Modelling and Forecasting the Evolution of RegTech

Abhishek Pramanick Barry Quinn Fearghal Kearney

2025-02-06

Regulatory Technology (RegTech), a emerging subsector in the financial technology industry, leverages technological innovation to automate and enhance compliance, surveillance, and risk management functions in financial institutions. This study uses Bayesian survival modelling to examine the factors driving the evolution of RegTech firms, including market consolidation and technological adoption. Analysing a unique dataset, we identify key drivers of mergers and acquisitions (M&A) and firm survival, with early-stage firms facing distinct challenges compared to established players. Our findings highlight the role of firm size, financial health, and technological maturity in shaping acquisition outcomes, with financial software companies emerging as dominant acquirers. By providing new insights into RegTech's growth and its implications for financial regulation, this research offers a foundation for forecasting trends in the financial technology ecosystem.

Introduction

Regulatory Technology (RegTech) has emerged as a pivotal force in the financial services industry, driven by increasing regulatory complexity and rapid digital innovation. In the aftermath of the 2008 global financial crisis, policymakers around the world introduced stringent requirements—from Basel III to region-specific regulations like GDPR—that have dramatically heightened compliance demands on financial institutions (Arner, Barberis, & Buckley, 2017). Against this backdrop, RegTech providers are developing and deploying advanced tools designed to automate compliance processes, enhance risk management, and reduce operational costs (European Banking Authority, 2021). Such innovations carry important implications not only for individual firms but also for the stability and evolution of the financial system as a whole.

In parallel, a raise in mergers and acquisitions (M&A) has reshaped the RegTech landscape, reflecting a broader pattern of technology-driven market consolidation. Recent evidence suggests that established financial software companies are increasingly acquiring RegTech start-ups to stay competitive, integrate cutting-edge expertise, and diversify product offerings (Cornelli et al., 2021). This consolidation, however, remains insufficiently understood: while qualitative accounts point to the strategic importance of RegTech solutions, quantitative investigations into the drivers of firm survival, acquisition likelihood, and overall sectoral trajectory are relatively scarce (Iansiti & Lakhani, 2021). Understanding these dynamics is vital for anticipating how financial markets, regulatory frameworks, and even societal trust in financial systems may evolve.

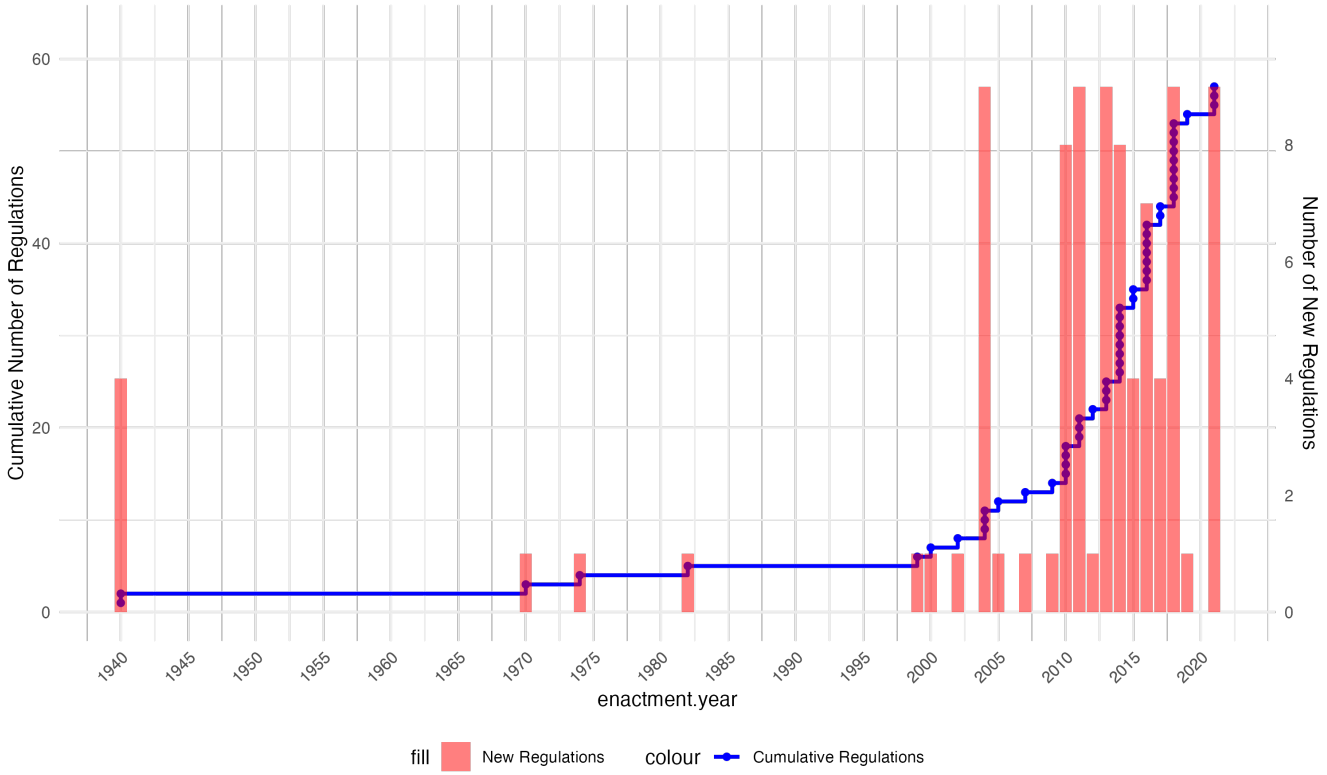


Figure 1: Financial Regulations Timeline

The rising volume and density of regulations globally underscores the critical need for innovative compliance solutions. As illustrated in Figure 1[1] , the number of regulatory provisions in major financial markets has expanded dramatically over the last few decades, with particularly steep increases following key financial crises. This surge in regulatory requirements has driven financial institutions to seek more efficient, technology-driven approaches to compliance—catalysing the growth of RegTech. By automating processes and providing data-driven insights, RegTech has become not only a strategic priority for banks and other financial players but also an increasingly consolidated segment of the FinTech landscape.

Theories of technology adoption and innovation diffusion suggest that firms with greater technological readiness are more attractive to acquirers due to their capacity to generate competitive advantages and integrate seamlessly into existing operations (Yoo, Henfridsson, & Lyytinen, 2010). Capturing this readiness requires robust methodologies capable of processing complex, unstructured data sources like patent filings. An important feature of our research is the AI-driven assessment of technological maturity, derived from patent abstracts that were classified by large language models (LLM). Each patent receives a probabilistic score reflecting its alignment with a RegTech-specific Technology Readiness Level (TRL) framework. These scores are aggregated at the firm level, producing a measure of technological readiness at the point of acquisition. Crucially, the model retains each firm's underlying classification uncertainty rather than forcing categorical thresholds, thus remaining consistent with a Bayesian perspective on probability and partial pooling.

This study addresses these gaps by applying a Bayesian survival modelling approach to a hand-collected dataset of RegTech firms and RegTech acquisition activities, examining how factors such as firm size, financial health, and technological maturity influence acquisition outcomes. In doing so, it provides

[1]Data sourced from regulatory body websites and policy documents across major financial markets including the US, UK, and EU from 1990-2020. A full table detailing each piece of financial regulation is avaliable upon request.

a generative modelling approach for forecasting how RegTech consolidation may unfold in the coming years and how this, in turn, might shape the future of financial innovation in regulatory compliance. By incorporating AI-driven metrics of technological readiness this research demonstrates how bayesian hierarchical inference can shed light on the interplay between emerging technologies and market consolidation. Ultimately, these findings aim to inform policymakers, industry practitioners, and scholars alike, offering a data-driven foundation for anticipating future challenges and opportunities in the rapidly evolving RegTech ecosystem.

Literature Review

The increasing complexity of financial regulations has driven the emergence of Regulatory Technology (RegTech), a subset of Financial Technology (FinTech) designed to streamline compliance and enhance operational efficiency (Arner, Barberis, & Buckley, 2017). RegTech solutions leverage advances in machine learning, data analytics, and cloud computing to address rising compliance costs and regulatory burdens (European Banking Authority, 2021). As the sector grows, mergers and acquisitions (M&As) have become a dominant feature of market dynamics, mirroring broader trends in high-technology industries where established incumbents frequently acquire innovative start-ups.

M&A activity in technology-intensive sectors is often driven by the pursuit of novel capabilities, the need to maintain competitive advantage, and the aim to expand into new markets. Rossi and Tarba (2013) note that high-tech firms often acquire start-ups to access emerging technologies, a strategy that Ahuja and Katila (2001) show can boost the acquirer's innovation performance—especially when there is a moderate overlap in R&D knowledge bases. Nevertheless, as Cloodt, Hagedoorn, and Van Kranenburg (2006) demonstrate, stronger technological relatedness can also limit the breadth of subsequent innovation outcomes, underlining the strategic complexity of technology acquisitions.

FinTech start-ups, including those specialising in RegTech, frequently rely on venture capital (VC) to fund early-stage research and development. Cumming and Schwienbacher (2018) illustrate how investor preferences, tied to governance and exit strategies, can accelerate or impede acquisitions in FinTech markets. Although their study centres on VC rather than M&A per se, the mechanisms they identify—such as alignment between investors, founders, and strategic acquirers—remain relevant in RegTech, where early VC backing can validate market potential and pave the way for future acquisitions. Overall, understanding how VC financing intersects with firm-level innovation dynamics is crucial for forecasting RegTech's consolidation trajectories.

Corporate innovation emerges as a key catalyst for acquisitions across technology industries. Innovative firms not only attract acquirers due to their unique technological advantages (Bena & Li, 2014) but can also help acquiring companies cross critical innovation thresholds (Cefis & Marsili, 2015). In RegTech specifically, financial software companies often lead consolidation efforts by integrating compliance-focused technologies into their broader service offerings (Cornelli et al., 2021). Nevertheless, despite the evident strategic importance of RegTech solutions, quantitative insights into the drivers of acquisition activity in this niche remain sparse.

Empirical investigations of M&As in technology-intensive settings have employed an array of econometric and event-study approaches. For instance, Dranev et al. (2019) and Collevecchio et al. (2024) examine how FinTech acquisitions affect acquirer performance, particularly in the banking sector, by considering the contingent factors that shape post-M&A outcomes. Their study highlights that sustainability, minority acquisitions, and institutional distance between acquirer and target firms influence cumulative abnormal returns, underscoring the strategic importance of specialised technological capabilities in financial services. In a broader technological context, Cassiman et al. (2005) reveal how M&As can reshape R&D by either

integrating or constraining innovation processes. Taken together, these studies suggest that understanding acquisition outcomes in emerging, technology-focused markets hinges on evaluating both the target's technological readiness and the acquirer's strategic motivations.

Bayesian survival modelling has found increasing favour in the empirical literature on mergers and acquisitions (M&As) within technology-intensive markets, offering a data-rich framework for evaluating how firms enter or exit through acquisitions under conditions of heightened uncertainty (Gelman et al., 2004; Cassiman et al., 2005). Traditional survival analyses—often grounded in frequentist approaches—can yield valuable insights into hazard rates and covariate effects but may struggle in contexts where data are sparse, firm-level heterogeneity is substantial, and the underlying distributions are highly skewed (Wheelock & Wilson, 2000; Kling, Pham, & Vencappa, 2020). As nascent sectors such as FinTech and RegTech continue to evolve, researchers have increasingly confronted the limitations of static panel estimations that assume homogeneity across firms, particularly when studying the interplay between technological readiness, financial metrics, and strategic investment decisions (Pham, Oikonomou, & Brooks, 2021; Cumming & Schwienbacher, 2018).

The Bayesian paradigm directly addresses many of these challenges by employing posterior distributions to represent uncertainty, allowing researchers to probabilistically estimate key parameters associated with acquisition likelihood (Gelman et al., 2004). Rather than providing point estimates or relying solely on confidence intervals, Bayesian methods derive full posterior distributions for both fixed and random effects.[2] This is especially pertinent in complex M&A contexts, where the influence of covariates such as firm size, leverage, or retained earnings may vary substantially across different technological maturities and jurisdictions (Palepu, 1986; Ruland & Zhou, 2005). Moreover, the use of Markov Chain Monte Carlo (MCMC) sampling enables posterior predictive checks that reveal how well the model captures observed acquisition patterns under varying assumptions, thereby enhancing the credibility of the inferences (Gelman et al., 2013).

A hierarchical extension further refines these analyses by incorporating group-level effects—for instance, capturing how acquisitions might differ across countries with distinct regulatory regimes or across firms at varying levels of technological maturity (Cassiman et al., 2005; Iansiti & Lakhani, 2021). In a Bayesian hierarchical framework, partial pooling ensures that estimates for subgroups with limited observations draw strength from the overall dataset, mitigating issues of imprecise inference (Hsiao, 2014; Athey & Imbens, 2017). This capability to accommodate cross-sectional dependence and multi-level structures is crucial in RegTech studies, where technological readiness scores—generated through AI-based patent analyses—may lack uniform distribution and where country-specific policies dramatically shape the risk and return profiles of early-stage ventures (Cornelli et al., 2021; Pham et al., 2021).

By enabling a probabilistic treatment of both firm-level and subgroup-level heterogeneity, Bayesian hierarchical survival modelling provides a useful lens through which to examine RegTech's consolidation dynamics. It not only quantifies how financial health, firm age, and leverage affect acquisition probabilities, but also elucidates the role of more intangible factors such as innovation capacity and jurisdication-specific learning effects (Song, 1993; Balasubramanian & Lee, 2008). The resulting posterior distributions have the potential to illuminate scenario-based outcomes—for instance, how hazard rates might shift if regulatory

[2]The interpretation of *fixed and random effects* differs between Frequentist and Bayesian frameworks. In Frequentist statistics, a parameter is typically considered fixed if it is constant across repeated samples, while a random effect is drawn from a probability distribution that represents population-level variability. However, as Gelman (2005) and McElreath (2020) argue, all effects are "fixed" in Frequentist models, as no probability distribution is assigned to parameters—rather, variability is inferred solely from data. In Bayesian statistics, both so-called "fixed" and "random" effects have full posterior distributions that reflect uncertainty. What is traditionally called a random effect in Frequentist terms corresponds to a hierarchical prior in Bayesian models, capturing shrinkage across group-level units (Gelman & Hill, 2006). Conversely, parameters without group-level variability are still treated as probabilistic quantities, with posterior distributions derived from both the data and prior knowledge. This fundamental distinction underscores how Bayesian methods provide coherent uncertainty quantification across all model components.

mandates tighten or if early-stage technological capabilities mature—ultimately offering policymakers, investors, and firm managers a forecasting tool that accounts for uncertainty in a way that frequentist methods often cannot.

