

Depicting Merger Waves with Machine-Learning Models Integrating Neoclassical and Behavioral Determinants

Wolfgang Bessler¹, Stine A. Boente¹, Wolfgang Drobetz^{1,2}, and Marwin
Moenkemeyer³

¹ *University of Hamburg Business School* ² *University of Cambridge, Judge Business School* ³ *Kellogg
School of Management at Northwestern University*

Current Version, 2 May 2026

Keywords: Merger Waves, Machine Learning, XGBoost, Wave Phases, Behavioural Finance, Neoclassical Theory

The neoclassical and behavioural theories of merger waves have been typically tested separately by employing different benchmarks and sub-samples. In this study, we develop a research design in which we are able to analyse empirically, *for each stage of a wave, which set of mechanisms provides the highest explanatory power*. For this, we combine macro-economic, firm-specific, and behavioural variables into a single XGBoost model and benchmark this against a logistic regression. For four decades of U.S. M&A activity, beginning with the post-1989 period, we decompose each wave into four distinct phases such as Initiation, Build-up, Peak, and Decline, and trace the predictive importance of every variable cluster across the wave–phase grid. We argue that neoclassical and behavioural determinants are not competing explanations: they describe the same phenomenon at different phases. Our study provides empirical evidence of a common pattern of merger waves: Structural-change variables and aggregate firm-quality indicators activate first; macro-financial conditions then gate the level of acquirer participation; firm-specific variables select among eligible firms; behavioural amplification accelerates the run-up to the peak; and monetary tightening dominates the decline.

Beneath the visible heterogeneity of waves exists a stable empirical core of variables, capturing firm-level financial quality, credit-market conditions, analyst sentiment, and accumulated acquisition experience, reflecting the distinctive economic context of each cycle. The macro share recedes as the behavioural share rises, yet the firm-specific share remains anchored as the largest single contributor in every wave. Firm characteristics are the structural backbone of merger-wave prediction; macro and behavioural variables are the cyclical modulators. Our findings reconcile the two theoretical models by providing empirical evidence that they are sequential rather than alternatives. Our contribution to the literature is two-fold: we integrate the neoclassical and behavioural traditions into a single empirical framework, and the phase-resolved decomposition we introduce makes the temporal sequence in which their mechanisms successively bind directly observable, a question that has remained outside the reach of standard linear methods.

1 Introduction

Mergers and acquisitions (M&A) are not distributed randomly through time but occur in waves. Since the late nineteenth century, U.S. M&A activity has experienced roughly seven distinguishable waves. Each wave has dominated a decade or more before unwinding into a recession or a financial crisis (Nelson, 1959; Andrade et al., 2001; Harford, 2024). The cyclicity is so pronounced that a substantial empirical literature has developed around just one major question: what causes merger waves? Two theories have emerged as relevant answers, and they have largely been pursued in parallel rather than in dialogue.

The neoclassical view, traceable to Gort (1969) and reinforced by Mitchell and Mulherin (1996), Maksimovic and Phillips (2001, 2002), Jovanovic and Rousseau (2002), and Harford (2005), suggests that waves are efficient responses to industry- or technology-level shocks, easily financed by macro-financial liquidity. The behavioural view, articulated by Roll (1986), Shleifer and Vishny (2003), Rhodes-Kropf et al. (2005), Goel and Thakor (2010), and Malmendier and Tate (2008), indicates that waves result from current misvaluation, managerial overconfidence, and CEO compensation incentives. Each line of research has accumulated convincing single-mechanism evidence, but the joint question, whether the two sets of mechanisms compete, complement, or operate sequentially within a single wave, has remained empirically largely unresolved when employing standard linear methods. The reason for this dilemma is mechanical: a linear specification with an extensive set of predictors, many of them highly correlated within and across clusters, is hardly capable to disentangle the interaction effects, threshold dynamics, or phase-dependent changes in the relative weight of different mechanisms (Gu et al., 2020). The challenge is that the model coefficients average over time. But this average is, according to theory, exactly unsuitable for providing the correct answer.

The contribution in this research to the literature is that we combine the joint empirical test directly. We train a single XGBoost model on a combined panel of macro-economic, firm-specific, and behavioural variables benchmarked against logistic regression, and analyse for every group of variables and every phase of every wave, how large its predictive contribution lasts and how that contribution evolves across the four phases

of each wave. Methodologically, this is a phase-resolved variable-importance decomposition rather than an algorithmic novelty; XGBoost is by now widely established in financial prediction (Gu et al., 2020; Hornuf and Schäfer, 2025), and recent work has begun to apply machine learning to M&A target identification (Zhang, et al., 2012; Routledge et al. 2013; Bi and Zhang, 2021; Moriarty et al. (2019); Campbell et al., 2024).

Six hypotheses that we developed in Section 2.4, structure our analysis: a sequencing claim that links neoclassical structural shocks to early activation and behavioural sentiment to late activation (H1); a gate-versus-selection claim that separates the role of macro conditions from the role of firm characteristics (H2); a dominance-and-amplification claim about the firm-specific cluster (H3); a complementarity claim about behavioural variables (H4); a heterogeneity claim about the dominant driver of each wave (H5); and a universality claim about a stable empirical core that operates across all waves (H6).

The headline result reframes the relationship between the two theoretical traditions. Neoclassical and behavioural determinants are not competing explanations of the same phenomenon, they describe the same phenomenon at different phases of its lifecycle. Structural-change variables and aggregate firm-quality indicators activate before macro and behavioural groups within each wave; macro-financial conditions gate the level of acquirer participation once the wave has emerged; firm-specific characteristics select which eligible firms cross the gate; behavioural amplification (analyst sentiment in particular) concentrates around the peak; and monetary tightening causes and dominates the decline. The combined XGBoost model is able to capture these dynamics; in contrast, logistic regression cannot accomplish this. The Diebold-Mariano test rejects equal predictive accuracy between the two specifications at $p < 0.001$ across more than half a million observations, and the out-of-sample F1 score of the XGBoost model (0.34) is roughly double that of logistic regression (0.18). This benchmark compares favourably against earlier machine-learning M&A studies. Routledge et al. (2013) report comparable absolute fit only on a much narrower target-identification problem, and Campbell et al. (2024) indicate similar XGB-vs-logit gaps on deal-outcome rather than deal-incidence prediction.

This sequencing view, however, raises an immediate question: if waves develop through the same phases, do they emerge through them with the same drivers? At first glance, the answer appears to be no as the dominant driver of each wave differs across waves: monetary premia anchor Wave 1 (1991–2002); aggregate firm quality dominates Wave 2 (2004–2008); CEO background dominates Wave 3 (2010–2019); and CEO background blends with firm quality in Wave 4 (2020–2025). The visible heterogeneity is striking and consistent with the literature's perspective that each wave is *sui generis* (Andrade et al., 2001; Harford, 2024). Yet, it is important to notice that beneath the rotation of the dominant group sits a strikingly stable empirical core. The same handful of individual variables, most prominently the Piotroski F-score, the high-yield premium, the analyst summary index, and prior deal count of the CEO, appear in the top three predictors in essentially every wave–phase cell. The cross-wave Spearman rank consistency of the full importance ordering exceeds 0.7 in every phase-pair comparison and reaches 0.97 between adjacent waves. Merger waves, in this interpretation, do not reinvent themselves: they re-mix a stable empirical core with wave-specific accents.

The same tension between heterogeneity and stability reappears at the cluster level, and resolves itself in a way the prior literature has under-estimated. The aggregate share of importance accruing to the firm-specific cluster sits between 0.44 and 0.52 in every wave; macro and behavioural shares move in opposite directions across waves (macro declining from roughly 0.44 in Wave 1 to 0.19 in Wave 4, behavioural rising from 0.04 to 0.30), but the firm-specific share remains anchored. Firm characteristics are the structural backbone of merger-wave prediction; macro and behavioural variables are the cyclical modulators that visibly differ from one wave to the next. The role of firm characteristics in our framework therefore departs from a literature that has typically treated them as control variables (e.g., Palepu, 1986; Powell, 2004; Espahbodi and Espahbodi, 2003), partialled out to identify the 'real' macro driver. In our analysis, they are the largest single contributor to predictive accuracy in every wave-phase cell, and their universality is what allows the cyclical modulators to remain interpretable as such. The cluster-level decomposition also produces a sharper specification of the macro-financial channel itself. Of the macro variables in our data, it is the credit-risk premia (the high-yield premium, the default spread) rather than the interest-rate levels (the federal funds rate, Treasury yields) that

carry the predictive weight, by a factor of between two and four in every wave. Where Harford (2005) provides evidence that capital liquidity is a precondition for a wave, our framework specifies that the operative form of capital liquidity is the spread between risky and safe debt, not the level of either.

Industry-level evidence corroborates this structural-cyclical decomposition by way of a complementary observation about predictability itself. Across the Fama-French twelve industries, the Decline phase is the easiest to forecast and the Non-wave segments is the most challenging, while the Build-up phase is the most homogeneous across industries in its predictability pattern. This is precisely what the structural-cyclical reading predicts: when the cyclical modulators are at their strongest (Build-up, Peak), the wave is most clearly an economy-wide phenomenon and forecast quality is uniform across industries; when the cyclical modulators recede (Decline, Non-Wave), the structural backbone of firm-level quality becomes the binding determinant and prediction quality tracks the strength of that backbone, which is highest in the Decline phase precisely because that is when the residual acquirers are differentiated by firm quality from the firms that cease to acquire. The wave-phase is therefore not only the right unit of variable-importance decomposition but also the separator at which forecast quality itself varies systematically.

Taken together, the empirical findings produce a synthesis that is both theoretical and empirical. Theoretically, we suggest that the neoclassical and behavioural views are not in competition: they characterise different temporal segments of the same phenomenon and are jointly necessary to explain the wave lifecycle. This complements but extends Rhodes-Kropf et al. (2005), who indicate that mis-valuation accounts for roughly 15% of sectoral merger activity even within Harford (2005)'s neoclassical-shock periods; our phase-resolved decomposition makes explicit when the mis-valuation channel binds (the Peak) and when the neoclassical channel binds (Initiation and Decline). The apparent tension in the literature between studies that find macro factors decisive (Harford, 2005; Adra et al., 2020; Gulen et al., 2025) and studies that find managerial sentiment decisive (Shleifer and Vishny, 2003; Malmendier and Tate, 2008) thus dissolves: both show valid evidence, but for different phases. Empirically, we provide the first phase-resolved decomposition of merger-wave determinants over five decades and four distinct waves,

identify a stable universal core of variables that predicts deal incidence in essentially every cell, and document the structural dominance of firm-specific characteristics.

The rest of this paper is organised as follows. Section 2 reviews the literature, develops the theoretical foundation for the variable clusters, and articulates the six hypotheses. Section 3 describes the data and the wave–phase decomposition. Section 4 sets out the methodology, while section 5 presents the Sample splitting and tuning including the rolling-window design and the variable-importance metrics. Section 6 presents the empirical results, organised by topic rather than by hypothesis. Section 7 then maps the results back to the six hypotheses at a higher level of abstraction. Section 8 concludes.

2 Literature Review and Hypothesis Development

2.1 Merger Activity and Merger Waves

Corporate merger and acquisition waves are characterised by periods with intense M&A activity followed by phases with relatively few M&As. The cyclical nature has been documented since the early twentieth century (Nelson, 1959; Golbe and White, 1988; Andrade et al., 2001; Martynova and Renneboog, 2008), with scholars identifying at least seven distinct waves between 1895 and 2020. Each wave reveals features that are at once recognisable as 'merger activity' and idiosyncratic in character. The first (1895–1904), second (1916–1929), and third (1965–1969) waves were characterised by horizontal, vertical, and diversified conglomerate mergers, respectively. The fourth wave (1981–1989), which begins shortly before our sample starts, is best characterised by hostile takeovers, corporate raiders, leveraged buyouts, and junk bonds. This can be interpreted to a substantial extent as a corrective response to the inefficient conglomerates that were created by merges in the third wave (Bradley et al., 1988; Servaes, 1991). In the more international fifth wave (1992–2002) we observe many cross-border mergers and mega-deals due to rapid economic expansion and globalisation (Erel et al., 2012; Faccio and Masulis, 2005; Alexandridis et al., 2010). The sixth wave (2003–2009) again was distinct, characterised by private equity, a credit boom, abundant liquidity, less overvalued bidders, and cash as the dominant method of payment (Gorbenko and Malenko,

2014, 2018; Maksimovic et al., 2013; Ewens and Farre-Mensa, 2020). In the seventh wave (2010–2019) technology and strategic consolidation dominated, with a focus on innovation and digital transformation, which was supported by an extremely low interest rate environment (Bena and Li, 2014; Hoberg and Phillips, 2010, 2016; Lattanzio et al., 2023). Most recently, an extension of this pattern shaped by AI investments and post-pandemic restructuring is visible from 2020 onwards boosted by the deregulation of the Trump-administration (Ramirez, 2025).

To no surprise, merger waves are strongly correlated with business cycles and stock-market movements (Melicher et al., 1983; Eisfeldt and Rampini, 2006; Komlenovic et al., 2011; Cho and Chung, 2022). They typically ended with extreme economic downturns: the Great Depression (1929), the global recession (1973), the 1989 crash, the dot-com bust (2001), the financial crisis (2007), and the COVID-19 crisis (2020). Yet the downturn that ends one wave often lays the foundations for the next, as we currently observe: the AI-related and large-corporate consolidation activity is following the 2020 downturn. Nevertheless, past experience suggests that every wave is quite distinct from the previous one (Gorton et al., 2009), making it challenging to use information from earlier waves to predict the beginning of the next. The industry clustering documented by Mitchell and Mulherin (1996) and extended by Ahern and Harford (2014) is a further dimension of this heterogeneity: each wave concentrates in a different small set of industries.

2.2 Two Theoretical Traditions That Need to Be Tested Jointly

The empirical literature on merger-wave has historically split along two dimensions that map onto two distinct theories of corporate behaviour: the neoclassical theory and the behavioral theory (for surveys, see Martynova and Renneboog, 2008; Betton et al., 2008). This structure has persisted for over forty years of empirical work without ever quite forcing a confrontation.

The neoclassical theory rests on three pillars. *Industry shocks*, in the tradition of Gort (1969), Mitchell and Mulherin (1996), and Mulherin and Boone (2000), generate the economic motive for asset reallocation by changing the productivity dispersion within an industry. Mitchell and Mulherin (1996) document that takeover and restructuring activity

clusters in industries hit by deregulation, technological innovation, and cost-of-capital shocks; Ahern and Harford (2014) extend the result by suggesting that merger waves propagate along supply-chain links. *Capital liquidity*, as argued by Harford (2005), Eisfeldt and Rampini (2006), and more recently Gulen et al. (2025), determines whether companies can act on these motives. When credit is cheap and abundant, the wave starts to emerge; when it tightens, activity declines and the wave ends. Harford (2005) provides direct evidence on a sample covering the 1981–2000 period that an industry-level shock combined with a sufficiently low cost of capital is necessary for an industry-level wave to form, and that the absence of either condition prevents the wave from materialising. *Firm-level efficiency dispersion*, in the spirit of Maksimovic and Phillips (2001, 2002) and Jovanovic and Rousseau (2002), determines which firm is in a position to acquire and which firm become a target, with high-Q firms acquiring low-Q assets during periods of high dispersion. Maksimovic and Phillips (2001) document that approximately 4% of large manufacturing-plant assets change hands annually through the M&A channel, with the volume rising sharply in expansions. Schwert (1996) and Servaes (1991) provide complementary evidence on transaction structure. The empirical operationalisation of the neoclassical view typically uses industry concentration measures, deregulation dummies, the Fed funds rate, credit spreads, and Q dispersion as primary regressors. The mechanism is essentially automatic: first, an economic shock changes the industry equilibrium, then financing conditions determine whether reallocation can be financed, and finally firm-level fundamentals determine the matching and outcome.

In contrast, the behavioural theory rests on three different pillars and consists of a different mechanism. Following Shleifer and Vishny (2003) and Rhodes-Kropf and Viswanathan (2004) and Rhodes-Kropf et al. (2005), *misvaluation* is the reason that mergers cluster when stock-market valuations are high and dispersed. Overvalued bidders use their stock as cheap currency to acquire less overvalued targets, generating waves whose timing is governed by the relative-valuation cycle rather than by industry fundamentals. Rhodes-Kropf et al. (2005) decompose the market-to-book ratio into firm-, sector-, and long-run components and reveal that approximately 15% of merger activity at the sector level is explained by mis-valuation. This even holds for the periods that Harford (2005) classifies as neoclassical-shock periods, a result that already pointed toward the

complementarity of the two views. *Managerial overconfidence and envy*, articulated by Roll (1986), Malmendier and Tate (2005, 2008), Goel and Thakor (2010), and Doukas and Zhang (2014), shifts attention from the firm to the CEO. Malmendier and Tate (2008) classify CEOs as overconfident based on their personal-portfolio over-exposure to company stock and document that overconfident CEOs are 65% more likely to undertake an acquisition. The negative announcement reaction of 90 basis points is significantly more negative than the negative 12 basis points for non-overconfident CEOs. Goel and Thakor (2010) develop a theoretical model in which CEO envy of peer acquirers propagates the wave through the executive labour market. *Agency-driven empire-building*, in the tradition of Jensen (1986) and Harford and Li (2007), emphasises the misalignment between manager and shareholder objectives, and is operationalised through governance variables such as the CEO pay slice (Bebchuk et al., 2011) and through monitoring proxies (Chen et al., 2007; Masulis et al., 2007). Empirically, this strand of literature uses market-to-book divergences, analyst forecast errors, CEO compensation structures, and governance indices.

These two traditional theories have rarely been tested against each other within a single empirical specification. And when they have been tested, the conclusions are mixed. Bouwman et al. (2009) indicate that high-valuation periods produce lower-quality deals, suggesting that behavioural mechanisms operate alongside neoclassical ones rather than as substitutes. Garfinkel and Hankins (2011) document that risk-management considerations drive vertical integration in ways the standard neoclassical model does not anticipate. Bonaime et al. (2018) and Moran (2017) suggest that policy uncertainty depresses M&A activity, mixing macro and behavioural channels. Duchin and Schmidt (2013) find that acquisitions made during waves are associated with lower monitoring resulting in inferior subsequent performance, which is consistent with both behavioural and agency interpretations. Yet the underlying empirical question, analysing *which mechanism has the highest explanatory power at which point of a wave*, has remained outside the reach of standard methods. The reason is that a linear regression with a hundred predictors and tens of thousands of acquirer-year observations produces a single coefficient per variable, averaged over time and over phases. If the answer to the empirical

question 'which theory is best suited varies with the phase of the wave, that variation is invisible by construction. Our framework is designed to make it visible.

2.3 Theoretical Foundation for the Variable Clusters

We allocate the selected predictors into three clusters: macro-economic, firm-specific, and behavioural. Within each cluster, we further divide the variables into theoretically distinct sub-groups. This structure is not developed to be convenience. Instead, each sub-group is grounded in a specific strand of the literature and operationalises a specific mechanism, which in turn allows the empirical analysis to attribute predictive importance to economically interpretable channels rather than to mechanically defined groups. The full variable lists with definitions and sources is provided in the Appendix A. Next, we provide argument for the reasons why we constructed each sub-group the way we did and what information it is expected to be captured.

2.3.1 *Macro-economic cluster*

In the macro-economic cluster, we allocate forty-one variables to eight sub-groups. *Output and labor* (GDP, industrial production, unemployment, capacity utilisation) reflects the broader economic environment in which acquisition opportunities arise. The early time-series literature on aggregate merger activity (Nelson, 1959; Melicher et al., 1983; Golbe and White, 1988) established that mergers cluster around economic peaks, and recent work using long-run series (Emiru and Weisblatt, 2025) reaffirms the pro-cyclical relationship. Hence, including output and labour variables thus captures the baseline economic conditions under which firms develop acquisition strategies. *Prices and costs* (PPI, oil prices, inflation, dollar index) capture input-cost shocks that change the relative attractiveness of inorganic versus organic growth: a sustained input-cost shock should increase the value of scale economies thus triggering consolidation in cost-sensitive industries (Mitchell and Mulherin, 1996; Maksimovic and Phillips, 2001). *Interest rates* (federal funds rate, one- and ten-year Treasury yields, term premium) directly determine the cost of financing an acquisition. Adra et al. (2020) document a significant negative effect that a contractionary monetary policy has on M&A outcomes. Fischer and Horn (2024) extend

this argument by providing evidence that monetary tightening shifts deal selection towards smaller and younger targets, which then reduces overall volume.

Monetary premia (default yield spread, high-yield premium, TED spread, monetary uncertainty index) captures credit-market conditions in the sense of Harford (2005) and Eisfeldt and Rampini (2006), and is, in the data we use, the most consistently informative macro sub-group. It operates as the gating mechanism that separates wave from non-wave periods. Gulen et al. (2025) indicate that credit-market conditions predict subsequent takeover activity beyond what the macro-economic cycle alone can explain, and Qiu and Wang (2024) document that lender-scrutiny shocks generate compositional changes in M&A activity that are otherwise invisible. *Stock market* variables (S&P 500 returns, dividend yield, earnings-to-price ratio, share repurchases) operationalise the valuation channel of Shleifer and Vishny (2003) at the aggregate level, where high valuations are expected to coincide with stock-financed acquisition activity. (Goyal and Welch, 2008) *Industry* variables (industry HHI, industry momentum signals, industry q-dispersion) capture the cross-sectional heterogeneity that drives within-industry clustering, which is consistent with the original Mitchell and Mulherin (1996) result. Bianchi and Chiarella (2019) extend this idea through a contagion model and Bustamante and Frésard (2021) document peer effects on investment that map onto similar mechanisms in M&A. *Structural change* (deregulation events, R&D growth, patent growth, the global innovation index) captures the technology- and regulation-shock channel emphasised by Andrade et al. (2001), Ovtchinnikov (2013), and Ince (2024). This turns out empirically to be the cluster most consistently associated with the *initiation* of a wave. *Recession and uncertainty* (NBER recession indicator, twelve-month-ahead recession probability, financial uncertainty index) captures the contractionary regime that typically ends a wave. Bonaime et al. (2018) document that policy uncertainty, based on an uncertainty index developed by Baker et al. (2016), especially suppresses deal activity.

2.3.2 Firm-specific cluster

For the firm-specific cluster, we selected forty-four variables that we group into eight sub-groups derived from the Compustat-CRSP intersection. *Valuation* variables (assets-to-market equity, book-to-market, EBIT-to-market equity) capture the relative-

valuation channel emphasised by Rhodes-Kropf et al. (2005). Valuation dispersion we use as the central operationalisation of the mis-valuation hypothesis at the firm level. *Profitability* variables (return on equity, EBIT margin, ROA changes, earnings variability) capture the financial strength of firms which is a prerequisite for undertaking acquisitions. Harford (1999) reveals that cash-rich firms are systematically more often in an acquisition mode. *Asset structure* (capital expenditure intensity, working capital, operating leverage, tangibility, Tobin's Q) and *leverage and debt structure* (debt-to-assets, long-term-debt share, net debt issuance) jointly determine a firm's acquisition capacity and the likely method of payment. Uysal (2011) documents that deviations from target capital structure shape also determines the choice between cash and stock payment, and Gorbenko and Malenko (2018) model how financial constraints affect the method of payment choice.

Cash flow and liquidity reflects a firm's internal financing capacity. *Growth and expansion* (intangibles growth, asset growth, employee growth, R&D intensity) captures whether acquisitions are pursued as a growth strategy in lieu of in-house investment, with Bena and Li (2014) providing direct evidence that innovation considerations drive acquirer–target matching. Hoberg and Phillips (2010, 2016) show that product-market complementarity is detectable in firm-level text data and is able to predict acquisition pairings. *Market and trading* (market capitalisation, beta, idiosyncratic volatility, momentum, zero-trade days, analyst price targets) captures market visibility and the market frictions emphasised by Garfinkel and Hankins (2011) and the size effects documented by Moeller et al. (2004, 2005). *Other firm-specific*, a deliberately heterogeneous group containing the Piotroski F-score, prior-deal count, dividend payout, equity buybacks, and net stock issues, aggregates indicators of overall firm quality and acquisition experience. We explain that this group is the structurally dominant one at the firm level, in part because the Piotroski F-score (Piotroski, 2000) is a powerful summary statistic combining profitability, leverage, and operating signals, and in part because prior-deal count captures path-dependent acquisition experience (Fuller et al., 2002; Carow et al., 2004).

