps (1972), Lynch (1988), and Mayer (2018)), this study employs robust econometric methods and provides statistical evidence that supports the significance of certain investment characteristics and disproves other commonly held beliefs based on intellectual speculation or anecdotal examples. Most notably, it found that earnings growth – in all forms: growth of earnings per share, sales, gross, operating, and net profit, cash flow, both year-on-year and longer-term 5-year cumulative growth, as well as 5-year CAGR rates – was statistically insignificant in predicting future multibagger returns (Section 6.3). These findings echo Tortoriello's (2008) observation that variables effective at explaining past stock performance frequently lose predictive power when modelling future returns. By providing quantitative support for popular qualitative assertions, this study fills a significant gap in the existing literature on multibagger stocks and advances our understanding of the true drivers of superior stock returns.

Inclusion of macroeconomic variables in the modelling framework: This study demonstrates that macroeconomic conditions, specifically, interest rate environments, have a substantial impact on multibagger stock returns – a factor, although well-known, often overlooked in traditional asset pricing models. Interest rates adjustments are the key monetary policy tool used by central banks, which transmit to the economy primarily by influencing costs of borrowing. These changes, in turn, influence incentives to save and invest, the future profitability of the corporate sector, and overall economic activity. Financial markets are at the centre of this transmission mechanism, with equity prices being particularly sensitive to interest rate fluctuations, as demonstrated by Bernanke and Kuttner (2005), among many others. However, factor model studies that explicitly incorporate interest rates or other macroeconomic variables to predict future stock returns are less common (Jensen and Mercer, 2002). By addressing this gap, this study provides deeper insights into how macroeconomic factors influence growth stock returns and reduces the risk of omitted variables bias in traditional multi-factor models.

Refining the impact of the investment factor: Contrary to the traditional Fama-French model (2015), which postulates a negative relationship between active investment and future stock returns, this study finds that aggressive investment can drive multibagger stock growth if it is accompanied by equivalent increases in EBITDA. Aggressive investment is only detrimental to stock returns when it exceeds the firm's financial capacity, highlighting that **affordability is more critical than the aggressiveness of the investment policy** itself. The investment patterns observed in multibagger stocks challenge the conventional propositions of multi-factor models and reveal a more nuanced interaction between investment spending and future stock prices. This could significantly influence how investment decisions are evaluated in both financial theory and practice.

Complex momentum effects and market efficiency: The discovery of complex non-linear dynamic effects, which demonstrate that multibagger stocks exhibit a term structure of returns with rapid trend reversals, challenges the simplistic application of momentum strategies in asset management. This finding indicates that the timing of trades plays a critical role in realising heavily outsized returns and necessitates a more sophisticated approach to stock selection and portfolio formation – one that explicitly identifies advantageous entry points. Moreover, the insights on characteristics distinguishing high-performing stocks provide a basis for developing exit strategies to mitigate the risk of rapid trend reversals, which is crucial for managing the volatility of high-growth stocks. Additionally, the term structure of returns identified in this study questions the extent to which markets efficiently incorporate past price information into current stock prices, thereby contributing to the academic debate on market anomalies.

Implications to investment practice

The insights derived from this study significantly enhance the toolkit available to investors and asset managers seeking to identify potential multibaggers. These findings have substantial implications for investment practice, particularly in the development and refinement of practical investment strategies and systematic stock selection methods.

The excellent forecasting performance of the estimated model, which accurately captures the performance of the multibagger portfolio both in-sample and out-of-sample, demonstrates that the observed fundamental characteristics of listed companies and their recent stock performance can reliably predict future stock returns. These findings explicitly challenge the Efficient Market Hypothesis by suggesting that the **U.S. stock market** does not fully price in publicly available information about listed stocks, thus indicating that it **is not entirely efficient.** This insight is promising for investors seeking alpha, as it implies that **market inefficiencies can be exploited to achieve abnormal returns.** Moreover, these results confirm the **effectiveness of both fundamental and technical analysis as valuable tools for practical investment** decision-making, stock selection, and portfolio management.