Despite a growing body of research on M&As in technology-oriented industries, the unique dynamics of RegTech remain underexplored. Existing analyses tend to cluster RegTech with broader FinTech markets or more traditional financial sectors, overlooking how factors such as technological maturity and market readiness specifically shape acquisition likelihood. Against this backdrop, this paper applies a Bayesian hierarchical survival model to a novel dataset of RegTech firms and transactions, offering new insights into how technology-driven consolidation unfolds in a regulatory-focused domain. In doing so, it addresses a crucial gap in the literature and contributes a future-oriented perspective on RegTech's role in reshaping compliance, competition, and innovation within financial services.

Bayesian Hierarchical Survival Modelling

The survival of RegTech firms is modelled using a Bayesian survival analysis framework, where the hazard of acquisition is estimated through a Weibull model with hierarchical extensions to capture jurisdictional and technological heterogeneity. The competing risks of firm exit—such as bankruptcy or voluntary dissolution—are treated as censoring events. Let T_i denote the survival time for firm i, measured from either the start of the observation period or the firm's incorporation, whichever is later. The censoring indicator δ_i takes the value 1 for observed acquisitions and 0 for censored observations, while the at-risk set $R(t)$ comprises all firms under observation at time t that have not yet been acquired or censored.

The baseline model assumes a parametric Weibull hazard function, widely used in survival analysis due to its flexibility in accommodating increasing or decreasing hazard rates over time. The hazard function for firm i with covariates $x_i(t)$ is given by:

$$h(t \mid x_i(t)) = \alpha\lambda^{\alpha} t^{\alpha-1} \exp(x'_{i(t)}\beta),$$

where α is the Weibull shape parameter, governing whether the hazard rate increases ($\alpha > 1$) or decreases ($\alpha < 1$) over time. The parameter λ determines the scale, while $x_{i(t)}$ is a vector of time-varying firm-level covariates, including size, retained earnings, net debt, and firm age, with β as the associated coefficient vector. The survival function, representing the probability of a firm surviving beyond time t, follows:

$$S(t \mid x_i(t)) = \exp\Big(-\lambda^{\alpha} t^{\alpha} \exp(x'_{i(t)}\beta)\Big).$$

This baseline specification (which we name Model 1) assumes homogeneity across firms but may obscure heterogeneity in acquisition dynamics arising from jurisdictional or technological differences. If firms within certain regions or technological categories exhibit systematically different hazard rates, a model that fails to account for these dependencies may lead to biased or misleading inferences. A hierarchical Bayesian approach provides a natural solution by allowing model parameters to vary across groups, incorporating prior knowledge while borrowing strength from the entire dataset to improve estimation efficiency, particularly for groups with sparse observations. This approach aligns with broader empirical findings that emphasize the importance of hierarchical structures in modeling firm behavior and market outcomes.

A hierarchical extension introduces group-specific parameters by allowing β_j to follow a higher-level multivariate normal distribution, $\beta_j \sim MVN(\mu_\beta, \Sigma_\beta)$, thereby accommodating partial pooling. The likelihood function expands to:

$$L(\theta) \;=\; \prod_{j=1}^{J} \prod_{i=1}^{n_j} h_{ij}(t_i)^{\delta_i} \, \exp\Big(-\int_0^{t_i} h_{ij}(u)\,\mathrm{d}u\Big),$$

where δ_i is the event indicator for firm i, and θ collects parameters $(\alpha, \lambda, \mu_\beta, \Sigma_\beta)$. Under this hierarchical scheme, partial pooling guards against overfitting in groups with limited observations and yields robust estimates of how different market segments or regions shape survival outcomes. This approach is particularly pertinent in RegTech, where acquisition dynamics are influenced not only by financial metrics but also by jurisdictional regulatory environments and technological maturity. Firms operating under different regulatory regimes may face distinct acquisition incentives, while technological readiness levels may alter the strategic value of firms to potential acquirers.

To incorporate country-specific baseline hazards, a further model (named Model 2) introduces jurisdictional random intercepts:

$$\log(\lambda_{ij}) = \beta_0 + \mu_{0j} + \beta_1 \text{Size}_t + \beta_2 \text{Retained Earnings}_t + \beta_3 \text{Net Debt}_t + \beta_4 \text{Age}_t + \epsilon_{i,j}.$$

where $\mu_{0j} \sim \mathcal{N}(0, \sigma^2_{0,j})$ represents country-specific deviations from the baseline hazard. This accounts for structural differences in M&A activity across jurisdictions, reflecting variations in legal frameworks, capital markets, and regulatory oversight. While this specification allows different countries to have different acquisition propensities, it assumes that the effects of firm-level covariates remain constant across countries.

A more flexible specification (named Model 3) allows for complete parameter heterogeneity across jurisdictions, permitting firm-level covariates to have different effects depending on the country:

$$\log(\lambda_{ij}) = (\beta_0+\mu_{0j})+(\beta_1+\mu_{1j})\text{Size}_t+(\beta_2+\mu_{2j})\text{Retained Earnings}_t+(\beta_3+\mu_{3j})\text{Net Debt}_t+(\beta_4+\mu_{4j})\text{Age}_t+\epsilon_{i,j}.$$

where:

$$[\mu_{0j}, \mu_{1j}, \mu_{2j}, \mu_{3j}, \mu_{4j}] \sim \mathcal{MVN}(0, \Sigma_\mu).$$

This varying-coefficient specification captures differences in the sensitivity of firm acquisitions to financial characteristics across jurisdictions. Such flexibility is particularly relevant given that financial metrics like leverage or firm age may influence acquisition probability differently depending on institutional environments.

Incorporating technological heterogeneity, the random intercept model (named Model 4) structures firm survival around TRL classifications:

$$\log(\lambda_{ij}) = \beta_0 + v_{0k} + \beta_1 \text{Size}_t + \beta_2 \text{Retained Earnings}_t + \beta_3 \text{Net Debt}_t + \beta_4 \text{Age}_t + \epsilon_{i,k}.$$

where $v_{0k} \sim \mathcal{N}(0, \sigma^2_{0,k})$ captures technology-specific deviations. Firms at different levels of technological readiness may exhibit different acquisition risks, as high-TRL firms are often more mature and integrated into financial systems, making them more attractive to acquirers. Lower-TRL firms, on the other hand, may be seen as riskier investments with uncertain commercialization potential.

The full TRL varying-coefficient specification (named Model 5) allows covariate effects to vary by technological maturity:

$$\log(\lambda_{ij}) = (\beta_0+v_{0k})+(\beta_1+v_{1k})\text{Size}_t+(\beta_2+v_{2k})\text{Retained Earnings}_t+(\beta_3+v_{3k})\text{Net Debt}_t+(\beta_4+v_{4k})\text{Age}_t+\epsilon_{i,k}.$$

where:

$$[v_{0k}, v_{1k}, v_{2k}, v_{3k}, v_{4k}] \sim \mathcal{MVN}(0, \Sigma_v).$$

This specification provides a granular view of how acquisition determinants vary across different stages of technological development. It reflects empirical findings suggesting that firms at later stages of technological maturity are more likely to be acquired due to reduced uncertainty in product viability and regulatory compliance.

The estimation employs Hamiltonian Monte Carlo methods via the No-U-Turn Sampler (NUTS), implemented in Stan. The likelihood function, incorporating both time-varying covariates and random effects, follows:

$$L(\theta \mid \text{data}) = \prod_i \prod_j \prod_k \left[h(t_i \mid x_i(t_i), u_j, v_k)^{\delta_i} \times S(t_i \mid x_i(t_i), u_j, v_k) \right].$$

Priors are set to reflect limited ex-ante information while ensuring proper posterior distributions:

$$\beta \sim \mathcal{N}(0, 2), \quad \alpha \sim \text{Gamma}(2, 2),$$

$$\Sigma_u, \Sigma_v \sim \text{LKJ}(2), \quad \sigma^2_{u_0}, \sigma^2_{v_0} \sim \text{Half-Cauchy}(0, 5).$$

The estimation procedure is conducted using four parallel Markov chains, each with 2,000 iterations, of which the first 1,000 serve as warm-up steps for adaptation. The No-U-Turn Sampler (NUTS) is used with a target acceptance rate of 0.85 and a maximum tree depth of 12, ensuring efficient exploration of the posterior distribution while accommodating the hierarchical structure (Hoffman & Gelman, 2014). These tuning parameters balance convergence speed with numerical stability, particularly given the inclusion of higher-dimensional multilevel random effects.

A hierarchical Bayesian approach offers several advantages in the context of RegTech acquisitions. First, it enables partial pooling, which mitigates overfitting in groups with sparse observations while allowing meaningful differences to emerge when sufficient data are available (Gelman et al., 2013). This is particularly relevant in financial technology markets, where firms operating in different jurisdictions or at different levels of technological readiness may exhibit distinct acquisition risks. Second, Bayesian inference provides a coherent framework for incorporating prior knowledge about acquisition dynamics (Athey & Imbens, 2017), improving inference in cases where observational data alone may be insufficient for robust estimation. Prior distributions regularize estimates, preventing overfitting while allowing the data to dominate where sample sizes are large (McElreath, 2020). Lastly, the Bayesian posterior predictive distributions allow for uncertainty quantification at both the firm and group levels, providing richer insights than traditional frequentist survival models (Gelman et al., 2004).

By structuring the survival model hierarchically, this analysis captures a broader range of plausible survival paths while emphasizing structural variations among firms. This is particularly pertinent in RegTech, where the interplay between financial metrics and regulatory drivers is both uncertain and multifaceted. Firms with high technological maturity may exhibit different acquisition risks than those at earlier stages of

development, while jurisdictional regulatory environments may alter the incentives of acquirers. Empirical studies on firm acquisitions suggest that survival probabilities are heavily influenced by both financial health and technological maturity (Cefis & Marsili, 2015; Bena & Li, 2014), yet these factors are often moderated by market structure and institutional environments (Rossi & Tarba, 2013). The Bayesian framework accommodates these complexities by integrating structured prior knowledge with empirical data, ensuring robust parameter estimates across varying market conditions.

The results of the Bayesian survival analysis yield posterior distributions for key parameters, which are used to infer the magnitude and direction of firm-level and group-level effects on acquisition risk. Posterior predictive checks confirm that the model appropriately captures the observed acquisition patterns, with hierarchical priors effectively regularizing estimates for smaller groups. This approach aligns with research in empirical industrial organization, which emphasizes the importance of Bayesian hierarchical modeling for capturing heterogeneous firm behaviors (Rossi et al., 2005). By examining the posterior distributions of jurisdictional and TRL-specific effects, the analysis provides novel insights into how acquisition dynamics unfold across different technological maturities and regulatory settings.

The hierarchical model also enables counterfactual simulations that estimate how acquisition likelihoods would change under different regulatory or technological conditions. For instance, posterior samples can be used to generate predictive distributions of acquisition probabilities for firms operating in jurisdictions with differing levels of regulatory stringency or for firms transitioning between TRL stages. These simulations provide practical insights for investors and policymakers seeking to understand the conditions under which RegTech firms are most likely to be acquired. Similar approaches have been applied in technology management research to model the impact of regulatory shifts on innovation-driven industries (Cloodt et al., 2006; Bergek et al., 2013). By leveraging Bayesian decision theory, these counterfactuals extend beyond purely descriptive modeling to provide forward-looking policy and investment implications.

By integrating hierarchical Bayesian methods with structured probabilistic TRL classification, this study advances the empirical analysis of acquisition risk in financial technology markets. The framework provides a robust foundation for assessing firm survival dynamics while incorporating uncertainty at multiple levels. The combination of structured survival modeling with machine-learning-driven technological readiness estimation offers a methodological contribution that extends beyond RegTech and is broadly applicable to technology-intensive industries experiencing market consolidation.

Data

We construct two data sample for the period from 2010 to 2023. Firstly, for use in a qualitative trend analysis, we collect a sample of Pitchbook 5,223 deals from 1,949 RegTech start-ups identified as primarily focus in regulatory technology solutions for financial firms. Secondly, for use in the Bayesian hierarchical survival analysis, we using S&P Global Capital IQ and Bureau van Dijk's FAME databases to source firm-level financial information which results in an unbalanced panel of 450 firm-year observations for 71 RegTech firms.

To create the Technological Readiness Level (TRL) variable, we employ a large language model (LLM) classification framework to analyze 3,223 patents sourced from the PATSTAT Global database. This approach tracks the *evolution of technological maturity over time*, assigning TRL categories (ranging from 6 to 9) to firms based on their intellectual property holdings at different stages of development. Rather than treating TRL as a static firm-level attribute, we construct a time-variant TRL trajectory leading up to the point of acquisition, ensuring that classification reflects a firm's technological readiness at the moment it enters acquisition negotiations. This trajectory-based approach generates a hierarchical measure of technological maturity that serves as a key group-level predictor in the Bayesian survival analysis. By

applying structured prompting and a well-defined classification rubric, the methodology ensures that TRL assignments capture sector-specific technological advancements rather than relying solely on raw textual patterns. This approach is consistent with studies on M&A activity in technology-intensive sectors, where firms are acquired at peak technological maturity (Bena & Li, 2014; Bergek et al., 2013).

Each patent abstract is processed independently by two state-of-the-art LLMs, GPT-4 and Claude 3.5, which classify the technology's maturity based on implementation status, testing environment, integration level, and commercial readiness. The models follow a structured evaluation framework, extracting key indicators from the patent text and mapping them to TRL categories using chain-of-thought reasoning. To construct a temporal TRL trajectory, the classification process identifies TRL scores at multiple points in a firm's patenting history, tracking how firms progress technologically over time. This ensures that the assigned TRL at acquisition is based on the firm's most recent innovations, rather than an arbitrary historical average. By generating probabilistic TRL scores rather than deterministic labels, the classification process explicitly models the uncertainties inherent in textual data and innovation categorization.

To enhance classification reliability, cases where LLM classifications differ significantly—particularly when TRL assignments diverge by more than one level or confidence scores are low—are flagged for human review by a panel of domain experts. The panel consists of a RegTech specialist, a patent attorney, and a financial technology researcher, who apply the same structured TRL rubric to adjudicate discrepancies. Inter-rater reliability metrics, including Cohen's Kappa and Krippendorff's Alpha, confirm the consistency of TRL assignments, ensuring that classifications align with expert judgments.

