SAMBHAV REKHAWAT

**PREDICTING ACQUISITION TARGETS IN THE EUROPEN UNION**

Academic integrity declaration

I declare that:

* the thesis submitted has been written by me alone.
* the thesis submitted has not been previously submitted to this university or any other.
* that all content, including primary and/or secondary data, is true to the best of my knowledge.
* that all quotations and references have been duly acknowledged according to the requirements of academic research.

I understand that to knowingly submit work in violation of the above statement will be considered by examiners as academic misconduct.

Abstract

This study develops predictive models to identify likely acquisition targets among publicly listed firms in Europe. Using a comprehensive dataset of over 8,000 companies spanning 2000-2023, financial indicators are extracted to forecast takeover probability based on theories like inefficient management and undervaluation. Logistic regression, lasso regression and random forest models are implemented and evaluated on out-of-sample validation data. The random forest model demonstrates superior discriminatory ability with an AUC of 0.6791, significantly exceeding logistic and lasso regressions. The findings provide updated evidence that inferior profitability, growth and valuation discount target firms amidst the ongoing M&A wave. Predictive modelling enables pre-emptive identification of potential targets while informing bidding company strategy. Overall, this study highlights the evolving significance of financial metrics for modelling acquisitions in Europe. The predictive models can aid investors and managers in capitalizing on deal opportunities.

Keywords:

Mergers and Acquisitions, M&A, Random Forest, Logistic Regression, Lasso logit, Prediction, Acquisitions, European Union.

Acknowledgements

The successful completion of this intensive dissertation signifies the culmination of my enriching time at Cranfield School of Management. While a thesis is the work of the author, it is never accomplished alone. Therefore, my deepest gratitude goes to my supervisor and mentor, Dr. Vineet Agarwal, for his invaluable guidance, tireless dedication, insightful suggestions, thought-provoking discussions, and unwavering positivity throughout this process. His wisdom and advice have been instrumental in shaping this high-quality, robust analysis.

I am also tremendously appreciative of all the exceptional professors at Cranfield, whose teachings imparted me with crucial practical knowledge and skills. My sincere thanks to my talented international cohort - we made an excellent team! The enriching time spent at Cranfield University, collaborating with brilliant minds from around the world, will remain a treasured memory and formative experience. This dissertation provides an ideal foundation and source of inspiration as I embark on the next chapter in my personal and professional development. I am truly thankful to everyone who contributed to this achievement - we were all lifelong students together on the rewarding path of knowledge.

**Table of Contents**

[Academic integrity declaration ii](#_Toc144944771)

[Abstract iii](#_Toc144944772)

[Acknowledgements iv](#_Toc144944773)

[**CHAPTER 1: INTRODUCTION** 7](#_Toc144944774)

[A. BACKGROUND AND CONTEXT 7](#_Toc144944775)

[B. RESEARCH QUESTION 9](#_Toc144944776)

[C. SCOPE AND LIMITATIONS 10](#_Toc144944777)

[D. SIGNIFICANCE OF STUDY 11](#_Toc144944778)

[E. STRUCTURE OF THIS STUDY 11](#_Toc144944779)

[**CHAPTER 2: LITERATURE REVIEW** 12](#_Toc144944780)

[A. OVERVIEW OF M&A 12](#_Toc144944781)

[B. PREVIOUS RESEARCH 13](#_Toc144944782)

[ EARLY EMPIRICAL STUDIES 13](#_Toc144944783)

[ PALEPU’S CONTRIBUTION 15](#_Toc144944784)

[ MODERN THEORIES 17](#_Toc144944785)

[C. THEORETICAL FRAMEWORK AND HYPOTHESIS 20](#_Toc144944786)

[D. CONTRIBUTION OF THIS THESIS 24](#_Toc144944787)

[**CHAPTER 3: RESEARCH METHODOLOGY** 25](#_Toc144944788)

[A. RESEARCH APPROACH 25](#_Toc144944789)

[B. DATA COLLECTION AND SOURCES 29](#_Toc144944790)

[C. FINANCIAL CHARACTERISTICS OF TARGETS 32](#_Toc144944791)

[D. FINANCIAL PERFORMANCE ANALYSIS AND INVESTMENT STRATEGIES 36](#_Toc144944792)

[**CHAPTER 4: RESULTS AND FINDINGS** 38](#_Toc144944793)

[A. DESCRIPTIVE STATISTICS 38](#_Toc144944794)

[B. ANALYSIS AND INTERPRETATION 40](#_Toc144944795)

[ RESULTS 40](#_Toc144944796)

[ ABNORMAL RETURNS 48](#_Toc144944797)

[**CHAPTER 5: CONCLUSION** 49](#_Toc144944798)

[A. SUMMARY OF STUDY 49](#_Toc144944799)

[B. IMPLICATIONS OF THE FINDINGS 49](#_Toc144944800)

[C. LIMITATIONS AND FUTURE DIRECTION 49](#_Toc144944801)

[**CHAPTER 6: REFERENCES** 50](#_Toc144944802)

[**CHAPTER 7: APPENDICES** 54](#_Toc144944803)

Table of Figures

[**Figure 1: Annual Announcement of Takeover** 31](#_Toc144960393)

[**Figure 2:ROC Curve for Logistic Regression Analysis** 43](#_Toc144960394)

[**Figure 3: ROC Curve for Lasso Logistic Regression Analysis** 44](#_Toc144960395)

[**Figure 4: Lasso path Plot** 45](#_Toc144960396)

[**Figure 5: ROC Curve for Random Forest Classification Analysis** 46](#_Toc144960397)

[Table 1 Research on Prediction Acquisition 19](#_Toc144969615)

[Table 2 Descriptive Summary of Sample 30](#_Toc144969616)

[Table 3 Year-wise Description of Database 31](#_Toc144969617)

[Table 4: Independent Target Variable Selection 36](#_Toc144969618)

[Table 5 The table indicates significant variables and descriptions of various variables used. \*\*\*, \*\* and \* significant at 1%, 5%, and 10% level, respectively. 40](#_Toc144969619)

# **CHAPTER 1: INTRODUCTION**

## BACKGROUND AND CONTEXT

The past decade has been characterized by robust merger and acquisition (M&A) activity globally, with the total value of deals more than doubling from $2.2 trillion in 2010 to $5.8 trillion in 2021 (IMAA (2022)). As a prime driver of this worldwide M&A boom, Europe saw deal value grow from $383 billion to $1.26 trillion over the same period. The European Union has been at the forefront, accounting for 28% of global deal value in 2021 (Mergermarket (2022)). The region witnessed strong convergence in M&A volume across major economies like the UK, Germany, France, Italy, and Spain.

Several factors have contributed to the mounting M&A wave in Europe. Accommodative monetary policies post the Eurozone crisis drove companies to use cheap financing for inorganic growth (Bartram et al. (2022)). The withdrawal of the UK from the EU led UK companies to refocus domestically, while Brexit uncertainty hampered deals in 2018-2020.

For bidders, European assets have become attractive as the region recovers from a decade of slow growth. Valuations remain lower compared to the US, especially in Southern Europe. Private equity has also played a major role, with buyout deals comprising 42% of European M&A value in 2021 (Mergermarket (2022)). As the deal frenzy continues, accurately predicting potential targets has become critical for investors aiming to profit from sizable acquisition premiums.

Since Palepu’s (1986) seminal work, modelling acquisition likelihood has relied heavily on financial indicators related to inefficient management, undervaluation, growth-resource imbalance and other inefficiencies that may motivate a takeover bid. More recent studies have augmented these models with technical factors like price momentum and trading volume (Brar et. al (2009)), along with macroeconomic and industry variables. Advanced statistical techniques like machine learning have also been incorporated into predictive modelling in the last decade (Andriosopoulos, Andriosopoulos and Hoque (2019).

However, most prior studies focusing specifically on the European market utilized data from before 2010 (Brar et. al (2009); Ouzounis, Gaganis and Zopounidis (2009)). The region’s economic landscape has changed significantly over the past decade, necessitating updated data. Modelling techniques have also evolved rapidly with exponential growth in computing power. Hence, this paper aims to develop the latest acquisition likelihood models for the EU leveraging data from 2010-2022.

Logistic regression will be implemented along with regularized forms like ridge, lasso and elastic net regression that perform automatic variable selection (Tibshirani (1996)). Ensemble methods like random forest will also be explored for non-linear modelling (Breiman (2001)). The predictive performance of these techniques will be evaluated on out-of-time validation data. Macroeconomic indicators will provide crucial context on the deal environment. The large sample size (~1300 deals) will facilitate robust validation across time periods, industries and economic conditions.

Overall, this study will provide important practical insights into the changing profile of European targets amid the ongoing M&A boom. The predictive models can aid investors in pre-emptive deal identification while also informing bidding company strategy. For researchers, the results will highlight the evolving significance of financial metrics, stock markets, and macro-economy in explaining European M&A activity since the financial crisis.

## RESEARCH QUESTION

With European M&A activity reaching new highs in 2022 amid economic recovery and regional consolidation trends, an intriguing research question arises - can machine learning techniques accurately model the financial profile of European acquisition targets and enable profitable trading strategies? Despite prior studies demonstrating the potential to forecast takeover probability and earn abnormal returns in the US and UK (Palepu, 1986; Powell, 2004), limited research has examined the continental European market across different macroeconomic environments using latest prediction methods. While historical analyses have relied on logistic modelling of accounting ratios and firm financials, advanced algorithms like random forest may better capture the non-linear relationships and interactions governing acquisition likelihood in today's complex markets.

By compiling recent data on European firms and deals from 2000-2023, this proposed research aims to harness the predictive strengths of logistic regression, the regularization and sparsity of lasso regression, and the ensemble learning of random forests.

The study should provide updated evidence on the financial indicators that help forecast acquisitions in Europe amidst the ongoing M&A boom. From a practical standpoint, the predictive models can aid investors and managers in identifying likely targets (Palepu, 1986). For researchers, the findings will highlight the evolving explanatory power of financial ratios in the context of modern machine learning and feature selection methods.

## SCOPE AND LIMITATIONS

This study has certain limitations that warrant acknowledgement. First, the sample comprises only publicly listed European firms, excluding privately held companies, which reduces the generalizability of the results (Rossi and Volpin (2004)). In addition, the exclusions of utilities and financial institutions, although necessary to mitigate industry-specific confounds, limit the diversity of sectors analysed (Jensen and Ruback (1983)). A key constraint with observational data is the inability to make causal claims regarding the predictive ability of financial indicators on future takeover likelihood. The models are also vulnerable to omitted variable bias by not incorporating all potential factors influencing takeover decisions, such as managerial incentives or private information (Palepu (1986)).

Since the analysis relies solely on pre-acquisition financial data, it does not account for how motives for acquisition may evolve over time as macroeconomic conditions change (Gorton, Kahl and Rosen (2009)). The use of announced deals regardless of outcome risks bias if failed takeover attempts differ systematically from completed transactions (Walkling, 1985). Additionally, transaction costs and strategic rationale behind acquisitions are not directly observable in secondary datasets (Comment and Schwert, (1995)). The choices involved in statistical modelling, including technique selection, variable specification, and tuning, necessitate exercising researcher degree of freedom that could potentially influence results (Lo (1986)).

The study's focus on predictive models gives limited insight into post-merger integration success, which depends on a wider range of factors beyond pre-deal financials (Tuch and O'Sullivan, 2007). Finally, the use of a single time-period for out-of-sample testing provides incomplete evidence on model robustness across different economic regimes (Levine and Aaronovitch, 1981). In summary, the limitations primarily concern sample construction, unobserved confounding, potential omitted variable and model specification bias, inability to make causal claims, and lack of insight into post-acquisition performance.Ohlson (1980)

## SIGNIFICANCE OF STUDY

This study offers important contributions to research on prediction modeling for corporate acquisitions, specifically concerning publicly-traded European firms. First, it provides updated empirical evidence from a extensive sample spanning 2000-2022. Many prior studies relied on pre-2010 data, despite major economic changes in Europe over the past decade. By leveraging recent data, this study reveals evolved insights into the financial profile of acquisition targets.

Second, the research explores an array of advanced statistical learning techniques including logistic regression, elastic net, and random forest models. The application of these methods to M&A prediction is still emerging in academic literature. Comparing multiple techniques on a large unified dataset advances understanding of their relative effectiveness for this application.

