The Pricing and Economic Impact of Legal Risk *

Dean Ryu†

August, 2025

Abstract

Legal risk is an inherent feature of the market, arising from the institutional foundations that enable economic activity, yet it remains difficult to observe systematically across firms. To address this gap, this paper constructs a text-based measure of firm-level legal risk (**LRISK**) using earnings call transcripts. **LRISK** predicts class action lawsuits and exhibits sharp spikes during periods of heightened legal exposure. Legal risk has real and financial consequences: a one standard deviation increase predicts a 3–7% decline in future investment and reduces firms' reliance on debt when facing financing needs, consistent with a precautionary motive. Legal risk is positively priced, particularly after 2010 when regulatory oversight intensified. The pricing of legal risk varies across legal origins, reflecting its institutional foundations.

Keywords: Legal Risk, Natural Language Processing, Law and Finance, Asset Pricing

JEL Classification: G12, G14, K22, K41, C55

*I am grateful for the valuable feedback received from Samuel Antill, Malcolm Baker, Fred Bereskin, Justin Birru, Pedro Bordalo, John Campbell, Tom Chang (discussant), Thummim Cho, Robin Greenwood, Umit Gurun, Tarek Hassan, Jonas Heese, Rustam Jamilov, Mathias Kruttli, Max Miller, Jaideep Oberoi (discussant), Ji Yeol Jimmy Oh, Andrea Pagliuca (discussant), Ludovic Phalippou, David Schoenherr, Markus Schwedeler, Vitalii Tubdenov, Ruosen Yang (discussant), Jaesun Yun (discussant), and seminar participants at the Allied Korean Finance Associations Joint Conference, Dauphine Finance Ph.D. Workshop, Finance Theory Group Summer School, GSU-MS AI and FinTech Conference, FMA 2024 & 2025, FMA 2025 European Doctoral Consortium, International Accounting & Finance Doctoral Symposium, LAPE Spring Workshop, Monash-Paris-Warwick-Zurich-CEPR Text-As-Data Workshop, Sungkyunkwan University, TADC, University of Oxford, and USC Marshall Ph.D. Conference. Above all, I am deeply indebted to Lauren Cohen and Mungo Wilson, whose insightful guidance and unwavering encouragement were instrumental in shaping this work.

†Saïd Business School, University of Oxford **Email:** dean.ryu@sbs.ox.ac.uk

1 Introduction

A well-functioning economy relies on the rule of law. Legal and regulatory institutions such as courts, enforcement agencies, and financial regulators support economic activity by enforcing contracts and protecting property rights. At the same time, these institutions introduce legal risk — stemming from uncertainty in how rules are applied and the financial burdens that follow. The salience of this risk has increased in tandem with the rising complexity of legal systems and the tightening of regulatory scrutiny.[1] Legal risk, in this sense, is an inherent feature of modern financial markets.

Despite its importance, legal risk remains difficult to measure systematically, as it often materializes through opaque and anticipatory channels shaped by evolving statutes, regulatory discretion, and strategic compliance behavior. To address this challenge, I develop a simple text-based measure of firm-level legal risk, **LRISK**, derived from earnings call transcripts. This measure, which captures how frequently firms reference legal-related terms such as litigation, regulatory investigations, or court proceedings, is used to advance the view that legal risk is a first-order determinant of both asset prices and corporate behavior.

Specifically, using **LRISK**, the paper presents three main findings. First, **LRISK** effectively predicts future class action lawsuits and captures firm-level attention to litigation, regulatory enforcement, climate regulations, and cybersecurity threats in a timely manner. This validates **LRISK** as a reliable measure of a firm's exposure to legal risk—contrasting with standard SEC filings such as 10-Ks or 10-Qs, which often mention such risks only with a delay due to higher disclosure thresholds and more restrictive reporting practices.

Second, I examine the asset pricing implications of legal risk, showing that it carries a positive risk premium: a long-short portfolio that buys firms with high legal risk and sells firms with no legal risk (the **LAW** factor) earns an annualized return of 4.7% (t-stat: 3.6) from July 2004 to June 2024. This premium increases to 7.1% (t-stat: 4.9) in the post-crisis period, when regulatory scrutiny intensified. Firms with high legal risk are more exposed to adverse legal events such as lawsuits and enforcement actions, leading investors to demand compensation for bearing this risk. The **LAW** factor captures a unique

[1] This trend is particularly evident in the United States. The number of securities class action lawsuits has more than doubled since the 1990s, even as the number of publicly listed firms has declined (see https://corpgov.law.harvard.edu/2021/03/11/recent-trends-in-securities-class-action-litigation/). Patent disputes have tripled since 2005 (Cohen et al. 2019), and regulatory enforcement activity has accelerated in the post-crisis period. Online Appendix Section A and Figure A.1 illustrates this trend, showing the steady expansion of litigation-prone industries over the past two decades.

dimension of firm-level legal uncertainty than existing text-based measures, subsumes the effect of cyber risk, and remains largely orthogonal to standard asset pricing factors. It also avoids the common "factor zoo" critique, underscoring its distinctiveness and relevance to the financial market.

The pricing effect of legal risk appears to be driven by both systematic and idiosyncratic components. Leveraging the flexibility of textual analysis, I separately construct a macro-legal risk factor and show that it carries significant pricing implications. I also outline three possible channels through which ostensibly firm-specific legal shocks can become market-relevant. Moreover, I rule out behavioral explanations for the **LAW** factor—such as mispricing, or investor mood—as primary drivers of the observed premium. I also construct an aggregate measure of legal risk by calculating the economy-wide coverage ratio of legal terms in earnings calls. This measure captures the evolving macro-level salience of legal concerns. Importantly, it is conceptually and empirically distinct from well-known uncertainty indices such as the Economic Policy Uncertainty (EPU) index (Baker et al. 2016) and the macroeconomic uncertainty measure of Jurado et al. (2015).

Third, **LRISK** has an economic impact. Faced with heightened legal or regulatory uncertainty, firms behave precautionarily—postponing irreversible or hard-to-redeploy projects and shying away from fixed commitments that amplify downside risk. Quantitatively, a one–standard-deviation increase in **LRISK** is associated with a 5–7% reduction in investment, with the largest pullback in intangibles (R&D), where option values of waiting are greatest. On the financing margin, firms with elevated **LRISK** are significantly less likely to issue debt when they face funding deficits, consistent with tighter debt capacity, higher required yields, and concerns about legal tail events interacting with rigid repayment obligations. Local-projection estimates indicate these effects persist for multiple years, pointing to durable, rather than, transitory, distortions in capital allocation. Together, the patterns are consistent with legal risk operating as a financing friction that depresses investment.

Taken together, this paper highlights the role of legal risk as a first-order determinant of both asset prices and corporate decision-making. While legal and regulatory institutions underpin market stability, they also introduce frictions and uncertainty that shape corporate decisions and investor behavior. Recognizing the role of legal risk is essential to understanding capital allocation and the functioning of modern financial markets.

Literature review: This paper intersects several strands of literature on the role of legal systems in finance. A first body of work examines stock price reactions to lawsuits, generally finding negative effects at both the firm and industry level (e.g., Kamma et al. 1988; Gande and Lewis 2009; Hadlock and Sonti 2012). While these studies rely on event-based methodologies, this paper adopts a calendar-time portfolio approach to show that legal risk is systematically priced across firms. Relatedly, Cohen et al. (2013) show that political signals such as senators' voting behavior can significantly affect industry-level returns.

This paper also contributes to the literature on the economic effects of legal institutions and uncertainty. While earlier studies have focused on the role of legal origins and judicial efficiency in shaping capital markets (La Porta et al. 1998; Glaeser and Shleifer 2002), recent work has emphasized the economic costs of legal uncertainty itself. Lee et al. (2024) develop a model that distinguishes between diversifiable and non-diversifiable legal uncertainty and show empirically that systematic legal uncertainty reduces credit supply and investment in Korea's bankruptcy system. Complementing this, Ash et al. (2025) find that more detailed and contingent legislation in U.S. states spurs economic growth, particularly in sectors requiring relationship-specific investments. These findings underscore the importance of legal clarity for capital allocation—an insight this paper builds upon by introducing a firm-level measure of legal risk in the U.S. and demonstrating its real and financial consequences.

In corporate finance, legal risk is known to influence firm behavior. It can lead to IPO underpricing (Hughes and Thakor 1992; Lowry and Shu 2002), shape acquisition strategies (Gormley and Matsa 2011), alter disclosure practices (Skinner 1994), and raise audit costs (Shu 2000). Legal proceedings also impose indirect costs through reputational damage, managerial distraction, and heightened financing constraints.

Finally, this paper adds to the growing literature on text-based methods in finance, which use unstructured data to measure risk and predict outcomes (e.g., Baker et al. 2016; Manela and Moreira 2017; Hassan et al. 2019). More recently, earnings call transcripts have been used to quantify firm-level exposures to climate risk, cyber threats, and extreme weather (Sautner et al. 2023; Jamilov et al. 2025; Kruttli et al. 2025). A growing subset of this literature focuses on legal and litigation risk. However, many existing measures lack firm-level precision, relying instead on industry or regional proxies. For example, Francis et al. (1994) identify high-litigation sectors, while Hossain et al. (2023) use judge ideology as a geographic proxy for litigation risk. Among studies that do construct firm-level measures, Bereskin et al. (2023) focus on patent litigation using matched samples, and Bennett et al.

(2025) extract litigious language from 10-K filings based on the Loughran and McDonald (2011) dictionary. In contrast, this study develops a more comprehensive and forward-looking firm-level measure, capturing a broad range of legal risks—including class action lawsuits, regulatory investigations, and Supreme Court cases—based on bigram and trigram expressions in earnings calls. As discussed in Section 4.1, this conversational-text approach provides a unique and real-time perspective on firm-specific legal risk.

Structure of the Paper: The remainder of the paper is organized as follows. Section 2 details the construction of the firm-level legal risk measure using textual analysis of earnings call transcripts and validates the measure by showing that **LRISK** predicts class action lawsuits, supplemented with case studies that illustrate how it captures meaningful firm-level variation in legal exposure. Section 3 presents the asset pricing results, demonstrating that legal risk is a priced factor in the cross-section of stock returns and addressing concerns related to data mining and the broader "factor zoo" critique. This section also discusses potential mechanisms underlying the pricing of legal risk. Section 4 evaluates the robustness and distinctiveness of the **LRISK** factor relative to other text-based measures, rules out behavioral explanations, and considers heterogeneity across legal origins. Section 5 explores the real effects of legal risk on corporate investment and financing decisions. Section 6 concludes.

2 LRISK: Construction and Validation

2.1 Extracting Legal Risk from Corporate Earnings Conference Calls

This paper constructs a text-based measure of ex-ante legal risk at the firm level, denoted as **LRISK**, using a simple bag-of-words approach applied to quarterly earnings conference call transcripts of U.S. public firms.[2] The measure captures how frequently firms mention legal terminology—such as litigation, regulatory investigations, or court proceedings—during calls with analysts and investors. By relying on the straightforward presence

[2]Typically, firms hold quarterly earnings calls aligned with their financial reporting cycle. These calls begin with a structured management presentation covering recent performance, financial metrics, and key developments, followed by a Q&A session where analysts and investors question management. The interactive nature of these calls allows real-time insights into management's concerns, strategic priorities, and potential risk exposures. In that vein, the (transcripts of) conference calls provide a natural context to learn about the risks firms face and market participants' views thereof (Hassan et al. 2019).

of legal-related terms rather than sentiment classification or complex embeddings, **LRISK** provides a transparent and tractable way to quantify firm-level legal exposure.

Why rely on a simple bag-of-words approach rather than large language models (LLMs) such as FinBERT (Huang et al. 2023)? First, legal risk mentions are sparse in earnings calls, so the central task is detecting whether firms raise these issues at all, rather than classifying subtle sentiment. Second, the presence of legal references is inherently meaningful—mentions of lawsuits, regulatory scrutiny, or court proceedings are material events regardless of tone. Third, bag-of-words methods are transparent, computationally efficient, and widely used in related literatures. For example, Kruttli et al. (2025) identify extreme weather uncertainty by counting hurricane-related terms in transcripts, while Jamilov et al. (2025) construct a global cyber risk measure based on validated keyword frequencies. These examples demonstrate that simple, contextually grounded frequency-based approaches can yield credible and informative measures of firm-level risk.

Why base the legal risk measure on earnings calls rather than SEC filings such as 10-Ks or 10-Qs? Calls typically disclose relevant developments sooner and with greater openness than formal reports, which are highly scripted, subject to legal review, and bound by higher materiality thresholds, which can delay disclosure of ongoing disputes or anticipated legal challenges. The Apple–Qualcomm litigation offers a clear example: the case began in January 2017 and settled in April 2019, yet Apple's 10-Q and 10-K did not reference it until August and November 2018—over a year after it started. By contrast, Apple's earnings call at the end of January 2017—held just a week after the lawsuit was filed—explicitly mentioned the dispute, illustrating the timeliness with which earnings calls can address such matters. Calls also enjoy a scheduling advantage, as they are generally held a few days before the corresponding 10-K or 10-Q is released, allowing key developments to reach the market earlier. Moreover, formal filings have become increasingly unwieldy: the average 10-K is now roughly six times longer than in 1995 (Cohen et al. 2020), creating a processing bottleneck for investors, and they are often laden with dense legal jargon (Garcia et al. 2023).

[INSERT FIGURE 1 HERE]

Figure 1 shows that the sample of firms analyzed in this study is highly representative of the overall U.S. stock market. The top panel tracks the number of firms publishing earnings call transcripts from 2002 to 2024. The black solid line indicates that coverage rises steadily from approximately 55% in 2002 to over 85% by the late 2010s, with a modest

decline during the COVID-19 period. The same panel shows that these firms have collectively accounted for about 85% of total U.S. market capitalization since 2010 (red solid line). The bottom panel confirms that virtually all of the top 100 and 500 firms by market capitalization regularly publish earnings calls. This broad coverage underscores the representativeness of the sample and supports the generalizability of the **LRISK** measure across the market.

Using this comprehensive sample of firms, I analyze the textual content of their earnings call transcripts, focusing on the presence of legal terminology and related language to construct **LRISK**. The objective is to capture how frequently legal issues are discussed during these calls—for example, references to regulatory investigations, litigation threats, or changes in legal exposure.

[INSERT TABLE 1 HERE]

Table 1 provides illustrative examples of how firms articulate legal risks during earnings calls. One additional example comes from Alphabet Inc. (Google) during a 2024 earnings call, which stated: *"But it looks like the way that the Google versus DOJ search trial is going, there's a decent likelihood that the Apple ISA contract and some of the Android pre-install contracts are going to be voided at some point in the future."* This refers to the U.S. Department of Justice's antitrust lawsuit, which poses a direct threat to Google's core business model. The firm's public acknowledgment of this risk highlights how legal uncertainty—reflected in earnings call discourse—can materially affect investor perception and firm value.

More broadly, the fact that firms openly discuss legal risks in high-stakes earnings calls where their words are scrutinized by analysts, investors, and regulators suggests that these mentions are far from mere "cheap talk" (Crawford and Sobel 1982; Farrell and Rabin 1996). Earnings conference calls are typically broadcast live over the internet, which encourages management to avoid excessive legal language or jargon, making the information more accessible and transparent.[3] Therefore, mentions of legal risk in earnings call transcripts are particularly meaningful, as their frequency reflects substantive concerns rather than strategic messaging or rhetorical hedging.

Central to the analysis is a curated dictionary of legal n-grams (covering unigram, bigrams, and trigrams) that captures common patterns in legal risk language. I initiate

[3]This is in line with the 10-K sample as documented in Loughran and McDonald (2011), where the authors find that the Management's Discussion and Analysis (MD&A) section tends to contain fewer litigation-related terms compared to the full 10-K document (see Table 2 of that paper).

this process using the Loughran–McDonald (LM) Dictionary (Loughran and McDonald 2011), specifically its "litigious (Fin-Lit)" category, which includes legal terminology and litigation-related terms. The 2023 edition of the LM Dictionary contains 86,553 keywords, 903 of which are classified as litigious. While the LM dictionary provides a valuable foundation, this paper refines and extends it by constructing a dynamic set of legal n-grams tailored to the context in which legal risk is discussed. In contrast to the traditional reliance on unigrams, my approach systematically incorporates bigrams and trigrams to better capture context-specific legal expressions, thereby reducing misclassification and improving measurement precision.[4]

[INSERT TABLE 2 HERE]

With the training library of legal terms $\mathbb{L}$ established (see Table 2 and Online Appendix B), I proceed to define a simple measure of **LRISK**, representing the legal or litigation risk of firm i at quarter t as:

$$\textbf{LRISK}_{it} = \frac{\sum_{b}^{B_{it}} \Big(\mathbb{1}[b \in \mathbb{L}] \Big)}{\text{No. of sentences}} \tag{1}$$

where $b = 0, 1, ...B_{it}$ are the bigrams contained in the call of firm i at time t, $\mathbb{1}[\cdot]$ is the indicator function, $\mathbb{L}$ is the set of bigrams contained in $\mathbb{L}$. In summary, the legal risk faced by each firm at any moment is expressed as a simple sum of legal-related content, scaled by the total number of sentences in that transcript. **LRISK** can be further divided into three categories: **LRISK_COURT**, which captures references to court proceedings and judicial actions; **LRISK_TRIAL**, which focuses on mentions of trials and litigation processes; and **LRISK_TERMS**, which includes broader legal terminology such as "legal costs" and "regulatory compliance."

2.2 LRISK Predicts Future Class Action Lawsuits

This section validates **LRISK** along multiple dimensions. First, I demonstrate that **LRISK** predicts future class action lawsuit filings, confirming its value as a forward-looking measure of firms' exposure to legal risk. In addition, detailed case studies illustrate how well-known firms and industries respond to emerging legal threats, with **LRISK** exhibiting

[4]Refer to Section 4.1 for a comparison highlighting how my approach differs from that of Loughran and McDonald (2011).

sharp spikes during periods of heightened legal concern related to their business environment. Taken in conjunction with the evidence of non-overlap with other textual constructs described in Section 4.1, these findings substantiate that **LRISK** measures legal risk with precision, avoiding conflation with broader narrative or risk-related language.

To validate my measure of legal risk, I estimate a logistic model of class action lawsuit filings.[5] If **LRISK** effectively captures firm-level legal risk, it should reflect forward-looking concerns—especially those related to potential or ongoing class action litigation—discussed in earnings call transcripts, as evidenced by a significant relationship with future lawsuit filings. The specification is given by:

$$\Pr(\text{Lawsuit}_{i,t+h} = 1) = \Phi(\beta_1 \cdot \mathbf{LRISK}_{it} + \beta \cdot \text{Controls}_{it} + \delta_q + \varepsilon_{it}), \tag{2}$$

where $\Phi(\cdot)$ denotes the cumulative distribution function of the standard normal distribution. The dependent variable is a binary indicator equal to one if a class action lawsuit was filed against firm i in quarter $t + h$, and zero otherwise. It is constructed for three different horizons. When $h = 1$, the indicator equals one if a class action lawsuit is filed in the next quarter. When $h = 2$ ($h = 4$), the indicator equals one if a lawsuit is filed in any of the subsequent two (four) quarters.[6] The key explanatory variable, **LRISK**, is a standardized measure of firm-level legal risk constructed in the previous section. The vector of control variables includes firm characteristics such as firm size (log of total assets), Tobin's Q, leverage, sales growth, asset tangibility, and cash flow, all measured in quarter t (see Table 8 for details on the variables). I also include a second set of variables directly related to litigation risk. These litigation-related controls include: (i) a binary indicator for litigation-prone industries—commonly referred to as FPS (Francis et al. 1994);

[5]The data on class action lawsuits are obtained from the Stanford Law School / Cornerstone Research Securities Class Action Clearinghouse (SCAC). For firms that initiate earnings calls during the sample period, I retain all subsequent firm-quarter observations, imputing zero legal risk when an earnings call is not available. I exclude all firm-quarters prior to the first observed call and drop firms that never publish earnings calls. This design avoids excessive sample attrition while limiting extrapolation beyond observable disclosures. Compared to the more restrictive approach of dropping all observations with missing earnings calls, my strategy preserves a substantially larger sample while still maintaining a credible interpretation of silence as the absence of legal risk discussion.

[6]That is, in quarters $t + 1$ or $t + 2$ for $h = 2$ and in quarters $t + 1$, $t + 2$, $t + 3$, or $t + 4$ for $h = 4$. In cases where earnings calls and class action lawsuit filings occurred in the same calendar quarter, I determine the correct temporal ordering based on the calendar date. If the earnings call precedes the lawsuit filing within the month, I treat the legal risk disclosure as predictive. For example, Tesla had class action lawsuits filed in November 2013, October 2017, August 2018, and February 2023. For the November 2013 case, the fourth-quarter earnings call occurred three days before the lawsuit filing, so I treat that quarter's legal risk disclosure as forward-looking. In the October 2017 case, the earnings call for the fourth quarter occurred on November 1 (i.e., after the lawsuit was filed on October 10). Accordingly, I use the previous earnings call (August 2, 2017) as the relevant predictive disclosure.

(ii) a dummy variable for whether a firm has ever been the subject of a class action lawsuit (which switches to one at the time of the first filing and remains one thereafter); (iii) cumulative stock return; and (iv) return volatility over the previous quarter. The model further includes calendar-quarter fixed effects δ_q to account for time-varying macroeconomic influences. Standard errors are clustered at the firm level.

