Private Equity-Backed Public-to-Private Transaction Activity: A Machine Learning Approach to Quarterly Deal Volume Prediction in the US

Sebastian Kwiedor

Dissertation written under the supervision of professor Dan Tran.

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

Title: Private Equity-Backed Public-to-Private Transaction Activity: A Machine Learning Approach to Quarterly Deal Volume Prediction in the US

Author: Sebastian Kwiedor

Abstract

This thesis addresses a critical gap in private equity research by developing quantitative models to forecast quarterly PE-backed public-to-private (P2P) transaction volumes in the United States. While extensive literature exists on firm-level takeover prediction, systematic approaches to forecasting aggregate PE deal activity remain largely unexplored. This study employs machine learning techniques to predict quarterly P2P deal counts using a comprehensive dataset spanning 1986-2024, incorporating 152 macroeconomic and financial variables.

The methodology implements three distinct modeling paradigms—Lasso regression, Random Forest, and XGBoost—within a rigorous time-series cross-validation framework. Variables were systematically processed through principal component analysis to address multicollinearity while preserving economic interpretability. Feature engineering incorporated temporal dynamics through Fourier transforms, momentum indicators, and multi-scale rolling statistics.

Results demonstrate modest but consistent predictive improvements over naive baselines, with mean absolute error reductions of 3-9.8% and R^2 values ranging from 0.145-0.254. Lasso regression emerged as the most balanced approach, maintaining interpretability while achieving competitive accuracy. The analysis reveals strong path dependence in PE activity, with lagged deal counts and momentum indicators consistently ranking as top predictors. Multi-scale temporal patterns, including seasonal and decade-long cycles, contribute significantly to forecast accuracy.

Keywords: Private Equity, Leveraged Buyout, Buyout, Public-to-Private, Going-Private, Machine Learning

Título: Atividade de Transações Public-to-Private Apoiadas por Private Equity: Uma Abordagem de Machine Learning para Previsão de Volume Trimestral de Transações nos Estados Unidos

Autor: Sebastian Kwiedor

Resumo

Esta tese aborda uma lacuna crítica na pesquisa de private equity ao desenvolver modelos quantitativos para prever volumes trimestrais de transações public-to-private (P2P) apoiadas por PE nos Estados Unidos. Embora exista extensa literatura sobre previsão de aquisições ao nível da empresa, abordagens sistemáticas para prever a atividade agregada de transações de PE permanecem largamente inexploradas. Este estudo emprega técnicas de machine learning para prever contagens trimestrais de transações P2P usando um conjunto de dados abrangente de 1986-2024, incorporando 152 variáveis macroeconômicas e financeiras.

A metodologia implementa três paradigmas de modelagem distintos—regressão Lasso, Random Forest e XGBoost—dentro de uma estrutura rigorosa de validação cruzada em séries temporais. As variáveis foram sistematicamente processadas através de análise de componentes principais para abordar a multicolinearidade preservando a interpretabilidade econômica. A engenharia de características incorporou dinâmicas temporais através de transformadas de Fourier, indicadores de momentum e estatísticas móveis multi-escala.

Os resultados demonstram melhorias preditivas modestas mas consistentes sobre baselines ingênuas, com reduções de erro absoluto médio de 3-9.8% e valores de R^2 variando de 0,145-0,254. A regressão Lasso emergiu como a abordagem mais equilibrada, mantendo interpretabilidade enquanto alcançava precisão competitiva. A análise revela forte dependência de trajetória na atividade de PE, com contagens defasadas de transações e indicadores de momentum consistentemente classificados como preditores principais. Padrões temporais multi-escala, incluindo ciclos sazonais e decenais, contribuem significativamente para a precisão das previsões.

Palavras-chave: Capital Privado, Aquisição Alavancada, Aquisição, Fechamento de Capital, Privatização, Aprendizado de Máquina

Table of Contents

List of Tables

List of Figures

List of Abbreviations

ACF	Autocorrelation Function
ADF	Augmented Dickey-Fuller Test
CRSP	Center for Research in Security Prices
CV	Cross Validation
FRED	Federal Reserve Economic Data
HP	Hyperparameter
IPO	Initial Public Offering
KPSS	Kwiatkowski-Phillips-Schmidt-Shin Test
MAE	Mean Absolute Error
M&A	Mergers & Acquisitions
LBOs	Leveraged Buyouts
OOS	Out of Sample
PACF	Partial Autocorrelation Function
PCA	Principal Component Analysis
PE	Private Equity
P2P	Public-to-private
SBO	Secondary Buyout
SDC	Securities Data Company
SE	Standard Error
STD	Standard Deviation

1. Introduction

1.1 Background and Context

Private equity (PE) grew from a niche corner of finance (Almufti et al., 2024) into a major force in global mergers and acquisitions (M&A) (BCG, 2024). A key demonstration of this influence is the prevalence of public-to-private (P2P) transactions, which are acquisitions in which a publicly traded company is bought out and taken private by PE investors often via a leveraged buyout (LBO) (DWF, 2020). These take-private deals account for some of the largest acquisitions in history, underscoring their economic importance. During the 1980s wave of Buyouts Jensen famously argued that LBOs represented a new organizational form that could better address agency problems than the traditional public corporation (M. C. Jensen, 1989; M. C. Jensen, 1986). Since then, the size of the PE market has grown significantly, with global capital commitments to buyout funds increasing from just a few hundred million dollars in the early 1980s to hundreds of billions by the 2000s (Kaplan & Strömberg, 2009). With assets under management expanding from $576 billion in 2000 (Fernández Tamayo et al., 2023) to $8.2 trillion by Q2 2023, representing a 22% annual growth rate in recent years (McKinsey Global Private Markets Review 2024, 2024), the historical rise of PE and prominence of P2P deals highlight the importance of understanding what drives this form of acquisition activity over time.

PE deal flow including P2P transactions, tends to be highly cyclical, closely tracking broader economic and financial conditions (Phalippou, 2017). Observers have long noted that buyout activity comes in waves, surging during periods of economic expansion and staggering during downturns. For example, the late 1980s LBO boom, the 2001-2006 buyout surge each concurred with favorable macroeconomic environments, characterized by strong stock markets, low interest rates, and readily available debt financing. Each boom cycle was followed by sharp declines in deal activity amid economic slowdowns or credit crises (Acharya et al., 2007).

Empirical research has linked these patterns to fundamental drivers e.g. cheap credit and high liquidity fuel larger and more frequent buyouts, whereas tight credit conditions constrain PE activity (Axelson et al., 2013). These boom-bust fluctuations indicate that P2P activity is not random but rather is systematically influenced by macroeconomic and financial market cycles.

1.2 Research Problem and Motivation

Despite PE activity's cyclicality and potential macro drivers, most academic research has examined takeovers at the firm level. Over the past few decades, a rich literature has predicted which companies will become takeover targets using firm-specific features and financial statistics. Palepu (1986) used logistic regression to identify potential acquisition targets. Many subsequent studies have refined such models (Brar et al., 2009), but often exclude non-target firms, potentially missing key determinants. Similar models predict P2P targets, emphasizing stable cash flows, undervaluation, and governance factors (Lehn & Poulsen, 1989; Weir et al., 2005). These contributions set the foundation of why certain firms are acquired, but they inherently treat takeovers as isolated rare events in a cross-sectional setting. Little attention has been paid to forecasting aggregate deal activity over time. Studies of merger and buyout waves (Becketti, 1986; Harford, 2005; Martynova & Renneboog, 2008) usually explain previous surges rather than anticipate future surges or downturns. Thus, systematic, quantitative approaches to anticipate macro-level P2P deal activity trends are lacking.

There is no comprehensive literature on "How many P2P deals will occur next quarter?" or "Will the total value of buyouts rise or fall in the coming months?" This gap is significant and addressing it would advance merger wave analysis to prediction. Predicting aggregate PE transaction flow is economically valuable. Predicting market trends can help PE firms plan fundraising, deployment, and investment. With record levels of dry powder (uncommitted capital) exceeding $2 trillion globally in 2024 (Edlich et al., 2025), PE firms are under pressure to find attractive prospects in competitive marketplaces. Forecasting transaction activity helps investment banks and advisory businesses similarly since it helps them to devote resources to areas of expected development. Furthermore, knowing the larger dynamics behind take-private waves will help public firm boards and executives to have insightful knowledge that might guide defensive actions or proactive discussions in times of increasing buyout activity.

1.3 Research Objectives and Questions

Instead of firm-level targets, this research predicts PE market trends, specifically quarterly P2P deals. The study uses theory- and industry-based buyout predictors. This study is guided by the central research question: Can machine learning techniques be employed to predict private equity-backed public-to-private (P2P) activity in the United States, and what macroeconomic, financial, and market variables drive quarterly fluctuations in such deals? How does model performance vary across different market regimes, including periods of financial distress and private equity booms? This study answers these questions using machine learning for time series forecasting. Analysis compares linear models with regularization (Lasso regression) to ensemble tree-based methods (Random Forest and XGBoost) to determine the best modeling method. Principal component analysis (PCA) reduces macroeconomic and financial predictor dimensionality while preserving content. Using a rigorous time-series cross-validation framework ensures that models are evaluated on predictive power, not historical pattern (over)fit. To my knowledge, this is one of the first studies to undertake predictive modeling of PE market activity at a quarterly frequency. Instead of focusing on why particular companies are taken private or on previous efforts to predict individual takeover targets, it was attempted to predict when the market will see more (or fewer) such take-private deals.

1.4 Thesis Structure

This thesis has seven chapters following this introduction. Chapter 2's extensive literature review outlines the identified strands in PE research and gives an overview of the academic state of M&A and PE transaction prediction. The chapter concludes by identifying gaps in existing literature this thesis addresses. Chapter 3 breaks down the sample construction of PE-backed P2P transactions utilizing the Securities Data Company (SDC) Platinum database. It discusses quarterly deal aggregation, PermID matching and keyword searches to verify PE-backed deals, and CRSP delisting data to validate transactions. Quarterly deal counts are compared to previous studies, seasonality and structural breaks are discussed, and feature selection is explained. Methodology, preprocessing, and feature engineering are in Chapter 4. Principal component analysis (PCA) on variable clusters reduces dimensionality, time-series cross-validation avoids look-ahead bias, and Lasso Regression, Random Forest, and XGBoost are described. The chapter concludes with forecast performance metrics. Chapter 5 presents results, beginning with model diagnostics verifying the forecasting approach validity. It provides detailed performance comparison, analyzing linear versus tree-based methods and assessing feature engineering impact on prediction quality. The chapter explores feature importance and key P2P activity drivers, interpreting findings economically. Visualizations illustrate how models track actual deal counts and perform during market extremes. Chapter 6 discusses findings critically, interpreting results within PE investment cycles and relating them to existing literature. The chapter acknowledges study limitations, addressing data constraints, methodological challenges, and potential omitted variables. Chapter 7 summarizes findings, discusses theoretical and methodological advances, and suggests future research.

2. Literature Review

2.1 Identified Strands in Private Equity Research

During the extensive literature review the author identified six main strands, each addressing various facets of the academic research in PE.

Strand 1: Value Creation and Structures in Private Equity

The first strand looks at how structural systems and investments made by PE firms create value. Analyzing whether they outperform their counterparts and looking at the structural elements of PE investments (including possible agency conflicts) this study stream investigates the performance and operational transformation of PE-backed companies. Kaplan & Strömberg (2009) provide evidence that LBOs generate operational gains through financial and governance engineering that concentrates ownership. Their work demonstrates that value creation occurs primarily through improved operational efficiency rather than solely financial engineering. Lerner et al. (2011) further support this by showing that PE ownership sustains long-term investments, evidenced by stable innovation outputs and increased high-impact patents post-buyout, challenging critics who argue that PE firms sacrifice long-term value for short-term gains. The role of leverage in value creation has been extensively studied within this strand. Guo et al. (2011) found that buyouts from 1990 to 2006, although less aggressively leveraged than their 1980s predecessors, still created value through tax benefits, market gains, and moderate operational improvements. However, Axelson et al. (2013) highlighted that PE value creation is significantly influenced by credit conditions, where easy access to debt enhances leverage and acquisition prices but also heightens agency conflicts between different stakeholders. Agency theory plays a crucial role in understanding PE structures. Phalippou et al. (2018) examined fees charged by PE firms to their portfolio companies, raising important questions about the alignment of interests and the role of discretionary compensation in governance structures. This research reveals potential conflicts between the stated objectives of value creation and the actual fee structures that may prioritize PE firm profits over portfolio company performance.

Strand 2: Criteria and Patterns in PE Target Selection

The second strand examines how PE firms choose companies based on financial measures, market positioning, and other factors. This basic study is needed to understand PE deal flow trends and what distinguishes successful targets from other investments. Early research by Le Nadant and Perdreau (2006) indicated that French LBO targets were less indebted and had higher business risk than their counterparts, suggesting PE funds prefer companies with leverage and risk-adjusted earnings. Predictive modeling is becoming more accurate at identifying takeover targets. Brar et al. (2009) improved European M&A forecasts by adding technical aspects to acquisition likelihood models. Their research revealed that combining technical market indicators with financial analysis can improve target identification. Recent target selection studies show evolving trends. After focusing on operational adjustments rather than financial restructuring, Cohn et al. (2022) found that PE firms target underperforming corporations with turnaround potential or private organizations with high growth potential but limited finance. This change reflects PE sector growth and competition for traditional goals. In this strand, frontier development is target selection using Machine Learning and advanced analytics. Schneider (2022) used machine learning to identify undervalued high-potential companies by analyzing industry performance and financial data, emphasizing the expanding use of technology in investment strategies and deal sourcing.

Strand 3: Performance, Risk, and Persistence in PE Funds

The third strand examines PE funds' performance itself, benchmarking their returns against public markets and assessing whether they consistently deliver superior risk-adjusted performance. Early studies produced mixed findings regarding PE performance relative to public markets. Kaplan and Schoar (2005) provided evidence of strong performance persistence in PE funds, finding that successful funds were more likely to raise larger follow-on funds and continue delivering superior returns. However, Phalippou and Gottschalg (2009) challenged these findings, revealing significant underperformance compared to public benchmarks due to biases in net asset value reporting and high fees. Harris et al. (2014) found that US buyout funds historically outperformed public markets by 20-27%, while L'her et al. (2016) benchmarked PE buyout funds against size-, sector-, and leverage-adjusted indices and found no significant risk-adjusted outperformance. These conflicting results highlight the importance of appropriate benchmarking and risk adjustment methodologies. The persistence of fund performance has emerged as a particularly important research question. Braun et al. (2017) observed a significant decline in performance persistence in buyout investments over time, attributing this to increased competition and commoditization of the PE market.

Strand 4: Exit Strategies and Financial Outcomes

The fourth strand investigates the various exit options used by PE firms and their financial outcomes, concentrating on the factors that impact whether firms pursue initial public offerings (IPO), strategic sales, or secondary buyouts (SBO). This research is crucial for understanding the complete PE investment cycle and the realization of returns. The choice of exit strategy significantly influences financial outcomes and is often driven by market conditions and company characteristics (S. N. Kaplan & Sensoy, 2015). Reverse leveraged buyouts tend to outperform traditional IPOs initially but may face sustainability challenges due to increased financial burdens, especially in highly leveraged cases (Cao & Lerner, 2008). Jenkinson & Sousa (2015) found that PE firms are more likely to pursue SBOs when debt conditions are favorable, leveraging cheap credit availability, while opting for IPOs during rising equity markets to capitalize on higher valuations. This strategic approach to exit timing demonstrates sophisticated market timing capabilities. The context and motivations behind exits play a crucial role in determining outcomes. Degeorge et al. (2016) analyzed SBO performance and found that transactions executed under investment pressure often destroy value.

Strand 5: Long-Term Societal Impact of PE

The fifth strand addresses the broader economic and social implications of PE ownership, analyzing how PE affects employment, productivity, and overall economic welfare. This research has gained increased attention as PE's influence on the economy has grown. Davis et al. (2014) analyzed U.S. PE buyouts and found that these transactions lead to modest net job losses but significant productivity gains, driven by accelerated closure of low-productivity units and increased job creation at high-productivity establishments. This research challenges simplistic narratives about PE's employment effects by revealing the complex reallocation processes involved. In healthcare settings, research has raised concerns about PE's impact on service quality. Gupta et al. (2021) found that PE ownership of nursing homes was associated with increased short-term Medicare patient mortality and reduced nurse availability, raising ethical questions about profit incentives in essential services.

Strand 6: Selection Criteria and Investment Decisions in PE Funds

The sixth strand focuses on the decision-making processes of institutional investors (limited partners) in selecting PE funds, examining factors that influence capital allocation decisions and the effectiveness of due diligence processes. Da Rin and Phalippou (2017) surveyed 249 limited partners and found that larger LPs engage in more specialized due diligence and monitoring activities, which may explain their better investment performance. This research highlights the importance of investor sophistication and resources in achieving superior outcomes. Goyal et al. (2021) studied LP selection of PE funds and found that LPs often chase past performance but also allocate capital to first-time and young general partners, suggesting a balance between proven track records and emerging opportunities in allocation decisions. Fernández Tamayo et al. (2023) notably developed a machine learning approach for PE fund selection.

2.2 Predicting M&A and PE Transactions: Empirical Insights

The academic literature addressing PE backed P2P transactions, their determinants, and their prediction encompasses multiple interconnected research areas. This comprehensive body of work provides the foundation for understanding both the phenomenon of P2P deals and the methodological approaches suitable for forecasting such activity.

The theoretical foundations for P2P activity rest on several key hypotheses. Jensen's (1989) influential argument that traditional public corporations were losing relevance positioned PE and LBOs as superior mechanisms for addressing agency conflicts and cash flow inefficiencies. This built on earlier work by Jensen and Meckling (1976) demonstrating how concentrated ownership and debt financing align managerial incentives with value creation. The undervaluation hypothesis suggests PE firms systematically target companies trading below intrinsic value, creating opportunities for operational improvements and profitable exits (Kaplan, 1989). Empirical research has revealed mixed support for these theoretical foundations.

Lehn and Poulsen (1989) found evidence supporting the free cash flow hypothesis, identifying firms with substantial cash generation but limited investment opportunities as attractive buyout candidates. Weir et al. (2005) documented different patterns in UK versus US markets. This heterogeneity suggests that P2P determinants may vary across time periods, market conditions, and institutional environments. Market-level drivers have emerged as increasingly important determinants of P2P activity. The availability and cost of credit play fundamental roles in enabling leveraged buyouts. Axelson et al. (2013) demonstrated that buyout leverage and pricing are strongly related to debt market conditions, with easy credit leading to higher leverage ratios and acquisition prices. Their analysis revealed that economy wide credit conditions are the primary determinant of leverage in buyouts, whereas firm-specific factors matter relatively little. The cyclical nature of PE investments has become well-documented, with deal flow tending to be highly procyclical, surging during periods of economic expansion and abundant capital and retreating during downturns. An important contribution comes from Haddad et al. (2017), finding that declines in the economy-wide risk premium are the primary driver of PE buyout waves. Their research revealed that the equity risk premium dominates credit conditions in explaining fluctuations in buyout activity, with low risk premiums increasing the present value of future performance gains and lowering the cost of illiquidity.

Predicting financial market activity and merger waves remains a longstanding challenge in academic finance, as researchers develop various approaches to forecast aggregate deal volumes and understand M&A cycle timing. Harford (2005) provided seminal insights into merger wave formation, arguing that industry-level waves are triggered by economic, regulatory, or technological shocks but only materialize when sufficient capital liquidity is available to finance transactions. His framework suggests that favorable financing environments serve as catalysts transforming potential merger opportunities into actual deal activity.

Recent research has begun developing explicit forecasting models for aggregate M&A activity. Bonaime et al. (2018) examined how policy and economic uncertainty affect merger activity, finding that heightened political and regulatory uncertainty strongly predicts lower M&A volume with approximately a one-year lag. More sophisticated approaches have emerged, with Ojea Ferreiro et al. (2022) proposing Mixed Data Sampling count models to forecast monthly M&A deal numbers, demonstrating significant improvements in forecast accuracy compared to traditional time-series models. Degen et al. (2024) utilized ChatGPT to analyze earnings call transcripts, extracting forward-looking M&A sentiment signals from management discussions. Their market-wide M&A Sentiment Score showed strong predictive power for future deal activity, even after controlling for traditional macroeconomic indicators, highlighting the potential for incorporating qualitative information and alternative data sources into forecasting models. The methodological landscape for studying and predicting PE activity has evolved significantly, reflecting advances in econometric techniques and increasing data availability. Traditional approaches centered around logistic regression models (compare Palepu 1986) for predicting individual buyout targets. Despite methodological advances, significant limitations remain in existing approaches. Many studies focus on explaining historical patterns rather than generating actionable forecasts. Sample sizes are often limited by the relative rarity of PE transactions, constraining model complexity that can be reliably estimated. The non-stationary nature of financial time series poses additional challenges, as relationships that hold historically may break down during structural changes or crisis periods.

2.3 Research Gap and Contribution

The literature review reveals several important gaps in the author's understanding of PE-backed P2P transaction patterns and forecasting methodologies. While individual components of this research area have received attention, the specific challenge of predicting aggregate quarterly P2P deal activity remains largely unaddressed in academic literature. The most significant gap lies in the lack of market-level predictive models for P2P deals specifically. The existing literature on PE cycles, while informative about the drivers of buyout activity, has been primarily descriptive rather than predictive. Studies like Haddad et al. (2017) and Axelson et al. (2013) provide valuable insights into what causes PE booms and busts, but they do not translate these insights into operational forecasting models that could be used to anticipate future deal activity. This gap is particularly important given the practical relevance of such forecasts for market participants. Methodologically, most prior research has relied on relatively simple statistical techniques or focused on cross-sectional analysis rather than time-series forecasting. The application of modern machine learning methods to PE research remains limited, despite their demonstrated success in other areas of financial forecasting. This thesis makes several theoretical and methodological contributions to literature. It extends perception of PE market dynamics by demonstrating which macroeconomic and financial market variables are most predictive of quarterly deal flow. The practical relevance of this research extends beyond academic contributions, as PE firms, institutional investors, and policymakers could benefit from improved ability to anticipate deal cycles, informing strategic decisions about fund-raising timing, deployment strategies, and allocation decisions.