Third, the macroeconomic context is explicitly incorporated into the modelling framework to account for deal environment factors. The models are evaluated across different time periods and economic conditions, rather than assuming temporal stability. This allows robust assessment of generalizability. The study covers the breadth of Europe across 27 countries and diverse industries. Many prior works concentrated on single countries. The expansive scope enhances the external validity of findings. The examination of post-prediction investment returns provides important evidence on the economic significance of the forecasting models. This analysis of risk-adjusted returns and performance impact demonstrates the practical utility of the approach to investors.

## STRUCTURE OF THIS STUDY

This paper is constructed as follows. Chapter 2 presents the theoretical framework of target firms characteristics, industry and macroeconomics determinants due to various acquisition hypotheses. The relevant literature on acquisition probability modelling and outlines the contribution. Chapter 3 details the comprehensive database used in this study. Chapter 4 describes the employed methodology with regards to the input variables, the takeover prediction modelling, the portfolio formation, and the return calculation. Chapter 5 presents findings ranging from form the summary statistics over classification accuracy in- and out-of-sample to the generation of abnormal returns.

# **CHAPTER 2: LITERATURE REVIEW**

## OVERVIEW OF M&A

The definition of targets and non-targets in the literature varies. In most studies, targets refer to firms that have been successfully taken over, while non-targets are those that have not been subject to takeover attempts. However, some researchers, such as Belkaoui (1978) and Barnes (1999), consider all firms that have faced takeover attempts, regardless of the outcome, as targets. There is also the argument made by Bartley, (1990) that firms facing investment attempts exceeding 5% of ownership should be considered targets and not included as non-targets. This threshold is based on the disclosure norms invoked by the Securities and Exchange Commission in the US.

The selection methods for non-targets have varied between matched sample designs and choice-based research designs. While the selection method for non-targets is still a subject of debate, many researchers have randomly chosen companies from the list of all non-targets. Others, such as Barnes (1990, 1999, 2000), have selected companies that match target firms' assets, sales, or market capitalization.

## PREVIOUS RESEARCH

### EARLY EMPIRICAL STUDIES

Early studies on acquisition and mergers focused primarily on the United States and aimed to identify the financial characteristics of companies targeted for takeovers. Taussig (1968) conducted research that refuted the hypothesis suggesting that companies with low investor transparency were more likely to be taken over. Through a univariate analysis of three accounting variables, they found that weak investment policies, low return on equity, and underperforming dividend payouts instead characterized takeover targets. Simkowitz (1971) expanded on this methodology by employing a multiple discriminant analysis. Their model, based on 24 financial variables, ultimately included six factors: profitability, growth, size, leverage, liquidity, and payout. The study in the US in 1968 revealed that targets had lower price-to-earnings(P/E) ratios, growth rates, and dividends, confirming theories of inefficient management and undervaluation.

Stevens (1973) conducted a study that examined a matched set of 40 targets from 1966 and non-targets from 1966-1970, matched by size. The model used by Stevens consisted of net working capital, long-term liabilities, sales in relation to total assets, and earnings before interest and taxes (EBIT) to sales. The results indicated a positive correlation with liquidity and a negative correlation with gearing, revenue, and profit. The accuracy of the tested model reached 70%.

Singh (1975) investigated a wave of takeovers in the United Kingdom from 1967 to 1970. By employing univariate and multivariate analyses on 112 targets and 463 counterparts from the engineering, textile, food, and drink sector, Singh identified 17 cluster determinants. The univariate analysis revealed that firms with higher profitability acquired smaller, less liquid, and leveraged targets. The multivariate analysis showed a success rate of 83%.

From 1960 to 1973, Belkaoui (1978) and Rege (1984) utilized multiple discriminant analyses for the Canadian market. Belkaoui randomly selected 25 acquired entities matched in industry and size with non-targets. Based on bankruptcy indicators, their model found that net income and cash flow to net worth had the best predictive ability. Similar to the inefficiency theory, targets exhibited unfavourable values in working capital, total assets, receivables, and sales two years before the takeover. The model was statistically significant only for the furthest periods, and the accuracy in significant years 3 to 5 ranged from 84% to 78% when validated with a holdout sample of 11 firms per side.

Rege (1984) conducted a joint hypothesis test and concluded that his five measures needed to be more significantly differentiating between domestic, foreign, acquired, and non-acquired companies. Based on his sample, which was linked by industry, size, and year, Rege inferred that historical accounting data was incorporated into stock prices, suggesting semi-strong market efficiency.

Dietrich and Sorensen (1984) refined previous studies by implementing a logit model that used industry-related financial data rather than raw financial data. They expanded the dataset to include legal and regulatory information, focusing on firm cash flows and net present value parameters. The study identified ten corporate features, including size, long-term debt, sales, capital expenditure (CAPEX), current ratio, profit margin, EBIT to interest, trading volume, dividend payout, and P/E ratio. Based on their model, smaller businesses with lower asset turnover, lower dividend payouts, and higher trading volume were found to be more likely takeover targets. The study demonstrated an explanatory power of 67% and an accuracy rate of 93% in predicting targets when validated with a holdout sample of 22 companies.

Hasbrouck (1985) used a logit model to assess the variations in financial features between acquired and non-acquired companies. The study involved 86 unregulated target firms, excluding financial institutions, matched with control groups of 172 non-targets (1977-1982) based on year, size, and industry. Tobin's Q for common equity and the entire firm, along with five items for size, leverage, and liquidity, were employed. The results indicated that targets tended to have significantly lower Qs and higher current liquidity. However, the study did not include a classification accuracy test to assess the model's performance.

Overall, the early studies in this field suffered from high multicollinearity, insignificant variables, limited consideration of theoretical foundations, poor models, and short observation periods. Logit models provided less restrictive data assumptions and allowed for more direct interpretation compared to linear analyses. However, these studies' high explanatory power and accuracy were based on small, biased samples, with limited out-of-sample testing and arbitrary cut-off points.

### PALEPU’S CONTRIBUTION

During the period from 1986 to 2002, research in the field of acquisitions aimed to develop theories for variable selection and enhance empirical techniques used in prediction models. Notably, Palepu's seminal study in 1986 identified three methodological shortcomings in earlier research that led to artificially inflated predictive accuracy.

Firstly, previous studies employed state-based samples by matching both takeover targets and non-targets. However, this approach introduced inherent bias due to sample selection, resulting in inconsistent estimates of model parameters and biased takeover probabilities. To address this issue, Palepu (1986) proposed the implementation of weighted and conditional maximum likelihood estimators to avoid overstating prediction ability.

Secondly, prior studies utilized contrived samples to assess the accuracy of prediction models. However, such samples did not ensure the generalizability of the results. The population under study was characterized by a smaller proportion of acquired companies, known as the "rare event problem." Consequently, equal-share samples overestimated the classification ability of the model. Palepu (1986) suggested using random samples or the entire population to mitigate this issue.

Thirdly, early research defined arbitrary cut-off probabilities for classifying firms as targets, often using a likelihood threshold of 0.5 or higher. Palepu (1986) emphasized that the optimal point depends on prior probabilities, contextual factors, and appropriate pay-off functions. The choice of threshold impacts the hit ratio to identify targets and the trade-off between minimizing Type I errors (predicting non-targets as targets) and maximizing abnormal returns.

Palepu's 1986 study constructed a logit model based on six hypotheses and nine factors using data from 163 targets and 256 non-targets listed on AMEX or NYSE between 1971 and 1979. Four models were developed using a conditional maximum likelihood procedure, revealing significance at the 1% level for size, average excess return, growth resource, and industry dummy variables. However, only the industry dummy variable exhibited a sign contrasting the related hypothesis. While the models were statistically significant, their predictive power was relatively low (7% to 13%), confirming Palepu's argument that previous studies had overestimated results due to methodological deficiencies.

A holdout sample consisting of 30 targets and 1,087 non-targets from 1980 was employed to validate the superior model. The overall accuracy rate of 45.60% resulted in high misclassifications of targets, leading to insignificant abnormal returns of -1.62% over 250 days, indicating inferior performance compared to the stock market.

In other related studies during this period, Bartley and Boardman (1990) used multiple discriminant analysis to investigate the optimization of prediction models in ranking firms as acquisition targets using current cost, constant dollar, and historical cost data.

Ambrose and Megginson (1992) assessed the effects of takeover defence, asset, and ownership structure on US acquisitions. Powell (1997) examined whether friendly and hostile acquisition targets required different prediction indicators. Barnes (1999, 2000); Fairclough and Hunter (1998); Powell (2001) also explored alternative prediction methodologies.

These studies yielded mixed results, with varying levels of explanatory power ranging from approximately 6% to 9%. While some models demonstrated high predictive accuracy for non-targets, the ability to forecast targets remained limited accurately. Powell (2001) criticized Palepu's low cut-off points, suggesting the need to maximize the number of targets (minimize Type II errors) rather than focusing solely on reducing overall misclassifications.

The studies during this period revealed methodological shortcomings, including biased selection, contrived samples, and arbitrary cut-off probabilities, which resulted in inflated predictive accuracy and limited generalizability. Samples introduced inherent bias, leading to inconsistent estimation of model parameters. The disproportionate representation of acquired companies created a "rare event problem," with equal-share samples overestimating model classification ability. Researchers recommended weighted and conditional maximum likelihood estimators to address these issues. Optimal cut-off probabilities depended on prior probabilities, contextual factors, and pay-off functions. While some models demonstrated high accuracy in classifying non-targets, accurately predicting targets remained challenging. The studies collectively concluded that statistically significant excess returns could not be achieved at the level of predictive accuracy attained, aligning with efficient market theory. The research emphasized the need to address methodological limitations and carefully consider sample selection and cut-off probabilities in acquisition prediction models.

### MODERN THEORIES

From 2003 to 2012, advanced studies focused on developing more accurate prediction models for acquisitions, leveraging new computational technology, and expanding the geographic scope of research. Espahbodi and Espahbodi (2003) conducted a comparative analysis of discriminant, logit, probit, and recursive partitioning models. To address the time dependency issue observed in previous models, the authors limited their sample to the last six months of 1997, resulting in a matched dataset of 133 targets and 385 non-targets. Their findings demonstrated the superior predictive ability of the recursive partitioning model, which achieved validation accuracies of 66% to 55% for logit/probit models and 53% for the discriminant model.

Building on the work of Powell (1997), Powell (2004) expanded the sample size and employed binomial and multinomial models to examine the characteristics of hostile and friendly targets. The results indicated that multinomial models outperformed in accurately predicting target firms, with the models focusing on hostile targets generating significant positive abnormal returns of 15% to 17% over a 24/36-month holding period. Doumpos, Kosmidou and Pasiouras (2004) reported higher overall accuracy for multicriteria decision aid methods and neural networks compared to discriminant analysis and logit models, indicating the absence of clear methodological preferences regarding target accuracy. Tsagkanos, Georgopoulos and Siriopoulos (2007) applied recursive partitioning techniques to predict targets in Greece, revealing that the accuracy of the models, although below 50%, was likely influenced by contextual and country-specific factors.

Ouzounis, Gaganis and Zopounidis (2009) found that targets were significantly larger, undervalued, and less profitable. Comparing different modelling techniques, they observed a minor superiority of non-parametric techniques such as UTADIS, neural networks, and support vector machines. Brar et. al (2009) examined a large sample of European and cross-border takeovers using logit models based on various firm, industry, and country factors. The study revealed that targets tended to be small, slow-growing, and positively correlated with dividends, momentum, trading volume, and industry disturbance. The study by Cremers et al. (2009) investigated the effects of takeover probability on firm valuation using logit models capturing various firm characteristics. Their findings suggested that portfolios based on acquisition vulnerability generated significant alphas relative to existing asset pricing models.

De and Jindra (2012) explored the operational, financial, and ownership characteristics of US firms acquired three years after their initial public offerings (IPOs). They found that acquired companies exhibited good operational performance, growth potential, strong interest from institutional investors, and superior share returns.

Brar et. al (2009) investigated a large sample of European and cross-border takeovers using four logit models in their study. These models incorporated various firm, industry, and country factors and technical measures such as momentum, trading volume, and market sentiment.