[INSERT TABLE 3 HERE]

The logistic regression results in Table 3 provide strong evidence for the predictive validity of the **LRISK** measure. Across all specifications, the coefficient on **LRISK** is positive and highly statistically significant, indicating that firms with greater legal risk language in earnings calls are more likely to face class action lawsuits in the near future. Given that class action lawsuits are relatively rare events in firm-quarter data, the model's performance is especially notable: the area under the ROC curve (AUC) is 0.746 (see Column (1) of the Table), indicating that the model correctly ranks a randomly selected sued firm above a non-sued firm approximately 75% of the time. Taken together, these results underscore that **LRISK** is a robust and forward-looking proxy for anticipated legal risk, consistent with both investor perception and future litigation activity.

The coefficient on the FPS variable—a binary indicator for litigation-prone industries such as pharmaceuticals, technology hardware, and electronics (see Online Appendix A)—is also positive and highly significant. This aligns with prior work by Francis et al. (1994) and Kim and Skinner (2012), who show that firms in certain sectors face heightened legal scrutiny. The past lawsuit indicator (a dummy variable that switches to one after a firm is first sued and remains one thereafter) is also strongly positive and significant, indicating that litigation risk tends to be persistent. Additionally, the coefficient on recent stock returns is negative and significant, consistent with the idea that lawsuits are more likely to follow steep stock price declines, which may trigger investor suspicion of managerial misconduct or disclosure failures. Related, return volatility also explains future lawsuit filings. Although not displayed in the table, firm size is consistently positive and significant across specifications, consistent with the "deep pockets" hypothesis that larger firms are more likely to be targeted given their perceived capacity to settle legal claims.

2.3 Case Study I: Brown firms and Environmental Litigation & Regulation

[INSERT FIGURE 2 HERE]

A growing body of research in finance examines the implications of climate change for financial markets, with particular attention to the risks associated with carbon emissions and regulatory oversight.[7] Stroebel and Wurgler (2021), in particular, emphasize that regulatory risk is perceived as the most immediate climate-related concern for businesses and investors. Given these insights, legal risk presents a crucial dimension of climate finance. *Brown* firms (those with high carbon footprints and heavy reliance on fossil fuels) face increasing scrutiny from regulators, litigation risks from environmental lawsuits, and reputational damage due to changing investor and consumer preferences.

Figure 2 illustrates standardized **LRISK** of major energy companies (*Brown* firms), including ConocoPhillips, ExxonMobil, and Halliburton, highlighting periods of heightened legal concerns related to climate regulations, relevant litigations, and firm-specific legal events such as M&A transactions. ConocoPhillips (upper panel) saw increased legal risk in 2002 (FTC approval of its merger with Phillips Petroleum), 2007 (Venezuelan expropriation of the Corocoro oil field that led to arbitration proceedings), 2012 (spinoff of Phillips 66 which involved complex regulatory approvals), 2015 (the company settled $11.5 million in environmental litigation over gasoline storage violations in California), 2018 (lawsuit with Oklahoma homeowners over soil and water contamination), and 2023 (scrutiny over oil sands investments led to exclusion from sustainability-focused portfolios). ExxonMobil (middle panel) saw spikes in legal risk in 2009–2010 (shareholder litigation following its $41B acquisition of XTO Energy), 2015 (a controversial $225 million settlement with New Jersey over decades of environmental damage at two refineries), and 2024 (a lawsuit filed by the State of California accusing the company of misleading consumers about the recyclability of plastic waste). Finally, Halliburton (lower panel) experienced heightened legal risk in 2004 (a multibillion-dollar settlement to resolve extensive asbestos-related litigation stemming from its former subsidiaries' use of asbestos in construction materials), 2013 (legal proceedings related to its role in the 2010 Deepwater Horizon oil spill, culminating in a $1.1B settlement in 2014), and 2016 (failed $34.6B merger with Baker Hughes,

[7] Studies such as Bolton and Kacperczyk (2021) find that firms with higher carbon emissions tend to earn higher stock returns, suggesting that investors demand compensation for exposure to carbon risk. Similarly, Pastor et al. (2022) highlight how green assets have delivered strong returns in recent years due to shifting investor preferences and regulatory developments, though their expected future returns may be lower.

which was blocked by regulators over antitrust concerns, resulting in a $3.5B termination fee).

Overall, the spikes in legal risk exposure in Figure 2 align with major legal disputes, validating the effectiveness of the firm-level legal risk measure in capturing periods of heightened regulatory and litigation risk.

2.4 Case Study II: Bank firms and Financial Regulation

The banking sector has long been a focal point for legal and regulatory risk, owing to its systemic importance and tight supervision. Legal exposure has intensified following the 2008 financial crisis, with post-crisis reforms and heightened enforcement actions increasing banks' vulnerability to litigation and reputational damage. Case studies of Citigroup, Bank of America, and Wells Fargo illustrate how legal risk has shaped firm strategy and investor perceptions over time. Notably, recent research in Fahlenbrach et al. (2012) finds that bank performance in past crises strongly predicts outcomes in future crises—suggesting that persistent risk culture and business models play a significant role in how banks navigate legal and systemic shocks.

[INSERT FIGURE 3 HERE]

This section examines Citigroup, Bank of America, and Wells Fargo to assess how legal risk has shaped their operations and investor perceptions over time. Citigroup faced major legal challenges in 2011 and 2014. In 2011, the SEC fined Citigroup $285M for misleading investors about a $1B collateralized debt obligation (CDO), concealing its role in asset selection while betting against it. In 2014, the DOJ imposed a $7B settlement for its role in selling faulty residential mortgage-backed securities (RMBS) pre-2009, including a record $4B civil penalty under FIRREA. These legal risks were reflected in earnings call transcripts, as shown in Figure 3 (Panel A). In Panel B, Bank of America's legal risk spiked during 2002–2005, 2012, and 2022. During this period, it was implicated in the Parmalat scandal, resulting in a $98.5M settlement, joined the $26B National Mortgage Settlement in 2012 over foreclosure abuses, and was fined $16.65B by the DOJ in 2014 for mortgage fraud. In Q4 2022, regulators imposed a $225M penalty for mishandling unemployment benefits during COVID-19.

Lastly, Wells Fargo has faced major legal challenges over the past decade. In 2012, as part of the National Mortgage Settlement, the bank agreed to a $5.4B settlement to address

allegations of improper foreclosure practices. In 2016, its fraudulent account scandal led to a $185M CFPB fine, followed by the Federal Reserve's 2017 asset cap. In December 2022, the CFPB imposed a $3.7B fine for consumer protection violations. In mid to late 2022, Wells Fargo faced a securities class action lawsuit (*SEB Investment Management AB v. Wells Fargo*) over allegations that it misled investors about its diversity hiring practices. These key legal events, captured in the lower panel of the Figure, corroborate that legal risk is effectively reflected in earnings conference calls.

3 The Pricing Impact of Legal Risk

3.1 Data

To examine the financial and real effects of legal risk, I match the firm-level legal risk exposure to stock market data from CRSP and accounting data from COMPUSTAT over the 2002–2024 period.[8] The CRSP dataset provides information on stock returns, market capitalization, and share characteristics, allowing for the construction of asset pricing tests such as portfolio returns and factor loadings. COMPUSTAT offers detailed firm-level fundamentals, including balance sheet and income statement items, which are essential for analyzing investment behavior, capital structure decisions, and control variables in regression analyses. This merged panel forms the empirical backbone of the paper.

3.2 Portfolio Construction and Asset Pricing Results

I assess the pricing impact of legal risk by constructing value-weighted, calendar-time portfolios sorted by firms' ex-ante legal risk exposure. To reduce noise from heterogeneous disclosure practices—where firms with stronger compliance may mention legal terms more often, while others with high latent risk may disclose less—I sort firms into

[8]While extending the analysis further back is not feasible due to the limited availability of transcripts (which begin in 2002), the 20-year sample offers a strong foundation for developing investment strategies, covering two major economic recessions (2008 and 2020). The strength of my sample lies in its coverage of the full cross-section of firms, as most publicly traded companies publish earnings calls. While Bereskin et al. (2023) focus on a carefully selected subset of firms involved in patent infringement cases using a propensity score matching algorithm, my approach complements theirs by leveraging a broader firm universe. Additionally, Figure A.1 illustrates the rising number of litigation-prone industries around the turn of the century, underscoring the growing relevance of legal risk since 2000.

terciles rather than standard sortings like quintiles or deciles. This grouping helps smooth out reporting extremes and provides a more reliable categorization of legal risk.

Following Florackis et al. (2023), portfolio 1 (L1) includes firms with no legal risk mentions in their earnings calls, capturing the extensive margin. Firms with non-zero legal risk are split into portfolios 2 (L2) and 3 (L3) based on the median value of the exposure metric. Portfolios are formed annually at the end of June and weighted by market capitalization as of the formation date. The sample begins in July 2004, using transcripts from the prior year to construct the signal.[9]

I adopt a (12-12) strategy: portfolios are constructed using legal risk information aggregated over the past 12 months (four quarters) and held for the subsequent 12 months. This approach not only aligns with standard asset pricing methodology for forming long-short portfolios but also accommodates the delayed or strategically timed disclosure of legal risks. Firms may postpone bad news to low-attention periods or prioritize other topics in earnings calls (Skinner 1994; deHaan et al. 2015). Aggregating disclosures over a one-year window mitigates such timing distortions and yields a more comprehensive measure of firm-level legal risk (See Online Appendix C for results using alternative portfolio formation strategies such as (12-3)). Consistent with convention, I exclude utilities (SIC 4900–4999) and financials (SIC 6000–6411, 6500–6553, 6700–6799) due to their distinct regulatory environments and disclosure practices. Notably, utilities represent a disproportionately large share of high legal risk mentions; excluding them enhances comparability with other cross-sectional asset pricing studies (however, unreported results show that the pricing effect becomes stronger when utilities and financial firms are included).

[INSERT TABLE 4 HERE]

Table 4 presents the results. In the baseline analysis, average monthly excess returns rise from 0.638% for low legal risk stocks to 1.035% for high legal risk stocks in value-weighted portfolios over the period from July 2004 to June 2024. The corresponding zero-cost long-short portfolio (L3 - L1) generates average return difference of 40 basis points per month (0.397% per month, or 4.7% per annum), with a Newey-West adjusted t-statistic of 3.66.[10] This strong cross-sectional relationship between legal risk and future stock returns is notable, especially considering that existing literature has not documented such a connec-

[9] While transcripts are available from 2002, their widespread adoption began in 2003. Using a one-year lookback aligns with standard practices and ensures adequate data coverage.

[10] I use 6 lags; the results remain consistent when using 3, 9, or 12 lags, and the implications remain unchanged.

tion. The premium associated with firms exposed to legal risk is economically significant, with the LAW factor exhibiting an annual Sharpe ratio of 0.713 over the sample period. Over the same period, the annual Sharpe ratios of other standard asset pricing factors, ordered by performance, are: RMW (0.672), MKT (0.621), MOM (0.077), CMA (0.007), SMB (–0.027), and HML (–0.082).[11]

Next, I examine whether existing asset pricing models account for the legal risk factor. The factor spanning tests presented in the same table indicate that the legal risk factor remains significant across various models, including the CAPM, Fama-French three-factor model, Fama-French three-factor model with momentum, Fama-French three-factor model with liquidity, and the q-factor model. The annualized alpha results suggest that a zero-cost portfolio on legal risk earns an abnormal return of 0.439% per month (5.2% per annum) when benchmarked against the CAPM. Similar results hold when using other asset pricing models, such as the Fama-French three-factor model and the q-factor model. To summarize, legal risk remains unexplained by standard asset pricing models, representing a distinct and priced source of systematic risk.

In Panel B of the same table, I demonstrate that firms generally exhibit balanced characteristics across portfolios. Firms with no exposure to legal risk tend to be smaller compared to those with some degree of legal risk. This is consistent with the previous literature that deep pockets are associated with legal risk (DuCharme et al. 2004). Finally, Panel C documents the frequency of firms positioned in each tercile. On average, 38% of companies that had no previous experience with legal risk saw a rise in their exposure within the subsequent year. Interestingly, around 33% of companies previously at the highest levels of legal risk also transitioned to other terciles. Overall, the transition matrix exercise helps alleviate concerns about certain firms consistently having higher or lower exposure to legal risk.

3.3 Addressing the Factor Zoo Critique

Although introducing a new risk factor is not the paper's main objective, it nonetheless touches upon the well-known "factor zoo" critique (Cochrane 2011). This section directly

[11] I also compute the marginal contribution of each factor to the maximum squared Sharpe ratio (MSSR) following the approach of Fama and French (2018), by comparing the full model (FF5 + MOM + LAW)'s MSSR to that of the model excluding each factor in turn. Over the full sample from 2004 to 2024, the **LAW** factor delivers the third-largest contribution to MSSR, following RMW and MKT. In the subsample from 2010 to 2024, **LAW** ranks second—behind only MKT. Interestingly, during this latter period of heightened regulatory scrutiny, RMW contributes virtually nothing to the model's MSSR.

confronts those concerns by providing rigorous evidence that the legal risk factor is not a spurious result of data mining, but instead captures economically meaningful and distinct variation in asset returns. A detailed set of supporting results is provided in Online Appendix C, with key findings briefly summarized in this section.

First, unlike standard accounting-based variables, legal risk has a clear and intuitive economic foundation, as it captures downside risk related to litigation, regulatory actions, and policy uncertainty. As such, it avoids the "Hypothesizing After Results are Known" (HARK) critique (Novy-Marx and Velikov 2025).

Second, the **LAW** factor is supported by strong statistical evidence, with a full-sample t-statistic of 3.66[12] over the 2004–2024 period, and rises to 4.91 in the post-2010 subsample (see Section 3.5), both comfortably exceeding the threshold of 3.0 recommended by Harvey et al. (2016) for identifying robust asset pricing anomalies. To further assess robustness, I conduct a series of spanning regressions of the **LAW** factor on the Fama–French three factors augmented with the momentum factor and one additional factor at a time from a zoo of 193 anomalies.[13] The t-statistic on the **LAW** factor exceeds 2.0 in all 193 regressions, and remains above 3.0 in 181 of them (see Online Appendix Figure A.4).

Third, while many accounting-based signals exhibit post-2000 decay due to declining information frictions (Chordia et al. 2014), the **LAW** factor is created from 2004 and onwards, and show robust performance during when other well-known anomalies such as SMB, HML, and MOM fail to provide robust power (see Online Appendix Figure A.3).

3.4 Isn't it Idiosyncratic?

Legal risk is often viewed in asset pricing as an idiosyncratic concern: firm-specific, transitory, and diversifiable across a broad portfolio. Under this view, lawsuits or regulatory

[12]This benchmark specification excludes firms in the utilities and financial sectors, consistent with standard asset pricing practice. When these sectors are included to maximize market coverage, the **LAW** factor delivers a slightly reduced annual premium of 4.1% over the same June 2004–July 2024 period, while maintaining strong statistical significance with a Newey-West t-statistic of 4.29.

[13]That is, I estimate the following time-series regression 193 times (with the time subscript t omitted for brevity): $\textbf{LAW} = \alpha + \beta_{MKT}\textbf{MKT} + \beta_{SMB}\textbf{SMB} + \beta_{HML}\textbf{HML} + \beta_{MOM}\textbf{MOM} + \beta_{ZOO}\textbf{ZOO}$ where **ZOO** refers to one of 193 benchmark anomaly factors sourced from Chen and Zimmermann (2022) (from the original set of 210+ signals, I exclude those with missing monthly return data over my sample period, resulting in 193 available risk factors). For each specification, I record the intercept term (α) and its associated Newey–West t-statistic which captures the component of the **LAW** factor that is not explained by the five-factor benchmark model augmented with the given anomaly. The collection of these alphas provides a measure of the extent to which the **LAW** factor is distinct from known pricing factors.

actions targeting individual firms do not warrant a risk premium. However, this paper argues that the pricing implications of legal risk reflect both idiosyncratic and systematic components. In particular, regulatory actions and legal developments may affect entire industries—or even the broader market—introducing persistent, non-diversifiable risk that is priced by investors.

[INSERT TABLE 5 HERE]

The text-based construction of the **LAW** factor enables flexible measurement of legal risk at both the macro and micro levels. To operationalize this, I decompose the long-side portfolio exposures using the legal content taxonomy in Table 2, which classifies legal language into three subcomponents: **LRISK_COURT** (references to courts and regulators), **LRISK_TRIAL** (litigation-specific terminology), and **LRISK_TERMS** (broader legal language). Separate long–short portfolios are then formed for each subcomponent to assess their distinct pricing implications.[14]

Panel A of Table 5 provides a diagnostic check of whether the macro- and micro-based classifications coincide. The joint distribution shows that just under half of the sample lies on the diagonal (29.4 + 8.5 + 10.5 = 48.4%), while a slight majority falls in the off-diagonal cells. These mismatches indicate that the two approaches frequently disagree: firms judged to have no macro-level legal risk are often flagged as medium or high risk by the micro measure, and vice versa. This systematic divergence underscores that the two dimensions are related but distinct—macro-level terms capture broad regulatory and institutional salience, whereas micro-level terms emphasize firm-specific litigation intensity.

Panel B then turns to the pricing evidence. Here, portfolios of firms with non-zero macro-level legal risk (**LAW_COURT**) earn significantly higher excess returns and positive alphas relative to the benchmark zero-risk group. Together, the results in Table 5 show that macro and micro measures are not interchangeable but complementary: Panel A demonstrates that they do not fully overlap, while Panel B confirms that at least the macro dimension is priced in asset markets.

In the rest of this section, I also outline three complementary mechanisms through which

[14] In forming subcomponent factors, the long side includes firms with any legal risk mentions exclusively in the relevant cluster (e.g., **LRISK_COURT**), while the short side recycles the benchmark no-risk group (L1)—firms with no legal risk mentions in any category. Holding L1 fixed across all constructions avoids contamination from unrelated exposure and ensures that return differences reflect only variation in the high-risk group. The resulting long–short portfolio on macro-level legal risk is labeled **LAW_COURT**.

idiosyncratic legal shocks can propagate more broadly or signal systemic developments, thereby generating economy-wide pricing implications.

Economic Spillovers Firm-level legal shocks can have broader economic consequences when they affect companies embedded in interconnected production or financial networks. Through input-output relationships, long-term contracts, and shared operational dependencies, legal disruptions may alter the behavior of customers, suppliers, and competitors. These indirect effects can amplify the original shock, leading to second-order impacts that extend well beyond the initiating firm. The severity of such spillovers depends on the structure of the network and the degree of frictions—such as limited substitutability, delayed adjustment, or coordination challenges—that constrain the system's ability to absorb and reallocate risk. In this vein, legal risk serves as a transmission channel through which localized events give rise to macro-relevant outcomes.

Belief Updating About Aggregate Legal Uncertainty Firm-specific legal events may also function as informational signals about broader shifts in the legal or regulatory environment. Following the framework of Savor and Wilson (2016), investors observing such events face a signal extraction problem: they must infer whether a lawsuit, investigation, or enforcement action reflects isolated firm behavior or emerging systemic trends. For instance, the U.S. Department of Justice's 2020 antitrust lawsuit against Google heightened concerns about a broader regulatory crackdown on Big Tech. In response, peer firms such as Amazon and Meta experienced valuation declines, suggesting that market participants revised their expectations about future legal scrutiny across the sector. In this way, legal shocks at the firm level may prompt belief updating about the distribution and intensity of legal risk economy-wide.

Granularity A third channel through which legal risk may generate systematic effects is granularity. As shown in Gabaix (2011), idiosyncratic shocks to the largest firms can influence aggregate outcomes in economies with fat-tailed firm size distributions. Legal risk becomes systemically relevant when it disproportionately affects these dominant firms. A class action lawsuit, regulatory enforcement action, or major litigation involving a firm like Amazon, Meta, or JPMorgan can have macroeconomic impacts simply due to the scale of the firm involved.

This granular mechanism also aligns with the observed return asymmetries in the data. Firms with high legal exposure exhibit more negatively skewed return distributions,[15] indicating a higher likelihood of extreme negative outcomes. Rational investors, anticipating this asymmetry, demand compensation for holding such firms. Legal risk thus operates as a granular shock—idiosyncratic in origin but macro-relevant in consequence—particularly when concentrated among the largest firms in the economy.

3.5 Time-variation of Aggregate Legal Risk

[INSERT FIGURE 4 HERE]

This section examines fluctuations in aggregate legal risk and their broader implications. To do so, I construct an *aggregate* measure of legal risk by classifying firms each quarter as either non-zero (mentioning legal risk in at least one of the past four earnings calls) or zero (no mentions). The coverage ratio—defined as the share of non-zero firms—captures the extensive margin of legal risk in the corporate sector.

Figure 4 presents two panels. The top panel applies the structural break procedure of Bai and Perron (2003) to the standardized coverage ratio (Aggregate LRISK), identifying four statistically significant breakpoints: December 2007, March 2013, June 2017, and June 2020. These align closely with key regulatory and macroeconomic events. From 2003 to 2007, aggregate legal risk trended upward amid both regulatory tightening (e.g., post-SOX compliance requirements) and deregulatory shifts (e.g., GLBA and Universal Demand laws).[16] Following the Global Financial Crisis, reforms such as Dodd-Frank's whistleblower provisions and heightened SEC enforcement sustained elevated legal risk.

[15] I test this by comparing the ex-ante return skewness of portfolios L1 (no legal risk) and L3 (high legal risk). Following Jondeau et al. (2019), I compute monthly skewness for each stock i as

$$Sk_{i,t} = \sum_{d=1}^{D_t} \left(\frac{r_{i,d} - \overline{r}_{i,t}}{\sigma_{i,t}} \right)^3,$$

where $r_{i,d}$ is the daily excess return, $\overline{r}_{i,t}$ is the monthly mean, and $\sigma_{i,t}$ is the standard deviation over D_t trading days. The results, suppressed for brevity, show that firms in L3 consistently exhibit lower skewness than those in L1, indicating greater downside risk. Importantly, legal risk is not merely a proxy for skewness: co-skewness-based models (Harvey and Siddique 2000) do not explain the pricing of the LAW factor, suggesting that legal risk captures distinct, priced information.