3. Data

3.1 Construction of the sample of PE backed P2P transactions

Because this research centers on PE-backed P2P transactions in the US, the initial step in data collection involves identifying these deals. Data access was facilitated through the Workspace platform provided by the London Stock Exchange Group. Specifically, the SDC Platinum database was employed, which is widely utilized in prior academic research, providing extensive historical transaction data spanning several decades. Among the various datasets available for M&A research, SDC Platinum is widely recognized as the leading source in empirical studies, owing to its extensive coverage and detailed categorization of corporate transaction types (Wharton Research Data, 2025). Researchers querying the database can select from six distinct universes including PE, Infrastructure and M&A. Importantly, each universe within the database offers a distinct set of available fields, requiring queries to be tailored to the specific data structure of each universe. The PE Universe, for instance, does not include the "Going-Private-Flag" or any equivalent variable that would allow for the identification of companies transitioning from public to private ownership. As such, it is not well suited for the purpose of this research. In contrast, the M&A Universe is particularly appropriate for identifying and classifying public-to-private transactions. It offers access to more than 2,000 variables, including explicit flags for P2P deals, company identifiers such as CUSIP, and detailed information about the acquiror at multiple levels of ownership. While Bargeron et al. (2008) exclude spin-offs, share repurchases, exchange offers, self-tenders, recapitalizations, and privatizations, and Doidge et al. (2017) requires 100% post-deal ownership, these criteria were not applied in this study. When tested, both approaches excluded too many valid transactions compared to unfiltered results, primarily due to incomplete data fields in SDC. Therefore, a more inclusive approach was adopted to capture a comprehensive sample. Tables 5 and 6 in the Appendix present the final query specifications and variable list.

Identification of PE backed deals

Distinguishing private equity involvement requires multiple methods due to inconsistent firm identification in SDC. Compiling a thorough list of PE companies came first, following Officer et al. (2010). But instead of limiting the sample to well-known PE companies found by internet searches or industry rankings, this study used the whole SDC Private Equity Fundraising Database, which records every money raised since 1960. This comprehensive approach ensures the identification of PE involvement across the full spectrum of PE firms. The fundraising database provides systematic coverage of the entire PE ecosystem, capturing Firm Names and 12,430 distinct PE firm identifiers (PermIDs). The identification process examines six ownership levels for each transaction:

- Acquiror
- Acquiror Immediate Parent
- Acquiror Ultimate Parent
- Investor
- Investor Immediate Parent
- Investor Ultimate Parent

This hierarchical approach captures complex ownership structures common in PE transactions, where the direct acquiror may be a special purpose vehicle while the actual PE sponsor appears at higher ownership levels. Initial matching against the PE firm universe flagged 2,384 transactions as PE-backed. However, manual inspection revealed systematic undercounting, particularly in earlier periods where PermID coverage is incomplete. To address this limitation, the identifier-based approach was supplemented with keyword analysis of firm descriptions, following Fidrmuc et al. (2013). The acquiror short business description columns were searched at all acquiror ownership levels for PE-related terms including "private equity firm," "private equity fund," and variant spellings identified through manual review. This text-based filter identified an additional 233 PE-backed transactions, increasing the PE-backed sample to 2,617 deals while 25,157 transactions showed no PE involvement.

Identification of Going Private Transactions

The SDC "Going Private Flag" suffers from documented reliability issues. As Officer et al. (2010) note, the flag's assignment criteria remain opaque, automatically marking any transaction involving financial sponsors as "going private" regardless of the target's actual delisting status. This overinclusive definition necessitates external validation through delisting records. SDC transactions were matched to the CRSP database using a multi-step procedure. First, a comprehensive linking table from the CRSP/Compustat Merged (CCM) database was constructed, mapping 6-digit NCUSIPs (CCM + CRSP Names Table Merge) to CRSP's permanent identifiers (PERMNO) while preserving temporal validity through link date restrictions. This temporal matching is critical, since CUSIP assignments change over time, and incorrect temporal alignment produces false matches.

The matching algorithm required:

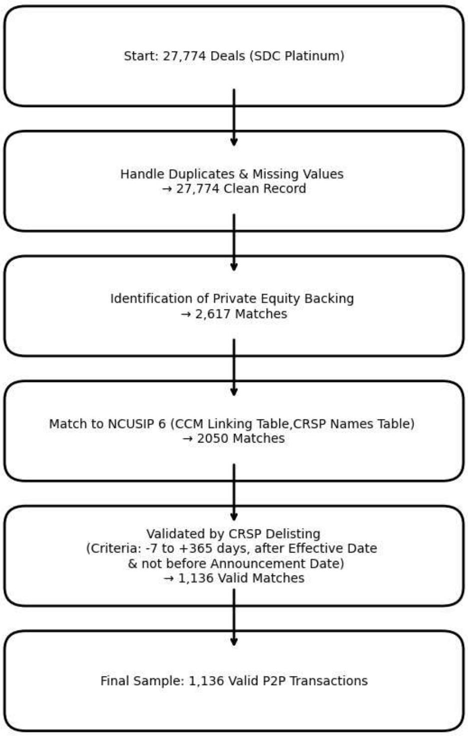
1. Exact 6-digit CUSIP correspondence between SDC target and CCM records
2. Deal announcement date falling within the CCM link validity period
3. Preservation of multiple PERMNOs for firms with complex capital structures

Delisting confirmation followed Gao et al. (2013) requiring CRSP delisting dates to occur:

- After the deal announcement (preventing retroactive misclassification)
- Within [-7, +365] days of the effective date (allowing for administrative variations)

The rigorous matching and validation process yielded 1,136 confirmed PE-backed public-to-private transactions. Figure 1 illustrates the data construction pipeline, showing the progressive refinement from initial SDC records to the validated P2P sample.

Figure 1: Target Label Construction Flowchart

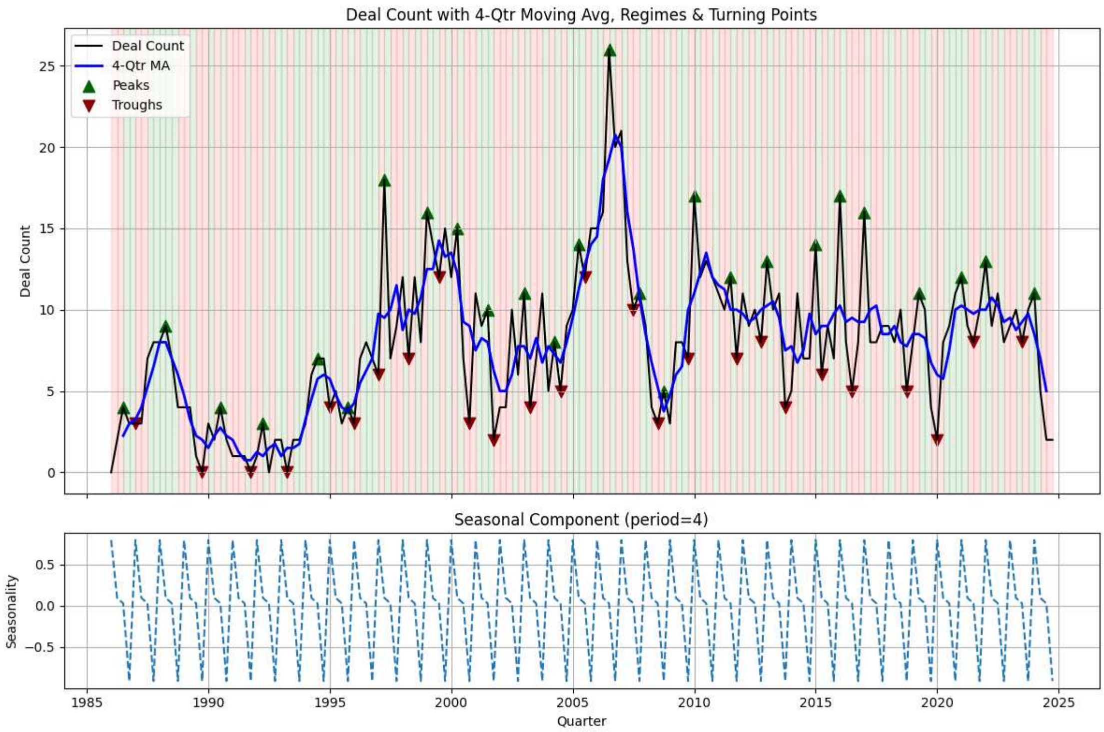


Note: The process begins with 27,774 completed M&A transactions from SDC Platinum and applies sequential filters including PE involvement identification through PermID, Names matching and keyword analysis, going-private flag validation, and CRSP delisting confirmation.

3.2 Sample Description and Time-Series Properties

The validated sample spans 156 quarters from Q1 1986 through Q4 2024, with deals aggregated by announcement date to capture market conditions at transaction initiation. The extended timeframe encompasses multiple economic cycles, including the late-1980s LBO boom, the dot-com era, the 2000s credit bubble, the global financial crisis, and the recent pandemic period. Figure 2 presents quarterly PE-backed P2P deal counts alongside a four-quarter moving average. The series exhibits pronounced cyclicality: subdued activity averaging 1-3 deals quarterly through 1994, acceleration to 10-12 deals during the late-1990s boom, a brief correction in 2001-2003, and dramatic expansion peaking above 20 deals in Q4 2007. The financial crisis precipitated a sharp contraction to near-zero activity by early 2009, followed by gradual recovery to 8-12 deals quarterly through 2020.

Figure 2: Quarterly Deal Count with Moving Average and Regime Indicators



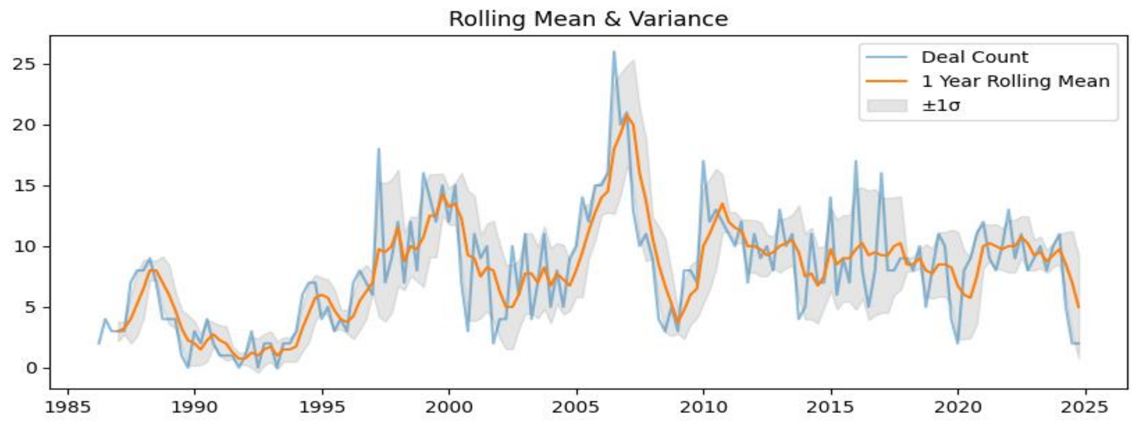
Note: Shows quarterly PE-backed P2P deals (1986-2024) with 4-quarter moving average. Peak activity in Q4 2007 (>20 deals) followed by near-zero during the 2009 crisis. Lower panel reveals Q1/Q3 seasonal strength, justifying Fourier features in the model.

The bottom panel isolates the seasonal component, revealing consistent intra-year patterns. Deal activity shows positive deviations in Q1 and Q3, with relative weakness in Q2 and Q4. This seasonality likely reflects PE fund deployment cycles and corporate planning calendars, necessitating seasonal adjustment in forecasting models. Time-series forecasting requires careful assessment of stationarity properties; hence complementary tests were applied to evaluate the series characteristics.

The Augmented Dickey-Fuller (ADF) test strongly rejects the unit root null hypothesis (test statistic = -3.80, p = 0.0029), suggesting stationarity in the classical sense. However, the Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test rejects its null of trend stationarity (test statistic = 0.647, p = 0.0184). This apparent contradiction signals potential structural breaks rather than pure non-stationarity.

Visual inspection confirms multiple regime changes coinciding with major economic events. Figure 3 displays 12-quarter rolling statistics, revealing substantial time variation in both mean and variance. The rolling mean rises steadily through 2007, drops precipitously during the financial crisis, and stabilizes at a lower post-crisis level. Rolling variance likewise shows regime-dependent behavior, with heightened volatility during crisis periods and market transitions.

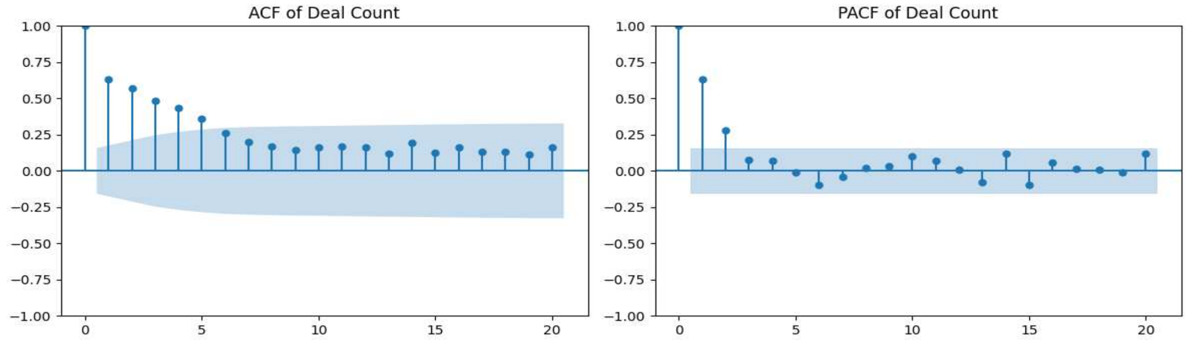
Figure 3:Quarterly Deal Count Rolling Mean & Variance



Note: 1-year rolling mean (orange line) with ±1 standard deviation bands (gray shading) around actual quarterly deal counts (blue). Shows increasing volatility during the 2007-2008 boom-bust cycle and persistent lower activity post-crisis, indicating regime changes in PE market dynamics.

These patterns are characteristic of financial time series subject to regulatory changes, credit cycles, and systemic shocks. The 2008 financial crisis represents a particularly clear structural break, fundamentally altering PE market dynamics through tighter credit conditions and enhanced regulatory scrutiny. Figure 4 presents the autocorrelation (ACF) and partial autocorrelation (PACF) functions. The ACF shows strong positive correlation at lag 1 ($\rho_1 \approx 0.65$) with gradual decay over 8-10 quarters. This persistence indicates that deal activity shocks propagate through multiple periods, consistent with the lumpy nature of PE fund deployment and multi-quarter transaction processes.

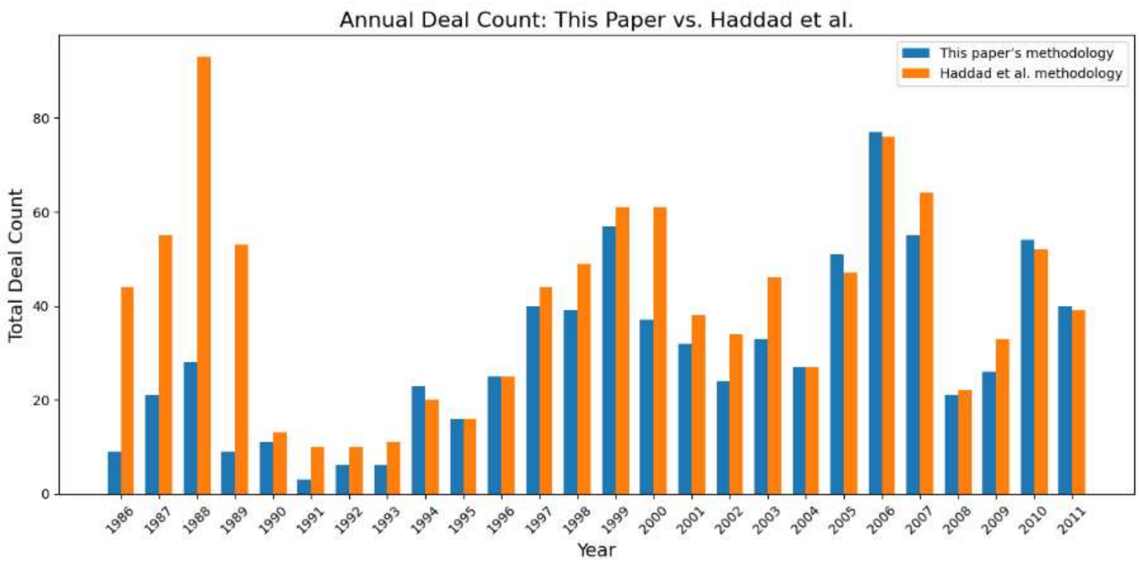
Figure 4: Autocorrelation and partial autocorrelation of the target label Deal Count



Note: ACF shows strong lag-1 correlation ($\rho_1 \approx 0.65$) with slow decay over 8-10 quarters. PACF cuts off after lag 1, suggesting AR(1) process, but spikes at lags 4 and 8 indicate quarterly/bi-annual cycles. Informed inclusion of lag1 and lag4 features.

The PACF cuts off sharply after lag 1, suggesting an AR(1) component captures most linear dependence. However, significant spikes at lags 4 and 8 indicate additional quarterly and bi-annual cycles, possibly reflecting PE fund vintage effects and macroeconomic cycles. To validate the sample construction methodology, the sample of PE-backed P2P counts was compared with existing studies.

Figure 5: Comparison of Annual PE-Backed P2P Deal Counts with Haddad et al. (2017)



Note: Validates sample construction with $\rho = 0.89$ correlation from 1994 onward. Pre-1994 divergence reflects SDC data limitations. Strong alignment during 2000s cycle confirms methodology reliability for main analysis period.

Figure 5 shows this thesis' series alongside data from Haddad, Loualiche, and Plosser (2017), with replication files accessed from Erik Loualiche's website. The correlation between the two series reaches 0.89 from 1994 onward, confirming methodological consistency during periods with reliable data coverage. Pre 1994 divergence reflects data limitations: early SDC records suffer from incomplete PE firm identification, missing PermIDs, and inconsistent naming conventions. The keyword-supplemented approach partially mitigates these issues but cannot fully reconstruct historical PE involvement.

3.3 Feature Selection and Data Retrieval

The comprehensive dataset was built using five primary sources: the Federal Reserve Economic Data (FRED) database for macroeconomic indicators, SDC Platinum for PE transaction and fundraising data, Datastream for financial market variables, Cambridge Associates for PE performance metrics (accessed via LSEG Workspace), and CRSP for equity market data. Both theoretical reasons and empirical data from past studies guide the choice of predictor variables for quarterly PE-backed P2P deal activity. Based on the credit market insights of Axelson et al. (2013) and the framework developed by Haddad et al. (2017), who showed that buyout waves react systematically to changes in aggregate risk premiums, this study combines a complete set of 121 macroeconomic and financial variables arranged into four different categories:

1. Macroeconomic (53 variables): GDP measurements; monetary indicators (federal funds rate, money supply); credit spreads (AAA/BAA bonds); term structure; labor market metrics and bankruptcy filings.
2. Stock Market (13 variables): S&P 500 measures, IPO activity (Ritter, 2015), international stock indexes, and high-yield credit indicators—capturing public market circumstances affecting P2P desirability.
3. Sentiment (24 variables): Consumer sentiment (Michigan), OECD business surveys, Baker et al. (2019) EMV trackers evaluating news-based uncertainty across policy dimensions (fiscal, monetary, regulatory).
4. Alternative Assets (31 variables): PE "dry powder" (uninvested capital), fundraising measures, exit trends, real estate indices (NCREIF) and commodities prices.

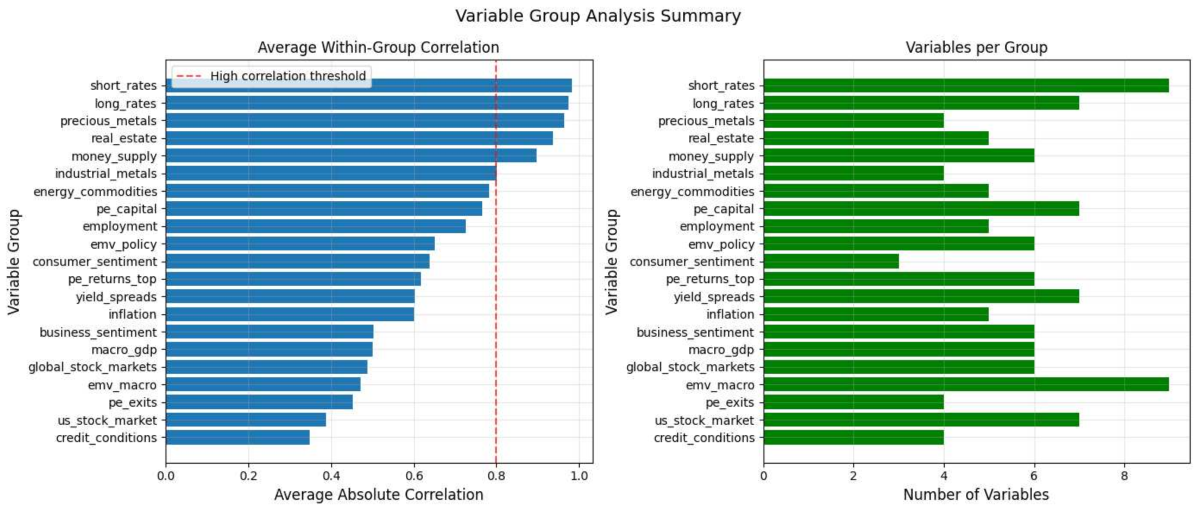
As discussed in the chapter on methodology, these four primary categories are further arranged into 21 sub-groups to support Principal Component Analysis. Higher-frequency data was adjusted to quarterly averages to reduce look-ahead bias; all variables were aggregated to match deal count data. Table 8 specifies all variables, sources, frequencies, and PCA subgroups classifications. This comprehensive data framework allows machine learning methods to find intricate linkages and catches several factors influencing PE investment decisions.

4. Methodology

4.1 Principal Component Analysis

To address multicollinearity and reduce dimensionality while preserving economic interpretability, a systematic approach combining variable selection and principal component analysis (PCA) was applied to the macro-financial variable groups. The methodology employed a two-stage decision process based on within-group correlation levels and the variance explained by the first principal component. First, the average absolute correlation within each of 21 predefined variable groups was computed. Figure 6 presents the average absolute correlations for twenty-one variable groups, revealing substantial heterogeneity in correlation patterns across macro-financial categories.