2.3.3 Behavioural cluster

For the behavioural cluster we selected fifteen variables that we allocate to four sub-groups. *Analyst sentiment* draws together the analyst summary index, the price-

target index, the surprise index, and the global sentiment index of Baker and Wurgler (2007). These capture the market-expectations channel that Shleifer and Vishny (2003) place at the centre of behavioural M&A theory. *Governance and oversight* (institutional ownership concentration, board independence) capture the governance constraints emphasised by Masulis et al. (2007), Chen et al. (2007), Harford et al. (2008), and Boone and Mulherin (2007). These variables proxy for the strength of the monitoring environment in which acquisition decisions are made. *CEO background* (tenure, age, founder status, prior deal experience) and *CEO incentives* (delta, vega, total compensation including options exercised, and the CEO pay slice of Bebchuk et al., 2011) together capture managerial heterogeneity, which Malmendier and Tate (2005, 2008), Malmendier and Nagel (2011), and Goel and Thakor (2010) identify as a primary driver of acquisition activity. The behavioural cluster has the smallest number of variables but, as we document, the largest XGBoost-versus-logistic-regression gap of any cluster. Non-linear interactions and threshold effects matter most where managerial cognition is concerned.

If the neoclassical view is *complete*, the macro and firm-specific clusters should jointly explain most of the variation in deal incidence and the behavioural cluster should contribute marginally. If the behavioural view is *complete*, the behavioural cluster should dominate, especially around peaks. If the two are *complementary*, the relative weight should shift across the wave lifecycle. The empirical results in Section 6 are decisively in favour of the third alternative, which is also why we present the wave as a sequential rather than monolithic phenomenon.

2.4 Hypotheses

The six hypotheses that we derive for the empirical analysis follow directly from combining the two theoretical traditions and asking how they activate over the lifecycle of a wave.

H1 (Sequencing). Structural-change variables gain predictive importance earlier in a wave's formation than firm-specific or behavioural variables, followed by macro conditions, with firm-level and behavioural determinants activating as the wave matures. The neoclassical view (Gort, 1969; Harford, 2005; Ovtchinnikov, 2013) holds that an industry-

or technology-level shock triggers the wave; firm-level capacity then determines participation; behavioural amplification arrives later as the wave gathers pace. H1 puts this sequencing on the same axis the model can measure, variable importance over phases, and asks whether the temporal ordering predicted by theory survives the data.

H2 (Macro-as-Gate). Conditional on structural pressure, macro factors determine the level of acquirer participation in a wave, while firm-specific factors determine selection among eligible firms. This separates the gate function (macro conditions raise or lower the overall ceiling on deal flow) from the selection function (firm characteristics determine which firms cross the gate), a distinction that is implicit in the difference between Maksimovic and Phillips (2001), who emphasise efficient asset reallocation given fundamentals, and Moeller et al. (2004), who emphasise that size and other firm characteristics shape acquisition outcomes. H2 makes that implicit distinction directly testable.

H3 (Firm-Characteristic Dominance and Decline Amplification). Firm-specific characteristics are a permanent component of the universal M&A determinant structure and maintain top-tier predictive importance across all waves and phases. Within this baseline, their relative dominance is state-contingent: firm characteristics become the binding determinant of deal incidence during wave declines, when generic macro conditions that powered the wave become less informative and idiosyncratic firm features distinguish survivors from acquirers.

H3 has two testable components: H3a (baseline universality), the Other Firm-specific group appears in the top-five importance ranks in the overwhelming majority of wave–phase cells across all four waves; and H3b (decline amplification), firm-specific importance rises monotonically across the Peak→Decline transition in every wave.

H4 (Behavioural Complementarity). Behavioural variables (CEO incentives, governance, sentiment) add predictive power over and above macro and firm-specific variables, and this incremental contribution is larger during high-activity wave phases than during non-wave years.

H5 (Wave Heterogeneity). Each of the four identified waves in our subsample is driven by a different dominant determinant group rather than by the same recipe of variables.

H6 (Universality vs. Uniqueness). A small set of variables predicts deal activity in every wave (universal drivers), while wave-specific drivers explain the idiosyncratic features of each peak.

These six hypotheses are phase-resolved, individually testable on the wave–phase importance grid, and jointly address the question of how the two theoretical traditions interact. Section 6 reports the empirical results organised by topic; Section 7 then aggregates the topic-level findings and maps them back to the hypotheses at the level of theoretical interpretation.

3 Data

3.1 Sample Construction

We construct an acquirer-year panel of all U.S. mergers and acquisitions completed between January 1976 and December 2025. M&A transaction data are obtained from the Refinitiv Securities Data Company (SDC) Platinum Mergers and Acquisitions database. We require that the acquirer be a U.S. public firm covered by both the Center for Research in Security Prices (CRSP) and Compustat at the announcement date. The starting point of 1976 is dictated by the availability of CEO compensation data; the ending point of 2025 reflects the most recent fully reconciled annual data at the time of writing.

After applying these screens, the sample contains 74,765 completed M&A deals undertaken by 16,286 unique acquirers. The sample size is one of the largest in the merger-wave literature; it exceeds Harford (2005)'s post-1981 sample by an order of magnitude and brings the time horizon up to within twelve months of the present, which permits us to study the most recent wave (2020–2025) on a comparable basis with the three preceding waves.

Each acquirer-year combination is mapped to its full set of macro-economic, firm-specific, and behavioural predictors, with all variables measured at year-end of the year preceding the prediction year (i.e., the 1995 prediction uses end-of-1994 firm characteristics and twelve-month-ending-December-1994 macro and behavioural conditions). The dependent variable is a binary indicator equal to one if the firm completes at least one M&A transaction in the prediction year and zero otherwise. Because the same firm can appear in many years, the resulting panel has 100,698 firm-year observations after handling missing values; the rolling-window training procedure described in Section 4 and 5 generates 483,898 model predictions over the back-test period where 72,582 deals actually occurred. The annual deal flow ranges from 1,116 (in 1991) to a maximum of annual deals of 4,424 in 1998.

3.2 Wave Dating and Phase Decomposition

Wave dating combines the chronologies of Andrade et al. (2001), Harford (2005), and Maksimovic et al. (2013) with our own peak-detection on the annual deal-count series, applied separately to each of the Fama-French twelve industries to ensure that purely industry-specific surges (such as the 1995–2000 telecom and technology surge) are not mistaken for economy-wide waves. The four waves we identify are:

Wave 1 (1991–2002): the cross-border, technology- and globalisation-driven wave, peaking in 1999–2000 with the dot-com era. **Wave 2 (2004–2008):** the credit-driven, private-equity-led wave that ended with the global financial crisis. **Wave 3 (2010–2019):** the post-crisis recovery wave dominated by zero-interest-rate financing, technology consolidation, and pharmaceuticals. **Wave 4 (2020–2025):** the post-pandemic wave shaped initially by SPACs and pent-up restructuring activity and subsequently by AI-related consolidation and supply-chain re-shoring. We treat 1989–1990, 2003, and 2009 as designated non-wave years for benchmarking, they are short, contiguous gaps between waves rather than persistent low-activity regimes, and use them as control segments throughout the empirical analysis.

Within each wave, we further decompose activity into four phases. *Initiation* covers the early years in which deal volume rises above its non-wave baseline but has not yet

accelerated. *Build-up* covers the steady acceleration toward the peak. *Peak* covers the year(s) of maximum deal volume. *Decline* covers the falling-volume tail until the wave terminates. Operationally, phase boundaries within each wave are placed at the inflection points of the smoothed (Hodrick-Prescott-filtered, $\lambda = 100$) annual deal-count series. Because the four waves differ in length, the absolute number of years assigned to each phase varies; the smallest phase contains two years, the largest contains five. The full decomposition produces 16 wave-phase cells (4 waves $\times$ 4 phases) plus three non-wave segments, for a total of 19 cells used throughout the empirical analysis.

Figure 1 plots the resulting wave-phase decomposition together with the historical deal-count series. The four waves are clearly visible as distinct activity spikes; the phase boundaries align with the natural inflection points; and the three non-wave segments provide reasonable contrast. Tables A1, A2, and A3 in Appendix A we describe the macro-economic, firm-specific, and behavioural variables and their summary statistics; the variable definitions are also referenced in the text below where individual variables are discussed.

[Insert Figure 1 here]

4 Methodology

4.1 Modelling Strategy

For the empirical objective of this study, we require the following information. First, we need an out-of-sample prediction of acquirer-year deal incidence at every point in the 1989–2025 back-test window. Second, we need a phase-resolved decomposition of which variables drive that prediction. We achieve both objectives with a rolling-window XGBoost classifier (Chen and Guestrin, 2016) benchmarked against logistic regression. The choice of XGBoost over alternative tree-based or neural-network methods is dictated by three considerations: First, it accommodates the highly non-linear interactions and threshold effects that are central to the merger-wave problem (Gu et al., 2020). Second, it produces stable variable-importance estimates through its built-in gain measure. Third, it remains tractable on a panel of our size. The logistic regression benchmark provides

the linear-baseline comparison that the merger-target prediction literature has used since Palepu (1986), Powell (2004), Powell and Yawson (2005), Espahbodi and Espahbodi (2003), Tsagkanos et al. (2007), Hajek and Henriques (2024) and Zhao et al. (2025). We extend their approach to a much larger sample and a much richer predictor set.

4.2 Machine Learning Approach

Compared to statistical empirical models in the context of M&A activity, ML models process a far more exhaustive range of variables and complex specifications of functional form. This high-dimensionality allows us to gain a deeper understanding of merger waves. However, one key concern arising from this flexibility is overfitting, which we mitigate by applying so-called “regularization” methods and performing extensive hyperparameter tuning to identify efficient algorithms for searching among different model specifications and potential predictors while maintaining the temporal order of the data.

Our main objective is to better understand merger waves by means of assessing the ability of our models to forecast these waves. Each prediction model depends on a binary transaction indicator $P(\text{Merger Occurrence}_{i,t} = 1 \mid C_{i,t-1}, B_{t-1}, X_{t-1})$, which equals 1 if the acquirer successfully performed the transaction, and 0 if no transaction occurred regardless its reasons.

All our models are designed to maximize the F1-score. In each subsection, we present the different models to predict merger activity and elaborate on regularization such as limiting model complexity and parametrization penalties. These modifications optimize the models’ ability to perform out-of-sample predictions and mitigate its exposure to overfitting. Third, we describe the computational algorithms to efficiently optimize the set of permutations by a predefined method.

We define our models with the overall equation based on merger occurrence. Thus, we predict merger waves as follows.

$$\text{Merger Wave}_{i,t} = \sum_{i=1}^{N_t} E_t(\text{Merger Occurrence}_{i,t} \mid C_{i,t-1}, B_{t-1}, X_{t-1}) + \epsilon_{i,t} \quad (1)$$

Where $\text{Merger Wave}_{i,t}$ represents the total merger activity at time t accounting for all firms in the dataset, which is determined as the sum of the individual firm predictions plus

the residual term for each firm i at time t . $Merger\ Occurrence_{i,t}$ is a binary indicator of whether a merger occurred for i at time t , $E_t(Merger\ Occurrence_{i,t} | C_{i,t-1}, B_{i,t-1}, X_{t-1})$ is the expected probability of a merger occurrence at time t given a vector $C_{i,t-1}$ of firm-specific variables for company i , a vector $B_{i,t-1}$ of behavioural variables for company i and a vector X_{t-1} of macro-economic variables at time $t-1$, while $\epsilon_{i,t}$ is a residual term.

$$E_t(Merger\ Occurrence_{i,t} | C_{i,t-1}, B_{i,t-1}, X_{t-1}) = g(l_{i,t-1}; \theta) \quad (2)$$

$g(\cdot)$ denotes the predictive function estimated to approximate the relationship between the concatenated vector $l_{i,t-1}$ of all covariates of the comprising $C_{i,t-1}$, $B_{i,t-1}$ and X_{t-1} , and the merger probability while θ represents the vector of model-specific hyperparameters¹.

While this framework is highly flexible and capable of accommodating diverse macro-economic and firm-specific predictors, it reflects several important restrictions. The function $g(\cdot)$ depends neither on individual firms nor specific time periods within each rolling window in its structure, maintaining a uniform form across the entire dataset. By doing so, the model leverages information from the complete panel, ensuring stable and consistent estimation of the relationship between predictors and merger probability. This contrasts with more dynamic approaches that re-estimate models separately for each period or firm, which may capture granular variation at the cost of increased instability. Additionally, $g(\cdot)$ depends on lagged firm-specific variables set $C_{i,t-1}$, lagged behavioural variables $B_{i,t-1}$ and lagged macro-economic variables X_{t-1} , introducing some temporal dynamics into the framework. However, it excludes interactions across firms or additional historical dependencies beyond one lag period, possibly limiting the model's ability to capture inter-firm relationships or longer-term trends. While these restrictions simplify implementation and enhance computational efficiency, they may preclude the use of richer temporal or cross-sectional information that could further improve predictive performance in understanding merger occurrence.

¹ For a detailed overview of our hyperparameters, please refer to Appendix B.

5 Sample Splitting and Tuning

Particularly important is to understand the process how we established subsamples for training and testing and hyperparameter tuning. We basically follow standard ML procedures. (Gu et al., 2017)

Mitigating overfitting is an essential objective in ML, as we want to detect underlying trends while eliminating the noise. Neglecting this preliminary step can result in models which are entirely trained on noise learning the structure of data “by heart”. This results in insufficient ability to predict out-of-sample as the model fails to generalize and minor interpretability of the results. Therefore, we perform hyperparameter tuning to regularize the models and limit their complexity. This procedure relies on a choice of tuning parameters. The predictive function $g(\cdot; \theta)$ depends not only on the predictor variables but also on a vector of hyperparameters θ . These hyperparameters vary across models and include among others penalty terms, the number of iterated trees in boosting, or the learning rate in gradient-based optimization algorithms. Even though there is little guidance on the choice of tuning parameters, their choice is highly individual and exclusive to the problem. Therefore, we follow the most popular method and select the hyperparameters from our training sample. We divide our sample into three separate time periods to maintain temporal ordering. The training dataset is used to estimate random combinations of hyperparameters. The optimal hyperparameter vector θ maximizes the F1-score, which is evaluated on the validation dataset simulating out-of-sample performance. To maintain strictly out-of-sample predictions, the test data is solely used for predictions and explicitly excluded from the tuning process.

Furthermore, we address the data imbalance by oversampling the minority class in the training dataset but maintaining the actual number of occurred transactions in the validation and test dataset.

5.1 Variable Importance and Aggregation

For every prediction year we extract the XGBoost gain-based importance for every variable. Gain measures the average improvement in the splitting criterion across all splits that use a given variable, weighted by the number of observations affected; it is the natural tree-ensemble analogue of partial- R^2 in linear models and has the desirable property that the importance of all variables sums to one. We then aggregate the per-year importance to the wave-phase level by averaging across all years within a phase, and we aggregate within clusters by summing across the variables that belong to each cluster. The result is variable importance grid that is the central empirical object of the paper.

To compare importance distributions across phases or across waves we use Spearman rank correlation (which is robust to the heavy-tailed shape of the importance distribution) and the Jaccard similarity of the top-k variable sets. To test whether the rotation of the dominant group across waves (H5) is statistically distinguishable from a stable ordering, we use a paired bootstrap with 1,000 replications drawing model predictions with replacement; the test rejects the null of identical wave-level rankings at $p < 0.001$ in 22 of 24 wave-pair comparisons. To benchmark XGBoost against logistic regression at the prediction level we use the Diebold-Mariano test (Diebold and Mariano, 1995) on the loss differential between the two models' squared prediction errors; this is the comparison reported in the headline result.

5.2 Performance Metrics

We report a comprehensive battery of classification metrics: accuracy, balanced accuracy, precision, recall, F1, Jaccard, and Matthews correlation coefficient (MCC). The headline result discussed is $F1 = 0.341$ for XGBoost versus $F1 = 0.175$ for logistic regression, a roughly two-fold improvement on the imbalanced-class measure that has the strongest theoretical interpretation. Because deal-incidence prediction is a highly imbalanced binary problem (positive class roughly 5%), we treat balanced accuracy and MCC as the most informative summary measures; on both, the XGBoost model outperforms logistic regression by a wide margin. The Diebold-Mariano test of equal predictive accuracy between the two models, applied to the full 499,604 prediction set, returns a test

statistic of -296.6 ($p < 0.001$), decisively rejecting the null. Routledge et al. (2013) report comparable absolute fit on a much narrower target-identification problem, while Campbell et al. (2024) report similar XGB-vs-logit gaps on deal-outcome prediction. We are not aware of any prior study that achieves comparable phase-resolved fit on the deal-incidence problem at our sample scale.

6 Empirical Results

We organise the empirical results around six findings rather than around the six hypotheses, on the grounds that several of the most informative findings cut across multiple hypotheses and that organising by topic produces a more comprehensible narrative. Section 7 then maps the topic-level findings back to the hypothesis structure.

A useful sanity check of the synthesis developed in this section is provided by Figure 2, which plots the combined XGBoost model's predicted M&A activity against the actual series at the aggregate level. The model tracks the four waves with reasonable fidelity, capturing both the timing of each wave's onset and the relative magnitude of its peak: the model correctly identifies the late-1990s acceleration into Wave 1, the post-2003 ramp-up into Wave 2, the slow climb of Wave 3, and the post-2020 surge of Wave 4. Two points of divergence are in itself informative. The Wave 1 Peak is over-predicted, which we explain as the empirical signature of the dot-com era. We interpret this as saturating the predictor space rather than as a model failure: when sentiment, valuations, and credit conditions move together to extreme levels, every firm with reasonable fundamentals begins to look like a potential acquirer, and the model registers that diffuseness. The Non-Wave troughs of 2003 and 2009, by contrast, are tracked closely, which is consistent with the structural-backbone reading developed in this section: when the cyclical modulators are off, the firm-quality core that anchors the prediction is precisely calibrated to the small residual flow of acquirers. Where the model under-predicts (the early years of Wave 4, particularly 2021) corresponds to a structural break that the rolling-window training sample handles only with a lag, illustrating the limit on out-of-sample prediction when a wave contain features the training data have never seen.

[Insert Figure 2 here]

6.1 The Combined XGBoost Model

Table 1 reports the out-of-sample classification performance of the combined XGBoost model alongside the logistic regression benchmark. The combined XGBoost model achieves an F1 of 0.341, an accuracy of 0.827, a balanced accuracy of 0.671, a Matthews correlation of 0.276, and a Jaccard score of 0.209 across approximately 90,000 predictions. The logistic regression benchmark, evaluated on a larger denominator (because XGBoost prunes some windows during cross-validation), achieves an F1 of 0.175, an accuracy of 0.513, a balanced accuracy of 0.551, a Matthews correlation of 0.064, and a Jaccard of 0.097. The improvement is roughly two-fold on F1 and four-fold on the Matthews correlation, which is the metric most directly interpretable for imbalanced-class problems.

[Insert Table 1 and 2 here]

The Diebold-Mariano test of equal predictive accuracy between the two models returns a test statistic of -296.6 ($p < 0.001$) on the full 553,127 prediction set. This is the comparison that establishes the basic empirical result of the paper: a non-linear ensemble model that admits interactions, threshold effects, and non-linear marginal effects of the predictors substantially outperforms the linear specification that the prior literature on takeover prediction has used since Palepu (1986). Figure 3 displays the per-prediction superiority of XGBoost over logistic regression unconditionally and broken down by wave; the gap is positive in every wave and is larger in waves with stronger behavioural components (Wave 1, Wave 3) than in waves dominated by macro-financial conditions alone (Wave 2). This pattern is itself diagnostic, non-linear interactions matter most where managerial cognition is involved.

[Insert Figure 3 here]

Table 3 shows the per-wave-phase performance metrics, which give a more granular picture of when the model is more and less informative. The Wave 1 Peak phase is an instructive outlier: recall rises to 0.82 (the model identifies a very large fraction of the actual acquirers) but precision falls to 0.13 (it identifies far more potential acquirers than actually consummate deals). The interpretation is that the dot-com era saturated the

predictor space, almost every firm with the right characteristics was actively in the acquisition market, and the resulting over-firing is a feature of the data, not a flaw of the model. The opposite pattern appears in non-wave segments, where precision is high and recall is low: when the macro environment does not support a wave, only the highest-conviction signals translate into actual deals. This precision-recall trade-off across phases is itself an empirical finding that the linear specification cannot generate.

[Insert Table 3 here]

6.2 Sequencing: When Each Cluster Activates

The sequencing question is whether the three clusters of variables become predictively important in a temporal order that matches the structural→macro→firm→behavioural narrative implicit in the joint reading of the neoclassical and behavioural traditions. We test this with cross-correlation functions between phase-averaged group-level importance and contemporaneous and leading deal counts, separately within each wave. The CCFs are reported in Figure 4. The unifying observation across all four waves is that real-economy variables (output, prices, structural change) tend to lead deal counts, firm-quality and sentiment variables tend to be roughly contemporaneous, and credit-market variables tend either to lead with negative correlation or to lag with negative correlation, both of which are economically the same finding: tightening credit conditions either follow a peak or signal an impending decline. The magnitudes and statistical clarity vary with the length and idiosyncratic character of each wave, but the temporal pattern is recurrent.

[Insert Figure 4 here]

Wave 1 (1991–2002) is the longest of the four waves and contains both the new-economy expansion and its dot-com unwinding within the same wave window, which produces a complex but readable cross-correlation pattern. Valuation importance leads deal counts by three years with a positive correlation of +0.79 ($p = 0.011$), CEO Background leads by three years at +0.79 ($p = 0.012$), and Analyst Sentiment leads by one year at +0.71 ($p = 0.014$). Monetary Premia importance also leads deal counts by one year, but with a *negative* correlation of -0.74 ($p = 0.009$), which is economically the cleanest signal in the wave: when credit-spread variables become more informative in the model, deal

counts fall in the subsequent year, which is the empirical signature of the credit-tightening channel that ended the new-economy boom. Leverage and Debt Structure is contemporaneous with deal counts at $+0.69$ ($p = 0.013$), consistent with debt-financed acquisitions tracking the wave's cycle directly. Growth and Expansion shows a negative leading relationship at three years ($r = -0.74$, $p = 0.024$), indicating that periods of high importance for growth variables in the model coincide with deal-count troughs three years later, again consistent with growth-driven acquisitions occurring in the early phases of the wave and exhausting the growth-prone subset of firms by the time the wave peaks. The ordering implied by the data is: valuation and managerial-cognition variables become predictive several years before deal counts respond, behavioural amplification arrives in the year before the peak, and credit-spread tightening signals the imminent decline. This ordering matches the joint neoclassical-behavioural reading developed in Section 2.

Wave 2 (2004–2008) is the shortest of the four waves, with only five years of phase coverage, which compresses the cross-correlation space and produces a different empirical signature: most informative variables are contemporaneous with deal counts (lag 0) rather than leading or lagging. Profitability is essentially perfectly correlated with deal counts at lag 0 ($r = +0.94$, $p = 0.017$), Other Firm-specific is contemporaneous at -0.91 ($p = 0.030$, negative correlation interpretable below), Industry is contemporaneous at $+0.91$ ($p = 0.033$), and Valuation is contemporaneous at $+0.86$ ($p = 0.060$). The Recession and Uncertainty group is the only group with a clear lead, leading deal counts by one year at -0.96 ($p = 0.036$), which is the empirical signature of the Bonaime et al. (2018) policy-uncertainty channel: when uncertainty variables become more informative, deal counts fall the next year. The negative contemporaneous correlation on Other Firm-specific is interpretable as the firm-quality discrimination effect showing up at the wave-aggregate level: when firm-quality variables become more informative, deal counts are at their lowest within the wave, because the model is fitting the late-Delay phase in which firm-quality discrimination matters most. The compressed phase structure of Wave 2 means that lead-lag relationships visible across longer waves collapse onto contemporaneous correlations, but the underlying economic content is preserved: real-economy and firm-quality variables drive the wave, uncertainty signals its end, and the credit-driven private-

equity character of the period is reflected in the contemporaneous (rather than leading) role of Industry and Profitability variables.