The findings of the study revealed several characteristics of the target firms. Targets were generally small in size and exhibited slow growth rates. They also positively correlated with dividend payments, momentum, trading volume, and industry disturbance.

However, the study did not find significant effects of country variables on the likelihood of takeovers. The findings suggest that country-specific factors, such as regulatory environments or cultural differences, did not substantially impact the acquisition decisions in the European and cross-border takeovers examined in the study.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **TABLE 1: RESEARCH ON PREDICTION ACQUISITION**  This table provides an overview of research studies conducted on predicting acquisition targets over different periods and in various countries. The studies used different methodologies, including univariate and multivariate analyses, logistic regression, neural networks, and more. | | | | | |
| ***Research*** | ***Area*** | ***Period*** | ***data*** | ***prediction accuracy*** | ***Methodology (VARIABLES)*** |
| Taussig & Hayes (1968) | USA | 1956-1967 | 50/50 | - | UNIVARIATE (3) |
| Simkowitz & Monroe (1971) | USA | 1968 | 25/25 | 63.20% | MDA (6) |
| Stevens (1973) | USA | 1966-1970 | 40/40 | 70% | MDA (20) |
| Singh (1975) | UK | 1967-1970 | 112/463 | 83% | UNI/MULTIVARIATE (17) |
| Belkaoui (1978) | CANADA | 1960-1968 | 36/36 | 78-84% | MDA (16) |
| Rege (1978) | CANADA | 1962-1973 | 55/55 | - | MDA (5) |
| Dietrich & Sorensen (1984) | USA | 1969-1973 | 30/60 | 91% | LOGIT (10) |
| Hasbrouck (1985) | USA | 1977-1982 | 86/172 | - | LOGIT (7) |
| Palepu (1986) | USA | 1971-1980 | 193/1343 | 80/45% | LOGIT (9) |
| Bartley & Boardman (1990) | USA | 1975-1981 | 41/153 | - | MDA (26) |
| Ambrose & Megginson (1992) | USA | 1981-1986 | 169/267 | - | LOGIT (19) |
| Powell (1997) | UK | 1984-1991 | 411/532 | 6-9% | BI/MULTINOMINAL LOGIT (8) |
| Fairclough & Hunter (1998) | UK | 1987-1996 | 140/1480 | - | NEURAL NETWORKS (17) |
| Barnes (1999/2000) | UK | 1991-1994 | 95/968 | - | LOGIT/MDA (17) |
| Cudd & Duggal (2000) | USA | 1987-1992 | 121/1534 | 76% | LOGIT (9) |
| Powell (2001) | UK | 1986-1996 | 473/1393 | 3% | LOGIT (8) |
| Espahbodi & Espahbodi (2003) | USA | 1993-1999 | 171/585 | - | MDA/LOGIT /PROBIT/ RP(4) |
| Powell (2004) | UK | 1986-1996 | 500/10391 | - | BI/MULTINATIONAL LOGIT (8) |
| Doumpous et al. (2004) | UK | 200-2002 | 76/76 | - | MCDA/MDA/LOGIT/NN (29) |
| Tsagkanos et at. (2007) | GERMANY | 1995-2002 | 51/200 | <50% | RP(17) |
| Ouzounis et al. (2009) | UK | 2001-2006 | 745/1489 | - | MCDA/MDA/SVM/NN(7) |
| Brar et al. (2009) | E.U | 1992-2003 | 530/1434 | 73,42,45% | LOGIT (33) |
| Cremers et al. (2009) | USA | 1981-2004 | 5457/78295 | - | LOGIT (10) |
| De and Jindra (2012) | USA | 1980-2006 | 932/5144 | - | LOGIT/ HAZARD (19) |

Table 1 Research on Prediction Acquisition

## THEORETICAL FRAMEWORK AND HYPOTHESIS

Theoretical frameworks comprise concepts, theories, models, and existing knowledge from relevant literature that inform the research design, methodology, and data analysis. They serve as a lens through which researchers examine their research questions, formulate hypotheses, and make sense of their findings.

The *Inefficient Management Hypothesis,* introduced by Marris (1964), proposes that acquisitions serve to transfer control of resources from inefficient management to more capable leadership. This hypothesis aligns with the concept of a market for corporate governance, as outlined by Manne (1965), where acquirers can benefit from acquiring firms with previously ineffective operations and subsequently improve their performance, leading to increased wealth generation.

The notion, inefficiently managed firms are more likely to become targets of takeover bids in the literature, as Palepu (1986) noted. Jensen (1986) defines poor management as a reluctance to distribute free cash flows, resulting in the funding of projects with negative net present value (NPV) and incurring agency costs. Consequently, stockholders penalize firms with low dividends and limited investment opportunities, leading to lower share prices and a higher probability of acquisition, as suggested by Simkowitz (1971) and Powell (1997).

Researchers have utilized various proxies, such as total assets, receivables, sales growth, profit before tax, and return on capital employed (ROCE), to assess managerial performance. Studies by Belkaoui, (1978), Barnes (1990), and Ouzounis, Gaganis and Zopounidis (2009) have found that target firms tend to exhibit lower accounting profitability. Powell (2007) incorporates market measures, precisely average abnormal returns from the previous two years, to capture future outcomes of managerial actions independent of regulatory influences. While the efficient market hypothesis suggests that stock prices reflect all available information, market anomalies, and inefficiencies can temporarily cause prices to deviate from their fundamental values, leading to abnormal returns. These abnormal returns may persist over time due to behavioural biases, market inefficiencies, and persistent factors that influence a company's performance. Their findings indicate declining values before takeovers, supporting the inefficiency hypothesis.

The hypothesis of inefficiency is further supported by research demonstrating inferior profitability of acquisition targets compared to their peer groups, as indicated by Tzoannos (1972) and Asquith and P. (1983). Brar et. al (2009)also find lower sales growth for European targets, consistent with these findings. Additionally, Rege (1984) suggests that firms with higher asset turnover are deemed attractive to acquirers, as they indicate high demand and output potential.

*The Growth-Resource Mismatch Hypothesis*, rooted in the neoclassical perspective on acquisitions and value creation through synergies, suggests that slow-growing acquiring companies with excess resources seek out targets with higher growth potential and weaker financial structures. This hypothesis highlights the disparity between growth and financial resources, indicating that companies with low growth rates, high liquidity, and low leverage, or vice versa, are appealing targets for acquisitions. Palepu (1986) incorporates a dummy variable in his prediction model that includes average sales growth, net liquid to total assets, and debt-equity ratio as indicators. Empirical studies such as Levine and Aaronovitch (1981) and Powell (2004) support the idea that the growth-resource imbalance theory contributes to an increased likelihood of takeovers.

*The Size Hypothesis* suggests that smaller companies are more prone to acquisition bids. This is due to the presence of various transaction costs associated with purchasing a business, which results in a reduced number of potential bidders as the size of the target firm increases. As identified by Palepu (1986), these transaction costs include long takeover battles, post-merger integration costs, and advisory services. Gorton, Kahl and Rosen (2009) support the notion of affordability as a rationale behind this hypothesis, while the empire-building theory of Mueller (1969) suggests that managers may prefer larger targets. Firm size is measured using different variables, primarily capturing the purchase price (net book value; Palepu (1986)) and transaction costs (market cap; (Cremers et al. (2009)). Singh (1975), (Dietrich and Sorensen, 1984), Bartley (1990), Song (1993), and (Brar et. al (2009) have found statistically significant evidence supporting the idea that smaller companies are more likely to be targeted for acquisitions.

*The Leverage and Liquidity Hypothesis* is derived from the imbalance and inefficient theories in the context of mergers and acquisitions (M&A). These hypotheses highlight the role of leverage and liquidity in influencing the likelihood of a takeover.

The Leverage Hypothesis suggests that low-leverage companies, which have a higher capacity to take on additional debt, are attractive to acquirers, particularly in the context of leveraged buyouts (LBOs). Studies such as Lewellen (1971), Barnes (1999), and (Brar et. al (2009)) have found evidence supporting this hypothesis. Erdogan (2012) includes the debt ratio as a variable and identifies a strong negative relationship between leverage and the likelihood of a takeover. Additionally, lower leverage levels are often associated with lower growth, which may be interpreted as inefficient management by excessively relying on debt structure (Banner, 1999).

*The Liquidity Hypothesis* suggests that high-liquidity firms are more likely to be involved in mergers. Firms with high liquidity have greater financial resources and flexibility, making them attractive to acquirers. Studies by Rege (1984), Hasbrouch (1985), and others have found a positive relationship between liquidity and the likelihood of a merger. Rege (1984) specifically links this hypothesis to the Growth-resources theory, highlighting that highly leveraged firms, which may lack funding for further expansion, become potential targets for acquisitions.

*The Asset Structure Hypothesis* suggests that higher levels of fixed assets in a target company serve as collateral and can reduce transaction costs in mergers and acquisitions. (Ambrose, B.W. and Megginson and W.L., 1992) have provided evidence supporting this hypothesis. They found that firms with higher fixed asset levels are more attractive to acquirers due to the availability of collateral, which can lower the perceived risk and costs associated with the acquisition.

In addition to the Asset Structure Hypothesis, other qualitative factors related to takeover defence and ownership structure have been found to influence the likelihood of a takeover. Sokolyk (2011) reveals that the presence of a golden parachute (a compensation arrangement for executives in the event of a takeover) has a positive impact on the likelihood of a takeover. On the other hand, the absence of a poison pill (a defence mechanism against hostile takeovers) and the presence of a staggered board (where only a portion of the board is up for election each year) also increase the likelihood of a takeover.

Furthermore, De and Jindra (2012) find that U.S. target companies tend to have a higher level of institutional ownership. This suggests that institutional investors play a significant role in the target selection process, possibly due to their influence on corporate governance and strategic decision-making. These findings contribute to understanding the various factors that influence the occurrence of mergers and acquisitions, including the role of asset structure, qualitative takeover defence mechanisms, and ownership structure.

*The Momentum Hypothesis* suggests that short-term technical factors, such as three-month price momentum and trading volume, can be indicative of potential takeovers. These factors reflect bearish share prices, rumours, or insider trades that may signal increased takeover activity. Brar et. al (2009) confirm statistically significant differences in these factors for target companies.

It is important to note that market proxies, including price momentum and trading volume, are influenced by future expectations and can lead to temporary distortions in market prices. In the case of high-momentum companies, there is a possibility of overestimation or inflated market prices due to heightened expectations or speculative trading. Therefore, while momentum factors can provide insights into potential takeover targets, they should be interpreted with caution, considering the influence of market expectations and possible distortions in pricing.

## CONTRIBUTION OF THIS THESIS

This thesis seeks to contribute to existing research on acquisition target prediction by investigating numerous characteristics that might be utilised to anticipate future acquisitions. The study demonstrates the model's forecasting abilities by doing tests on historical financial data. It also emphasises the distinctions between the variables utilised in this thesis and those usually used in past acquisition studies, emphasising the benefits of the variables chosen. The study statistically evaluates the explanatory power of the new variables to those used in previous studies. Furthermore, the thesis analyses the model's effectiveness in generating profitable and sustainable investment strategies, particularly in the context of acquisition events. It discusses the suitability of implementing an acquisition-driven investment strategy and provides insights into the patterns and trends observed in acquisition targets, offering recommendations for future research in this area. Ultimately, the research aims to provide valuable information to executive directors, helping them identify whether their firm is at risk of being targeted for acquisition.

# **CHAPTER 3: RESEARCH METHODOLOGY**

## RESEARCH APPROACH

This section describes the research design and approach used in this thesis. Researchers have been constantly researching new approaches and methods to improve existing models as the discipline of predicting acquisition targets has evolved. Multivariate analysis approaches such as discriminant, logit, and probit models were frequently used in the early phases. However, Palepu (1986) introduced new methodologies, resulting in the widespread use of cutting-edge techniques such as neural networks, rough sets, and decision trees. These new methodologies shifted the research environment, providing up new opportunities for anticipating acquisition candidates.