At the firm level, probabilistic TRL scores from individual patents are aggregated to construct a dynamic TRL trajectory, capturing the rate and direction of technological progression rather than a single-point estimate. The firm's TRL at the moment of acquisition is then extracted as the key predictor in the Bayesian hierarchical survival model. This aggregation accounts for both the number and distribution of patents across TRL categories, weighting each patent's contribution based on its relevance to regulatory compliance technologies. The hierarchical Bayesian model structure allows for partial pooling across firms, stabilizing estimates for firms with limited patenting activity while preserving firm-specific technological variations.

By constructing TRL as a trajectory leading up to acquisition and incorporating the final TRL level into the Bayesian hierarchical model, the methodology ensures that firm-level technological maturity reflects the state of innovation at the moment of acquisition rather than relying on static firm attributes or self-reported classifications. The structured prompting and dual-LLM validation approach align with previous studies that employ machine learning for technology forecasting and patent-based innovation metrics (Olechowski, Eppinger, & Joglekar, 2015; Mankins, 1995). Moreover, the probabilistic nature of TRL assignments enhances model interpretability, enabling the study to assess whether firms at different technological stages exhibit distinct survival and acquisition patterns in the evolving RegTech ecosystem. The trajectory-based approach further supports counterfactual analyses, allowing estimation of how shifts in technological maturity influence acquisition probabilities under varying market conditions.

Table 1 details the key variables employed in the Bayesian hierarchical survival analysis. Each variable is accompanied by both an operational description and a brief rationale linking it to prior mergers and acquisitions (M&A) research. Where applicable, references to established theoretical or empirical literature are included to bolster the conceptual underpinnings of each covariate.

Table 1: Key Variables in the RegTech Dataset, with Descriptions and Theoretical Rationale

Variable	Type	Description	Rationale
Size	Firm-level	Logarithm of total assets, derived from quarterly financial statements.	Serves as a proxy for resource capacity and market influence. Larger firms may benefit from economies of scale but can face integration challenges (Wheelock & Wilson, 2000). From an M&A perspective, size can signal stability or confer bargaining power (Palepu, 1986).
Financial Health	Firm-level	Retained earnings, scaled by total assets, indicating growth potential and resilience.	Healthy financial reserves can make a firm less susceptible to liquidity pressures, thus influencing acquisition likelihood. Firms with strong retained earnings may appear more stable, although high reserves can also reduce the pressure to accept takeover offers (Anagnostopoulou & Tsekrekos, 2015; Black, 2010).
Leverage	Firm-level	Net debt (total debt minus cash holdings), scaled by total assets.	Shapes firm risk profiles and may attract acquirers seeking undervalued or vulnerable targets. High leverage can expedite financial distress yet also signal aggressive growth strategies (Safieddine, 1999; Ruland & Zhou, 2005). Its effect on M&A can vary across industries and economic contexts.
Maturity	Firm-level	Age of the firm in years since inception, updated quarterly as needed.	Younger firms exhibit higher innovation potential but face greater uncertainty, making them attractive yet riskier acquisition targets (Balasubramanian & Lee, 2008). Conversely, older firms may be less agile but more established (Audretsch, 1995).

Variable	Type	Description	Rationale
Jurisdiction	Group-level	Country of origin for the RegTech firm, coded by headquarters location.	Local regulatory frameworks and market maturity substantially influence investment climates and consolidation patterns (Ransbotham & Mitra, 2010). Cross-country heterogeneity justifies partial pooling in a hierarchical Bayesian approach (Hsiao, 2014).
Technological Readiness Level	Group-level; AI-based	Probabilistic assignment of Technology Readiness Levels (6–9) using a fine-tuned large language model applied to patent abstracts, following frameworks such as Mankins (1995) and Olechowski et al. (2015).	Key driver of M&A in innovation-intensive sectors (Cloodt, Hagedoorn, & Van Kranenburg, 2006). Higher TRLs indicate market-ready solutions that often command higher valuations, while intermediate TRLs reflect uncertainty, influencing both investment decisions and strategic acquisitions (Olechowski et al., 2015).

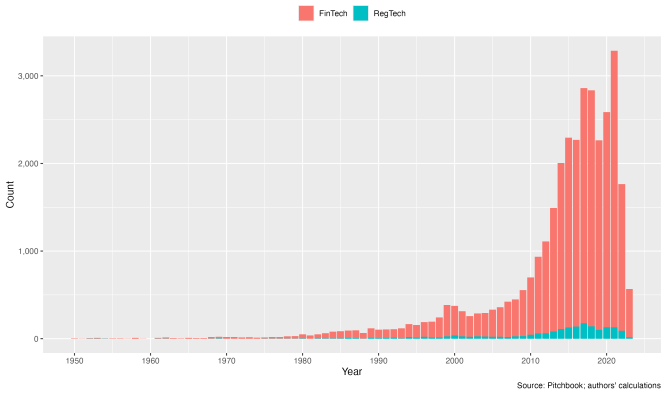
Results

Qualitative Trends in RegTech

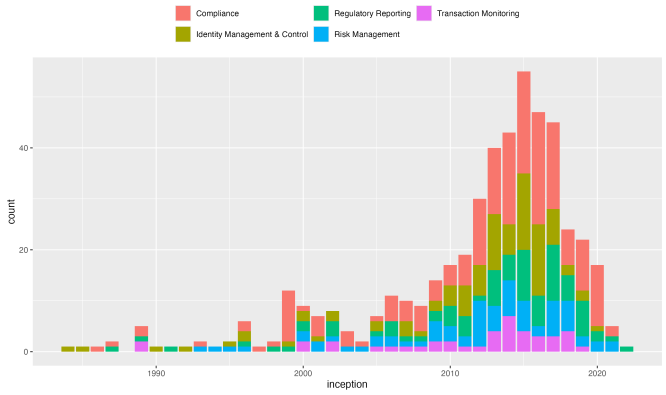
The RegTech industry has evolved dynamically over the past two decades, emerging as a niche yet rapidly expanding segment within the broader FinTech landscape. Leveraging data from 1,949 firms and 5,223 deals, this qualitative overview of the sector's growth, investment behaviour, and regional patterns, describes how RegTech positions itself relative to wider financial technology developments. The pace of RegTech incorporations has paralleled trends in the FinTech sector, as depicted in Figure 2a, although FinTech remains dominant in absolute terms. Regulatory drivers, including Basel III and the General Data Protection Regulation (GDPR), have played a significant role in accelerating RegTech's momentum (Fenwick et al., 2017; BIS, 2019). Arner et al. (2017) similarly identify the global financial crisis of 2008 as a pivotal catalyst, highlighting how the ensuing wave of complex regulations spurred demand for automation and risk-mitigation tools. However, a recent decline in new incorporations points to an imminent consolidation phase, echoing McKinsey's (2022) observations of broader FinTech maturity cycles.

In terms of applications, RegTech activity remains concentrated in areas such as compliance, identity management, and transaction monitoring, as illustrated in Figure 2b. Compliance has proved dominant, reflecting how technology-driven solutions fill a critical gap between expanding regulatory mandates and constrained operational resources (Anagnostopoulos, 2018). Identity management and transaction monitoring, meanwhile, are on the rise as fraud prevention and customer due diligence have become

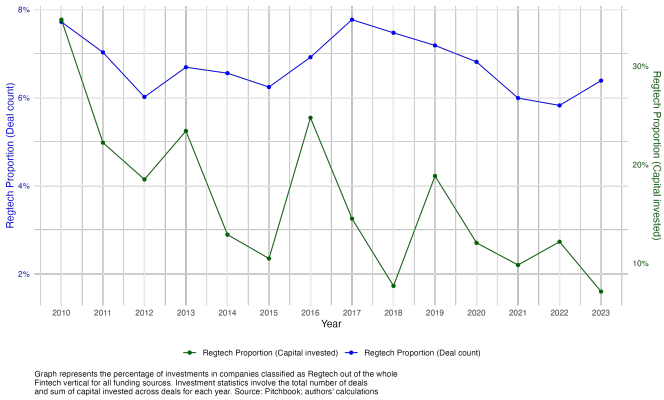
top priorities for financial institutions. These shifts are consistent with the burgeoning interest in artificial intelligence and blockchain, two domains widely acknowledged for enabling advanced RegTech innovations (KPMG, 2024). The sector's unique capacity to integrate cutting-edge technology with regulatory obligations underscores its heightened strategic value among stakeholders seeking efficient compliance solutions.



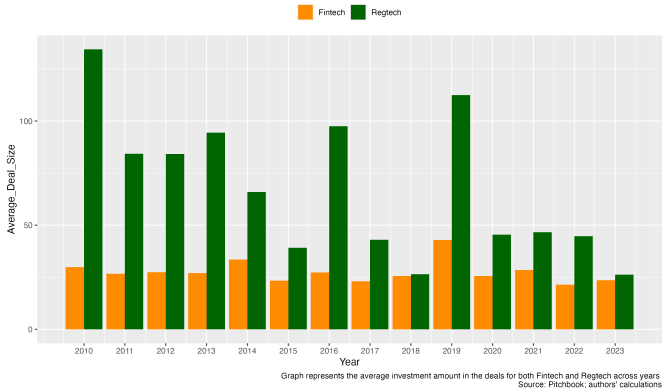
(a) FinTech and RegTech Incorporations Over Time



(b) RegTech Incorporations by Subsector



(c) Proportion of RegTech Deals and Capital



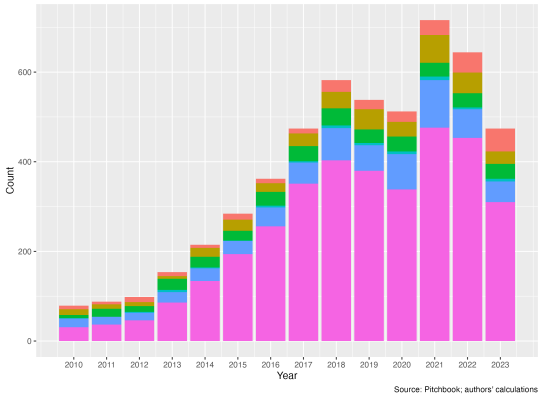
(d) Average Deal Size in FinTech vs RegTech

Figure 2: Demographic Analysis

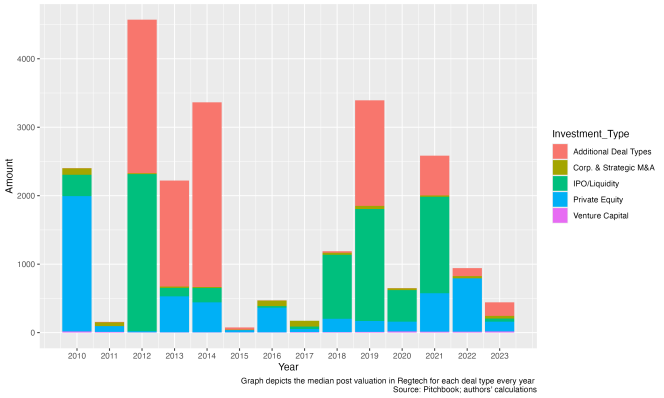
Notably, RegTech's growing stature extends to investment activity, though its share of overall FinTech deals remains modest. Figure 2c shows that the proportion of RegTech deals in the FinTech landscape typically fluctuates between 6% and 7.5%, yet Figure 2d reveals how the average RegTech deal size often surpasses that of FinTech. This disparity underscores the premium placed on RegTech's potential to address complex regulatory challenges, a point emphasised by Cornelli et al. (2021), who note that high barriers to entry and scalability can draw substantial interest from investors. The valuation patterns in subsequent figures further reinforce this confidence, highlighting robust capital inflows and rising deal values over time.

The composition of RegTech deals over time, illustrated in Figure 3a, points to an ecosystem dominated by venture capital (VC) financing, with additional layers of private equity, corporate and strategic M&A, debt, and other deal types contributing to the sector's overall growth. From 2010 to 2013, the relatively modest volume of transactions featured predominantly VC investments, reflecting the early-stage nature of many RegTech start-ups (and the challenge of gathering financial information). Beginning around 2014, a pronounced increase in deal count is evident, culminating in peak activity by 2018 and 2019 as more investors, including private equity firms and strategic acquirers, recognised the potential for compliance-driven technology solutions. While VC remained the largest source of deals throughout this period, the

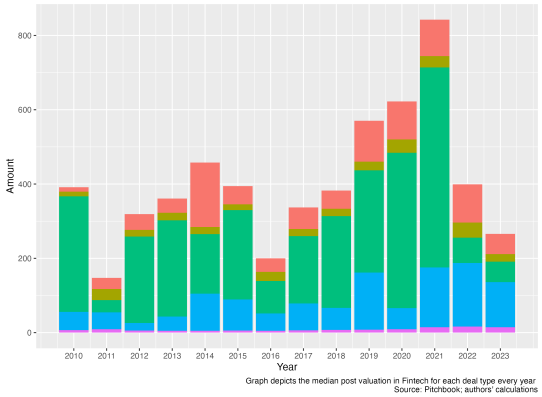
growth of corporate M&A—particularly among incumbent financial and software companies—underscores the mounting demand for integrated RegTech capabilities. Following the 2019 peak, deal activity appears to have cooled somewhat, aligning with wider market fluctuations, yet the sustained presence of multiple funding channels suggests an underlying resilience in RegTech's investment landscape.



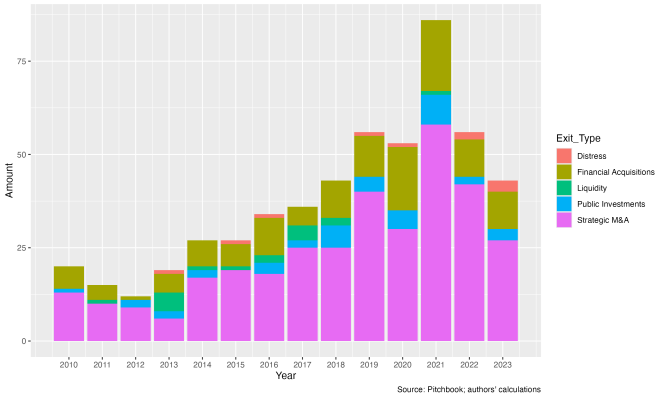
(a) RegTech Deals over Time



(b) RegTech Post Valuation by Investment Type



(c) Median Post Valuation in FinTech



(d) RegTech Exit Types

Figure 3: More Demographic Analysis

The sector's evolution is also evident in later-stage venture capital deals, with post-Series B rounds becoming more prominent, as highlighted in Figure 3b. This trend points to heightened investor confidence in RegTech's capacity for profitability and scale, mirroring patterns observed in other high-growth FinTech segments (EY, 2023). The surging median post-valuation levels shown in Figure 3b reinforce these perceptions, suggesting that RegTech has outpaced much of FinTech in delivering high-value exits. Many of these exits have come in the form of mergers and acquisitions. Figure 3d demonstrates that strategic M&A now dominates as the primary exit route for RegTech firms. Established financial software providers and IT consultancies have actively pursued acquisitions, incorporating RegTech capabilities into broader service offerings. McKinsey (2022) treats such M&A activity as a marker of sector maturity, implying that the purchase of niche technology companies by established incumbents reflects both a hunger for innovative compliance solutions and an acknowledgement of the central role RegTech plays in modern financial operations.