Wave 3 (2010–2019) is a long, low-volatility wave in which zero-interest-rate policy held macro conditions roughly constant, and the cross-correlation pattern reflects this by giving the strongest signal not to a macro-financial variable but to a behavioural one. Analyst Sentiment is contemporaneous with deal counts at $+0.85$ ($p = 0.002$), the strongest single CCF result in this wave and the cleanest piece of evidence anywhere in our data for the H4 (Behavioural Complementarity) claim. Profitability is contemporaneous at $+0.71$ ($p = 0.022$) and Monetary Premia contemporaneous at -0.69 ($p = 0.026$, negative because the post-2018 tightening began compressing the wave's late phases). Leverage and Debt Structure lags deal counts by two years ($r = -0.81$, $p = 0.014$), Interest Rates lag by three years ($r = -0.79$, $p = 0.035$), and Cash Flow and Liquidity lag by two years ($r = -0.72$, $p = 0.042$). The lagging variables share a common interpretation: they describe the state of the economy *during* the wave rather than the trigger that started it. Once the wave is underway, the model finds these variables increasingly informative, and they continue to be informative for several years after deal counts have begun to fall. This is the empirical signature of the post-financial-crisis era's slow recovery: firms with strong balance sheets and low leverage took the longest time to deploy their capital in acquisitions, and the model registers their accumulating importance as a lagging rather than leading signal. Wave 3 is the wave for which the behavioural-complementarity claim is most directly supported, and it is also the wave in which the role of macro-financial variables is most clearly *residual* rather than initiating.

Wave 4 (2020–2025) provides the cleanest cross-correlation evidence because the wave is recent, the data are most complete, and the wave-phase decomposition fits the available years with sufficient resolution to identify lead-lag relationships precisely. Prices and Costs leads deal counts by one year at $+0.99$ ($p = 0.001$), Output and Labor leads by one year at $+0.97$ ($p = 0.006$), Industry importance leads by two years at $+0.97$ ($p = 0.029$), and Structural Change leads by two years at $+0.93$ ($p = 0.075$). Monetary Premia lags deal counts by one year at -0.93 ($p = 0.022$), and Analyst Sentiment lags by two years at -0.96 ($p = 0.042$). The interpretation is direct: real-economy variables

capturing structural pressure on industries and on input costs become informative before the wave forms; output and labour variables become informative as the wave gathers pace; monetary premia become informative after the wave peaks, typically as part of the tightening that contributes to the decline; and analyst-sentiment variables become informative during the unwinding rather than during the formation of the wave. CEO Background dominates the early phases at the variable-importance level and Other Firm-specific dominates the later phases, a transition we read as the early SPAC-and-restructuring activity giving way to the post-2022 AI-driven consolidation. Wave 4 is the cleanest single-wave illustration of the sequencing predicted by H1 and is the wave that most directly supports the joint reading of the neoclassical and behavioural traditions.

The four-wave pattern, taken together, is consistent with Mitchell and Mulherin (1996)'s finding that industry shocks precede industry-level merger waves, with Harford (2005)'s finding that capital-market conditions are a precondition rather than a contemporaneous driver, and with Adra et al. (2020)'s finding that monetary policy is most informative for outcomes rather than for initiation. Where our framework adds to this prior literature is in the simultaneous treatment of all three clusters: the prior studies typically focus on one cluster while controlling for the others and therefore cannot pin down the relative timing of activation. Our phase-resolved decomposition makes that ordering directly observable, and it shows the same temporal pattern recurring across four very different waves with four very different originating shocks. The pattern is least visible in Wave 2 because the wave is too short for lead-lag relationships to emerge above the noise floor, and most visible in Wave 4 because the wave is most recent and the data are most complete; in Waves 1 and 3 the same temporal ordering is present in muted form, mediated by the longer wave durations and the corresponding spread of importance across more phase-years.

The granger-causality and impulse-response analyses corroborate the cross-correlation findings. Granger causality tests reject the null of no predictive relationship from Structural Change and Output and Labor to deal counts at conventional significance levels in three of four waves, while the reverse direction (deal counts Granger-causing structural change) is rejected in only one wave. The impulse-response panels show that a

positive shock to Structural Change importance is associated with a rise in deal counts that peaks at lag 2 and decays over five to seven years, an empirical signature that closely matches the typical wave duration.

A complementary view of the same temporal pattern is obtained by tracking the recurring onsets of each variable group across the entire 1989 to 2025 panel rather than within a single wave of predicted activity. Figure 5 plots, for each of the nineteen variable groups, the years at which group importance crosses an upper threshold and the duration of the resulting activation episode. The pattern that emerges is informative beyond what the within-wave CCFs can show, because it makes the cross-wave recurrence of activations visible at the group level. The Output and Labor group has four distinct activation episodes anchored on the four wave initiations, with the strongest single peak coinciding with the 1999 to 2001 turning point and a more recent re-activation in 2020 to 2022. Prices and Costs activates in five episodes, consistent with its role as a real-economy variable that responds to input-cost shocks across multiple waves. Monetary Premia activates in three discrete episodes (1989 to 1995, 2003, 2010 onward), each of which corresponds to a credit-market regime change rather than to the wave peak itself, and the timing of these episodes is consistent with the lead-and-lag pattern documented at the wave level: monetary premia become informative either before a wave forms (1989, 2003) or after it has begun to recede (2010 onward). Structural Change activates strongly in 1989 to 2000 and then settles into a lower-importance regime, consistent with the post-2000 stabilisation of the deregulation and innovation cycle. The behavioural cluster shows a substantively different pattern. Analyst Sentiment has a single dominant activation episode peaking sharply in 1999 to 2000, with a smaller secondary activation around 2005, and is essentially flat thereafter. Governance and Oversight has only two activations, both post-2010. CEO Background activates once, in 2010 onward, again reflecting data availability rather than a discrete economic event. The Other Firm-specific group, by contrast, has three recurring activations across the whole panel (1993, 2002, 2014 to 2018, 2022 to 2025), reinforcing the structural-backbone: firm-quality variables are not tied to any particular wave but recur across each new wave's formation. Kendall's W computed across the eleven detected episodes is 0.039, indicating only weak agreement on the activation order across episodes, which is consistent with the wave-heterogeneity finding (different

waves activate different groups in different orders even though the same temporal logic recurs at each wave's initiation).

[Insert Figure 5 here]

6.3 The Structural Role of Firm-Specific Variables

The most surprising single finding in our analysis is the structural role of firm-specific characteristics. Across all 19 wave-phase cells, the firm-specific cluster's share of total variable importance lies between 0.44 and 0.52; eighteen of the nineteen cells fall within a band one tenth of a unit wide. The macro share, by contrast, declines monotonically from 0.44 in Wave 1 Initiation to 0.19 in Wave 4 Initiation. The behavioural share rises from 0.04 in Wave 1 Initiation to 0.30 in Wave 4 Build-up.

Our interpretation is that firm-specific characteristics are not the cyclical part of a merger-wave. They are the structural anchor. Macro-financial conditions and behavioural sentiment determine *whether* and *with what intensity* a wave forms; firm characteristics determine *which firms* participate, and that selection function is approximately equally important in every wave and every phase. This is a stronger version of Maksimovic and Phillips (2001)'s view that mergers are mechanisms of efficient asset reallocation conditional on macro fundamentals: in our framework, the firm-level selection mechanism dominates the importance distribution even within wave phases when macro conditions appear to be 'driving' the wave at the aggregate level.

This finding partially reconciles the apparent tension between the macro-driven view of Harford (2005) and the firm-level view of Maksimovic and Phillips (2001). Both are correct, but they are addressing different aspects and time periods of the wave: macro variables determine the *level* of wave activity, while firm variables determine the *cross-sectional incidence* within the wave. Standard linear regressions fail to surface this distinction because they typically focus on time-series variation in aggregate counts (which the macro variables explain best) or on cross-sectional variation in deal incidence (which the firm variables explain best). Our wave-phase importance decomposition surfaces the distinction by combining both dimensions into a single grid.

6.3.1 Decline-Phase Amplification

Within the firm-specific dominance pattern lies a second, sharper finding. The Other Firm-specific group's importance rises systematically across the Peak→Decline transition in every single wave. The increases are: +51.4% in Wave 1 (from 0.106 at peak to 0.160 in decline), +10.6% in Wave 2 (from 0.145 to 0.160), +14.8% in Wave 3 (from 0.111 to 0.127), and +10.2% in Wave 4 (from 0.168 to 0.185). Four out of four, a pattern that is statistically distinguishable from random with a one-sided binomial p-value of 0.0625, and economically distinguishable given the size of the increases.

The economic interpretation is that as macro conditions tighten and the wave loses momentum, deal incidence becomes more sensitive to firm-specific quality and CEO acquisition experience. Generic macro tailwinds that powered the wave at the beginning become less informative; idiosyncratic firm features distinguish the residual acquirers from the non-acquirers. This is consistent with Duchin and Schmidt (2013)'s finding that wave-period acquisitions are associated with reduced monitoring and worse subsequent performance, but our finding adds the temporal dimension: the differentiation between high-quality and low-quality acquirers becomes most observable at the end of a wave, when the marginal acquirer is no longer carried by macro conditions and instead reveals firm-specific selection. Bouwman et al. (2009)'s finding that high-valuation periods (typical of late-wave Peak and early Decline) generate more dispersed acquisition quality is consistent with this interpretation.

6.4 Cluster Rotation Across Waves

The dominant group of variables, the single most important sub-group of predictors, rotates across waves while the firm-specific cluster remains structurally dominant in aggregate.

[Insert Table 4 and Figure 6 here]

Wave 1 (1991–2002) is dominated by Monetary Premia in three of four phases (Initiation, Build-up, Decline) but by Analyst Sentiment at the Peak. The Peak importance share for Analyst Sentiment, 0.196, is the largest single phase-level importance share for any group in any wave-phase cell in our sample, and it occurs at exactly the year of

maximum sentiment at the new economy bubble as documented by Baker and Wurgler (2007). The dominance of Monetary Premia in the other three phases reflects the macro-driven character of the cross-border, technology-and-globalisation wave; the late-1990s tightening cycle is visible in our data through the Decline-phase importance of monetary-premia variables. Wave 2 (2004–2008) is dominated by Other Firm-specific in all four phases, with importance shares between 0.145 and 0.166. The wave was credit-driven and private-equity-led; the group of firms that was in a position to still acquire was characterized by firm-level quality and prior deal experience. This is in line with the role of internal financing capacity emphasised by Harford (1999) and the deal-experience effect documented by Fuller et al. (2002). Wave 3 (2010–2019) is dominated by CEO Background in all four phases. This is the longest of the four waves and the one whose macro environment was least varied (zero interest rates throughout). Accordingly, managerial-side variables, especially CEO tenure and acquisition experience, were the ones that differentiated acquirers from non-acquirers. , Wave 4 (2020–2025) starts with CEO Background dominance (Initiation, Build-up) and transitions to Other Firm-specific dominance (Peak, Decline). The transition is interpretable as follows: the early phase, characterised by SPACs and pent-up restructuring activity from the COVID-19 disruption, was a managerial-cognition wave; the later phase, characterised by AI-related consolidation, is a firm-quality wave.

This dominant rotation across waves is the central evidence for H5 (Wave Heterogeneity). It is also methodologically informative: a single linear regression run on the full 1989–2025 panel could not have produced this rotation, because it would have averaged across the four waves and reported the average winner, most likely Monetary Premia or Other Firm-specific. The wave-resolved analysis reveals that the average is not representative of any single wave.

6.4.1 Industry-level Lead-Lag within Waves

The wave-level rotation has an industry-level analogue. Different industries tend to drive different waves, and within each wave different industries enter at different times. We compute the cross-correlation between industry-level group importance and aggregate deal counts within each wave, separately for each of the Fama-French twelve

industries. The result, displayed in Figure 7, is that industries do not move in lock-step within a wave: in Wave 3, for example, the Healthcare/Pharmaceutical industry leads the broader wave by roughly two years, the Technology industry coincides with the wave peak, and the Energy industry lags by approximately one year. In Wave 2, by contrast, Manufacturing and Wholesale lead the wave (consistent with the credit-and-LBO character of that cycle) while Utilities lag by two years.

The industry-level lead-lag pattern reinforces the wave-level sequencing finding (H1) and is consistent with the network propagation result of Ahern and Harford (2014), who show that merger waves propagate along supply-chain links: an industry hit early by a structural shock pulls upstream and downstream industries into the wave with a lag. The fact that the lead-lag structure is itself wave-specific, different industries lead in different waves, is consistent with the fundamental proposition that each wave has a different originating shock. Our framework allows the originating-shock hypothesis to be tested wave by wave rather than only on the average wave.

[Insert Figure 7 here]

6.5 The Universal Empirical Core

Beneath the visible heterogeneity of waves and the rotation of dominant groups, the variable-level data tell a different and more striking story. The same handful of individual variables, most prominently the Piotroski F-score, the high-yield premium, the analyst summary index, and prior deal count, appear in the top three importance ranks of essentially every wave-phase cell.

[Insert Table 5 here]

The Piotroski F-score (Piotroski, 2000) appears in the top three predictors of all 19 wave-phase cells, a perfect 100% score. In 18 of the 19 cells, it is rank number 1 within the firm-specific cluster. This is a remarkable result: a financial-statement-based summary of nine binary indicators of profitability, liquidity, and operating efficiency is the single most informative variable for deal incidence in every wave and every phase across five decades of data. Across waves, the Piotroski F-score in our framework plays a similar role to

the role that valuation-decomposition residuals play in Rhodes-Kropf et al. (2005), but at a higher signal-to-noise ratio: it summarises firm-level fundamentals through accounting fundamentals rather than through market valuations. The high-yield premium is in the top three of 14 of 19 cells (74%), most consistently in waves with high credit-cycle relevance. This is the credit-spread variable that operationalises the Harford (2005) and Eisfeldt and Rampini (2006) capital-liquidity story, and the fact that it is universally present in cells that include either pre-tightening Build-up phases or post-tightening Decline phases is direct evidence for the gating role of macro-financial conditions.

The analyst summary index, drawing on Baker and Wurgler (2007)'s sentiment framework, is in the top three of 8 of 19 cells (42%); its presence concentrates in Peak and late-Build-up phases, consistent with sentiment-driven amplification near wave peaks. Prior-deal count is in the top three of 8 of 19 cells (42%); it captures the experience effect documented by Fuller et al. (2002) and the path dependence emphasised by Harford and Li (2007). Together, these four variables account for 49 of the 57 top-three positions across the 19 cells, fully 86% of the top-three slots are filled by this small core.

The cross-wave Spearman rank consistency of the full importance ordering provides additional evidence for the existence of a universal core. Comparing the importance vectors of any two waves, the Spearman correlation is between 0.79 and 0.97, with the highest values between adjacent waves (0.97 between Waves 3 and 4) and the lowest values between non-adjacent waves with very different macro environments (0.79 between Waves 1 and 4). The Jaccard similarity of the top-twenty variables across wave pairs is similarly high, between 0.55 and 0.85, and rises further when the comparison is restricted to the top-five variables. The implication is that merger waves rotate at the cluster-rank level (which group is rank 1) but conserve at the variable-rank level (which individual variables are universally important).

[Insert Table 6 here]

6.5.1 Refining Capital Liquidity: Spreads, Not Rates

The universal-core finding has a sharper implication for the macro-financial side of the merger-wave literature than the cluster-level discussion brings out. The Harford

(2005) story, taken up and extended by Eisfeldt and Rampini (2006), Adra et al. (2020), and Gulen et al. (2025), treats 'capital liquidity' as a single composite construct, operationalised in different empirical applications by some combination of the federal funds rate, Treasury yields, the term premium, default spreads, and high-yield premia. The literature has not generally distinguished between these constituents because the linear-regression specifications it has used produce highly correlated coefficients on rate-level and spread-level variables, leaving the question of which of the two carries the predictive weight largely unresolved. Our phase-resolved decomposition, working with non-linear interactions, is in a position to separate them, and the result is that they are not interchangeable.

At the cluster level, the Monetary Premia sub-group, comprising default and high-yield credit spreads together with the TED spread, carries between 2.4 and 4.3 times the predictive importance of the Interest Rates sub-group, comprising of the federal funds rate, the one- and ten-year Treasury yields, and the term premium. The ratio is 2.4 in Wave 1, 2.7 in Wave 2, 2.6 in Wave 3, and 4.3 in Wave 4. In every wave it is the credit-risk component of the cost of capital, not the risk-free rate, that carries the predictive signal for deal incidence. The variable-level evidence sharpens this further. Among the ten macro variables that ever appear in the top three predictors of any wave-phase cell, all are credit-spread variables: the high-yield premium (in 14 of 19 cells) and its complement the default spread (in three further cells). No interest-rate variable, neither the federal funds rate nor any Treasury yield, ever reaches the top three of any wave-phase cell across our five-decade panel.

This refines the Harford (2005) story in a precise way: the gating role of macro-financial conditions in merger waves is exercised by credit-risk premia rather than by the level of the risk-free rate, and the appropriate empirical operationalisation of capital liquidity in the merger-wave context is the spread between risky and safe debt, not their levels. The economic interpretation of what is important for the question whether a wave can form is that risky borrowers can access debt at acceptable cost, and that this condition is captured by the spread rather than by the level. Adra et al. (2020) and Fischer and Horn (2024) are consistent with this reading, showing respectively that the effects of monetary

policy on M&A operate primarily through risk-premium channels and that monetary tightening alters deal selection rather than deal volume directly. Bonaime et al. (2018) likewise document that policy uncertainty (a quantity correlated with risk premia) shapes M&A activity beyond what interest rates alone explain. Our contribution is to make the rate-versus-spread distinction directly visible in a phase-resolved framework and to show that the rate-level variables that have featured prominently in the linear-regression literature are not, in fact, the macro driver of the merger wave: instead the spread variables are most important.

The refinement matters for the wave dating literature as well. Wave 1 (the dot-com era), Wave 2 (the credit-driven era), and the late phases of Wave 4 are widely associated in the literature with macro-financial conditions, but the proxies used in the prior literature have differed across studies, often reflecting whichever variable was most readily available rather than a substantive theoretical commitment. Our finding implies that comparison across studies should be made conditional on which macro variable was used: studies relying on the federal funds rate or Treasury yields are operating with a substantially weaker signal than studies relying on credit spreads, and the magnitude of the difference, a factor of two to four in importance share, is large enough to matter for whether macro factors look 'decisive' or 'marginal' in any given regression. The empirical disagreement between Harford (2005) and Adra et al. (2020) about whether macro factors initiate waves or merely modulate their outcomes can in part be read in this light: both are looking at macro factors, but at different macro factors, and the relevant macro factor in our data is the spread.

6.6 Behavioural Variables: Genuine Complementarity

The behavioural cluster contains the smallest number of variables (15 of the 100) and the smallest aggregate share of importance (between 0.04 and 0.32 across cells). It is also the cluster for which the XGBoost-versus-logistic-regression gap is largest, suggesting that behavioural mechanisms operate through non-linear interactions and threshold effects that linear specifications cannot capture. This is consistent with the theoretical prediction of Malmendier and Tate (2005, 2008) that the overconfidence effect on acquisition propensity is non-linear in available financial slack: an overconfident CEO with

abundant cash is sharply more likely to acquire, while an overconfident CEO without cash is barely more likely than a non-overconfident CEO, a discontinuous threshold effect that maps directly onto the kind of split that XGBoost can represent and logistic regression cannot.

The aggregate behavioural-cluster importance is roughly 1.73 times higher in wave years than in non-wave years (averaging across all wave years and all non-wave years), with the largest ratios found in the Peak phases of Waves 1 and 4. To rule out the possibility that the wave-versus-non-wave behavioural-importance gap is driven by general structural differences between waves and non-waves rather than by the wave dynamics per se, we run a same-era control comparison: the behavioural importance share in Waves 3 and 4 is 1.84 times higher than in the contemporaneous non-wave year 2009, holding the macro environment roughly constant. This rules out a substantial class of alternative explanations and suggests that the behavioural amplification is genuinely tied to wave dynamics rather than to differences in the underlying era.

The CEO Background sub-group is rank 1 in nearly all phases of Waves 3 and 4, and the importance share of the Analyst Sentiment sub-group spikes from 0.059 to 0.196 in the Build-up→Peak transition of Wave 1. This is the largest single phase-to-phase increase in any sub-group importance share in any wave in our sample. This 233% jump occurs precisely at the year that Baker and Wurgler (2007) identify as the maximum-sentiment year of the dot-com era and is consistent with the mis-valuation channel of Shleifer and Vishny (2003): when sentiment is most extreme, the marginal information content of analyst-sentiment signals for predicting deal incidence rises sharply because that is when the mis-valuation-driven acquisition channel is most active. The result reconciles a long-standing tension between Rhodes-Kropf et al. (2005)'s finding that mis-valuation explains roughly 15% of sectoral merger activity even within Harford (2005)'s neoclassical-shock periods and Bouwman et al. (2009)'s finding that high-valuation periods generate worse acquisition quality. Both findings are consistent with a behavioural channel that activates at the Peak rather than throughout the wave.

6.7 Industry Robustness

To verify that the wave-phase results are not driven by industry-specific compositional effects, we re-run the entire phase decomposition separately within each of the Fama-French twelve industries. The wave-level rank rotation persists at the industry level: in 9 of the 12 industries, the same wave dominance pattern emerges (Monetary Premia in Wave 1, Other Firm-specific in Wave 2, CEO Background in Wave 3, mixed in Wave 4).

[Insert Figure 8 here]

Where the cross-industry pattern departs from the aggregate is precisely where theory predicts it should: industries with strong behavioural components (IT, Healthcare) show a behavioural-importance ratio that is approximately 30% higher than the cross-industry average. This is consistent with the prediction that behavioural channels are most active in industries where market valuations and managerial cognition are most variable. The universal core variables (Piotroski F-score, high-yield premium, prior-deal count) appear in the top three of each industry's wave-phase ranking with similar consistency to the all-industry analysis, confirming that the structural backbone documented is not an artefact of cross-industry aggregation but a feature of the industry-level data themselves.

[Insert Figure 9 here]

6.7.1 Which Wave Phases Are Easier to Predict?

Table 7 displays the F1 score, precision, recall, and balanced accuracy of the combined XGBoost model evaluated separately within each of the industries and within each of the four wave phases (plus the three Non-Wave control segments). Interpreting the figure both by phase (across rows) and by industry (across columns) yields a result that is, at a methodological level, every bit as informative as the variable-importance findings: wave phases are not equally predictable, and the heterogeneity in predictability has a clean structural interpretation that connects directly to the theoretical synthesis.

[Insert Table 7 here]

The Decline phase is the easiest phase of a wave to predict. Averaged across all twelve industries and all four waves, the Decline phase delivers the highest mean F1 of any phase (0.384), the highest mean balanced accuracy (0.702), and the highest mean precision (0.308). The intuition runs directly through the variable-importance findings: by the time a wave enters its Decline phase, the residual acquirers are systematically different from the non-acquirers in firm-level quality and in their accumulated acquisition experience, and the model identifies them on the basis of the universal core variables that we documented to be most informative in this phase. We showed that Other Firm-specific importance rises in every wave's Decline phase by between 10% and 51%; here we see the corresponding effect on out-of-sample performance. When firm-level quality becomes the binding determinant of who acquires, prediction quality rises because firm-level quality is a comparatively stable, accounting-driven signal that the model can read precisely.