In the context of mergers and acquisitions (M&A) data analysis, logistic regression offers a suitable choice due to its ability to predict binary or categorical outcomes. As M&A targeting often involves assessing the likelihood of a target company being acquired, logistic regression can estimate the probability of such an event occurring. Additionally, logistic regression can accommodate both categorical and continuous independent variables, making it flexible for incorporating a range of predictor variables relevant to M&A transactions(Ambrose, B.W. and Megginson and W.L. (1992); Brar et. al (2009); Powell (2004)). By estimating coefficients, logistic regression provides insights into the significance and direction of relationships between predictor variables and the likelihood of M&A targets.

In equation 1, is the logit probability the firm will be acquired and takes the number one if the firm is acquired and 0 if it is not acquired in the prescribed time period. The vector form of the characteristics of the firms examined are denoted by the term . The rationale of the model is that the acquisition of the firm is influenced by various characteristics of the firm as well as the motives, sentiment, and number of bidders. A three-month lag is used to avoid hindsight bias and availability of adequate data. The dependent variable is assigned the value of one if the firm gets a bid in the next period, otherwise the responding variable is coded zero.

The suitability of prediction parameters is assessed with a ROC analysis and the classification results shown in contingency tables are tested by chi-square statistics:

Where is the observed frequency at level of targets and of non-targets and is the

expected frequency at level of targets and of non-targets .

There are numerous techniques that primarily differ in the theory of Type I and II mistake costs. Palepu (1986) obtains cut-offs by lowering total misclassifications, whereas Barnes (1999) seeks to maximise returns. Since neither strategy is clearly superior, this study employs an alternative by investing in the top 30% of enterprises with the highest likelihood, as Brar et. al (2009) did with the top 10%.

Lasso regression serves as a valuable technique in M&A prediction models due to its ability to perform variable selection and reduce model complexity. With a potentially large number of predictor variables in M&A analysis, Lasso regression helps identify the most important variables by shrinking less relevant predictors to zero. By promoting sparsity, Lasso regression provides a concise set of predictors that significantly contribute to predicting M&A targets, thus enhancing model interpretability.

Here λ denotes the regularization tuning parameter that controls the sparsity. Examination of coefficient trajectories along the regularization path provided insights into feature selection behaviour.

The randomforest algorithm operates by constructing an ensemble of decision trees, each trained on a random subset of the data, and then averaging their predictions to yield overall robust forecasts. A key advantage of random forests is reducing the variance of single decision tree models, which are high variance estimators, while retaining low bias. Specifically, the random forest prediction process involves bootstrapping or randomly sampling with replacement from the training data to create multiple subsets of the data. Decision trees are then trained individually on each data subsample. However, a key difference from bagging methods is that only a random subset of candidate predictor variables is made available for splitting at each node in the tree creation process.

By training each tree on a slightly different data sample and set of variables, the resulting decision trees vary and will make somewhat different predictions. Conceptually, each tree provides a "voice" in classifying a new data point, and the wisdom of the crowd of decision trees overall gives a more robust prediction than any individual tree. Mathematically, given a training dataset with N observations and p features, a bootstrap sample of size N is drawn to grow each tree. At each node split in the tree, rather than evaluating all p predictors, a random subset of m << p variables is chosen as candidates. This has the effect of decorrelating the trees, such that they are not overly like one another. Each fully grown tree provides a class prediction for new data points, and the predictions of all n trees in the forest are averaged to obtain the overall random forest prediction.

Tuning parameters such as the number of trees n, the subsample size used to train each tree, and m the number of split variables considered at each node allow trading off computational expense versus model performance. Variable importance measures can also be extracted to understand which predictors had the most marginal impact on accuracy.

The prediction from a random forest model, for an input data point x, can be written as:

E[Θ(x; Θk)]

Where:

f(x) is the prediction for data point x

Θk represents the parameters (tree structure, split points, etc.) of the kth decision tree

Θ is the set of parameters for all the decision trees in the forest

E[] denotes the expectation over the random forest ensemble

So the prediction is the average of the predictions from the individual decision trees (Θk), where each tree depends on a different random sampling of the data and parameters.

The expectation is approximated as the average prediction over the n trees in the forest:

Where Θk(x) is the prediction of the kth decision tree for input x. So in summary, the random forest prediction is the average of the predictions from n decision trees, where each tree is trained on a bootstrap sample of the data and parameters. The averaging reduces the variance and improves the overall predictive accuracy.

The key parameters of a random forest include:

n - the number of trees

m - the number of variables randomly sampled at each split

s - the sample size drawn with replacement (bootstrapping)

So the full random forest prediction equation captures how the individual decision trees are trained and then averaged to generate robust predictions.

## DATA COLLECTION AND SOURCES

Creating a robust and reliable database is a crucial foundational step in conducting thorough empirical investigations within the realm of finance research (Powell (1997)). This research effort has meticulously compiled an extensive dataset concerning European enterprises, aiming to delve into the financial attributes of companies marked for corporate takeovers before these acquisitions occur. The dataset initially encompasses a total of 13,347 firms hailing from 27 European nations, all of which were listed on their respective national stock exchanges during the span between January 2000 and June 2023. The selection encompasses diverse industries except for utilities and financials, as these sectors possess distinct regulatory and accounting practices (Brar et. al (2009)). The data is further filter to have only primary and major listings in Europe. The dataset incorporates intricate firm-level accounting information and monthly stock prices, both sourced from Datastream in Euro currency. Additionally, the corporate takeover announcements from January 2000 to July 2023 are seamlessly integrated through the employment of the SDC Platinum database by Thomson Reuters.

|  |  |  |  |
| --- | --- | --- | --- |
| **Table 2: Country wise distribution of Targets and NonTargets**  The table shows the Country wise proportion of target to non-target data. This explains the Merger activities happening in the European union. | | | |
| **Countries** | **Non-Targets(NT)** | **Targets(T)** | **prop. T to NT** |
| Austria | 98 | 23 | 23.4694% |
| Belgium | 192 | 37 | 19.2708% |
| Bulgaria | 312 | 9 | 2.8846% |
| Croatia | 327 | 14 | 4.2813% |
| Cyprus | 138 | 9 | 6.5217% |
| Czech Republic | 19 | 4 | 21.0526% |
| Denmark | 209 | 62 | 29.6651% |
| Estonia | 26 | 3 | 11.5385% |
| Finland | 156 | 62 | 39.7436% |
| France | 983 | 162 | 16.4802% |
| Germany | 1917 | 152 | 7.9291% |
| Greece | 171 | 50 | 29.2398% |
| Hungary | 61 | 9 | 14.7541% |
| Ireland | 55 | 70 | 127.2727% |
| Italy | 424 | 67 | 15.8019% |
| Latvia | 25 | 3 | 12.0000% |
| Lithuania | 90 | 7 | 7.7778% |
| Luxembourg | 34 | 13 | 38.2353% |
| Malta | 20 | 5 | 25.0000% |
| Netherlands | 128 | 114 | 89.0625% |
| Poland | 786 | 74 | 9.4148% |
| Portugal | 42 | 15 | 35.7143% |
| Romania | 919 | 14 | 1.5234% |
| Slovakia | 192 | 3 | 1.5625% |
| Slovenia | 167 | 15 | 8.9820% |
| Spain | 167 | 50 | 29.9401% |
| Sweden | 1026 | 228 | 22.2222% |
| **TOTAL** | **8684** | **1274** | **14.6707%** |

Table Descriptive Summary of Sample

This sample encompasses both successful and unsuccessful takeover endeavours, without differentiation between friendly and hostile transactions, in line with Walkling's perspective (1985) suggesting that bid outcomes often hinge on managerial resistance rather than mere fundamentals. This inclusive approach, as advised by (Ambrose, B.W. and Megginson and W.L. (1992) and (Brar et. al (2009), to incorporate both successful and unsuccessful bids bolster the generalizability and accuracy of predictive models. The primary focus of the M&A sample revolves around control transfer announcements, aiming to examine the Inefficient Management Hypothesis (Cornett et al., 2011). Excluded from consideration are other deal types like stake purchases, exchange offers, privatisations, recapitalisations, repurchases, rumours, self-tenders, and spinoffs.

By incorporating accounting and market data until June 2023, the risk of hindsight bias is mitigated in prediction models for takeovers and tests related to asset pricing, such as the Five Factor Model. To sum up, this extensive and intricate dataset comprising comprehensive financial data from European firms over an extended temporal span will greatly facilitate empirical analyses pertaining to the characteristics of entities prior to their acquisition, all the while alleviating potential challenges stemming from sample selection biases. In the next step, the firm’s year observation of the merged database are reduced due to missing and implausible data.

Table 2 shows the structure of annual sample. Companies are identified as targets if they obtained a bid of acquisition during the prediction period.

Moreover, a multicollinearity analysis is undertaken in order to ascertain whether there are any substantial intercorrelations amongst the autonomous variables (Doumpos et al., 2004). A multicollinearity analysis enables the detection of potential collinearity issues that could adversely impact the outcomes of the regression analysis. Furthermore, each individual variable is either winsorized at the 1% or 5% level or mathematically transformed, with the exception of indicator variables. Winsorizing the outliers rather than outright removing them assists in reducing the bias in the regression results. High degrees of multicollinearity contribute to a deterioration in the explanatory capacity of variables that are relevant to financial traits. Thus, it is imperative to pinpoint the most statistically significant variables and eliminate those that are less impactful (Barnes, (2000); Palepu (1986)).

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Table 3: descrptive construction of database  Table 2 shows the structure of annual sample. Companies are identified as targets if they obtained a bid of acquisition during the prediction period. | | | | |
|  |  |  |  |  |
| YEARS | TOTAL | NT | T | TARGET/TOTAL |
| 2000 | 6719 | 6639 | 80 | 1.19% |
| 2001 | 5663 | 5589 | 74 | 1.31% |
| 2002 | 4885 | 4830 | 55 | 1.13% |
| 2003 | 4866 | 4779 | 87 | 1.79% |
| 2004 | 5030 | 4965 | 65 | 1.29% |
| 2005 | 4871 | 4813 | 58 | 1.19% |
| 2006 | 4846 | 4756 | 90 | 1.86% |
| 2007 | 5010 | 4937 | 73 | 1.46% |
| 2008 | 5063 | 4987 | 76 | 1.50% |
| 2009 | 5108 | 5066 | 42 | 0.82% |
| 2010 | 4833 | 4796 | 37 | 0.77% |
| 2011 | 4603 | 4567 | 36 | 0.78% |
| 2012 | 4511 | 4475 | 36 | 0.80% |
| 2013 | 4302 | 4270 | 32 | 0.74% |
| 2014 | 4124 | 4082 | 42 | 1.02% |
| 2015 | 3988 | 3941 | 47 | 1.18% |
| 2016 | 3933 | 3886 | 47 | 1.20% |
| 2017 | 3787 | 3747 | 40 | 1.06% |
| 2018 | 3658 | 3615 | 43 | 1.18% |
| 2019 | 3575 | 3534 | 41 | 1.15% |
| 2020 | 3741 | 3697 | 44 | 1.18% |
| 2021 | 3448 | 3401 | 47 | 1.36% |
| 2022 | 3276 | 3234 | 42 | 1.28% |
| 2023 | 2678 | 2642 | 36 | 1.34% |
| TOTAL | 106518 | 105248 | 1270 | **1.192%** |

Table Year-wise Description of Database

**Figure 1: Annual Announcement of Takeover**

This graph shows the year-by-year acquisitions announcements

## FINANCIAL CHARACTERISTICS OF TARGETS

A critical first step in creating forecast models is to find variables with high discriminate capacity (Espahbodi (2003)). The thesis focuses on firm- and industry-specific characteristics that were chosen based on their frequency and statistical importance in previous research (Palepu, 1986). Although explanatory coverage is necessary for effectively distinguishing between acquisition targets and non-targets, the inclusion of multiple variables testing similar causes can lead to issues of multicollinearity and reduced practicality due to the computational costs involved (Barnes (1999); Powell (2004)).