Refinement of investment strategies linked to business cycles: The impact of macroeconomic conditions, particularly interest rates, on stock performance reinforces the need for a dynamic asset allocation approach that actively adjusts to economic cycles. Asset managers and individual investors may benefit from altering their portfolio exposure to growth vs. value stocks based on predicted economic conditions, potentially enhancing their return profiles. However, it should be noted that although portfolios with a focus on multibagger stocks experience reduced returns during periods of rising interest rates, they have been shown to outperform the market across all economic conditions (Section 6.4). This suggests that careful selection of stocks based on their fundamental and technical factors can still generate alpha, despite an adverse macroeconomic environment.

Development of a practical stock screener: The dynamic panel data model developed in this study establishes a theoretically sound and empirically validated quantitative framework for devising an effective stock screening strategy, aimed at identifying potential future stock market winners and maximising capital gains. The insights into the factors that drive multibagger returns can be incorporated into a usable stock screener model that surpasses traditional financial metrics, such as those proposed by Piotroski (2000) and Mohanram (2004). By screening for companies that exhibit characteristics similar to historical multibaggers, investors can systematically identify stocks with the potential to yield returns significantly above market averages. The development of such a screener will be the subject of the author's future research in this field. Thus, this research makes a significant contribution to investment practice by bridging the gap between theoretical discussions on asset pricing and practical investment decision-making.

Limitations and directions for further research

Although this study has made significant progress in our understanding of the unique features of multibagger stocks, like all research, it possesses inherent limitations. Numerous intriguing questions remain unaddressed, representing a fruitful field for future investigation that could further enrich both the academic and practical knowledge base, thereby advancing the field of empirical asset pricing.

Global validation of findings: The focus on U.S. stock exchanges might restrict the applicability of the predictive model in other markets, particularly in countries with differing economic systems or regulatory environments. Future studies could investigate whether the drivers of American multibagger returns maintain their significance across global markets, especially in emerging economies.

Sector-specific studies: Given the varying dynamics across different industries, sector-specific studies could provide deeper insights into the drivers of multibagger outperformance within specific industries and market segments. For example, the technology sector may exhibit distinct characteristics that require adjustments to the model. Conversely, industrials, healthcare, or consumer cyclicals might respond to unique drivers of stock performance that are not relevant for tech companies. Furthermore, financials, such as banks and asset management companies, which have distinct balance sheet compositions, require different metrics of fundamental analysis compared to non-financial sectors. Investigating these variances can refine predictive models, more accurately account for industry-specific risks, and tailor investment strategies, potentially boosting portfolio returns.

Impact of disruptive technological innovations: Given the rapid pace of technological transformation that disrupts traditional industries and business practices, the factors currently identified as key drivers of stock market outperformance may evolve over time. As existing companies give way to new market leaders that offer innovative products and services, the significance of traditional metrics, such as asset growth in driving company stock performance, might diminish as firms' operations become increasingly digitalised. Simultaneously, new factors – such as the author's 'R&D propensity' or marketing expenditure – may gain greater explanatory power. Therefore, future longitudinal studies should investigate how technological advancements alter the characteristics of multibagger stocks and periodically update the estimated parameters of the model.

Integration with artificial intelligence methods: Leveraging AI and machine learning techniques could significantly enhance the predictive power of the stock screening model. As Shmueli (2011) explains, there are fundamental differences between explanatory and predictive modelling, leading to completely distinct research paths – from variations in data collection to differing techniques of model validation and optimal model selection, necessitating the use of specific statistical methods tailored to the research aim. Since this study primarily aimed to identify factors that drive (i.e., explain and cause) multibagger stock returns, a dynamic panel regression framework was utilised. This approach was chosen due to ease of interpretation of estimated coefficients and the opportunities for theory building based on the results.