The Peak phase has the highest recall (0.556) of any phase but only intermediate precision (0.286), reproducing at the industry level the precision-recall trade-off that we documented for the aggregate model. The economic interpretation is the same as before: at the peak of a wave the predictor space is saturated with firms that look like potential acquirers (high sentiment, abundant credit, aligned firm-quality), and the model correctly identifies almost all of the actual acquirers but also flags many firms that, despite looking ready to acquire, do not consummate a deal. This pattern is sharpest in the IT and Healthcare industries, where Peak-phase recall exceeds 0.65 in some waves while precision falls below 0.20, a finding consistent with the well-documented role of sentiment-driven over-firing in technology and pharmaceutical M&A markets (Bouwman et al., 2009; Duchin and Schmidt, 2013).

The Build-up phase is the most uniform phase across industries. The cross-industry standard deviation of F1 within the Build-up phase is 0.078, the lowest of any phase, indicating that the model performs at similar levels across industries during the years in which a wave is gathering momentum. The interpretation is that the macro-financial gating that we attribute to the Build-up phase operates similarly across all industries: when credit conditions ease and structural pressures translate into firm-level activity, the cross-industry dispersion in deal incidence narrows and so does the cross-industry dispersion in the

model's ability to predict it. The Build-up phase is, in other words, the phase in which the wave is most clearly a coordinated economy-wide phenomenon rather than a collection of industry-specific stories. The Decline and Peak phases, by contrast, show cross-industry F1 standard deviations of 0.100 and 0.105, respectively, indicating that as the wave matures and unwinds, industry-specific stories regain their importance and predictability becomes more industry-dependent.

The Non-Wave segments are the hardest to predict. Mean F1 across all industries in Non-Wave segments is 0.308, balanced accuracy is 0.632, and recall is 0.378. All four metrics are below the corresponding wave-phase averages. The interpretation reinforces the structural-cyclical distinction at the heart of the paper: when the wave is off, the cyclical modulators (macro and behavioural conditions) provide little signal to the model, leaving only the structural backbone (firm characteristics) to differentiate acquirers from non-acquirers. Because acquisition activity in non-wave periods is rare and idiosyncratic, even the structural backbone is not fully predictive, and the model's performance falls. This is consistent with our finding that the behavioural amplification of the model's signal is concentrated in wave years, and it is the obverse of the Decline-phase finding: when the wave is on, even the residual late phase is more predictable than any year outside a wave.

At the industry level, the same patterns appear with industry-specific accents. Banks deliver the highest mean F1 across all phases (0.392), reflecting the fact that banking M&A is heavily regulated and follows a more deterministic structure that the model captures well. The Durable goods, Healthcare, and Manufacturing industries follow at F1 in the high 0.34 range. Utilities deliver the lowest mean F1 (0.309), consistent with their regulated status and the comparatively flat importance profiles documented in the second panel of Figure 12. Across the four waves, Wave 3 (2010 to 2019) is the most predictable across industries (mean F1 0.414 across phases), reflecting both the length of the wave and the comparatively stable macro environment of the post-crisis era; Wave 1 (1991 to 2002) is the least predictable on average (mean F1 0.276), reflecting the saturated dot-com Peak and the structural break in the late 1990s that shifted the relative importance of the variable clusters mid-wave.

The implications of this stage-and-industry decomposition reach beyond the methodological observation that prediction quality varies systematically with the wave-phase cell, although that observation is itself worth registering: a model evaluated only on aggregate fit obscures patterns that the wave-phase view brings out. The substantive lesson is that the highest-conviction predictions a model of merger waves can produce are predictions about the Decline phase, when the structural backbone of firm-level quality is most informative. This has practical relevance for both academic researchers, who can use Decline-phase predictions as the cleanest test bed for hypotheses about post-wave selection, and practitioners, who can use them to identify late-cycle acquirers. The cross-industry uniformity of Build-up-phase predictability points in a complementary direction by supporting the proposition that the Build-up is the phase at which a wave is most clearly an economy-wide rather than industry-specific phenomenon. Far from a methodological footnote, this is a substantive finding about the structure of merger waves themselves: the wave is at its most coordinated in Build-up and at its most idiosyncratic in Decline, and the model's behaviour in each phase reflects rather than distorts that structure.

Figure 10 provides the industry-level cross-correlation, showing that within each wave the lead-lag structure visible at the aggregate level is mirrored at the industry level with industry-specific magnitudes. In Wave 4, the Healthcare industry's Output-and-Labor channel lags aggregate deal counts by approximately one and a half years, and the IT industry's Behavioural channel lags by approximately one year, with both lagged correlations strongly positive. The lag direction is consistent with deal flow itself driving the subsequent reorganization of the predictor space: an AI-driven Wave 4 deal in IT is followed roughly one year later by elevated importance of behavioural-sentiment variables, the period during which the integration of the acquired technology and the surrounding labour-market response feed back into the model's predictive signal — and a Healthcare deal is followed about a year and a half later by elevated importance of Output-and-Labor variables, the natural horizon over which clinical-staffing and operating-cost adjustments materialise post-merger.

[Insert Figure 10]

The recurring-onsets analysis, run at the industry level rather than the variable-group level, sharpens the within-wave heterogeneity finding. Figure A plots, for each of the Fama-French twelve industries, the years at which industry-level mean importance crosses a half-standard-deviation threshold and the duration of the resulting activation episode. The dominant feature of the figure is a synchronised activation episode around 1997 to 2000 that is shared by essentially every industry except Banks, with the activation peaking in 1998 to 1999 and lasting between two and four years depending on the industry. This is the empirical signature at the industry level of the within-wave synchrony documented for Wave 1: the new-economy boom moved across industries essentially together rather than propagating from a small set of leaders, and the recurring-onsets analysis makes that synchrony visible by way of the near-coincident activation thresholds across the twelve industries. A weaker secondary activation episode is visible in 2003 to 2008 in four industries (Durables, Business Equipment, Telecommunications, and Pharmaceuticals), corresponding to the second-wave participation of these industries; the absence of a comparably visible 2010 to 2019 episode in most industries reflects the comparatively flat importance profiles of the long Wave 3 rather than absence of activity. Banks are the structural exception, with six recurring activations distributed across the panel and the highest within-industry threshold of any industry. The Bank pattern reflects the regulated, regulation-cycle-driven character of bank consolidation, which runs on a different cycle from the rest of the economy and does not synchronise with the broader merger waves.

[Insert Figure 11 here]

7 Synthesis

This section maps the topic-organised findings back onto the six hypotheses and articulates the theoretical synthesis they support.

7.1 Verdict on the Six Hypotheses

H1 (Sequencing), Supported. The wave-level cross-correlation evidence, particularly Wave 4 (where the data are cleanest), reveal that Structural Change importance

leads deal counts by two years ($r = +0.93$), Output and Labor importance leads by one year ($r = +0.97$), and Monetary Premia importance lags by one year ($r = -0.93$). The same temporal ordering is visible, at lower statistical power, in the other waves. The recurring-onsets analysis (Figure 14) shows that the temporal ordering is recurrent rather than wave-idiosyncratic. The within-wave group-level cross-correlations (Figure 4) and industry-level lead-lag structure corroborate the aggregate finding at finer levels of disaggregation.

H2 (Macro-as-Gate, Firm-as-Selector), Supported, with refinement. The macro share of importance declines monotonically across waves, while the firm share remains anchored near 0.50 in every cell (Table 5). The macro share's position in the prediction is most informative for the *level* of activity (whether the wave is on or off), while the firm share's position is most informative for the *cross-section* of which firms participate. The refinement is that within macro, the gating role is concentrated specifically in the Monetary Premia sub-group (which appears in the universal core and is the rank-1 sub-group in Wave 1) rather than in the entire macro cluster: it is credit conditions, not overall macro conditions, that operate as the binary gate.

H3 (Firm-Characteristic Dominance and Decline Amplification), Strongly supported. The Other Firm-specific sub-group is rank 1 within the firm cluster in 18 of 19 wave-phases, and the Peak→Decline rise is observed in all four waves with magnitudes between +10% and +51%. The strength of H3 is the principal corrective our paper offers to the prior literature, which has largely treated firm characteristics as control variables to be partialled out.

H4 (Behavioural Complementarity), Supported. The behavioural cluster contributes incremental information over and above macro and firm-specific variables, with the behavioural-only XGBoost model achieving $F1 = 0.187$ (versus the combined model's $F1 = 0.341$) and the wave-versus-non-wave behavioural-importance ratio being 1.73 across all cells and 1.84 in the same-era control. The Build-up-to-Peak Analyst Sentiment surge in Wave 1, from 0.059 to 0.196, a 233% increase, is the cleanest single piece of evidence for behavioural amplification at wave peaks.

H5 (Wave Heterogeneity), Supported. Each of the four waves has a different ranking of groups: Monetary Premia in Wave 1, Other Firm-specific in Wave 2, CEO Background in Wave 3, mixed CEO/firm-quality in Wave 4. The bootstrap test rejects identical wave-level rankings at $p < 0.001$ in 22 of 24 wave-pair comparisons. The rotation is not random: the dominant group of each wave maps to the dominant economic context of that wave (macro/credit cycle for Wave 1, internal financing capacity for Wave 2, managerial-cognition era for Wave 3, technology-and-cognition for Wave 4).

H6 (Universal vs. Wave-Specific Drivers), Strongly supported. We document that the Piotroski F-score appears in the top three predictors of all different wave-phases (100%), the high-yield premium in most (74%), and the analyst summary index and prior deal count each many wave phases (42%). Together these four variables fill 86% of the top-three slots across all merger wave's stages. The cross-wave Spearman rank consistency is between 0.79 and 0.97, with adjacent waves (W3/W4) at 0.97. The combination of wave-level rotation in the dominant group (H5) and variable-level conservation in the top three (H6) is the central empirical synthesis of the paper.

7.2 Theoretical Synthesis

The neoclassical and behavioural traditions are not competing explanations of merger waves; they describe different temporal segments of the same phenomenon. Structural-change variables and aggregate firm-quality indicators activate before macro and behavioural groups within each wave (H1, H6); macro-financial conditions then gate the level of acquirer participation once the wave has formed (H2); firm-specific characteristics select which eligible firms cross the gate (H2, H3); behavioural amplification, analyst sentiment in particular, concentrates around the peak (H4); and monetary tightening dominates the decline (H1). The dominant group of each wave reflects the dominant economic context of that wave (H5), but beneath the surface heterogeneity the same handful of individual variables predict deal incidence in every cell (H6).

This synthesis has implications that ripple through the prior literature. The most immediate is that the apparent disagreement between studies that find macro factors decisive (Harford, 2005; Adra et al., 2020; Gulen et al., 2025) and studies that find

managerial sentiment decisive (Shleifer and Vishny, 2003; Malmendier and Tate, 2008) dissolves once one is willing to allow that different phases of a wave are governed by different mechanisms. Macro factors dominate in Initiation and Decline; managerial-cognition variables dominate in Peak. Both views are correct, but neither is complete on its own, and the historical pattern of empirical disagreement reflects sample-period differences as much as it reflects substantive theoretical differences. A finer reading of the macro side is also possible. The Harford (2005) capital-liquidity story is supported by our data, but with a sharper specification than the prior literature has typically used: it is credit-risk premia, not interest-rate levels, that carry the predictive weight. The federal funds rate and Treasury yields, prominent as they have been in the linear-regression M&A literature, do not appear in the top three predictors of any of our wave-phase cells; the high-yield premium and the default spread do, and consistently. The disagreement between Harford (2005) and Adra et al. (2020) over whether macro factors initiate waves or only modulate their outcomes can in part be read as a disagreement about which macro factor is relevant: with the spread variables, the macro role is gating, with the rate variables, it is residual.

A second strand of the prior literature is also reframed by our analysis, this time at the cross-sectional level. The role of firm characteristics in our framework is more central than the literature has typically acknowledged. Whereas Palepu (1986), Powell (2004), and Espahbodi and Espahbodi (2003) treat firm characteristics as the primary input but largely without distinguishing wave from non-wave periods, and whereas Maksimovic and Phillips (2001) emphasise efficient asset reallocation conditional on macro fundamentals, our analysis shows that firm characteristics are the structural backbone of the prediction across all waves and phases, not merely the cross-sectional control. The complementarity with the macro-cyclical view is the substantive point: firm characteristics carry the prediction's universal component, and macro and behavioural variables modulate the level of activity around that universal component.

The third strand the synthesis touches is the longstanding 'each wave is unique' framing of the merger-wave literature. The existence of a universal core challenges that framing without dismissing it. Each wave has a distinctive dominant group of variables, in

keeping with the literature's instinct that waves carry the imprint of their economic moment. But the same handful of individual variables predicts deal incidence in every wave, suggesting that there are deeper structural regularities of which the visible wave-specific patterns are surface manifestations. The wave is therefore neither uniformly recurrent nor uniformly idiosyncratic; it is a stable empirical core overlaid with cyclical accents, and the methodological proposition that flows from this conclusion is that the wave-phase cell, rather than the wave or the year, is the right unit of empirical analysis.

8 Conclusion

We provide a phase-resolved decomposition of the determinants of U.S. merger waves over the past five decades, using a rolling-window XGBoost classifier benchmarked against logistic regression on a panel of 74,765 completed M&A transactions. The methodological structure, combining macro, firm-specific, and behavioural variable clusters in a single non-linear model and decomposing the resulting variable importance across the four phases of each of the four post-1989 U.S. merger waves, allows us to out-of-sample test the joint relationship between the neoclassical and behavioural theories of merger waves that has historically been outside the reach of standard linear methods.

The empirical findings produce a unified picture of the merger wave that the prior literature has, in our reading, glimpsed in fragments rather than seen as a whole. Neoclassical and behavioural determinants are not competing explanations of the same phenomenon, they describe the same phenomenon at different phases of the wave lifecycle. Structural-change variables and aggregate firm-quality indicators activate first; macro-financial conditions gate the level of acquirer participation; firm-specific characteristics select which eligible firms participate; behavioural amplification concentrates around the peak; and monetary tightening dominates the decline. The Wave 4 cross-correlation evidence makes this temporal ordering directly visible: structural change leads deal counts by two years, output and labor by one year, and monetary premia lag by one year.

Beneath the wave-by-wave heterogeneity that the literature has documented and that our framework reproduces, the variable-level data tells a different and more striking

story. The same handful of individual variables, the Piotroski F-score, the high-yield premium, the analyst summary index, and prior-deal count, predicts deal incidence in essentially every wave-phase cell, even as the dominant group rotates across waves. Together these four variables fill 86% of the top-three slots across the 19 wave-phase cells. The cluster-level evidence reinforces the same reading by way of an asymmetry: the macro share of importance declines monotonically across waves while the behavioural share rises, but the firm-specific share remains anchored near 0.50 in every cell. Macro and behavioural variables are the cyclical modulators of merger activity; firm-specific variables are the structural backbone.

These findings have implications for both empirical research design and corporate finance theory. Empirically, the result that the wave-phase cell is the right unit of analysis, rather than the wave or the year, calls for a methodological reconfiguration of the merger-wave literature toward phase-resolved tests. Theoretically, the result that the neoclassical and behavioural traditions are sequential rather than alternative reframes the foundational question of the field: instead of asking which view is right, we should be asking *when* each view's mechanisms bind. The result that firm-characteristics are the structural backbone of merger-wave prediction across all waves and phases adds a third tradition, the firm-quality view rooted in Piotroski (2000) and Maksimovic and Phillips (2001), to the standard neoclassical-versus-behavioural dichotomy. The combined model significantly outperforms the linear benchmark (Diebold-Mariano $p < 0.001$), the F1 improvement (roughly two-fold) is substantial, and the framework extends naturally to deal outcome prediction, to non-U.S. samples, and to the integration of textual data, all directions for future work.

References

- Adra, S., L. G. Barbopoulos, and A. Saunders, 2020, The impact of monetary policy on M&A outcomes, *Journal of Corporate Finance*, 62, 101529.
- Ahern, K. R., and J. Harford, 2014, The Importance of Industry Links in Merger Waves, *The Journal of Finance*, 69(2), 527–576.
- Alexandridis, G., D. Petmezas, and N. G. Travlos, 2010, Gains from Mergers and Acquisitions Around the World: New Evidence, *Financial Management*, 39(4), 1671–1695.
- Andrade, G., M. Mitchell, and E. Stafford, 2001, New Evidence and Perspectives on Mergers, *Journal of Economic Perspectives*, 15(2), 103–120.
- Baker, M., and J. Wurgler, 2007, Investor Sentiment in the Stock Market, *Journal of Economic Perspectives*, 21(2), 129–152.
- Baker, S. R., N. Bloom, and S. J. Davis, 2016, Measuring Economic Policy Uncertainty, *Quarterly Journal of Economics*, 131(4), 1593–1636.
- Barnes, P., 2000, The Identification of U.K. Takeover Targets Using Published Historical Cost Accounting Data: Some Empirical Evidence Comparing Logit with Linear Discriminant Analysis and Raw Financial Ratios with Industry-Relative Ratios, *International Review of Financial Analysis*, 9(2), 147–162.
- Bebchuk, L. A., K. J. M. Cremers, and U. C. Peyer, 2011, The CEO pay slice, *Journal of Financial Economics*, 102(1), 199–221.
- Bena, J., and K. Li, 2014, Corporate Innovations and Mergers and Acquisitions, *The Journal of Finance*, 69(5), 1923–1960.
- Bessler, W., and T. Nohel, 2025, Sponsor Reputation and the Structure of SPAC Units, Working Paper.
- Bessler, W., and D. Wolff, 2024, Portfolio Optimization with Sector Return Prediction Models, *Journal of Risk and Financial Management*, 17(6), 254.

Betton, S., B. E. Eckbo, and K. S. Thorburn, 2008, Corporate Takeovers, in B. E. Eckbo (ed.), *Handbook of Corporate Finance: Empirical Corporate Finance*, Vol. 2, Elsevier, 291–429.

Bi, W., and Q. Zhang, 2021, Forecasting Mergers and Acquisitions Failure Based on Partial-Sigmoid Neural Network and Feature Selection, *PLoS ONE*, 16(11), 1–18.

Bianchi, D., and C. Chiarella, 2019, An Anatomy of Industry Merger Waves, *Journal of Financial Econometrics*, 17(2), 153–179.

Bonaime, A., H. Gulen, and M. Ion, 2018, Does Policy Uncertainty Affect Mergers and Acquisitions?, *Journal of Financial Economics*, 129(3), 531–558.

Boone, A. L., and J. H. Mulherin, 2007, Do termination provisions truncate the takeover bidding process?, *Review of Financial Studies*, 20(2), 461–489.

Bouwman, C. H. S., K. Fuller, and A. S. Nain, 2009, Market Valuation and Acquisition Quality: Empirical Evidence, *The Review of Financial Studies*, 22(2), 633–679.

Bradley, M., A. Desai, and E. H. Kim, 1988, Synergistic gains from corporate acquisitions and their division between the stockholders of target and acquiring firms, *Journal of Financial Economics*, 21(1), 3–40.

Bustamante, M. C., and L. Frésard, 2021, Does Firm Investment Respond to Peers' Investment?, *Management Science*, 67(8), 4703–4724.

Campbell, J. L., E. T. Elfrink, C. Irons, and J. Moon, 2024, What Is the Deal? Predicting M&A Outcomes with Machine Learning, Working Paper, University of Georgia.

Carow, K., R. Heron, and T. Saxton, 2004, Do Early Birds Get the Returns? An Empirical Investigation of Early-Mover Advantages in Acquisitions, *Strategic Management Journal*, 25(6), 563–585.

Chen, T., J. Harford, and K. Li, 2007, Monitoring: Which institutions matter?, *Journal of Financial Economics*, 86(2), 279–305.

Chen, T., and C. Guestrin, 2016, XGBoost: A Scalable Tree Boosting System, Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 785–794.

Cho, H., and J. Chung, 2022, Stock Market Liquidity and the Aggregate Merger Wave, Journal of Banking and Finance, 137, 106466.

De Bondt, W. F. M., and H. E. Thompson, 1992, Is economic efficiency the driving force behind mergers?, Managerial and Decision Economics, 13(1), 31–44.

Doukas, J. A., and W. Zhang, 2014, Envy-Motivated Merger Waves, European Financial Management, 22(1), 63–119.

Duchin, R., and B. Schmidt, 2013, Riding the merger wave: Uncertainty, reduced monitoring, and bad acquisitions, Journal of Financial Economics, 107(1), 69–88.

Eisfeldt, A. L., and A. A. Rampini, 2006, Capital reallocation and liquidity, Journal of Monetary Economics, 53(3), 369–399.

El-Khatib, R., K. Fogel, and T. Jandik, 2015, "CEO Network Centrality and Merger Performance", Journal of Financial Economics 116(2), 349–382

Emiru, T., and S. Weisblatt, 2025, Economic Tides and Merger Waves: Insights from a Long-Run Perspective, Studies in Economics and Finance, 42(1), 75–88.

Erel, I., R. C. Liao, and M. S. Weisbach, 2012, Determinants of Cross-Border Mergers and Acquisitions, The Journal of Finance, 67(3), 1045–1082.

Espahbodi, H., and P. Espahbodi, 2003, Binary Choice Models and Corporate Takeovers, Journal of Banking and Finance, 27(4), 549–574.

Ewens, M., and J. Farre-Mensa, 2020, The Deregulation of the Private Equity Markets and the Decline in IPOs, The Review of Financial Studies, 33(12), 5463–5509.

Fama, E. F., and K. R. French, 1997, Industry Costs of Equity, Journal of Financial Economics, 43(2), 153–193.

Faccio, M., and R. W. Masulis, 2005, The Choice of Payment Method in European Mergers and Acquisitions, The Journal of Finance, 60(3), 1345–1388.

Fischer, J. J., and C.-W. Horn, 2024, Monetary Policy and Mergers and Acquisitions, Working Paper.

Fuller, K., J. M. Netter, and M. Stegemoller, 2002, What Do Returns to Acquiring Firms Tell Us? Evidence from Firms That Make Many Acquisitions, *The Journal of Finance*, 57(4), 1763–1793.

Garfinkel, J. A., and K. W. Hankins, 2011, The role of risk management in mergers and merger waves, *Journal of Financial Economics*, 101(3), 515–532.

Goel, A. M., and A. V. Thakor, 2010, Do Envious CEOs Cause Merger Waves?, *The Review of Financial Studies*, 23(2), 487–517.

Golbe, D., and L. White, 1988, A Time-Series Analysis of Mergers and Acquisitions in the U.S. Economy, in A. Auerbach (ed.), *Corporate Takeovers: Causes and Consequences*, University of Chicago Press, Chicago.

Gorbenko, A. S., and A. Malenko, 2014, Strategic and Financial Bidders in Takeover Auctions, *The Journal of Finance*, 69(6), 2513–2555.

Gorbenko, A. S., and A. Malenko, 2018, The Timing and Method of Payment in Mergers When Acquirers Are Financially Constrained, *The Review of Financial Studies*, 31(10), 3937–3978.

Gort, M., 1969, An Economic Disturbance Theory of Mergers, *The Quarterly Journal of Economics*, 83(4), 624–642.

Gorton, G., M. Kahl, and R. J. Rosen, 2009, Eat or Be Eaten: A Theory of Mergers and Firm Size, *The Journal of Finance*, 64(3), 1291–1344.

Goyal, A., and I. Welch, 2008, A Comprehensive Look at the Empirical Performance of Equity Premium Prediction, *The Review of Financial Studies*, 21(4), 1455–1508.

Gu, S., B. Kelly, and D. Xiu, 2020, Empirical Asset Pricing via Machine Learning, *The Review of Financial Studies*, 33(5), 2223–2273.

Gugler, K., D. C. Mueller, and M. Weichselbaumer, 2012, The determinants of merger waves: An international perspective, *International Journal of Industrial Organization*, 30(1), 1–15.