[16] While SOX increased disclosure obligations and board accountability, deregulatory reforms reduced litigation exposure. On balance, the compliance effects of SOX appear to have prevailed.

A brief dip around 2012–2013 may reflect regulatory adjustment, but legal risk rose again from 2018 onward. Key developments include the Supreme Court's *Cyan (2018) ruling*, which enabled forum shopping in securities class actions, and the growing legal salience of cybersecurity—driven by evolving SEC rules and major data breaches. The COVID-19 shock further intensified legal exposure.

The bottom panel compares this legal risk measure (LRISK) with two established uncertainty indices: Economic Policy Uncertainty (Baker et al. 2016) and Macroeconomic Uncertainty (Jurado et al. 2015). LRISK shows only moderate correlation with these indices (0.3–0.6), indicating that it captures a distinct dimension of firm-level risk.

[INSERT TABLE 6 HERE]

Consistent with an increase in aggregate legal risk following the Global Financial Crisis (GFC), Table 6 reports a stronger pricing effect of legal risk in the 2010–2024 subsample. The **LAW** factor now earns an annualized return of 7.1% (t-stat: 4.91), driven by the strong performance differential across legal risk portfolios. Notably, the enhanced performance of the strategy is largely attributable to the short leg of the portfolio: firms with no disclosed legal risk (L1) exhibit significantly negative alphas. This suggests that as aggregate legal risk increased, investors became more willing to accept lower returns from firms perceived to have minimal legal exposure. This pricing dynamic reinforces the notion that legal risk has become a salient dimension of cross-sectional return variation in the post-GFC era.

4 Discussion

This section compares the text-based legal risk exposure measure with existing firm-level text-based metrics to establish its distinct role in capturing legal risk-related components. It also rules out potential behavioral explanations. Finally, the section adopts a comparative economics perspective to examine how legal origins shape the pricing of legal risk, highlighting differences between common law and civil law systems.

4.1 Comparison with Other Text-based Measures

How the legal risk measure relates to the Loughran-McDonald measure: My work builds on the pioneering research of Loughran and McDonald (2011) (henceforth LM). While legal terms such as litigation and plaintiff appear in both the "Fin-Neg" and "Fin-Lit" categories proposed by LM, I tailor the approach to the spoken dynamics of earnings calls in two major ways, and these refinements provide a more precise and contextually relevant measure of legal language in financial discourse.

Many keywords in LM's Fin-Lit dictionary are formal and specialized: they are terms that are more likely to be found in legal contracts or court rulings rather than in real-time corporate discussions. To illustrate this distinction, consider a few examples. The LM dictionary includes complex legal terms such as *"appurtenance"* (referring to a right, privilege, or improvement associated with a piece of property) and *"conveniens"* (derived from *forum non conveniens*, a legal doctrine that determines the most appropriate or convenient forum for a case). While these terms may be useful in legal scholarship, they are unlikely to appear in earnings calls, where executives communicate in a more accessible and practical manner. Similarly, complex legal verbs such as *"exculpate"*, *"indemnify"*, and *"usurp"* are rarely used in corporate discussions and thus are excluded from the legal dictionary.

Furthermore, my methodology highlights the importance of bigrams (or trigrams) over unigrams in analyzing legal language. While both the LM dictionary and my legal dictionary include unigrams such as antitrust, lawsuit, litigation, plaintiff, prosecutor, and settlement (all of which clearly signal legal risks associated with a firm), many legal terms related to judicial entities, legal proceedings, and legal concepts (the three broad categories in the keyword library) are often overlooked when relying solely on unigrams. For instance, mentioning the term *"Department of Justice"* during earnings calls may suggest potential regulatory scrutiny or legal challenges related to a firm's future operations. However, a unigram-based analysis would fail to capture the phrase in its entirety, missing the broader legal context it conveys. Additionally, relying solely on unigrams can lead to misinterpretation, as they may fail to distinguish between a positive statement about a valuable *"patent"* and a legal concern involving *"patent infringement,"* thereby distorting the assessment of legal risk.[17] To conclude, these two key refinements enhance the ability to measure legal risk more effectively, contributing to the pricing of legal risk in this work, in contrast to prior studies.[18]

[17] For a related discussion on the advantages of using bigram or trigram approaches over unigrams, see Caldara and Iacoviello (2022).

[18] On page 55 of Loughran and McDonald (2011), the authors document that a calendar-time portfolio

[INSERT FIGURE 5 HERE]

How the legal risk measure relates to the Political Risk and Regulatory Burden Measure: In a related study, Hassan et al. (2019) construct several text-based measures of risk exposure, including political risk (**PRISK**), non-political risk (**NPRISK**), and overall risk (**RISK**), each accompanied by a corresponding sentiment measure: **PSENT** (political sentiment), **NPSENT** (non-political sentiment), and **SENT** (overall sentiment).

To assess whether **LRISK** captures a unique dimension of legal risk, I compute its correlation with these related measures, presenting the results in Figure 5. First, I find that the composite measure **LRISK** is highly correlated with its key components, **LRISK_COURT** and **LRISK_TRIAL**, with correlations between 0.80 and 0.90 (**LRISK_TERMS**, which includes general legal phrases such as "legal costs" and "legal dimension", exhibits more modest correlations). Most importantly, **LRISK** and its components show low correlations with other text-based risk measures such as Political Risk / Sentiment (Hassan et al. 2019) and regulatory burdens (Kalmenovitz 2023), reinforcing that my legal risk metric captures a distinct legal dimension that firms face in the market.

How the legal risk measure relates to the Cyber Risk measure: In a series of recent works, researchers have developed firm-level measures of cyber risk to assess its impact on financial markets. Florackis et al. (2023) analyze 10-K filings, Jiang et al. (2024) use machine learning techniques on firm characteristics, and Jamilov et al. (2025) apply natural language processing to earnings calls to construct firm-level cyber risk measures and assess their financial impacts.

I compare the legal risk measure with the cyber risk measure developed by Jamilov et al. (2025). Since both measures are derived from the same textual source (earnings call transcripts), their cyber risk search terms serve as a natural benchmark for my study. Specifically, I construct a cyber risk dictionary, $\mathbb{C}$, based on Table 2 of Jamilov et al. (2025), along with a law-specific cyber risk dictionary $\mathbb{C}_{\text{LAW}}$ to obtain **Cyber Risk** and **Cyber Risk Law**, respectively.

[INSERT TABLE 7 HERE]

strategy based on negative words ("Fin-Neg") does not generate significant abnormal returns: *"We calculate the Fama and French (1993) four-factor portfolio returns generated by taking a long position in stocks with a low negative word count and a short position in stocks with a high negative count. [...] [N]one of the [alpha] values are statistically significant. Hence, after controlling for various factors, the relation between 1-year returns and negative word counts is not enough to warrant active trading by investors."* Similarly, Loughran and McDonald (2013) find that legal word lists are not significantly associated with first-day IPO returns.

First, I observe that **LRISK** has essentially no correlation with **Cyber Risk** but exhibits a modest correlation of 0.42 with **Cyber Risk Law**, as shown in Panel A of Table 7. This result is expected, as cyber risk and legal risk capture distinct dimensions. In the Online Appendix E, I also provide firm-level evidence using SolarWinds, which was the target of major cyberattacks, to illustrate that legal risk exposure and cyber risk exposure represent separate aspects of risk.

The legal components of cyber risk, represented by **Cyber Risk Law**, share many common keywords—such as bankruptcy court, lawsuit, legal claim, and plaintiff—which naturally explains this correlation. In Panel B of the same table, I implement a long-short portfolio strategy, as documented in Section 3, and find that both **LRISK** (replicating the results from Table 4) and **Cyber Risk** are priced, whereas **Cyber Risk Law** is not. The pricing of **Cyber Risk** independently affirms the findings of Florackis et al. (2023) and Jiang et al. (2024) that cybersecurity matter is a priced source of risk.

Importantly, Panel C presents the results of a factor spanning test, where the **LAW** factor (constructed from firms' exposure to **LRISK**) explains both factors constructed from **Cyber Risk** and **Cyber Risk Law**. Crucially, it is not subsumed by the presence of these measures under the Fama-French three-factor model augmented with momentum. Equation (1) shows that the legal risk factor loads positively on both **Cyber Risk** and **Cyber Risk Law**, yet remains distinct and not fully captured by these factors. Conversely, Equations (2) and (3) in the same panel demonstrate that the presence of the law factor (**LRISK**) not only loads positively on **Cyber Risk** and **Cyber Risk Law** as expected, but also subsumes these factors. Given that cyber risk is a type of operational risk (Kamiya et al. 2021), which in turn is a sub-component of legal risk, it follows that legal risk is a broader concept and thus encompasses the effects of cyber risk.

4.2 Behavioral Explanations?

One potential concern is that the observed legal risk premium may reflect behavioral biases rather than compensation for systematic risk. However, my empirical analyses point against this interpretation.

Stambaugh et al. (2012) argue that anomaly returns are significantly influenced by investor sentiment—particularly that short-leg returns are lower and long-short spreads narrower following high-sentiment periods, consistent with mispricing and limits to arbitrage. In contrast, I find that neither the long-leg, short-leg, nor long-short portfolio

returns of the LAW factor exhibit sensitivity to market-wide sentiment, as measured by Huang et al. (2015). This suggests that the LAW premium is unlikely to be driven by sentiment-induced mispricing. Moreover, I show that existing behavioral and mispricing-based factors—such as PEAD and FIN from Daniel et al. (2020), and MGMT and PERF from Stambaugh (2017)—are unable to subsume the LAW factor. Even after controlling for the aggregate disagreement index proposed by Huang et al. (2021), which captures belief dispersion by aggregating 24 individual disagreement measures using a partial least squares (PLS) method, the alpha on LAW remains economically large and statistically significant. These findings, presented in Table A.2 in the Online Appendix, suggest that legal risk, as captured by the LAW factor, is unrelated to sentiment, mispricing, or belief heterogeneity as proxied by analyst disagreement.

Furthermore, if the returns to the LAW factor are primarily driven by mispricing linked to investor mood, as suggested by Birru (2018), then the LAW factor should exhibit stronger returns on Fridays.[19] However, my empirical results in Table A.3 in the Online Appendix indicate the opposite pattern: LAW returns are significantly positive on Mondays, with no statistically significant improvement on Fridays. Limited attention also plays a limited role here. According to

Finally, if the return premium on legal risk were purely driven by mispricing or behavioral biases, one would not expect firms to alter their real economic decisions in response. However, the evidence to be presented in the next Section shows that firms internalize legal risk as real and costly. These real effects reinforce the interpretation that legal risk is a priced source of systematic risk, rather than a behavioral anomaly.

4.3 Legal Origins and the Cross-Country Pricing of Legal Risk

Legal systems around the world broadly fall into two traditions: common law, as seen in the United States, and civil law, as found in countries like Germany and France. Common law systems emphasize judicial precedent, decentralized adjudication, and strong private enforcement mechanisms, including class actions and broad discovery rights. In contrast, civil law systems rely heavily on codified statutes, state-controlled judges, and limited private litigation avenues, with a greater emphasis on public enforcement. As a result, firms

[19] Birru (2018) documents that investor mood improves throughout the week, leading speculative stocks to outperform on Fridays when mood is highest. Applying this framework, if high legal risk firms represent the speculative leg—given that lawsuits, regulatory actions, and litigation outcomes introduce significant uncertainty into their future cash flows and valuations, making them more difficult to price and more sensitive to shifts in investor sentiment—then LAW returns should be elevated at the end of the week.

in common law countries tend to face more frequent and visible legal disputes, while those in civil law systems often operate under stricter procedural rules and more bureaucratic forms of legal accountability.

This institutional divergence has implications for how legal risk is perceived and priced in financial markets. Drawing on Glaeser and Shleifer (2002), the civil law tradition developed as a way to protect state-controlled judges from local coercion, often at the cost of legal flexibility and transparency. Unlike the U.S., where independent juries and private lawsuits can produce large, uncertain outcomes (i.e., legal tail risk), France and Germany channel disputes through formal and predictable procedures. Furthermore, as La Porta et al. (1998) and Djankov et al. (2003) show, civil law countries generally offer weaker investor protection, making it harder for shareholders to sue or extract information about pending legal exposure. These factors—less uncertainty, reduced disclosure, and lower shareholder litigation intensity—diminish both the visibility and materiality of legal risk to outside investors, potentially muting its impact on asset prices.

Consistent with these institutional differences, my empirical results show that legal risk is priced in the United States, but not in civil law countries such as Germany or France (results are unreported but available upon request). In these civil law countries, legal risk does not appear to be a salient dimension of priced risk in cross-sectional stock returns. This finding supports the view that differences in legal origin and enforcement structures shape how legal risk enters into firm valuation, and that legal risk may simply not be priced in countries where it is procedurally muted and economically less consequential.

5 The Economic Impact of Legal Risk

Firms often respond to heightened uncertainty by deferring or reducing investment, particularly when projects are irreversible and involve long-term commitments. In such settings, the option to wait becomes valuable, as premature capital commitment entails the loss of flexibility.[20] Consistent with these insights, legal risk represents a novel and economically significant channel through which uncertainty influences firm behavior. By introducing ambiguity into the regulatory and litigation environment, legal risk increases

[20] McDonald and Siegel (1986) show that firms optimally delay investment until expected returns substantially exceed costs, reflecting this option value. Pindyck (1988) similarly finds that uncertainty raises the opportunity cost of investment, prompting firms to scale back capacity expansion and place greater weight on growth options. Bernanke (1983) emphasizes that when information arrives gradually, uncertainty about future conditions can depress investment by increasing the value of waiting.

perceived volatility, raises the cost of capital, and may lead firms to delay investment, reallocate resources, or reduce innovative activity. Online Appendix G provides simple real options frameworks to show that legal risk increases the value of waiting and thereby discouraging investment.

[INSERT TABLE 8 HERE]

To empirically assess the relationship between investment and legal risk at the firm level, I estimate

$$\text{INV}_{i,t+1} = \alpha + \beta\, \text{LRISK}_{i,t} + \gamma\, \mathbf{X}_{i,t} + \theta_i + \eta_t + \varepsilon_{i,t+1} \tag{3}$$

where $\text{INV}_{i,t}$ denotes investment for firm i at year t, $\text{LRISK}_{i,t}$ represents firm-level legal risk exposure, measured as the average across all quarters within a given year. $\mathbf{X}_{i,t}$ is a vector of control variables. I add firm fixed effects (θ_i), year fixed effects (η_t). This procedure ensures that the results are not attributed to time-invariant firm characteristics, or common macroeconomic shocks. Also, independent variables are lagged one year to limit concerns of reverse causality.

For the dependent variable (INV) I use three distinct measures of investment, each scaled by beginning-of-year total assets. First, CAPX + AQC captures tangible investment, including capital expenditures and acquisitions, and reflects the firm's spending on physical assets and M&A activity. Second, R&D captures intangible investment, measuring expenditures related to innovation, product development, and technological advancement. Finally, I construct a total investment measure by summing CAPX, AQC, and R&D, thereby capturing both tangible and intangible components of investment. These broader metrics provide a more comprehensive view of firm-level investment activity (see Table 8 for details). For ease of interpretation, all investment variables are expressed in percentage terms, and legal risk is standardized.

[INSERT TABLE 9 HERE]

Table 9 presents the results. Panel A examines the benchmark specification. Across all three measures of firm-level investment—tangible (CAPX + AQC), intangible (R&D), and total investment—the regressions reveal a consistent pattern: higher legal risk significantly predicts lower future investment. For economic interpretation, consider column

(1), where the dependent variable is tangible investment. A one standard deviation increase in legal risk is associated with an approximate 3.8% decline in tangible investment in the following year.[21] The effect is even more pronounced when focusing on intangible investment. Column (2) shows that a one standard deviation increase in legal risk predicts a 6.6% decline in R&D expenditures.[22] Finally, column (3) shows that legal risk is associated with a 5.0% reduction in total investment.[23]

Panel B addresses potential endogeneity by controlling for the persistence of investment dynamics. Following Eberly et al. (2012), I include up to three lags of the dependent variable. The coefficients on legal risk remain negative and statistically significant, though slightly attenuated in magnitude, suggesting that the effect is not solely driven by past investment momentum. Panel C replaces firm fixed effects with two-digit SIC industry fixed effects ($\delta_{s(i)}$), where $s(i)$ maps firm i to its industry. The negative relationship between legal risk and investment becomes even stronger in magnitude across all specifications. However, the R^2 values decline, reflecting the loss of within-firm variation captured by the more granular firm fixed effects. Finally, Panel D tests robustness to alternative proxies for financial constraints by replacing the Whited and Wu (2006) index (WW) with Ohlson's O-score. The results remain qualitatively and quantitatively similar, underscoring that the observed investment effects of legal risk are not mechanically driven by the choice of control for firm-level financial vulnerability.

[INSERT FIGURE 6 HERE]

To investigate investment dynamics, I estimate impulse response functions using a panel data adaptation of the local projection method developed by Jordà (2005), as shown in Figure 6. Specifically, for each yearly forecast horizon $h = 1, ..., 5$, I estimate the panel regression of the form: $\text{INV}_{i,t+h} = \alpha + \sum_{j=0}^{2} \beta_j \, \text{LRISK}_{i,t-j} + \gamma \, \mathbf{X}_{i,t} + \sum_{j=0}^{2} \delta_j \, \text{INV}_{i,t-j} + \theta_i + \eta_t + \varepsilon_{i,t+h}$, with standard errors clustered at the firm level. The impulse response of total investment shows that the negative effect of a rise in legal risk persists for up to three years, highlighting a prolonged dampening of firm-level investment activity. This pattern suggests that legal risk generates uncertainty around the future business environment, leading firms to delay or scale back capital expenditures—a response consistent with models of irreversible investment under uncertainty (Bernanke 1983).

[21] The sample mean of tangible investment (CAPX + AQC) is 8.45, implying an economic effect of $-0.277/8.45 \approx -3.79\%$.

[22] The sample mean of R&D investment is 5.11, implying an economic effect of $-0.336/5.11 \approx -6.57\%$.

[23] The sample mean of total investment is 12.48, implying an economic effect of $-0.616/12.48 \approx -4.93\%$.

To further examine the economic consequences of legal risk, I assess whether it also affects the mode of financing. Following Malmendier et al. (2011), I estimate the following regression:

$$\text{Finance Mode}_{i,t} = \alpha + \beta_1 \, \text{FD}_{i,t} + \beta_2 \, \text{LRISK}_{i,t} + \beta_3 \, (\text{FD}_{i,t} \times \text{LRISK}_{i,t}) + \gamma \, \mathbf{X}_{i,t} + \theta_i + \eta_t + \varepsilon_{i,t} \quad (4)$$

In this specification, $\text{Finance Mode}_{i,t}$ denotes either net debt issuance (Debt) or net equity issuance (Equity) for firm i in year t, each scaled by lagged total assets. The variable $\text{FD}_{i,t}$ captures the financing deficit, defined following Frank and Goyal (2003), and measures the extent to which a firm's expenditures exceed its internal funds and thus necessitating external financing (see Table 8 for variable definitions). The coefficient of interest is β_3, which captures the interaction between the financing deficit and legal risk (FD $\times$ LRISK), and tests whether legal uncertainty moderates the firm's reliance on debt financing when internal resources are insufficient.

[INSERT TABLE 10 HERE]

Panel A in Table 10 presents strong evidence that legal risk systematically shapes corporate financing behavior, particularly when firms face a financing deficit (FD). To start with, across all model specifications, the coefficient on FD is consistently positive and highly significant, including in the baseline model in Column (1). This indicates that, in general, firms respond to funding shortfalls by raising debt, in line with traditional pecking-order theory: when internal funds are insufficient, firms first turn to debt markets before issuing equity. In contrast, LRISK is positive but statistically insignificant across all specifications. This suggests that legal risk on its own does not systematically increase or decrease a firm's propensity to issue debt when there are no pressing funding needs.

However, the interaction between FD and LRISK tells a different story. The coefficient on the interaction term LRISK × FD, β_3, is consistently negative and statistically significant at the 1% level across all models. This implies that when firms do experience a financing deficit, those with higher legal risk are significantly less likely to meet their funding needs via debt issuance. In other words, legal uncertainty introduces financing frictions that constrain debt capacity during periods of external funding demand. This effect is robust to the inclusion of controls for firm size, leverage, growth opportunities, and asset structure, as well as to firm, year, and industry fixed effects.

Panel B of Table 10 examines whether firms facing legal constraints on debt issuance turn instead to equity markets. The results show that the coefficient on the interaction term LRISK $\times$ FD is small and statistically insignificant across all specifications. This suggests that firms with elevated legal risk do not offset reduced debt access by substituting into equity financing. Rather than reallocating across financing instruments, these firms appear constrained in their overall access to external capital, reinforcing the view that legal risk operates as a broad financing friction.

To summarize, the firm behavior observed in this section is consistent with a precautionary financing motive. Firms exposed to greater legal risk may deliberately avoid taking on fixed obligations such as debt, given that legal liabilities (like lawsuits or regulatory penalties) are often large, opaque, and non-diversifiable. When such firms face a financing deficit, issuing debt can be especially costly: creditors may demand higher yields due to elevated tail risk, and rigid repayment structures reduce financial flexibility. These conditions increase the likelihood of debt overhang or distress, particularly if adverse legal events materialize.

6 Conclusion

This paper identifies legal risk as a distinct, priced, and economically meaningful dimension of firm-level uncertainty. Using textual analysis of earnings call transcripts, I construct a firm-level measure of legal risk (**LRISK**) that captures forward-looking concerns related to litigation, regulatory scrutiny, and legal exposure. This measure offers a scalable and dynamic proxy that reflects how both firms and investors perceive legal uncertainty in real time.