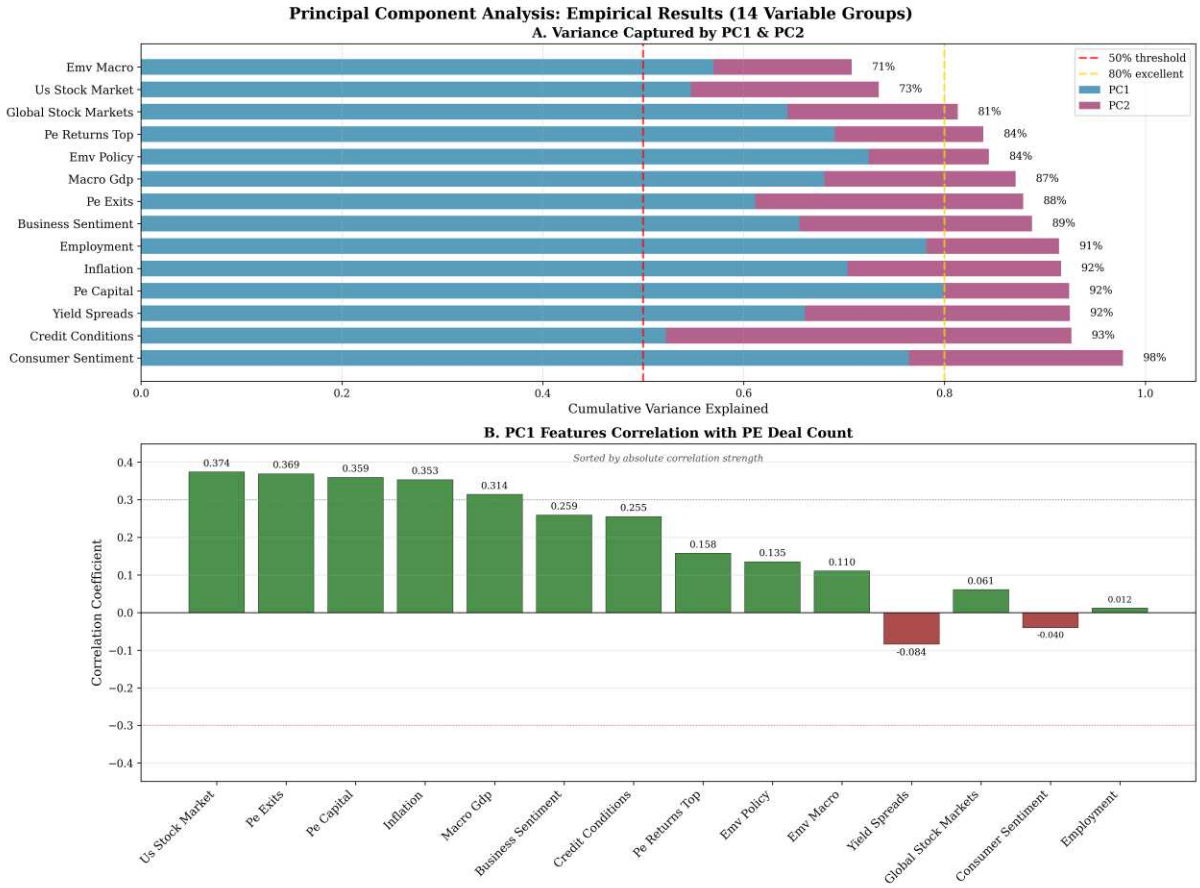
Figure 6: Pre PCA Variable Group Analysis



Note: Average within-group correlations identify multicollinearity. Groups >0.80 (e.g. short rates: 0.983, long rates: 0.974) require single representatives. Moderate correlations (0.40-0.70) justify PCA transformation.

For groups exhibiting high average absolute correlations (>0.80), single representative selection was implemented to eliminate redundancy. This approach was applied to six groups: short-term rates (0.983), long-term rates (0.974), precious metals (0.964), real estate (0.937), money supply (0.897), and industrial metals (0.801). Representatives were selected based on maximum within-group variance, ensuring the most informative variable was retained: DFF for short-term, DGS3 for long-term rates, SP Precious Metal Index for precious metals, Property Industrial for real estate, M2SL for money supply, and Tin for industrial metals. Because manual inspection showed the energy commodity series exhibited highly similar dynamics and redundant information, the SP Energy Index was selected as the sole representative. For the remaining fourteen variable groups with average correlations below 0.80, suitability for dimensionality reduction through PCA was assessed.

Figure 7: Principal Component Analysis: Empirical Results

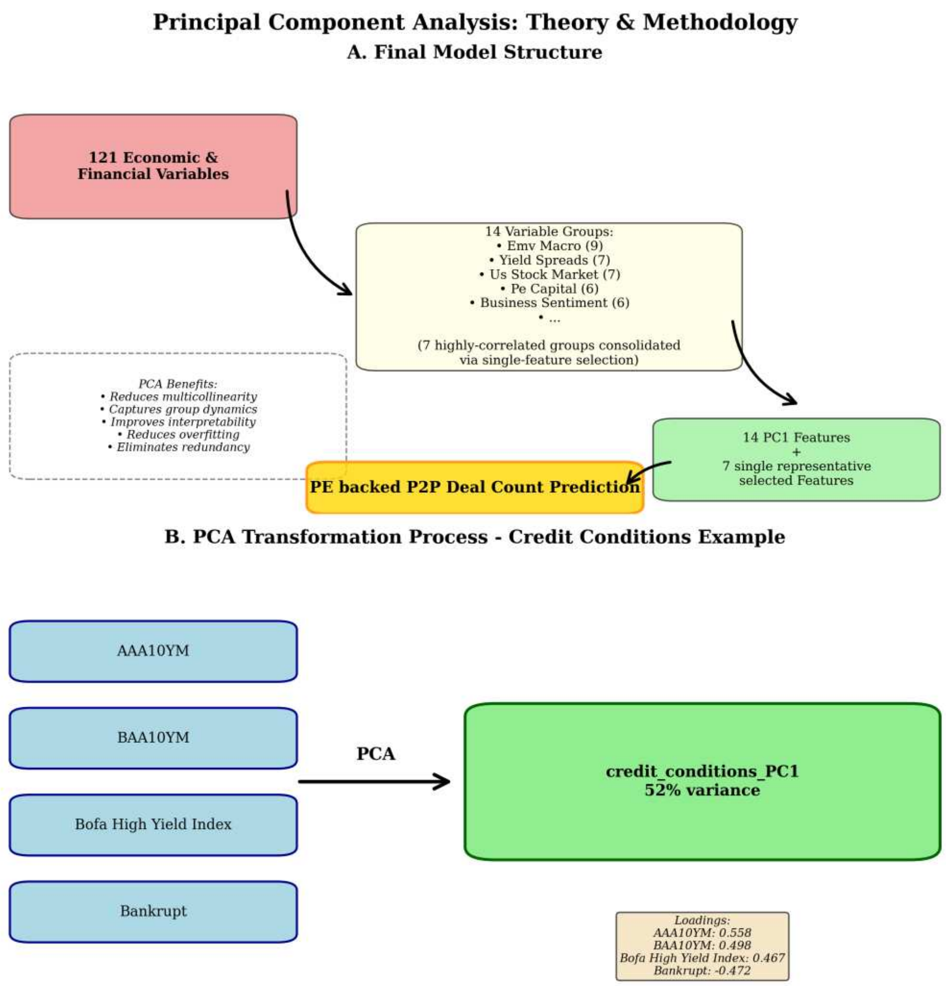


Note: PC1 explains 50-80% variance across groups. Strongest correlation with target: stock markets (0.374), PE exits (0.369), capital availability (0.359).

Figure 7 Panel A illustrates the variance explained by the first principal component (PC1) across fourteen variable groups that underwent PCA transformation. An empirical threshold requiring PC1 to explain at least 50% of total group variance was established to justify PCA transformation, a threshold ensuring that the extracted component captures meaningful common variation rather than noise (indicated by the red dashed line). The results demonstrate that all groups exceed this threshold, with PC1 capturing between 50% and 80% of total group variance. Consumer sentiment (77%), credit conditions (59%), and employment indicators (71%) exhibit particularly strong common factors, while environmental macro variables (50%) and US stock market indicators (51%) barely surpass the threshold, suggesting greater heterogeneity within these groups. Panel B shows the correlation structure between the retrieved first principal components and the goal variable (PE-backed P2P deal count) by absolute correlation strength. US stock market conditions (ρ = 0.374), private equity exits (ρ =

0.369), PE capital availability (ρ = 0.359), and inflation (ρ = 0.353) had the largest positive correlations with transaction activity. Theories predict that favorable equity markets, exit environments, and money availability will encourage P2P transactions. A weak negative connection (ρ = -0.084) between yield spreads and credit spreads suggests tighter funding conditions, perhaps limiting leveraged buyout activity. The generally modest correlation magnitudes underscore the complex, multifaceted nature of P2P deal determinants, justifying the comprehensive approach to feature engineering. Figure 8 Panel A summarizes the final variable selection outcomes across all 21 groups. Panel B illustrates the PCA transformation process using the credit conditions group as an example, where four credit spread variables are consolidated into a single principal component that captures 52% of the total variance, with the component loadings revealing how each original variable contributes to the extracted factor.

Figure 8: Principal Component Analysis: Theory & Methodology



Note: Panel A: 7 groups → single variables, 14 groups → PCA, deal variables preserved. Panel B: Credit conditions example shows 4 variables → 1 PC capturing 52% variance. Reduces 121 variables to tractable set.

The resulting variable set (See Table 7) balances statistical efficiency with economic interpretability, providing a robust foundation for predicting PE-backed P2P transaction volumes.

4.2 Preprocessing and Time Series Feature Engineering

The forecasting methodology employs advanced time series feature engineering to model the complex temporal dynamics of PE deal activity. The feature engineering approach draws from established practices in machine learning competitions, specifically following the framework outlined in the Kaggle Learn Time Series course (Kaggle, 2025). Fourier transform features were created using trigonometric functions to represent seasonal patterns and longer-term business cycles. Generation of seasonal harmonics based on:

$$sin\_annual_h = sin\left(\frac{2\pi ht}{4}\right), \quad cos\_annual_h = cos\left(\frac{2\pi ht}{4}\right)$$

, where h represents the harmonic number, t is the time index, and the period of 4 captures quarterly seasonality. Business cycle components were similarly constructed with a 40-quarter period to capture decade-long economic cycles:

$$sin_{cycle} = sin\left(\frac{2\pi t}{40}\right), \quad cos_{cycle} = cos\left(\frac{2\pi t}{40}\right)$$

The 40-quarter (10-year) cycle period reflects empirical evidence of business cycle lengths in financial markets and private equity activity specifically. These Fourier features provide a flexible, non-parametric approach to modeling cyclical patterns without imposing rigid seasonal constraints, allowing the models to learn complex temporal relationships naturally. Simple lag features capture direct autoregressive dependencies that reflect persistence and seasonal patterns in deal activity. The implementation includes both short-term and medium-term lags:

$$X_{lag1,t} = y_{t-1}, \; X_{lag4,t} = y_{t-4}$$

, where Y_{t-1} represents the previous quarter's deal count, capturing short-term persistence and momentum effects, while Y_{t-4} captures year-over-year autoregressive patterns that reflect annual business cycles. These features embody the economic intuition that deal activity exhibits both quarter-to-quarter continuity and seasonal patterns driven by business calendar effects, regulatory cycles, and fund timing considerations. To capture medium-term trends while maintaining temporal integrity, rolling statistics were computed and appropriately lagged to prevent data leakage.

This approach provides smoothed trend indicators that reduce noise while preserving directional information:

$$X_{roll\_mean_{w,t}} = \frac{1}{w}\sum_{i=1}^{w} y_{t-1-i}\,, \qquad X_{roll\_std_{w,t}} = \sqrt{\frac{1}{w-1}\sum_{i=1}^{w}(y_{t-1-i} - \bar{y}_{w,t-1})^2}$$

, where w represents window sizes of 4, 8, and 12 quarters respectively. The one-period lag ensures that only historical information informs current predictions, maintaining the integrity of the forecasting framework. Multiple growth and momentum indicators were constructed to capture the rate, direction, and acceleration of changes in deal activity. Quarter-over-quarter and year-over-year growth rates were calculated as:

$$QoQ\_pct_t = \frac{y_t - y_{t-1}}{|y_{t-1}| + \varepsilon}, \quad YoY\_pct_t = \frac{y_t - y_{t-4}}{|y_{t-4}| + \varepsilon}, \quad vs\_trend_t = \frac{y_t - \bar{y}_{4,t-1}}{|\bar{y}_{4,t-1}| + \varepsilon}$$

, where $\varepsilon = 10^{-5}$ prevents division by zero when deal counts are near zero. To handle extreme values, percentage changes are clipped to ±1000% to prevent outliers from dominating model training. Binary rising indicators capture directional momentum:

$$rising_t = 1(y_t > y_{t-1})$$

This indicator equals 1 when deal activity increases from the previous quarter and 0 otherwise, providing a simple but effective measure of directional change that can capture turning points in deal cycles. Acceleration measures capture second-order changes in growth momentum:

$$accel_t = (y_t - y_{t-1}) - (y_{t-1} - y_{t-2})$$

This acceleration metric measures the change in the rate of change, capturing whether deal activity is accelerating or decelerating beyond simple growth measures. Positive acceleration indicates increasing momentum, while negative acceleration suggests slowing growth or increasing decline. Additional momentum features included momentum streak counters that measure consecutive periods of positive or negative growth, providing information about the persistence of directional trends. All momentum features were lagged by one period to maintain temporal integrity and prevent look-ahead bias. Following feature engineering, a standardized preprocessing pipeline was applied consistently across all models using scikit-learn's preprocessing modules. Quarter seasonality was extracted using pandas datetime functionality and one-hot encoded using scikit-learn's OneHotEncoder with predefined categories (Q1, Q2, Q3, Q4). This two-stage pipeline transforms temporal information into four binary seasonal indicators, allowing models to learn each quarter's unique seasonal patterns independently without imposing ordinal relationships between quarters.

All numerical variables, including macroeconomic indicators, PCA components, and engineered time series features, were standardized using scikit-learn's StandardScaler:

$$X_{standardized=}\frac{X-\mu_X}{\sigma}$$

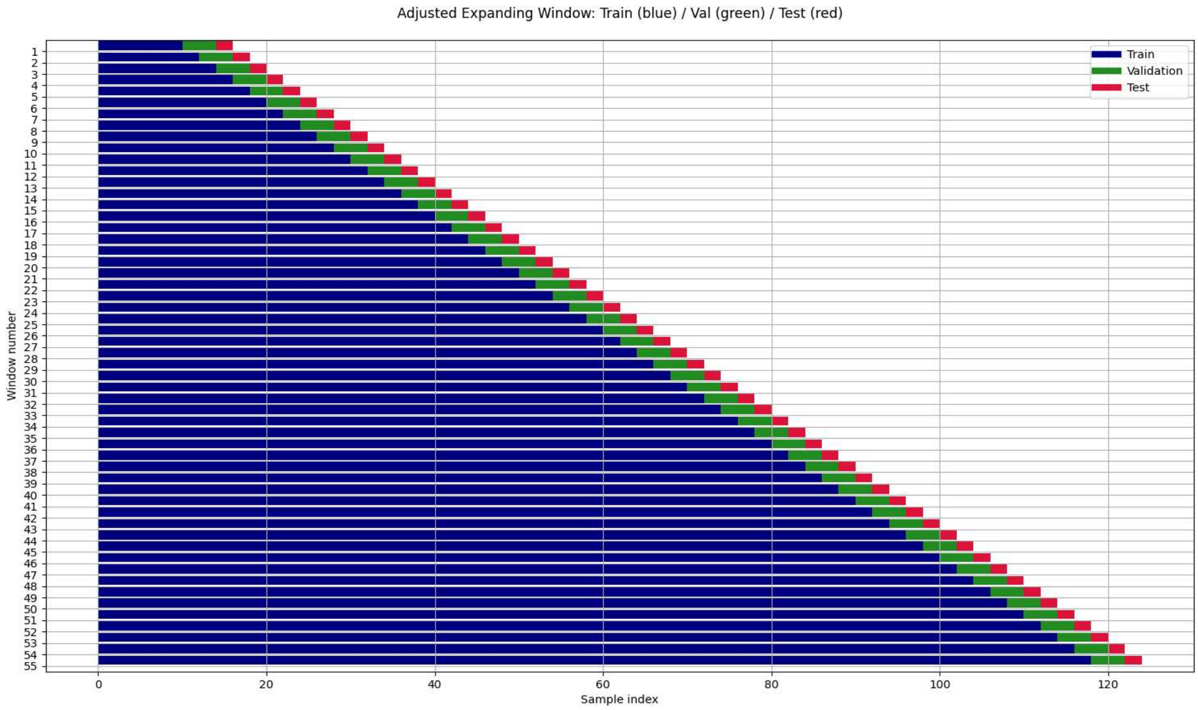
, where μ and σ represent the sample mean and standard deviation computed exclusively on training data. This standardization approach ensures that all features contribute equally to regularized models like Lasso regression while preventing scale differences from biasing tree-based models. The StandardScaler was fitted only on training data and then applied to validation and test sets, maintaining the separation necessary for unbiased performance evaluation.

The complete preprocessing pipeline employed scikit-learn's ColumnTransformer, which handles different feature types appropriately. The pipeline applies quarter-specific transformations to temporal features while standardizing all numerical variables, ensuring consistent preprocessing.

4.3 Time-Series Split and Cross Validation

Using a nested time-series CV technique that acknowledges temporal ordering and produces robust out-of-sample performance estimates, model evaluation. This methodology solves the special difficulties of financial time series analysis in which traditional cross-validation methods might induce overstatement of model performance and look-ahead bias. The nested structure serves certain methodological purposes on two different levels. Using an expanding window technique and TimeSeriesSplit from scikit-learn, the outer loop does model assessment. While test sets consist of two quarters for multi-step-ahead prediction, each fold uses all available historical data for training, guaranteeing models have access to the maximum amount of information. This design evaluates the model's ability to forecast both immediate (t+1) and near-term (t+2) deal activity, providing a more comprehensive assessment of predictive performance. This design maintains strict temporal ordering to prevent data leakage while providing realistic forecasting scenarios that mirror practical implementation conditions. The inner loop handles hyperparameter optimization through a fixed validation window of 8 quarters (2 years) that facilitates systematic grid search over predefined hyperparameter spaces. Best parameters are selected based on validation Mean Absolute Error (MAE), chosen for its direct interpretability in the context of deal count forecasting. This nested structure ensures that hyperparameter selection does not bias final performance estimates, as the test data remains completely isolated from the hyperparameter optimization process. Figure 9 illustrates the nested CV splits.

Figure 9: Adjusted Expanding Window for Nested CV splits



Note: Expanding windows for evaluation with 2-quarter test periods fixed 8-quarter validation for hyperparameter selection. Prevents look-ahead bias while maximizing data usage.

To evaluate model stability over time and assess temporal variations in predictive accuracy, a post-hoc rolling window R^2 analysis was implemented on test predictions generated from the nested cross-validation procedure. A naive model serves as the performance baseline, employing a "previous value" forecasting strategy that represents the simplest possible forecasting assumption:

$$\hat{y}_{naive,t} = y_{t-1}$$

This benchmark assumes that the best prediction for the current period is the observed value from the immediately preceding period. For the first prediction in any sequence, the naive model uses the last known training value. This analysis examines performance across different periods while maintaining the integrity of the temporal validation framework, providing insights into model robustness across varying market conditions. During nested CV, all out-of-sample test predictions were systematically collected and stored with their corresponding quarter index, actual target values, model predictions, naive baseline predictions, and cross-validation fold identifiers. This comprehensive collection resulted in a temporally ordered sequence of test predictions spanning the entire evaluation period, with each prediction representing genuine out-of-sample forecasts that models had never encountered during training. The rolling R^2 analysis employed 8-quarter windows with 4-quarter roll-forward intervals. This configuration ensures sufficient data points for reliable R^2 calculation while providing annual updates to capture evolving model performance patterns. Window R^2 was calculated as:

$$R_w^2 = 1 - \frac{\sum_{i=t}^{t+7}(y_i - \hat{y}_i)^2}{\sum_{i=t}^{t+7}(y_i - \bar{y}_w)^2}$$

, where y_i represents actual target values, $\hat{y}_i$ represents model predictions, $\bar{y}_w$ represents the mean of actual values within the window, and summation occurs over the 8-quarter window period.

4.4 Model Specifications and Hyperparameter Optimization

Three distinct machine learning paradigms were selected to capture different aspects of the forecasting problem, each representing fundamentally different approaches to the bias-variance tradeoff that governs predictive modeling. The selection spans linear regularization, ensemble bootstrapping, and gradient boosting methodologies, providing comprehensive evaluation across the spectrum of machine learning approaches suitable for financial time series prediction. Table 1 presents the complete model specifications and hyperparameters explored during the optimization process.

Table 1: Machine Learning Model Specifications and Hyperparameter Configuration

Method	Algorithm	Hyperparameters
Regularization	Lasso Regression	• alpha • tol • max_iter • selection
Boosting	XGBoost	• n_estimators • learning_rate • max_depth • min_child_weight • gamma • subsample • colsample_bytree • reg_alpha • reg_lambda
Bootstrapping	Random Forest	• n_estimators • max_depth • min_samples_split • min_samples_leaf • max_features

Note: For detailed results of the hyperparameter optimization process see Figures 13,14 and 15

Lasso Regression employs L1 regularization through scikit-learn's Lasso class for automatic feature selection. The L1 penalty provides predictive power and model parsimony that makes interpretation easier by automatically identifying critical predictors and setting irrelevant coefficients to zero. In addition to tolerance criteria, maximum iterations, and coordinate descent selection techniques, hyperparameter optimization using grid search investigated regularization strengths (alpha) ranging from 1e-5 to 1.0 on a logarithmic scale.

Random Forest represents the ensemble bootstrapping approach, implemented through scikit-learn's RandomForestRegressor, combining multiple decision trees trained on bootstrap samples of both observations and features. This method addresses the bias-variance tradeoff through bagging (bootstrap aggregating), which reduces variance by averaging predictions across multiple models while maintaining relatively low bias through flexible tree structures. The ensemble nature provides robust predictions while built-in feature randomization reduces overfitting tendencies common in high-dimensional financial data. The hyperparameter space encompassed the number of trees (50-300), maximum depth (none-30), minimum samples for node splitting (2-10) and leaf nodes (1-4), and feature sampling strategies (sqrt, log2, and none).

XGBoost implements gradient boosting with advanced regularization and optimization techniques through the XGBoost package. This sequential ensemble method represents the boosting paradigm, which iteratively reduces bias by correcting prediction errors from previous models while controlling variance through sophisticated regularization mechanisms. For XGBoost hyperparameter optimization, the Optuna package (Akiba et al., 2019) was employed to maximize R^2 through efficient Bayesian optimization rather than exhaustive grid search, reflecting the large hyperparameter space and computational complexity of gradient boosting methods. The search space included the number of boosting rounds (100-800 in steps of 50), learning rate (0.01-0.3 on a logarithmic scale), maximum tree depth (3-10), minimum child weight (1-7), and gamma regularization (0-2). Additional regularization parameters included L1 (alpha: 0-3) and L2 (lambda: 0-3) penalties, along with subsampling ratios for both observations (0.6-1.0) and features (0.6-1.0). The optimization also incorporated feature selection, with the number of selected features ranging from 5 to 15. All hyperparameter optimization procedures maintained strict temporal ordering within the nested cross-validation framework, ensuring that future information never influenced parameter selection. Detailed results of the optimization process are presented in Figures 13-15, illustrating how optimal parameters evolved across different market regimes.