Additionally, it is essential to account for the possibility that predicted events are already priced into the market, necessitating the reduction of highly correlated variables (Cremers et al. (2009)). To address these challenges, the research employs principal component analysis, which considers both common and unique variance among the variables (Doumpos et al. (2004)). Despite challenges arising from overlapping factors and contradictory relationships, the study considers different implications of positive and negative signs under various hypotheses, such as the Inefficient Management Hypothesis (Palepu (1986)), allowing for a more comprehensive exploration of potential target firms and their characteristics.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **INDEPENDENT TARGET VARIABLES SELECTION**  Table 4 details the independent variables applied to represent each takeover theory (dependent factors: 1 represents target, 0 represents non-targets). Independent variables of the study are calculated using the components of accounting and stock market data obtained from DataStream. | | | | | |
| ***HYPOTHESIS*** | ***VARIABLES*** | ***EMPIRICAL RESEARCH*** | ***COMPONENTS*** | ***DS CODE*** | ***EXPECTED SIGN*** |
| INEFFICIENT MANAGEMENT | NET PROFIT MARGIN [NPM] | Dietrich and Sorensen (1984), Tsagkanos et al. (2007), Brar et al. (2009) | NET INCOME | WC01751 | - |
| NET SALES | WC01001 |
| ASSET TURNOVER [ATO] | Rege (1984), Bartley and Boardman (1990), Barnes (1999/2000), Doumpos et al. (2004), Brar et al. (2009) | NET SALES | WC01001 | -/+ |
| TOTAL ASSETS | WC02999 |
| EBITDA/Total Assets [ETA] | De and Jindra (2012); EBIT: Castagna and Matolcsy (1985), Barnes (1999/2000); NI: Cremers et al. (2009) | EBITDA | WC18191 | - |
| TOTAL ASSETS | WC02999 |
| EBIT/(TA - CURRENT LIAB.) [ROCE] | Powell (2001/2004), Ouzounis et al. (2009); ROE: Belkaoui (1978), Palepu (1986), Cudd and Duggal (2000), Tsagkanos et al. (2007), Brar et al. (2009) | EBIT | WC18191 | - |
| CURRENT LIABILITIES | WC03101 |
| TOTAL ASSETS | WC02999 |
| (Operating CF - CAPEX)/TA [CFTA] | Belkaoui (1978), Powell (1997/2001/2004), Ouzounis et al. (2009), Brar et al. (2009) | NET OP. CF | WC04860 | - |
| CAPEX | WC04601 |
| TOTAL ASSETS | WC02999 |
| DIVDEND YIELD [DY] | Ouzounis et al. (2009), Brar et al. (2009) | DIVIDEND YIELD | DY | - |
| GROWTH RESOURCES HYPOTHESIS | NET SALES GROWTH [NSG] [1Y] | Espahbodi and Espahbodi (2003), Powell and Yawson (2007), Ouzounis et al. (2009), Brar et al. (2009) | NET SALES | WC01001 | + |
| NET LIQUID ASSETS/TA [LATA] | Palepu (1986), Ambrose and Megginson (1992), Powell (1997/2001/2004), Powell and Yawson (2007), Cremers et al. (2009); Cash/Capital: Brar et al. (2009) | CASH & ST investments | WC02001 | - |
| CURRENT LIABILITIES | WC03101 |
| TOTAL ASSETS | WC02999 |
| GROWTH DUMMY [DRG] | Palepu (1986), Ambrose and Megginson (1992), Cudd and Duggal (2000) | 1- LOW GROWTH,HIGH LIQUIDITY, LOW LEVERAGE, ELSE 0 |  | + |
|
| CURRENT RATIO [CR] | Dietrich and Sorensen (1984), Castagna and Matolcsy (1985), Doumpos et al. (2004), Tsagkanos et al. (2007) | CURRENT ASSETS | WC02201 | -/+ |
| CURRENT LIABILITIES | WC03101 |
| R&D/ TOTAL ASSETS [RDTA] | De and Jindra (2012); R&D/Firm Value: Espahbodi and Espahbodi (2003) | R&D | WC01201 | + |
| TOTAL ASSETS | WC02999 |
| SIZE | TOTAL ASSETS [TA] | Palepu (1986), Powell (1997/2001/2004), Espahbodi and Espahbodi (2003), Powell and Yawson (2007) | TOTAL ASSETS | WC02999 | - |
| MARKET CAPITALIZATION [MCAP] | Palepu (1986), Powell (1997/2001/2004), Espahbodi and Espahbodi (2003), Powell and Yawson (2007) | MARKET VALUE(CAPITAL) | MV | - |
| UNDERVALUATION | TOBIN'S Q ((MV+TL)/TA) [TQ] | Palepu (1986), Powell (1997/2001/2004), Espahbodi and Espahbodi (2003), Powell and Yawson (2007) | MARKET VALUE (CAPITAL) | MV | - |
| TOTAL LIABILITIES | WC03351 |
| TOTAL ASSETS | WC02999 |
| EBITDA/Firm Value (EBITDA/(MV+TD+MI+PS-CE))[EFV] | Bartley and Boardman (1990), Espahbodi and Espahbodi (2003) | EBITDA | WC18198 | - |
| MARKET VALUE (CAPITAL) | MV |
| TOTAL DEBT | WC03255 |
| MINORITY INTEREST | WC03425 |
| PREFERED STOCK | WC03451 |
| CASH & EQ. | WC02005 |
| LEVERAGE | INTEREST COVER [IC] | Dietrich and Sorensen (1984); Barnes (1999/2000);Brar et al. (2009) | EBITDA | WC18198 | + |
| Interest Exp. On Debt | WC01251 |
| TOTAL LIABILITIES/TA [TLTA] | Stevens (1973), Doumpos et al. (2004), Brar et al. (2009), Cremers et al. (2009), De and Jindra (2012) | TOTAL LIABILITIES | WC03551 | - |
| TOTAL ASSETS | WC02999 |
| LIQUIDITY | WORKING CAPITAL/TA [WCTA] | Stevens (1973), Belkaoui (1978), Bartley and Boardman (1990), Barnes (1999/2000) | CURRENT ASSETS | WC02201 | - |
| CURRENT LIABILITIES | WC03101 |
| TOTAL ASSETS | WC02999 |
| ASSET STRUCTURE | TANGIBLE ASSETS(PPE)/TA [PPETA] | Ambrose and Megginson (1992), Powell (2001/2004), Powell and Yawson (2007), Cremers et al. (2009); Intangible Assets/Total Assets: De and Jindra (2012) | TANGIBLE ASSETS (PPE) | WC02201 | + |
| TOTAL ASSETS | WC02999 |
| MOMENTUM | PRICE MOMENTUM 3M [PM] | Brar et al. (2009) | CLOSING PRICE | RI | -/+ |

Table : Independent Target Variable Selection

## FINANCIAL PERFORMANCE ANALYSIS AND INVESTMENT STRATEGIES

A pivotal application of predictive modelling is translating model outputs into implementable trading strategies, allowing empirical assessment of whether forecasted probabilities yield significant risk-adjusted returns Palepu (1986). This study conducts such tests using portfolio simulation approaches common in the empirical finance literature.

Specifically, at the end of each fiscal year, the fitted randomforest classification model generates estimated takeover probabilities for all sample firms. Firms are then assigned to portfolios based on these predictive scores (Espahbodi (2003)). In the "top 25% portfolio", the quartile of stocks with the highest modelled takeover likelihood are selected, following previous studies (Brar et. al (2009)). This portfolio is held for a 12-month holding period with monthly rebalancing, as implemented by Cremers et al. (2009).

To estimate abnormal returns while accounting for potential additional redistributions, this study utilizes monthly total return data from the DataStream database. The monthly returns for each portfolio are calculated using an equally weighted average. The abnormal returns generated by the investment strategy are evaluated through the Fama French 5-Factor model. Screens are applied to the target portfolio to assess whether the inclusion of highly distressed non-target firms (type II errors) reduces investment returns.

The factors are obtained directly from Kenneth French's widely used database French, (2022). While originally derived for the US, note these factors often translate to international markets. As highlighted by Arbel et al. (1983), converting returns to a common currency enables performance comparisons. The Fama French 5-Factor model builds on the original Fama French 3-Factor model (Fama (1993)) by incorporating two additional factors - investment and profitability - derived from the dividend discount model, as proposed by Fama and French in 2014.

Where is the equally weighted return on the portfolio; is the risk-free rate; is the excess return; represents the market risk premium for the month t; is the difference in the portfolio returns between small ad big firms which is an indicator of size premium; is the difference in returns between high value firms and growth firms which is an indicator of value premium; is the difference in portfolio returns between robust and weak firms which indicates profitability premium; and is the difference between conservative and aggressive investments for portfolio returns which signifies investment premium.

# **CHAPTER 4: RESULTS AND FINDINGS**

## DESCRIPTIVE STATISTICS

Table\_ provides descriptive statistics and difference of means hypothesis tests evaluating possible prognosticators of a binary outcome variable (targetdummy) across two stratified groups (targetdummy=0 vs targetdummy=1). A variety of financial ratios and operational measures, including asset turnover (ATO), earnings yield (ETA), dividend yield (DY), and capital ratio (CR), among others, show statistically significant differences across the clusters at p0.01. These findings are consistent with the body of academic research that demonstrates how to distinguish between insolvent and solvent firms with accuracy using profitability, leverage, liquidity, and valuation proportions ((Altman (1968); Ohlson (1980); Zmijewski (1984)).

Analysis of the summary statistics reveals several salient insights into the potential discriminatory power of the examined variables between acquisition targets and non-targets. Multiple metrics provide empirical affirmation for established theoretical perspectives on the pre-acquisition financial characteristics of targets.

In alignment with the inefficient management hypothesis advanced in seminal works by Barnes (2000); and Palepu (1986), targets exhibit significantly inferior profitability and operating performance. This is evidenced by highly significant disparities in means for return on capital employed (ROCE; t=-3.09, p<0.01) as well as the net profit margin ratio (NPM; t=-3.61, p<0.01). As suggested by prior scholars, the results imply acquirers consciously seek out poorly managed companies as prospective targets to create value through superior oversight.

However, contrary to certain predictions on the relevance of firm size (Cremers et al. (2009); Powell (2004)), the analysis indicates negligible variation between targets and non-targets for asset-based proxies such as total assets and market capitalization. Nevertheless, multiple indicators of financial leverage and capital structure align closely with existing hypotheses on the role of financial distress as an antecedent of acquisition ((Barnes, 1999; Palepu, 1986). Metrics such as the debt-to-equity ratio (t=-6.03, p<0.01) and dividend yield (t=-5.04, p<0.01) exhibit significantly higher means for the target cohort, consistent with the notion that heightened leverage renders firms susceptible to takeover.

Furthermore, in congruence with similar empirical studies on valuation discounts (Brar et. al (2009); De and Jindra (2012)), targets demonstrate substantial undervaluation relative to industry peers prior to acquisition across several market-based proxies. For instance, the earnings forecast error (t=-3.67, p<0.01) and sales-per-share growth ratio (t=-2.65, p<0.01) show sizable discounts for acquired firms. Collectively, the results provide credible affirmation for theoretical perspectives positing that acquirers capitalize on underpriced targets to capture valuation gains.