Alternative predictive model building approaches, such as neural networks, random forests, data compression methods, boosting, and ensemble methods, while challenging to interpret, may deliver superior predictive accuracy. If the primary research objective were to forecast rather than explain future multibagger returns, AI algorithms could be trained on a larger dataset to detect subtle patterns and correlations that may not be evident through traditional econometric models. Thus, the application of AI to the multibagger dataset could yield further promising insights.

Use of alternative data sources and inclusion of investor sentiment: Incorporating non-traditional data sources and analytical approaches, such as sentiment analysis from social media and news trends (potentially with the use of AI), along with additional explanatory variables that represent investor psychology, could significantly enhance the model's predictive capabilities and provide a more precise understanding of the factors influencing multibagger stock prices.

The widespread adoption of online investment platforms and mobile apps, such as Robinhood, Webull, Charles Schwab, and Interactive Brokers, after the COVID pandemic has democratised access to financial markets, amplifying the impact of retail investor sentiment on stock prices, particularly noticeable in "glamour", “meme” or "Reddit" stocks. As more individuals engage in stock trading via these platforms[9], the collective mood, emotions (fear/greed/panic), herding influences, and psychological biases captured through unconventional channels may become significant predictors of market movements. This trend towards an increasing role of retail investors underscores the need to include market sentiment regressors in the quantitative modelling framework in future research.

Conclusion

To sum up, this study significantly enriches the field of financial economics by providing new empirical evidence that challenges and refines existing asset pricing theories. The novel incorporation of stock-specific fundamental characteristics, past pricing information, and macroeconomic factors into traditional models offers a more sophisticated understanding of the drivers of extraordinary stock returns. This study not only refines existing asset pricing theories but also lays a solid foundation for future research and the practical application to development of investment strategies, aimed at identifying high-growth opportunities and generating alpha in the stock market.

[9] According to Reuters (2021), individual investors accounted for over 25% of the U.S. equity trading volume in 2020. There were over 100 million retail users at just six of the most popular online brokerages. Furthermore, total client assets at the two leading retail-focused brokerages amounted to $15.5 trillion – compared to total capitalisation of the U.S. stock market of approximately $40.7 trillion (Index Mundi).

8. References

Akaike, H. (1974) A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19 (6): 716-723

Alta Fox Capital (2021) The Makings of a Multibagger. *Alta Fox Capital Research Report*.

Anderson, T.W. and Hsiao, C. (1982) Formulation and estimation of dynamic models using panel data. *Journal of Econometrics*, 18: 47-82.

Ang, R. and Chng, V. (2013) *Value investing in growth companies: How to spot high growth businesses and generate 40% to 400% investment returns.* Singapore: John Wiley & Sons Singapore Pte.

Arellano, M. and Bond, S. (1991) Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. *The Review of Economic Studies*, 58(2): 277-297.

Arellano, M. and Bover, O. (1995) Another look at the instrumental variable estimation of error-components models. *Journal of Econometrics*, 68(1): 29-51.

Asness, C., Moskowitz, T., and Pedersen, L. H. (2013) Value and momentum everywhere. *Journal of Finance*, 68(3): 929-985.

Baltagi, B. H. (2021) *Econometric analysis of panel data*. Springer, 6th ed.

Bernanke, B.S., and Kuttner, K.N. (2005) What explains the stock market's reaction to Federal Reserve policy? *The Journal of Finance*, 60(3): 1221-1257.

Blundell, R. and Bond, S. (1998) Initial conditions and moment restrictions in dynamic panel data models. *Journal of Econometrics*, 87(1): 115-143.

Campos, J., Ericsson, N. and Hendry, D. (2005) General-to-specific modeling: An overview and selected bibliography. *International Finance Discussion Paper*. 838, 1-94.

Carhart, M. M. (1997) On persistence in mutual fund performance. *Journal of Finance*, 52(1): 57-82.

Chan, K. C. (1988) On the contrarian investment strategy. *The Journal of Business*. 61(2): 147-163.