At the firm-level, I show that higher legal risk is associated with significantly lower future investment (particularly in intangible assets) and reduced reliance on debt financing during periods of funding need. These findings suggest that legal risk acts as a friction in both capital allocation and financing decisions, consistent with a precautionary motive. At the aggregate-level, legal risk is robustly priced in the cross-section of stock returns. A long-short portfolio that buys high-legal-risk firms and sells no-risk firms earns an annualized return of 4.7% over the full sample (2004–2024), rising to 7.1% post-2010—a period characterized by intensified regulatory enforcement. The return premium is not explained by standard factor models, ESG proxies, sentiment measures, or existing anomaly variables.

Further, the legal risk premium survives the factor zoo critique, and is closely tied to legal tail events, suggesting that investors demand compensation for downside risk uniquely associated with legal exposure. In comparative perspective, I show that legal risk is priced in common law countries—like the United States—but not in civil law systems such as France or Germany, where legal procedures are more codified and litigation risk less salient. This cross-country variation reinforces the idea that institutional foundations shape how markets perceive and price legal uncertainty.

Together, these findings establish legal risk as a first-order determinant of both firm behavior and expected returns. Understanding how legal systems contribute to firm-level risk, investor behavior, and capital market outcomes is central to the broader agenda of law and finance.

References

Ash, E., Morelli, M. and Vannoni, M.: 2025, More Laws, More Growth? Evidence from US States, *Journal of Political Economy* **126**(2), 850–901.

Bai, J. and Perron, P.: 2003, Computation and Analysis of Multiple Structural Change Models, *Journal of Applied Econometrics* **18**(1), 1–22.

Baker, S. R., Bloom, N. and Davis, S. J.: 2016, Measuring Economic Policy Uncertainty, *Quarterly Journal of Economics* **131**(4), 1593–1636.

Bennett, B., Ham, C., Milbourn, T. and Wang, Z.: 2025, Firm-Level Litigation Risk: Measurement and Effects, *Working Paper* .

Bereskin, F., Hsu, P.-H., Latham, W. and Wang, H.: 2023, So Sue Me! The cross section of stock returns related to patent infringement allegations, *Journal of Banking and Finance* **148**, 106740.

Bernanke, B. S.: 1983, Irreversibility, Uncertainty, and Cyclical Investment, *Quarterly Journal of Economics* **98**(1), 85–106.

Birru, J.: 2018, Day of the week and the cross-section of returns, *Journal of Financial Economics* **130**(1), 182–214.

Bolton, P. and Kacperczyk, M.: 2021, Do investors care about carbon risk?, *Journal of Financial Economics* **142**(2), 517–549.

Caldara, D. and Iacoviello, M.: 2022, Measuring Geopolitical Risk, *American Economic Review* **112**(4), 1194–1225.

Chen, A. Y. and Zimmermann, T.: 2022, Open source cross-sectional asset pricing, *Critical Finance Review* **27**(2), 207–264.

Chordia, T., Subrahmanyam, A. and Tong, Q.: 2014, Have capital market anomalies attenuated in the recent era of high liquidity and trading activity?, *Journal of Accounting and Economics* **58**(1), 41–58.

Cochrane, J. H.: 2011, Presidential Address: Discount Rates, *Journal of Finance* **66**(4), 1047–1108.

Cohen, L., Diether, K. and Malloy, C.: 2013, Legislating stock prices, *Journal of Financial Economics* **110**(3), 574–595.

Cohen, L., Gurun, U. G. and Kominers, S. D.: 2019, Patent Trolls: Evidence from Targeted Firms, *Management Science* **65**(12), 5461–5486.

Cohen, L., Malloy, C. and Nguyen, Q.: 2020, Lazy Prices, *Journal of Finance* **75**(3), 1371–1415.

Crawford, V. P. and Sobel, J.: 1982, Strategic Information Transmission, *Econometrica* **50**(6), 1431–1451.

Daniel, K., Hirshleifer, D. and Sun, L.: 2020, Short- and Long-Horizon Behavioral Factors, *Review of Financial Studies* **33**(4), 1673–1736.

deHaan, E., Shevlin, T. and Thornock, J.: 2015, Market (in)attention and the strategic scheduling and timing of earnings announcements, *Journal of Accounting and Economics* **60**(1), 36–55.

Djankov, S., La Porta, R., Lopez-De-Silanes, F. and Shleifer, A.: 2003, Courts, *Quarterly Journal of Economics* **118**(2), 453–517.

DuCharme, L. L., Malatesta, P. H. and Sefcik, S. E.: 2004, Earnings management, stock issues, and shareholder lawsuits, *Journal of Financial Economics* **71**(1), 27–49.

Eberly, J., Rebelo, S. and Vincent, N.: 2012, What explains the lagged-investment effect?, *Journal of Monetary Economics* **59**(4), 370–380.

Fahlenbrach, R., Prilmeier, R. and Stulz, R. M.: 2012, This Time Is the Same: Using Bank Performance in 1998 to Explain Bank Performance during the Recent Financial Crisis, *Journal of Finance* **67**(6), 2139–2185.

Fama, E. F. and French, K.: 2018, Choosing factors, *Journal of Financial Economics* **128**(2), 234–252.

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

Farrell, J. and Rabin, M.: 1996, Cheap Talk, *Journal of Economic Perspectives* **10**(3), 103–118.

Florackis, C., Louca, C., Michaely, R. and Weber, M.: 2023, Cybersecurity Risk, *Review of Financial Studies* **36**(1), 351–407.

Francis, J., Philbrick, D. and Schipper, K.: 1994, Shareholder Litigation and Corporate Disclosures, *Journal of Accounting Research* **32**(2), 137–164.

Frank, M. Z. and Goyal, V. K.: 2003, Testing the pecking order theory of capital structure, *Journal of Financial Economics* **67**(2), 217–248.

Gabaix, X.: 2011, The Granular Origins of Aggregate Fluctuations, *Econometrica* **79**(3), 733–772.

Gande, A. and Lewis, C. M.: 2009, Shareholder-Initiated Class Action Lawsuits: Shareholder Wealth Effects and Industry Spillovers, *Journal of Financial and Quantitative Analysis* **44**(4), 823–850.

Garcia, D., Hu, X. and Rohrer, M.: 2023, The colour of finance words, *Journal of Financial Economics* **147**(3), 525–549.

Glaeser, E. L. and Shleifer, A.: 2002, Legal Origins, *Quarterly Journal of Economics* **117**(4), 1193–1229.

Gormley, T. A. and Matsa, D. A.: 2011, Growing Out of Trouble? Corporate Responses to Liability Risk, *Review of Financial Studies* **24**(8), 2781–2821.

Hadlock, C. J. and Sonti, R.: 2012, Financial Strength and Product Market Competition: Evidence from Asbestos Litigation, *Journal of Financial and Quantitative Analysis* **47**(1), 179–211.

Harvey, C. R., Liu, Y. and Zhu, H.: 2016, ... and the Cross-Section of Expected Returns, *Review of Financial Studies* **29**(1), 5–68.

Harvey, C. R. and Siddique, A.: 2000, Conditional Skewness in Asset Pricing Tests, *Journal of Finance* **55**(3), 1263–1295.

Hassan, T. A., Hollander, S., van Lent, L. and Tahoun, A.: 2019, Firm-Level Political Risk: Measurement and Effects, *Quarterly Journal of Economics* **134**(4), 2135–2202.

Hossain, A., Rjiba, H. and Zhang, D.: 2023, Ex-ante litigation risk and firm-level climate-change exposure, *Journal of Economic Behavior and Organization* **214**, 731–746.

Huang, A. H., Wang, H. and Yang, Y.: 2023, FinBERT: A Large Language Model for Extracting Information from Financial Text, *Contemporary Accounting Research* **40**(2), 806–841.

Huang, D., Jiang, F., Tu, J. and Zhou, G.: 2015, Investor Sentiment Aligned: A Powerful Predictor of Stock Returns, *Review of Financial Studies* **28**(3), 791–837.

Huang, D., Li, J. and Wang, L.: 2021, Are disagreements agreeable? Evidence from information aggregation, *Journal of Financial Economics* **141**(1), 83–101.

Hughes, P. J. and Thakor, A. V.: 1992, Litigation Risk, Intermediation, and the Underpricing of Initial Public Offerings, *Review of Financial Studies* **5**(4), 709–742.

Jamilov, R., Rey, H. and Tahoun, A.: 2025, The Anatomy of Cyber Risk, *Working Paper* .

Jiang, H., Khanna, N., Yang, Q. and Zhou, J.: 2024, The Cyber Risk Premium, *Management Science* **70**(12), 8791–8817.

Jondeau, E., Zhang, Q. and Zhu, X.: 2019, Average skewness matters, *Journal of Financial Economics* **134**(1), 29–47.

Jordà, O.: 2005, Estimation and Inference of Impulse Responses by Local Projections, *American Economic Review* **95**(1), 161–182.

Jurado, K., Ludvigson, S. C. and Ng, S.: 2015, Measuring Uncertainty, *American Economic Review* **105**(3), 1177–1216.

Kalmenovitz, J.: 2023, Regulatory Intensity and Firm-Specific Exposure, *Review of Financial Studies* **36**(8), 3311–3347.

Kamiya, S., Kang, J.-K., Kim, J., Milidonis, A. and Stulz, R. M.: 2021, Risk management, firm reputation, and the impact of successful cyberattacks on target firms, *Journal of Financial Economics* **139**(3), 719–749.

Kamma, S., Weintrop, J. and Wier, P.: 1988, Investors' perceptions of the Delaware Supreme Court decision in Unocal V. Mesa, *Journal of Financial Economics* **20**, 419–430.

Kim, I. and Skinner, D. J.: 2012, Measuring securities litigation risk, *Journal of Accounting and Economics* **53**(1-2), 290–310.

Kruttli, M. S., Roth Tran, B. and Watugala, S. W.: 2025, Pricing Poseidon: Extreme Weather Uncertainty and Firm Return Dynamics, *Journal of Finance* **Forthcoming**.

La Porta, R., Lopez-de Silanes, F., Shleifer, A. and Vishny, R. W.: 1998, Law and Finance, *Journal of Political Economy* **106**(6), 1113–1155.

Lee, J., Schoenherr, D. and Starmans, J.: 2024, The Economics of Legal Uncertainty, *Working Paper* .

Loughran, T. and McDonald, B.: 2011, When Is a Liability Not a Liability? Textual Analysis, Dictionaries, and 10-Ks, *Journal of Finance* **66**(1), 35–65.

Loughran, T. and McDonald, B.: 2013, IPO first-day returns, offer price revisions, volatility, and form S-1 language, *Journal of Financial Economics* **109**(2), 307–326.

Lowry, M. and Shu, S.: 2002, Litigation risk and IPO underpricing, *Journal of Financial Economics* **65**(3), 309–335.

Malmendier, U., Tate, G. and Yan, J.: 2011, Overconfidence and Early-Life Experiences: The Effect of Managerial Traits on Corporate Financial Policies, *Jounal of Finance* **66**(5), 1687–1733.

Manela, A. and Moreira, A.: 2017, News implied volatility and disaster concerns, *Journal of Financial Economics* **123**(1), 137–162.

McDonald, R. and Siegel, D.: 1986, The Value of Waiting to Invest, *Quarterly Journal of Economics* **101**(4), 707–728.

Novy-Marx, R. and Velikov, M.: 2025, AI-Powered (Finance) Scholarship, *Working Paper* .

Pastor, L., Stambaugh, R. F. and Taylor, L. A.: 2022, Dissecting green returns, *Journal of Financial Economics* **146**(2), 403–424.

Pindyck, R. S.: 1988, Irreversible Investment, Capacity Choice, and the Value of the Firm, *American Economic Review* **78**(5), 969–985.

Sautner, Z., Van Lent, L., Vilkov, G. and Zhang, R.: 2023, Firm-Level Climate Change Exposure, *Journal of Finance* **78**(3), 1449–1498.

Savor, P. and Wilson, M.: 2016, Earnings Announcements and Systematic Risk, *Journal of Finance* **71**(1), 83–138.

Shu, S. Z.: 2000, Auditor Resignations: clientele effects and legal liability, *Journal of Accounting and Economics* **29**(2), 173–205.

Skinner, D. J.: 1994, Why Firms Voluntarily Disclose Bad News, *Journal of Accounting Research* **32**(1), 38–60.

Stambaugh, R. F.: 2017, Mispricing Factors, *Review of Financial Studies* **30**(4), 1270–1315.

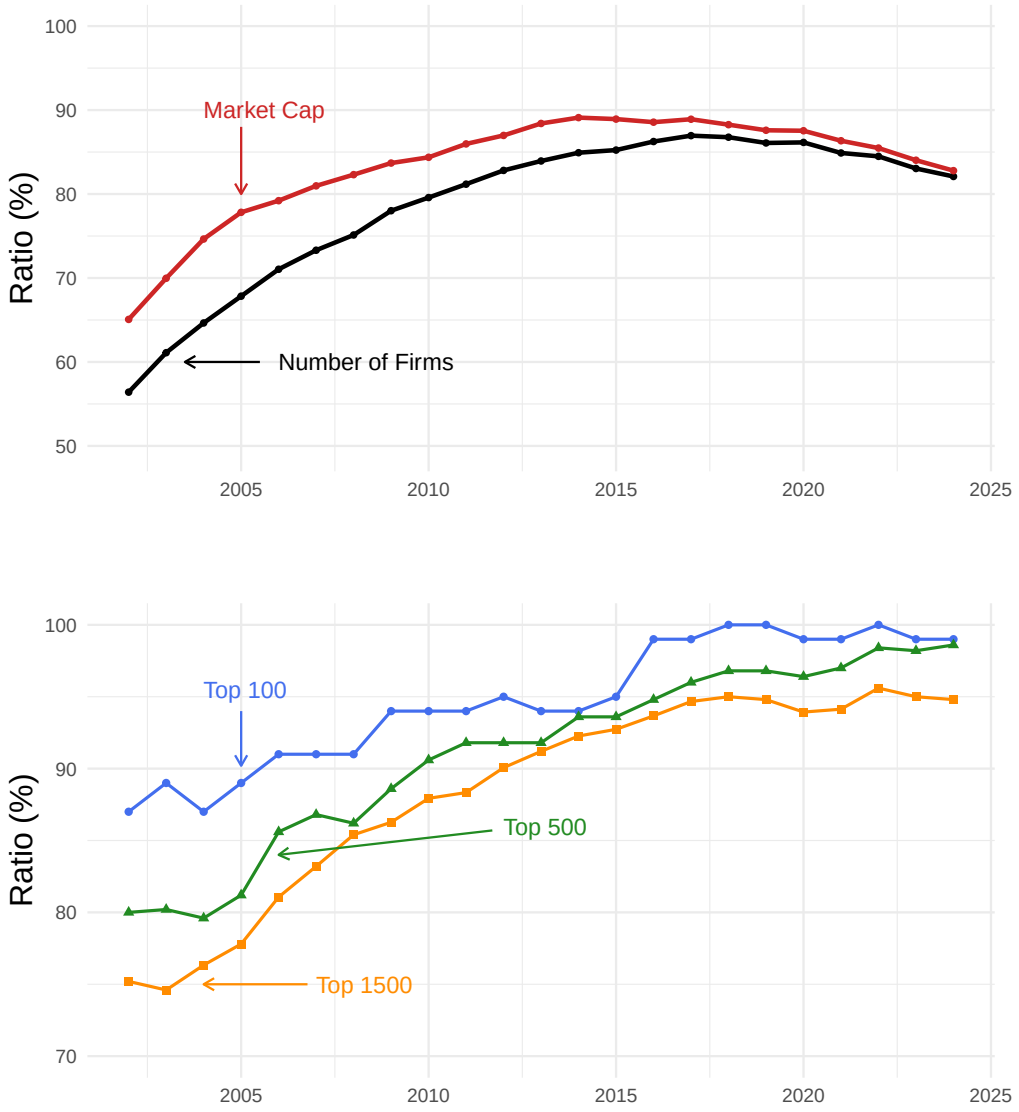
Stambaugh, R. F., Yu, J. and Yuan, Y.: 2012, The short of it: Investor sentiment and anomalies, *Journal of Financial Economics* **104**(2), 288–302.

Stroebel, J. and Wurgler, J.: 2021, What do you think about climate finance?, *Journal of Financial Economics* **142**(2), 487–498.

Whited, T. M. and Wu, G.: 2006, Financial Constraints Risk, *Review of Financial Studies* **19**(2), 531–559.

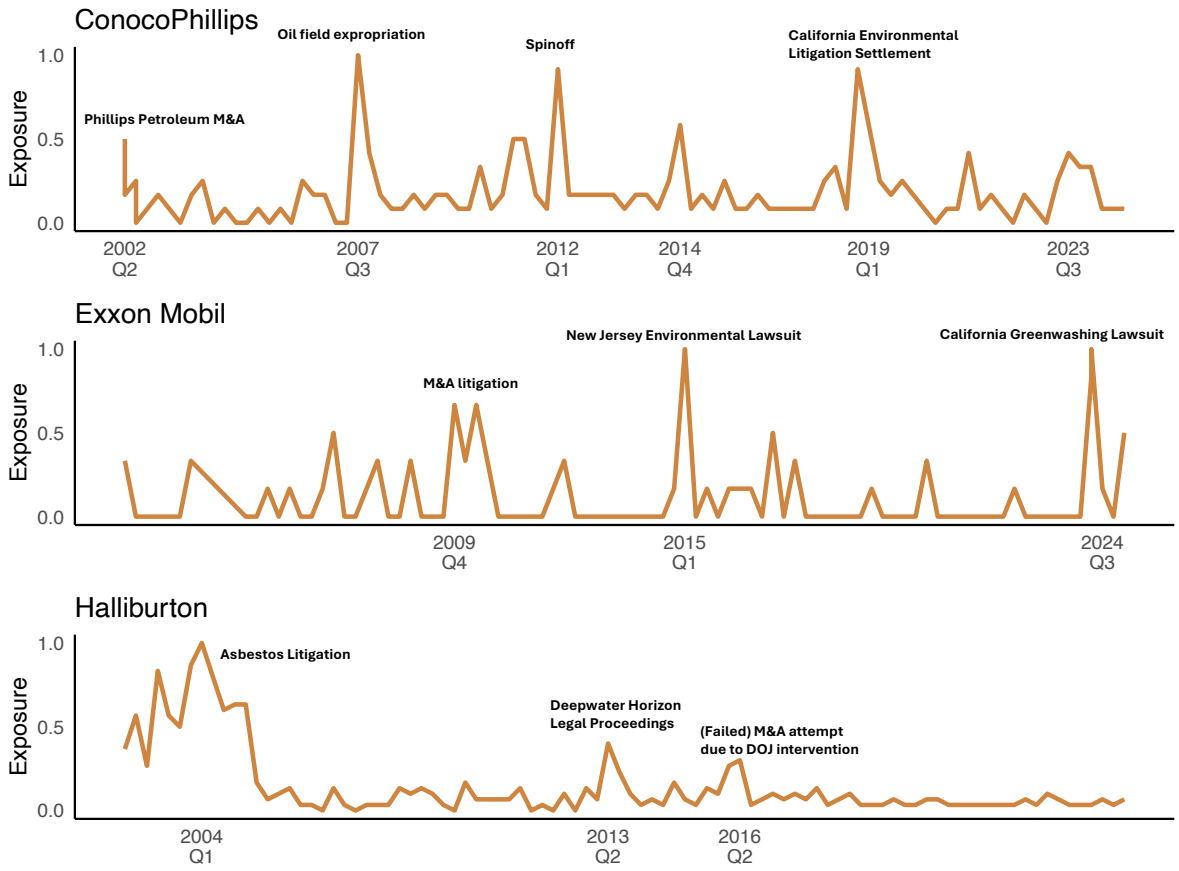
List of Figures

Figure 1: **Summary Statistic on U.S. Firms Publishing Earnings Calls Transcripts**



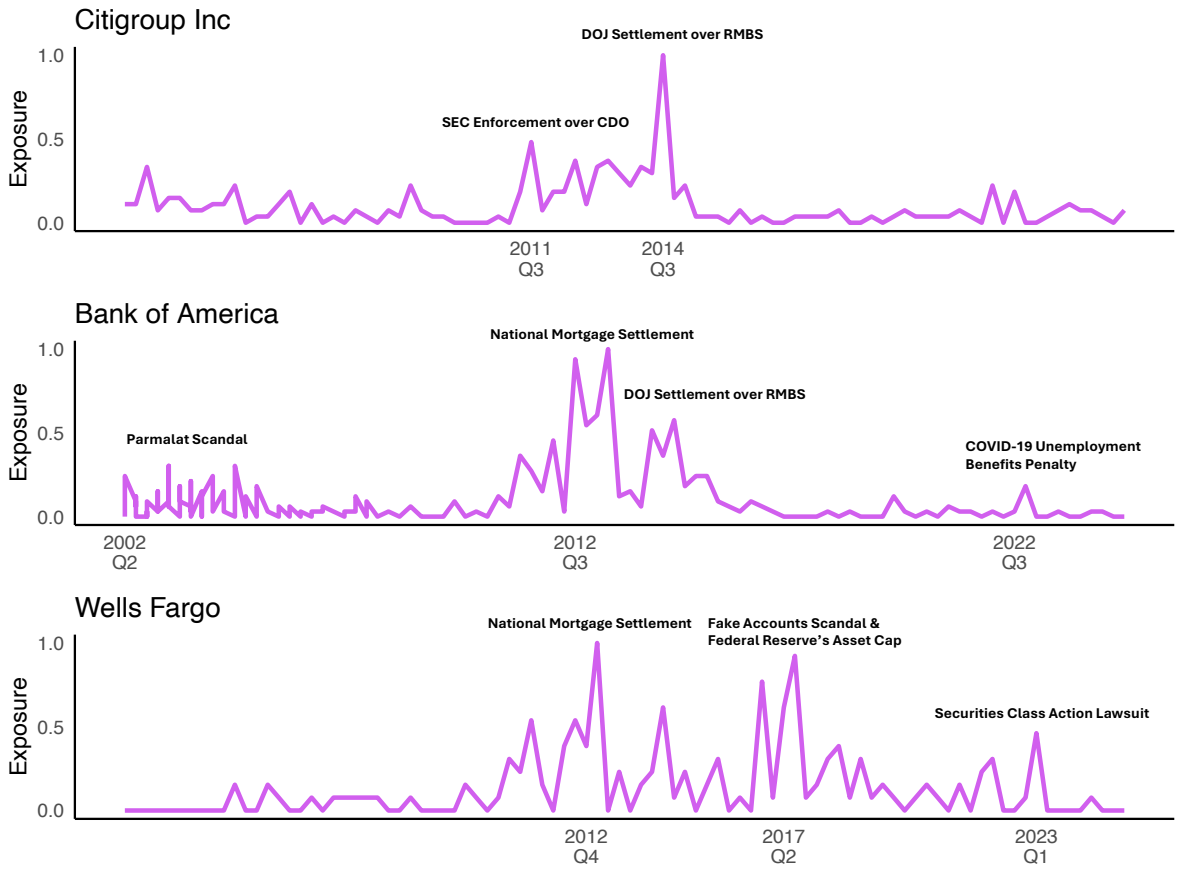
Note: The figures display the following: the proportion of firms publishing earnings call transcripts (black line in the top panel), the proportion of market capitalization of those firms that publish earnings calls (red line in the top panel), and the proportion of the top 100, top 500, and top 1500 firms by market capitalization that publish earnings call transcripts (bottom panel). The ratio is calculated by dividing the number of relevant firms by the total number of firms listed on the three major U.S. stock exchanges—NYSE, NASDAQ, and NYSE American (formerly AMEX)—with ordinary shares (CRSP share codes 10 or 11).