4.5 Evaluation Metrics and Performance Assessment

Performance evaluation employed multiple complementary metrics to assess predictive capabilities while maintaining practical interpretability for investment professionals. R^2 quantifies the proportion of variance in deal count explained by model predictions, providing a standardized measure of overall predictive power that facilitates comparison across different modeling approaches. Mean Absolute Error (MAE) serves as the principal evaluation metric due to its direct interpretability in the business context of deal count forecasting. Unlike percentage-based or squared error metrics, MAE provides exact quantification of forecasting accuracy in the original unit of measurement: the number of deals. This allows practitioners to understand precisely how many deals model predictions deviate from actual outcomes. For instance, an MAE of 2.8 directly indicates that model forecasts are, on average, off by approximately 3 deals per quarter, enabling investment professionals to assess the practical utility of predictions. To assess whether model improvements over the naive baseline are statistically significant rather than due to chance, a rigorous testing framework was implemented that evaluates the consistency of performance gains across temporal windows. This approach tests whether the model systematically outperforms the naive baseline across different time periods, rather than merely achieving better aggregate performance. For each rolling window i = 1, 2, (...), n, the adjusted differences between model and baseline performance:

R^2 Difference:

$$H_0 : \mu_{\Delta R^2} = 0 \;\; vs. \;\; H_1 : \mu_{\Delta R^2} \neq 0$$

MAE Difference:

$$H_0 : \mu_{\Delta \mathrm{MAE}} = 0 \;\; vs. \;\; H_1 : \mu_{\Delta MAE} \neq 0$$

For MAE, the difference was calculated as naive minus model since lower MAE indicates better performance, ensuring positive differences indicate model improvement for both metrics. Afterwards one-sample t-tests were performed on these difference series to determine whether the mean performance differences are statistically significant across rolling windows:

For R^2 Differences:

$$t_{R^2} = \frac{\overline{\Delta R^2}}{SE(\Delta R^2)} = \frac{\overline{\Delta R^2}}{s_{\Delta R^2}/\sqrt{n}}$$

For MAE Differences:

$$t_{MAE} = \frac{\overline{\Delta MAE}}{SE(\Delta MAE)} = \frac{\overline{\Delta MAE}}{s_{\Delta MAE}/\sqrt{n}}$$

Where:

- $\overline{\Delta R^2} = \frac{1}{n}\sum_{i=1}^{n} \Delta R_i^2$ is the mean R^2 difference across windows
- $\overline{\Delta MAE} = \frac{1}{n}\sum_{i=1}^{n} \Delta MAE_i$ is the mean MAE difference across windows
- $s_{\Delta R^2}$ and
  $s_{\Delta MAE}$ are the sample standard deviations of the respective difference series
- $SE(\cdot)$ denotes the standard error of the mean

The null hypothesis for both tests is that the mean difference equals zero (H_0: μ_Δ = 0), indicating no systematic improvement over the baseline. A significant positive t-statistic provides evidence that the model consistently outperforms the naive baseline across different market conditions and time periods, rather than achieving sporadic improvements in isolated periods. This statistical framework ensures that reported performance improvements represent genuine predictive value rather than random variation, providing practitioners with confidence that the additional complexity introduced by machine learning methods translates into meaningful and consistent improvements in forecasting accuracy.

5. Results

5.1 Model Development and Training Performance

The nested CV framework provided robust out-of-sample performance estimates across all three modeling approaches over the 39-year evaluation period spanning Q1 1986 to Q4 2024. The evaluation encompassed 50 cross-validation folds, with each fold representing a two-step-ahead quarterly forecast. All models demonstrated meaningful predictive capability for PE-backed P2P deal counts, though with varying degrees of accuracy and consistency. The Lasso regression model achieved a cross-validated mean MAE of 3.23 deals per quarter with an overall test R^2 of 0.234. The L1 regularization enforced sparsity in the feature space, while maintaining competitive predictive performance. This parsimonious approach enhanced model interpretability without substantially sacrificing accuracy. The temporal stability of Lasso's performance is illustrated in Figure 16 in the Appendix, which shows the rolling R^2 analysis demonstrating periods of outperformance and underperformance relative to the naive baseline. Random Forest demonstrated robust ensemble performance with a mean MAE of 3.28 deals and the highest R^2 of 0.254 among all models. The bootstrap aggregation approach effectively captured complex interactions between macroeconomic variables, though at the cost of increased computational requirements and reduced interpretability. The model's performance remained stable across different market regimes, suggesting effective variance reduction through ensemble averaging (Compare Appendix Figure 18). XGBoost, despite its sophisticated gradient boosting architecture, achieved a mean MAE of 3.28 deals but recorded the lowest R^2 of 0.145. This unexpected underperformance relative to simpler methods suggests potential overfitting to training data patterns despite extensive hyperparameter tuning. The model's complexity may have captured noise rather than genuine signal in the relatively small quarterly dataset (Compare Appendix Figure 20). All sophisticated models outperformed the naive baseline, which achieved a mean MAE of 3.38 deals and R^2 of 0.179. This represents a 3-9.8% improvement in MAE, demonstrating that incorporating macroeconomic information provides modest but meaningful gains over the naive approach. The feature importance showed which economic indicators consistently drove forecasts across modeling methodologies. The Lasso model's most popular CV characteristics were the lagged deal count, macroeconomic GDP principal component, and 4-quarter rolling mean of deal activity. This consistent selection pattern shows that PE deal activity is strongly autoregressive and that macroeconomic conditions are predictive. Cyclical features dominated the Random Forest model, with the annual cosine transformation (cos_annual_2) having the highest mean permutation importance of 0.064 (SD = 0.032), followed by the employment principal component (0.045, SD = 0.021) and the 10-year business cycle cosine feature (0.036, SD = 0.018 It appears that Random Forest accurately predicted transaction activity using seasonal patterns and longer-term economic cycles. XGBoost had a distinct feature significance profile. The algorithm ranked fundamental economic variables first, with inflation, business sentiment, and credit conditions as top predictors. Notably, XGBoost has a severely skewed feature importance distribution, with 80% of features near-zero. This trend shows that Optuna's Bayesian optimization method yielded an implicitly aggressive feature selection mechanism that used only a small set of predictors. In all models, temporal features were crucial. All three techniques have deal count momentum indicators in the top 20 features, including consecutive quarters of positive/negative growth (deal_count_streak) and year-over-year percentage changes. This broad awareness of momentum effects suggests PE deal activity has lasting tendencies beyond autoregressive patterns. To verify the significance of reported performance gains, paired t-tests were performed on R^2 and MAE differences over 24 rolling 2-year intervals. The rolling window examination assessed model stability and consistency across market situations. Figures 16, 18, and 20 in the Appendix show the temporal evolution of model performance compared to the baseline for Lasso, XGBoost, and Random Forest.

Table 2: Statistical Significance Tests Across Rolling Windows (Training)

Model/Metric	Mean Diff.	SD	SE	t-statistic	p-value	95% CI	Sig.
Lasso							
R^2	0.2	1.071	0.219	0.913	0.371	[-0.253, 0.652]	ns
MAE	-0.057	1.396	0.285	-0.199	0.844	[-0.646, 0.533]	ns
XGBoost							
R^2	-0.05	1.608	0.328	-0.153	0.88	[-0.729, 0.629]	ns
MAE	-0.044	1.272	0.26	-0.168	0.868	[-0.581, 0.494]	ns
Random Forest							
R^2	0.345	1.155	0.236	1.463	0.157	[-0.143, 0.833]	ns
MAE	-0.136	1.333	0.272	-0.5	0.622	[-0.699, 0.427]	ns

Note: CI = confidence interval; ns = not significant. **p* < .05. ***p* < .01. ****p* < .001.

Statistical tests revealed that none of the models achieved statistically significant improvements over the naive baseline at the $\alpha = 0.05$ level. For Lasso regression, the mean R^2 improvement of 0.200 (SE = 0.219) yielded a t-statistic of 0.913 ($p = 0.371$), while the MAE reduction of -0.057 deals produced a t-statistic of -0.199 ($p = 0.844$). Random Forest showed the largest mean R^2 improvement of 0.345 (SE = 0.236) but remained statistically insignificant ($t = 1.463$, $p = 0.157$). XGBoost underperformed the baseline on average, with a negative mean R^2 difference of -0.050 ($t = -0.153$, $p = 0.880$).

These results indicate substantial temporal variability in model performance. While the sophisticated models demonstrated superior average performance, their advantages were inconsistent across different time periods. The high standard errors relative to mean differences (coefficient of variation exceeding 5 for most metrics) suggest that model performance was highly dependent on the specific market regime. The rolling R^2 visualizations clearly illustrate these regime-dependent performance variations, with models alternating between periods of substantial outperformance and underperformance relative to the naive baseline.

The rolling window analysis revealed distinct periods of model outperformance and underperformance. All models showed enhanced predictive capability during stable market conditions (2012-2019) but struggled during transition periods, particularly around the 2008 financial crisis. This pattern suggests that while macroeconomic variables contain predictive information, their relationships with PE deal activity may be non-stationary or subject to structural breaks during market disruptions as the ADF and KPSS tests during sample inspection indicated. The lack of statistical significance does not negate the practical value of these models. The consistent directional improvement in MAE across all sophisticated approaches, combined with meaningful R^2 gains for Lasso and Random Forest, suggests that these models capture genuine predictive patterns despite high variance.

5.2 Final Model Evaluation

The final evaluation of the holdout test set, which spans the period from Q1 2017 to Q4 2024, provides an unbiased OOS test. Within this eight-year period, financial markets exhibited high volatility during COVID-19 disruptions in 2020. Visual inspection of the holdout predictions (Figures 22-28) reveals that all models captured the general downward trend in deal activity from the 2017 peak through the 2024 trough, though with varying degrees of accuracy. The actual deal count exhibited substantial volatility, ranging from a high of 16 deals in Q1 2017 to lows of 2 deals in both Q4 2020 and Q4 2024. This extreme variability presented significant forecasting challenges, as evidenced by all models' tendency to underpredict during activity spikes and overpredict during sharp contractions.

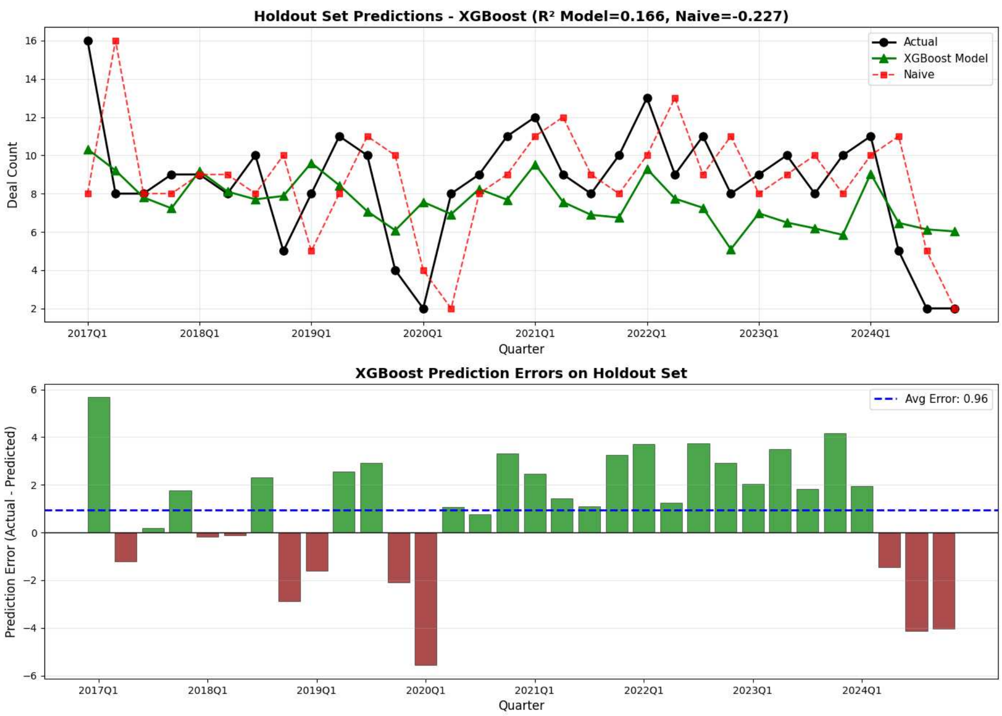
Table 3: Holdout Set Performance Comparison

Model	***MAE***	***R²***
Naive	2.62	-0.227
Lasso	2.30	0.026
XGBoost	2.41	0.166
Random Forest	2.38	-0.151

Note: XGBoost gives the best hold-out performance (R^2), followed by Lasso.

XGBoost achieved the highest overall R^2 of 0.166 on the holdout set, compared to -0.227 for the naive baseline, representing a 39.3 percentage point improvement. However, this aggregate performance metric masks considerable temporal variation. The model's MAE of 2.41 deals represented only a modest 0.21-deal improvement over the naive baseline (2.63 deals), and statistical testing revealed this difference was not significant (t = -0.495, p = 0.624).

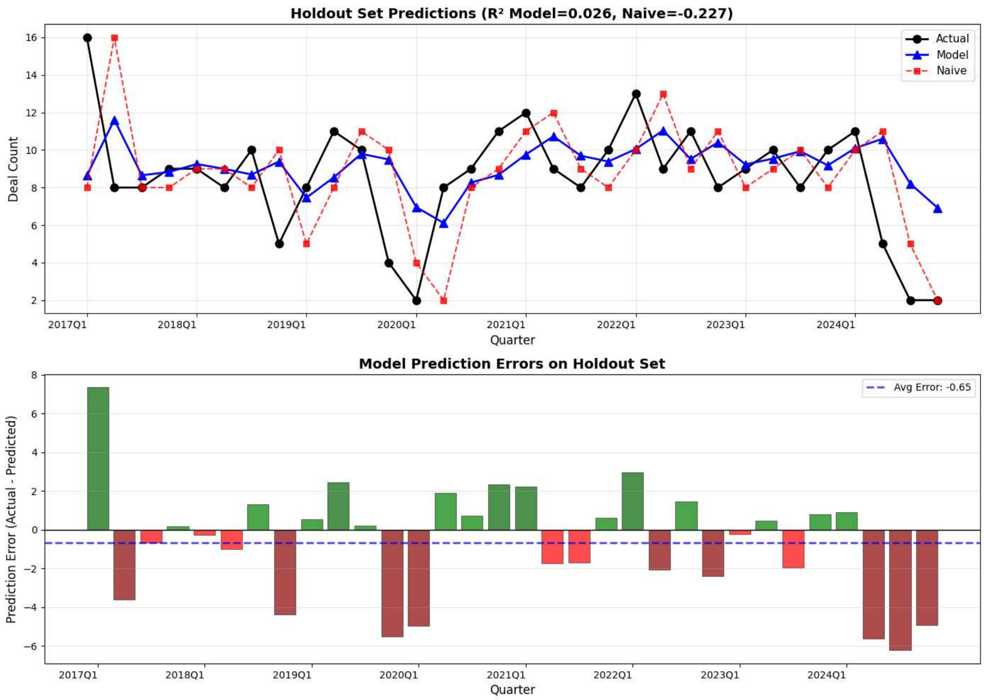
Figure 10: XGBoost OOS Predictions



Note: R^2 = 0.166, MAE = 2.41. Captures general trends but underpredicts Q1 2017 spike, overpredicts during pandemic. Tree-based limitations evident in extreme value handling.

Lasso regression demonstrated the most consistent performance characteristics, achieving an overall R^2 of 0.026 and MAE of 2.30 deals, the lowest absolute error among all models. While the MAE improvement of 0.33 deals over the naive baseline was not statistically significant ($t = -1.034$, $p = 0.309$), the model showed statistically significant R^2 improvements when evaluated over rolling 8-quarter windows ($t = 4.346$, $p < 0.001$). This suggests that Lasso's linear framework provided stable, if modest, predictive value across varying market conditions.

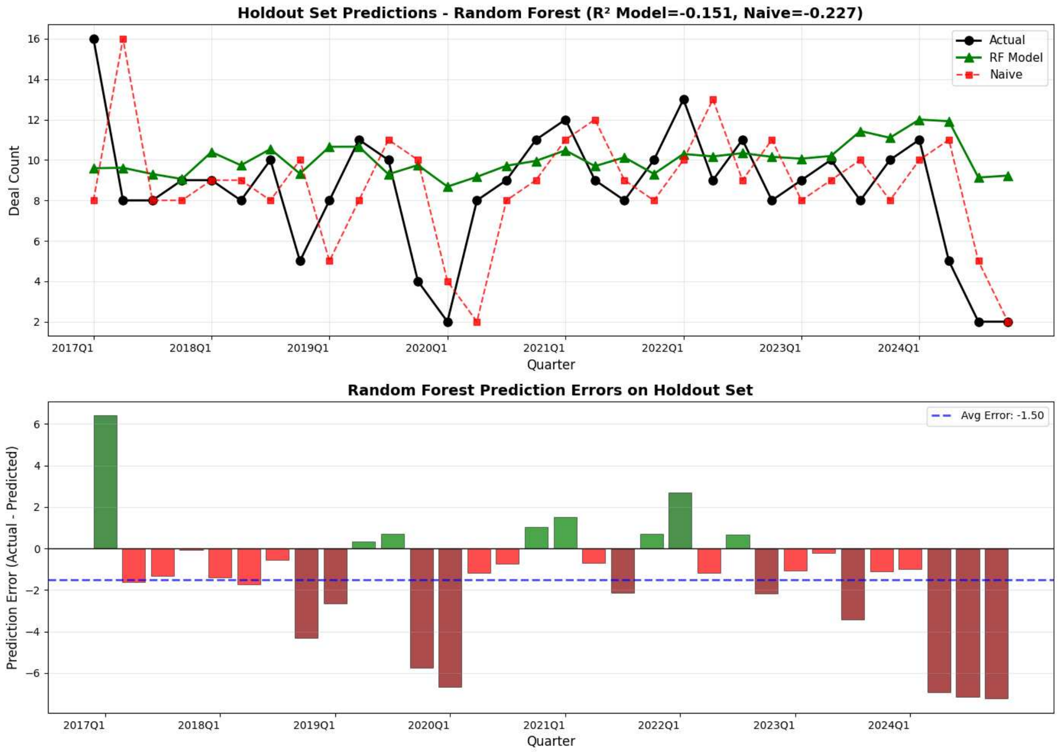
Figure 11: Lasso OOS Predictions



Note: R^2 = 0.026, MAE = 2.30 (lowest). Smooth linear predictions track medium-term trends. Most stable and practical despite modest R^2.

Random Forest produced mixed results, with an overall negative R^2 of -0.151 indicating predictions that increased variance relative to simply using the mean. Despite this aggregate underperformance, the model achieved an MAE of 2.38 deals and demonstrated statistically significant R^2 improvements over rolling windows ($t = 3.354$, $p = 0.003$). This apparent contradiction suggests that Random Forest performed well during stable periods but was particularly susceptible to prediction errors during the extreme market events that dominated the holdout period.

Figure 12: Random Forest OOS Predictions



Note: R^2 = -0.151, MAE = 2.38. Negative R^2 indicates harmful predictions. Persistent overprediction 2020-2024 shows failure to adapt to structural changes.

The error pattern analysis reveals systematic challenges faced by all models. The Q1 2017 spike to 16 deals was substantially underpredicted by all approaches, with errors ranging from 4-6 deals. Similarly, the sharp contractions in 2020 and 2024 were generally overpredicted, as models failed to fully capture the severity of these downturns. These systematic errors during extreme events highlight the inherent difficulty in forecasting rare, high-impact market disruptions using historical patterns. Statistical significance testing using paired t-tests on prediction errors provided nuanced insights into model performance.

Table 4: Statistical Significance Tests Across Rolling Windows (OOS)

Model/Metric	Mean Diff.	SD	SE	t-statistic	p-value	95% CI	Sig.
Lasso							
R^2	0.842	0.949	0.19	4.346	<0.001	[0.450, 1.233]	***
MAE	-0.326	1.753	0.31	-1.034	0.309	[-0.958, 0.306]	ns
XGBoost							
R^2	0.267	1.389	0.278	0.941	0.356	[-0.300, 0.834]	ns
MAE	-0.214	2.4	0.424	-0.495	0.624	[-1.079, 0.652]	ns
Random Forest							
R^2	0.768	1.123	0.225	3.354	0.003	[0.311, 1.226]	**
MAE	-0.243	2.472	0.437	-0.548	0.588	[-1.134, 0.648]	ns

*Note: CI = confidence interval; ns = not significant. *p < .05. **p < .01. ***p < .001.*

While no model achieved statistically significant MAE improvements at the $\alpha = 0.05$ level, both Lasso and Random Forest demonstrated significant R^2 improvements when evaluated over rolling windows. This divergence between point-error metrics (MAE) and variance-explained metrics (R^2) suggests that the models provided value in capturing directional movements and medium-term trends, even when absolute prediction accuracy remained limited.

The holdout period's characteristics include unprecedented monetary policy shifts, the COVID-19 pandemic, and subsequent market normalization and represented a particularly challenging test environment. The fact that all models maintained directional accuracy and avoided systematic bias despite these extreme conditions provides evidence of robustness. The consistency between cross-validation performance rankings and holdout results further validates the stability of the modeling approach across different market regimes.

6. Discussion and Limitations

6.1 Interpretation of Results and Comparison with Prior Research

Random Forest achieved the highest cross-validation R^2 (0.254) but performed poorly on holdout data (-0.151), suggesting overfitting. Conversely, XGBoost showed the lowest CV R^2 (0.145) but best holdout performance (0.166), indicating its regularization prevented overfitting. Lasso regression demonstrated the most consistent performance with lowest MAE (2.30 deals) and modest holdout R^2 (0.026). The results suggest linear models with regularization may offer better stability than complex ensemble methods for PE deal flow prediction, balancing accuracy with reliability for practitioners.

The stark contrast between CV and holdout performance, particularly for Random Forest, highlights a critical challenge in financial time series prediction: models that excel at capturing historical patterns may fail catastrophically when market regimes shift. The 2017-2024 holdout period encompassed multiple structural changes from late-cycle exuberance through pandemic disruption to monetary tightening that likely violated the stationarity assumptions implicit in tree-based ensemble methods.

The models' reliance on specific predictor categories provides empirical support for established theoretical frameworks in private equity research. Credit market conditions emerged as consistent predictors across all modeling approaches, aligning with Axelson et al.'s (2013) findings that economy-wide credit conditions dominate firm-specific factors in determining buyout leverage and pricing. The importance of credit spreads and yield curve features in these models confirms that PE activity responds systematically to debt market accessibility, as periods of compressed spreads coincide with increased deal flow.

Equity market indicators demonstrated even stronger predictive power, particularly measures capturing market valuation and volatility. This finding directly supports Haddad et al.'s (2017) argument that declines in the aggregate equity risk premium serve as the primary driver of buyout waves. The principal component analysis revealed that equity market variables, namely S&P 500 earnings, dividends, and valuation metrics consistently ranked among the most important features. During quarters with rising equity valuations and subdued volatility, models correctly anticipated increased P2P activity, reflecting PE sponsors' tendency to capitalize on optimistic market sentiment.