Chopra, N., Lakonishok, J., and Ritter, J. R. (1992) Measuring abnormal performance: Do stocks overreact? *Journal of Financial Economics*, 31(2): 235-268.

Chow, G.C. (1960) Tests of equality between subsets of coefficients in two linear regression models. *Econometrica*, 28: 591-605.

Gunasekaran, S., Rajpoot, V., Kumar, V., Yadav, V. and Kakkar, R. (2024) An empirical study on factors influencing multi-bagger returns in stocks listed in Indian Stock Exchange. *International Journal of Creative Research Thoughts*. 12(2).

De Bondt, W. F. M., and Thaler, R. H. (1985) Does the stock market overreact? *Journal of Finance*. 40(3): 793-805.

Fama, E.F. (1970) Efficient capital markets: A review of theory and empirical work. *The Journal of Finance*, 25(2): 383-417.

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

Fama, E.F. and French, K.R. (2015) A five-factor asset pricing model. *Journal of Financial Economics,* 116: 1-22.

Fama, E.F. and French, K.R. (2017) International tests of a five-factor asset pricing model. Journal of Financial Economics, 123(3): 441-463.

Foye, J. (2018) A comprehensive test of the Fama-French five-factor model in emerging markets. *Emerging Markets Review*, 37: 199-222.

Graham, B. and Dodd, D. (1934) *Security Analysis*. New York: McGraw-Hill.

Hansen, C.K and Schjunk, D.D. (2017) *The stock market's greatest winners: Picking the stocks that move the soonest, fastest and farthest in every bull cycle.* Copenhagen Business School.

Hendry, D. (1995). *Dynamic econometrics.* Oxford: Oxford University Press.

Hou, K., Mo, H., Xue, C., and Zhang, L. (2021) An augmented Q-factor model with expected growth. *Review of Finance*, 25(1): 1-41.

Jegadeesh, N. and Titman, S. (1993) Returns to buying winners and selling losers: Implications for stock market efficiency. *Journal of Finance*, 48(1): 65-91.

Jensen, G.R, and Mercer, J.M. (2002) Monetary policy and the cross-section of expected stock returns. *The Journal of Financial Research*, 25(1): 125-139.

Lynch, P. and Rothchild, J. (1988) *One up on wall street. How to use what you already know to make money in the market.* New York: Simon & Schuster.

Lynch, P. and Rothchild, J. (1993) *Beating the street.* Simon & Schuster.

Martelli, K. (2014) *10x return stocks in the last 15 years.* Martek Partners.

Mayer, C.W. (2018) *100 baggers: Stocks that return 100-to-1 and how to find them.* Laissez-Faire Books.

Mohanram, P.S. (2004) Separating winners from losers among low book-to-market stocks using financial statement analysis [pdf]. Columbia Business School Working Paper. Available at: http://dx.doi.org/10.2139/ssrn.403180 [Accessed 22 February 2025].

Oswal M. (2014) *19th annual wealth creation study (2009-2014). 100x: The power of growth in wealth creation*. Motilal Oswal. Available at http://ftp.motilaloswal.com/emailer/Research/WC19-20141212-MOSL-2009-14-PG048.pdf, accessed 13 February 2025.

Padmavathy, M. (2024) Behavioral finance and stock market anomalies: Exploring psychological factors influencing investment decisions. *Shanlax International Journal of Management*, 11(1): 1-15.

Phelps, T.W. (1972, reprinted 2015) *100 to 1 in the stock market: A distinguished security analyst tells how to make more of your investment opportunities.* Echo Point Books & Media.

Ren, F. (2024) A comprehensive analysis of behavioral finance and its impact on investment decisions. *Highlights in Business, Economics and Management*, 32: 72–77.

Reuters (2021) Factbox: The U.S. retail trading frenzy in numbers, by John McCrank. Available at https://www.reuters.com/article/us-retail-trading-numbers/factbox-the-u-s-retail-trading-frenzy-in-numbers-idUSKBN29Y2PW/ , accessed 21 February 2025.