Figure 2: **Case Studies: Brown Firms**



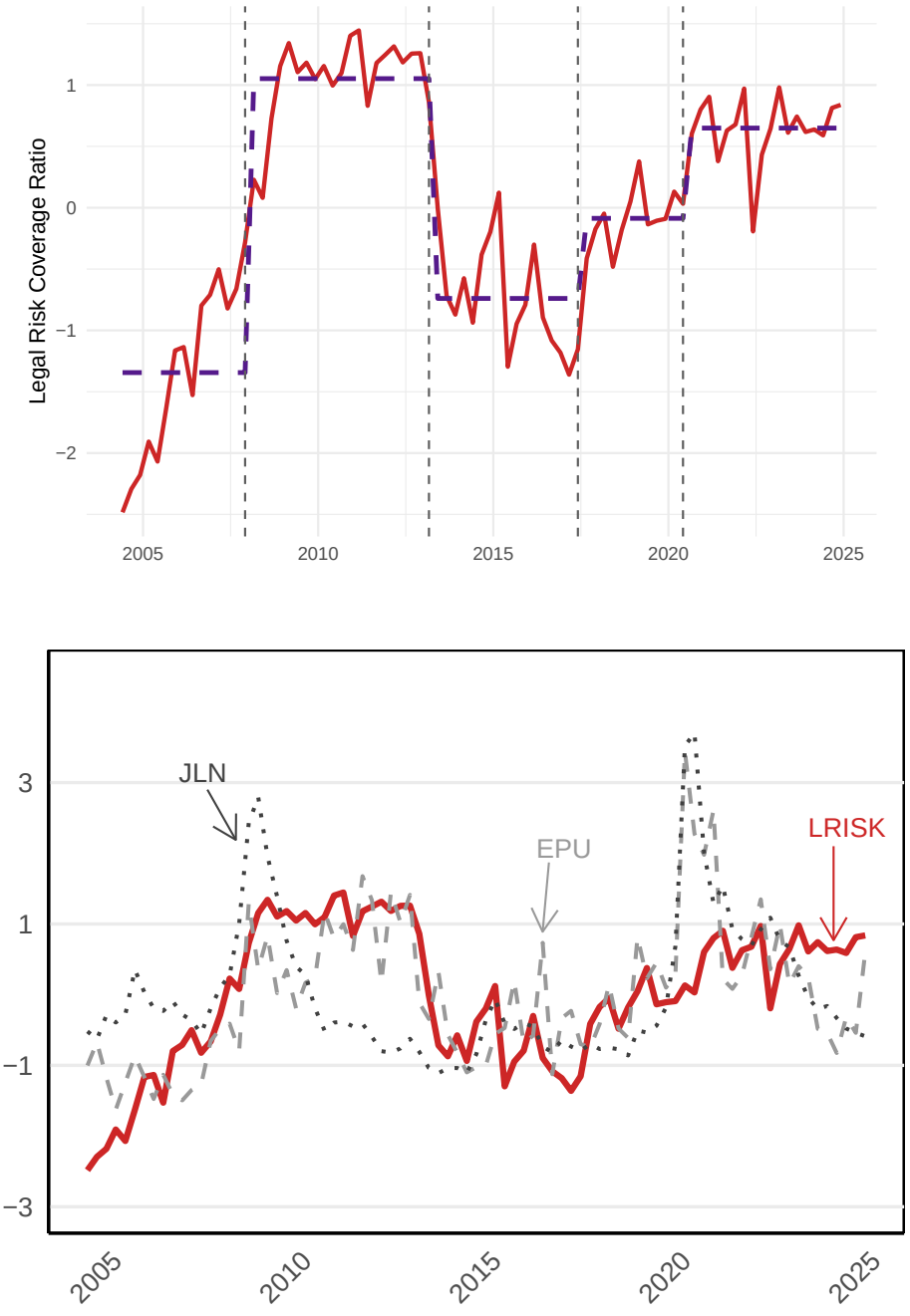
Note: This figure illustrates the firm-level legal risk exposure of three representative companies in the oil industry: ConocoPhillips, ExxonMobil, and Halliburton. All three are constituents of the S&P 500. The brown solid line in each figure represents the standardized firm-level legal risk exposure, as measured from earnings call transcripts.

Figure 3: **Case Studies: Bank Firms**



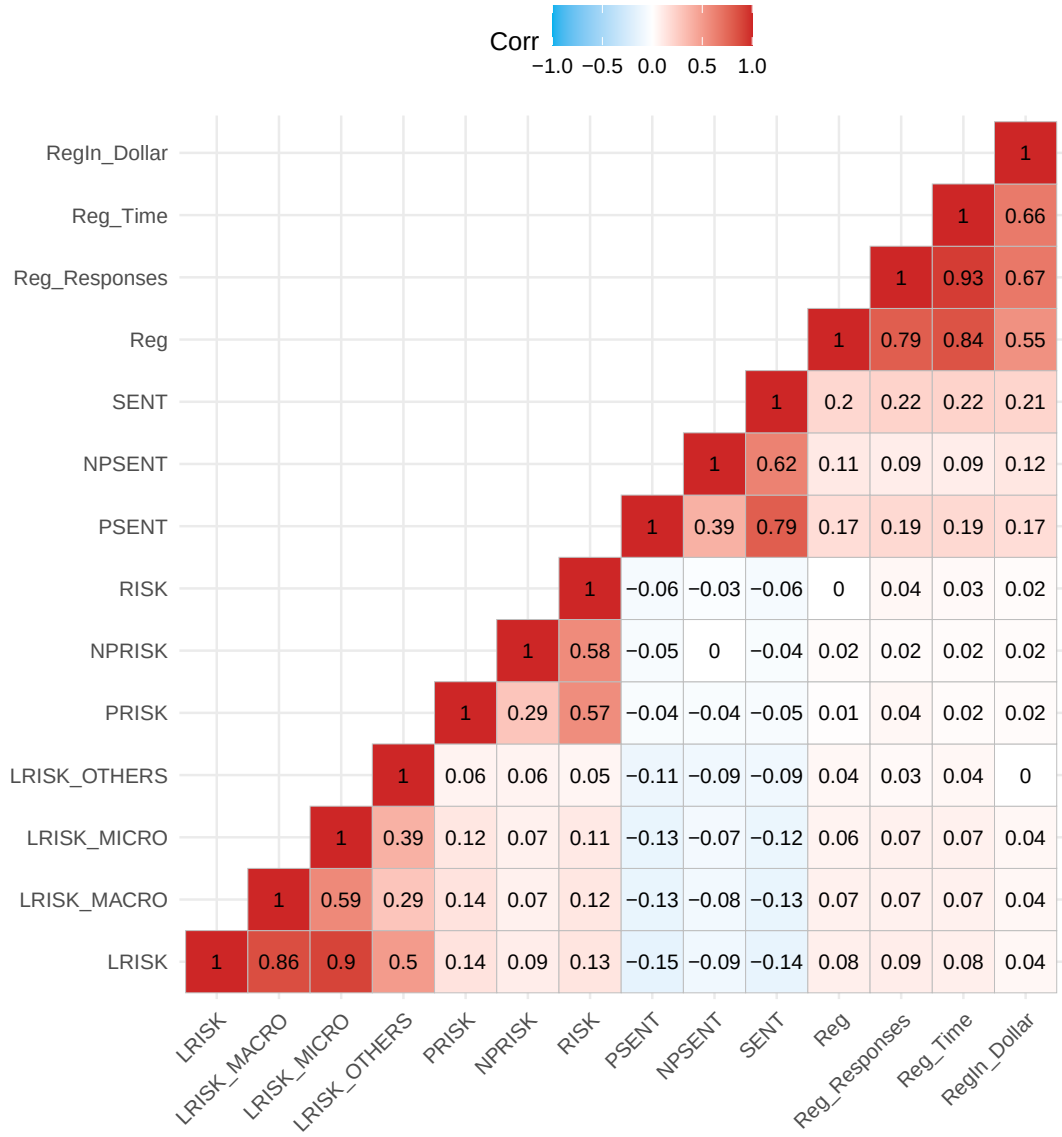
Note: This figure illustrates the firm-level legal risk exposure of three representative companies in the banking industry: Citigroup, Bank of America, and Well Fargo. All three are constituents of the S&P 500. The purple solid line in each figure represents the standardized firm-level legal risk exposure, as measured from earnings call transcripts.

Figure 4: **Aggregate Measure of Legal Risk**



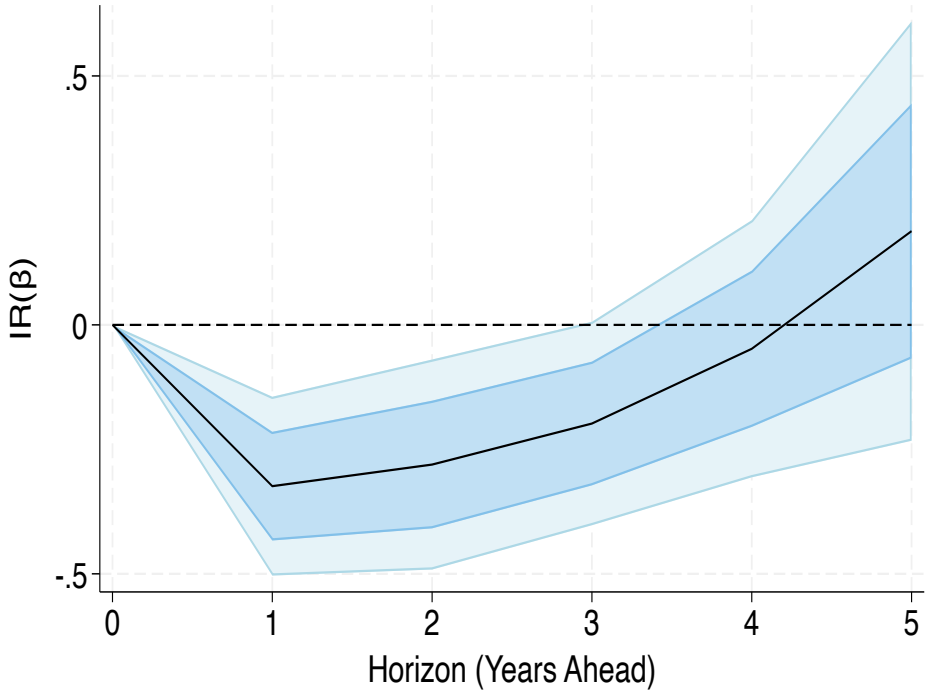
Note: The figures display an aggregate measure of legal risk, defined as the proportion of firms that reference legal risk in their earnings call transcripts over the past four quarters. Each quarter, firms are classified as either non-zero (mentioning legal risk) or zero (not mentioning it), and the coverage ratio is computed accordingly. To highlight deviations from the historical norm, the series is standardized using z-scores. In the top panel, the aggregate legal risk measure is segmented using the Bai and Perron (2003) multiple structural break method, with segment means plotted to illustrate shifts in the underlying trend. The bottom panel compares the standardized legal risk measure (LRISK) with well-known indicators of aggregate macroeconomic uncertainty: the Economic Policy Uncertainty (EPU) index and the Jurado, Ludvigson, and Ng (JLN) index.

Figure 5: **Correlation between LRISK and other Text-based Firm-level Exposure Measures**



Note: This figure shows the Pearson correlations between **LRISK** and its three components (**LRISK_COURT**, **LRISK_TRIAL**, and **LRISK_TERMS**) alongside other relevant text-based measures. Political and non-political risk and sentiment are captured by **PRISK**, **NPRISK**, **RISK**, **PSENT**, **NPSENT**, and **SENT**, as constructed by Hassan et al. (2019). The final three measures—**Reg_Responses**, **Reg_Time**, and **RegIn_Dollar**—track firms' regulatory burden, following Kalmenovitz (2023). Correlations are computed at the aggregate level; however, grouping by firm or year prior to calculation does not materially affect the main conclusions.

Figure 6: **Investment Impulse Response Function**



Note: This figure presents impulse response functions estimated using a panel data adoption of Jordà's (2005) local projection method. For each yearly forecast horizon $h = 1, ..., 5$, I estimate the panel regression of the form: $\text{INV}_{i,t+h} = \alpha + \sum_{j=0}^{2} \beta_j \text{ LRISK}_{i,t-j} + \gamma \mathbf{X}_{i,t} + \sum_{j=0}^{2} \delta_j \text{ INV}_{i,t-j} + \theta_i + \eta_t + \varepsilon_{i,t+h}$, with standard errors clustered at the firm level.The y-axis shows the percentage response of investment ("total investment" as defined in Table 8) to a one-standard-deviation increase in legal risk, and the x-axis reports the horizon in years. The projections include firm and year fixed effects, and standard errors are clustered at both the firm and year levels. Shaded areas represent 68% and 90% confidence intervals, respectively.

List of Tables

Table 1: **Examples of U.S. Firms mentioning Litigation-related terms in their Transcripts**

No	Companies	Date	SIC	Texts
1	NV Energy Inc	Feb 10, 2003	4911	The **lawsuit** filed *against* us in **Federal Bankruptcy Court** by Enron is an ongoing matter.
2	Metro International SA	Feb 13, 2007	2711	However, over the years, the **case law** has developed and at the end of 2005, it became apparent that certain deductions could be challenged by the Swedish tax agency.
3	Associated Materials LLC	Nov 17, 2011	3089	But in any event, they went to the same **lawyers** that were handling the first-**class action** suit, and in order to avoid an issue in commonality they set up a second suit against us.
4	Hong Kong Technology Venture Co Ltd	Nov 21, 2012	4813	They won't be able to grant me a license of 2015 because, as I mentioned earlier, in order to protect our investor interest, if we don't get our license by end of the year we will seriously consider to raise the **judicial review** against the government and get a **court** to push the government to issue the license in a reasonable timeframe.
5	Odyssey Marin Exploration Inc	March 16, 2015	8732	The **oral argument** for this final claim is scheduled for early May, that claim was dismissed at the **trial court** level and we expect the dismissal to be upheld.
6	ParkerVision Inc	November 14, 2016	3663	It's also important to note that we filed last month an **infringement case** against Apple in Munich, citing the same patent as in the LG case, and we've just recently learned that the **hearing** for that case has been set for May 4, 2017, less than 6 months from now.
7	Digital Ally Inc	Apr 1, 2019	3663	And quite frankly, we think even the **patent infringement** that we're seeing out there is becoming more and more brazen, not only with Axon and WatchGuard, who we have active **lawsuits** with, but we believe potentially even other competitors out there.
8	Johnson & Johnson	July 20, 2023	2834	The **bankruptcy judge** is expected to rule by August 2 on the motion to dismiss hearing that took place in the last week of June.

Note: This table shows the corporate earnings calls transcripts of some U.S. companies that mention legal- or litigation-related terms. The bold red text indicates legal-related terms

Table 2: **Word Choices to Measure Litigation Risk**

Categories	Key terms in this category
Judicial Entities and Key Participants (**LRISK_COURT**)	**Judicial Entities**: Bankruptcy Court, Department of Justice, Federal Circuit, Supreme Court, Trial Court, ... **Key Participants**: Administrative Law Judge, Judges, Jury, Plaintiffs, Prosecutors, ...
Legal Processes and Disputes (**LRISK_TRIAL**)	**Legal Processes**: Appeals, Arbitration, Bankruptcy Proceeding, Class Action, Oral Argument, Trial Date, ... **Legal Disputes**: Antitrust Case, Civil Litigation, Patent Infringement, ...
Legal Terms and Concepts (**LRISK_TERMS**)	Adverse Ruling, Case Law, Legal Costs, Sub Judice, ...

Note: This table shows the list of legal terms contained in the legal dictionary $\mathbb{L}$. Broadly put, there are three distinct categories, terms related to (1) Judicial Entities and Key Participants; (2) Legal Processes and Disputes; and (3) Legal Terms and Concepts. Refer to the online appendix for the full list of keywords and word choices.

Table 3: **Logistic Regression Predicting Class Action Lawsuits**

	(1) Lawsuit in $t+1$	(2) Lawsuit within $t+1$ to $t+2$	(3) Lawsuit within $t+1$ to $t+4$
LRISK	**0.104*** (5.68)**	**0.085*** (4.76)**	**0.075*** (3.92)**
FPS	0.416*** (5.68)	0.412*** (5.75)	0.415*** (5.68)
Past Lawsuit	0.189*** (2.53)	0.253*** (3.48)	0.300*** (4.02)
Return	-2.814*** (-17.43)	-1.848*** (-16.07)	-1.222*** (-13.67)
Return Vol	0.007*** (7.64)	0.006*** (7.75)	0.005*** (7.05)
Constant?	Yes	Yes	Yes
Controls?	Yes	Yes	Yes
Firm Fixed Effects?	Yes	Yes	Yes
Observations	174,563	174,563	174,563
No. of Lawsuit Filings	1,437	1,437	1,437
AUC	0.746	0.707	0.685
Pseudo R^2	0.078	0.054	0.044

Note: This table reports logistic regression results where the dependent variable is an indicator for whether a class action lawsuit is filed in quarter $t+1$ (column 1), within quarters $t+1$ to $t+2$ (column 2), or within quarters $t+1$ to $t+4$ (column 3), using a quarterly sample from 2002 to 2024. LRISK is the standardized measure of firm-level legal risk. FPS is a dummy for litigation-prone industries. Past Lawsuit is an indicator for whether the firm has previously been sued. Return is the cumulative stock return over the past quarter, and Return Vol is the return volatility over the same period. Robust z-statistics clustered at the firm level are reported in parentheses. Firm-level controls, which are suppressed for brevity, include firm size, Tobin's Q, leverage, sales growth, asset tangibility, and cash flow, all measured in quarter t. ***/**/* denote the statistical significance at 1%/5%/10% level.

Table 4: **Legal Risk Portfolio Results**

Panel A: Value-weighted future excess returns (%)

	L1 No legal risk	L2 Middle legal risk	L3 High legal risk	L3 - L1 Long - Short
Excess return	0.638	0.912	1.035	**0.397*** **
	(2.02)	(3.23)	(3.37)	**(3.66)**
CAPM alpha	-0.207	0.094	0.232	**0.439*** **
	(-2.17)	(1.50)	(3.79)	**(4.15)**
FF3 alpha	-0.193	0.084	0.205	**0.398*** **
	(-1.92)	(1.39)	(4.56)	**(3.87)**
FF3 + Mom alpha	-0.157	0.095	0.215	**0.372*** **
	(-1.67)	(1.62)	(4.89)	**(3.72)**
FF3 + Liquidity alpha	-0.198	0.100	0.195	**0.392*** **
	(-1.96)	(1.65)	(4.46)	**(3.81)**
FF5 alpha	-0.193	0.028	0.178	**0.372*** **
	(-1.88)	(0.536)	(3.79)	**(3.43)**
FF5 + Mom alpha	-0.162	0.037	0.189	**0.351*** **
	(-1.65)	(0.739)	(4.01)	**(3.31)**
q-factor alpha	-0.124	0.063	0.204	**0.328*** **
	(-1.18)	(1.01)	(3.81)	**(2.96)**

Panel B: Firm characteristics

	L1 No legal risk	L2 Middle legal risk	L3 High legal risk
Num. of Firms	701	895	895
Market value (log)	5.5	7.16	6.69
Book-to-market ratio	0.72	0.60	0.64
Gross Profit	0.23	0.32	0.31
Investment	0.33	0.26	0.29
Momentum	0.13	0.12	0.11
Asset Tangibility	0.20	0.22	0.19
Long-term Debt / Total Liabilities	0.27	0.33	0.30

Panel C: Transition Matrix (% of firms moving from one portfolio to another)

	(to) L1	(to) L2	(to) L3	Total
(from) L1	62.0	26.6	11.3	100
(from) L2	17.8	53.8	28.4	100
(from) L3	7.5	26.3	66.2	100

Note: Panel A presents the average future excess returns (2004:07 - 2024:06) alongside the alpha results derived from portfolios constructed based on ex-ante legal risk measures. Firms in L1 (No legal risk) are those with no exposure to legal risk. Firms with non-zero exposure to legal risk are split into two groups, placed in L2 or L3 according to their level of exposure. Utility firms (SIC codes 4900–4999) and financial firms (SIC codes 6000 - 6411, 6500 - 6553, and 6700 - 6799) are excluded due to the highly restrictive nature of government regulation. The numbers in parenthesis are Newey-West t-statistics with 6 lags. Panel B records the firm characteristics of companies analyzed in the portfolio analysis. Panel C reports the one-year transition probabilities between portfolio groups based on legal risk. Each row shows the percentage of firms originally in portfolio L1, L2, or L3 that migrate to each group in the subsequent portfolio formation period. ***/**/* denote the statistical significance at 1%/5%/10% level.