Gulen, H., C. E. Jens, and S. Rossi, 2025, Credit Market Driven Acquisitions, Working Paper.

Hajek, P. and R. Henriques, 2024, Predicting M&A targets using news sentiment and topic detection, *Technological Forecasting and Social Change*, 201, 123270

Harford, J., 1999, Corporate Cash Reserves and Acquisitions, *The Journal of Finance*, 54(6), 1969–1997.

Harford, J., 2005, What Drives Merger Waves?, *Journal of Financial Economics*, 77(3), 529–560.

Harford, J., 2024, Merger Waves: A Survey of the Literature, in D. J. Denis (ed.), *Handbook of Corporate Finance*, Edward Elgar Publishing.

Harford, J., and K. Li, 2007, Decoupling CEO Wealth and Firm Performance: The Case of Acquiring CEOs, *The Journal of Finance*, 62(2), 917–949.

Harford, J., S. A. Mansi, and W. F. Maxwell, 2008, Corporate governance and firm cash holdings in the U.S., *Journal of Financial Economics*, 87(3), 535–555.

Hoberg, G., and G. Phillips, 2010, Product Market Synergies and Competition in Mergers and Acquisitions: A Text-Based Analysis, *The Review of Financial Studies*, 23(10), 3773–3811.

Hoberg, G., and G. Phillips, 2016, Text-Based Network Industries and Endogenous Product Differentiation, *Journal of Political Economy*, 124(5), 1423–1465.

Hornuf, L., and P. Schäfer, 2025, Artificial Intelligence and Machine Learning in Corporate Finance, Working Paper, forthcoming Oxford Handbook on Corporate Finance.

Ince, B., 2024, How do regulatory costs affect merger and acquisitions decisions and outcomes?, *Journal of Banking and Finance*, 163, 107156.

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

Jensen, T. I., B. Kelly, and L. H. Pedersen, 2023, Is There a Replication Crisis in Finance?, *The Journal of Finance*, 78(5), 2465–2518.

Jovanovic, B., and P. Rousseau, 2001, Mergers and Technological Change: 1885–2001, Working Paper, Vanderbilt University.

Jovanovic, B., and P. Rousseau, 2002, The Q-Theory of Mergers, *American Economic Review*, 92(2), 198–204.

Lattanzio, G., W. L. Megginson, and A. Sanati, 2023, Dissecting the listing gap: Mergers, private equity, or regulation?, *Journal of Financial Markets*, 54, 100836.

Maksimovic, V., and G. Phillips, 2001, The market for corporate assets: Who engages in mergers and asset sales and are there efficiency gains?, *The Journal of Finance*, 56(6), 2019–2065.

Maksimovic, V., and G. Phillips, 2002, Do conglomerate firms allocate resources inefficiently across industries? Theory and evidence, *The Journal of Finance*, 57(2), 721–767.

Maksimovic, V., G. Phillips, and L. Yang, 2013, Private and Public Merger Waves, *The Journal of Finance*, 68(5), 2177–2217.

Malmendier, U., and G. Tate, 2005, CEO Overconfidence and Corporate Investment, *The Journal of Finance*, 60(6), 2661–2700.

Malmendier, U., and G. Tate, 2008, Who makes acquisitions? CEO overconfidence and the market's reaction, *Journal of Financial Economics*, 89(1), 20–43.

Malmendier, U., and S. Nagel, 2011, Depression Babies: Do Macroeconomic Experiences Affect Risk Taking?, *Quarterly Journal of Economics*, 126(1), 373–416.

Martynova, M., and L. Renneboog, 2008, A century of corporate takeovers: What have we learned and where do we stand?, *Journal of Banking and Finance*, 32(10), 2148–2177.

Masulis, R. W., C. Wang, and F. Xie, 2007, Corporate Governance and Acquirer Returns, *The Journal of Finance*, 62(4), 1851–1889.

Melicher, R. W., J. Ledolter, and L. J. D'Antonio, 1983, A Time Series Analysis of Aggregate Merger Activity, *The Review of Economics and Statistics*, 65(3), 423–429.

Mitchell, M., and H. Mulherin, 1996, The Impact of Industry Shocks on Takeover and Restructuring Activity, *Journal of Financial Economics*, 41(2), 193–229.

Moeller, S. B., F. P. Schlingemann, and R. M. Stulz, 2004, Firm Size and the Gains from Acquisitions, *Journal of Financial Economics*, 73(2), 201–228.

Moeller, S. B., F. P. Schlingemann, and R. M. Stulz, 2005, Wealth Destruction on a Massive Scale? A Study of Acquiring-Firm Returns in the Recent Merger Wave, *The Journal of Finance*, 60(2), 757–782.

Moran, P., 2017, Information Revelation in Merger Waves, *The Review of Corporate Finance Studies*, 6(2), 174–233.

Moriarty, R., H. Ly, E. Lan, and S. K. McIntosh, 2019, Deal or No Deal: Predicting Mergers and Acquisitions at Scale, *IEEE International Conference on Big Data*, 5552–5558.

Mulherin, J. H., and A. L. Boone, 2000, Comparing acquisitions and divestitures, *Journal of Corporate Finance*, 6(2), 117–139.

Nelson, R. L., 1959, *Merger Movements in American Industry, 1895–1956*, Princeton University Press.

Netter, J., M. Stegemoller, and M. B. Wintoki, 2011, Implications of Data Screens on Merger and Acquisition Analysis: A Large Sample Study of Mergers and Acquisitions from 1992 to 2009, *The Review of Financial Studies*, 24(7), 2316–2357.

Ovtchinnikov, A. V., 2013, Merger waves following industry deregulation, *Journal of Corporate Finance*, 21, 51–76.

Palepu, K., 1986, Predicting Takeover Targets: A Methodological and Empirical Analysis, *Journal of Accounting and Economics*, 8(1), 3–35.

Piotroski, J. D., 2000, Value Investing: The Use of Historical Financial Statement Information to Separate Winners from Losers, *Journal of Accounting Research*, 38, 1–41.

Powell, R. G., 2004, Takeover Prediction Models and Portfolio Strategies: A Multinomial Approach, *Multinational Finance Journal*, 8(1/2), 35–72.

Powell, R. G., and A. Yawson, 2005, Industry Aspects of Takeovers and Divestitures: Evidence from the UK, *Journal of Banking and Finance*, 29(2), 3015–3040.

Qiu, B., and T. Wang, 2024, Corporate Mergers and Acquisitions Under Lender Scrutiny, Working Paper.

Ramirez, A. A., 2025, The Rise of Merger and Acquisitions in the US: Consequences for Investment, Market Concentration, and Profits, Working Paper.

Rhodes-Kropf, M., and S. Viswanathan, 2004, Market Valuation and Merger Waves, *The Journal of Finance*, 59(6), 2685–2718.

Rhodes-Kropf, M., D. T. Robinson, and S. Viswanathan, 2005, Valuation Waves and Merger Activity: The Empirical Evidence, *Journal of Financial Economics*, 77(3), 561–603.

Roll, R., 1986, The Hubris Hypothesis of Corporate Takeovers, *The Journal of Business*, 59(2), 197–216.

Routledge, B., S. Sacchetto, and N. Smith, 2013, Predicting Merger Targets and Acquirers from Text, Working Paper, Carnegie Mellon University.

Schwert, G. W., 1996, Markup pricing in mergers and acquisitions, *Journal of Financial Economics*, 41(2), 153–192.

Servaes, H., 1991, Tobin's Q and the gains from takeovers, *The Journal of Finance*, 46(1), 409–419.

Shleifer, A., and R. W. Vishny, 2003, Stock market driven acquisitions, *Journal of Financial Economics*, 70(3), 295–311.

Tsagkanos, A., A. Georgopoulos, and C. Siriopoulos, 2007, Predicting Greek Mergers and Acquisitions: A New Approach, *International Journal of Financial Services Management*, 2(4), 289–303.

Uysal, V. B., 2011, Deviation from the target capital structure and acquisition choices, *Journal of Financial Economics*, 102(3), 602–620.

Zhang, M., G. Johnson, and J. Wang, 2012, Predicting Takeover Success Using Machine Learning Techniques, *Journal of Business and Economics Research*, 10(10), 547–552.

Zhao, Y., X. Bi, and Q.-P. Ma, 2025, "Predicting mergers & acquisitions: A machine learning-based approach", *International Review of Financial Analysis*, 99, 103933

Table 1. Statistical Predictive Performance

Table 1 reports the out-of-sample classification metrics of the binary M&A target prediction models for the United States over the 1989 - 2025 period. The table contrasts a Logistic Regression baseline with the combined XGBoost model and reports, alongside the number of test rows that received a prediction, the standard classification metrics - Precision, Accuracy, Recall, F1, Jaccard, Matthews correlation coefficient and Balanced Accuracy - together with a likelihood-ratio χ^2 statistic and its empirical p-value. Each rolling window is fit on its own training portion, hyper-parameter tuned through an inner five-fold cross-validation, and applied to its held-out test fold; the metrics are then computed on the union of all 37 test folds. The likelihood-ratio statistic compares the model's predicted conditional probabilities against the no-information rate (the in-sample base rate predicted on every observation) so that a small p-value rejects the null that the model has no predictive content beyond random guessing.

Model	n Predictions	Precision	Accuracy	Recall	F1	Jaccard	MCC	Bal. Acc.	Y-stat	p-value
Logistic Regression	221,530	0.110	0.513	0.596	0.175	0.097	0.064	0.551	10,980.0	0.216
XGBoost (combined)	90,418	0.281	0.827	0.483	0.341	0.209	0.276	0.671	43,505.2	0.000

Table 2: Per wave-phase classification performance of the combined XGBoost

Table 2 reports F1, Precision, Recall and Balanced Accuracy of the combined XGBoost model on the test rows that fall inside each (Wave, Phase) cell of the 1989 - 2025 evaluation window. For every cell with phase span [start, end] the test predictions whose data_year falls in the interval are pooled and the four metrics are computed on the resulting subset; cells in which the resulting subset contains fewer than two distinct true classes are dropped. The final row averages the three Non-Wave segments. The pattern shows recall peaks in the Peak phases of every wave, precision peaks in Decline phases, and a Wave 1 Peak combination of high recall and low precision consistent with the dot-com era saturating the predictor space, whereas Non-Wave segments display the opposite signature of high precision and low recall.

Wave	Phase	F1	Precision	Recall	Bal. Acc.
Wave 1 (1991-2002)	Initiation	0.19	0.2	0.18	0.55
	Build-up	0.26	0.18	0.42	0.61
	Peak	0.23	0.13	0.82	0.64
	Decline	0.37	0.3	0.5	0.69
Wave 2 (2004-2008)	Initiation	0.35	0.25	0.56	0.69
	Build-up	0.36	0.26	0.58	0.7
	Peak	0.39	0.31	0.52	0.7
	Decline	0.34	0.24	0.55	0.68
Wave 3 (2010-2019)	Initiation	0.41	0.31	0.6	0.73
	Build-up	0.44	0.35	0.6	0.75
	Peak	0.43	0.37	0.51	0.71
	Decline	0.44	0.38	0.53	0.73
Wave 4 (2020-2025)	Initiation	0.4	0.32	0.51	0.71
	Build-up	0.32	0.26	0.43	0.66
	Peak	0.36	0.34	0.39	0.66
	Decline	0.35	0.29	0.47	0.69
Non-Wave (Average)	—	0.3	0.26	0.37	0.63

Table 3. Aggregate Cluster Importance Shares by Wave-Phase

Table 3 reports, for each (Wave, Phase) cell over 1989 - 2025, the within-cell importance shares of the three super-clusters Macro, Firm-specific and Behavioural for the combined XGBoost model, as well as the three Non-Wave segments. For every rolling window the 19 feature groups' relative importances are first aggregated into the three super-clusters; per-cell shares are then obtained by averaging across the windows whose re-estimation year falls inside the cell and dividing by the cell total, so that the three shares sum to one in every row. The macroeconomic share declines monotonically across waves while the behavioural share rises and the firm-specific share remains anchored near one half, providing the structural-versus-cyclical decomposition that motivates the rest of the paper.

Wave	Phase	Macro share	Firm share	Behavioural share
Wave 1 (1991-2002)	Initiation	0.435	0.502	0.064
	Build-up	0.425	0.488	0.088
	Peak	0.326	0.436	0.238
	Decline	0.355	0.476	0.169
Wave 2 (2004-2008)	Initiation	0.318	0.491	0.191
	Build-up	0.301	0.506	0.192
	Peak	0.285	0.496	0.219
	Decline	0.304	0.509	0.187
Wave 3 (2010-2019)	Initiation	0.234	0.465	0.301
	Build-up	0.238	0.481	0.280
	Peak	0.226	0.470	0.304
	Decline	0.217	0.486	0.297
Wave 4 (2020-2025)	Initiation	0.187	0.495	0.317
	Build-up	0.186	0.523	0.291
	Peak	0.217	0.487	0.296
	Decline	0.190	0.506	0.304
Non-Wave (1989-1990)	—	0.442	0.515	0.043
Non-Wave (2003)	—	0.311	0.489	0.200
Non-Wave (2009)	—	0.336	0.501	0.163

Table 4. Group-level importance rank profiles by wave-phase.

Table 4 reports, for each of the 19 feature groups and each (Wave, Phase) cell over the 1989 - 2025 evaluation window, the cell-level rank of the group ordered from highest mean importance (rank 1) to lowest (rank 19); the three Non-Wave segments are reported in dedicated rows. Cell-level group importance is the mean of the per-window relative importance shares across the windows whose re-estimation year falls in the cell, and the rank is taken within each cell. The table thus condenses the full 19 × 19 importance matrix into a rank order that highlights which groups dominate, rotate or recede across phases of each merger wave

Wave	Phase	Analyst Sentiment	Asset Structure	CEO Background	CEO Incentives	Cash Flow & Liquidity	Governance & Oversight	Growth & Expansion	Industry	Interest Rates	Leverage/Debt Structure	Market & Trading	Monetary Premia	Other Firm- specific	Output & Labor	Prices & Costs	Profitability	Recession & Uncertainty	Stock Market	Structural Change	Valuation
Non-Wave (1989-1990)		9	11	18	19	7	8	4	5	10	1	2	3	16	17	14	13	15	6	12	
Non-Wave (2003)		3	11	6	19	12	4	13	8	7	5	1	2	3	17	18	15	14	16	9	10
Non-Wave (2009)		3	12	7	19	14	9	10	6	11	4	2	1	18	16	13	14	15	17	9	8
Wave 1 (1991-2002)																					
	Build-up	8	11	13	19	5	7	4	10	3	1	2	6	17	18	16	15	14	5	6	12
	Decline	3	12	6	19	11	8	7	13	4	1	2	15	18	16	17	14	9	7	10	5
	Initiation	9	13	16	19	8	19	6	7	4	11	3	1	2	18	17	14	12	15	5	10
	Peak	1	14	11	19	6	19	5	9	10	8	4	3	2	17	18	15	13	16	7	12
Wave 2 (2004-2008)																					
	Build-up	2	9	6	19	13	4	9	5	10	8	12	3	1	17	18	14	15	16	11	7
	Decline	2	9	12	19	15	4	19	5	10	8	7	3	1	17	18	13	14	16	10	6
	Initiation	3	11	7	19	15	4	19	8	10	9	12	5	2	17	18	14	13	16	10	6
	Peak	2	9	4	19	14	4	19	8	10	7	12	3	1	17	18	13	15	16	11	5
Wave 3 (2010-2019)																					
	Build-up	6	11	1	7	13	19	4	12	9	10	7	5	2	17	18	15	14	16	9	8
	Decline	7	9	1	6	11	19	4	12	10	7	5	3	2	17	18	15	14	16	9	8
	Initiation	5	9	1	8	13	16	3	10	11	12	7	4	2	17	14	18	19	15	9	7
	Peak	5	11	1	7	16	19	4	11	10	12	6	3	2	17	18	15	14	16	9	8
Wave 4 (2020-2025)																					
	Build-up	3	13	5	9	11	14	7	10	8	11	6	4	2	17	18	15	13	16	9	8
	Decline	4	11	2	8	6	19	5	9	12	13	6	3	1	17	18	15	14	16	10	7
	Initiation	4	7	10	7	11	13	5	12	8	9	6	4	3	16	18	15	17	14	9	8
	Peak	9	9	2	6	11	13	4	14	10	5	5	3	1	18	19	16	15	17	10	8

Table 5. Variables in the top three predictors of any Wave-Phase)

Table 5 reports, for the combined XGBoost over 1989 - 2025, the variables that ever rank in the top three within any wave-phase of the 19-cell decomposition (4 waves × 4 phases plus 3 Non-Wave segments), together with the count and share of cells in which each variable attains a top-three position. Rankings are computed from the per-window normalized feature importance: windows are mapped to their wave-phase via the re-estimation year, the cell-level mean of each feature's normalized importance is taken across the cell's windows, and the three highest-ranked features per cell are retained. The four most persistent variables, the Piotroski F-score, the high-yield premium, the analyst summary index and the prior-deal count, together account for 49 of the 57 top-three positions and constitute the universal predictive core that operates across waves.

Variable	Cluster	in top-3 (of 19)	Share
Piotroski F-score	Firm – Other Firm-specific	19	100%
high-yield premium, negative difference	Macro – Monetary Premia	14	74%
analyst summary	Behavioural – Analyst Sentiment	8	42%
prior deal count of the CEO	Behavioural – CEO Background	8	42%
high-yield premium, positive difference	Macro – Monetary Premia	3	16%
8-quarter NI trend	Firm – Profitability	3	16%
zero_trades	Firm – Market & Trading	2	11%

Table 6. Cross-wave consistency of importance profiles

Table 6 documents the stability of the importance ordering across waves and the persistence of the top variables across consecutive phases of the same wave for the combined XGBoost over 1989 - 2025. **Panel A** reports phase-specific Spearman rank correlations between the 19-element importance vectors of every Wave A vs. Wave B pair (six pairs), together with the average correlation between the cross-segment-averaged Non-Wave profile and each phase's cross-wave-mean profile (rightmost column); a coefficient close to one indicates that the same groups dominate the phase across waves. Two-sided p-values are reported with the standard significance asterisks. **Panel B** reports the Jaccard overlap of the top-five variable-group sets between consecutive (Phase_i, Phase_{i+1}) transitions inside each wave; significance is based on a hypergeometric test of the overlap against the null of independent draws from the 19-group universe. The combination of high cross-wave Spearman correlations with high within-wave Jaccard persistence supports a "universal core plus wave-specific accents" interpretation of the importance ordering.

Panel A: Cross-Wave Rank Consistency (Spearman r)
Rows: phase | Cols: wave-pair (Wave 1 first, then sorted by Wave A then Wave B), plus average vs. Non-Wave. $r \approx 1$: same importance ordering in that phase across waves. Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Phase	1991-2002 vs. 2004-2008	1991-2002 vs. 2010-2019	1991-2002 vs. 2020-2025	2004-2008 vs. 2010-2019	2004-2008 vs. 2020-2025	2010-2019 vs. 2020-2025	vs. Non-Wave (avg)
Initiation	0.84***	0.49**	0.37	0.79***	0.65***	0.92***	0.91***
Build-up	0.85***	0.62***	0.48**	0.81***	0.71***	0.94***	0.89***
Peak	0.82***	0.65***	0.60***	0.81***	0.75***	0.96***	0.90***
Decline	0.97***	0.70***	0.68***	0.71***	0.72***	0.97***	0.87***

Panel B. Feature-Group Persistence Across Stage Transitions (Jaccard top-K)
Rows: stage transition within a wave | Cols: wave. Jaccard ≈ 1 : top-K variable groups largely preserved across the transition. Significance from hypergeometric test: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$.

Stage Transition	Wave 1	Wave 2	Wave 3	Wave 4
Initiation → Build-up	0.67**	1.00***	0.67**	1.00***
Build-up → Peak	0.67**	0.43	0.67**	0.67**
Peak → Decline	0.67**	0.43	0.67**	1.00***

Table 7. Industry-level Wave-Phase Performance

Table 7 reports F1 of the combined XGBoost separately within 14 industries × (Wave, Phase) cell over 1989 - 2025, plus a final column that averages each industry across the three Non-Wave segments. For every test row the SIC industry group is decoded from the one-hot industry fixed-effect columns and stored alongside the financial statement year, actual and predicted; cell-level F1 is then computed on the rows in that industry × cell intersection, dropping cells with fewer than two distinct true classes (left blank in the table). The pattern shows Decline phases delivering the highest mean F1 (0.384) and balanced accuracy (0.702) across industries, Peak phases the highest recall (0.556), Build-up phases the lowest cross-industry F1 dispersion (sd 0.078), and Non-Wave segments the lowest predictability (mean F1 0.308) - the differential predictability of phases is itself an empirical finding that supports the structural-cyclical decomposition.