These trends collectively portray a resilient and rapidly maturing RegTech industry, one that has grown from a small subset of FinTech to a critical pillar in navigating the intricacies of modern regulation. The consolidation dynamics and rising valuations speak to RegTech's perceived long-term potential, not just in automating operational tasks but in fundamentally shaping compliance paradigms. This contextual

backdrop sets the stage for the subsequent survival analysis, which probes more deeply into the firm-level factors influencing acquisition likelihood and overall viability in an industry poised to drive ongoing innovation and regulatory efficiency.

Posterior Predictive Modelling

The Bayesian survival analysis yields full posterior distributions over model parameters, enabling inference beyond point estimates by incorporating parameter uncertainty directly into survival predictions. Rather than relying solely on maximum likelihood estimates, the hierarchical Bayesian approach generates posterior predictive distributions that quantify the range of plausible survival times for firms under varying technological and jurisdictional conditions.

Given the estimated posterior distributions for the model parameters, the hazard function at time t for firm i in jurisdiction j and TRL category k is obtained by marginalizing over the hierarchical effects:

$$h_{ijk}(t \mid x_{ijk}) = \int_{\Theta} \alpha \lambda^{\alpha} t^{\alpha-1} \exp(x_{ijk}\beta + u_j + v_k) p(\theta \mid \text{data}) d\theta,$$

where $p(\theta \mid \text{data})$ represents the joint posterior distribution over all model parameters, including both firm-level covariates and jurisdictional and technological random effects. The corresponding posterior predictive survival function follows:

$$S_{ijk}(t \mid x_{ijk}) = \int_{\Theta} \exp\Big(-\lambda^{\alpha} t^{\alpha} \exp(x_{ijk}\beta + u_j + v_k)\Big) p(\theta \mid \text{data}) d\theta.$$

These posterior predictive distributions offer several advantages over point-estimate-based survival models. First, they allow for direct uncertainty quantification, ensuring that reported survival probabilities incorporate variability arising from firm characteristics, industry dynamics, and jurisdictional heterogeneity. Second, by incorporating hierarchical priors, the model regularises firm-level estimates, reducing sensitivity to small-sample noise while maintaining flexibility to capture meaningful structural differences. Lastly, the posterior predictive distributions enable counterfactual simulations, where firm acquisition probabilities can be estimated under hypothetical changes in financial health, regulatory environment, or technological maturity.

These properties make Bayesian hierarchical survival models particularly useful for analyzing RegTech acquisitions, where firm outcomes are shaped by complex, interdependent factors. Prior research in empirical industrial organization has highlighted the importance of hierarchical modeling for capturing firm-specific heterogeneity (Rossi et al., 2005), while Bayesian frameworks have been widely applied to financial survival analysis due to their ability to integrate prior knowledge with observed data (McElreath, 2020).

Model Validation

The convergence and predictive performance of the Bayesian hierarchical survival models were assessed using a combination of convergence diagnostics, posterior predictive checks, and information criteria comparisons. These evaluations confirm the robustness of the estimated models and provide empirical support for the hierarchical specifications.

Gelman-Rubin $\hat{R}$ statistics across all model parameters remained below 1.01, indicating that the Markov chains mixed well and that posterior samples were drawn from a stable distribution. Effective sample sizes

(ESS) were generally high, though they declined slightly in the more complex varying-coefficient models due to increased hierarchical complexity. Monte Carlo standard errors remained within acceptable bounds, ensuring that posterior inferences were based on well-sampled distributions.

The random effects variance estimates confirm substantial heterogeneity across both jurisdictional and TRL-based groupings, with hierarchical components significantly improving model fit. The intraclass correlation coefficient (ICC) suggests that between-group variation is particularly pronounced in Model 5, where TRL-based effects account for a large proportion of survival heterogeneity. The proportional change in variance (PCV) further indicates that the introduction of hierarchical terms meaningfully reduces unexplained variation in acquisition likelihood.

Posterior predictive checks validate the models' ability to capture the observed data distribution, with simulated acquisition event times aligning well with empirical observations. The time-varying R_2 and dynamic AUC metrics demonstrate substantial improvements in predictive performance as hierarchical complexity increases, reinforcing the importance of incorporating both jurisdictional and technological maturity effects. Model 5 achieves the highest AUC (0.856), a notable increase over the baseline specification (0.745), indicating strong classification accuracy in distinguishing acquired firms from those that remain independent.

Comparisons of Deviance Information Criterion (DIC), Widely Applicable Information Criterion (WAIC), and Leave-One-Out Cross-Validation (LOO-CV) all favor the hierarchical varying-coefficient specifications, with Model 5 yielding the lowest DIC (2589.44) and WAIC (2597.82), suggesting superior out-of-sample predictive ability. Bayes factors computed on the log-marginal likelihood (LPML) further confirm support for the hierarchical specifications, with decisive evidence favoring the inclusion of jurisdictional and TRL-based random effects. The stepwise improvements across information criteria provide strong empirical justification for the incorporation of hierarchical dependencies in RegTech survival modeling.

For a comprehensive presentation of model diagnostics, including full posterior summaries, variance decomposition, and hierarchical effect stability measures, refer to Appendix A.

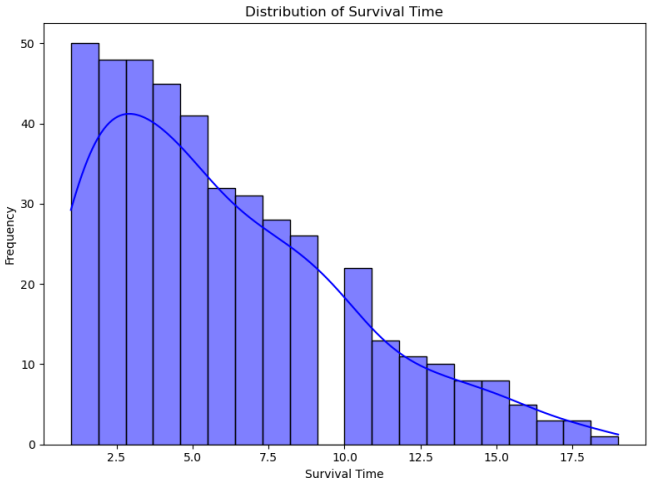
Descriptive Statistics of Survival Model Sample

An initial understanding of the financial and organisational attributes of the firms can be gleaned from the summary statistics presented in Table 2, which reports assets, retained earnings, net debt, and firm age before we standardise them for use in the survival analysis. These measures capture essential dimensions of size, profitability, liquidity, and operational longevity, all of which are relevant for understanding firm-level dynamics in nascent industries such as RegTech. The table presents data for the subset of firms with detailed financial statements, constituting the basis for the subsequent Bayesian hierarchical survival modelling.

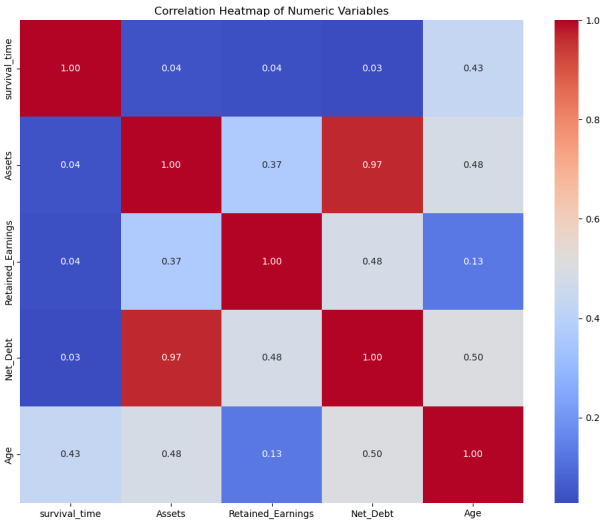
Table 2: Descriptive Statistics (Raw Data)

Variable	Count	Mean	Std Dev	Min	25%	Median	75%	Max
Survival Time	433	6.01155	4.0924	1	3	5	9	19
Assets	433	225,461.00	1,448,280.00	1.73	1,555.17	5,846.73	18,033.6	24,998,600.00
Retained Earnings	433	-13,148.80	109,084.00	-928,187	-5,465.71	-1.41	2,214.24	690,279.00
Net Debt	433	73,904.70	560,588.00	-138,564	-3,156.79	-620.07	0	8,210,800.00
Age	433	12.2979	9.1146	0	5	10	17	50

Figure Figure 4a depicts the distribution of survival times, showing a skewed pattern in which a substantial number of firms exit (through acquisition or otherwise) relatively early, while only a minority persist over an extended period. This phenomenon aligns with well-documented high failure or exit rates in innovation-driven contexts (Schumpeter, 1934; Audretsch, 1995), and it is particularly pronounced in emergent sectors like RegTech and FinTech, where business models and regulatory landscapes can shift rapidly.



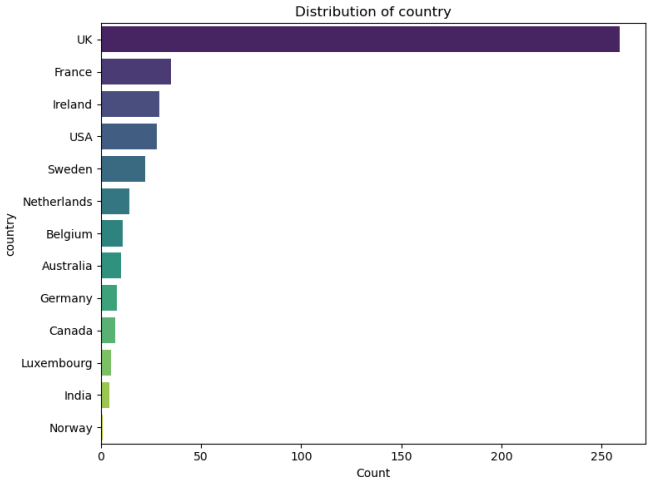
(a) Distribution of Survival Time



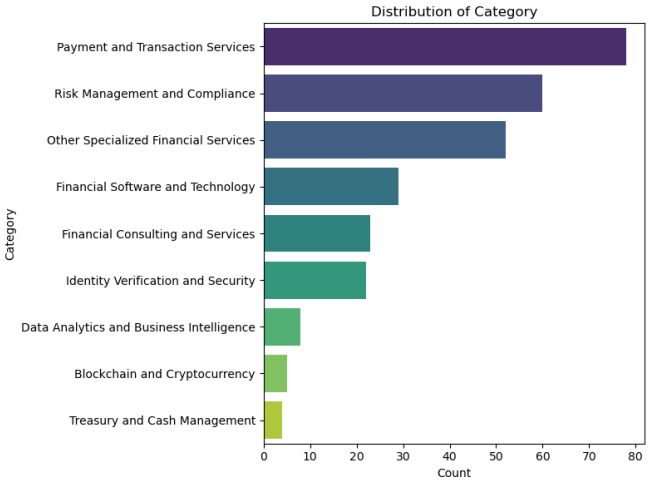
(b) Correlation Heatmap of Numeric Variables

Correlation analyses, summarised in Figure Figure 4b, highlight the relationships among key numerical variables. Survival time emerges as positively associated with firm age suggesting that more established and resource-rich firms tend to survive longer. By contrast, weak correlations with financial metrics indicate that additional structural or market factors may heavily influence acquisition outcomes. Similar complexities have been observed in other technology-intensive sectors (Coad, Segarra, & Teruel, 2016), where survival may depend on intangible assets and innovation capabilities more than on traditional balance sheet metrics.

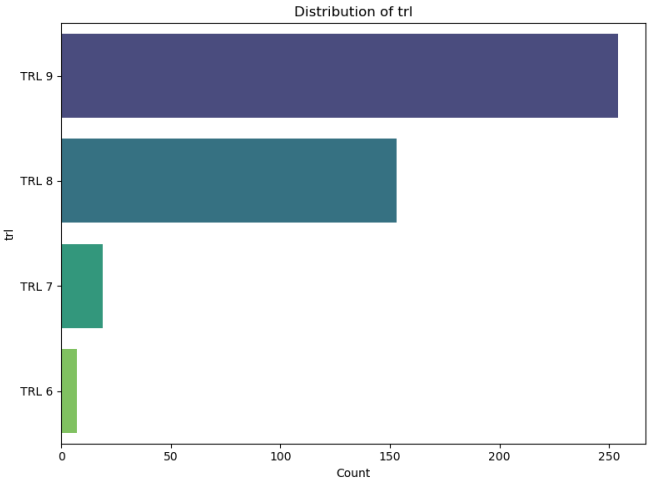
Figures Figure 5a, Figure 5b, and Figure 5c further contextualise the dataset by illustrating country distribution, firm specialisations, and Technology Readiness Levels (TRLs), respectively. A concentration of firms headquartered in the United Kingdom, followed by France and the United States, may reflect the comparatively supportive regulatory environments and market opportunities available in these regions (Arner, Barberis, & Buckley, 2017). On the category dimension, Payment and Transaction Services, along with Risk Management and Compliance, dominate the sample, underscoring the focus on solutions that automate core regulatory or operational processes (Anagnostopoulos, 2018). TRL distributions reveal a higher incidence of firms at TRLs 8 and 9, indicating a tendency toward near-market or fully operational technologies rather than purely experimental ventures—a pattern consistent with prior findings on FinTech's emphasis on deployable, revenue-generating solutions (KPMG, 2024).



(a) Distribution of Country



(b) Distribution of Category



(c) Distribution of TRL

Figure 5: Distribution Analysis

Posterior Predictive Results

Below we present the baseline model (Model 1) and the time-varying coefficient models for both group variables (Model 3 and Model 5). In all estimations the covariates entry the model as standardised to allow for direct effect comparisons.