Table 5: **Asset Pricing Effects of Macro-Level Legal Risk**

Panel A: Cross-sorting of Macro vs. Micro Legal Risk (%)

	Micro = 1 (No micro-level legal risk)	Micro = 2 (Medium micro-level legal risk)	Micro = 3 (High micro-level legal risk)
Macro = 1 (No macro-level legal risk)	29.4	15.4	11.4
Macro = 2 (Medium macro-level legal risk)	5.3	8.5	8.2
Macro = 3 (High macro-level legal risk)	5.1	6.3	10.5

Panel B: Macro-level Legal Risk (LAW_COURT) (%)

	L1 Zero legal risk	L2 Non-zero macro-level legal risk	L2 - L1 Long - Short
Excess return	0.638	0.962	**0.323*** **
	(2.01)	(3.24)	**(3.18)**
FF3 + mom alpha	-0.157	0.146	**0.303*** **
	(-1.67)	(3.12)	**(3.13)**
FF5 + mom alpha	-0.162	0.096	**0.257*** **
	(-1.65)	(2.11)	**(2.48)**

Note: Panel A reports the joint distribution of firm–date observations by macro- and micro-level legal risk sorts (in terciles). Each cell reports the percentage of total observations, with all entries summing to 100%. Diagonal cells indicate agreement across the two measures; off-diagonals capture disagreement. Panels B reports average excess returns and alphas for the macro-level legal risk factor (**LAW_COURT**). The low-risk group (L1) is held constant and consists of firms entirely free of legal content, ensuring that pricing differences reflect only variation in the high-risk group. The sample covers July 2004 to June 2024. T-statistics are Newey–West adjusted with 6 lags. Utility and financial firms are excluded due to sector-specific regulatory structures. ***/**/* indicate statistical significance at the 1%/5%/10% level, respectively.

Table 6: **Post-2010 Results**

Post-2010 Result (%)

	P1 No legal risk	P2 Middle legal risk	P3 High legal risk	P3 - P1 Long - Short
Excess return	0.841	1.173	1.434	**0.593*** **
	(2.91)	(4.47)	(5.12)	**(4.91)**
CAPM alpha	-0.371	-0.013	0.282	**0.653*** **
	(-3.64)	(-0.21)	(3.93)	**(5.58)**
FF3 alpha	-0.370	-0.027	0.0.236	**0.606*** **
	(-3.53)	(-0.432)	(4.91)	**(5.41)**
FF3 + Mom alpha	-0.319	-0.001	0.244	**0.563*** **
	(-3.28)	(-0.023)	(4.96)	**(5.18)**
FF3 + Liquidity alpha	-0.374	-0.015	0.212	**0.587*** **
	(-3.42)	(-0.24)	(4.57)	**(5.03)**
FF5 alpha	-0.339	-0.057	0.212	**0.551*** **
	(-3.26)	(-1.00)	(4.16)	**(4.93)**
FF5 + Mom alpha	-0.292	-0.034	0.223	**0.515*** **
	(-2.99)	(-0.60)	(4.30)	**(4.71)**
q-factor alpha	-0.280	-0.046	0.230	**0.509*** **
	(-2.84)	(-0.68)	(3.99)	**(4.43)**

Note: This table presents the average future excess returns alongside the alpha results derived from portfolios constructed based on ex-ante legal risk measures, spanning post-2010 sample (2010:07 - 2024:06). Firms in P1 (No litigation risk) are those with no exposure to legal risk. Firms with non-zero exposure to legal risk are split into two groups, placed in P2 or P3 according to their level of exposure. The portfolio return data spans from July 2003 to December 2023. Utility firms (SIC codes 4900–4999) and financial firms (SIC codes 6000 - 6411, 6500 - 6553, and 6700 - 6799) are excluded due to the highly restrictive nature of government regulation. The numbers in parenthesis are Newey-West t-statistics with 6 lags. ***/**/* denote the statistical significance at 1%/5%/10% level.

Table 7: **Comparing Legal Risk and Cyber Risk**

Panel A: Correlation

	LRISK (LAW)	**Cyber Risk**	**Cyber Risk Law**
LRISK (LAW)	1.00		
Cyber RISK	-0.05	1.00	
Cyber Risk Law	0.42	-0.01	1.00

Panel B: Risk Factor Performance

	LRISK (LAW)	**Cyber Risk**	**Cyber Risk Law**
Excess return	**0.397*** **	**0.360** **	0.142
(% per month)	**(3.66)**	**(2.05)**	(1.19)
FF5 + Mom alpha	**0.351*** **	**0.275*** **	0.070
(% per month)	**(3.31)**	**(2.25)**	(0.62)

Panel C: Factor Spanning Test

	(1)	(2)	(3)
→ Dependent variable:	**LRISK (LAW)**	**Cyber Risk**	**Cyber Risk Law**
↓ Independent variable:			
Intercept	**0.277*** **	0.164	-0.093
	(2.85)	(1.49)	(-0.96)
LRISK (LAW)		**0.244** **	**0.256** **
		(2.24)	**(2.16)**
Cyber Risk	**0.203** **		**0.361*** **
	(2.12)		**(5.24)**
Cyber Risk Law	**0.281** **	**0.475*** **	
	(2.08)	**(4.59)**	
MKT	0.004	0.005	0.007
	(0.14)	(0.21)	(0.29)
SMB	**-0.120** **	-0.029	0.061
	(-2.56)	(-0.53)	(1.37)
HML	0.06	**-0.353*** **	0.065
	(1.10)	**(-5.68)**	(1.32)
MOM	**0.05***	-0.002	0.041
	(1.70)	(-0.04)	(1.44)

Note: This table compares Legal Risk (**LRISK** or **LAW** factor) constructed in this paper with **Cyber Risk** and its component measure, **Cyber Risk Law**. All three measures are derived from earnings call transcripts, covering the sample period from 2004:07 to 2024:06. Panel A reports the Pearson correlation at the firm level. Panel B presents the performance of a long-short portfolio strategy over the sample period. Panel C shows the factor spanning test, where the dependent variable is one of the three measures, and the benchmark model is the Fama-French five-factor model augmented with momentum, as well as the two remaining measures not used as the dependent variable. In Panels B and C, numbers in parentheses represent Newey-West t-statistics with six lags. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Table 8: **Variable Definitions and Descriptive Statistics**

Panel A: Variable Definitions

Variable	Definition
Legal Risk	
LRISK	Standardized firm-level legal risk exposure, constructed as the annual average of quarterly textual exposure measures to legal risk from earnings calls, and standardized across all firm-year observations
Control Variables	
Size	Natural logarithm of total assets: $\log(\texttt{at})$
Tobin's Q	Book value of equity plus long-term debt over total assets: $(\texttt{at} + \texttt{mve} - \texttt{ceq} - \texttt{txdb})/\texttt{at}$ where $\texttt{mve} = \texttt{csho} \times \texttt{prcc_f}$
Leverage	Ratio of total debt (long-term + current) to total assets: $(\texttt{dltt} + \texttt{dlc})/\texttt{at}$
Sales Growth	Annual growth in sales: $(\texttt{sale} - \texttt{L1.sale})/\texttt{sale}$
Tangibility	Asset tangibility, defined as the ratio of net property, plant, and equipment to total assets: $\texttt{ppent}/\texttt{at}$
CF	Cash Flow: $(\texttt{ib} + \texttt{dp})$ / $\texttt{at}$
WW	Whited-Wu index to measure firms' financial constraints
FD	Financing Deficit (%): Following Frank and Goyal (2003), this variable is defined as the sum of cash dividends (dvc), capital expenditures (capx), acquisitions (aqc), investment in other assets (ivch), and the change in working capital (act−lct), minus operating cash flow (oancf); scaled by lagged total assets (at). All components are set to zero where missing. The variable proxies for a firm's reliance on external financing
Net Debt Issues	The net amount of long-term debt issued by the firm, defined as debt issuance (dltis) minus debt repayments (dltr), scaled by lagged total assets. Missing values are set to zero to retain sample size. This variable captures a firm's reliance on long-term debt financing relative to its size
Net Equity Issues	The net amount of equity issued by the firm, defined as sales of common stock (sstk) minus stock repurchases (prstkc), scaled by lagged total assets. Missing values are set to zero to retain sample size. This variable captures a firm's reliance on external equity financing relative to its size
Investment variables	
CAPX + AQC	Tangible capital investment (%): $100 \times [\texttt{capx} + \texttt{aqc}]$ / lagged total assets. Missing values in capx and aqc are imputed with zeros
R&D	Intangible capital investment using R&D intensity (%): $100 \times$ xrd / lagged total assets. Missing values in xrd are imputed with zeros
Total Investment	Total investment (%): $100 \times [\texttt{capx} + \texttt{aqc} + \texttt{xrd}]$ / lagged total assets. Missing values are imputed with zeros for all components

Panel B: Descriptive Statistics

Variable	N	Mean	SD	Min	P25	P50	P75	Max
LRISK	50,612	0.00	1.00	-0.59	-0.46	-0.26	0.05	30.28
SIZE	60,149	6.80	2.21	1.05	5.24	6.77	8.27	14.34
Tobin's Q	59,896	0.62	0.25	-1.89	0.53	0.69	0.79	0.98
Leverage	59,899	0.24	0.25	0.00	0.04	0.19	0.36	2.63
Sales Growth	53,427	-0.05	1.70	-42.76	-0.03	0.06	0.16	1.00
Tangibility	59,728	0.22	0.23	0.00	0.04	0.13	0.31	0.95
CF	58,629	-0.02	0.33	-4.12	-0.00	0.06	0.11	0.60
WW	51.980	-0.02	0.33	-4.12	-0.43	-0.34	-0.25	0.60
FD	60,149	0.15	0.51	-0.78	-0.03	0.02	0.14	7.45
Net Debt Issues	60,149	0.03	0.19	-6.74	-0.01	0.00	0.01	11.40
CAPX + AQC	59,935	7.24	14.48	-14.11	0.88	3.22	7.84	204.84
R&D	59,935	6.25	15.26	0.00	0.00	0.00	5.89	194.45
Total	59,935	13.55	21.46	-10.64	2.32	7.22	16.34	278.00

Note: Panel A describes the variables used in this paper. Panel B provides summary statistics. LRISK is standardized across all firm-year observations. All other variables are winsorized at the 0.1% and 99.9% levels.

Table 9: **Legal Risk and Firm Investment**

	(1) CAPX + AQC	(2) R&D	(3) Total Inv
Panel A: Benchmark Specification			
LRISK	**-0.277*****	**-0.336*****	**-0.616*****
	(-2.99)	**(-3.65)**	**(-4.53)**
Controls?	Yes	Yes	Yes
Firm Fixed Effects?	Yes	Yes	Yes
Year Fixed Effects?	Yes	Yes	Yes
Observations	43,741	43,741	43,741
R^2	0.257	0.769	0.429
Panel B: Controlling for Lagged Investment			
LRISK	**-0.230****	**-0.175****	**-0.465*****
	(-2.25)	**(-2.43)**	**(-3.43)**
Controls?	Yes	Yes	Yes
Firm Fixed Effects?	Yes	Yes	Yes
Year Fixed Effects?	Yes	Yes	Yes
Observations	36,389	36,389	36,389
R^2	0.257	0.803	0.430
Panel C: Industry Fixed Effects			
LRISK	**-0.480*****	**-0.366*****	**-0.853*****
	(-6.53)	**(-4.18)**	**(-7.29)**
Controls?	Yes	Yes	Yes
Industry Fixed Effects?	Yes	Yes	Yes
Year Fixed Effects?	Yes	Yes	Yes
Observations	41,681	41,681	41,681
R^2	0.101	0.546	0.247
Panel D: Controlling for Ohlson's O-score			
LRISK	**-0.254*****	**-0.352*****	**-0.611*****
	(-2.48)	**(-3.56)**	**(-4.14)**
Controls?	Yes	Yes	Yes
Firm Fixed Effects?	Yes	Yes	Yes
Year Fixed Effects?	Yes	Yes	No
Observations	39,072	39,072	39,072
R^2	0.240	0.765	0.407

Note: This table presents panel regressions of investment on lagged firm-level legal risk (LRISK), using data from 2002 to 2024. The dependent variables are: (i) capital expenditures and acquisitions (CAPX + AQC), (ii) R&D expenditures, and (iii) total investment (CAPX + AQC + R&D), all scaled by beginning-of-year total assets. LRISK is a text-based measure of firm-level legal risk. Controls (suppressed to save space) include firm size, Tobin's Q, leverage, sales growth, asset tangibility, cash flow, and the Whited-Wu index (a proxy for financial constraints). Standard errors are clustered at the firm level. Panel (A) shows the baseline results. Panel (B) includes up to three lags of the dependent variable to account for the strong persistence in investment dynamics, as emphasized by Eberly, Rebelo, and Vincent (2012). Panel (C) replaces firm fixed effects with two-digit SIC industry fixed effects. Panel (D) substitutes the WW index with Ohlson's O-score. t-statistics are in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Table 10: **Legal Risk and External Financing Decisions**

	Panel A: Net Debt Issues			**Panel B: Net Equity Issues**		
	(1)	(2)	(3)	(4)	(5)	(6)
FD	0.152***	0.156***	0.176***	0.480***	0.473***	0.462***
	(15.82)	(15.43)	(15.09)	(19.83)	(19.20)	(17.78)
LRISK	0.001	0.002	0.002	0.002	0.002	0.001
	(0.97)	(1.40)	(1.33)	(0.69)	(0.66)	(0.70)
LRISK × FD	**-0.016*** **	**-0.017*** **	**-0.019*** **	-0.013	-0.012	-0.013
	(-3.49)	**(-3.88)**	**(-4.08)**	(-0.80)	(-0.75)	(-0.75)
SIZE			0.027***			-0.023***
			(10.70)			(-5.91)
Tobin's Q			0.035***			-0.121***
			(3.01)			(-7.10)
Leverage			0.223***			-0.150***
			(19.18)			(-11.55)
Sales Growth			0.001			0.001
			(0.82)			(0.72)
Tangibility			-0.023			0.031
			(-1.26)			(1.13)
Intangibility			0.088***			0.052***
			(6.79)			(2.91)
Z-score			-0.002***			0.004***
			(-3.73)			(4.65)
Firm Fixed Effects?	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects?	No	Yes	Yes	No	Yes	Yes
Industry Fixed Effects?	No	Yes	Yes	No	Yes	Yes
Observations	50,308	47,665	40,728	47,083	46,446	39,727
R^2	0.280	0.293	0.361	0.702	0.700	0.698

Note: This table combines regression results for debt and equity issuance. FD denotes the financing deficit. LRISK is a standardized firm-level legal risk measure. All dependent variables are scaled by lagged assets. Fixed effects vary by specification. T-statistics (in parentheses) are clustered at the firm level. *, **, and *** denote significance at the 10%, 5%, and 1% levels.

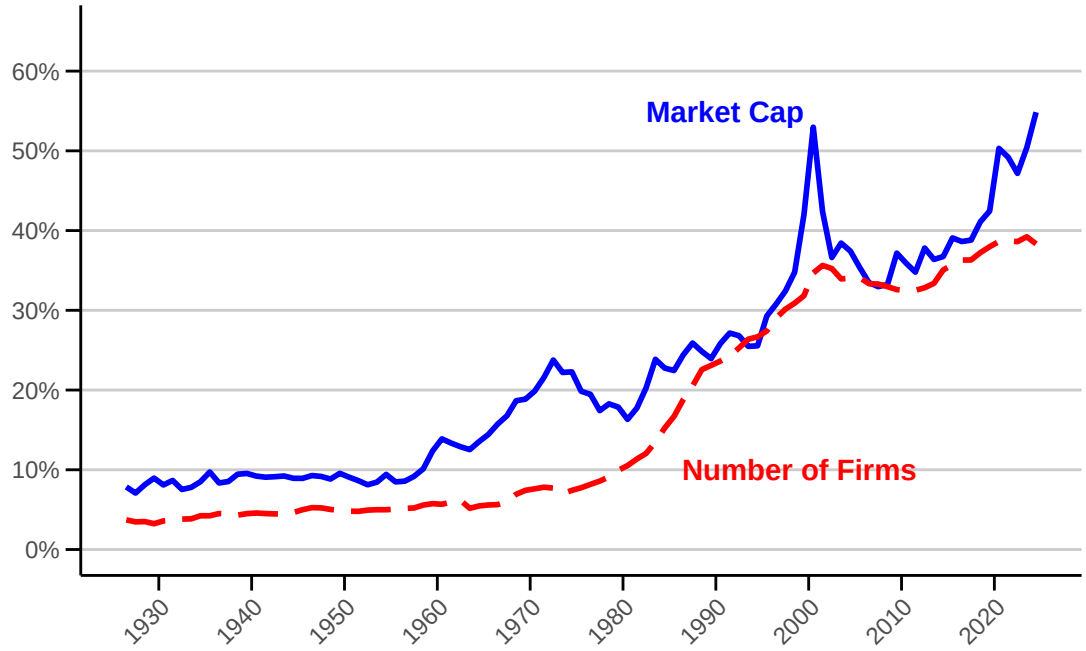
Legal Risk

Online Appendix

- **Section A**: Growth of Legal Risk through Litigation-Prone Industries
- **Section B**: Earnings Call Signals of Legal Exposure
- **Section C**: Additional Asset Pricing Results and Data Mining Concerns
- **Section D**: Aggregate LRISK vs. Uncertainty Measures
- **Section E**: Firm-level Evidence: Case studies using SolarWinds
- **Section F**: Legal Shifts: Implications for Firms and Markets
- **Section G**: Modeling Legal Risk in a Real Options Framework

A Growth of Legal Risk through Litigation-Prone Industries

Figure A.1: **Growth of Litigation-Prone Industries**



Note: In this figure, the red dotted line plots the ratio of firms in high-litigation industries to all publicly traded firms on the NYSE, NASDAQ, and NYSE American (formerly AMEX) from 1926 to 2024. The blue solid line plots the corresponding ratio of market capitalization. The industries prone to litigation include pharmaceuticals and biotechnology (SIC 2833 - 2836), computer and office equipment (SIC 3570 - 3577), software and IT services (SIC 7370 - 7374), and electronic and electrical equipment (SIC 3600 - 3674). These industries are frequently involved in litigation due to intellectual property disputes, regulatory challenges, and product liability risks.

The growth of litigation-prone industries over the past two decades highlights an important structural source of rising legal risk in the U.S. economy. Figure A.1 illustrates this trend by tracking the number of publicly listed firms operating in high-litigation industries—such as pharmaceuticals and biotechnology (SIC 2833 - 2836), computer and office equipment (SIC 3570 - 3577), software and IT services (SIC 7370 - 7374), and electronic and electrical equipment (SIC 3600 - 3674)—as well as their aggregate market capitalization relative to the overall market. These sectors are frequently exposed to legal disputes in-

volving intellectual property, regulatory compliance, antitrust, and product liability, making them natural incubators for firm-level legal risk. While the absolute number of public firms has declined since the late 1990s, the share of firms operating in these litigation-intensive sectors has grown steadily, as has their combined market value.

This structural reallocation of capital toward more legally exposed industries amplifies the relevance of legal risk for asset pricing and corporate decision-making. As the figure shows, firms in these sectors now account for a substantially larger fraction of public market capitalization than they did in prior decades, especially post-2000. This compositional shift implies that aggregate market exposure to legal uncertainty is not merely cyclical but reflects deeper changes in the industrial landscape. Legal risk, in this sense, is increasingly embedded in the technological and regulatory complexity of the modern economy.

These patterns, along with the evidence in Table 3 showing that firms in litigation-prone industries are significantly more likely to face class action lawsuits, reinforce the core motivation of this paper: legal risk is not an episodic concern confined to a narrow set of firms, but a pervasive and growing feature of the public firm landscape. Understanding how legal risk shapes valuation, investment, and financing decisions is therefore essential to analyzing firm behavior and return premia in modern capital markets.

B Earnings Call Signals of Legal Exposure

Extracting keywords from earnings call transcripts is essential for identifying underlying signals of litigation risk faced by firms. This section of the online appendix provides additional details to clarify the robustness of my analysis and ensure its accuracy.

To begin, I adopt a conservative approach and deliberately exclude the keyword "litigation" and closely related terms. This decision is guided by two key considerations. First, most firms are required to include disclaimers at the start of their earnings calls, such as:

"Before we begin, may I draw your attention to the disclaimer on our presentation and company announcement regarding forward-looking statements as defined in the U.S. Private Securities Litigation Reform Act of 1995."