The relative importance of equity versus credit factors evolved across the sample period. While credit conditions dominated predictions during the 2008-2009 crisis and its immediate aftermath, equity market signals showed superior predictive power during the subsequent recovery and expansion phases. This temporal variation reconciles apparently conflicting findings in the literature: both credit availability and risk appetite matter, but their relative importance depends on the prevailing market regime.

The results both confirm and extend findings from the broader M&A forecasting literature. The modest improvements over naive baselines averaging 9.8% reduction in MAE and align with Bonaime et al.'s (2018) observation that even sophisticated models struggle to substantially outperform simple benchmarks in aggregate deal prediction. However, the specific importance of PE-related variables such as dry powder levels and exit market conditions distinguishes P2P forecasting from general M&A activity, supporting the need for specialized modeling approaches.

The temporal instability of model performance resonates with documented challenges in financial forecasting. Like Ojea Ferreiro and Gregori's (2022) models for cross-border M&A, the models applied in this thesis demonstrated regime-dependent accuracy, performing well during stable periods but struggling around structural breaks. This pattern underscores the inherent difficulty in forecasting rare events subject to regime changes.

For practitioners, even modest forecasting improvements carry significant value. The ability to anticipate quarterly deal flow with 2-3 deal accuracy enables more informed decisions about resource allocation, fundraising timing, and market exposure. PE firms could use these forecasts to optimize dry powder deployment strategies, accelerating investments ahead of predicted activity surges or conserving capital during anticipated downturns.

The models' varying performance across market regimes suggests a portfolio approach to forecasting. During stable economic conditions, ensemble methods like Random Forest may provide superior predictions by capturing complex variable interactions. However, during periods of heightened uncertainty or structural change, simpler linear models with strong regularization appear more reliable. This insight argues for maintaining multiple models and weighting their predictions based on prevailing market conditions.

6.2 Limitations of the Study

Several methodological restrictions in this work limit the validity and generalizability of its results. None of the machine learning models attained consistent, statistically significant gains over the naive baseline according to statistical significance testing. Although the models showed somewhat reduced MAEs, significant temporal variance stopped null hypothesis of no improvement from being rejected, so posing basic questions on the applicability of machine learning techniques for aggregate private equity deal flow forecasting.

Data reliability poses significant challenges. The identification of transactions before 1994 suffers from systematic undercounting, hence perhaps introducing biases that machine learning algorithms could magnify rather than reduce. The author investigated several temporal beginning points throughout the research process but saw no improvement in model performance. Furthermore, given six quarters of zero transactions and great variability in deal activity, percentage change measures for quarterly deal counts proved inappropriate for study. The focus on raw quarterly deal counts ignores significant structural changes in capital markets. Raw transaction counts become an increasingly poor indicator of PE market activity as the denominator of possible targets has contracted dramatically and the population of US listed companies is declining over the sample period. The 40-year sample period includes several structural breaks violating basic stationarity presumptions under machine learning approaches. Significant economic upheavals that fundamentally changed market dynamics and regulatory systems included the COVID-19 pandemic and the 2008 financial crisis. Random Forest's performance declines from $R^2 = 0.254$ in cross-valuation to $R^2 = -0.151$ on holdout data, so illustrating the models' limited transferability of historical patterns to changing market environments. At last, strong overfitting seen in ensemble techniques points to limited capacity to derive generalizable patterns from past data. Either too restrictive (Lasso, XGBoost) or too permissive (Random Forest), the feature selection process produced unstable results. Given the possibility that future market conditions will differ greatly from historical precedents, the significant generalization failure between CV and holdout evaluation indicates that models caught historical idiosyncrasies rather than robust predictive relationships.

7. Conclusion

7.1 Summary of Findings

This thesis addressed a significant gap in the private equity literature by developing and evaluating quantitative models to forecast quarterly PE-backed public-to-private transaction volumes in the United States. Through systematic application of machine learning techniques to a comprehensive dataset spanning 1986 to 2024, the research demonstrated both the potential and limitations of data-driven approaches to predicting aggregate deal flow.

The empirical analysis yielded several key findings. First, all tested models achieved modest but consistent improvements over naive baselines, with mean absolute error reductions of 3-5% and R^2 values ranging from 0.145 to 0.254. While these gains fell short of statistical significance at conventional levels, they represent meaningful progress in a domain characterized by extreme volatility and rare events. Second, the research revealed strong path dependence in PE deal activity, with lagged deal counts and momentum indicators consistently ranking among the most important predictors across all modeling approaches. Third, the analysis uncovered multi-scale temporal patterns, with both seasonal and decade-long cycles contributing to forecast accuracy, supporting theoretical perspectives on the cyclical nature of PE investment.

The comparative evaluation of modeling paradigms provided additional insights. Linear models with regularization (Lasso) offered the most stable and interpretable results. Ensemble methods captured important nonlinear relationships, with Random Forest achieving the highest overall R^2, though at the cost of interpretability. Gradient boosting approaches unexpectedly underperformed, suggesting that model complexity must be carefully balanced against data sparsity in PE applications.

7.2 Contribution to Knowledge and Practical Implications

This research makes several contributions to academic knowledge. Methodologically, it demonstrates the feasibility of applying modern machine learning techniques to PE market forecasting while highlighting the importance of appropriate feature engineering and temporal validation frameworks. The systematic comparison across linear, bagging, and boosting paradigms provides guidance for future researchers navigating the bias-variance tradeoff in financial count data prediction.

Theoretically, the findings support and extend existing frameworks for understanding PE market dynamics. The dominant role of autoregressive features confirms strong momentum effects, while the importance of macroeconomic principal components validates links between broader economic conditions and deal activity. The identification of multi-scale cyclical patterns contributes new evidence for adaptive market perspectives on PE investment timing.

For practitioners, the research offers actionable insights despite modest absolute performance gains. PE firms can leverage even small improvements in deal flow forecasting to optimize fundraising timing, adjust deployment pacing, and manage dry powder more effectively. The models' ability to capture directional movements, if not precise counts, provides valuable input for strategic planning and resource allocation decisions. Investment banks and advisory firms may similarly benefit from improved visibility into future deal flow when staffing teams and pursuing mandates. The identified feature importance patterns also carry practical implications. The strong predictive power of momentum indicators suggests that practitioners should monitor recent deal flow trends closely when forming market outlook expectations.

7.3 Future Research Directions

Several promising avenues emerge to extend this research. Incorporating higher frequency data is a valuable direction. Monthly or even weekly deal counts, while noisier, could reveal short-term dynamics masked by quarterly aggregation. Geographic expansion beyond the US market would enhance generalizability and enable cross-country comparisons. Developing region-specific models for European and Asian markets, then exploring transfer learning or multi-task learning approaches, could identify universal versus market-specific drivers of P2P activity. This international perspective would be particularly valuable given the increasing globalization of PE investment. Finally, extending beyond aggregate deal counts to model deal characteristics would provide richer insights. Predicting not just the number but also the size distribution, sector composition, or leverage characteristics of future deals would enable more nuanced strategic planning. Hierarchical or multi-output models could simultaneously forecast multiple aspects of deal flow, providing a comprehensive view of future market activity.

This thesis has demonstrated that while perfect prediction of PE deal flow remains elusive, systematic application of machine learning techniques can extract meaningful signals from complex market data. As data availability improves and methodological innovations continue, the frontier of feasible and valuable PE Market forecasting will undoubtedly expand, supporting more informed decision-making across the investment ecosystem.

Appendix

Table 5: SDC Platinum Query

Results	Type	Description	Operator	Value
	Content Set	Mergers & Acquisitions		M&A
1,554,166	Data Items	Date Announced		All Dates (1 Jan 1962 - 11 May 2025)
441,940	Data Items	Target Nation	Include	United States
177,068	Data Items	Deal Value (Curn=USD, Scale=6)	Greater Than Or Equal	1
56,283	Data Items	Public Status (DealPartRole=T)	Include	Values: Public
27,774	Data Items	Deal Status	Include	Values: Completed
27,774	Output	Report: MnA_Report_#2_withgoingprivateflag (Columnar Grid)		

Note: The SDC Platinum query was designed to be as unrestrictive as possible to maximize data capture, as many fields in the database are not consistently maintained. Applying strict filters or flags on potentially empty fields would exclude valid records, so a more inclusive query approach proved most effective for comprehensive data retrieval.

Table 6: SDC Platinum retrieved Variables

Variable Name	Variable Description
Date Effective	Date Effective: Date when the entire transaction is completed and effective. In a two-step transaction this is the date when the second-step merger is completed. See also DUNCON (Date Unconditional) when searching targets headquartered in the United Kingdom, Australia, and New Zealand.
Date Announced	Date Announced: The date one or more parties involved in the transaction makes the first public disclosure of common or unilateral intent to pursue the transaction (no formal agreement is required). Among other things, Date Announced is determined by the disclosure of discussions between parties, disclosure of a unilateral approach made by a potential bidder, and the disclosure of a signed Memorandum of Understanding (MOU) or other agreement. For transactions prior to 2006 this date is set to equal Rank Date. Date Announced must be dated either on or before the Rank Date. Therefore, in cases where the first public announcement of a transaction is made after the transaction has completed, Date Announced should equal both Rank Date and Date Effective/Unconditional.
Date Originally Announced	Date Originally Announced: The date when the target company is first publicly disclosed as a possible takeover candidate. DAO is used for the calculation of stock premiums. When multiple bidders exist, the DAO is recorded in the following cases: (1) If acquiror changes from 'Seeking Buyer' or 'Undisclosed Acquiror' to an actual entity. (2) Competing bids are announced. (3) Competing stakes are announced. (4) A defensive transaction is announced. In most cases DAO should be dated before Date Announced. However, in cases where the first public announcement of a transaction was made after the transaction has completed, the DAO should be dated after the Date Effective/Unconditional.
Target Short Name	Target Short Name: Short target company name; up to 30 characters.
Target Full Name	Target Full Name: Full target company name on 1 line; up to 75 characters.

Deal Synopsis	Deal Synopsis: A 600 character text field summarizing the events of the transaction. Includes: (1) Parties involved (2) Explanation of consideration, including charges and analyst estimates (3) Challenging bids (4) Acquisition Techniques, if important (5) Attitude, if important (6) Defensive Tactics (7) Lockup Description Depending upon importance to deal, may also include: (1) Bid History (2) Related Deals (3) Financial Advisors (4) Litigation (5) Regulatory Agencies
Deal Value (USD Millions)	Deal Value: Total value of consideration paid by the acquiror, excluding fees and expenses. The dollar value includes the amount paid for all common stock, common stock equivalents, preferred stock, debt, options, assets, warrants, and stake purchases made within six months of the announcement date of the transaction. Liabilities assumed are included in the value if they are publicly disclosed. Preferred stock is only included if it is being acquired as part of a 100% acquisition. If a portion of the consideration paid by the acquiror is common stock, the stock is valued using the closing price on the last full trading day prior to the announcement of the terms of the stock swap. If the exchange ratio of shares offered changes, the stock is valued based on its closing price on the last full trading date prior to the date of the exchange ratio change. For public target 100% acquisitions, the number of shares at date of announcement (CACT) is used.
Acquiror Short Business Description	Acquiror Short Business Description : Indicates the primary business of the Acquiror
Acquiror Immediate Parent Short Business Description	Acquiror Immediate Parent Short Business Description : Indicates the primary business of the Acquiror Immediate Parent
Acquiror Ultimate Parent Short Business Description	Acquiror Ultimate Parent Short Business Description : Indicates the primary business of the Acquiror Ultimate Parent
Acquiror Primary Venture Economics Industry	Acquiror Primary Venture Economics Industry: An acquiring company's main line of business using the Venture Economics Industry Code (VEIC) classification as defined by Company. VEIC's, the industry standard in the private equity market, are detailed industry classifications that focus on emerging technologies, life sciences and other industries that private equity firms invest in.
Acquiror Immediate Parent Primary Venture Economics Industry	Acquiror Immediate Parent Primary Venture Economics Industry: Pertaining to the immediate parent company of an acquiror; the main line of business using the Venture Economics Industry Code (VEIC) classification as defined by Company. VEIC's, the industry standard in the private equity market, are detailed industry classifications that focus on emerging technologies, life sciences and other industries that private equity firms invest in.
Acquiror Ultimate Parent Primary Venture Economics Industry	Acquiror Ultimate Parent Primary Venture Economics Industry: Pertaining to the ultimate parent company of an acquiror; the main line of business using the Venture Economics Industry Code (VEIC) classification as defined by Company. VEIC's, the industry standard in the private equity market, are detailed industry classifications that focus on emerging technologies, life sciences and other industries that private equity firms invest in.
Acquiror PermID	Acquiror PermID: Permanent Identifier of an Acquiror.
Acquiror Immediate Parent PermID	Acquiror Immediate Parent PermID : Unique Permanent Identifier of Acquiror Immediate Parent
Acquiror Ultimate Parent PermID	Acquiror Ultimate Parent PermID : Unique Permanent Identifier of Acquiror Ultimate Parent
Investor PermID	Investor PermID : Unique Permanent Identifier of Investor
Investor Immediate Parent PermID	Investor Immediate Parent PermID : Unique Permanent Identifier of Investor Immediate Parent
Investor Ultimate Parent PermID	Investor Ultimate Parent PermID : Unique Permanent Identifier of Investor Ultimate Parent

Acquiror is a Leveraged Buyout Firm	Acquiror is an LBO Firm Flag: Retrieves deals where the acquiror company in a transaction is a buyout firm regardless of the nature of the transaction itself. Use LBO to find leveraged buyout transactions.
Acquiror Ultimate Parent is an LBO Firm	Acquiror Ultimate Parent is an LBO Firm Flag: A True/False flag set to True when the acquiror's ultimate parent is an LBO firm.
Leveraged Buyout Flag	Leveraged Buyout Flag: Retrieves leveraged buyout transactions. Company includes transactions in which management forms a part of the investor group in this definition, as well as transactions that are identified as an LBO in the financial press if a majority interest of the target company is acquired.
Deal Type	Deal Type: In the Mergers & Acquisition database, you can use the Deal Type window to specify the type of transaction to select or exclude in your search criteria. You can indicate whether you want to select or exclude Disclosed Value Mergers & Acquisitions, Undisclosed Value Mergers & Acquisitions, or both. You can also select and/or exclude one or more specific types of transactions, such as leveraged buyouts and/or tender offers. Transaction Type Code: Code number for the type of transaction (e.g. 1=DI): 1 = Disclosed Value; 2 = Undisclosed Value; 3 = Leveraged Buyouts; 4 = Tender Offers; 5 = Spinoffs; 6 = Recapitalizations; 7 = Self-Tenders; 8 = Exchange Offers; 9 = Repurchases; 10 = SP: indicates all deals in which a company is acquiring a minority stake (i.e. up to 49.99% or from 50.1% to 99.9%) in the target company; 11 = Acquisitions of Remaining Interest; 12 = Privatizations
Acquisition Techniques	Acquisition Techniques: Acquisition technique code number, e.g. 8 (Divestiture): A yes/no flag which indicates significant characteristics about the transaction: (…) Acquiror Includes Management: 'Y' indicates that the management of the target company is taking an equity interest in the target company as part of the acquisition. Acquiror is an Investor Group: 'Y' indicates that the acquiror is an investor group. (…) Going Private Flag: 'Y' indicates that a private acquiror ('private' meaning that none of the acquiror's ultimate parentage is public either) is acquiring a public target and upon completion, it will become a private company. Institutional Buyout: Yes/No flag set to 'Y' for a highly leveraged transaction where one or more institutional investors act together to lead or initiate a buyout deal. (…) LBO + Management + Employee: Yes/No flag set to 'Y' for a highly leveraged transaction where employees in conjunction with existing target management, backed by one or more institutional investors, lead or initiate a buyout deal. Leveraged Buyout Flag: 'Y' indicates that the transaction is a leveraged buyout. An "LBO" occurs when an investor group, investor, or firm offers to acquire a company, taking on an extraordinary amount of debt, with plans to repay it with funds generated from the company or with revenue earned by selling off the newly acquired company's assets. TR considers an LBO if the investor group includes management or the transaction is identified as such in the financial press and a majority interest of the company is acquired. (…) Privatization Flag: 'Y' indicates a government or government controlled entity sells shares or assets to a non-government entity. Privatizations include both direct and indirect sales of up to a 100% stake to an identifiable buyer and floatations of stock on a stock exchange. The former is considered an M&A transaction and will be included in the quarterly rankings; the latter will not. (…)
Acquiror Immediate Parent PermID	Acquiror Immediate Parent PermID : Unique Permanent Identifier of Acquiror Immediate Parent
Acquiror Ultimate Parent PermID	Acquiror Ultimate Parent PermID : Unique Permanent Identifier of Acquiror Ultimate Parent
Target 6-digit CUSIP	Target 6-digit CUSIP : The 6-digit CUSIP of the Target. The CUSIP is a six character, unique identifier for every company. Every U.S. company with publicly traded securities is

	listed in the Standard & Poors CUSIP directory and has a CUSIP assigned to it. When companies are sorted by their CUSIPs, the listing will be approximately alphabetical. In cases where a CUSIP has not been assigned, SDC will estimate one according to the rules specified by S&P.
Target Cusip State	Target Cusip State : States the Cusip Type, A for Actual Cusip assigned by Cusip.com and E for Estimated Cusip.
Acquiror Short Name	Acquiror Short Name: Acquiring company's short name on 1 line, up to 30 characters.
Acquiror Full Name	Acquiror Full Name: Acquiring company's full name on 1 line, up to 75 characters.
Acquiror Immediate Parent Name	Acquiror Immediate Parent Name : Short company name; up to 30 characters.
Acquiror Ultimate Parent Name	Acquiror Ultimate Parent Name : Short company name; up to 30 characters.
Investor Name	Investor Name : Short company name; up to 30 characters.
Target SIC	Target SIC: Description of SIC codes for target's line of business.
Target Primary SIC	Target Primary SIC: Description of SIC code for target's primary line of business.
Target Major SIC Group	-
Target NAIC 2007	Target NAIC 2007: North American Industrial Classification description. Describes which industry the target is a part of, as classified by the government.
Target NAIC 2022	Target NAIC 2022: North American Industrial Classification description. Describes which industry the target is a part of, as classified by the government.
SDC Deal No	SDC Deal Number: Unique nine digit number assigned to every individual transaction.
Going Private Flag	Going Private Flag: 'Y' indicates that a private acquiror or a financial sponsor is acquiring a public target and upon completion, the target will no longer have any of its shares traded on the public market. In cases where an investor group is acquiring a public target, the Going Private Flag will be set to Y when there is buyside financial sponsor activity. The Going Private Flag will also be set to 'Y' even if the public target, one that is originally intended to be taken private, remains a public entity upon deal completion.

Note: The Data Item Browser (accessible from the Workspace app) was used to retrieve the variable descriptions.

Table 7: Final Model Features and Principal Component Composition

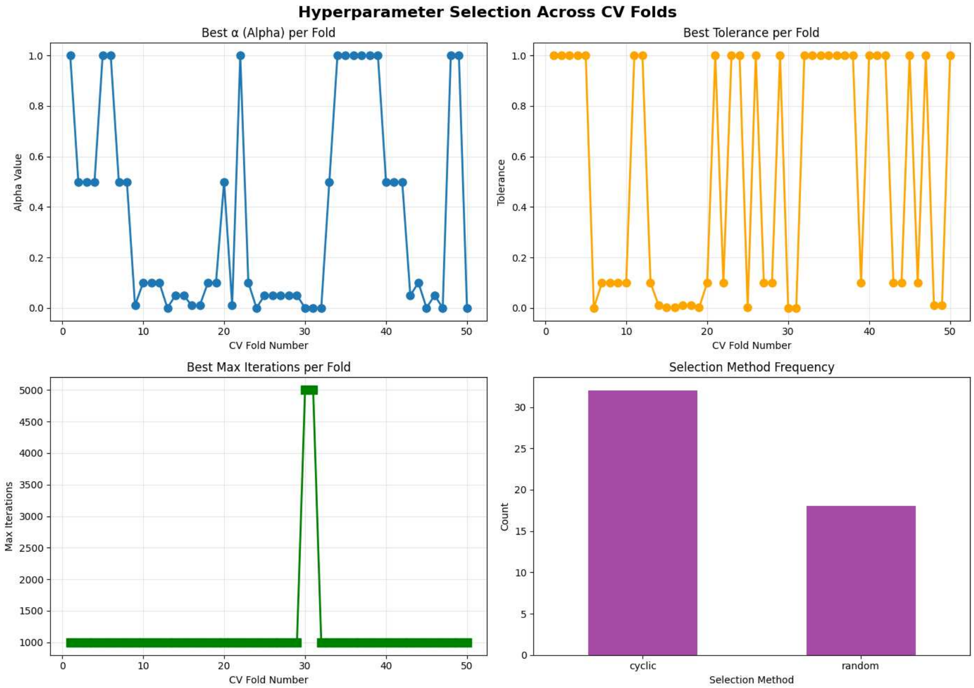
Feature Name	Description
macro_gdp_PC1	Macroeconomic GDP Indicators: GDP (Gross Domestic Product); PINCOME (Personal Income); BORROW (Total Borrowings from Fed); GNP (Gross National Product); GDPC1 (Real GDP); BBKMGDP (Brave-Butters-Kelley Real GDP Nowcast)
employment_PC1	Labor Market Conditions: UNRATE (Unemployment Rate); SAHMCURRENT (Sahm Rule Recession Indicator); LREM64TTUSM156S (Employment Rate 15-64 Years); LNS12032195 (Part-Time for Economic Reasons - All Industries); LNS12032198 (Part-Time for Economic Reasons - Nonagricultural)
yield_spreads_PC1	Term Structure and Yield Spreads: T10YFFM (10Y Treasury - Fed Funds); T5YFFM (5Y Treasury - Fed Funds); T1YFFM (1Y Treasury - Fed Funds); TB3SMFFM (3M Treasury - Fed Funds); TB6SMFFM (6M Treasury - Fed Funds); T10Y3MM (10Y-3M Spread); T10Y2YM (10Y-2Y Spread)
credit_conditions_PC1	Credit Market Conditions: AAA10YM (Aaa Corporate Bond Spread); BAA10YM (Baa Corporate Bond Spread); Bofa High Yield Index (High Yield Spread); Bankrupt (Bankruptcy Filings)