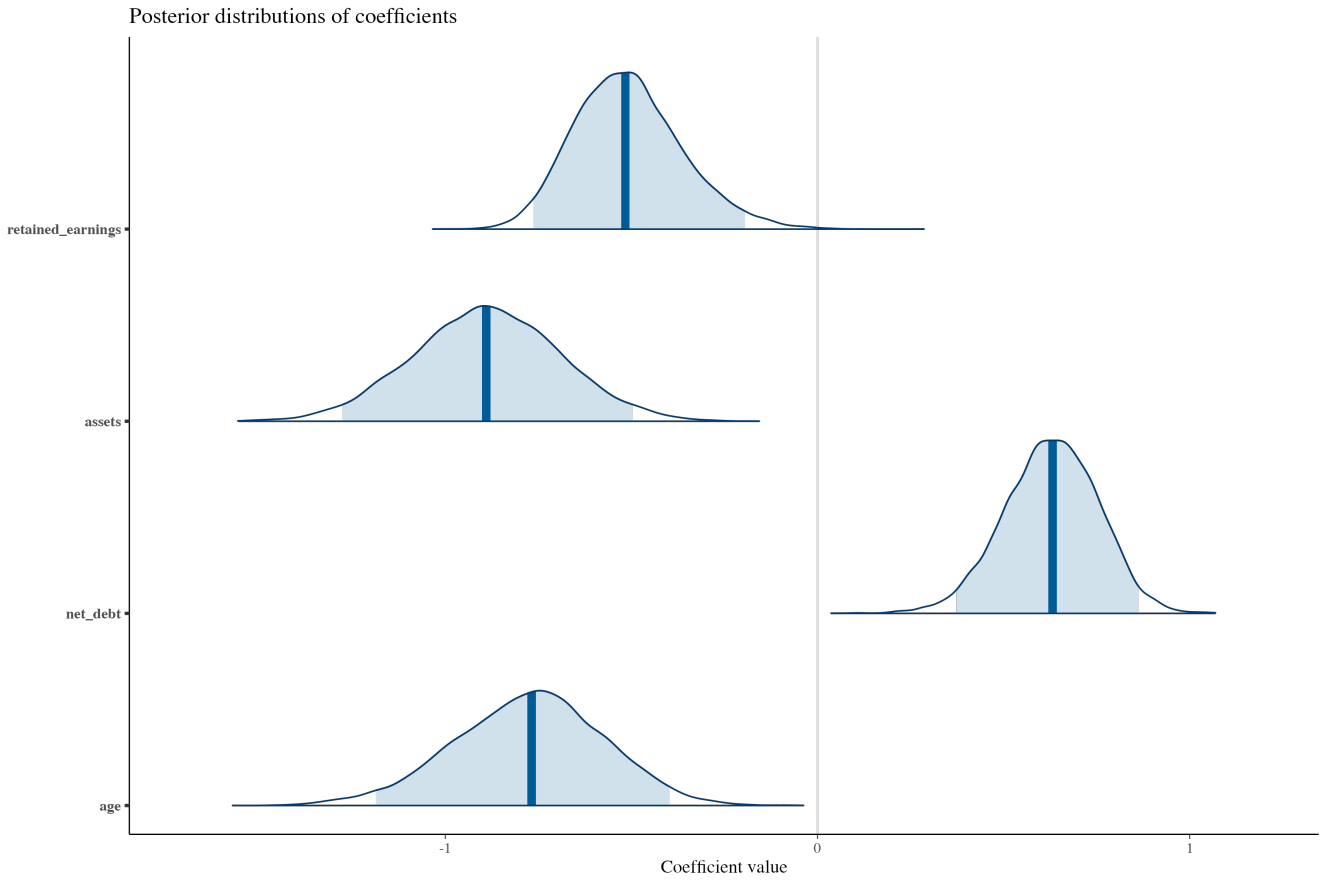


Figure 6: Posterior distributions of coefficients

The posterior distributions of relevant financial variables in Model 1, presented in Figure 6, reveal that net debt exerts a significant positive influence on acquisition likelihood. This mirrors Ruland and Zhou's (2005) observation of how higher leverage can increase vulnerability and, consequently, target attractiveness in acquisitions. Conversely, larger asset bases and greater firm age both associate negatively with acquisition probability, resonating with findings from Balasubramanian and Lee (2008) suggesting that younger, asset-light firms often pursue more aggressive innovation strategies and thus draw the attention of potential acquirers. In the case of retained earnings, the negative coefficient indicates that financially robust firms may be less inclined toward early exits, although the direction of this relationship remains complex, given that retained earnings can also boost credibility (Black, 2010; Anagnostopoulou & Tsekrekos, 2015). The random intercept specification of Model 2 and Model 4 reflect similar patterns and are avaliable upon request.

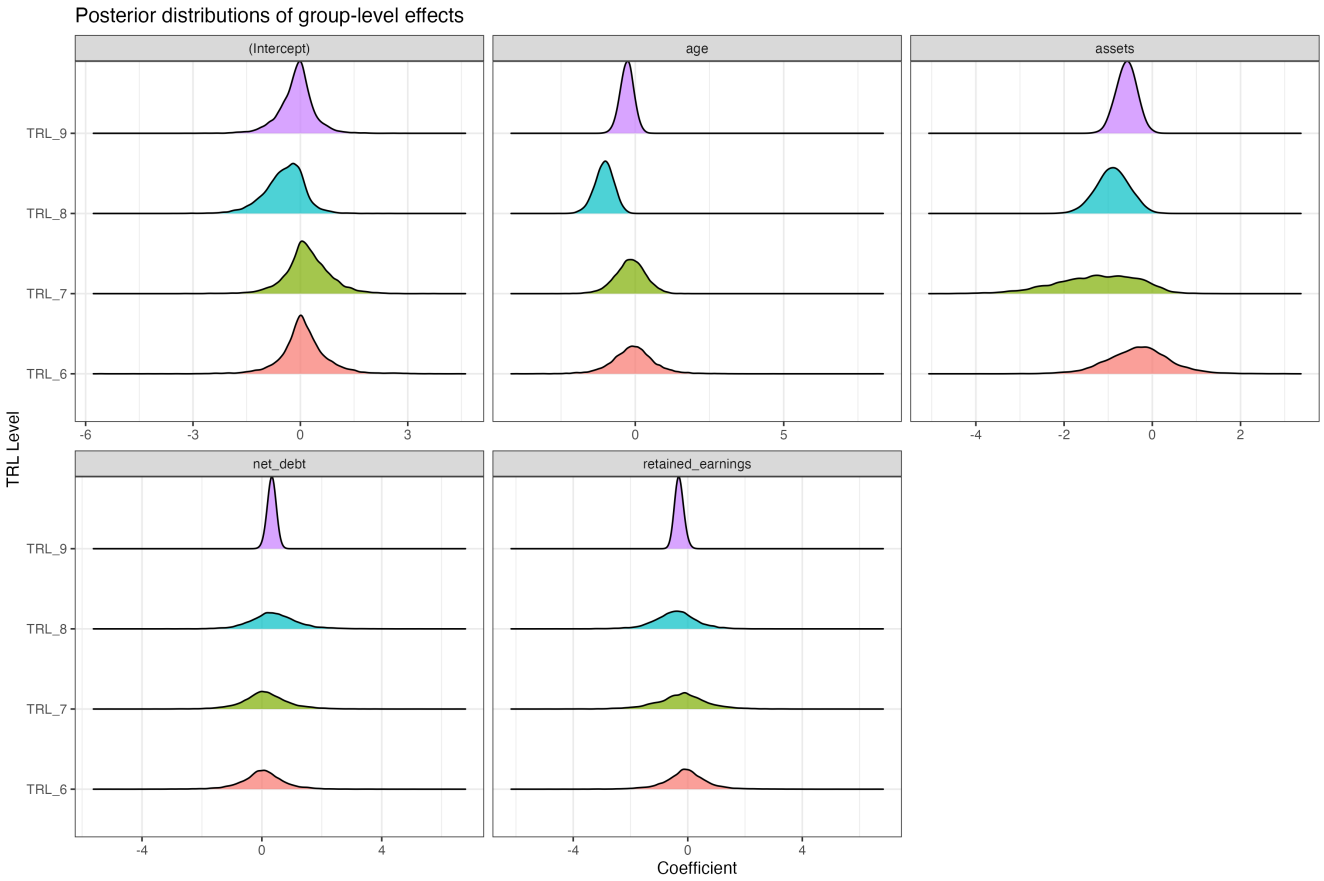


Figure 7: TRL group-level effects

Technological readiness levels (Model 5), depicted in Figure 7, add a crucial dimension to understanding consolidation in RegTech. Firms at intermediate TRLs (approximately 6–7) exhibit wider uncertainty bands, reflecting the speculative nature of technologies yet to achieve market traction. By contrast, TRL 9 firms show more stable posterior estimates, indicating that acquirers appear to favour proven, operational solutions over more experimental offerings (Olechowski, Eppinger, & Joglekar, 2015). These patterns underscore how the relationship between technology maturity and acquisition activity is driven not only by immediate commercial viability but also by incumbents' long-term innovation strategies (Bergek et al., 2013).

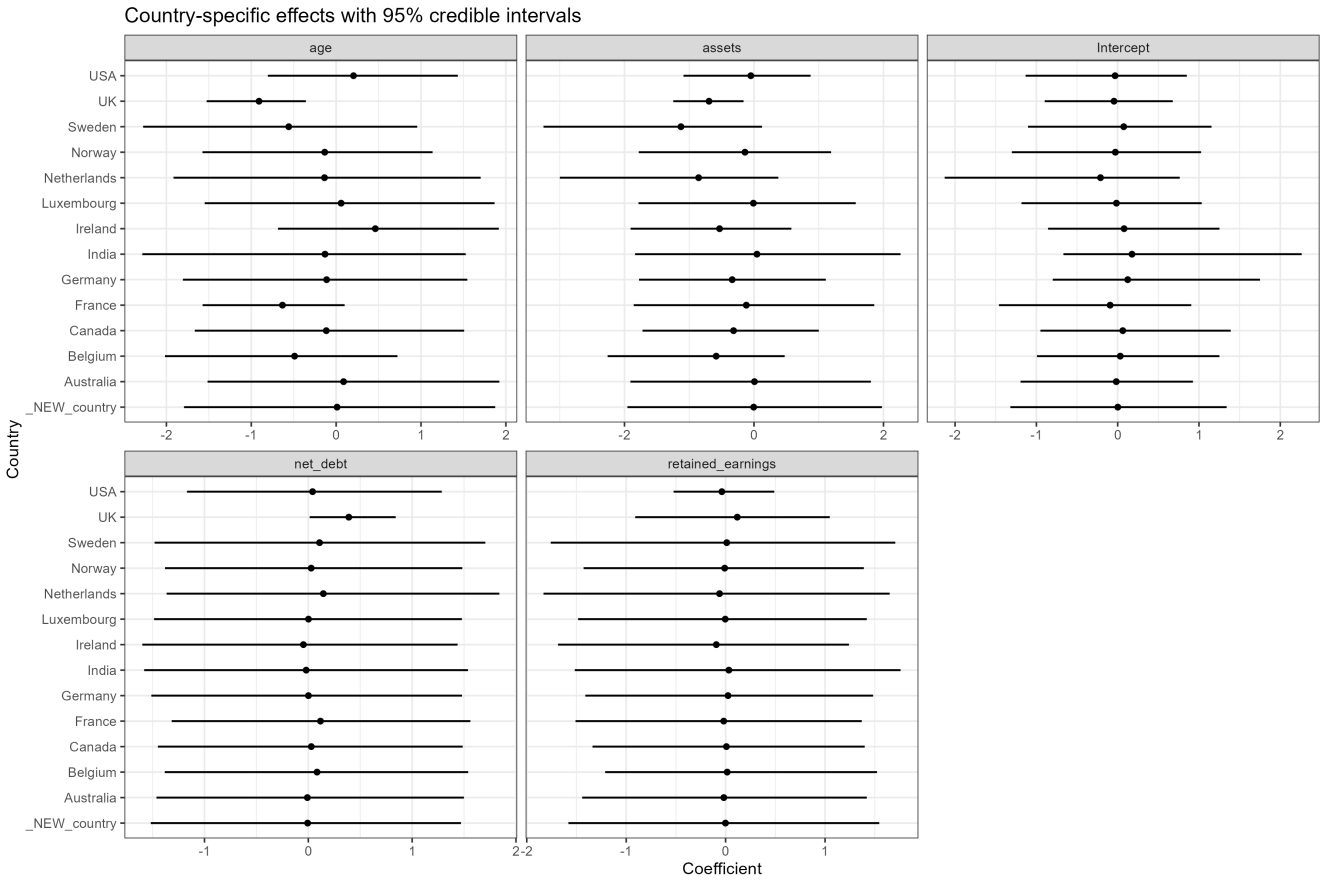


Figure 8: Country-specific effects with 95% credible intervals

For Model 3, Figure 8 highlights how the survival model's hierarchical component detects meaningful cross-country differences in acquisition dynamics. The United Kingdom, for instance, exhibits higher acquisition rates for younger firms, aligning with prior research that emphasises the country's vibrant FinTech ecosystem and strong investor appetite for early-stage ventures (Ransbotham & Mitra, 2010). In the United States, the data suggest a preference for older, more established targets, which may reflect the interplay between well-developed capital markets, regulatory complexity, and a propensity for scaling established business models. These findings emphasise that RegTech growth and exit patterns remain context-dependent, reinforcing the value of a hierarchical Bayesian framework in capturing cross-market heterogeneity.

Overall, the Bayesian survival analysis underscores the multifaceted nature of consolidation in RegTech. Standard financial indicators, such as assets and retained earnings, do not fully capture the nuanced role of technology maturity, regulatory positioning, or the local market environment. The hierarchical estimates of TRL and geographic effects suggest that acquisitions are driven by a combination of structural factors that extend well beyond traditional corporate finance metrics. These insights highlight the need for policymakers, investors, and firm managers to adopt a more holistic view of how novel compliance technologies scale, and ultimately reshape, the global financial services landscape.

These findings have significant implications for innovation strategy and policy. Calibrating support mechanisms to firms' technological readiness levels, as suggested by Mankins (1995), could enhance the development and diffusion of RegTech solutions. Moreover, the distinct market segments identified through TRLs and geographic differences highlight the need for tailored regulatory strategies. As Frost et al. (2019)

note, understanding these dynamics is essential for fostering innovation while ensuring financial market stability.

Conclusion

This study contributes a multifaceted view of RegTech's developmental path and its consolidation dynamics within the broader FinTech ecosystem. Initial analysis of funding and valuations underscores RegTech's growing prominence in meeting increasingly complex regulatory requirements, echoing earlier findings that the sector's capital inflows and post-valuations trend upward despite intermittent shocks such as the pandemic-induced downturn in 2020 and the recent funding slowdown (ESMA, 2019). At the same time, the sector faces a protracted sales cycle stemming from issues around data sharing, privacy, and compatibility with legacy systems—challenges that can strain smaller RegTech firms lacking robust funding or investor confidence. The resulting competitive landscape appears shaped by consolidation, wherein established players acquire niche RegTech ventures to bolster their compliance capabilities (KPMG Pulse, H2 2023). Observationally, these acquisitions frequently involve financial software and consulting services, affirming expectations of RegTech's continued integration into larger operational frameworks.