In summary, the analysis yields multiple statistically significant findings that corroborate established hypotheses on the key financial criteria differentiating targets from the broader firm population. The results offer a robust foundation for further predictive modelling aimed at precisely identifying European companies most likely to attract acquisition bids.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **TABLE 5: Descriptive Statistics.**  Table 5 shows the relevant acquisition determinants and their values between 2000 to 2023. | | | | | | | |
|  | **Mean(ALL)** | **Standard Deviation** | **Mean (targetdummy=0)** | **Mean (targetdummy=1)** | **Maximum** | **Minimum** | **t-statistics** | **significance** |
| **GRD(dummy)** | 0.11524187 | 0.319314446 | 0.114745795 | 0.191558442 | 1 | 0 | -8.417673634 | \*\*\* |
| **NPM\_w** | -0.007423474 | 0.091807352 | -0.007386248 | -0.013150446 | 0.12935768 | -0.3339432 | 2.196666834 | \*\* |
| **ATO\_w** | 0.357897499 | 0.550807037 | 0.355618794 | 0.708455581 | 1.7307692 | 0 | -22.44110103 | \*\*\* |
| **ETA\_w** | 0.027713762 | 0.067789388 | 0.027563748 | 0.050792146 | 0.18978814 | -0.10464502 | -11.99274184 | \*\*\* |
| **ROCE\_w** | 0.013002122 | 0.065308698 | 0.012908406 | 0.027419492 | 0.18553739 | -0.14495755 | -7.774903017 | \*\*\*\* |
| **CFTA\_w** | 0.000158005 | 0.049811014 | 0.000150469 | 0.001317343 | 0.10583841 | -0.14321926 | -0.819589357 |  |
| **DY\_w** | 0.648954415 | 1.416484677 | 0.646974812 | 0.953498377 | 4.97 | 0 | -7.572068289 | \*\*\* |
| **NSG\_w** | 0.031448954 | 0.118852807 | 0.031325577 | 0.050429299 | 0.39951086 | -0.16961198 | -5.623954691 | \*\*\* |
| **LATA\_w** | -0.051501675 | 0.185036233 | -0.051149435 | -0.105690575 | 0.54891941 | -0.83218347 | 10.31539366 | \*\*\* |
| **CR\_w** | 0.53787421 | 0.954167334 | 0.535708236 | 0.871089552 | 3.2950802 | 0 | -12.30222904 | \*\*\* |
| **RDTA\_w** | 0.002059295 | 0.007201817 | 0.002043749 | 0.004450899 | 0.03045986 | 0 | -11.6980666 | \*\*\* |
| **MCAP\_w** | 818.4187749 | 3544.18903 | 816.9103473 | 1050.476648 | 28248.58 | 0 | -2.305667268 | \*\* |
| **TQ\_w** | 0.180563093 | 0.276589837 | 0.17941401 | 0.357339198 | 0.79013295 | 0 | -22.53594851 | \*\*\* |
| **EFV\_w** | 0.037279 | 0.229647965 | 0.037057047 | 0.071424509 | 0.74325793 | -0.47638231 | -5.236165233 | \*\*\* |
| **IC\_w** | 169.5045979 | 1607.682802 | 169.729347 | 134.9289891 | 15812 | -1214 | 0.757323225 |  |
| **TLTA\_w** | 0.218654227 | 0.296938708 | 0.217466262 | 0.401411952 | 0.83578307 | 0 | -21.69978497 | \*\*\* |
| **WCTA\_w** | 0.049039762 | 0.117680548 | 0.048911416 | 0.068784691 | 0.41104527 | -0.04457838 | -5.908833836 | \*\*\* |
| **PPETA\_w** | 0.196343016 | 0.287538659 | 0.195252545 | 0.364102164 | 0.828 | 0 | -20.56752915 | \*\*\* |
| **PM3m\_w** | 0.008104699 | 0.109746813 | 0.007974689 | 0.028105607 | 0.29412304 | -0.2363304 | -6.418238386 | \*\*\* |

Table The table indicates significant variables and descriptions of various variables used. \*\*\*, \*\* and \* significant at 1%, 5%, and 10% level, respectively.

## ANALYSIS AND INTERPRETATION

### RESULTS

**CORRELATION RESULTS**

Correlation analysis is commonly employed to study the interrelationships between ratios, providing insights into potential redundancies, synergies and areas warranting further investigation (Gujarati (2009)). Specifically, correlation coefficients quantify the linear associations between ratios, with higher absolute values indicating stronger correlations. However, caution is essential, as correlation does not imply causation as mentioned by Aldrich (1995). Analysis of a dataset of 19 key financial ratios reveals several notable correlations. For instance, gross and net profit margins exhibit a strong correlation, affirming their linkage as profitability metrics. Meanwhile, a moderate correlation between return on assets and equity highlights their distinct insights. Overall, correlation analysis delivers a valuable snapshot into the interconnectivity between ratios, facilitating the identification of complementary versus redundant metrics.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| ***Table 5: CORRELATON MATRIX***  *This table outlines the correlation between each variables.* | | | | | | | | | | | | | | | | | | | |
|  | *GRD* | *NPM\_w* | *ATO\_w* | *ETA\_w* | *ROCE\_w* | *CFTA\_w* | *DY\_w* | *NSG\_w* | *LATA\_w* | *CR\_w* | *RDTA\_w* | *MCAP\_w* | *TQ\_w* | *EFV\_w* | *TLTA\_w* | *WCTA\_w* | *PPETA\_w* | *PM3m\_w* | *IC\_w* |
| GRD | 1.00 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| NPM\_w | -0.28 | 1.00 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| ATO\_w | 0.15 | -0.18 | 1.00 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| ETA\_w | -0.11 | 0.64 | 0.58 | 1.00 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| ROCE\_w | -0.07 | 0.69 | 0.41 | 0.89 | 1.00 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| CFTA\_w | -0.29 | 0.76 | -0.04 | 0.59 | 0.60 | 1.00 |  |  |  |  |  |  |  |  |  |  |  |  |  |
| DY\_w | -0.23 | 0.02 | 0.07 | 0.05 | 0.05 | 0.08 | 1.00 |  |  |  |  |  |  |  |  |  |  |  |  |
| NSG\_w | 0.02 | -0.11 | 0.38 | 0.23 | 0.14 | -0.19 | -0.22 | 1.00 |  |  |  |  |  |  |  |  |  |  |  |
| LATA\_w | 0.15 | 0.09 | -0.85 | -0.55 | -0.47 | -0.07 | -0.23 | -0.29 | 1.00 |  |  |  |  |  |  |  |  |  |  |
| CR\_w | 0.84 | -0.50 | 0.51 | -0.02 | -0.04 | -0.45 | -0.18 | 0.19 | -0.21 | 1.00 |  |  |  |  |  |  |  |  |  |
| RDTA\_w | 0.20 | -0.47 | 0.31 | -0.12 | -0.22 | -0.35 | -0.12 | 0.05 | -0.21 | 0.37 | 1.00 |  |  |  |  |  |  |  |  |
| MCAP\_w | -0.35 | -0.02 | -0.07 | -0.05 | -0.08 | 0.03 | 0.28 | -0.16 | -0.16 | -0.32 | 0.01 | 1.00 |  |  |  |  |  |  |  |
| TQ\_w | -0.03 | -0.39 | 0.88 | 0.35 | 0.18 | -0.21 | 0.16 | 0.32 | -0.88 | 0.40 | 0.35 | 0.11 | 1.00 |  |  |  |  |  |  |
| EFV\_w | -0.19 | -0.26 | 0.24 | -0.01 | -0.02 | -0.25 | 0.02 | 0.04 | -0.47 | 0.04 | 0.07 | 0.03 | 0.43 | 1.00 |  |  |  |  |  |
| TLTA\_w | -0.02 | -0.39 | 0.89 | 0.36 | 0.20 | -0.21 | 0.12 | 0.37 | -0.87 | 0.41 | 0.35 | 0.09 | 0.99 | 0.42 | 1.00 |  |  |  |  |
| WCTA\_w | 0.84 | -0.44 | 0.39 | -0.08 | -0.06 | -0.41 | -0.20 | 0.11 | -0.04 | 0.96 | 0.35 | -0.36 | 0.23 | -0.08 | 0.25 | 1.00 |  |  |  |
| PPETA\_w | 0.11 | -0.39 | 0.82 | 0.37 | 0.15 | -0.25 | 0.05 | 0.42 | -0.76 | 0.46 | 0.33 | 0.05 | 0.90 | 0.38 | 0.92 | 0.25 | 1.00 |  |  |
| PM3m\_w | -0.20 | 0.07 | -0.08 | -0.03 | 0.03 | 0.01 | -0.11 | 0.00 | -0.05 | -0.22 | -0.19 | -0.12 | -0.09 | -0.11 | -0.08 | -0.21 | -0.11 | 1.00 |  |
| IC\_w | 0.29 | 0.59 | -0.39 | 0.18 | 0.24 | 0.37 | -0.14 | -0.17 | 0.58 | -0.06 | -0.23 | -0.23 | -0.73 | -0.60 | -0.75 | 0.10 | -0.64 | -0.01 | 1.00 |

**LOGISTIC REGRESSION**

The receiver operating characteristic (ROC) curve constitutes a salient graphical approach for evaluating the discriminatory performance of binary classification models across varying discrimination thresholds (Fawcett (2006)). A key advantage lies in its independence from the underlying response distribution, rendering it well-suited for imbalanced classification tasks (Bradley (1997)). This study employs ROC analysis to assess the predictive capacity of a logistic regression model in identifying prospective acquisition targets.

A logistic regression model was constructed to predict the dichotomous target variable based on financial predictors identified as theoretically relevant in prior academic literature (Barnes (2000); Cremers et al. (2009); Palepu (1986)). The model's classification ability was scrutinized across diverse thresholds by charting the ROC curve, which plots the true positive rate (TPR) against the false positive rate (FPR) at each threshold.

The area under the ROC curve (AUC) serves as a global metric of model discrimination, with an AUC of 0.5 and 1.0 implying random and perfect discrimination respectively. The obtained AUC of 0.5747 indicates modest discriminatory ability, with the model outperforming random guessing. However, substantial room remains for improving model performance and approaching complete separation of targets and non-targets. The AUC's narrow confidence interval [0.5531, 0.5963] further confirms the statistically significant, albeit limited predictive capacity (z stat: 6.7909; se: 0.0110)

**Figure 3:ROC Curve for Logistic Regression Analysis**

**A graph with a line

Description automatically generated**

**LASSO REGRESSION**

The lasso logistic regression augments the standard logistic model with an L1 regularization penalty to compel sparsity in the estimated coefficients (Tibshirani (1996)). This shrinkage mechanism can potentially improve out-of-sample generalization while inherently performing feature selection (Hastie et al. (2009)). The model predicts the binary target variable based on financial indicators identified as relevant in prior academic studies (Palepu (1986); Powell (2004)).

The ROC curve delineates the inverse relationship between the true positive and false positive rates across the spectrum of classification thresholds. The area under the curve (AUC) provides a composite measure of discriminatory ability, with an AUC of 0.5 and 1 implying random and perfect separation, respectively.

The lasso logistic model attained an AUC of 0.5924, denoting appreciable predictive capacity exceeding random guessing. However, substantial scope remains for enhancing sensitivity and specificity. The narrow confidence band [0.5737, 0.6111] further substantiates the statistically significant, albeit limited forecasting proficiency. (z stat: 9.7263: se: 0.0095)

**Figure 4: ROC Curve for Lasso Logistic Regression Analysis**

**A graph with a line

Description automatically generated**

Analysis of the regularization path plot, commonly termed the lasso path, provides critical insights into the variable selection process induced by the lasso penalty in regression models. As the regularization strength λ increases, the estimated coefficients are progressively shrunk toward zero, with spurious predictors driven to zero while relevant variables are retained (Tibshirani (1996)). Examination of the coefficient trajectories as λ varies facilitates the identification of essential predictors with strong explanatory power versus redundant variables contributing noise. Specifically, predictors whose coefficients persist even at higher λ values are deemed important by the lasso algorithm. In contrast, variables reaching zero and remaining at zero have been removed from the model due to a lack of predictive value. Furthermore, intersecting coefficient paths are indicative of correlation and potential collinearity between predictors. In summary, the lasso path plot enables nuanced revelations into the dynamics of automatic variable selection and discarding of uninformative features. Careful examination of these regularization trajectories supports parsimonious model building by highlighting significant predictors and pruning redundant variables to improve generalization performance.

**Figure 5: Lasso path Plot**

**A graph with numbers and lines

Description automatically generated**

Overall, while the lasso logistic model demonstrates non-trivial predictive validity, its capabilities remain confined. Further refinement of the model specification and predictor space seems necessary to attain robust identification of prospective acquisition targets. Therefore, we use the random forest classification model to attain the desired results.

**RANDOM FOREST CLASSIFICATION**

The receiver operating characteristic (ROC) curve constitutes an insightful graphical technique for evaluating the discriminatory capacity of binary classification models across the spectrum of decision thresholds (Fawcett (2006)). This study employs ROC analysis to gauge the predictive proficiency of a random forest classifier in identifying prospective acquisition targets.

As an ensemble technique, the random forest aggregates predictions across a multitude of decision trees constructed on random samples of the training data (Breiman, (2001)). This technique can yield enhanced predictive accuracy and robustness compared to individual models. The model predicts the binary target variable using key financial indicators identified by prior academic research (Cremers et al. (2009); Powell (2004)).

The ROC curve delineates the inverse relationship between true positive rate (TPR) and false positive rate (FPR) at varying decision thresholds. The area under the curve (AUC) provides a singular metric of overall discriminatory ability, with an AUC of 0.5 and 1 indicating random guessing and perfect discrimination respectively.