us_stock_market_PC1	US Equity Market Indicators: SP DIV YLD (S&P 500 Dividend Yield); SP CAPE (S&P 500 CAPE Ratio); SP EPS (S&P 500 Earnings Per Share); SP CASH DIV (S&P 500 Cash Dividend); spindx (S&P 500 Price Level); spindx_qoq (S&P 500 Quarterly Return); ipo_count (Ritter IPO Count)
global_stock_markets_PC1	International Equity Markets: SPASTT01USM657N (US Share Prices Growth); SPASTT01FRM657N (France Share Prices Growth); SPASTT01GBM657N (UK Share Prices Growth); SPASTT01DEM657N (Germany Share Prices Growth); SPASTT01JPM657N (Japan Share Prices Growth); MSCIWORLD (MSCI World Index)
consumer_sentiment_PC1	Consumer Confidence Indicators: MICH (Michigan Inflation Expectation); UMCSENT (Michigan Consumer Sentiment); CSINFT02USM460S (Consumer Price Future Tendency)
business_sentiment_PC1	Business Confidence and Activity: BSCICP03USM665S (Composite Business Confidence); BSCICP02USM460S (Manufacturing Confidence); BSPRTE02USM460S (Production Tendency); BSOITE02USM460S (Orders Inflow Tendency); BSCURT02USM160S (Capacity Utilization); AS % Bear (Advisors Sentiment - Bearish %)
emv_macro_PC1	Equity Market Volatility - Macroeconomic News: EMVMACRONEWS (Overall Macro News); EMVMACROBROAD (Broad Quantity Indicators); EMVMACROINFLATION (Inflation News); EMVMACROFININD (Financial Indicators); EMVMACROLABORMKT (Labor Markets); EMVMACRORE (Real Estate); EMVMACROTRADE (Trade); EMVMACROBUS (Business Investment); EMVMACROCONSUME (Consumer Spending)
emv_policy_PC1	Equity Market Volatility - Policy Related: EMVPOLRLTDEMV (Overall Policy Related); EMVFISCALPOL (Fiscal Policy); EMVTAXESEMV (Tax Policy); EMVGOVTSPEND (Government Spending); EMVWELFARE (Entitlement Programs); EMVMONETARYPOL (Monetary Policy)
pe_exits_PC1	Private Equity Exit Activity: number_of_exits (Total PE Exit Count); number_of_IPO_exits (IPO Exit Count); exit_duration (Average Exit Duration in Years); exit_proceeds_related_equity (Exit Proceeds in USD Millions)
pe_capital_PC1	Private Equity Capital Availability: drypowder (Uninvested PE Capital); Total Capitalization (GP + LP Capital); fundraise_funds (Number of Funds Raised); fundraise_companies_invested (Portfolio Companies); fundraise_avg_size_musd (Average Fund Size); fundraise_amt_range_musd (Total Amount Raised)
pe_returns_top_PC1	Private Equity Performance Metrics: Pooled Return (LP) (%); Top 5% (LP) (%); Upper Quartile (LP) (%); Equal Weighted (LP) (%); Average (LP) (%); mPME Index IRR (LP) (%)
inflation_PC1	Inflation and Price Indicators: CPIAUCSL (CPI All Urban Consumers); APU000072620 (Utility Gas Price); APU000072610 (Electricity Price); IR10000 (Import Price Index - Crude Oil); CCRETT01USM661N (Real Effective Exchange Rate)
DFF	Federal Funds Effective Rate (Daily)
DGS3	3-Year Treasury Constant Maturity Rate
SP Precious Metal Index	S&P GSCI Precious Metals Total Return Index
Property Industrial	NCREIF Property Index - Industrial Real Estate
M2SL	M2 Money Supply
Tin	LME Tin Cash Price (USD/Metric Tonne)
SP Energy Index	S&P GSCI Energy Total Return Index

sin_annual_1	Annual Seasonality - Sine Component (Harmonic 1)
cos_annual_1	Annual Seasonality - Cosine Component (Harmonic 1)
sin_annual_2	Annual Seasonality - Sine Component (Harmonic 2)
cos_annual_2	Annual Seasonality - Cosine Component (Harmonic 2)
sin_cycle	Business Cycle - Sine Component (10-Year Period)
cos_cycle	Business Cycle - Cosine Component (10-Year Period)
deal_count_lag1	Lagged Deal Count (t-1)
deal_count_lag4	Lagged Deal Count (t-4, Year-over-Year)
deal_count_roll_mean_4_lag	4-Quarter Rolling Mean (Lagged)
deal_count_roll_std_4_lag	4-Quarter Rolling Standard Deviation (Lagged)
deal_count_roll_mean_8_lag	8-Quarter Rolling Mean (Lagged)
deal_count_roll_std_8_lag	8-Quarter Rolling Standard Deviation (Lagged)
deal_count_roll_mean_12_lag	12-Quarter Rolling Mean (Lagged)
deal_count_roll_std_12_lag	12-Quarter Rolling Standard Deviation (Lagged)
deal_count_qoq_pct	Quarter-over-Quarter Percentage Change
deal_count_yoy_pct	Year-over-Year Percentage Change
deal_count_vs_trend	Deviation from Trend
deal_count_rising	Binary Rising Indicator (1 if Increasing)
deal_count_streak	Consecutive Quarters of Growth/Decline
deal_count_qoq_abs	Absolute Quarter-over-Quarter Change
deal_count_accel	Deal Count Acceleration (Second Derivative)

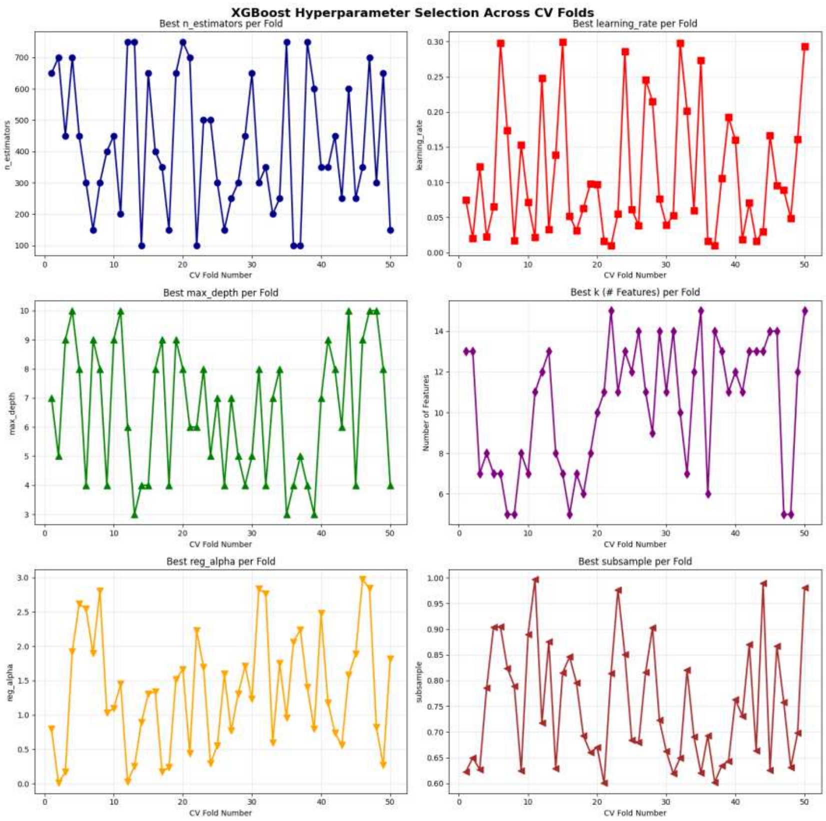
Note: For detailed descriptions of each variable or series—including its source, category, frequency, and units see Table 8.

Figure 13: Lasso Hyperparameter Selection Across CV Folds



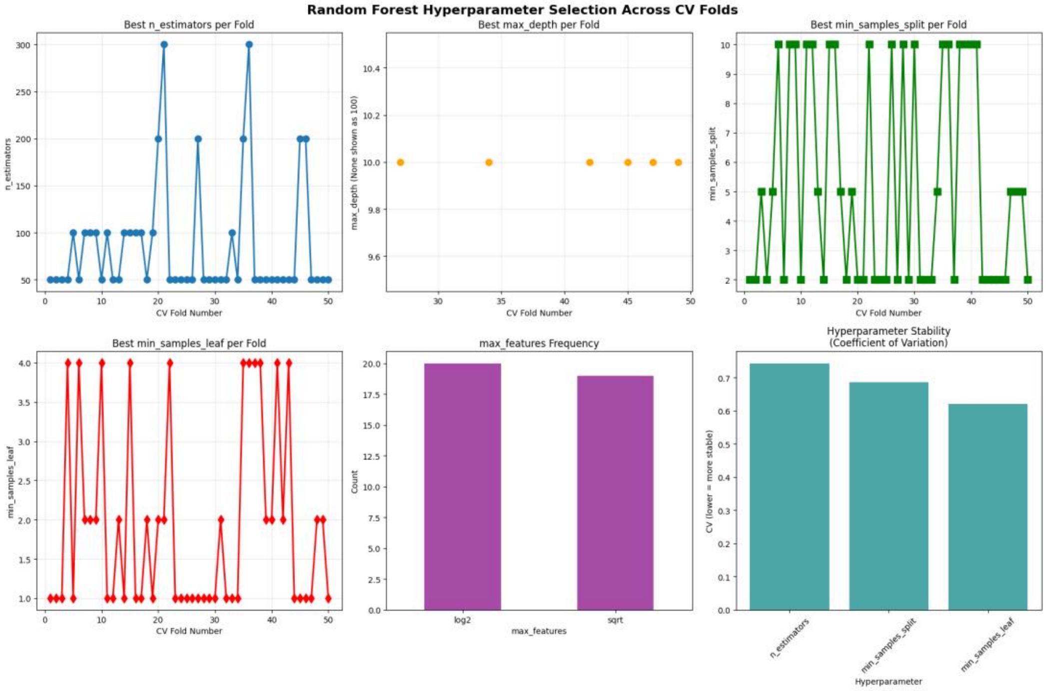
Note: Alpha decreases over time (0.1→0.01) as more data allows weaker regularization. Stable convergence parameters. 'Cycler' coordinate descent outperforms 'random'.

Figure 14: XGBoost Hyperparameter Selection Across CV Folds



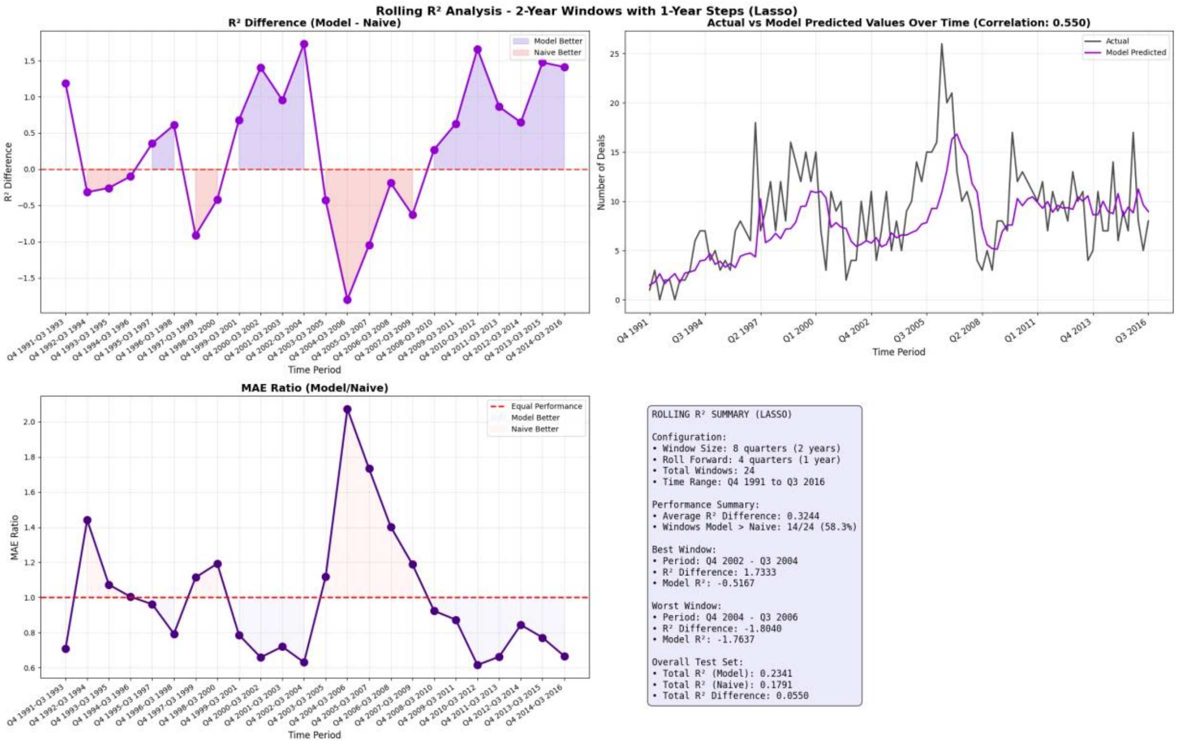
Note: High variability in feature selection and tree depth across folds. Unstable optimal configurations suggest period-dependent patterns, explaining poor generalization.

Figure 15: Random Forest Hyperparameter Selection Across CV Folds



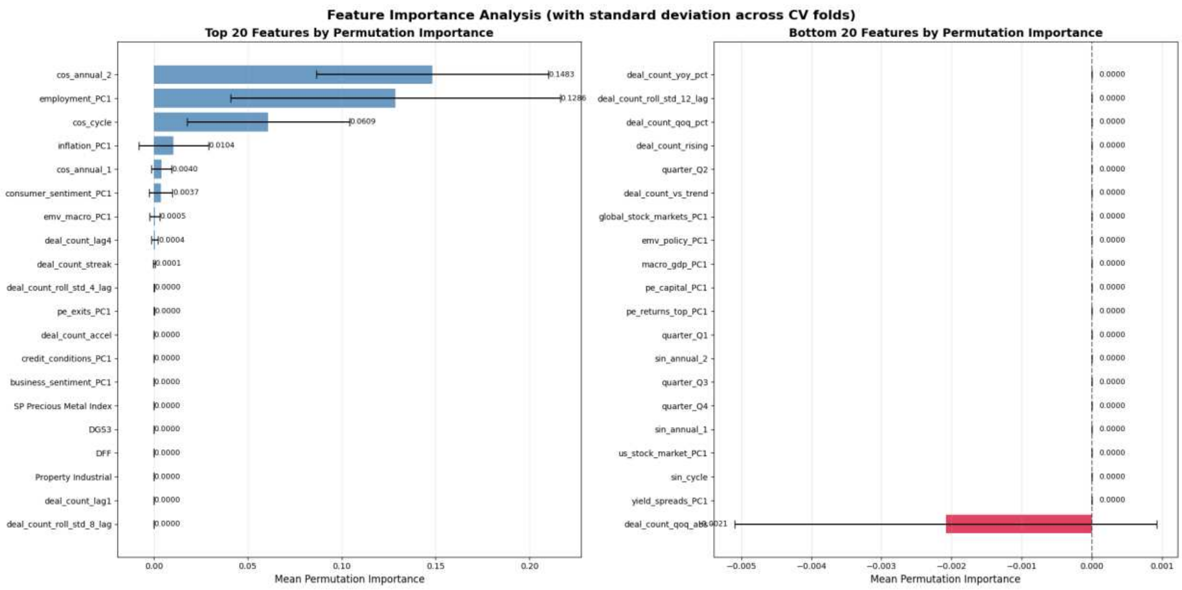
Note: Moderate stability: estimators increase (50→200), depth varies (5-15). More stable than XGBoost but still shows adaptation challenges for regime shifts.

Figure 16: Lasso Training Performance



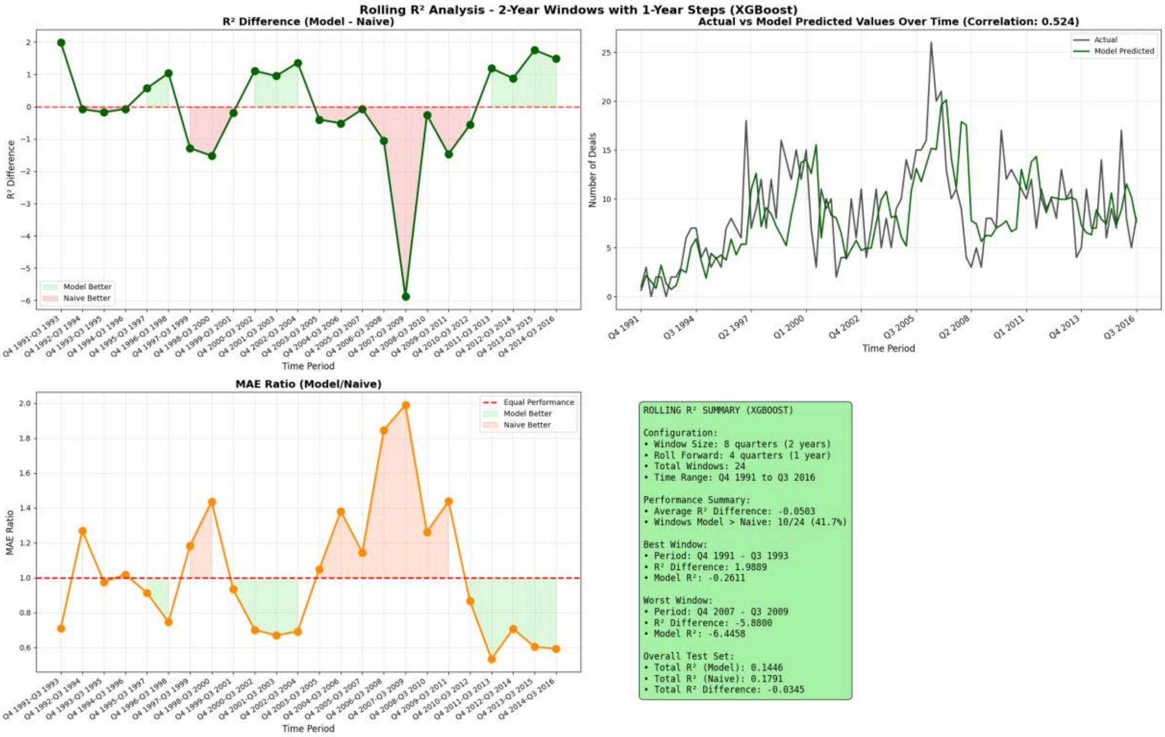
Note: Upper panel shows R² improvements 0-2% above naive baseline with MAE reductions reaching 40% in favorable periods. Performance peaks during high-activity periods (2006-2007), where linear relationships dominate. Errors concentrate at regime transitions but recover quickly.

Figure 17: Lasso Feature Importance Training



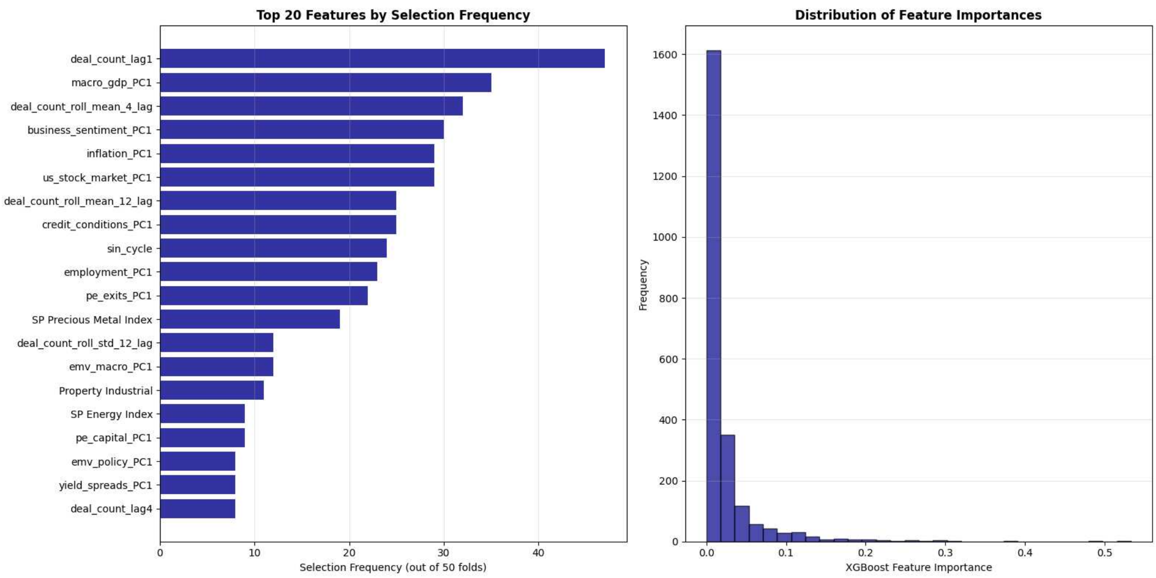
Note: Left panel shows top 20 features by permutation importance. Cyclical features dominate (cos_annual_2: 0.12±0.04), followed by employment_PC1 (0.10±0.05) and cos_cycle (0.09±0.04). Right panel reveals deal_count_qoq_abs has negative importance (-0.005), indicating this volatility measure degrades predictions. Error bars show importance stability across CV folds.

Figure 18: XGBoost Training Performance



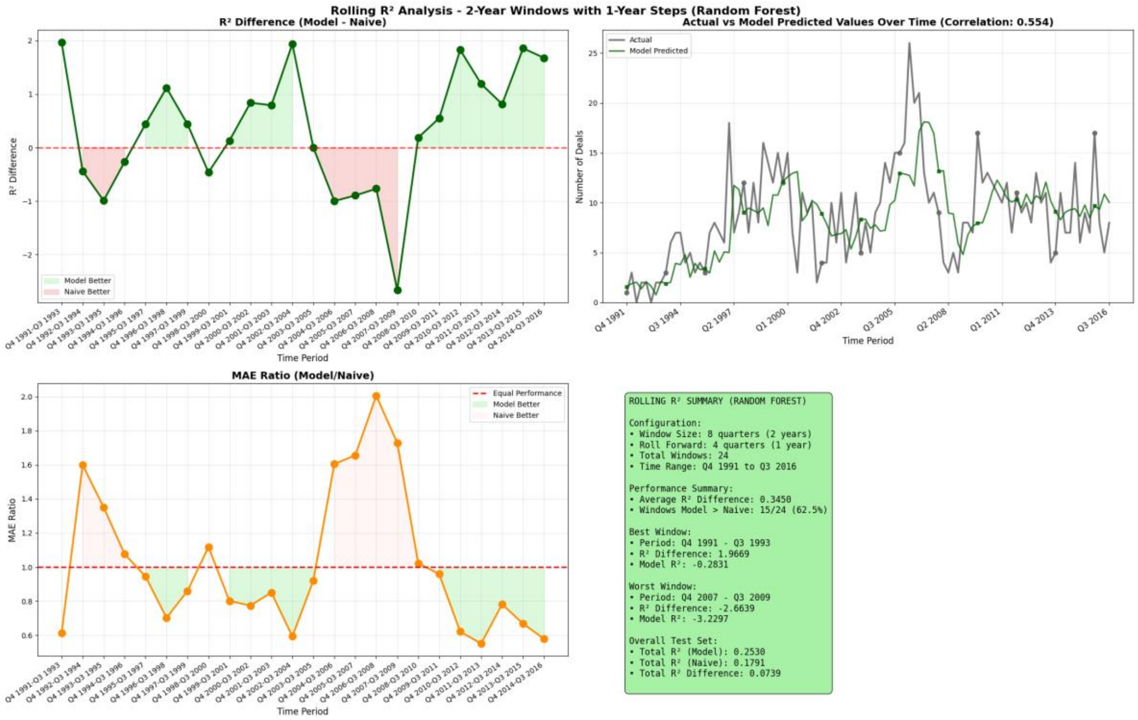
Note: XGBoost's training performance exhibits high volatility, with R² differences ranging from -6 to +2 relative to baseline. While average performance is positive, substantial negative periods indicate overfitting during certain market regimes.