Roodman, D.M. (2009) A note on the theme of too many instruments. *Oxford Bulletin of Economics and Statistics*, 71: 135-158.

Roodman, D.M. (2009) How to do xtabond2: An introduction to difference and system GMM in Stata. *The Stata Journal*, 9(1): 86-136.

Schwarz, G. (1978) Estimating the dimension of a model. *The Annals of Statistics*, 6(2): 461-464.

Singh, A., and Kaur, P. (2024) Overreaction hypothesis and contrarian profits: Evidence from the emerging markets. *Global Finance Journal*, 50: 1-18.

Shmueli, G. (2010) To explain or to predict? *Statistical Science*, 25(3): 289-310.

Tortoriello, R. (2008) *Quantitative strategies for achieving alpha: The Standard and Poor's approach to testing your investment choices.* McGraw-Hill Education.

Wooldridge, J.M. (2010) *Econometric analysis of cross section and panel data.* MIT Press.

Wooldridge, J.M. (2022) *Introductory econometrics: A modern approach.* Cengage, 7th ed.

Wright Research (2021) The Multibaggers of Momentum. *Wright Research Blog*. Available at https://www.wrightresearch.in/blog/multibaggers-momentum/, accessed 14 February 2025.

Wyckoff, R. D. (1931) The Richard D. Wyckoff method of trading and investing in stocks: A course of instruction in stock market science and technique.

Zhang, Y., and Li, J. (2024) Stock market overreaction and contrarian profits: Evidence from the U.S. market. *Journal of Financial Markets*, 58: 1-22.

9. Appendix

List of abbreviations

ADR	American depositary receipt
AIC	Akaike information criterion
B/M	Book to market (ratio)
BSE	Bombay stock exchange
CAGR	Compound annual growth rate
CAPEX	Capital expenditure
CMA	Conservative minus aggressive (investment factor)
EBITDA	Earnings before interest, taxes, depreciation and amortisation
EV	Enterprise value
FCF	Free cash flow
FE	Fixed effects
FD	First differences
FF5	Fama-French (five-factor model)
GLS	Generalized Least Squares
HML	High minus low (value factor)
MAE	Mean absolute error
OLS	Ordinary Least Squares
P/E	Price to earnings (ratio)
P/S	Price to sales (ratio)
PEG	Price/earnings-to-growth (ratio)
RHS	Right hand side (of equation)
RMSE	Root mean squared error
RMW	Robust minus weak (profitability factor)
ROA	Return on assets
ROCE	Return on capital employed
ROE	Return on equity
ROIC	Return on invested capital
S&P	Standard and Poor's
SBC	Schwarz Bayesian criterion
SMB	Small minus big (size factor)
TEV	Total enterprise value

Table A1. Average forecasting performance in various interest rate environments

Increasing Fed rate	2000	2001	2005	2006	2007	2011	2013	2015	2016	2017	2018	2019	2023	2024	Average
SP500 return	**3.44**	**-6.88**	**2.01**	**3.99**	**7.38**	**19.62**	**14.08**	**11.91**	**-3.04**	**16.95**	**22.48**	**-6.58**	**-14.29**	**13.68**	**6.05**
Actual return	**60.43**	**-3.37**	**27.12**	**9.22**	**26.93**	**51.38**	**38.40**	**23.78**	**13.33**	**22.90**	**40.62**	**-2.70**	**-19.52**	**24.86**	**22.38**
v1 forecast	-	-	25.31	9.21	22.42	47.12	57.64	26.53	8.88	24.93	39.63	-6.40	-34.96	18.62	19.91
v2 forecast	-	-	36.74	12.15	36.28	65.32	63.90	35.00	3.88	21.45	45.35	-10.13	-54.43	5.09	21.72
v3 forecast	-	-	-	-	-	54.06	52.48	27.40	3.44	23.26	42.15	-2.83	-34.59	19.02	20.49
v6 forecast	-	-	31.62	10.07	30.22	54.06	55.53	30.87	8.63	23.12	42.15	-5.42	-39.87	8.43	20.78
v7 forecast	-	-	31.05	10.47	29.92	52.96	54.71	30.07	8.50	23.44	41.66	-4.78	-39.31	8.86	20.63
Av. forecasted return	**-**	**-**	**31.18**	**10.47**	**29.71**	**54.70**	**56.85**	**29.97**	**6.67**	**23.24**	**42.19**	**-5.91**	**-40.63**	**12.00**	**20.71**
Av. forecast error	**-**	**-**	**4.07**	**1.25**	**2.78**	**3.32**	**18.46**	**6.19**	**-6.67**	**0.34**	**1.57**	**-3.21**	**-21.11**	**-12.86**	**-1.68**