This standardized disclaimer, ubiquitous across firms, could inadvertently introduce noise into the analysis by classifying firms with no genuine litigation exposure into portfolios of firms with such risks. Including the term "litigation" could therefore compromise the

integrity of the portfolio sorting process. Second, the terms "litigation" and "legal" are often used interchangeably in earnings calls. Since "legal" is already incorporated into my analysis, excluding "litigation" does not result in any substantive loss of information or analytical rigor.

Another important aspect of keyword extraction is dealing with terms that have multiple meanings, which can vary significantly based on context. Take, for example, the term "judge." As a verb, it could simply mean making a decision or forming an opinion, which may have no legal relevance. However, the term "Judge" (i.e., used with a capital letter) specifically refers to a legal authority figure. To maintain the precision of the dataset, I include "Judge" but exclude the lowercase "judge" to ensure that I capture relevant legal contexts. While a similar approach could be applied to the term "court" (which could imply an attempt to entice someone), I include both "court" and "Court" because it is unlikely that "court" would refer to social enticement given the corporate context of earnings calls.

Litigations related to patents are quite common, making patent-related terms an important consideration in this analysis. However, to maintain precision, I include only terms with clear legal implications, such as "patent infringement," while excluding broader terms like "patent." The rationale for this distinction is that technology firms often reference their patents in contexts unrelated to legal issues, such as discussing innovation or intellectual property strategies. Including the term "patent" indiscriminately could introduce noise, as it might capture mentions that have no legal significance, thereby diluting the accuracy of the analysis. By focusing on terms with explicit legal linkage, I ensure the analysis remains targeted and relevant. Ambiguity in language poses a significant challenge in text analysis, particularly when extracting legal nuances from corporate communications. For example, the term "complaint" could refer to a legal filing but may also be used in non-legal contexts, such as describing a customer's dissatisfaction with a product. Such ambiguity can obscure the legal relevance of the term if not carefully managed.

To address this, terms like "complaint" when used in isolation are excluded so that I focus on more legally specific phrases, such as "amended complaint," which refers to a formal revision of a plaintiff's original filing. Similarly, ambiguous terms such as "claim"—which could refer to insurance claims or warranty claims in non-legal contexts—are excluded unless they appear in explicitly legal contexts, such as "legal claim," "counterclaim," or terms associated with clear legal proceedings. Additionally, terms like "settle" are carefully evaluated, as they may relate to financial contexts rather than legal settlements. By excluding

these ambiguous or context-dependent terms unless they are part of a legally meaningful phrase, the keyword extraction process is refined to more accurately identify instances of litigation risk while minimizing noise from irrelevant mentions.

As a result, the final set of keywords used to identify firm-level legal risk is listed below. Keywords frequently appearing in corporate earnings calls are highlighted in italics.

- **Judicial Entities and Key Participants (LRISK_COURT)**
  - **Judicial Entities**: Administrative Court, Appeal Court, Apppellate Court, Arbitration Court, Arbitration Tribunal, Bankruptcy Court, ***Commission***, Constitutional Court, ***Court(s)***, ***Department of Justice*** **(*DOJ*)**, District court, Eastern District, European Court, European Court of Justice (ECJ), Federal Court, Fifth Circuit, High Court, Lower Court, Northern District, Oversight Board, Patent Office, Southern District, ***Supreme Court***, Trial Court,
  - **Key Participants**: Administrative Law Judge (ALJ), Arbitrator, Attorney General, Defendants, Federal Judge, ***Judge***, Jury, Plaintiff(s), Prosecutor(s)
- **Legal Processes and Disputes (LRISK_TRIAL)**
  - **Legal Processes**: ***Appeals***, ***Arbitration***, Bankruptcy Process, Binding Arbitration, Class Action, Court Approval, Court Hearing, Court Process, Court Proceedings, Evidentiary Hearings, ***Hearing(s)***, IPR Process, Judicial Process, Judicial Review, Legal Proceeding, Legal Process, Markman, Mediation Process, Next Hearing, Oral Argument, Oral Hearing, Pending Cases, Rebuttal Testimony, Retrial, Trial Date, Full Trial, ***Petition***, Procedural Schedule, ***Settlement***, ***Settlement Agreement***
  - **Legal Disputes**: Adverse Ruling, Antitrust Case, Arbitration Case, Class Action Lawsuit, Civil Cases, Civil Litigation, Infringement, ***Injunction***, IP Litigation, ***Lawsuit(s)***, Legal Challenges, Legal Claim, Legal Disputes, Patent Case, Patent Claims, Patent Dispute, Patent Infringement, ***Ruling***, Settlement Agreement, Tentative Settlement
- **Legal Terms and Concepts (LRISK_TERMS)**
  - Amended Complaint, Case Law, Compliance Filing, Constitutionality, Counterclaims, Injunctive Relief, Legal Basis, ***Legal Costs***, ***Legal Fees***, ***Legal Expenses***, Sub Judice, Regulatory Issue

C Additional Asset Pricing Results and Data Mining Concerns

This section presents additional analyses to assess the robustness and distinctiveness of the **LAW** factor. I begin with alternative calendar-time portfolio tests including different formation windows. I then address data mining concerns to directly confront the "factor zoo" critique. I further examine whether the **LAW** factor is subsumed by broader behavioral explanations, including market sentiment, mispricing-based factors, and belief disagreement. I also test whether short-term mood fluctuations—proxied by day-of-the-week return patterns—can account for the legal risk premium. Lastly, I explore its relationship with ESG-based risk factors. Together, these exercises reinforce the robustness and economic distinctiveness of the **LAW** factor and mitigate concerns about data mining.

C.1 Robustness Checks on the Calendar-time Portfolio Analysis

This section extends the empirical analysis by introducing alternative portfolio formation strategies to test the robustness of the main findings. The baseline specification in the main text adopts a (12-12) strategy, where portfolios are constructed at the end of June each year using legal risk exposures measured from the past 12 months (i.e., four quarters) of earnings call transcripts. Stocks are value-weighted using market capitalization as of that date, and the portfolio is held for 12 months, ending in June of the following year. This approach follows standard conventions in the asset pricing literature for constructing long-short anomaly portfolios.

To take advantage of the quarterly frequency of earnings calls, Table A.1 presents results for four alternative strategies with more frequent rebalancing: (12-3), (12-6), (9-9), and (18-6).[1] These strategies vary in both the formation window and holding period, while maintaining a consistent tercile-sorting methodology based on ex-ante legal risk. In each strategy, firms with no mentions of legal risk are assigned to L1; firms with positive legal risk exposure are split at the median into L2 (medium risk) and L3 (high risk). The "L3 – L1" column reports returns from a tradable long-short portfolio that buys the high legal risk group and sells the zero-risk group.

[1] In each strategy label (e.g., (12-3)), the first number denotes the formation window in months, and the second number indicates the rebalancing frequency in months. For example, the (12-3) strategy uses the past 12 months (i.e., 4 quarters) of earnings call transcripts to measure firms' legal risk exposure and rebalances the portfolio every 3 months (i.e., quarterly).

All portfolios are value-weighted using market capitalization at the time of formation. To isolate the effect of formation and rebalancing frequencies, the original sample restrictions are preserved: utility and financial firms are excluded due to regulatory constraints. The table reports average monthly excess returns and Fama–French three-factor plus momentum alphas, along with Newey–West t-statistics computed using six lags. Across all four strategies, the long-short portfolios as presented in Table A.1 deliver statistically and economically significant returns. These results reinforce the robustness of the main findings to alternative portfolio formation schemes.

In unreported robustness checks, I confirm that the results remain qualitatively unchanged under the following variations: (1) including utility and financial firms; (2) capping the market capitalization of the earnings call sample at the 95th percentile, which roughly corresponds to the 80th percentile NYSE breakpoint used by Jensen et al. (2023); and (3) excluding the so-called "Magnificent 7" firms—Apple, Microsoft, Amazon, Alphabet (Google), Meta (Facebook), Tesla, and Nvidia—which collectively represent a substantial share of total market capitalization in recent years. Importantly, as illustrated in Figure A.2, these well-known mega-cap firms also migrate across tercile groups over time, reinforcing the notion that legal risk exposure is both dynamic and widespread.[2]

[2]Figure A.2 shows that firms frequently transition across legal risk categories. For instance, Ford Motor Company experienced an escalation from moderate to high legal risk during the mid-2000s, linked to lawsuits involving employment discrimination, loan practices, and product liability. Similarly, Coca-Cola saw increased legal risk exposure in 2020–2021, driven by lawsuits such as *Earth Island Institute v. Coca-Cola Co.* over alleged environmental greenwashing and *Swartz v. Coca-Cola Co.* regarding misleading recyclability claims.

Table A.1: **Long-Short Strategy Results using Different Formation Period**

Panel A: (12-3) strategy (%)

	L1 No legal risk	L2 Middle legal risk	L3 High legal risk	L3 - L1 long - short
Excess return	0.626	0.948	0.973	**0.347*** **
	(2.02)	(3.26)	(3.29)	**(3.34)**
FF3 + mom alpha	-0.151	0.119	0.162	**0.313*** **
	(-1.90)	(2.41)	(2.90)	**(3.08)**

Panel B: (12-6) strategy (%)

	L1 No legal risk	L2 Middle legal risk	L3 High legal risk	L3 - L1 Long - Short
Excess return	0.670	0.916	0.998	**0.328*** **
	(2.14)	(3.18)	(3.35)	**(3.66)**
FF3 + mom alpha	-0.129	0.095	0.185	**0.315*** **
	(-1.66)	(1.89)	(3.35)	**(3.35)**

Panel C: (9-9) strategy (%)

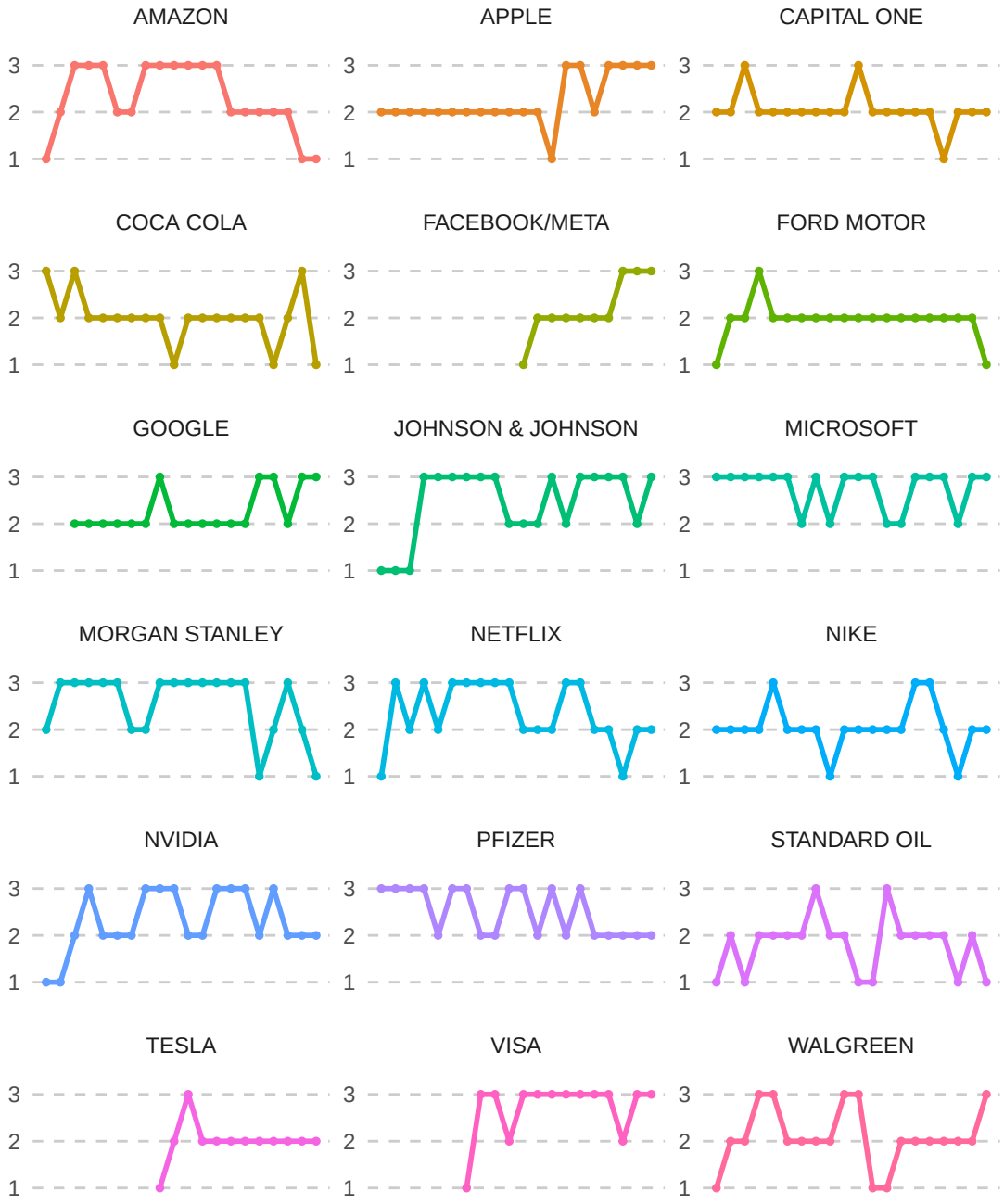
	L1 No legal risk	L2 Middle legal risk	L3 High legal risk	L3 - L1 Long - Short
Excess return	0.686	0.929	1.010	**0.325*** **
	(2.20)	(3.21)	(3.43)	**(2.94)**
FF3 + mom alpha	-0.096	0.103	0.196	**0.291*** **
	(-1.14)	(2.26)	(2.85)	**(2.68)**

Panel D: (18-6) strategy (%)

	L1 No legal risk	L2 Middle legal risk	L3 High legal risk	L3 - L1 Long - Short
Excess return	0.697	0.868	1.020	**0.324*** **
	(2.14)	(2.99)	(3.46)	**(3.17)**
FF3 + mom alpha	-0.123	0.062	0.200	**0.323*** **
	(-1.37)	(1.20)	(3.95)	**(3.13)**

Note: Each panel reports average future excess returns (2004:07 - 2024:06) and FF3 + momentum alphas from portfolios formed based on ex-ante legal risk measures. The strategy in each panel is defined by two parameters in parentheses: the portfolio formation window and the rebalancing frequency, both in months. For example, Panel A shows results for the (12-3) strategy, where firms are sorted based on legal risk observed over the past 12 months (4 quarters) and portfolios are rebalanced every 3 months (1 quarter). Panels B–D present analogous strategies: (12-6), (9-9), and (18-6), respectively. In each formation period, firms are sorted into terciles based on their legal risk exposure. L1 includes firms with no legal risk mentions. Firms with nonzero legal risk are split at the median into L2 (medium risk) and L3 (high risk). The "L3 – L1" column shows long-short portfolio results (high minus no legal risk). Utility firms (SIC 4900–4999) and financial firms (SIC 6000–6411, 6500–6553, and 6700–6799) are excluded due to heavy regulatory constraints. Newey-West t-statistics with 6 lags are shown in parentheses. ***/**/* indicate statistical significance at the 1%/5%/10% level, respectively.

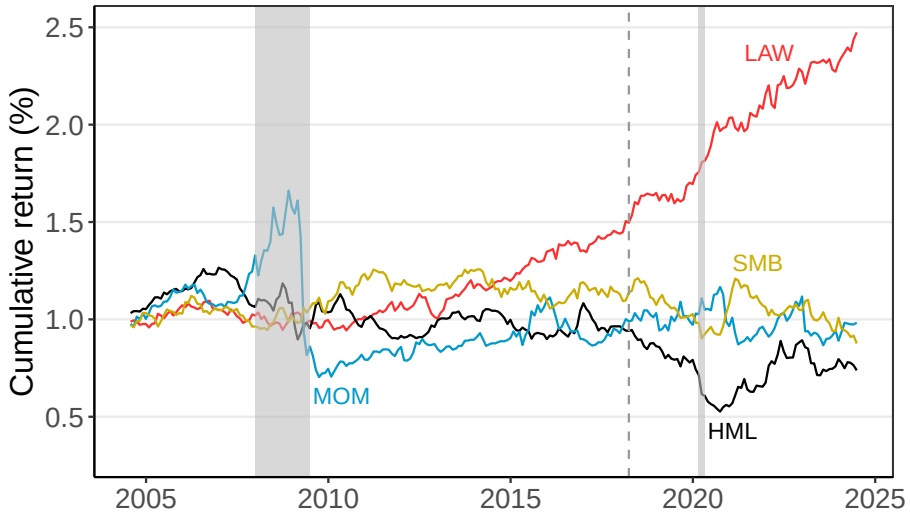
Figure A.2: **Legal Exposure of Well-known Firms**



Note: This figure illustrates the legal risk exposure of prominent corporate entities analyzed in this study. The x-axis shows the years over which portfolio sorting based on firm-level legal risk exposure was conducted, spanning from 2003 to 2023. The y-axis categorizes the legal risk into three levels: 1 indicates no legal risk, 2 represents moderate legal risk, and 3 denotes high legal risk.

C.2 Data Mining Concerns

Figure A.3: **Cumulative Return Performance of Legal Risk Factor (LAW)**



Note: This figure plots the cumulative return performance of $1 invested in each of the four risk factors (LAW, SMB, HML, and MOM) at the end of 2004:06. The shaded area are NBER economic recession periods, and the dotted vertical line represents when *Cyan, Inc. v. Beaver County Employees Retirement Fund (2018)* decision was released.

A recent number of articles address the potential data mining concerns. This section briefly details three reasons why I believe my **LAW** factor result is not data-mined.

First, my t-statistic result substantially exceeds the new threshold of 3.0 proposed by Harvey et al. (2016), providing strong evidence against data-mining concerns (3.66 for the full 20-year sample spanning 2004–2024, and 4.91 for the post-2010 sample). Second, unlike many anomalies that emerge from mechanical combinations of accounting variables (such as those simulated in Chen and Dim 2024 and Novy-Marx and Velikov 2025), my **LAW** factor is constructed from novel, forward-looking text-based information extracted from earnings call transcripts.

Second, while many traditional anomalies exhibit sharp decay in predictability after the early 2000s (Chordia et al. 2014), the **LAW** factor remains strong and stable throughout the post-2004 period (in fact, my sample starts in 2004). Figure A.3 plots the cumulative

return to a $1 investment in the long-short **LAW** portfolio alongside standard factors. Two key inflection points stand out: the 2018 surge aligns with the Supreme Court's Cyan, Inc. v. Beaver County decision, which increased forum-shopping risk and litigation costs; the 2020 spike coincides with the onset of COVID-19, which amplified legal uncertainty across sectors. Over the full period, $1 invested in the LAW factor grows to $2.38, while SMB, HML, and MOM shrink to $0.91, $0.78, and $0.97, respectively.[3]

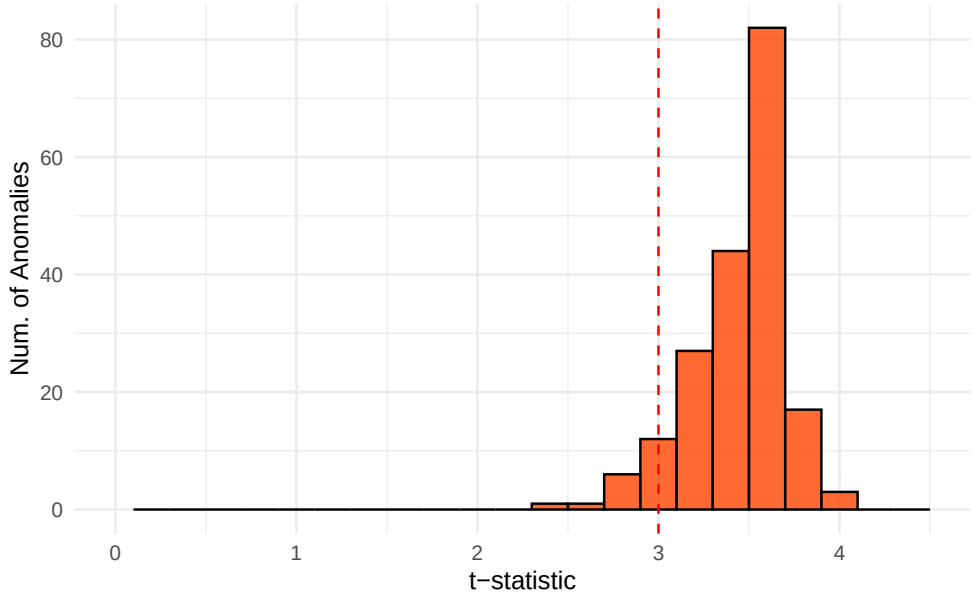
Importantly, the **LAW** factor demonstrates resilience during economic downturns, including the 2008 Global Financial Crisis and the 2020 COVID shock. Legal risks become particularly salient in these periods, as firms and investors are already under financial stress. Investors tend to value firms whose legal risk remains stable—or demand higher returns from those with greater exposure—driving the outperformance of the long-short strategy in downturns. By contrast, other risk factors, particularly momentum, tend to experience sharp declines during or shortly after recessions (Daniel and Moskowitz 2016).

Finally, Figure A.4 further underscores the uniqueness of the legal risk factor (**LAW**) within the broader factor zoo. Across 193 spanning regressions—each controlling for the Fama–French three factors, momentum, and one anomaly factor at a time—the **LAW** factor retains a statistically significant intercept, with Newey–West adjusted t-statistics exceeding 2.0 in all cases and remaining above 3.0 in 181 of them (93.7% of the factors zoo). These results indicate that **LAW** is not subsumed by any known return predictor and thus represents a robust, economically meaningful source of priced risk. The sample period spans July 2004 to December 2023, and the anomaly factors are obtained from Chen and Zimmermann (2022).[4]

[3] Individually, firms in the lowest legal risk tercile (q1) grow from $1 to $3.23, while q2 and q3 portfolios grow to $6.31 and $8.40, respectively.