Figure 19: XGBoost Feature Selection: Frequency of Selection Across Training Iterations



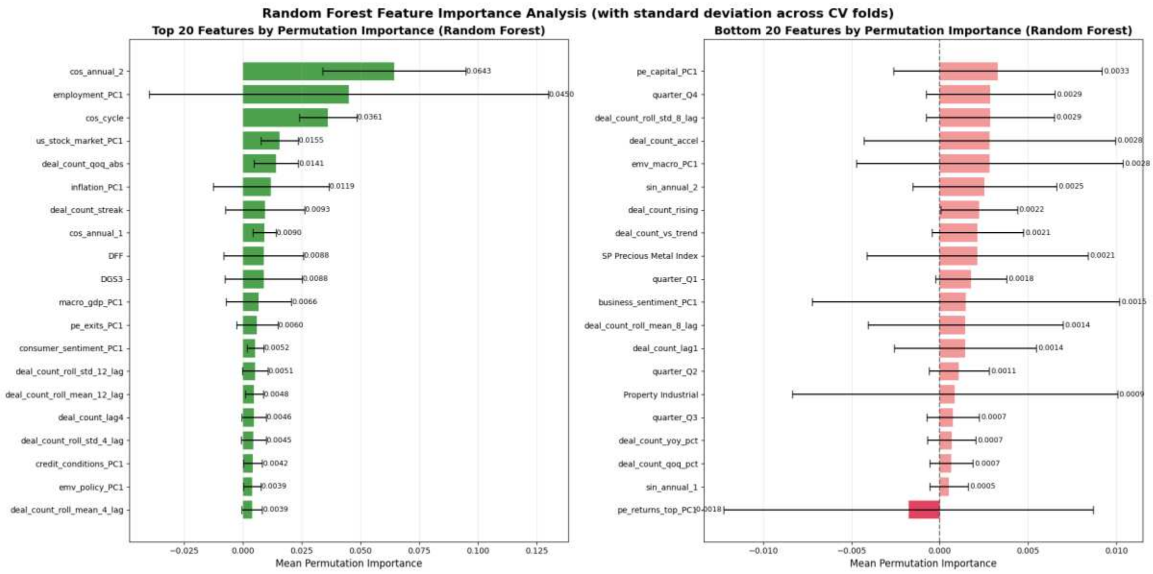
Note: Histogram reveals extreme feature concentration: ~80% of features have <5% usage frequency. The top 10 features account for >70% of splits, dominated by deal momentum indicators and lagged values.

Figure 20: Random Forest Training Performance



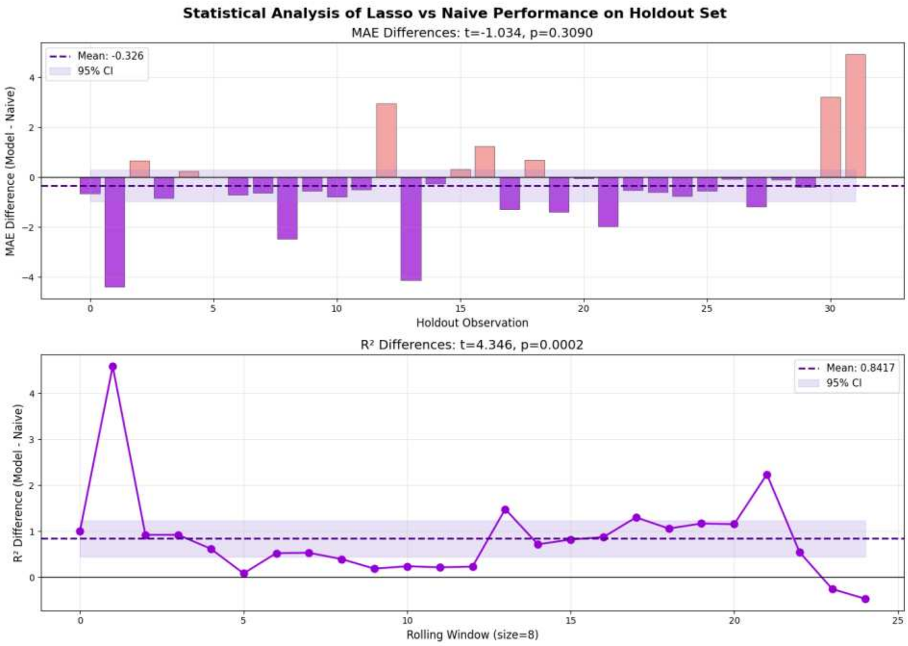
Note: Rolling R² analysis shows model outperforms naive baseline in 62.5% of 2-year windows. MAE ratios (bottom left) below 1.0 indicate consistent error reduction. Model performs best during stable periods but struggles with extreme values.

Figure 21: Random Forest Feature Importance Training



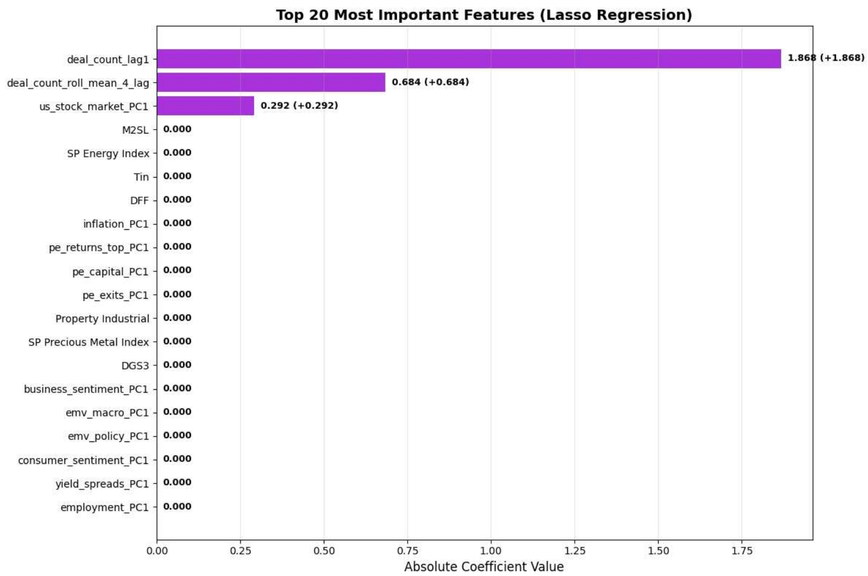
Note: Permutation importance reveals balanced feature usage: cos_annual_2 leads (0.064±0.032), followed by employment_PC1 (0.045±0.021) and cos_cycle (0.036±0.018). Negative importance for volatility measures (deal_count_qoq_abs: -0.002) indicates these features add noise.

Figure 22: Lasso Holdout Set Statistical Analysis



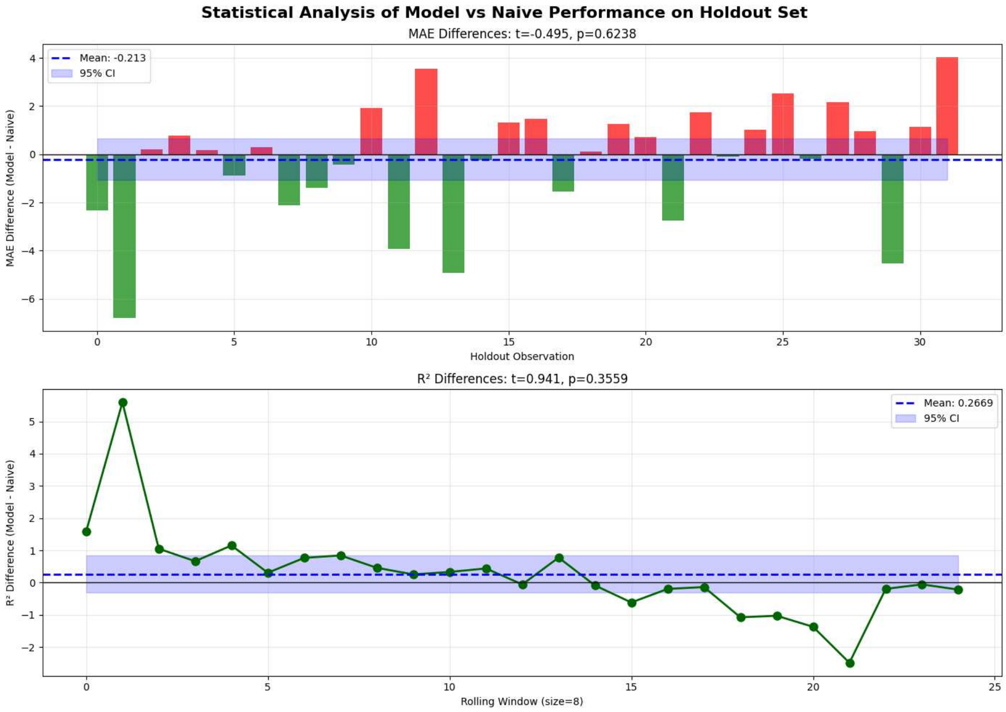
Note: Top panel shows MAE differences (Model - Naive) across 32 holdout observations. Most bars are negative (purple), indicating Lasso outperforms naive, though not significantly (t=-1.034, p=0.309). Bottom panel displays R^2 differences over rolling 8-quarter windows, showing significant improvement (t=4.346, p=0.0002) with mean difference of 0.8417. The discrepancy between MAE and R^2 significance suggests Lasso better captures variance patterns than point predictions.

Figure 23: Lasso Holdout Set Feature Importance



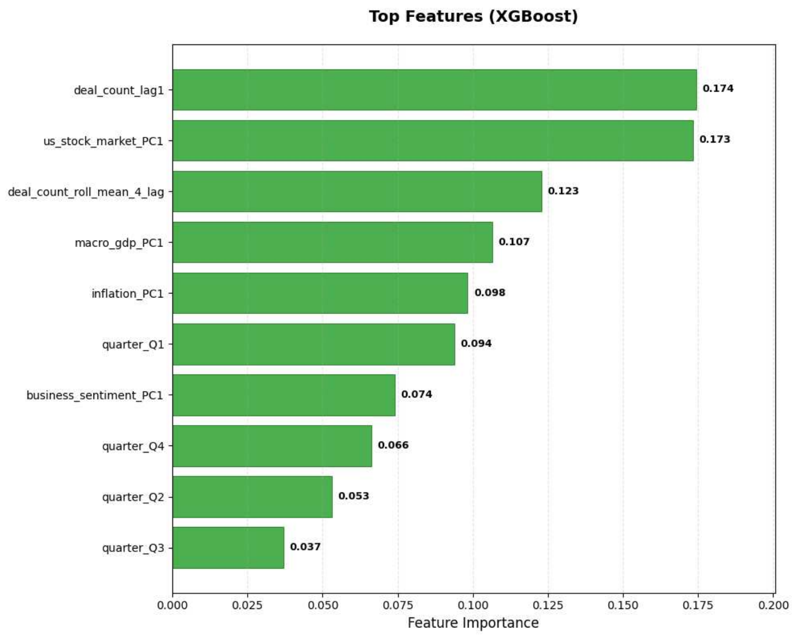
Note: Shows absolute coefficient values for top 20 features in final Lasso model. Only three features have non-zero coefficients: deal_count_lag1 (1.868±1.868), deal_count_roll_mean_4_lag (0.664±0.664), and us_stock_market_PC1 (0.292±0.292). All other features zeroed out by L1 regularization.

Figure 24: XGBoost Holdout Set Statistical Analysis



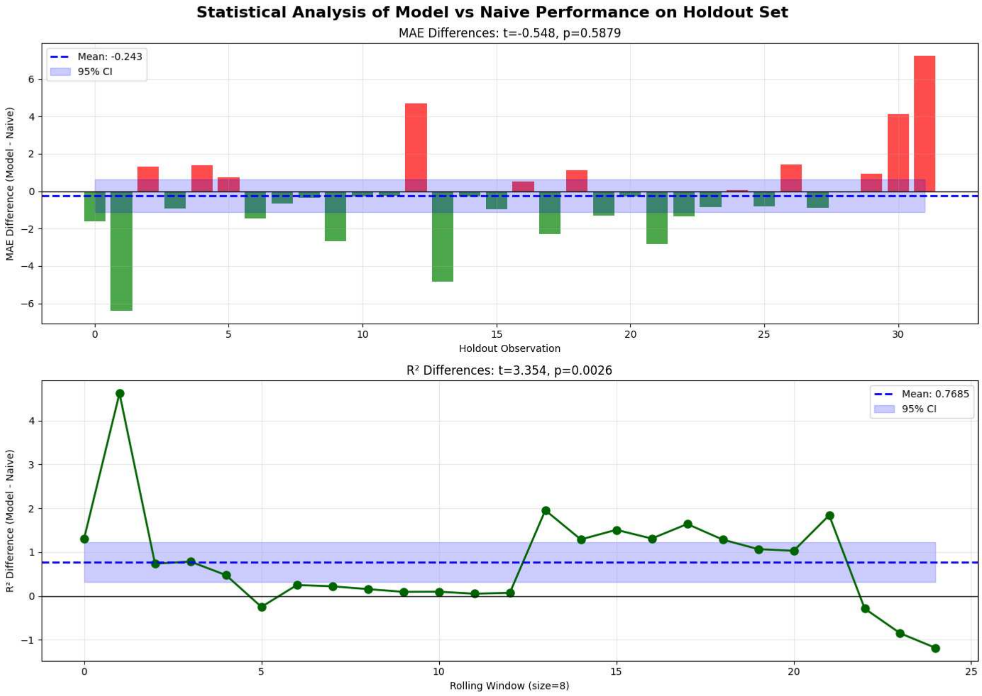
Note: Presents t-test results comparing XGBoost vs naive on holdout set. Top panel: MAE differences across 32 observations show high variability (σ=2.40), with mean difference of -0.214 (model better) but not significant (t=-0.495, p=0.624). Bottom panel: R^2 differences over 25 rolling 8-quarter windows show positive mean of 0.267 but also not significant (t=0.941, p=0.356).

Figure 25: XGBoost Holdout Set Feature Importance



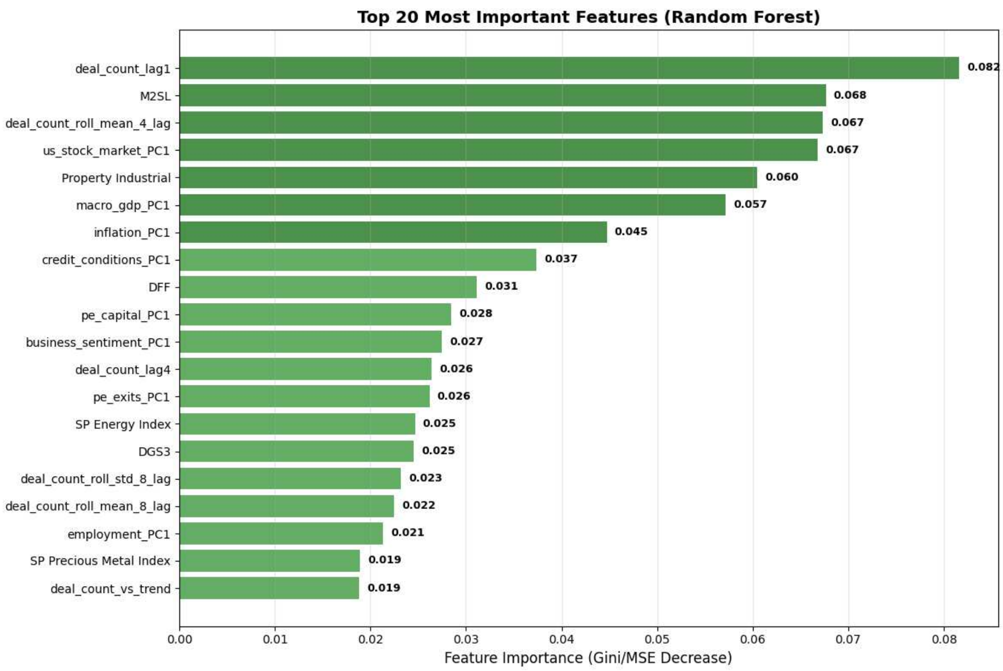
Note: Lagged deal count and US stock market principal component are the two most influential features, both with nearly equal importance scores around 0.17.

Figure 26: Random Forest Holdout Set Statistical Analysis



Note: Top panel: MAE differences show mean of -0.243 (model slightly better) but not significant (t=-0.548, p=0.588) with high variability across 32 observations. Bottom panel: R^2 differences over rolling windows achieve statistical significance with mean 0.768 (t=3.354, p=0.003), indicating consistent variance explanation improvement.

Figure 27: Random Forest Holdout Set Feature Importance



Note: Top feature (deal_count_lag1) having only 0.082 importance versus much lower values, suggesting the model relies on a broader set of features rather than being dominated by a few. The top 20 features encompass deal-related metrics (various lags and rolling means), macroeconomic indicators (M2 money supply, inflation, GDP components), and financial market indices, indicating the model captures diverse economic signals.

Table 8: Complete Variable List

Database	Category	Database Identifier	Title	Frequency	Units
Datastream	advisor_sentiment	USIIBER	ADVISORS SENTIMENT BEARISH	Daily	Percent
Fred	business_sentiment	BSCICP03USM665S	Composite Leading Indicators: Composite Business Confidence Amplitude Adjusted for United States	Monthly	Normalised (Normal=100)
Fred	business_sentiment	BSCICP02USM460S	Business Tendency Surveys (Manufacturing): Confidence Indicators: Composite Indicators: National Indicator for United States	Monthly	Percent
Fred	business_sentiment	BSOITE02USM460S	Business Tendency Surveys: Orders Inflow: Economic Activity: Manufacturing: Tendency for United States	Monthly	Percentage balance
Fred	business_sentiment	BSPRTE02USM460S	Business Tendency Surveys: Production: Economic Activity: Manufacturing: Tendency for United States	Monthly	Percentage balance
Fred	business_sentiment	BSCURT02USM160S	Business Tendency Surveys: Rate of Capacity Utilisation: Economic Activity: Manufacturing: Current for United States	Monthly	Percentage balance
Fred	consumer_sentiment	UMCSENT	University of Michigan: Consumer Sentiment	Monthly	Index 1966:Q1=100
Fred	consumer_sentiment	MICH	University of Michigan: Inflation Expectation	Monthly	Percent
Fred	consumer_sentiment	CSINFT02USM460S	Consumer Opinion Surveys: Consumer Prices: Future Tendency for United States	Monthly	Percentage balance
Fred	credit_conditions	AAA10YM	Moody's Seasoned Aaa Corporate Bond Yield Relative to Yield on 10-Year Treasury Constant Maturity	Monthly	Percent
Fred	credit_conditions	BAA10YM	Moody's Seasoned Baa Corporate Bond Yield Relative to Yield on 10-Year Treasury Constant Maturity	Monthly	Percent
Datastream	credit_conditions	MLHMACL	ICE Bank Of America United States High Yield Index	Daily	Total Return Index (USD)
Datastream	credit_conditions	USBNKRPTP	BANKRUPTCY FILINGS - TOTAL BUSINESS (12 MO ENDING)	Quarterly	Volume Actual Bankruptcy Filings
Fred	employment	UNRATE	Unemployment Rate	Monthly	Percent

Fred	employment	SAHMCURRENT	Sahm Rule Recession Indicator	Monthly	Percentage Points
Fred	employment	LNS12032195	Employment Level - Part-Time for Economic Reasons, Slack Work or Business Conditions, All Industries	Monthly	Thousands of Persons
Fred	employment	LNS12032198	Employment Level - Part-Time for Economic Reasons, Slack Work or Business Conditions, Nonagricultural Industries	Monthly	Thousands of Persons
Fred	employment	LREM64TTUSM156S	Infra-Annual Labor Statistics: Employment Rate Total: From 15 to 64 Years for United States	Monthly	Percent
Fred	emv_macro	EMVMACRONEWS	Equity Market Volatility Tracker: Macroeconomic News And Outlook	Monthly	Index
Fred	emv_macro	EMVMACROBROAD	Equity Market Volatility Tracker: Macroeconomic News and Outlook: Broad Quantity Indicators	Monthly	Index
Fred	emv_macro	EMVMACROINFLATION	Equity Market Volatility Tracker: Macroeconomic News and Outlook: Inflation	Monthly	Index
Fred	emv_macro	EMVMACROFININD	Equity Market Volatility Tracker: Macroeconomic News and Outlook: Other Financial Indicators	Monthly	Index
Fred	emv_macro	EMVMACROLABORMKT	Equity Market Volatility Tracker: Macroeconomic News and Outlook: Labor Markets	Monthly	Index
Fred	emv_macro	EMVMACRORE	Equity Market Volatility Tracker: Macroeconomic News and Outlook: Real Estate Markets	Monthly	Index
Fred	emv_macro	EMVMACROTRADE	Equity Market Volatility Tracker: Macroeconomic News and Outlook: Trade	Monthly	Index
Fred	emv_macro	EMVMACROBUS	Equity Market Volatility Tracker: Macroeconomic News and Outlook: Business Investment And Sentiment	Monthly	Index
Fred	emv_macro	EMVMACROCONSUME	Equity Market Volatility Tracker: Macroeconomic News and Outlook: Consumer Spending And Sentiment	Monthly	Index
Fred	emv_policy	EMVPOLRLTDEMV	Equity Market Volatility Tracker: Policy Related	Monthly	Index
Fred	emv_policy	EMVFISCALPOL	Equity Market Volatility Tracker: Fiscal Policy	Monthly	Index
Fred	emv_policy	EMVTAXESEMV	Equity Market Volatility Tracker: Taxes	Monthly	Index
Fred	emv_policy	EMVGOVTSPEND	Equity Market Volatility Tracker: Government Spending Deficits And Debt	Monthly	Index
Fred	emv_policy	EMVWELFARE	Equity Market Volatility Tracker: Entitlement And Welfare Programs	Monthly	Index
Fred	emv_policy	EMVMONETARYPOL	Equity Market Volatility Tracker: Monetary Policy	Monthly	Index