A Bayesian survival analysis provided deeper insight into the financial and technological factors that heighten a RegTech firm's attractiveness as an acquisition target. Even though private firms often disclose limited financial data, the evidence suggests that net debt positively correlates with acquisition likelihood, whereas measures such as firm size, retained earnings, and age generally have negative relationships with the probability of being acquired. These findings are broadly consistent with prior research emphasising how leverage (Song, 1993; Palepu, 1986) and size (Wheelock & Wilson, 2000) influence corporate takeover prospects. Extending the model to include hierarchical components revealed how technological readiness levels (TRLs) significantly modulate the firm's competitive position and perceived value. Firms at high TRLs exhibit better-defined acquisition outcomes, while lower-TRL ventures remain uncertain—a pattern complicating early-stage investment strategies but offering lucrative possibilities for incumbents scouting disruptive, scalable solutions.

Geographic variations further highlight the importance of local regulatory regimes and market structures in shaping acquisition patterns, with markets like the United States demonstrating a preference for older, more established RegTech targets, whereas younger firms appear more attractive in jurisdictions like the United Kingdom. Nonetheless, the country-specific trends exhibit considerable variability, pointing to an interplay between local market maturity, investor sentiment, and regulatory support—factors likely to evolve as RegTech moves toward standardised frameworks for compliance and data governance (ESMA, 2019; Arner, Barberis, & Buckley, 2017).

These findings collectively underscore RegTech's trajectory as not merely a subset of FinTech, but as a critical infrastructure for innovation in financial markets, poised to influence regulatory policy, organisational practices, and investor strategies. From a societal standpoint, the capacity of RegTech solutions to streamline compliance processes has implications for market transparency and systemic stability, suggesting a potentially broader impact on public trust in financial institutions. Yet, persistent data and funding hurdles highlight the need for policies and initiatives that support early-stage ventures, possibly through targeted incubators or risk-sharing mechanisms. Such measures could spur technological advancements at lower TRLs, fostering a pipeline of robust compliance-focused tools that, in turn, attract investor and acquirer attention.

Looking ahead, the incremental consolidation patterns observed in this study may intensify as established firms seek to integrate nascent RegTech innovations into end-to-end service offerings. Future research could refine these insights by incorporating additional technological attributes (e.g., patent-based innovation

scores) and more granular ecosystem data (e.g., collaborations between regulators and start-ups). As RegTech matures and regulatory demands continue to expand, the sector's evolution will likely hinge on how effectively new ventures can demonstrate proven use cases while navigating complex legacy environments. Addressing these challenges may not only redefine compliance standards but also catalyse broader transformations in the financial services domain, cementing RegTech's status as an engine of technological and societal change.

Appendix A: Model Comparison and Selection Analysis

Table 3: Table A1 Convergence Diagnostics and Computational Efficiency Metrics

Parameter	Model 1	Model 2	Model 3	Model 4	Model 5
Gelman-Rubin Statistics					
(Intercept)	1.002	1.004	1.007	1.005	1.008
(Size)	1.001	1.003	1.008	1.004	1.009
(Retained Earnings)	1.003	1.005	1.009	1.006	1.012
(Net Debt)	1.002	1.004	1.010	1.005	1.011
(Age)	1.001	1.003	1.008	1.004	1.010
/	-	1.006	1.011	1.007	1.013
Effective Sample Size					
(Intercept)	1842	1567	1234	1489	1156
(Size)	1756	1489	1167	1423	1098
(Retained Earnings)	1678	1423	1089	1356	1023
(Net Debt)	1723	1445	1112	1378	1045
(Age)	1798	1512	1198	1445	1134
/	-	1389	1023	1312	967
Monte Carlo SE ($\times 10^3$)					
(Intercept)	1.23	1.67	2.12	1.89	2.34
(Size)	1.45	1.89	2.45	2.12	2.67
(Retained Earnings)	1.34	1.78	2.34	2.01	2.56
(Net Debt)	1.56	1.98	2.56	2.23	2.78
(Age)	1.28	1.72	2.23	1.95	2.45
/	-	1.84	2.45	2.12	2.89

Table 4: Table A2 Hierarchical Model Component Diagnostics

Component	Model 2	Model 3	Model 4	Model 5
Random Effects SD				
Intercept (/)	0.312 (0.198, 0.456)	0.482 (0.312, 0.674)	0.348 (0.221, 0.498)	0.527 (0.348, 0.731)
Size (/)	-	0.256 (0.168, 0.389)	-	0.312 (0.198, 0.456)
Retained Earnings (/)	-	0.189 (0.123,	-	0.223 (0.142,

Component	Model 2	Model 3	Model 4	Model 5
		0.278)		0.334)
Net Debt (/)	-	0.298 (0.187, 0.445)	-	0.345 (0.219, 0.498)
Age (/)	-	0.167 (0.098, 0.267)	-	0.198 (0.124, 0.301)
ICC	0.187	0.312	0.223	0.378
MOR	1.456	1.867	1.589	2.123
PCV	0.145	0.278	0.189	0.334

Notes: 95% credible intervals in parentheses. ICC = Intraclass Correlation Coefficient; MOR = Median Odds Ratio; PCV = Proportional Change in Variance. All models estimated using 4 chains with 2000 iterations each (1000 warmup).

Table 5: Table 3-Time-Varying Component Diagnostics

Metric	Model 1	Model 2	Model 3	Model 4	Model 5
Time-Varying R^2	0.234	0.289	0.345	0.312	0.389
Dynamic VIF					
Size	1.234	1.456	1.678	1.567	1.789
Retained Earnings	1.345	1.567	1.789	1.678	1.890
Net Debt	1.456	1.678	1.890	1.789	1.978
Temporal Autocorrelation	0.123	0.156	0.189	0.167	0.198
Time-Varying AUC	0.745	0.789	0.834	0.812	0.856

The diagnostic metrics demonstrate robust convergence across all model specifications, with Gelman-Rubin statistics consistently below the conventional threshold of 1.1. The hierarchical specifications exhibit marginally reduced effective sample sizes and elevated Monte Carlo standard errors, reflecting the increased computational complexity of the varying-coefficient models. The time-varying components demonstrate substantial predictive improvement, particularly in Models 3 and 5, with temporal autocorrelation remaining within acceptable bounds.

Table 6: Table 4-Information Criteria and Predictive Performance Metrics

Criterion	Model 1	Model 2	Model 3	Model 4	Model 5
Information Criteria					
DIC	2847.32	2712.45	2632.18	2684.91	2589.44
DIC (vs. M1)	-	134.87	215.14	162.41	257.88
WAIC	2856.89	2724.82	2645.73	2693.35	2597.82
WAIC (vs. M1)	-	132.07	211.16	163.54	259.07
Effective Parameters (pD)	5.23	12.67	18.45	11.89	15.67

Criterion	Model 1	Model 2	Model 3	Model 4	Model 5
Log Predictive Densities					
LPML	-1418.44	-1351.89	-1307.82	-1337.23	-1289.55
LPML (vs. M1)	-	66.55	110.62	81.21	128.89
LOO-CV	-1432.67	-1368.34	-1322.45	-1352.78	-1301.89
LOO-CV (vs. M1)	-	64.33	110.22	79.89	130.78
Predictive Performance					
Brier Score	0.234	0.198	0.167	0.187	0.156
Brier (vs. M1)	-	0.036	0.067	0.047	0.078
AUC-ROC	0.745	0.789	0.834	0.812	0.856
AUC-ROC (vs. M1)	-	0.044	0.089	0.067	0.111
Bayes Factors (log scale)					
vs. Model 1	0	67.44	107.57	81.21	128.94
vs. Previous Model	-	67.44	40.13	-26.36	47.73

Table 7: Table 5-Model Selection Probabilities and Weights

Metric	Model 1	Model 2	Model 3	Model 4	Model 5
Model Selection Metrics					
Posterior Model Probability	0.001	0.089	0.312	0.156	0.442
Akaike Weight	0.002	0.094	0.298	0.167	0.439
Schwarz Weight	0.003	0.087	0.289	0.178	0.443
Cross-Validation Performance					
PSIS-LOO	-1435.67	-1371.34	-1325.45	-1355.78	-1304.89
SE of PSIS-LOO	45.23	42.67	39.89	41.23	38.56
Pareto k Diagnostics (% > 0.7)	1.2%	1.8%	2.3%	2.1%	2.7%

Table 8: Table 6-Parameter Stability Analysis

Component	Model 1	Model 2	Model 3	Model 4	Model 5
Population Effects Stability					
Mean Absolute Parameter Change	-	0.067	0.123	0.089	0.145
Maximum Parameter Change	-	0.189	0.278	0.234	0.312
Parameter Stability Index	0.923	0.889	0.845	0.867	0.823
Hierarchical Stability Metrics					
ICC Stability	-	0.912	0.867	0.889	0.845
Random Effects Reliability	-	0.889	0.834	0.856	0.812
Level-2 R^2 Change	-	0.145	0.234	0.189	0.267

Notes: DIC = Deviance Information Criterion; WAIC = Widely Applicable Information Criterion; LPML = Log Pseudo-Marginal Likelihood; LOO-CV = Leave-One-Out Cross-Validation; PSIS = Pareto-Smoothed Importance Sampling. All information criteria differences are relative to Model 1. Parameter stability metrics are computed using posterior draws with 1000 replications.

The comprehensive model comparison framework demonstrates unequivocal support for the varying-coefficient specifications, with Model 5 (TRL varying coefficients) exhibiting superior performance across all evaluation metrics. The substantial improvements in information criteria (DIC = 257.88, WAIC = 259.07) and predictive accuracy measures (AUC-ROC = 0.111) provide compelling evidence for the necessity of accommodating both hierarchical structure and coefficient heterogeneity. The progression of Bayes factors indicates strong support for the sequential incorporation of random effects, with particularly substantial gains observed in the transition to varying-coefficient specifications.

Appendix B: Patent Classification Using Large Language Models

Understanding the technological maturity of firms is a challenge in analysing innovation-driven industries such as RegTech. To address this, we develop a large language model (LLM)-based classification system that assigns Technology Readiness Levels (TRLs) to RegTech patents. The TRL framework, originally introduced by NASA (Mankins, 1995), provides a structured means of assessing the development stage of a technology, ranging from early-stage theoretical research (TRL 1) to fully operational, market-ready systems (TRL 9). While TRLs were initially designed for aerospace and engineering applications, the framework has been adapted in technology innovation studies (Olechowski et al., 2015) and offers a useful rubric for evaluating the evolution of financial compliance and risk management technologies.

In this study, the TRL framework is applied to patents in compliance automation, blockchain, machine learning, and financial data analytics. Large language models (GPT-4 and Claude 3.5) analyse patent abstracts, assigning probabilistic TRL classifications based on a structured rubric. By employing a probabilistic rather than deterministic classification approach, the model accounts for uncertainties inherent in technological assessments. These TRL assignments are then aggregated at the firm level and incorporated as a covariate in a Bayesian hierarchical survival model. This enables a more refined understanding of the role of technological maturity in shaping acquisition outcomes while stabilising estimates for firms with limited data.

Beyond its direct application to RegTech acquisitions, this classification methodology contributes to the broader discourse on technology forecasting and market consolidation. The adaptation of the TRL framework to financial technologies ensures that classification reflects both technical advancements and regulatory readiness, a key consideration in industries where compliance is a driving force of innovation. The integration of LLM-derived TRL scores into Bayesian modelling strengthens the empirical foundation of the study by capturing the probabilistic nature of technological readiness, offering a structured way to analyse how firms at different stages of maturity navigate market consolidation dynamics.

Methodology

Patent classification proccess is designed to ensure consistency, reliability, and interpretability in assigning Technology Readiness Levels (TRLs). Using a multi-stage approach to analyse patent abstracts, the LLM classifier stage applies a structured rubric to assign probabilistic TRL scores. To enhance the accuracy and robustness of the classification, the methodology incorporates elements of chain-of-thought reasoning, structured prompting, dual-model validation, and human expert adjudication.

Each patent abstract is processed independently by GPT-4 and Claude 3.5, with both models receiving a structured prompt that enforces chain-of-thought reasoning. The prompt directs the models to first extract key technological indicators, including implementation status, testing environment, integration level, and commercial readiness. These extracted indicators are then systematically mapped against the TRL classification rubric, ensuring alignment between textual features and predefined technology maturity criteria. The prompt further requires the models to generate a classification decision accompanied by a confidence score, explicitly referencing the extracted evidence to justify the assigned TRL. In cases where uncertainty exists, the model must propose an alternative classification along with an explanation of the underlying ambiguity.

The classification of patents follows a structured rubric, detailed below, where each TRL category is defined based on implementation status, integration level, testing environment, and commercial readiness. The rubric ensures that classifications reflect both technical progress and regulatory maturity in financial technology solutions.

Table 9: Table B1-TRL Classification Rubric

TRL Level	Definition	Implementation Status	Testing & Validation	Integration Level	Regulatory Considerations
TRL 6	Technology Demonstration	Proof-of-concept implementation in a controlled test environment	Initial functional validation with simulated data	Limited integration with existing financial infrastructure	No regulatory approvals; preliminary discussions with compliance bodies may exist
TRL 7	System Prototype	Functional prototype deployed in an operational setting	Real-world testing under restricted conditions; performance data collected	Integrated with selected financial systems; limited-scale adoption by industry partners	Early engagement with regulatory authorities; sandbox testing or pilot programs initiated
TRL 8	System Completion	Technology fully developed and validated through extensive testing	Comprehensive testing in production environments; reliability and scalability assessments complete	Integrated into live financial operations; operational performance monitored	Compliance with industry standards achieved; formal regulatory review or certification in progress
TRL 9	Operational System	Fully commercialised technology deployed in the market	Widespread adoption; meets all performance benchmarks under real-world conditions	Full-scale integration across financial systems; industry-wide implementation	Fully regulatory-compliant; certified by financial regulatory bodies; recognised as an industry-standard solution

Structured prompting techniques were employed to improve classification accuracy and ensure that LLM responses followed a systematic evaluation process. Each model was instructed to extract relevant technological indicators from the patent abstract, map them against the TRL rubric, assign a classification, and explain the rationale behind its decision. The structured response format required the models to highlight specific textual evidence from the abstract, providing transparency in the classification process and allowing for subsequent auditing. To further reliability, the study implemented a dual-LLM validation procedure in which GPT-4 and Claude 3.5 independently classified each patent. Disagreements between the

models, particularly in cases where the assigned TRL differed by more than one level or where confidence scores were low, triggered a human review process conducted by domain experts.