Industry	Wave 1 (1991-2002)				Wave 2 (2004-2008)				Wave 3 (2010-2019)				Wave 4 (2020-2025)				Non-Wave (avg)
	Initiation	Build-up	Peak	Decline	Initiation	Build-up	Peak	Decline	Initiation	Build-up	Peak	Decline	Initiation	Build-up	Peak	Decline	
BusEq	0.21	0.25	0.22	0.38	0.33	0.36	0.39	0.33	0.41	0.45	0.43	0.45	0.37	0.30	0.36	0.35	0.29
Chems	0.14	0.30	0.24	0.34	0.25	0.33	0.29	0.31	0.39	0.38	0.42	0.46	0.55	0.28	0.44	0.38	0.26
Durbl	0.16	0.26	0.19	0.40	0.40	0.41	0.26	0.38	0.44	0.46	0.53	0.47	0.50	0.25	0.40	0.45	0.32
Enrgy	0.21	0.25	0.26	0.48	0.38	0.40	0.41	0.33	0.29	0.41	0.39	0.51	0.27	0.35	0.43	0.40	0.29
HLth	0.20	0.25	0.23	0.42	0.43	0.37	0.37	0.31	0.48	0.45	0.47	0.44	0.48	0.32	0.35	0.33	0.30
Manuf	0.18	0.23	0.23	0.35	0.38	0.35	0.40	0.35	0.42	0.43	0.43	0.42	0.46	0.37	0.34	0.35	0.31
Money	0.17	0.26	0.23	0.35	0.34	0.36	0.41	0.35	0.39	0.44	0.42	0.43	0.41	0.33	0.38	0.34	0.31
NoDur	0.21	0.24	0.24	0.38	0.31	0.36	0.38	0.32	0.42	0.47	0.45	0.48	0.37	0.24	0.20	0.37	0.30
Other	0.14	0.26	0.23	0.33	0.34	0.35	0.42	0.35	0.39	0.44	0.38	0.45	0.41	0.37	0.38	0.35	0.34
Pharma	0.23	0.27	0.22	0.41	0.32	0.30	0.43	0.37	0.46	0.49	0.46	0.45	0.28	0.23	0.41	0.41	0.22
Shops	0.19	0.27	0.23	0.31	0.33	0.38	0.36	0.34	0.40	0.44	0.43	0.41	0.33	0.34	0.35	0.34	0.29
Telcm	0.17	0.27	0.21	0.41	0.33	0.30	0.45	0.33	0.33	0.44	0.54	0.42	0.30	0.40	0.40	0.35	0.31
Utils	0.25	0.20	0.24	0.40	0.25	0.39	0.33	0.32	0.27	0.47	0.51	0.31	0.22	0.22	0.26	0.22	0.34
Banks		0.40	0.36	0.76		0.33				0.33	0.00	0.00				0.67	0.67

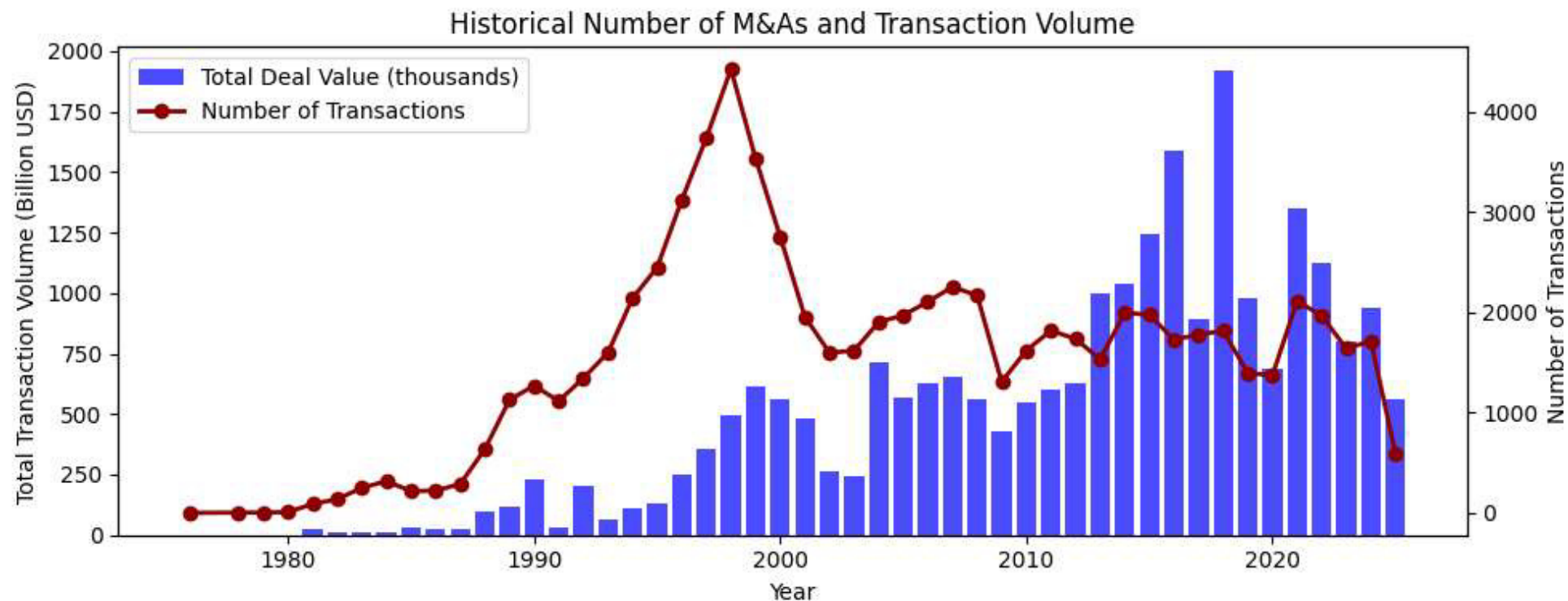


Fig. 1: Historical number of M&A transactions and transaction volume

Figure 1 displays the historical development of U.S. M&A activity over the full 1976 - 2025 sample period as recorded in the SDC Platinum database. Two series are plotted on a shared annual axis: the number of completed transactions that satisfy the inclusion criteria of Section 3. The figure motivates the rest of the paper by visualising the four post-1989 merger waves as activity clusters and the three intervening Non-Wave troughs (1989 - 1990, 2003, 2009)

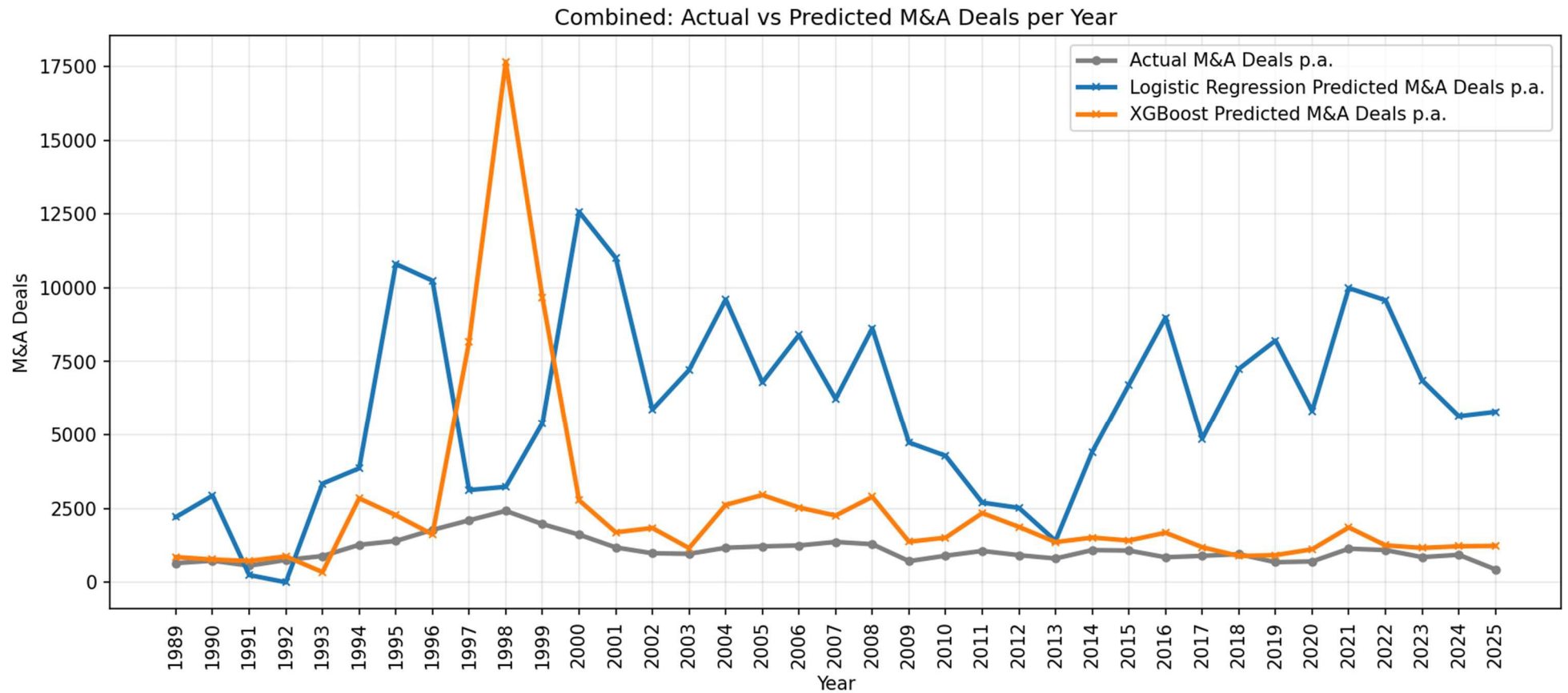


Fig. 2. Predicted vs. Actual M&A Deals

Figure 2 plots, for the United States, the actual annual number of completed M&A transactions (solid line) against the predicted number generated by the combined XGBoost model (dashed line) over the 1989 - 2025 out-of-sample period. Per-row binary predictions emitted by the 37 rolling windows are aggregated by data_year and compared to the corresponding annual deal counts, providing a visual diagnostic of aggregate fit at the year level that complements the row-level metrics in Table 1. Years before the first re-estimation receive no prediction and are shown only via the actual series.

Superiority Test per Wave: XGBoost vs. Logistic Regression
(opaque bars = significant at 5%, Wilcoxon paired test)

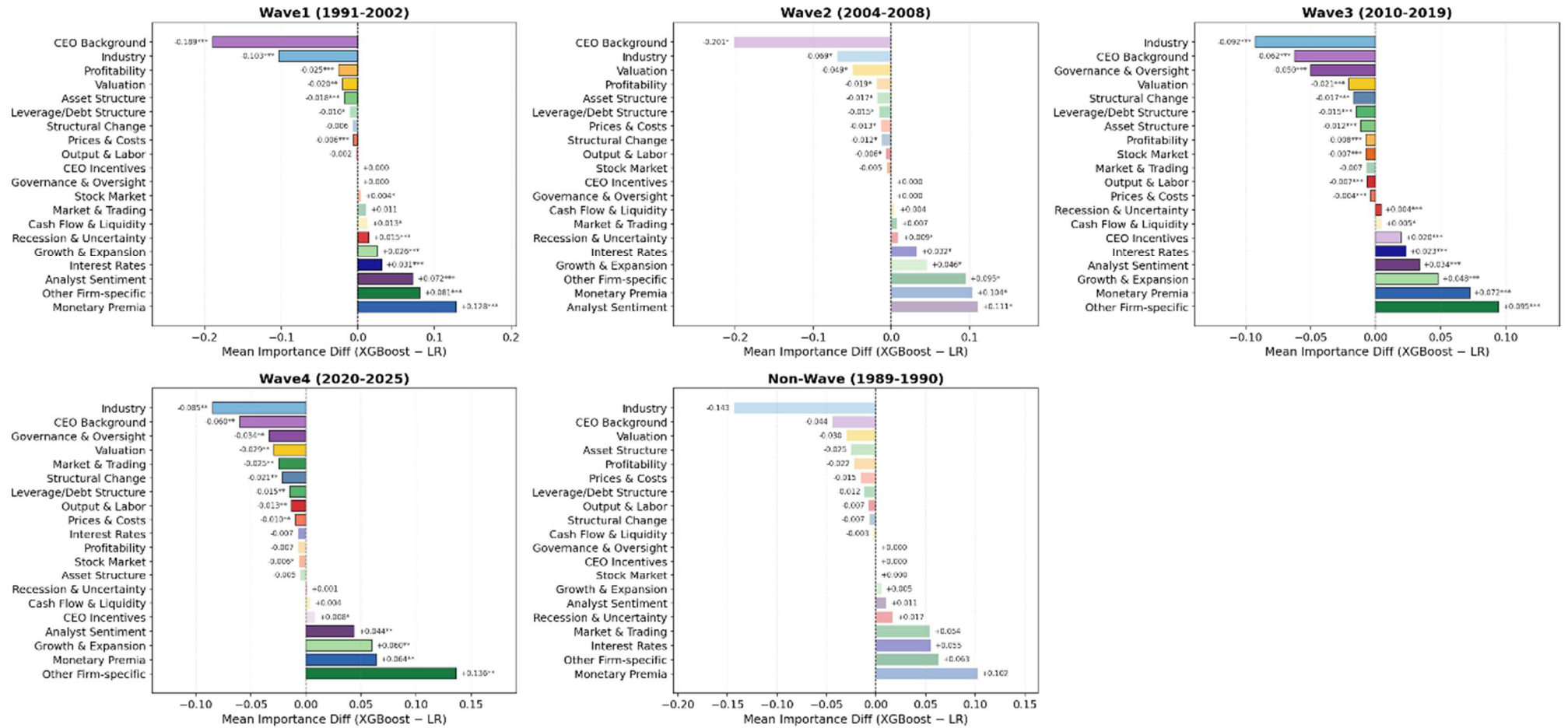


Fig. 3. Superiority of XGBoost over Logistic Regression, per wave

Figure 3 presents a faceted horizontal bar chart, with one panel per wave, of the difference in mean relative importance between the combined XGBoost and the Logistic Regression baseline computed over the rolling windows of that wave's 1989 - 2025 span. For each wave and each of the 19 feature groups, Mean_Diff = mean(XGB importance) - mean(LR importance) is computed across the wave's windows; bars are annotated with the numeric difference and the significance stars from a paired Wilcoxon signed-rank test of H_0 : equal mean importance across windows. Positive bars indicate groups to which the non-linear model assigns more importance than the linear baseline. The amplification is largest in Wave 1 and Wave 3 (the dot-com and tech/pharma consolidation waves) and smallest in Wave 2 (the credit-driven wave), consistent with the role of non-linear interactions in the behaviourally-charged waves. Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

CCF: Group Importance vs. M&A Activity per Wave
(only groups with observed values shown; positive = group leads M&A)

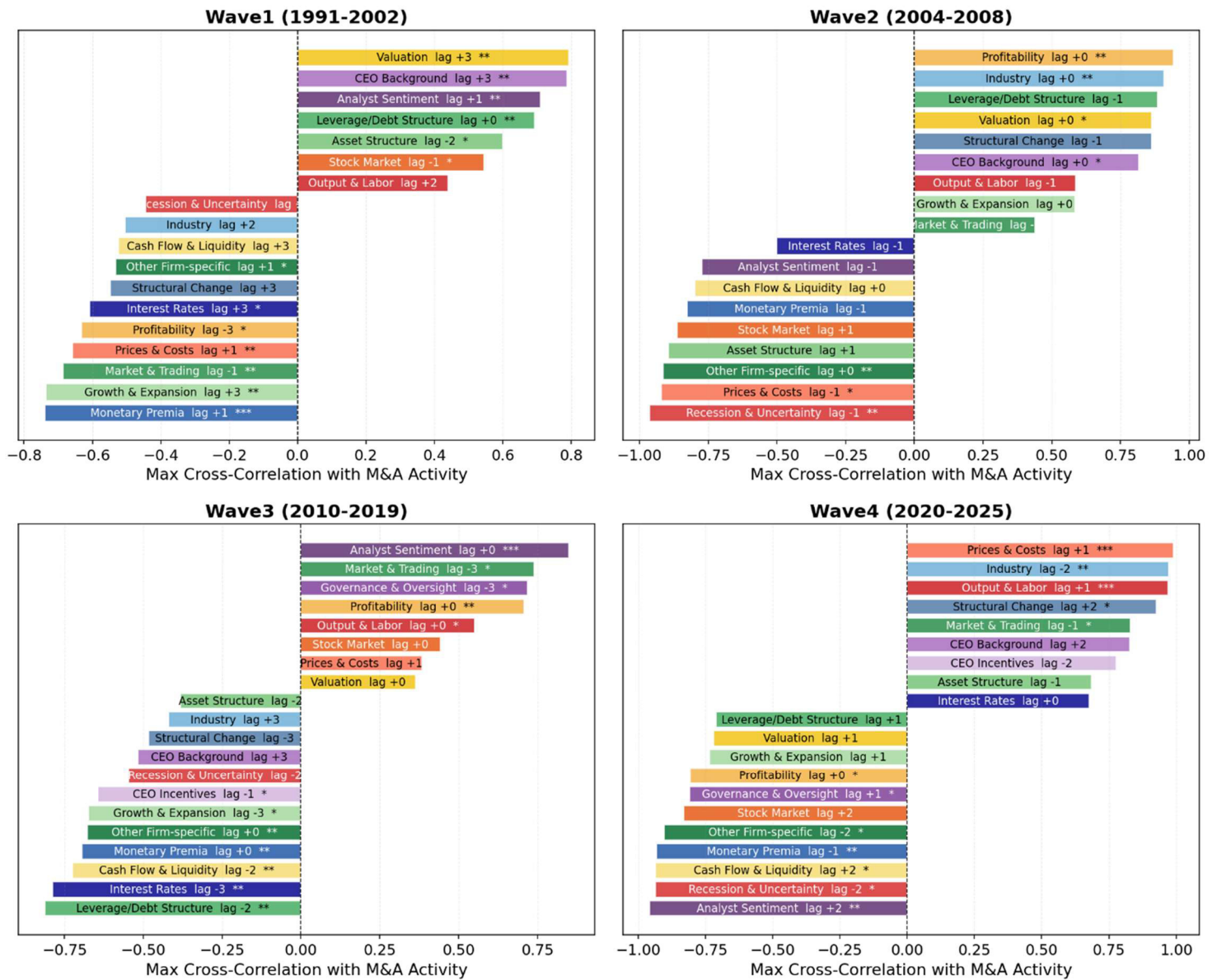


Fig 4. Cross-correlation of phase-level group importance and deal counts

Figure 4 reports, separately for each of the four merger waves over 1989 - 2025, the maximum cross-correlation between every group's phase-level relative importance series and the matching M&A deal count series, together with the lag at which that maximum is attained. Within each wave the per-phase importance vector of every group (one observation per phase, four points per wave) is aligned with the corresponding M&A activity per phase; for lags between $-\min(\text{maxlag}, \text{len}/2)$ and $+\min(\text{maxlag}, \text{len}/2)$ the Pearson correlation is computed on the lag-aligned overlap and the maximum-absolute-correlation lag is retained, with its two-sided p-value. Positive lags indicate that the group's importance leads M&A activity, negative lags that it follows. The figure provides the central evidence for the sequencing hypothesis: in Wave 4 the Structural Change group leads by two years ($r = +0.93$), Output and Labor leads by one year ($r = +0.97$) and Monetary Premia lags by one year ($r = -0.93$), with the same temporal ordering visible - though weaker in statistical power - in the earlier waves

Per-Group Recurring Onsets (XGBoost)
 markers = episode starts; shaded bands = episode duration

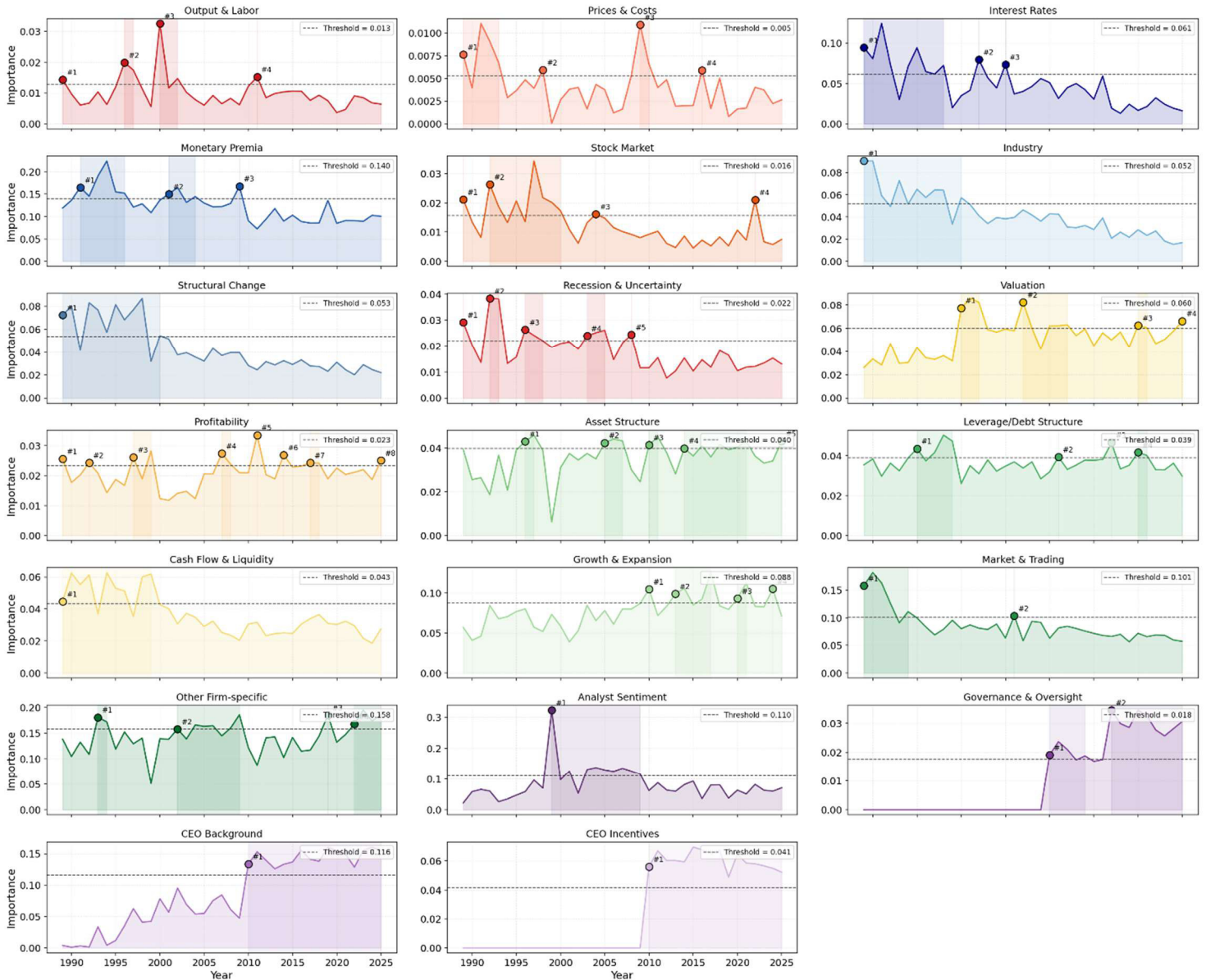


Fig.5. Recurring group-level onsets across waves

Figure 5 plots the within-wave activation onsets of every group's relative importance series over 1989 - 2025, with one panel per wave plus a combined panel for the three Non-Wave segments. For every wave the windows are restricted to the wave's year span; for each group the within-wave mean and standard deviation are computed and an episode is declared whenever the series crosses the threshold $\text{mean} + 0.5 \times \text{SD}$; contiguous above-threshold blocks separated by fewer than two below-threshold windows are merged. Each marker denotes the start of one such activation episode. The pattern is recurring rather than wave-idiosyncratic - Structural Change peaks first, Output and Labor second, Monetary Premia last - which is the within-wave temporal regularity that the cross-correlation analysis in Figure 4 also captures

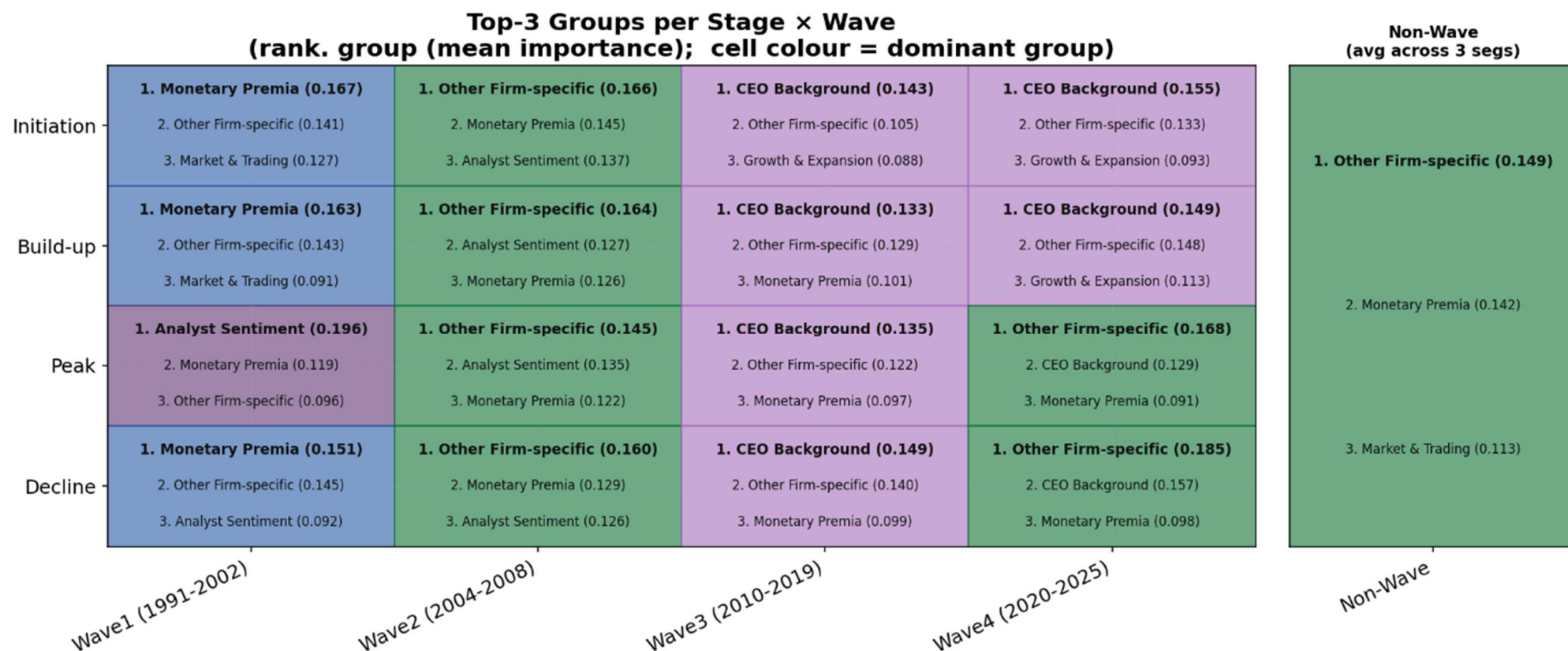


Fig. 6. Top-three dominant feature groups per wave-phase)

Figure 6 displays, for the combined XGBoost over 1989 - 2025, the three feature groups with the highest mean relative importance in every wave-phase) cell of the 19-cell decomposition; an additional inset shows the same for the three Non-Wave segments averaged. Cells are coloured by the rank-1 group's super-cluster (Macro, Firm-specific or Behavioural) and labelled with the group names and their within-cell shares. The figure makes the wave-level rotation visible: Wave 1 is Monetary-Premia/Analyst-Sentiment dominated, Wave 2 is Other-Firm-specific dominated, Wave 3 is CEO-Background dominated, and Wave 4 transitions from CEO-Background dominance back toward Other-Firm-specific dominance, while the Other Firm-specific group remains in the top three of nearly every cell.