Fred	energy_commodities	WTISPLC	Spot Crude Oil Price: West Texas Intermediate (WTI)	Monthly	Dollars per Barrel
Fred	energy_commodities	APU000074714	Average Price: Gasoline, Unleaded Regular (Cost per Gallon/3.785 Liters) in U.S. City Average	Monthly	U.S. Dollars
Fred	energy_commodities	APU00007471A	Average Price: Gasoline, All Types (Cost per Gallon/3.785 Liters) in U.S. City Average	Monthly	U.S. Dollars
Fred	energy_commodities	APU000072511	Average Price: Fuel Oil #2 per Gallon (3.785 Liters) in U.S. City Average	Monthly	U.S. Dollars
Datastream	energy_commodities	GSENTOT	Standard and Poors Goldman Sachs Commodity Index(GSCI) Energy Total Return	Daily	USD per Points
Fred	global_stock_markets	SPASTT01FRM657N	Share Prices: All Shares/Broad: Total for France	Monthly	Growth rate previous period
Fred	global_stock_markets	SPASTT01USM657N	Share Prices: All Shares/Broad: Total for United States	Monthly	Growth rate previous period
Fred	global_stock_markets	SPASTT01GBM657N	Share Prices: All Shares/Broad: Total for United Kingdom	Monthly	Growth rate previous period
Fred	global_stock_markets	SPASTT01JPM657N	Share Prices: All Shares/Broad: Total for Japan	Monthly	Growth rate previous period
Fred	global_stock_markets	SPASTT01DEM657N	Share Prices: All Shares/Broad: Total for Germany	Monthly	Growth rate previous period
Datastream	global_stock_markets	MSWRLD$	MSCI WORLD	Daily	Total Return Index (USD)
Datastream	industrial_metals	LCPCASH	LME-Copper Grade A Cash U$/MT	Daily	Metric Tonne
Datastream	industrial_metals	LAHCASH	LME-Aluminium 99.7% Cash U$/MT	Daily	Metric Tonne
Datastream	industrial_metals	LTICASH	LME-Tin 99.85% Cash U$/MT	Daily	Metric Tonne
Datastream	industrial_metals	LZZCASH	LME-SHG Zinc 99.995% Cash U$/MT	Daily	Metric Tonne

Fred	inflation	CPIAUCSL	Consumer Price Index for All Urban Consumers: All Items in U.S. City Average	Monthly	Index 1982-1984=100
Fred	inflation	CCRETT01USM661N	Financial Market: Real Effective Exchange Rates: CPI Based for United States	Monthly	Index 2015=100
Fred	inflation	APU000072610	Average Price: Electricity per Kilowatt-Hour in U.S. City Average	Monthly	U.S. Dollars
Fred	inflation	APU000072620	Average Price: Utility (Piped) Gas per Therm in U.S. City Average	Monthly	U.S. Dollars
Fred	inflation	IR10000	Import Price Index (End Use): Crude Oil	Monthly	Index 2000=100
Fred	long_rates	DGS3	Market Yield on U.S. Treasury Securities at 3-Year Constant Maturity, Quoted on an Investment Basis	Daily	Percent
Fred	long_rates	DGS5	Market Yield on U.S. Treasury Securities at 5-Year Constant Maturity, Quoted on an Investment Basis	Daily	Percent
Fred	long_rates	DGS10	Market Yield on U.S. Treasury Securities at 10-Year Constant Maturity, Quoted on an Investment Basis	Daily	Percent
Fred	long_rates	DGS20	Market Yield on U.S. Treasury Securities at 20-Year Constant Maturity, Quoted on an Investment Basis	Daily	Percent
Fred	long_rates	DGS7	Market Yield on U.S. Treasury Securities at 7-Year Constant Maturity, Quoted on an Investment Basis	Daily	Percent
Datastream	long_rates	TRUS10T	United States Government Eval Benchmark Bid Yield 10 Years	Daily	USD
Datastream	long_rates	TRUS10C	United States Government Benchmark CM Bid Yield 10 Years	Daily	USD
Fred	macro_gdp	BORROW	Total Borrowings from the Federal Reserve	Monthly	Millions of Dollars
Fred	macro_gdp	PINCOME	Personal Income	Quarterly	Billions of Dollars
Fred	macro_gdp	GNP	Gross National Product	Quarterly	Billions of Dollars
Fred	macro_gdp	GDPC1	Real Gross Domestic Product	Quarterly	Billions of Chained 2017 Dollars
Fred	macro_gdp	GDP	Gross Domestic Product	Quarterly	Billions of Dollars
Fred	macro_gdp	BBKMGDP	Brave-Butters-Kelley Real Gross Domestic Product	Monthly	Annualized Percent Change

					from Preceding Period
Fred	money_supply	CURRSL	Currency Component of M1	Monthly	Billions of Dollars
Fred	money_supply	M1SL	M1	Monthly	Billions of Dollars
Fred	money_supply	M2SL	M2	Monthly	Billions of Dollars
Fred	money_supply	DEMDEPSL	Demand Deposits	Monthly	Billions of Dollars
Fred	money_supply	TOTRESNS	Reserves of Depository Institutions: Total	Monthly	Billions of Dollars
Fred	money_supply	NONBORRES	Reserves of Depository Institutions: Nonborrowed	Monthly	Billions of Dollars
SDC Platinum	pe_capital	Estimated Equity Available (USD, Millions)	The estimated equity available for investment. It is derived from subtracting the fund's known equity invested from the fund size.	Quarterly	United States Dollars (millions)
SDC Platinum	pe_capital	No. of Firms	The number of firms.	Quarterly	Number of firms
SDC Platinum	pe_capital	No. of Funds	The number of funds managed by a firm.	Quarterly	Number of funds
SDC Platinum	pe_capital	No. of Companies Invested In	Number of companies invested in.	Quarterly	Number of companies
SDC Platinum	pe_capital	Avg Fund Size (USD, Millions)	Average Fund size per quarter.	Quarterly	United States Dollars (millions)
SDC Platinum	pe_capital	Amount Raised in Range (USD, Millions)	Amount the fund raised during the quarter.	Quarterly	United States Dollars (millions)

Cambridge Assoicates	pe_capital	Total Capitalization	Total Capitalization GP + LP	Quarterly	United States Dollars (millions)
SDC Platinum	pe_exits	Number of Deals	Total number of exits.	Quarterly	Number of Exits
SDC Platinum	pe_exits	No. Of IPO Exits	Total number of IPO type exits.	Quarterly	Number of IPO Exits
SDC Platinum	pe_exits	Exit Duration	Indicates time in Years between company's first investment received post previous vs the fund's exit date.	Quarterly	Years
SDC Platinum	pe_exits	Related Equity: Proceeds Amount All Markets	Proceeds Amount Incl Overallotment Sold All Markets : Total proceeds amount for the entire transaction plus overallotment amount (or green shoe) sold. This figure represents all tranches of the transaction.	Quarterly	United States Dollars (millions)
Cambridge Assoicates	pe_returns_top	Pooled Return (LP) (%)	Pooled Return (LP) (%)	Quarterly	Percent
Cambridge Assoicates	pe_returns_top	Top 5% (LP) (%)	Top 5% Returns (LP) (%)	Quarterly	Percent
Cambridge Assoicates	pe_returns_top	Upper Quartile (LP) (%)	Upper Quartile (LP) (%)	Quarterly	Percent
Cambridge Assoicates	pe_returns_top	Equal Weighted (LP) (%)	Equal Weighted (LP) (%)	Quarterly	Percent
Cambridge Assoicates	pe_returns_top	Average (LP) (%)	Average (LP) (%)	Quarterly	Percent
Cambridge Assoicates	pe_returns_top	PME Index IRR (LP) (%)	PME	Quarterly	Percent IRR with mPME: Russell 2000® Index (1979 Q1) with 300 bps premium
Fred	precious_metals	IQ12260	Export Price Index (End Use): Nonmonetary Gold	Monthly	Index 2000=100
Datastream	precious_metals	GOLDHAR	Gold H&H Bullion Price	Daily	USD per Troy Ounce

Datastream	precious_metals	SILVERH	Silver H&H Bullion Price	Daily	USD per Troy Ounce
Datastream	precious_metals	GSPMTOT	Standard and Poors Goldman Sachs Commodity Index(GSCI) Precious Metal Total Return	Daily	USD per Points
Datastream	real_estate	USNPIRN.F	NCREIF PROPERTY INDEX: VALUE OF RETURN - NATIONAL	Quarterly	Price index (United States Dollars)
Datastream	real_estate	USNPIRH.F	NCREIF PROPERTY INDEX: VALUE OF RETURN - HOTEL	Quarterly	Price index (United States Dollars)
Datastream	real_estate	USNPIRI.F	NCREIF PROPERTY INDEX: VALUE OF RETURN - INDUSTRIAL	Quarterly	Price index (United States Dollars)
Datastream	real_estate	USNPIRO.F	NCREIF PROPERTY INDEX: VALUE OF RETURN - OFFICE	Quarterly	Price index (United States Dollars)
Datastream	real_estate	USNPIRR.F	NCREIF PROPERTY INDEX: VALUE OF RETURN - RETAIL	Quarterly	Price index (United States Dollars)
Fred	short_rates	DTB3	3-Month Treasury Bill Secondary Market Rate, Discount Basis	Daily	Percent
Fred	short_rates	FEDFUNDS	Federal Funds Effective Rate	Monthly	Percent
Fred	short_rates	DFF	Federal Funds Effective Rate	Daily, 7-Day	Percent
Fred	short_rates	DPRIME	Bank Prime Loan Rate	Daily	Percent
Fred	short_rates	DTB6	6-Month Treasury Bill Secondary Market Rate, Discount Basis	Daily	Percent
Fred	short_rates	DTB1YR	1-Year Treasury Bill Secondary Market Rate, Discount Basis	Daily	Percent
Fred	short_rates	DGS1	Market Yield on U.S. Treasury Securities at 1-Year Constant Maturity, Quoted on an Investment Basis	Daily	Percent
Fred	short_rates	DGS3MO	Market Yield on U.S. Treasury Securities at 3-Month Constant Maturity, Quoted on an Investment Basis	Daily	Percent

Fred	short_rates	DGS6MO	Market Yield on U.S. Treasury Securities at 6-Month Constant Maturity, Quoted on an Investment Basis	Daily	Percent
Datastream	us_stock_market	USSPDIVY	STANDARD AND POORS' 500 COMPOSITE - DIVIDEND YLD	Monthly	Percent
Datastream	us_stock_market	USSPRPER	STANDARD AND POORS' 500 COMPOSITE - CAPE RATIO P/E10	Monthly	PE Ratio
Datastream	us_stock_market	USSPREPS	STANDARD AND POORS' 500 COMPOSITE - 12M AS REPORTED EPS	Monthly	United States Dollars
Datastream	us_stock_market	USSPMDPS	STANDARD AND POORS' 500 COMPOSITE - 12M CASH DIVIDEND PER SHARE	Monthly	United States Dollars
CRSP	us_stock_market	spindx	SPINDX is the level of the Standard & Poor's 500 Composite Index (prior to March 1957, 90-stock index) at the end of the trading day or month.	Monthly	Index Level (Points)
CRSP	us_stock_market	NaN	Calculated from spindx quarterly aggregated	Quarterly	Percent
Jay R. Ritter	us_stock_market	NaN	Professor Jay R. Ritter IPO Data, IPO Count aggregated Quarterly	Quarterly	Number of IPOs
Fred	yield_spreads	T10YFFM	10-Year Treasury Constant Maturity Minus Federal Funds Rate	Monthly	Percent
Fred	yield_spreads	T1YFFM	1-Year Treasury Constant Maturity Minus Federal Funds Rate	Monthly	Percent
Fred	yield_spreads	T5YFFM	5-Year Treasury Constant Maturity Minus Federal Funds Rate	Monthly	Percent
Fred	yield_spreads	TB3SMFFM	3-Month Treasury Bill Minus Federal Funds Rate	Monthly	Percent
Fred	yield_spreads	TB6SMFFM	6-Month Treasury Bill Minus Federal Funds Rate	Monthly	Percent
Fred	yield_spreads	T10Y2YM	10-Year Treasury Constant Maturity Minus 2-Year Treasury Constant Maturity	Monthly	Percent
Fred	yield_spreads	T10Y3MM	10-Year Treasury Constant Maturity Minus 3-Month Treasury Constant Maturity	Monthly	Percent

Note: Cambridge Associates and SDC Platinum series were accessed via Refinitiv Workspace, Datastream via the Workspace Excel add-in, FRED series via API, and Jay R. Ritter's IPO data via his website (excel).

References

Acharya, V. V., Franks, J., & Servaes, H. (2007). Private Equity: Boom and Bust? *Journal of Applied Corporate Finance*, *19*(4), 44–53. https://doi.org/10.1111/j.1745-6622.2007.00158.x

Akiba, T., Sano, S., Yanase, T., Ohta, T., & Koyama, M. (2019). Optuna: A Next-generation Hyperparameter Optimization Framework. *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 2623–2631. https://doi.org/10.1145/3292500.3330701

Almufti, A., Kahn, R. N., & Kazdin, J. (2024). Systematic Insights into Private Equity Investing. *The Journal of Portfolio Management*, *51*(1), 201–212. https://doi.org/10.3905/jpm.2024.51.1.201

Baker, S. R., Bloom, N., Davis, S. J., & Kost, K. J. (2019). Policy News and Stock Market Volatility. *NBER WORKING PAPER SERIES*.

Bargeron, L. L., Schlingemann, F. P., Stulz, R. M., & Zutter, C. J. (2008). Why do private acquirers pay so little compared to public acquirers? *Journal of Financial Economics*. https://doi.org/10.1016/j.jfineco.2007.11.005

BCG. (2024). *2024 M&A Report*. https://web-assets.bcg.com/fb/58/0e68c51c444e9f4691a1f2f7527c/m-and-a-report-oct-2024.pdf

Becketti, S. (1986). *Corporate mergers and the business cycle*. https://fedinprint.org/item/fedker/31094

Bonaime, A., Gulen, H., & Ion, M. (2018). Does policy uncertainty affect mergers and acquisitions? *Journal of Financial Economics*, *129*(3), 531–558. https://doi.org/10.1016/j.jfineco.2018.05.007

Brar, G., Giamouridis, D., & Liodakis, M. (2009). Predicting European Takeover Targets. *European Financial Management*, *15*(2), 430–450. https://doi.org/10.1111/j.1468-036X.2007.00423.x

Braun, R., Jenkinson, T., & Stoff, I. (2017). How persistent is private equity performance? Evidence from deal-level data. *Journal of Financial Economics*, *123*(2), 273–291.

Cao, J., & Lerner, J. (2008). The performance of reverse leveraged buyouts☆. *Journal of Financial Economics*, S0304405X08002018. https://doi.org/10.1016/j.jfineco.2008.03.002

Cohn, J. B., Hotchkiss, E. S., & Towery, E. M. (2022). Sources of Value Creation in Private Equity Buyouts of Private Firms. *Review of Finance*, *26*(2), 257–285. https://doi.org/10.1093/rof/rfac005

Da Rin, M., & Phalippou, L. (2017). The importance of size in private equity: Evidence from a survey of limited partners. *Journal of Financial Intermediation*, *31*, 64–76.

Davis, S. J., Haltiwanger, J., Handley, K., Jarmin, R., Lerner, J., & Miranda, J. (2014). Private equity, jobs, and productivity. *American Economic Review*, *104*(12), 3956–3990.

Degen, D., Kengelbach, D. J., Kim, D., Sievers, S., & Wang, Y. (2024). *Large Language Models and M&A: Can ChatGPT help forecast M&A activity?* (SSRN Scholarly Paper No. 4862121). Social Science Research Network. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4862121

Degeorge, F., Martin, J., & Phalippou, L. (2016). On secondary buyouts. *Journal of Financial Economics*, *120*(1), 124–145.

Doidge, C., Karolyi, G. A., & Stulz, R. M. (2017). The U.S. listing gap. *Journal of Financial Economics*, *123*(3), 464–487. https://doi.org/10.1016/j.jfineco.2016.12.002

DWF. (2020). *The Rise of Public to Private Transactions*. DWF. https://dwfgroup.com/en/news-and-insights/insights/2020/7/the-rise-of-public-to-private-transactions

Edlich, A., Croke, C., Dahlqvist, F., & Teichner, W. (2025). *Global Private Markets Report 2025: Private equity emerging from the fog*. McKinsey & Company.

Fernández Tamayo, B., Braun, R., Lopez-de-Silanes, F., Phalippou, L., & Sigrist, N. (2023). *Limited Partners versus Unlimited Machines; Artificial Intelligence and the Performance of Private Equity Funds* (SSRN Scholarly Paper No. 4490991). Social Science Research Network. https://doi.org/10.2139/ssrn.4490991

Fidrmuc, J. P., Palandri, A., Roosenboom, P., & van Dijk, D. (2013). When Do Managers Seek Private Equity Backing in Public-to-Private Transactions?*. *Review of Finance*, *17*(3), 1099–1139. https://doi.org/10.1093/rof/rfs021

Gao, X., Ritter, J. R., & Zhu, Z. (2013). Where Have All the IPOs Gone? | Journal of Financial and Quantitative Analysis. *Journal of Financial and Quantitative Analysis*. https://doi.org/10.1017/S0022109014000015

Goyal, A., Wahal, S., & Yavuz, M. D. (2021). Picking partners: Manager selection in private equity. *Swiss Finance Institute Research Paper*, *21–86*.

Guo, S., Hotchkiss, E. S., & Song, W. (2011). Do buyouts (still) create value? *The Journal of Finance*, *66*(2), 479–517.

Gupta, A., Howell, S. T., Yannelis, C., & Gupta, A. (2021). *Does Private Equity Investment in Healthcare Benefit Patients? Evidence from Nursing Homes* (SSRN Scholarly Paper No. 3785329). Social Science Research Network. https://doi.org/10.2139/ssrn.3785329

Haddad, V., Loualiche, E., & Plosser, M. (2017). Buyout Activity: The Impact of Aggregate Discount Rates. *The Journal of Finance*, *72*(1), 371–414. https://doi.org/10.1111/jofi.12464

Harford, J. (2005). What drives merger waves? *Journal of Financial Economics*, *77*(3), 529–560. https://doi.org/10.1016/j.jfineco.2004.05.004

Harris, R. S., Jenkinson, T., & Kaplan, S. N. (2014). Private equity performance: What do we know? *The Journal of Finance*, *69*(5), 1851–1882.

Jenkinson, T., & Sousa, M. (2015). What determines the exit decision for leveraged buyouts? *Journal of Banking & Finance*, *59*, 399–408. https://doi.org/10.1016/j.jbankfin.2015.06.007

Jensen, C., M. (1989). *Eclipse of the Public Corporation* (SSRN Scholarly Paper No. 146149). Social Science Research Network. https://doi.org/10.2139/ssrn.146149

Jensen, M. C. (1986). Agency Costs of Free Cash Flow, Corporate Finance, and Takeovers. *The American Economic Review*, *76*(2), 323–329.

Jensen, M. C., & Meckling, W. H. (1976). Theory of the firm: Managerial behavior, agency costs and ownership structure. *Journal of Financial Economics*, *3*(4), 305–360. https://doi.org/10.1016/0304-405X(76)90026-X

Kaggle. (2025). *Time Series*. Kaggle Learn. https://www.kaggle.com/learn/time-series

Kaplan, S. (1989). The effects of management buyouts on operating performance and value. *Journal of Financial Economics*, *24*(2), 217–254. https://doi.org/10.1016/0304-405X(89)90047-0

Kaplan, S. N., & Schoar, A. (2005). Private equity performance: Returns, persistence, and capital flows. *The Journal of Finance*, *60*(4), 1791–1823.

Kaplan, S. N., & Sensoy, B. A. (2015). Private equity performance: A survey. *Annual Review of Financial Economics*, *7*(1), 597–614.

Kaplan, S. N., & Strömberg, P. (2009). Leveraged buyouts and private equity. *Journal of Economic Perspectives*, *23*(1), 121–146.

Le Nadant, A.-L., & Perdreau, F. (2006). Financial Profile of Leveraged Buyout Targets: Some French Evidence. *Review of Accounting and Finance*, *5*, 370–392. https://doi.org/10.1108/14757700610712444

Lehn, K., & Poulsen, A. (1989). Free Cash Flow and Stockholder Gains in Going Private Transactions. *The Journal of Finance*, *44*(3), 771–787. https://doi.org/10.1111/j.1540-6261.1989.tb04390.x

Lerner, J., Sorensen, M., & Strömberg, P. (2011). Private equity and long-run investment: The case of innovation. *The Journal of Finance*, *66*(2), 445–477.

L'her, J.-F., Stoyanova, R., Shaw, K., Scott, W., & Lai, C. (2016). A Bottom-Up Approach to the Risk-Adjusted Performance of the Buyout Fund Market. *Financial Analysts Journal*, *72*, 36–48. https://doi.org/10.2469/faj.v72.n4.1

Martynova, M., & Renneboog, L. (2008). A century of corporate takeovers: What have we learned and where do we stand? *Journal of Banking & Finance*, *32*(10), 2148–2177. https://doi.org/10.1016/j.jbankfin.2007.12.038

McKinsey Global Private Markets Review 2024. (2024). McKinsey & Company. https://www.mckinsey.com/~/media/mckinsey/industries/private%20equity%20and%20principal%20investors/our%20insights/mckinseys%20private%20markets%20annual%20review/2024/mckinsey-global-private-markets-review-2024.pdf

Officer, M. S., Ozbas, O., & Sensoy, B. A. (2010). Club deals in leveraged buyouts. *Journal of Financial Economics*, *98*(2), 214–240. https://doi.org/10.1016/j.jfineco.2010.05.007

Ojea Ferreiro, J., Gregori, W., & Nardo, M. (2022). *Forecasting M&A deals with MIDAS count model*. European Commission.

Palepu, K. G. (1986). Predicting takeover targets: A methodological and empirical analysis. *Journal of Accounting and Economics*, *8*(1), 3–35. https://doi.org/10.1016/0165-4101(86)90008-X

Phalippou, L. (2017). *Private Equity Laid Bare*. Createspace Independent Publishing Platform.

Phalippou, L., & Gottschalg, O. (2009). The performance of private equity funds. *The Review of Financial Studies*, *22*(4), 1747–1776.

Phalippou, L., Rauch, C., & Umber, M. (2018). Private equity portfolio company fees. *Journal of Financial Economics*, *129*(3), 559–585.

Ritter, J. R. (2015). Growth Capital-Backed IPOs. *Financial Review*, *50*(4), 481–515. https://doi.org/10.1111/fire.12075

Schneider, J. (2022). *Private Equity Target Selection Using Artificial Intelligence* [CFA QuantAwards, Trinity College Dublin]. https://higherlogicdownload.s3.amazonaws.com/CFAI/35cafeb9-1a8a-4b92-b6e8-946c844e211d/UploadedImages/Quant_Awards_Winners_2022.pdf

Axelson, U., Jenkinson, T., Strömberg, P., & Weisbach, M. S. (2013). Borrow Cheap, Buy High? The Determinants of Leverage and Pricing in Buyouts. *Journal of Finance*, *68*(6), 2223–2267. https://doi.org/10.3386/w15952

Weir *, C., Laing, D., & Wright, M. (2005). Undervaluation, private information, agency costs and the decision to go private. *Applied Financial Economics*, *15*(13), 947–961. https://doi.org/10.1080/09603100500278221

Wharton Research Data. (2025). *M&A Databases and Research Guide*. WRDS. https://wrds-www.wharton.upenn.edu/pages/wrds-research/applications/programming-examples-and-other-topics/wrds-guide-mergers-acquisitions-m-databases-and-research/