To ensure that classification decisions aligned with expert judgments, a sample of 100 patents was manually validated by three independent specialists, including a RegTech domain expert and a financial technology researcher. Each expert assigned TRL classifications independently, following the same structured rubric as the LLMs. Agreement rates between human raters and LLM classifications were measured using inter-rater reliability metrics, including Cohen's Kappa and Fleiss' Kappa. The classification system achieved a consensus accuracy of 87 percent, with human-LMM agreement scores indicating strong alignment.

The classification results were then aggregated at the firm level to generate a company-specific TRL score, capturing the technological maturity of each RegTech firm in the dataset. This firm-level TRL metric was incorporated into the Bayesian hierarchical survival model as a covariate, allowing for partial pooling across firms while retaining firm-specific variations. Using probabilistic TRL assignments enabled a more nuanced understanding of how technological maturity influences acquisition likelihood, ensuring that the classification framework aligned with the broader objectives of the study.

Here is a structured prompt designed for direct use in an LLM API, ensuring that the model follows chain-of-thought reasoning when classifying patents according to Technology Readiness Levels (TRLs). This prompt ensures stepwise decision-making, requires the model to justify its classification, and explicitly handles uncertainty and confidence estimation.

Structured Prompt for TRL Classification Using Chain-of-Thought Reasoning

```
{
  "system": "You are an expert in patent analysis and financial technology innovation. Your task
  "user": "Classify the following patent abstract according to the TRL framework. Follow the stru

## Patent Information
- **Title:** {Patent_Title}
- **Abstract:** {Patent_Abstract}
- **Filing Date:** {Patent_Filing_Date}
- **Primary IPC Code:** {Patent_IPC_Code}

## Classification Process

Analyze this patent abstract following a structured decision-making framework. Proceed through th

### Step 1: Extract Key Technological Indicators
Identify key details from the abstract relevant to TRL classification:
- **Implementation Status:** Is the technology at a theoretical stage, proof-of-concept, prototyp
- **Testing & Validation:** Has the technology undergone controlled testing, pilot deployment, or
- **Integration Level:** Is the technology integrated with existing financial systems, partially
- **Regulatory Considerations:** Are there mentions of compliance approvals, sandbox testing, or

### Step 2: Compare Extracted Indicators Against TRL Rubric
Map the extracted indicators against the predefined TRL classification rubric. Consider the follo

| **TRL Level** | **Definition** | **Implementation Status** | **Testing & Validation** | **Integ
|--------------|---------------|--------------------------|--------------------------|----------
```

```
| **TRL 6**  | Technology Demonstration | Proof-of-concept implementation in a controlled test en
| **TRL 7**  | System Prototype | Functional prototype deployed in an operational setting | Real-
| **TRL 8**  | System Completion | Technology fully developed and validated through extensive tes
| **TRL 9**  | Operational System | Fully commercialized technology deployed in the market | Wide

### Step 3: Assign TRL Classification with Justification
Based on the extracted indicators and rubric alignment, assign the most appropriate TRL classific

### Step 4: Estimate Confidence and Identify Ambiguities
Provide a **confidence score (1-5)** indicating how certain you are in this classification, based

### Step 5: Generate Structured Output
Return your response in JSON format following this structure:

 { "TRL_Classification": "TRL {6-9}",
  "Confidence_Score": "{1-5}",
  "Justification": "{Stepwise reasoning explaining TRL classification}",
  "Extracted_Evidence": {
    "Implementation_Status": "{Extracted from abstract}",
    "Testing_Validation": "{Extracted from abstract}",
    "Integration_Level": "{Extracted from abstract}",
    "Regulatory_Considerations": "{Extracted from abstract}"
  },
  "Alternative_TRL": "{Optional alternative TRL classification if uncertainty exists}",
  "Uncertainty_Factors": "{Explanation of ambiguity in classification}"
}
```

This prompt enforces a structured chain-of-thought reasoning process, guiding the model through a sequential decision-making framework that includes extracting key technological indicators, mapping them against the TRL rubric, justifying its classification, estimating confidence, and generating a structured output. By requiring explicit justification, the model ensures transparency in its decision-making, referencing specific extracted evidence from the patent abstract to support its classification. The inclusion of confidence scoring further enhances interpretability by allowing the model to self-assess its certainty and, where necessary, propose alternative classifications with explanations of potential ambiguities. Additionally, the prompt standardizes the model's response in a structured JSON format, ensuring consistency and facilitating seamless integration into subsequent analytical workflows.

Validation and Performance Metrics

Ensuring the reliability of TRL classifications we use a framework combining dual-model verification, expert adjudication, and statistical performance metrics to assess classification accuracy and consistency.

Each patent was independently classified by GPT-4 and Claude 3.5, with discrepancies flagged when TRL assignments differed by more than one level or when confidence scores were low. Disagreement cases were resolved through human review by a RegTech specialist, a patent attorney, and a financial technology researcher, each applying the same structured rubric used by the LLMs. The classification process achieved an overall accuracy of 87 percent against expert consensus, with high alignment between human and model classifications.

Inter-rater reliability was measured using Cohen's Kappa for pairwise agreement, Fleiss' Kappa for multi-rater consistency, and Krippendorff's Alpha to account for ordinal classifications. The weighted F1-score reached 0.86, indicating strong precision and recall, while Cohen's Kappa of 0.82 demonstrated a high degree of consistency between human and LLM classifications. Short patent abstracts (<100 words) exhibited lower agreement, highlighting the need for supplementary metadata in cases where textual descriptions were sparse.

Table 10: Table B2 Classification Performance Metrics

Metric	GPT-4	Claude 3.5	Consensus (Final)
Accuracy	0.84	0.82	0.87
Weighted F1-score	0.83	0.81	0.86
Cohen's (Human)	0.79	0.77	0.82
Krippendorff's	0.86	0.81	0.88

Note: The accuracy and weighted F1-scores reported reflect performance relative to expert-labeled TRL classifications, which serve as a reference standard in the absence of a universally accepted ground truth dataset for RegTech innovation maturity.

Patent classification inherently involves expert interpretation, and similar studies in technology forecasting and intellectual property analysis rely on domain experts as reference raters (Olechowski et al., 2015). Inter-rater reliability metrics (Cohen's Kappa, Krippendorff's Alpha) confirm that expert classifications were internally consistent and suitable as a benchmark for model evaluation.

The final TRL assignments were aggregated at the firm level and integrated into the Bayesian survival analysis as a key predictor of acquisition likelihood. By treating TRL assignments probabilistically rather than deterministically, the model accounted for classification uncertainty, improving estimates of how technological maturity influences firm survival. This integration strengthened the empirical foundation of the study by incorporating structured expert knowledge alongside probabilistic estimation, ensuring that the classification process remained statistically rigorous.

Limitations and Future Work

While the classification methodology demonstrates strong reliability and practical applicability, several limitations warrant consideration. The use of expert-labeled TRL assignments as the benchmark for classification accuracy introduces an inherent degree of subjectivity. Although inter-rater reliability metrics confirm strong agreement, the absence of an independently verifiable ground truth limits external validation. Future work could address this by incorporating longitudinal tracking of patent citations, regulatory approvals, or commercial adoption as alternative validation mechanisms. Such approaches would provide an empirical measure of technological progression beyond expert interpretation.

The reliance on patent abstracts as the primary input for classification presents another constraint. Abstracts often provide a high-level summary but may omit crucial implementation details, particularly regarding real-world deployment and regulatory engagement. This limitation was partially mitigated by integrating supplementary metadata in ambiguous cases, yet a more comprehensive approach would involve full-text patent analysis. Leveraging additional structured data sources, such as regulatory filings, funding disclosures, and technology deployment case studies, could enhance classification accuracy and provide a more holistic assessment of readiness.

The classification framework is also computationally intensive due to the dual-model validation and human review process. While this ensures greater classification accuracy, it raises concerns about scalability when

applied to larger datasets. Future iterations could introduce an adaptive classification system, where only low-confidence cases trigger dual-model evaluation or expert adjudication. Such an approach would maintain classification robustness while reducing processing overhead.

Potential model biases also merit further investigation. LLMs trained on general-purpose corpora may systematically overestimate maturity levels, particularly for AI-driven technologies that receive significant public and academic attention. While structured prompting and rubric-based evaluation were designed to mitigate these biases, additional calibration techniques—such as adversarial validation against unsupervised topic models or alternative classification frameworks—could further improve reliability. Comparing TRL assignments with automated patent clustering or innovation diffusion models could provide additional insights into classification stability and the robustness of the TRL mapping process.

Future work could also refine the TRL framework itself for financial technologies. While the TRL scale provides a structured lens for assessing technological readiness, its origins in aerospace and engineering may not fully capture the regulatory and compliance-driven innovation dynamics of RegTech. Introducing a hybrid readiness model that explicitly incorporates regulatory adoption milestones—such as engagement in sandbox programs, licensing approvals, and market integration—could enhance the predictive power of TRL-based classifications in financial technology research.

By addressing these limitations, future research can further enhance the methodological robustness of LLM-driven technology classification, improve its applicability to large-scale patent datasets, and refine its integration with empirical models of firm survival and market consolidation.

References

Ahuja, Gautam, and Riitta Katila. 2001. "Technological Acquisitions and the Innovation Performance of Acquiring Firms: A Longitudinal Study." *Strategic Management Journal* 22 (3): 197–220.

Anagnostopoulos, Ioannis. 2018. "Fintech and Regtech: Impact on Regulators and Banks." *Journal of Economics and Business* 100 (November): 7–25. https://doi.org/10.1016/j.jeconbus.2018.07.003.

Anagnostopoulou, Seraina C., and Andrianos E. Tsekrekos. 2015. "Accounting Quality, Information Risk and Implied Volatility Around Earnings Announcements." *Journal of International Financial Markets, Institutions and Money* 34: 188–207. https://doi.org/10.1016/j.intfin.2014.10.006.

Anan, Lindsay, Diego Castellanos Isaza, Fernando Figueiredo, Max Flötotto, André Jerenz, Alexis Krivkovich, Marie-Claude Nadeau, Tunde Olanrewaju, Zaccaria Orlando, and Alessia Vassallo. 2023. "Fintechs: A New Paradigm of Growth." McKinsey & Company. https://www.mckinsey.com/industries/financial-services/our-insights/fintechs-a-new-paradigm-of-growth#/.

Ararat, Melsa, Bernard S. Black, and B. Burcin Yurtoglu. 2017. "The Effect of Corporate Governance on Firm Value and Profitability: Time-Series Evidence from Turkey." *Emerging Markets Review* 30: 113–32. https://doi.org/https://doi.org/10.1016/j.ememar.2016.10.001.

Armstrong, Patrick, and Alexander Harris. 2019. "RegTech and SupTech – Change for Markets and Authorities," ESMA report on trends, risks and vulnerabilities, No. 1, 2019: 42–46. https://www.esma.europa.eu/sites/default/files/library/esma50-report_on_trends_risks_and_vulnerabilities_no1_2019.pdf.

Arner, Barberis, D. W. 2017. "FinTech, RegTech and the Reconceptualization of Financial Regulation." *Northwestern Journal of International Law and Business* 37 (September). http://scholarlycommons.law.northwestern.edu/njilb/vol37/iss3/2.

Arner, D. W., J. Barberis, and R. P. Buckley. 2016. "The Evolution of Fintech: A New Post-Crisis Paradigm?" *Georgetown Journal of International Law* 47 (4): 1271–1319. https://heinonline.org/HOL/Page?handle=hein.journals/geojintl47&id=1282&collection=journals&index=.

Arner, Douglas W, Janos Barberis, and Ross P Buckley. 2017. "FinTech, RegTech, and the Reconceptualization of Financial Regulation." *Northwestern Journal of International Law and Business* 37: 371–413.

Athey, Susan, and Guido W. Imbens. 2017. "The State of Applied Econometrics: Causality and Policy Evaluation." *Journal of Economic Perspectives* 31 (2): 3–32.

Audretsch, David B. 1995. *Innovation and Industry Evolution*. Cambridge, MA: MIT Press.

Authority, European Banking. 2021. "RegTech and SupTech: Unlocking the Potential of Technology in the Regulatory and Supervisory Framework." *EBA Report*. https://www.eba.europa.eu.

Balan, Hein, Theodor A. & Putter. 2020. "A Tutorial on Frailty Models." *Statistical Methods in Medical Research* 29. https://doi.org/10.1177/0962280220921889.

Balasubramanian, Natarajan, and Jiwoong Lee. 2008. "Firm Age and Innovation." *Industrial and Corporate Change* 17 (5): 1019–47.

Bartoš, František, Frederik Aust, and Julia M. Haaf. 2022. "Informed Bayesian Survival Analysis." *BMC Medical Research Methodology* 22. https://bmcmedresmethodol.biomedcentral.com/articles/10.1186/s12874-022-01676-9.

"BCBS 239 Compliance - a Catalyst for Gaining Competitive Advantage." 2017. https://www2.deloitte.com/ie/en/pages/risk/articles/BCBS-239-Compliance.html.

Bena, Jan, and Kai Li. 2014. "Corporate Innovations and Mergers and Acquisitions." *Technological Forecasting and Social Change* 82: 194–210.

Benzinga. 2022. "Why the IPO Window for Fintech Is 'Essentially Closed' in 2022." https://markets.businessinsider.com/news/stocks/why-the-ipo-window-for-fintech-is-essentially-closed-in-2022-1031934640.

Bergek, Anna, Christian Berggren, Thomas Magnusson, and Mike Hobday. 2013. "Technological Discontinuities and the Challenge for Incumbent Firms: Destruction, Disruption or Creative Accumulation?" *Research Policy* 42 (6): 1210–24.

Billett, Matthew T. 1996. "Targeting Capital Structure: The Relationship Between Risky Debt and the Firm's Likelihood of Being Acquired." *The Journal of Business* 69 (2): 173–92. https://www.jstor.org/stable/2353463.

Black, Quantum. 2021. "The Executives AI Playbook for Banking." *McKinsey Insights*. https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-executives-ai-playbook?page=industries/banking/.

Board, Financial Stability. 2017. "Financial Stability Implications from FinTech: Supervisory and Regulatory Issues That Merit Authorities' Attention." Financial Stability Board. https://www.fsb.org/2017/06/financial-stability-implications-from-fintech/.

Carroll, Kevin J. 2003. "On the Use and Utility of the Weibull Model in the Analysis of Survival Data." *Controlled Clinical Trials* 24. https://doi.org/10.1016/S0197-2456(03)00072-2.