The random forest model attained an AUC of 0.6791, denoting appreciable predictive capacity transcending random chance. The narrow confidence interval [0.6623, 0.6958] further substantiates the statistically significant predictive proficiency. (z-stat: 21.0705; se: 0.0085)

**Figure 6: ROC Curve for Random Forest Classification Analysis**

**A graph with a line

Description automatically generated**

Comparative examination of the receiver operating characteristic curves and associated metrics reveals salient insights into the relative predictive proficiency of the logistic regression, lasso-regularized logistic regression, and random forest models for acquisition target identification. On the cardinal metric of the area under the ROC curve (AUC), the random forest classifier attains superior discrimination, with an AUC of 0.6791 versus 0.5747 and 0.5924 for the baseline logistic and lasso logistic models respectively. This denotes the random forest's superior capacity to differentiate between prospective targets and non-targets.

Further evidence of the random forest's dominance emerges from the Z-statistic, which quantifies the deviation of the AUC from the null hypothesis of 0.5 or random guessing in standard error units. The random forest yields a Z-value of 21.0705, notably exceeding the logistic (6.7909) and lasso logistic (9.7263) models.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| ***Table 6 : Model description and significance***  *Table 6 describes the number of observations in each model, the ROC area, the Standard errors, and the significance of the model.* | | | | | | |
|  | ROC |  |  | ASYMPTOTIC NORMAL |  |  |
| model | OBS | AREA | Std.error | [95% conf. interval] | | z-stat |
| logit | 58009 | 0.5747 | 0.011 | 0.55308 | 0.59627 | 6.7909 |
| lasso\_logit | 92081 | 0.5924 | 0.0095 | 0.57367 | 0.61106 | 9.7263 |
| ranfomforest | 99799 | 0.6791 | 0.0085 | 0.66234 | 0.69578 | 21.0705 |

A salient capability of random forest models lies in assessing predictor relevance through a technique known as permutation feature importance (Altmann (2010); Breiman (2001)). By quantifying the decline in model performance upon randomly shuffling a given predictor, this approach reveals variables that are critical for enhancing predictive accuracy. This study employs permutation importance analysis to elucidate the relative significance of financial indicators in random forest-based forecasting of acquisition targets.

The dataset encompasses an array of financial ratios for companies spanning 2000-2022, with a binary target denoting acquired firms. Predictors were selected based on economic rationale and prior academic research (Palepu, 1986; Powell, 2004)). The trained random forest model yields an importance ranking wherein larger values indicate greater relevance of the predictor to model accuracy.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| *Table 7: Statistical description of variables in the Random Forest Classification*  *Table 7 describes the variables of the 23rd model where the model trained for 22 years(2000-2022) shows the mean, standard deviation and feature importance of each variable.* | | | | | |
|  | Mean | Standard Deviation | Minimum | Maximum | Feature Importance |
| GRD | 0.11524187 | 0.319314446 | 0 | 1 | 0.00421762 |
| NPM\_w | -0.007423474 | 0.091807352 | -0.3339432 | 0.12935768 | 0.05313947 |
| ATO\_w | 0.357897499 | 0.550807037 | 0 | 1.7307692 | 0.061033775 |
| ETA\_w | 0.027713762 | 0.067789388 | -0.10464502 | 0.18978814 | 0.048185696 |
| ROCE\_w | 0.013002122 | 0.065308698 | -0.14495755 | 0.18553739 | 0.031726652 |
| CFTA\_w | 0.000158005 | 0.049811014 | -0.14321926 | 0.10583841 | 0.047767638 |
| DY\_w | 0.648954415 | 1.416484677 | 0 | 4.97 | 0.038952815 |
| NSG\_w | 0.031448954 | 0.118852807 | -0.16961198 | 0.39951086 | 0.049267959 |
| LATA\_w | -0.051501675 | 0.185036233 | -0.83218347 | 0.54891941 | 0.045005148 |
| CR\_w | 0.53787421 | 0.954167334 | 0 | 3.2950802 | 0.059065885 |
| RDTA\_w | 0.002059295 | 0.007201817 | 0 | 0.03045986 | 0.017132242 |
| MCAP\_w | 818.4187749 | 3544.18903 | 0 | 28248.58 | 0.137231306 |
| TQ\_w | 0.180563093 | 0.276589837 | 0 | 0.79013295 | 0.055276378 |
| EFV\_w | 0.037279 | 0.229647965 | -0.47638231 | 0.74325793 | 0.024470082 |
| TLTA\_w | 0.218654227 | 0.296938708 | 0 | 0.83578307 | 0.065318466 |
| WCTA\_w | 0.049039762 | 0.117680548 | -0.04457838 | 0.41104527 | 0.034857275 |
| PPETA\_w | 0.196343016 | 0.287538659 | 0 | 0.828 | 0.056776779 |
| PM3m\_w | 0.008104699 | 0.109746813 | -0.2363304 | 0.29412304 | 0.117923711 |
| ln\_IC\_w | 2.810148718 | 2.018683328 | -8.686381 | 9.668525 | 0.052651103 |

Several insights emerge Market capitalization and Price Momentum attain the highest importance scores, highlighting their criticality as valuation anchors in forecasting acquisition likelihood. Conversely, gross debt and profitability ratios receive low importance, suggesting limited marginal predictive value beyond other indicators. However, importance signifies correlation but not necessarily a direct mechanistic relationship (Strobl et al. (2007)).

Overall, permutation importance analysis grants vital revelations into predictor utility for enhancing random forest model performance on this forecasting task. It facilitates the selective retention of salient variables and pruning of expendable predictors to boost parsimony, efficiency and generalizability. Future research should combine these results with economic theory and domain expertise to garner nuanced perspectives on relationships between financial indicators and acquisition probabilities.

### ABNORMAL RETURNS

To evaluate the ability of the predictive models to generate excess returns, the main investment strategy is to buy the target and hold for 12 months. In addition, portfolio returns are calculated as equally weighted monthly return average. The top 25% of the predictions was taken to form the portfolio.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| YEAR | TARGETS | NON-TAREGTS | ACTUAL T | ACTUAL NT |
| 2000 | 18.50% | 7.54% | 13.77% | 14.70% |
| 2001 | -5.30% | -23.04% | 8.98% | 30.67% |
| 2002 | 3.14% | -18.73% | 8.32% | 19.29% |
| 2003 | -9.03% | -2.65% | 47.76% | 17.64% |
| 2004 | 17.54% | -5.86% | 7.47% | 8.50% |
| 2005 | 2.66% | -6.99% | 5.25% | -4.97% |
| 2006 | 27.07% | 8.03% | 32.04% | -9.72% |
| 2007 | 18.54% | 15.30% | 13.54% | -9.85% |
| 2008 | 39.87% | 8.19% | 7.18% | -15.00% |
| 2009 | -2.45% | 6.64% | 14.94% | 22.33% |
| 2010 | -5.53% | 14.93% | -6.29% | 9.20% |
| 2011 | 9.62% | -28.94% | -10.90% | 12.87% |
| 2012 | 16.72% | -8.47% | -1.66% | 6.45% |
| 2013 | 18.22% | 27.67% | 22.50% | -8.23% |
| 2014 | -5.31% | 12.82% | -17.12% | -5.39% |
| 2015 | 10.74% | 25.11% | -30.77% | -20.17% |
| 2016 | -3.21% | 41.18% | 25.54% | 28.49% |
| 2017 | -10.74% | 29.46% | 72.62% | 31.23% |
| 2018 | 12.01% | -3.96% | 8.47% | -7.02% |
| 2019 | 14.78% | -9.79% | 33.74% | -4.05% |
| 2020 | 17.20% | -0.32% | -5.58% | -5.06% |
| 2021 | 3.65% | -13.60% | -7.45% | 6.98% |
| 2022 | 5.69% | 28.47% | -4.21% | 28.05% |
| 2023 | 9.98% | 8.03% | 3.67% | 5.90% |
| AVERAGE | **8.52%** | **4.63%** | **10.08%** | **6.37%** |

The targets portfolio averaged annual returns of 8.52%, compared to only 4.63% for non-targets. The arbitrage strategy yielded excess returns averaging 3.89% per year (t-statistic = 1.27, p-value = 0.21). Although this was not statistically significant at conventional levels, this aligns with prior findings that takeover prediction strategies tend not to produce reliably positive abnormal returns after accounting for risk, as stock prices typically incorporate available information on acquisition likelihood (Cremers et al. (2009); Palepu (1986)).

However, when isolating only the actual ex-post targets and non-targets within the top 25% portfolio, arbitrage returns averaged 10.08% and were significant at the 1% level (t-statistic = 3.18, p-value = 0.004). This substantial outperformance highlights the potential gains from successfully identifying true targets while minimizing misclassified non-targets. The inclusion of type II errors likely diluted the profitability of the overall strategy.

The results suggest the models hold promise in isolating probable takeover targets, although limitations remain in consistently generating alpha for investors. As Brar et al. (2009) note, technical factors like price momentum may provide additional signals to supplement financial ratios. Moreover, integrating macroeconomic and industry data could potentially enhance predictive accuracy over time (Gorton et al. (2009)).

|  |  |  |
| --- | --- | --- |
|  |  |  |
| VARIABLES | TARGETS | NON-TARGETS |
| INTERCEPT | 1.91% | 0.85% |
|  | (4.9)\*\*\* | (-1.26) |
| MKT-RF | 0.56 | 1.0438 |
|  | (2.15)\*\* | (7.59)\*\*\* |
| SMB | -0.053 | 0.682 |
|  | (-0.29) | (2.01)\*\* |
| HML | -0.0504 | -0.2046 |
|  | -0.58 | (-0.36) |
| RMW | -0.385 | -0.2046 |
|  | (-2.57)\*\* | (-2.08)\*\* |
| CMA | 0.176 | 0.0098 |
|  | (-0.64) | -0.05 |
| ADJUSTED R2 | 0.719 | 0.659 |
| OBSEVATION | 279 | 279 |
| F-VALUE | 142.3 | 254.2 |

After adjusting for risk factors, the intercepts show the monthly returns that each portfolio earned above average. At the 1% level, the target portfolio intercept, which is 1.91%, is statistically significant. This shows that investing in takeover targets that were projected to happen produced a monthly abnormal return of 1.91%. The non-target portfolio, in comparison, did not produce appreciable excess returns.

Both portfolios have positive and significant market risk premium (MKT-RF) beta values, indicating exposure to broad market swings. Smaller companies appeared to do better in that portfolio, according to the size factor (SMB), which is positive and significant only for non-targets. For both portfolios, the profitability component (RMW) and the value factor (HML) are inconsequential. CMA, the investment element, is likewise inconsequential. This implies these risk factors did not drive differential performance between the target and non-target investments.

The higher adjusted R-squared for non-targets indicates the model explains a greater proportion of return variation for that portfolio. But the large and significant alpha for targets shows the potential to earn abnormal returns by investing in predicted takeover targets.

# **CHAPTER 5: CONCLUSION**

## SUMMARY OF STUDY

This research presents a significant advancement in the study of modelling acquisition likelihood for European firms, underpinned by a meticulous examination of over 8,000 companies spanning from 2000-2023. By harnessing sophisticated techniques like random forest modelling, the research not only evaluates predictive accuracy but also scrutinizes investment returns. This comprehensive approach yields compelling evidence that financial metrics remain pivotal in forecasting takeover targets, especially in the context of the surging M&A wave in Europe.

One of the standout findings is the superior discriminatory prowess of the random forest model, which notably outperforms logistic and lasso regression alternatives, as corroborated by an impressive AUC of 0.6791. Further testament to the model's efficacy is its investment application: portfolios constructed based on model scores consistently eclipse market performance.

## IMPLICATIONS OF THE FINDINGS

The results of this study have significant effects for both theoretical and practical M&A research. Most significantly, the random forest model's better performance to conventional regression techniques highlights the expanding importance and effectiveness of advanced machine learning techniques in foretelling M&A activity. This shows that academic research methodologies may undergo a paradigm shift that emphasises the incorporation of data-driven procedures. Additionally, the model's predictive scores' demonstrable economic value highlights the real advantages of fusing theoretical research with real-world application, potentially revolutionising the way investment strategies are developed in industries with high M&A activity.