Non-increasing Fed rate	2002	2003	2004	2008	2009	2010	2012	2014	2020	2021	2022	Average
SP500 return	**-18.98**	**-25.44**	**31.29**	**-6.02**	**-40.31**	**29.96**	**1.99**	**18.98**	**17.77**	**15.10**	**21.40**	**4.16**
Actual return	**17.56**	**1.43**	**104.82**	**3.30**	**-50.54**	**167.75**	**11.87**	**56.84**	**42.67**	**41.56**	**46.01**	**40.30**
v1 forecast	-9.42	1.60	71.52	-0.17	-46.93	106.32	3.99	48.26	40.44	15.59	40.53	24.70
v2 forecast	8.20	-11.30	99.76	6.85	-45.96	110.48	12.66	61.20	46.38	6.40	37.73	30.22
v3 forecast	-	-	-	-	-	109.52	12.22	49.50	46.02	12.33	42.82	45.40
v6 forecast	6.21	-1.67	74.25	9.48	-43.46	90.04	12.26	52.75	40.84	9.26	36.32	26.03
v7 forecast	6.38	-2.21	72.55	8.86	-44.23	87.24	12.75	52.59	40.07	9.40	37.45	25.53
Av. forecasted return	**2.84**	**-3.39**	**79.52**	**6.26**	**-45.14**	**100.72**	**10.78**	**52.86**	**42.75**	**10.60**	**38.97**	**30.38**
Av. forecast error	**-14.72**	**-4.83**	**-25.30**	**2.96**	**5.40**	**-67.03**	**-1.09**	**-3.98**	**0.08**	**-30.96**	**-7.04**	**-9.92**

Table A2. Selected descriptive statistics for the multibagger sample (464 enduring multibaggers)

- Average share price growth over 15-years observation period: 26-fold (21.4% CAGR), including 24 100-baggers
- Size (in 2009 at the start of observation period): small
  - Median market cap in 2009: $348 m,
  - Median revenue in 2009: $702 m
- Median growth rates over 15 years (2009-2024): reasonably high but not spectacular (apart from net profit and EPS):
  - revenue: 11.1% CAGR
  - gross profit: 12.0% CAGR
  - operating profit: 17.3% CAGR
  - net profit: 22.9% CAGR
  - earnings per share: 20.0% CAGR
  - R&D expenditure: 15.1% CAGR
- Valuation (in 2009 at the start of observation period): low
  - median P/S 0.6; P/B 1.1; forward P/E 11.3; PEG 0.8
- Profitability (in 2009 at the start of observation period): average
  - gross profit margin 34.8%; operating profit margin 3.9%; ROE 9.0%; ROC 6.5%

Table A3. Time taken to achieve tenfold share price increase and CAGR growth rates

Median values for fastest 10x companies	Total share price growth (%)	Price growth, times (X-factor)	Share price growth (CAGR %)	Time to 10x (months)
Top 10	12143.9	122.4	37.6%	7.5
Top 25	5340.0	54.4	30.5%	11.0
Top 50	4283.8	43.8	28.7%	20.0
Top 100	3433.0	35.3	26.8%	40.5
Whole sample	1725.1	18.3	21.4%	107.0
Benchmark: Nasdaq100 (QQQ)	1277.0	13.8	19.1%	140.0