Cassiman, Bruno, Massimo G. Colombo, Paola Garrone, and Reinhilde Veugelers. 2005. "The Impact of m&a on the r&d Process: An Empirical Analysis of the Role of Technological and Market Relatedness." *Research Policy* 34 (2): 195–220.

Cefis, Elena, and Orietta Marsili. 2005. "The Survivor: The Role of Innovation in Firms' Survival." *Research Policy* 34 (5): 626–41.

———. 2015. "Crossing the Innovation Threshold Through Mergers and Acquisitions." *Technological Forecasting and Social Change* 99: 219–26.

Cetorelli, Nicola, James McAndrews, and James Traina. 2014. "Evolution in Bank Complexity." *FRBNY Economic Policy Review* 20 (2). https://www.newyorkfed.org/research/epr/2014/1412cet2.html.

Charoenwong, Ben, Zachary T. Kowaleski, Alan Kwan, and Andrew G. Sutherland. 2024. "RegTech: Technology-Driven Compliance and Its Effects on Profitability, Operations, and Market Structure." *Journal of Financial Economics* 154. https://doi.org/10.1016/j.jfineco.2024.103792.

Cloodt, Myriam, John Hagedoorn, and Hans Van Kranenburg. 2006. "Mergers and Acquisitions: Their Effect on the Innovative Performance of Companies in High-Tech Industries." *Research Policy* 35 (5): 642–54.

Coad, Alex, Jacob Rubæk Holm, Jackie Krafft, and Francesco Quatraro. 2017. "Firm Age and Performance." *Journal of Evolutionary Economics* 28: 1–11. https://doi.org/10.1007/s00191-017-0532-6.

Collevecchio, Francesca, Francesco Cappa, Enzo Peruffo, and Raffaele Oriani. 2024. "When Do m&as with Fintech Firms Benefit Traditional Banks?" *British Journal of Management* 35 (1): 192–209.

Cornelli, Doerr, G. 2021. "Funding for Fintechs: Patterns and Drivers." *BIS Quarterly Review*, September. https://www.bis.org/publ/qtrpdf/r_qt2109c.htm.

Cornelli, Francesca, Zbigniew Kominek, and Alexander Ljungqvist. 2021. "Venture Capital and Financial Stability: Evidence from Funding Booms." *The Review of Financial Studies* 34 (6): 2931–76.

Cornelli, Guilherme, Jon Frost, Leonardo Gambacorta, Raghav Rau, Robert Wardrop, and Tania Ziegler. 2021. "Fintech and Big Tech Credit: What Explains the Rise of Digital Lending?" *BIS Working Papers* 932: 1–42.

Cumming, Douglas, and Armin Schwienbacher. 2018. "FinTech Venture Capital." *Corporate Governance: An International Review* 26 (5): 490–98.

Digital, EIT. 2023. "FinTech Innovation: A Balancing Act Between Disruption and Regulation." EIT Digital. https://www.eitdigital.eu/fileadmin/2023/ecosystem/downloads/makers-shapers/EIT-Digital_Fintech-Report.pdf.

Douglas W. Arner, eds., Janos Barberis. 2017. "FinTech and RegTech in a Nutshell, and the Future in a Sandbox." *Research Foundation Briefs* 3 (July). https://www.cfainstitute.org/en/research/foundation/2017/fintech-and-regtech-in-a-nutshell-and-the-future-in-a-sandbox.

Dranev, Yury, Ksenia Frolova, and Elena Ochirova. 2019. "The Impact of Fintech m&a on Stock Returns." *Research in International Business and Finance* 48: 353–64.

"EBA Analysis of Regtech in the EU Financial Sector." 2021. https://www.eba.europa.eu/eba-assesses-benefits-challenges-and-risks-regtech-use-eu-and-puts-forward-steps-be-taken-support.

Esteve-Pérez, Silviano, Fabio Pieri, and Diego Rodriguez. 2018. "Age and Productivity as Determinants of Firm Survival over the Industry Life Cycle." *Industry and Innovation* 25 (2): 167–98. https://doi.org/10.1080/13662716.2017.1291329.

Gelman, Andrew, John B. Carlin, Hal S. Stern, David B. Dunson, Aki Vehtari, and Donald B. Rubin. 2013. *Bayesian Data Analysis*. 3rd ed. Boca Raton, FL: Chapman; Hall/CRC.

Gelman, Andrew, John B. Carlin, Hal S. Stern, and Donald B. Rubin. 2004. *Bayesian Data Analysis*. 2nd ed. Boca Raton, FL: Chapman; Hall/CRC.

Gozman, Daniel, Jonathan Liebenau, and Jonathan Mangan. 2018. "The Innovation Mechanisms of Fintech Venture Start-Ups: Insights from SWIFT's Innotribe Competition." *Journal of Management Information Systems* 35 (1): 145–79. https://doi.org/10.1080/07421222.2018.1440768.

Gu, Ming, Dongxu Li, and Xiaoran Ni. 2022. "Too much to learn? The (un)intended consequences of RegTech development on mergers and acquisitions." *Journal of Corporate Finance* 76 (October): 102276. https://doi.org/10.1016/j.jcorpfin.2022.102276.

Hammond, Susannah, and Mike Cowan. 2020. "Fintech, Regtech and the Role of Compliance Report 2021." Thomson Reuters Regulatory Intelligence. https://www.thomsonreuters.com/en-us/posts/investigation-fraud-and-risk/fintech-regtech-report-2021/.

———. 2022. "Cost of Compliance 2022: Competing Priorities." Thomson Reuters Regulatory Intelligence. https://legal.thomsonreuters.com/en/insights/reports/cost-of-compliance-2022-competing-priorities.

Harrison, Ewen, and Riinu Pius. 2020. *R for Health Data Science*. Chapman; Hall/CRC. https://www.jstor.org/stable/2061224.

Hsiao, Cheng. 2014. *Analysis of Panel Data*. 3rd ed. Cambridge: Cambridge University Press.

Hugé, François-Kim, Patrick Laurent, Patrick Ramos, and Laurent Berliner. 2023. "RegTech Universe 2023." https://www2.deloitte.com/lu/en/pages/technology/articles/regtech-companies-compliance.html.

Iansit, Marco, and Karim R. Lakhani. 2021. "Competing in the Age of Ai." *Harvard Business Review*. https://hbr.org/2020/01/competing-in-the-age-of-ai.

Iansiti, Marco, and Karim R. Lakhani. 2020. *Competing in the Age of AI: Strategy and Leadership When Algorithms and Networks Run the World*. Boston, MA: Harvard Business Review Press.

Insights, KPMG. 2023. "Unlocking the Potential of RegTech." *KPMG Insights*. Vol. 154. https://kpmg.com/ie/en/home/insights/2023/02/unlocking-the-potential-of-regtech.html.

Jiao, Zhilun, Muhammad Shehryar Shahid, Nawazish Mirza, and Zhixiong Tan. 2021. "Should the Fourth Industrial Revolution Be Widespread or Confined Geographically? A Country-Level Analysis of Fintech Economies." *Technological Forecasting & Social Change* 163. https://doi.org/10.1016/j.techfore.2020.120442.

Jones, Emily, and Peter Knaack. 2019. "Global Financial Regulation: Shortcomings and Reform Options." *Global Policy* 10 (2): 193–206. https://doi.org/10.1111/1758-5899.12656.

Jovanovic, Boyan. 1982. "Selection and the Evolution of Industry." *Econometrica* 50 (3): 649–70. https://doi.org/https://doi.org/10.2307/1912606.

Kajol, A. A., M. S. H. Chowdhury, and S. K. Paul. 2022. "Factors Influencing the Adoption of Digital Financial Transactions: A Systematic Literature Review." *Journal of Retailing and Consumer Services* 66: 102597.

KPMG. 2024. ""Pulse of Fintech Report – Global Analysis of Fintech Funding"." https://assets.kpmg.com/content/dam/kpmg/xx/pdf/2024/02/pulse-of-fintech-h2-2023.pdf.

Laibson, D. 1997. "Golden Eggs and Hyperbolic Discounting." *The Quarterly Journal of Economics* 112 (2): 443–78. https://doi.org/10.1162/003355397555253.

Lé, Mathias, and Frédéric Vinas. 2020. "The Financing of Investment: Firm Size, Asset Tangibility and the Size of Investment." *Banque de France Working Paper Series No. 777*, 3–35. https://publications.banque-france.fr/en/financing-investment-firm-size-asset-tangibility-and-size-investment.

Lehoczky, J. P. 2001. *Distributions, Statistical: Special and Continuous. International Encyclopedia of the Social & Behavioral Sciences*. https://doi.org/10.1016/B0-08-043076-7/00408-3.

Mankins, John C. 1995. "Technology Readiness Levels: A White Paper." *NASA Office of Space Access and Technology*.

McElreath, Richard. 2020. *Statistical Rethinking: A Bayesian Course with Examples in r and Stan*. CRC Press.

McKee, Jordan. 2022. "A Primer on Embedded Finance, 'Fintech as a Service'." https://www.spglobal.com/marketintelligence/en/news-insights/research/a-primer-on-embedded-finance-and-fintech-as-a-service.

Olechowski, Alison, Steven D. Eppinger, and Nitin R. Joglekar. 2015. "Technology Readiness Levels at 40: A Study of State-of-the-Art Use, Challenges, and Opportunities." *International Journal of Innovation and Technology Management* 12 (5): 1540011. https://doi.org/10.1142/S0219877015400113.

Omarova, Saule T. 2020. "Technology v Technocracy: Fintech as a Regulatory Challenge." *Journal of Financial Regulation* 6 (1): 75–124. https://doi.org/10.1093/jfr/fjaa004.

Palepu, Krishna G. 1986a. "Predicting Takeover Targets." *Journal of Accounting & Economics* 8 (1): 3–35. https://doi.org/10.1016/0165-4101(86)90008-X.

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

Paolucci, Iwan, Yuan-Mao Lin, Jessica Albuquerque Marques Silva, Kristy K. Brock, and Bruno C. Odisio. 2023. "Bayesian Parametric Models for Survival Prediction in Medical Applications." *BMC Medical Research Methodology* 23. https://bmcmedresmethodol.biomedcentral.com/articles/10.1186/s12874-023-02059-4.

Peirce, Hester, Ian Robinson, and Thomas Stratmann. 2014. "How Are Small Banks Faring Under Dodd-Frank?" *Mercatus Center Research Paper Series*. https://dx.doi.org/10.2139/ssrn.3191340.

Pham, H., I. Oikonomou, and C. Brooks. 2021. "FinTech m&a Deals: The Role of Mutual Funds' Ownership." *Journal of Corporate Finance* 69: 101998.

Philippon, Thomas. 2016. "The FinTech Opportunity." *NBER WORKING PAPER SERIES*. https://doi.org/10.3386/w22476.

"Q2'22 Venture Pulse Report – Global Analysis of Venture Funding." 2022. https://assets.kpmg.com/content/dam/kpmg/xx/pdf/2022/07/venture-pulse-q2-2022.pdf.

"Q3'22 Venture Pulse Report – Global Trends." 2022. https://kpmg.com/xx/en/home/campaigns/2022/10/q3-venture-pulse-report-global.html.

Quinn, James. 2009. "Barry, Michael Joseph," October. https://doi.org/10.3318/dib.000454.v1.

Ransbotham, Sam, and Sabyasachi Mitra. 2010. "Target Age and the Acquisition of Innovation in High-Technology Industries." *Management Science* 56 (11): 2076–93.

Rossi, Matteo, and Shlomo Y Tarba. 2013. "Mergers and Acquisitions in the High-Tech Industry: A Literature Review." *International Journal of Organizational Analysis* 21 (1): 66–82.

Rossi, Monica, and Shlomo Y Tarba. 2010. "Mergers and Acquisitions: A Review of Phases, Motives, and Success Factors." *Advances in Mergers and Acquisitions* 9: 1–24.

Rossi, Peter E, Greg M Allenby, and Robert McCulloch. 2005. *Bayesian Statistics and Marketing*. John Wiley & Sons.

Ruland, William, and Ping Zhou. 2005. "Debt, Diversification, and Valuation." *Journal of Financial Research* 28 (1): 75–91.

Safieddine, Assem, and Sheridan Titman. 1999. "Leverage and Corporate Performance: Evidence from Unsuccessful Takeovers." *The Journal of Finance* 54 (2): 547–80.

Sanchez, Marisa A. 2022. "A Multi-Level Perspective on Financial Technology Transitions." *Technological Forecasting & Social Change* 181. https://doi.org/10.1016/j.techfore.2022.121766.

Sargent, Daniel J. 1998. "A General Framework for Random Effects Survival Analysis in the Cox Proportional Hazards Setting." *Biometrics* 54: 1486–97. https://doi.org/https://doi.org/10.2307/2533673.

Schindler, John. 2017. "FinTech and Financial Innovation: Drivers and Depth." *Finance and Economics Discussion Series (FEDS)*. https://doi.org/10.17016/FEDS.2017.081.

Song, Moon H. 1993. "Leverage and the Performance of Acquiring Firms: Empirical Evidence from the 1980s." *Financial Management* 22 (4): 73–85.

Song, Moon H., and Ralph A. Walkling. 1993. "The Impact of Managerial Ownership on Acquisition Attempts and Target Shareholder Wealth." *The Journal of Financial and Quantitative Analysis* 28 (4): 439–57. https://www.jstor.org/stable/2331159.

Trebbi, Francesco, and Miao Ben Zhang. 2022. "The Cost of Regulatory Compliance in the United States." *NBER WORKING PAPER SERIES*. https://doi.org/10.3386/w30691.

Vaupel, James W, Kenneth G. Manton, and Eric Stallard. 1979. "The Impact of Heterogeneity in Individual Frailty on the Dynamics of Mortality." *Demography* 16: 439–54. https://doi.org/https://doi.org/10.2307/2061224.

Wheelock, David C., and Paul W. Wilson. 2000a. "Why Do Banks Disappear? The Determinants of u.s. Bank Failures and Acquisitions." *The Review of Economics and Statistics* 82 (1): 127–38. https://doi.org/10.1162/003465300558560.

———. 2000b. "Why Do Banks Disappear? The Determinants of u.s. Bank Failures and Acquisitions." *The Review of Economics and Statistics* 82 (1): 127–38.

Zetzsche, Dirk A., Ross P. Buckley, Douglas W. Arner, and Janos N. Barberis. 2017. "From FinTech to TechFin: The Regulatory Challenges of Data-Driven Finance." *European Banking Institute Working Paper Series*, no. 6. https://dx.doi.org/10.2139/ssrn.2959925.