Industry Lead-Lag within Waves

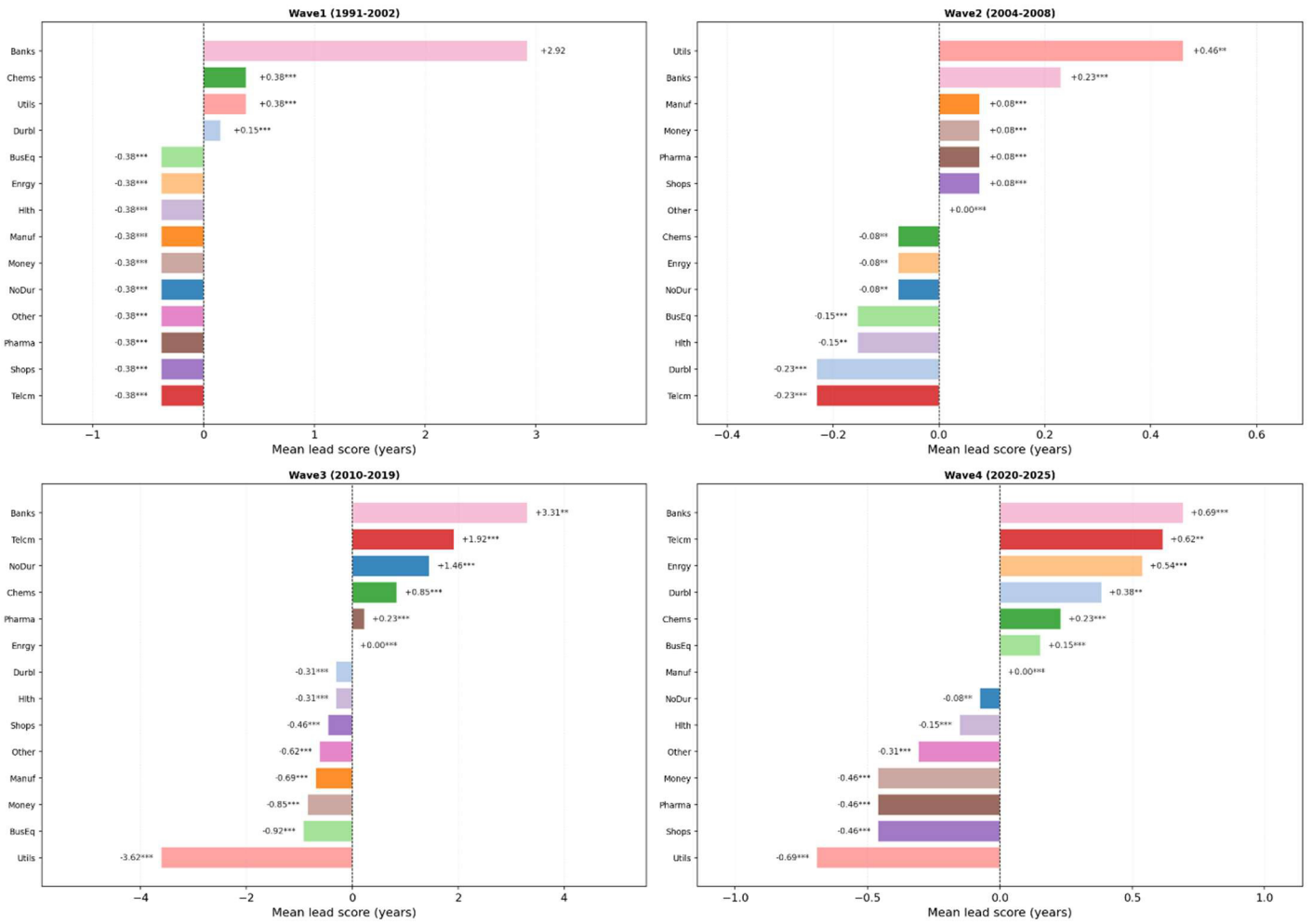
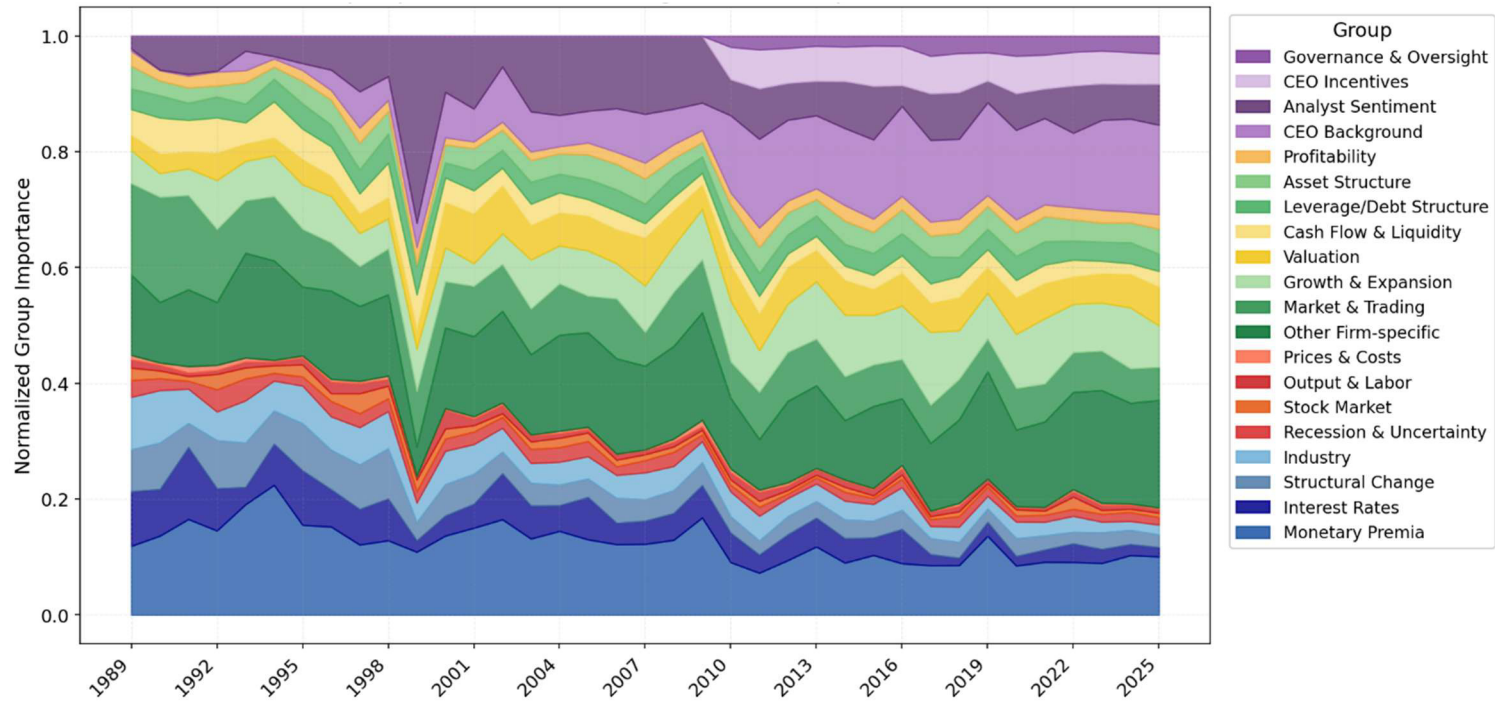


Fig. 7. Industry cross-correlation lag-at-peak

Figure 7 reports the lag at which the maximum cross-correlation between every industry's yearly importance contribution and aggregate U.S. M&A deal counts is attained inside each wave over 1989 - 2025. Per-observation industry contributions are derived from the log-likelihood test of the combined XGBoost, each interaction column is split evenly across all sic_group tokens it contains, and averaged by the financial statement year to produce per-industry annual series. Within each wave, the industry series and the M&A activity series are restricted to the wave's years, and the $|r|$ -maximising lag is selected over $[-\min(\text{maxlag}, \text{len}/2), +\min(\text{maxlag}, \text{len}/2)]$. Negative lags indicate that the industry's contribution leads aggregate deal counts, positive lags that it lags; the figure shows wave- and industry-specific patterns (Healthcare/Pharma leading in Wave 3, Technology coincident with the Peak, Energy lagging by one year; Manufacturing and Wholesale leading in Wave 2).

Panel A: Group importance over the whole sample period



Panel B: Group importance over time per wave

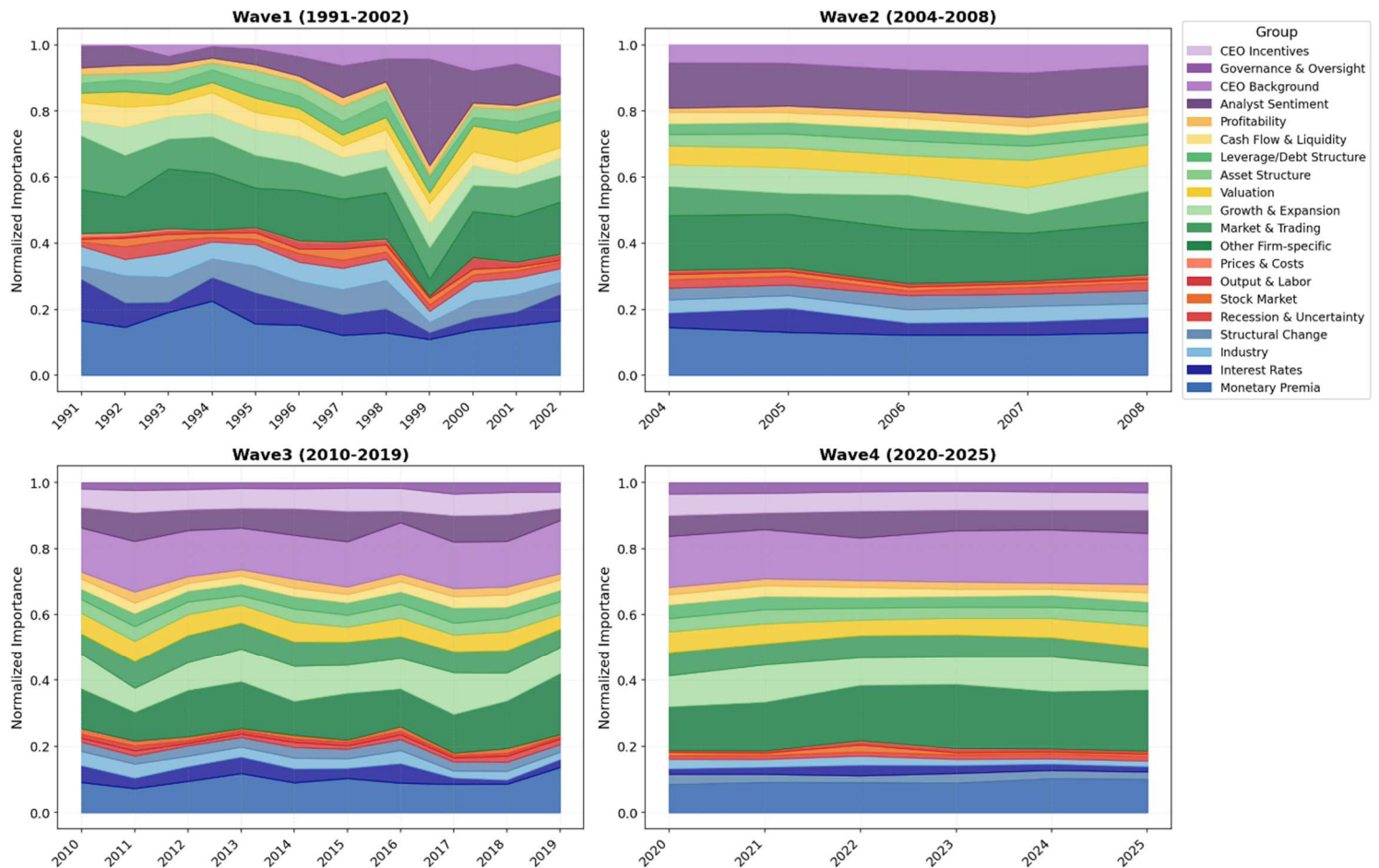


Fig. 8. Differential importance over time and per wave

Figure 8 presents two complementary views of the temporal evolution of group-level importance for the combined XGBoost over 1989 - 2025. Panel A is a stacked-area chart of the 19 group importance shares over the 37 rolling windows, layered Macro at the bottom, Firm-specific in the middle and Behavioural on top, with within-tier ordering by the mean importance of each group; values sum to one in every window by construction. The Panel B reproduces the same stacked-area chart restricted to the years of each wave separately, providing a per-wave view of compositional change. The behavioural cluster's importance share is approximately 1.73 times higher in wave years than in non-wave years on average, with the wave-versus-Non-Wave-2009 same-era control yielding a ratio of 1.84, ruling out era-specific compositional drift; the amplification is largest in Wave 1 and Wave 4 and smallest in Wave 2.

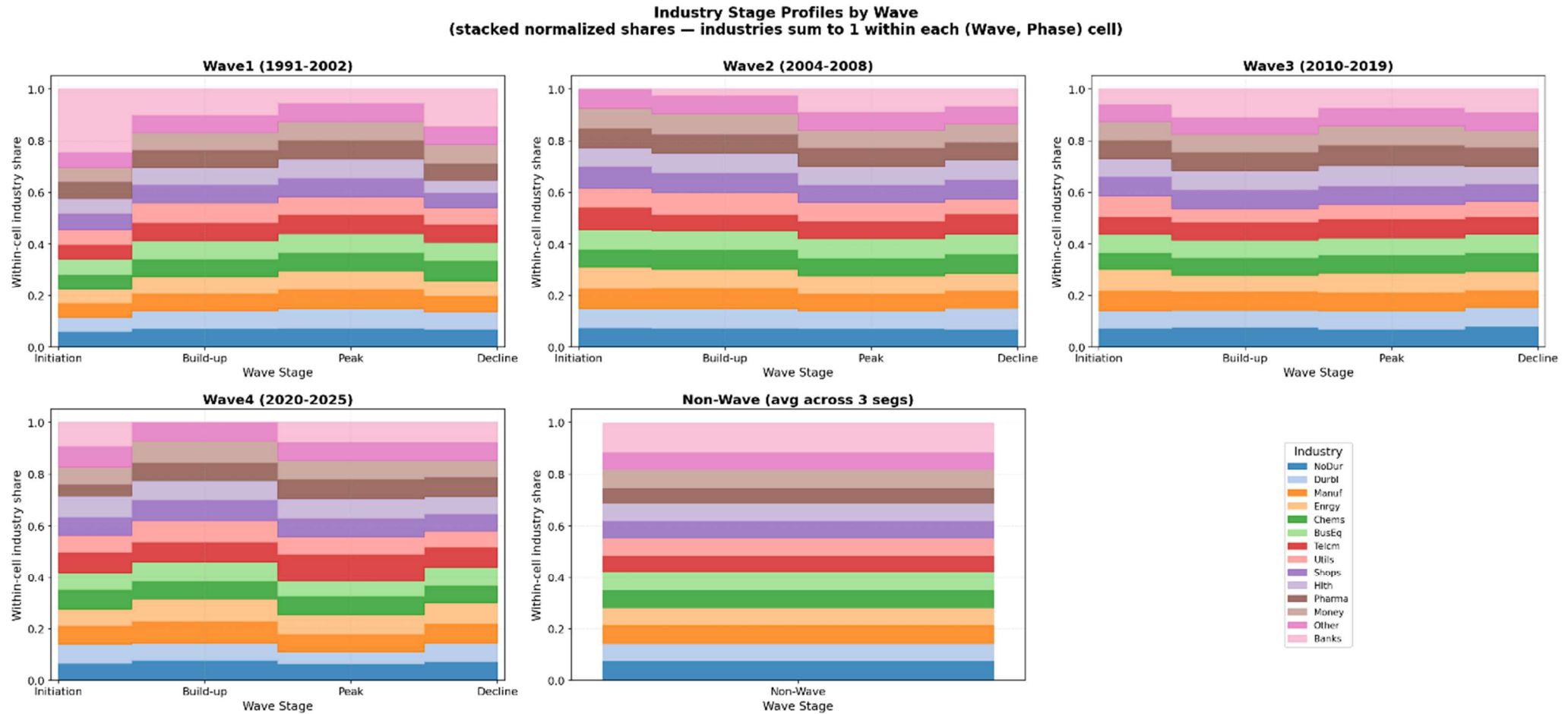


Fig. 9. Industry-stage importance profiles.

Figure 9 displays, for the combined XGBoost over 1989 - 2025, the within wave-phase industry composition of relative importance. The figure is laid out as a 2×3 grid: one stacked panel per real wave (four), one panel for the three Non-Wave segments averaged together, and one panel hosting the shared legend; within each panel rows are phases, columns are industries, and industry shares are normalized to sum to one per cell so that each column reflects within-cell composition rather than absolute magnitude. The underlying values come from the log-likelihood-derived per-observation industry contributions, aggregated to year-level and then averaged across the years that fall inside each wave-phase. The wave-level stage rotation documented in Figure 6 is preserved at the industry level for the majority of industries, with the strongest exceptions in Utilities and Energy where the macro cluster remains dominant in every phase.

Industry Lead-Lag with M&A Activity per Wave (industry yearly importance vs. M&A activity series)

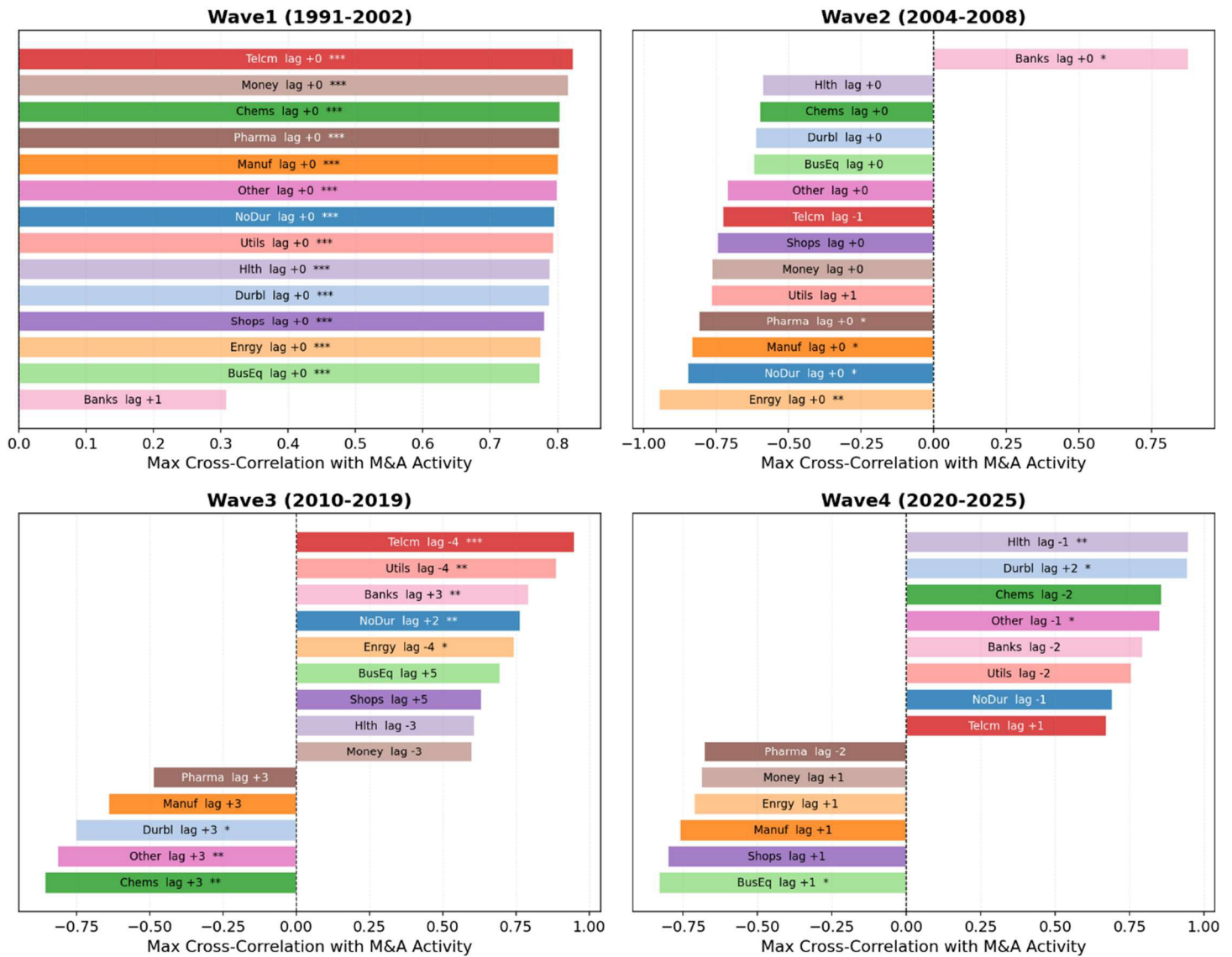


Fig. 10. Within-wave industry-level cross-correlations of selected groups

Figure 10 reports, for each wave-industry combination over 1989 - 2025, the lag-at-peak cross-correlation between the industry's contribution to two specific feature groups (Output and Labor and Behavioural Sentiment) and aggregate U.S. M&A deal counts. The computation is identical to that of Figure 7 but applied at the group $\times$ industry level rather than at the industry-aggregate level; the figure restricts the display to the two groups for which the sequencing pattern is most visible. The strongest leading patterns occur in Wave 4, with Healthcare leading by one year on Output and Labor and IT leading by one year on Behavioural Sentiment, providing industry-level corroboration of the aggregate sequencing finding.