## LIMITATIONS AND FUTURE DIRECTION

Even though this study makes valuable contributions, there are several drawbacks that need to be noted. According to Rossi and Volpin (2004), the sample was limited to publicly traded businesses exclusively, which made it difficult to generalise to other private businesses. Though it was required to remove sector-specific confounds, leaving out utilities and financials diminished industry variety (Jensen & Ruback, 1983). The inability to infer causal linkages from predictive associations using observational data is a significant limitation. By not including all possible takeover decision-makers, such as management incentives or private knowledge, the models are also vulnerable to bias caused by missing variables (Palepu (1986)).

The study did not take into consideration the developing strategic justification for acquisitions as macroeconomic conditions changed since it only used pre-acquisition financials (Gorton et al. (2009)). Regardless of the outcome, using announced deals carries the danger of completion bias (Walkling (1985)). Time-series data surrounding deal announcements should be analysed in future studies to better understand the dynamics of changing acquisition risks. Alternative methods, such as neural networks, could show more predictive ability. Forecasting business transactions may become more practical as datasets multiply, and computational methods improve.

# **CHAPTER 6: REFERENCES**

Aldrich, J. (1995) ‘Correlations genuine and spurious in Pearson and Yule.’, *Statistical Science*, 10(4), pp. 364–376.

Altman, E.I. (1968) ‘Financial ratios, discriminant analysis and the prediction of corporate bankruptcy.’, *The journal of finance*, 23(4), pp. 589–609.

Altmann, A., T.L., S.O., & L.T. (2010) ‘Permutation importance: a corrected feature importance measure.’, *Bioinformatics*, 26(10), pp. 1340–1347.

Ambrose, B.W. and Megginson and W.L. (1992) ‘The role of asset structure, ownership  structure, and takeover defenses in determining acquisition likelihood ’, *Journal of  Financial and Quantitative Analysis, vol. 27, no. 4, pp. 575-589*, 27(4), pp. 575–589.

Andriosopoulos, D., Andriosopoulos, K. and Hoque, H. (2019) ‘The determinants of shareholder value creation in merger and acquisition transactions. ’, *International Review of Financial Analysis*, 62, pp. 174–189.

Arbel, A., C.S., and S.P. (1983) ‘Giraffes, institutions and neglected firms.’, *Financial Analysts Journal, 39(3), 57-63.*, 39(3), pp. 57–63.

Asquith and P. (1983) ‘Merger bids, uncertainty, and stockholder returns’, *Journal of  Financial Economics*, 11(1–4), pp. 51–83.

Barnes, P. (1999) ‘Predicting UK takeover targets: some methodological issues and  an empirical study ’, *Review of Quantitative Finance and Accounting*, 12(3), pp. 283–303.

Barnes, P. (2000) *The identification of U.K. takeover targets using published historical cost accounting data Some empirical evidence comparing logit with linear discriminant analysis and raw financial ratios with industry-relative ratios*.

Barnes, P. (1990) *THE PREDICTION OF TAKEOVER TARGETS IN THE U.K. BY MEANS OF MULTIPLE DISCRIMINANT ANALYSIS*.

Bartley, J.W. and B.C.M. (1990) ‘The relevance of inflation adjusted  accounting data to the prediction of corporate takeovers’, *Journal of Business Finance and Accounting*, 17(1), pp. 53–72.

Bartram, S.M., Griffin, J.M., Lim, Y. and Ng, D.T. (2022) ‘The European Central Bank’s Corporate Sector Purchase Program: Effects on bonds and their issuers.’, *The Review of Financial Studies*, 35(2), pp. 632–673.

Belkaoui, A. (1978) ‘Financial ratios as predictors of Canadian takeovers’, *Journal  of Business Finance and Accounting*, 5(1), pp. 93–107.

Bradley, A.P. (1997) ‘The use of the area under the ROC curve in the evaluation of machine learning algorithms. ’, *Pattern recognition*, 30(7), pp. 1145–1159.

Brar, G., Giamouridis, D. and Liodakis, M. (2009) ‘Predicting European takeover targets’, *European Financial Management*, 15(2), pp. 430–450.

Breiman, L. (2001a) ‘Random forests.’, *Machine learning*, 45(1), pp. 5–32.

Breiman, L. (2001) (2001b) ‘*Random forests. Machine learning, 45(1), 5-32.*’, 45(1), pp. 5–32.

Comment, R. and Schwert, G.W. (1995) ‘Poison or placebo? Evidence on the deterrence and wealth effects of modern antitakeover measures.’, *Journal of Financial Economics, 39(1), 3-43.*, 39(1), pp. 3–43.

Cremers, K.J.M., Nair, V.B. and John, K. (2009) ‘Takeovers and the cross-section  of returns’, *Review of Financial Studies*, 22(4), pp. 1409–1445.

De, S. and Jindra, J. (2012) ‘Why newly listed firms become acquisition targets’, *Journal of Banking and Finance*, 36(9), pp. 2616–2631.

Dietrich, J.K. and and Sorensen, E. (1984) ‘An application of logit analysis to prediction  of merger targets’, *Journal of Business Research*, 12(3), pp. 393–402.

Doumpos, M., Kosmidou, K. and Pasiouras, F. (2004) *Em~etprlotCtK/1 ’Epgovct / Operational Research*.

Espahbodi, H. and E.P. (2003) ‘Binary choice models and corporate  takeover’, *Journal of Banking and Finance*, 27(4), pp. 549–574.

Fairclough, D. and Hunter, J. (1998) ‘The ex-ante classification of takeover targets  using neural networks, published in: Decision Technologies for Computational  Finance: Proceedings of the fifth International Conference Computational Finance (at London Business School, 1997) ’, *London Business School, 1997*, , pp. 381–388.

Fama, E.F., & F.K.R. (1993) ‘ Common risk factors in the returns on stocks and bonds.’, *Journal of financial economics*, 33(1), pp. 3–56.

Fawcett, T. (2006) ‘An introduction to ROC analysis. ’, *Pattern recognition letters*, 27(8), pp. 861–874.

French, K.R. (2022) *French, K. R. (2022). Detail for 5 Factors 2x3. Retrieved from https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data\_library.html*., *https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data\_library.html*

Gorton, G., Kahl, M. and Rosen, R. (2009) ‘Eat or be eaten: a theory of mergers  and firm size ’, *Gorton, G., Kahl, M. and Rosen, R. (2009), Eat or be eaten: a theory of mergers  and firm size, Journal of Finance, vol. 64, no. 3, pp. 1291-1344.*, 64(3), pp. 1291–1344.

Gujarati, D.N. (2009) *Basic econometrics.* . Tata McGraw-Hill Education.

Hasbrouck, J. (1985) ‘The characteristics of takeover targets: Q and other  measures’, *Journal of Banking and Finance*, 9(3), pp. 351–362.

Hastie, T., T.R., & F.J. (2009) ‘The elements of statistical learning: data mining, inference, and prediction.’, *Springer Science & Business Media*

IMAA (2022) *IMAA. (2022). M&A Statistics - Worldwide, Regions, Industries & Countries. Institute for Mergers, Acquisitions and Alliances (IMAA). https://imaa-institute.org/m-and-a-statistics/*., *IMAA. (2022). M&A Statistics - Worldwide, Regions, Industries & Countries. Institute for Mergers, Acquisitions and Alliances (IMAA). https://imaa-institute.org/m-and-a-statistics/*

Jensen, M.C. (1986) ‘Agency costs of free cash flow, corporate finance, and  takeovers’, *American Economic Review*, 76(2), pp. 320–330.

Jensen, M.C. and Ruback, R.S. (1983) ‘The market for corporate control: The scientific evidence.’, *Journal of Financial economics*, 11(1–4), pp. 5–50.

Lo, A.W. (1986) ‘Logit versus discriminant analysis: A specification test and application to corporate bankruptcies.’, *Journal of Econometrics*, 31(2), pp. 151–178.

Manne, H.G. (1965) ‘Mergers and the market for corporate control’, *Journal of  Political Economy*, 73(2), pp. 110–120.

Marris, R. (1964) The economic theory of managerial capitalism, *Free Press of  Glencoe*,

Mergermarket. (2022) *Global and Regional M&A Report 2021*.

Mueller, D.C. (1969), A. theory of conglomerate mergers, Q.J. of E. vol. 83, no. 4, pp. 643-659. (1969) ‘A theory of conglomerate mergers’, *Quarterly Journal of  Economics*, 83(4), pp. 643–659.

Ohlson, J.A. (1980) ‘Financial ratios and the probabilistic prediction of bankruptcy.’, *Journal of accounting research*, , pp. 109–131.

Ouzounis, G., Gaganis, C. and Zopounidis, C. (2009) ‘Prediction of acquisitions and portfolio returns’, *International Journal of Banking, Accounting and Finance*, 1(4), pp. 381–406.

Palepu, K.G. (1986) *PREDICTING TAKEOVER TARGETS A Methodological and Empirical Analysis*.

Powell, R.G. (2004) ‘Takeover prediction models and portfolio strategies: a  multinomial approach’, *Multinational Finance Journal*, 8(1–2), pp. 35–74.

Powell, R.G. (2001) ‘Takeover prediction and portfolio performance: a note’, *Journal of Business Finance and Accounting*, 28(7–8), pp. 993–1011.

Powell, R.G. (1997) ‘Modelling takeover likelihood’, *Journal of Business Finance and  Accounting*, 24(7–8), pp. 1009–1030.

Powell, R.G. and Y.A. (2007) ‘ Are corporate restructuring events driven by  common factors? Implications for takeover prediction’, *Journal of Business Finance  and Accounting*, 34(7–8), pp. 1169–1192.

Rege, U.P. (1984) ‘Accounting ratios to locate takeover targets’, *Journal of Business  Finance and Accounting* , 11(3), pp. 301–311.

Rossi, S. and Volpin, P.F. (2004) ‘Cross-country determinants of mergers and acquisitions.’, *Journal of Financial Economics.*, 74(2), pp. 277–304.

Simkowitz, M. and M.R.J. (1971) ‘A discriminant analysis function for  conglomerate targets’, *Southern Journal of Business*, 6(1), pp. 1–15.

Singh, A. (1975) ‘Takeovers, economic natural selection, and the theory of the firm:  evidence from the post-war United Kingdom experience’, *Economic Journal*, 85(339), pp. 497–515.

Song, M.H. and W.R.A. (1993) ‘The impact of managerial ownership on  acquisition attempts and target shareholder wealth’, *Journal of financial and  quantitative analysis*, 28(4), pp. 439–458.

Stevens, D.L. (1973) ‘Financial characteristics of merged firms: a multivariate  analysis’, *Journal of Financial and Quantitative Analysis*, 8(2), pp. 149–158.

Strobl, C., Boulesteix, A. L., Zeileis A. and Hothorn, T. (2007) ‘Bias in random forest variable importance measures: Illustrations, sources and a solution.’, *BMC bioinformatics*, 8(1), pp. 1–21.

Taussig, R.A. and H.S.L., I. (1968) ‘Cash take-overs and accounting  valuations’, *Accounting Review*, 4(1), pp. 68–74.

Tibshirani, R. (1996) ‘Regression shrinkage and selection via the lasso’, *Journal of the Royal Statistical Society: Series B (Methodological)*, 58(1), pp. 267–288.

Tsagkanos, A., Georgopoulos, A. and Siriopoulos, C. (2007) ‘Predicting Greek  mergers and acquisitions: a new approach’, *International Journal of Financial  Services Management*, 2(4)

Tzoannos, J. and S.M. (1972), M. and takeovers: the financial characteristics of the companies involved, J. of B.F. vol. 4, no. 3, pp. 5-16. (1972) ‘Mergers and takeovers: the financial  characteristics of the companies involved’, *Journal of Business Finance*, , pp. 5–16.

Walkling, R.A. (1985) ‘Predicting tender offer success: A logistic analysis. ’, *Journal of Financial and Quantitative analysis, 20(4), 461-478.*, 20(4), pp. 461–478.

Zmijewski, M.E. (1984) ‘Methodological issues related to the estimation of financial distress prediction models. ’, *Journal of Accounting research*, , pp. 59–82.

# **CHAPTER 7: APPENDICES**

[www.contactout.com](http://www.contactout.com/)

finalscout.com

mailscoop.io