Figure A4. Multibaggers' industry and sector distribution

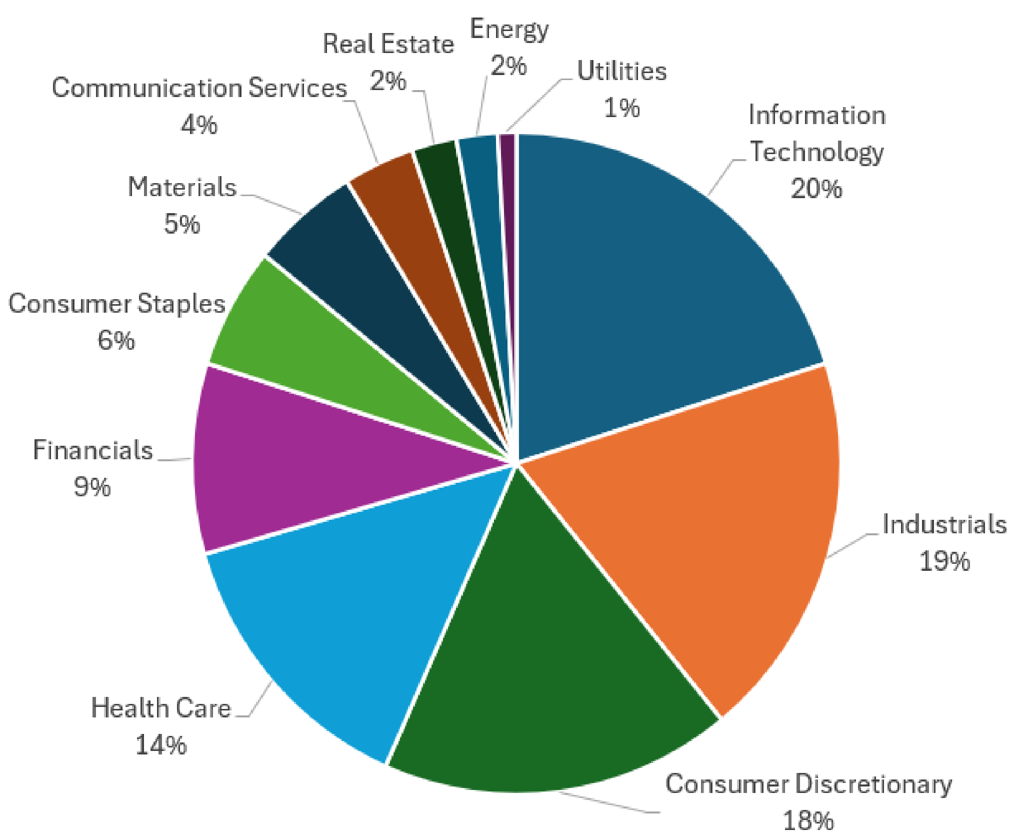

Table A4. Multibaggers' industry and sector distribution: further details

Primary sector	Number of companies	Number of companies (%)	Out of which (notable observations)
Information Technology	108	20.1%	Software 26 (4.8%), semiconductors 42 (7.8%), equipment and hardware 22 (4.1%)
Industrials	102	19.0%	Various machinery 23 (4.3%), logistics and transportation 17 (3.2%)
Consumer Discretionary	94	17.5%	Apparel 16 (3%), auto-related 26 (4.8%), home 16 (3%), leisure 18 (3.4%)
Health Care	75	14.0%	Biotechnology 11 (2%), healthcare tech, tool and equipment 37 (6.9%)
Financials	50	9.3%	Diversified banks 15 (2.8%) - all in developing countries
Consumer Staples	32	6.0%	Various food and drinks 20 (3.7%)
Materials	29	5.4%	
Communication Services	19	3.5%	10 US, 4 Europe, 5 other
Real Estate	12	2.2%	
Energy	11	2.0%	
Utilities	5	0.9%	All developing markets
Total	537	100.0%	