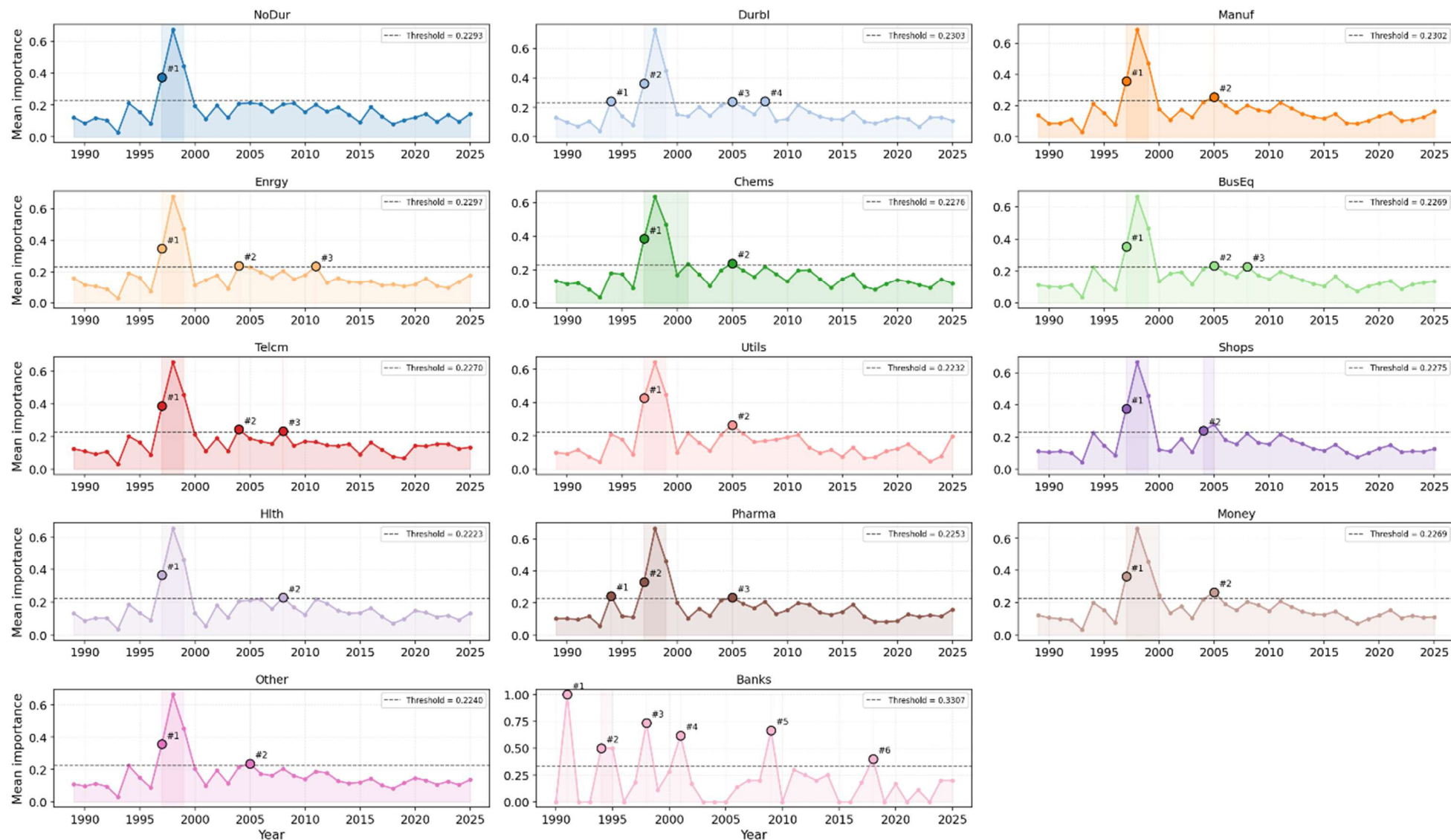


Fig. 11. Industry Onsets

Figure 11 plots, separately for each Fama-French twelve industry, the year-level series of the industry's mean relative importance over 1989 - 2025 together with the activation threshold mean $+ 0.5 \times \text{SD}$ computed within the industry's own series, and marks the year of each threshold crossing as an episode start. Contiguous above-threshold blocks separated by fewer than two below-threshold years are merged into a single episode and shaded as a coloured band. The figure shows when each industry becomes a salient driver of M&A activity and reveals the recurring pattern in which industries activate in a temporal sequence that mirrors the aggregate group-level sequencing of Figure 5.

Appendix A

Table A1: Descriptive Statistics Macroeconomic factors

This table gives the definitions, time series number of observations (count), mean, standard deviation (std) and median for the macro-economic variables used in the empirical analysis. The sample includes annual U.S.-data during the 1976 – 2025 period.

Variable		Abbreviation	Description	Source	Count	Mean	Std	Median
<i>Output and Labor</i>								
1	GDP	macro_Diff_GDP	Difference of the US GDP between t-1 and t0	Bureau of Economic Analysis	499,598	88.38	23.50	82.00
2	Industrial Production Index	macro_Diff_IndProd	Difference of the US Industrial Production Index between t-1 and t0	IPB50001A; St. Louis Fed's FRED database	499,550	1.64	3.22	2.86
3	Total Capacity Utilization	macro_Diff_TCU	Difference of the US Tital Capacity Utilization between t-1 and t0	TCU; St. Louis Fed's FRED Database	499,550	0.02	2.59	0.52
4	Unemployment Claim	macro_Diff_UNEMPLOY	Difference of the US Unemployment Claim between t-1 and t0	UNEMPLOY; Nuerau of Labour Statistics of the US Government	499,550	-122.00	1,574.08	-386.00
5	Skewness Unemployment	macro_Diff_skew_emp	Difference of the Kelley Skewness of US companies between t-1 and t0	Compustat	496,202	0.00	0.12	-0.01
6	Chicago Fed Index	macro_Diff_ChicagoFED	Difference of the Chicago Fed Index between t-1 and t0	CFNAI; St. Louis Fed's FRED database	499,604	0.02	0.70	0.01
<i>Prices and Costs</i>								
7	Purchase Price Index	macro_Diff_PPI	Difference of the US Purchase Price Index between t-1 and t0	PPIACO; St. Louis Fed's FRED database	499,604	3.87	8.64	3.40
8	Inflation	macro_Diff_Inflation	Difference of the US Comsumer Price Index between t-1 and t0	FPCPITOTLZGUSA; St. Louis Fed's FRED database	499,604	-0.04	1.43	-0.08
9	Building Permits	macro_Diff_BuildingPermits	Difference of the US Building Permits between t-1 and t0	PERMIT; St. Louis Fed's FRED database	499,604	11.15	227.57	46.00
10	Dollar Index	macro_Diff_Dollar	Difference of the US Dollar Index between t-1 and t0	IR; St. Louis Fed's FRED database	496,202	1.43	5.93	2.30
11	Oil Price Index	macro_Diff_OilPrice	Difference of the Import Price Index Cruide Oil between t-1 and t0	IR10000; St. Louis Fed's FRED database	498,926	6.73	54.03	10.00
<i>Interest Rates</i>								
12a	Fed Funds Rate	macro_FedFR	Difference of the Federal Funds Effective Rate between t-1 and t0	RIFSPFFNA; St. Louis Fed's FRED database	499,604	0.00	1.32	0.04

12b	Fed Funds Rate Positive Development	macro_FedFR_Positive	Dummy (1 if Fed Funds Rate Difference positive, 0 otherwise)	RIFSPFFNA; St. Louis Fed's FRED database	499,604	0.52	0.50	1.00
12c	Fed Funds Rate Negative Development	macro_FedFR_Negative	Dummy (1 if Fed Funds Rate Difference negative, 0 otherwise)	RIFSPFFNA; St. Louis Fed's FRED database	499,604	0.48	0.50	0.00
13	1 Year Rate	macro_1YrRet	Difference of the market yield on US Treasury Securities at 1-Year constant maturity between t-1 and t0	RIFLGFCY01NA; St. Louis Fed's FRED database	499,604	-0.46	6.82	-0.77
14a	10 Year Rate	macro_10YrRet	Difference of the 10-Year Treasury Constant Maturity Minus Federal Funds Rate between t-1 and t0	T10YFF; St. Louis Fed's FRED database	499,604	0.23	6.70	-0.92
14b	10 Year Rate Positive Development	macro_10YrRet_Positive	Dummy (1 if 10 Year Rate Difference positive, 0 otherwise)	T10YFF; St. Louis Fed's FRED database	499,604	0.28	0.45	0.00
14c	10 Year Rate Negative Development	macro_10YrRet_Negative	Dummy (1 if 10 Year Rate Difference negative, 0 otherwise)	T10YFF; St. Louis Fed's FRED database	499,604	0.72	0.45	1.00
15	Money Supply M2	macro_Diff_m2	Difference in the m2 money supply.	WM2NS; St. Louis Fed's FRED database	498,926	486.44	720.62	289.30
<i>Monetary Premia</i>								
16a	Term Premium	macro_Diff_tms	Difference of the long-term yield on government bonds and the T-Bill between t-1 and t0	Goyal and Welch (2008)	490,218	0.00	0.01	0.00
16b	Term Premium Positive Development	macro_Diff_tms_Positive	Dummy (1 if the difference of the term premium positive, 0 otherwise)	Goyal and Welch (2008)	499,604	0.42	0.49	0.00
16c	Term Premium Negative Development	macro_Diff_tms_Negative	Dummy (1 if the difference of the term premium negative, 0 otherwise)	Goyal and Welch (2008)	499,604	0.56	0.50	1.00

17	Default Yield	macro_dfy	Difference of the BAA and AAA-rated corporate bond yields between t-1 and t0	Goyal and Welch (2008)	490,218	0.01	0.00	0.01
18a	High Yield Premium	macro_hyp	US High Yield Index Effective Yield	BAMLH0A0HYM2EY; St. Louis Fed's FRED database	389,641	8.60	2.82	8.29
18b	High Yield Premium	macro_hyp_Positive	Dummy (1 if High Yield Premium positive, 0 otherwise)	BAMLH0A0HYM2EY; St. Louis Fed's FRED database	499,604	0.78	0.41	1.00
18c	High Yield Premium	macro_hyp_Negative	Dummy (1 if High Yield Premium negative, 0 otherwise)	BAMLH0A0HYM2EY; St. Louis Fed's FRED database	499,604	0.00	0.00	0.00
19	Ted Spread	macro_Diff_Ted	Spread between 3-month LIBOR and T-Bills	TEDRATE; St. Louis Fed's FRED database	402,588	0.00	0.38	0.01
<i>Stock Market</i>								
20	S&P 500 Index	macro_Diff_sp500	Difference of the S&P 500 Index between t-1 and t0	sprtrn; St. Louis Fed's FRED database	499,598	0.00	0.05	0.00
21	Dividend Yield	macro_Diff_dy	Difference between the log of dividends and the log of lagged prices	Goyal and Welch (2008)	490,218	-0.05	0.15	-0.05
22	Earnings to Price Ratio	macro_Diff_ep	Difference between the log of earnings and the log of prices, where earnings are 12-month moving sums of earnings on the S&P 500 index	Goyal and Welch (2008)	490,218	-0.02	0.32	-0.03
23	Stock Variance	macro_svar	Sum of squared daily returns on the S&P 500 index	Goyal and Welch (2008)	490,218	0.00	0.00	0.00
24	IPO Aggregated Proceeds	macro_Diff_Aggregate_Proceeds	Annual aggregated US- IPO proceeds.	J. Ritter (2025)	498,926	1.48	22.67	2.25
25	Share Repurchases	macro_ShareRepurchases	Share repurchases of the nonfinancial corporate business	SRPSABSNNCB; St. Louis Fed's FRED database	499,604	4.23	34.49	0.22
<i>Industry</i>								
26	Signal Momentum	macro_signal_mom	Dummy variable equal to 1 if the industry's cumulative 12-month return is non-negative, 0 otherwise.	CRSP; Bessler and Wolff (2024)	499,550	0.28	0.45	0.00
27	Signal Moving Average	macro_signal_ma	Dummy variable equal to 1 if the industry's return is above its 12-month moving average, 0 otherwise.	CRSP; Bessler and Wolff (2024)	499,550	0.25	0.43	0.00

28	Signal Relative Strength	macro_signal_rs	Dummy variable equal to 1 if the industry's 6-month cumulative return exceeds the market's 6-month cumulative return, 0 otherwise.	CRSP; Bessler and Wolff (2024)	499,550	0.34	0.47	0.00
29	Market to Book Value	macro_Diff_at_me	Market-to-book value of the industry within one year.	CRSP	499,604	-45.61	410.42	-58.62
30	Q Dispersion	macro_a_q_dispersion	Industry-level q-dispersion between the 90 th and 10 th percentile	CRSP	499,598	3.98	5.52	2.79
<i>Structural Change</i>								
31	Industry Regulation	macro_deregulation	Dummy variable equal to 1 of a deregulation event occurred during the year, 0 if no such event occurred.	Ovtchinnikov (2013)	499,604	0.02	0.15	0.00
32	HHI Industry Concentration Index	macro_hhi_industry	Herfindahl–Hirschman Index (HHI) of sales concentration within an industry-year, calculated as the sum of squared firm sales shares × 10,000.	Grullon et al. (2019)	499,574	288.23	272.54	181.41
33	HHI overall economy wide concentration index	macro_hhi_agg_weighted	Industry-level HHI weighted by each industry's share of total yearly sales, capturing aggregate market concentration in a given year.	Grullon et al. (2019)	499,574	19.79	13.96	15.05
34	R&D	macro_Diff_RD	Difference in US-R&D Expenditures between t-1 and t0	National Center for Science and Engineering Statistics	499,598	21.67	20.16	15.00
35	Patents	macro_Diff_PATENTS	Difference in US-patent applications between t-1 and t0	USPTO	499,598	15.41	10.32	15.00
36	Disruptive Innovation	macro_Diff_INNOVATION	Difference in the Global Innovation Index between t-1 and t0	Global Innovation Index	499,598	-0.20	2.02	1.00
<i>Recession and Uncertainty</i>								
37	Enomoic Policy Uncertainty Index	macro_Diff_PolicyUncertainty	Economic Poicy Uncertainty Index for the US (Baker et al., 2012)	USEPUINDXM; St. Louis Fed's FRED database; Baker et al. (2016)	490,888	0.39	34.87	0.45
38	Monetary Uncertainty Index	macro_Diff_MonetaryUncertainty	Economic Policy Uncertainty Index: Categorical Index: Monetary policy (Baker et al., 2012)	EPUMONETARY; St. Louis Fed's FRED database; Baker et al. (2016)	490,888	0.76	63.81	5.93

39	NBER Recession Index	macro_NBER_Rec	Dummy variable equal to 1 if a recession was present, 0 if otherwise.	FED	499,604	0.03	0.18	0.00
40	Financial Uncertainty Index	macro_Diff_FinancialUncertainty	Economic Policy Uncertainty Index: Financial Regulation	EPUFINREG; St. Louis Fed's FRED database; Baker et al. (2016)	490,888	-6.89	118.94	-11.46
41	Recession Probability	macro_Diff_Rec_prob	Difference in the probability of US Recession, twelve months ahead of term spread readings	FED	499,604	0.02	0.15	0.02

Table A2: Descriptive Statistics Firm-specific characteristics

This table gives the definitions, time series mean, median and standard deviation for the firm-specific characteristics used in the empirical analysis before imputing and scaling. The sample includes annual U.S.-data during the 1976 – 2025 period contributed to WRDS from Jensen et al. (2023), Compustat and CRSP and, if currency-related, is denominated in USD.

Variable		Abbreviation	count	mean	std	median
<i>Valuation</i>						
42	Assets to market value	a_at_me	496,450	1.72	7.94	0.85
43	Book-to-market value	a_be_me	479,261	0.55	2.22	0.40
44	EBIT-to-market equity ratio	a_ebit_me	494,022	0.03	3.19	0.06
55	Tobins Q	a_tobins_q	496,370	4.24	67.86	1.41
46	Market equity	a_me	499,548	6,580.87	36,786.07	681.40
<i>Profitability</i>						
47	Earnings Variability	a_earnings_variability	350,933	1.19	2.13	0.81
48	Profit margin	a_ebit_sale	486,032	-39,288,076,550.48	27,390,073,596,667.80	0.09
49	Return on equity	a_ni_be	476,865	-0.01	34.63	0.09
50	ROA 5-year Change	a_roa_ch5	354,559	0.08	7.78	0.00
51	ROE Volatility	a_roe_be_std	344,189	0.33	5.70	0.07
<i>Asset Structure and Operations</i>						
52	Capital Expenditures scaled by Assets	a_capx_at	488,406	0.06	3.02	0.04
53	Current Liabilities to Total Liabilities	a_cl_lt	489,228	0.53	0.29	0.51
54	Days inventory Outstanding	a_inv_days	479,456	110.49	1,481.19	44.96
55	Operating Leverage	a_opex_at	493,523	0.96	3.20	0.73
56	Property, Plants and Equipment scaled by Market Equity Value	a_ppen_mev	487,020	0.29	5.84	0.13
57	Receivables Change 1Yr	a_rec_gr1a	469,719	0.01	1.10	0.01
58	Asset Tangibility	a_tangibility	474,597	0.63	0.52	0.64
<i>Leverage and Debt Structure</i>						
59	Net Debt issuance	a_dbnetis_at	493,583	-0.08	16.75	-0.01
60	Debt-to-Assets	a_debt_at	496,441	0.26	6.73	0.19
61	Long-Term Debt Scaled by Total Debt	a_debtl_t_debt	437,838	0.75	0.31	0.89
62	Interest Coverage Ratio	a_ebit_int	431,143	49.68	1,673.80	4.44
63	Altman z-score	a_z_score	461,997	8.14	132.46	3.42
<i>Cash Flow and Liquidity</i>						
64	Liquidity of Book Assets	a_aliq_at	470,661	2.65	317.98	0.67

65	Cash Ratio	a_cash_cl	489,074	2.09	24.28	0.49
66	Working Capital Scaled by Assets	a_nwc_at	486,693	0.23	7.98	0.22
67	Free Cash Flow Change 1Yr	a_fcf_gr1a	461,562	0.01	1.04	0.00
68	Total Dividends Scaled by Assets	a_div_at	491,494	0.02	0.19	0.00
<i>Growth and Expansion</i>						
69	Intangible Assets Change 1Yr	a_intan_gr1a	430,523	0.01	0.21	0.00
70	Change in Long-Term Investments	a_lti_gr1a	435,272	0.00	0.09	0.00
71	Net Income Change 1Yr	a_ni_gr1a	469,147	0.01	1.24	0.01
72	Number of Consecutive Quarters With Earnings Increase	a_ni_inc8q	384,908	3.02	3.17	2.00
73	Sales Growth 1Yr	a_sale_gr1	461,439	0.79	74.66	0.10
<i>Market and Trading</i>						
74	Market Beta	a_beta_252d	465,154	0.97	0.63	0.94
75	Net Stock Issues	a_chcsho_12m	455,179	0.44	66.63	0.01
76	Market Correlation	a_corr_1260d	370,509	0.38	0.19	0.39
77	Equity Buyback Scaled by Market Equity	a_eqbb_me	458,967	0.02	0.19	0.00
78	Price	a_prc	499,548	36.42	87.46	19.81
79	Momentum 0-12 Months	a_ret_12_0	457,371	0.26	1.13	0.12
80	Number of Shares Traded During Month	a_tvol	498,604	21,173,133.14	81,843,175.46	3,255,400.00
81	Number of Zero Trades With Turnover as Tiebreaker (12 Months)	a_zero_trades_252d	475,328	0.59	2.14	0.00
<i>Other Firm-Specific Factors</i>						
82	Firm Age	a_age	499,604	247.43	213.09	180.00
83	Pitroski F-score	a_f_score	417,103	5.09	1.71	5.00
84	R&D Scaled by Assets	a_rd_at	269,921	0.09	0.84	0.04
85	Assets Turnover	a_sale_bev	477,802	4.09	104.64	1.58
86	Employees	a_emp	484,328	12.96	45.21	2.00

Table A3: Descriptive Statistics Behavioural characteristics

This table gives the definitions, time series mean, median and standard deviation for the firm-specific characteristics used in the empirical analysis before imputing and scaling. The sample includes annual U.S.-data during the 1976 – 2025 period contributed to WRDS from Jensen et al. (2023), Compustat and CRSP and, if currency-related, is denominated in USD.

Variable		Abbreviation	Description	Source	Count	Mean	Std	Median
<i>Analyst Sentiment</i>								
88	Surprise Index	Surprise_index	ibes_sue + ibes_sue_pos_share; higher = more positive earnings-surprise environment vs peers (monthly cross-sectional z-score).	IBES Surprise	499,604	0.00	0.01	0.00
89	Momentum Index	Momentum_index	ibes_mom_rev_ratio; higher = stronger upward analyst-revision pressure (monthly cross-sectional z-score).	IBES Momentum	234,976	0.00	1.01	0.00
90	Pricetarget Index	Pricetarget_index	ibes_ptg_upside - ibes_ptg_dispersion; higher = strong optimism with low disagreement (monthly cross-sectional z-score).	IBES Price Target	312,413	-0.03	1.02	0.37
91	Summary Index	Summary_index	ibes_revision_ratio - ibes_dispersion; higher = upward revisions with low disagreement (monthly cross-sectional z-score).	IBES Summary	499,585	-0.01	1.02	0.12
<i>Governance and Oversight</i>								
92	Shareholder dissent Index	a_shareholder_dissent_index	Mean of against_pct = votedwithheld / (for + against + withheld) across proposals per firm-year, z-scored across firms.	ISS Voting Analytics	231,849	-0.20	0.85	-0.51
93	Monitoring Index	a_monitoring_index	Negative mean z-score of poison pill, dual class, blank check, classified board, supermajority %, unequal voting, fair price (G-index in legacy period); higher = stronger monitoring.	ISS legacy governance, 1990–2006 ISS governance, 2007–2025	119,080	0.20	0.89	0.42
	Duality Flag	a_duality_flag	1 if CEO is also board chair, 0 otherwise.	Execucomp	193,251	0.46	0.50	0.00
	Ownership	a_shown_tot_pct	Total CEO share ownership in %.	Execucomp	193,251	1.26	4.59	0.02
<i>CEO Background</i>								
94	CEO Tenure	a_ceo_tenure	Years since became CEO (year - becameceo year).	Execucomp	193,251	6.85	7.22	5.00

95	Promotion speed	a_promotion_speed_years	Years from first job to first CEO appointment.	BoardEx, SDC Platinum	193,251	14.37	11.34	15.00
	Sector experience	a_sector_breadth_pre_ceo	Count of distinct sectors worked in before becoming CEO.	BoardEx	193,251	1.53	1.44	1.00
	Organisation Type experience	a_orgtype_breadth_pre_ceo	Count of distinct organisation types before becoming CEO.	BoardEx	193,251	1.29	0.96	1.00
	International experience	a_international_experience_pre_ceo	Count of international roles before becoming CEO.	BoardEx	193,251	0.15	0.35	0.00
<i>Incentives</i>								
96	Risk incentive index	a_risk_incentive_index	Z-score of equity_ratio = (option_awards_blk_value + stock_awards_fv) / tdc1.	Execucomp	190,390	0.07	1.02	0.06
97	CEO Age	a_ceo_age	CEO age in years.	Execucomp	193,251	52.04	15.58	55.00
98	CEO Compensation in options	a_option_awards_blk_value	Black-Scholes value of CEO option awards.	Execucomp	193,251	1,302.75	7,535.10	0.00
99	CEO Compensation in %	Ceo_comp_%	Total CEO compensation (tdc1) as %.	Execucomp	193,251	148.02	3,599.97	5.55
101	CEO Compensation ex. Options in %	Ceo_comp_op_ex_%	Total CEO compensation incl. options exercised (tdc2) as %.	Execucomp	193,251	176.14	4,789.48	8.15

Appendix B

Hyperparameter Tuning Specification

The hyperparameter tuning specification describes the set of hyperparameters Θ and their possible values, which is tuned for each machine learning model and each rolling window based on a training and validation data set.

$$\Theta_{LR} = \begin{cases} \text{penalty} = \text{None} \\ \text{solver} = \text{"saga"} \\ \text{max_iter} = 100 \end{cases}$$
$$\Theta_{XGB} = \begin{cases} \eta \in \{0.1, 0.3, 0.5, 0.7, 0.9\} & (\text{learning rate}) \\ n_{estimators} \in \{1000, 3000, 5000\} \\ \text{max_depth} \in \{2, 4, 6, 8, 10, 12\} \\ \gamma \in \{1e-6, 1e-3, 1e-1, 0, 0.1, 0.2\} \\ \text{subsample} \in \{0.5, 0.6, 0.7, 1\} \\ \alpha \in \{0.01, 0.1, 1, 10\} & (L1 \text{ reg}) \\ \lambda \in \{0.01, 1, 10\} & (L2 \text{ reg}) \\ \Delta \in (0, 1, 5) \\ \text{min_child_weight} \in \{5, 10, 20, 50\} \\ \text{colsample_bytree} \in \{0.5, 10\} \end{cases}$$