[4] Additionally, I find that the Pearson correlations between the **LAW** factor and standard asset pricing factors are modest in magnitude: **MKT** (–0.14), **SMB** (–0.26), **HML** (–0.19), **MOM** (0.25), **RMW** (0.10), **CMA** (0.04), and **LIQ** (–0.02). Notably, none of the 193 anomaly factors exhibit an absolute correlation with **LAW** greater than 0.4, further reinforcing the distinctiveness of legal risk as a priced dimension in the cross-section of returns.

Figure A.4: **Spanning Regressions: Distinctiveness of the LAW Factor**



Note: This figure displays the distribution of Newey–West adjusted (6 lags) t-statistics for the intercept term (α) from 193 spanning regressions of the **LAW** factor on the Fama–French three factors, the momentum factor, and one of 193 additional anomaly factors from the finance literature. Each regression takes the form: $\textbf{LAW} = \alpha + \beta_{MKT}\textbf{MKT} + \beta_{SMB}\textbf{SMB} + \beta_{HML}\textbf{HML} + \beta_{MOM}\textbf{MOM} + \beta_{ZOO}\textbf{ZOO}$, where **ZOO** represents one of the benchmark anomalies. The t-statistic on α exceeds 2.0 in all cases and remains above 3.0 in 181 out of 193 specifications, indicating that the **LAW** factor remains largely unexplained by known return predictors. The sample spans July 2004 to June 2023.

C.3 Market Sentiment and Mispricing explanations

In the behavioral finance literature, there has been a sustained effort to explain return predictability and risk premia through the lens of mispricing and investor biases. Prominent examples include models that attribute anomalies to underreaction, overreaction, or limits to arbitrage, such as the PEAD and FIN factors proposed by Daniel et al. (2020), and the MGMT and PERF factors developed by Stambaugh (2017), which aggregate information across known anomalies related to managerial behavior and firm performance, respectively. A related strand of research emphasizes investor belief dispersion as a source of pricing distortions, culminating in the construction of a disagreement index by Huang et al. (2021), which combines 24 individual disagreement measures using a partial least squares (PLS) approach.

Against this backdrop, I examine the **LAW** factor and find that it is not subsumed by these existing behavioral constructs. The results are presented in Table A.2. First, the returns to the **LAW** factor – whether measured on the long leg, short leg, or as a long-short portfolio – do not exhibit sensitivity to market-wide sentiment, as proxied by the index of Huang et al. (2015), in contrast to the pattern observed for traditional anomalies (Stambaugh et al. 2012). Second, regressions controlling for PEAD, FIN, MGMT, PERF, and the aggregate disagreement index of Huang et al. (2021) all leave the alpha on LAW economically meaningful and statistically significant. These findings suggest that the premium associated with legal risk reflects a distinct pricing channel—one that is not captured by current models of mispricing, sentiment, or disagreement.

Finally, to examine the role of mood in explaining the legal risk result, following the approach of Birru (2018), I estimate separate regressions of the daily return on the legal risk factor (LAW) for each weekday. Specifically, for each day $d \in \{1, 2, 3, 4, 5\}$ corresponding to Monday through Friday, I estimate:

$$\text{LAW}_t = \alpha_d + \varepsilon_t, \quad \text{for all } t \text{ such that } \text{Day}_t = d, \tag{A.1}$$

where LAW_t denotes the daily return on the legal risk factor, α_d captures the mean return on day d, and ε_t is the residual. Additionally, I estimate a pooled regression for midweek days by restricting the sample to $d \in \{2, 3, 4\}$ (Tuesday through Thursday).

The results, reported in Table A.3, show that the **LAW** factor earns significantly positive returns on Mondays through Thursdays. In contrast, Friday returns are smaller in magnitude and not statistically significant. This pattern is inconsistent with the mood-based mispricing hypothesis of Birru (2018), which predicts that returns for speculative long-leg factors should peak on Fridays when investor sentiment is highest. Instead, the evidence indicates that the LAW factor generates strong and persistent returns throughout the week, without a meaningful end-of-week increase. These findings suggest that legal risk exposure is unlikely to be driven by short-term mood fluctuations, and more plausibly reflects compensation for bearing systematic risk.

Table A.2: **Sensitivity of the LAW Factor to Market-Wide Sentiment and Mispricing-Based Risk Factors**

Panel A: Predictive regression results

Regression is of the form $LAW_{t+1}^{tercile} = \alpha + \beta s_t + \varepsilon_{t+1}$, where s_t is the Huang et al. (2015) market sentiment

Dep. Variable	β	$t(\beta)$	R-squared (%)	Constant?
$\mathbf{LAW^{Long-Short}}$	0.205	(1.22)	0.3	Yes
$\mathbf{LAW^{L1}}$	-0.537	(-0.91)	0.3	Yes
$\mathbf{LAW^{L3}}$	-0.331	(-0.58)	0.1	Yes

Panel B: Mispricing Factors and Disagreement

	(1)	(2)	(3)	(4)
Dependent variable:	LAW	LAW	LAW	LAW
↓ Independent variable:				
Intercept	**0.377*** (3.59)**	**0.358** (2.48)**	**0.312*** (3.01)**	**0.396*** (2.65)**
PEAD	-0.077 (-1.03)			-0.067 (-0.75)
FIN	0.019 (0.34)			-0.177** (-2.44)
MGMT		0.175** (2.09)		0.308*** (3.49)
PERF		-0.212*** (-3.91)		-0.139** (-2.27)
Disagreement			-0.286* (-1.83)	-0.087 (-0.44)
FF3 + mom controlled?	Yes	Yes	Yes	Yes
Sample	2004:07 - 2023:12	2004:07 - 2016:12	2004:07 - 2020:12	2004:07 - 2016:12
Adj. R-squared	0.090	0.172	0.111	0.177

Note: This table examines whether the LAW factor is subsumed by market-wide investor sentiment and mispricing-based risk factors. Panel A presents predictive regressions of LAW factor returns across tercile-sorted portfolios (low, high, long-short) on lagged monthly sentiment from Huang et al. (2015). Panel B reports time-series regressions of the LAW factor on combinations of mispricing-based factors: PEAD and FIN from Daniel et al. (2020), and MGMT and PERF from Stambaugh (2017), along with a proxy for belief disagreement (Huang et al. 2021). All regressions include controls for the Fama-French three factors and the momentum factor (FF3 + MOM), but their coefficients are omitted for brevity. The sample periods vary by factor availability and are indicated in each column. T-statistics are Newey-West adjusted with 6 lags, in parentheses. ***/**/* denote the statistical significance at 1%/5%/10% level.

Table A.3: **Day-of-the-Week Effects on LAW Returns**

	Monday	Tuesday	Wednesday	Thursday	Friday
Estimate	**0.027*** **	**0.016*** **	**0.017*** **	**0.017*** **	0.009
t-stat	**(5.07)**	**(2.89)**	**(3.17)**	**(3.07)**	(1.57)
N	939	1,032	1,034	1,018	1,010

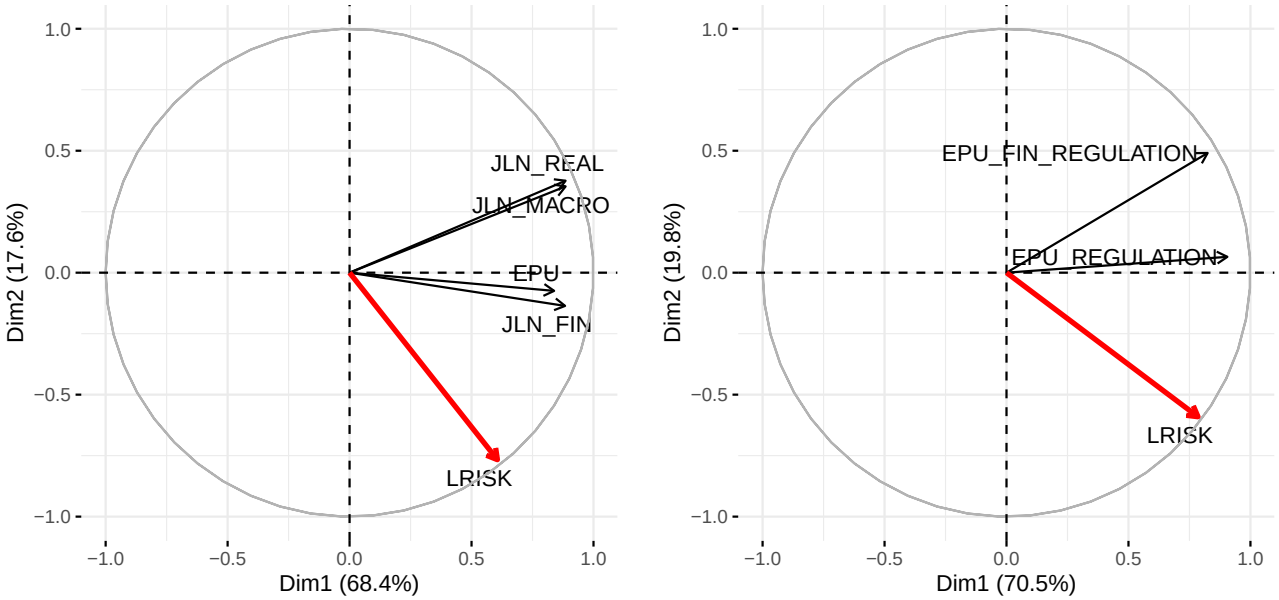
Note: Notes: This table reports the estimated mean returns and Newey-West t-statistics (lag = 6) from regressions of the daily LAW factor return on a constant, estimated separately for each weekday. ***/**/* denote statistical significance at the 1%, 5%, and 10% levels, respectively. The number of observations (N) varies by day.

C.4 LAW Factor vs. ESG Factors

In Table A.4, I conduct spanning regression tests using ESG factors from Pedersen et al. (2021) and Pastor et al. (2022). The results in Panel B indicate that the LAW factor is not subsumed by any of the ESG-related factors, underscoring its distinct role in capturing priced risk. Conversely, some ESG factors—such as GMB in column (5) of Panel A—retain significance when used as dependent variables, suggesting partial overlap between ESG and legal risk. Overall, these results imply that while ESG and legal risk factors may share some common ground, the LAW factor captures a unique dimension of financial market risk premia.

D Aggregate LRISK vs. Uncertainty Measures

Figure A.5: **Principal Component Loadings of Legal Risk and Other Uncertainty Measures**



(a) LRISK and JLN & EPU Index

(b) LRISK and EPU Subcategories

Note: This figure presents PCA loadings for the aggregate legal risk measure (LRISK) alongside other widely used measures of uncertainty. Panel (a) compares LRISK with broad indices, including the Economic Policy Uncertainty (EPU) index and three components of the Jurado-Ludvigson-Ng (JLN) framework: Macroeconomic, Financial, and Real Uncertainty. Panel (b) focuses on a narrower comparison, displaying PCA loadings for LRISK alongside two subcomponents of the EPU index—EPU Regulation and EPU Financial Regulation. In both panels, PCA is conducted using standardized variables, and the loadings are derived from the correlation matrix. The direction and length of each arrow represent the variable's contribution and alignment with the principal components. The distinct orientation of LRISK relative to the other measures indicates that it captures a unique and orthogonal dimension of uncertainty not accounted for by macroeconomic or policy-based indices.

This section examines how my measure of aggregate legal risk—constructed as a coverage ratio, as described in Section 3.5—compares to established measures of economic uncertainty, including the Economic Policy Uncertainty (EPU) index of Baker et al. (2016) and the Macro Uncertainty index of Jurado et al. (2015) (JLN). To analyze the relationships among these variables, I conduct a principal component analysis (PCA) using quarterly data spanning 2004 to 2024. Since the EPU and JLN indexes are originally available at a monthly frequency, I convert them to quarterly by selecting the end-of-quarter observa-

tion. For the JLN measure, I use the $h = 3$ horizon to align with the quarterly frequency. All variables are standardized to have zero mean and unit variance prior to estimation. PCA is then applied to the correlation matrix to extract the primary dimensions of variation across the uncertainty measures.

Figure A.5 displays the resulting factor loadings for LRISK and the alternative indices. In both panels, the direction and length of each arrow represent the contribution and alignment of each variable with the principal components. Variables that point in similar directions are highly correlated, while those that are orthogonal reflect distinct sources of variation. The noticeably different orientation of LRISK relative to EPU and JLN measures suggests that it loads onto a separate dimension. This visual evidence supports the interpretation that legal risk captures an independent and previously unaccounted for component of uncertainty in the cross-section of firms and over time.

Table A.4: **Comparing LAW factor and ESG-relevant factors**

Panel A: ESG factors are dependent variables

	PFP factors (1 through 4)				PST factors (5 through 7)		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
→ Dependent variable:	**ESG**	**E**	**S**	**G**	**GMB**	**G**	**B**
↓ Independent variable:							
Intercept	-0.087	0.333	-0.163	-0.119	**0.289***	**0.124***	-0.165
	(-0.51)	(1.61)	(-1.09)	(-0.67)	**(1.69)**	**(1.74)**	(-1.34)
LAW	**0.211****	0.082	-0.050	-0.028	**0.134***	**0.115*****	-0.019
	(2.10)	(0.86)	(-0.56)	(-0.32)	**(1.86)**	**(2.59)**	(-0.39)
Sample	2007:02 - 2019:03	2009:06 - 2019:03	2004:07 - 2019:03	2004:07 - 2019:03	2009:01 - 2023:12	2009:01 - 2023:12	2009:01 - 2023:12
FF5 + mom controlled?	Yes	Yes	Yes	Yes	Yes	Yes	Yes

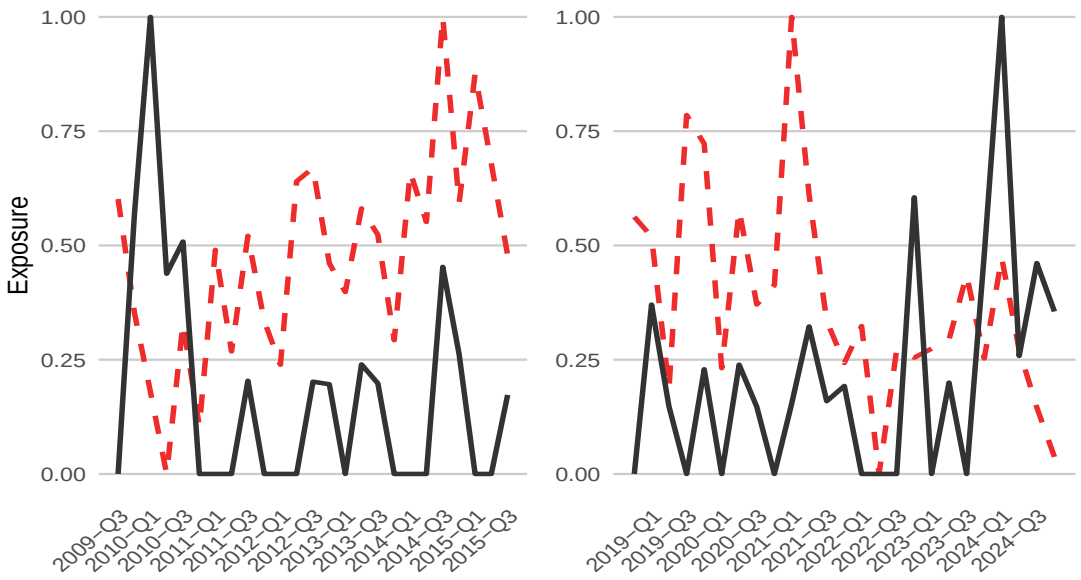
Panel B: LAW factor is dependent variable

	LAW factor						
↓ Independent variable:							
Intercept	**0.330****	**0.412****	**0.328****	**0.332****	**0.391*****	**0.356*****	**0.431*****
	(2.26)	**(2.44)**	**(2.32)**	**(2.31)**	**(3.40)**	**(3.09)**	**(3.79)**
(1) ESG	**0.225****						
	(2.57)						
(2) E		0.055					
		(0.95)					
(3) S			-0.036				
			(-0.56)				
(4) G				-0.022			
				(-0.31)			
(5) GMB					**0.131***		
					(1.77)		
(6) G						**0.465****	
						(2.34)	
(7) B							-0.035
							(-0.39)
Sample	2007:02 - 2019:03	2009:06 - 2019:03	2004:07 - 2019:03	2004:07 - 2019:03	2009:01 - 2023:12	2009:01 - 2023:12	2009:01 - 2023:12
FF5 + mom controlled?	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Note: This table presents spanning tests using Fama-French 5 factors augmented with momentum factors. Panel A uses ESG factors from Pedersen et al. (2021) (PFP factors) and GMB factors from Pastor et al. (2022) (PST factors) as dependent variables. Panel B uses LAW factor as dependent variable. Loadings on FF5 and momentum factors are suppressed. The numbers in parentheses represent Newey-West t-statistics with six lags. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

E Firm-level Evidence: Case studies using SolarWinds

Figure A.6: **Case study - SolarWinds**



Note: This figure presents the legal risk exposure (black solid line) and cyber risk exposure (red dotted line) for SolarWinds Inc. Both text-based exposure measures are normalized to range from 0 to 1 within each subperiod: 2009-Q3 to 2015-Q3 and 2018-Q4 to 2024-Q4. SolarWinds first went public in May 2009, was taken private in October 2015, and subsequently returned to the public market in October 2018.

SolarWinds[5] provides a compelling case study to discuss legal risk exposure of that firm. The study of SolarWinds' cyber risk exposure, pioneered in Florackis et al. (2023) and Jamilov et al. (2025) provide an interesting benchmark to illustrate the legal risk inherent in the cyberrisk-related issues. My analysis builds on these findings by showing that firms responding to cyber risk often face legal risk as a related but distinct mechanism. Notably, SolarWinds first encountered a securities class action lawsuit in 2010, which resulted in a Notice of Voluntary Dismissal. Despite this resolution, the textual analysis of earnings call transcripts during this period shows a noticeable increase in legal risk-related discussions

[5]The company, a provider of IT infrastructure management software, first went public in May 2009. However, in October 2015, it was taken private in a $4.5 billion buyout by private equity firms Silver Lake Partners and Thoma Bravo. In October 2018, SolarWinds re-entered the public markets with its second initial public offering (IPO). More recently, the company announced its intention to delist again in 2025 following an acquisition agreement with Turn/River Capital. Because of this transition between public and private ownership, my analysis presents two separate graphs to illustrate how legal risk exposure and market perception evolved over these distinct periods.

(black solid line in left panel of Figure A.6). This trend reemerged in 2015 when the company was again subject to a class action lawsuit. In this instance, the case concluded with a Court's Order of Dismissal, yet the company's legal risk exposure remained heightened. These early signals suggest that, even when lawsuits do not directly impact financial outcomes, firms anticipate and communicate about potential legal risks in their disclosures, which influences investor perceptions and stock price behavior.

A defining moment for SolarWinds occurred in December 2020 when it became the target of the SolarWinds supply chain attack (SUNBURST attack) that affected thousands of government and private-sector clients. As expected, my analysis shows a sharp spike in the firm's cyber risk measure (denoted by the red dotted line on the right panel) in response to the attack. This event triggered cascading legal consequences, including a class action lawsuit from shareholders and customers who alleged that SolarWinds had failed to implement adequate cybersecurity measures. In addition to these lawsuits, regulatory scrutiny intensified. In October 2023, the U.S. Securities and Exchange Commission (SEC) filed charges against SolarWinds and its Chief Information Security Officer (CISO) for allegedly misleading investors about the company's cybersecurity practices before and during the breach. The SEC's charges significantly increased SolarWinds' legal risk exposure, as reflected in a pronounced uptick in the legal risk measure. This exposure culminated in early 2024, when SolarWinds agreed to a $26 million settlement to resolve investor claims stemming from the 2020 breach.

F Legal Shifts: Implications for Firms and Markets

This section examines the evolution of the U.S. legal landscape throughout the 21st century and its implications for firms and markets. Specifically, I explore how the early 2000s were characterized by a mix of deregulatory measures and heightened regulatory oversight, while the post-Global Financial Crisis (GFC) era introduced a wave of stringent reforms that reshaped the regulatory framework.

F.1 The Rise of Aggregate Legal Risk in the Early 2000s

Background: The early 2000s were characterized by a series of mixed legal developments that, in aggregate, contributed to an increase in firm-level legal risk and shaped the regulatory climate leading up to the Global Financial Crisis (GFC). On one hand, deregulatory

measures such as the repeal of key provisions of the Glass-Steagall Act via the Gramm-Leach-Bliley Act (GLBA) in November 1999 facilitated financial consolidation. Originally enacted in 1933 to separate commercial and investment banking, the Glass-Steagall Act aimed to limit conflicts of interest and systemic risk. Its repeal under the GLBA allowed financial institutions to operate across banking, securities, and insurance sectors, accelerating the formation of financial conglomerates such as Citigroup. This deregulatory shift expanded the operational scope of financial firms, but also exposed them to more complex and potentially risk-laden legal environments.

Simultaneously, legal reforms at the state level contributed to a more insulated corporate governance regime. In particular, the widespread adoption of Universal Demand (UD) laws during the 1980s and 1990s, and their ongoing effects into the 2000s, substantially weakened shareholder litigation as a mechanism for corporate accountability. These laws required shareholders to make a formal demand to the board before initiating derivative lawsuits, thereby granting directors increased discretion to block litigation. The long-lasting effect of UD laws was a decline in shareholder lawsuits and reduced external monitoring of managerial behavior. Firms incorporated in UD states, as shown by Houston et al. (2018), experienced significantly fewer lawsuits, even in instances of governance failures—effectively lowering litigation risk for managers while raising latent risks for investors and stakeholders.

In contrast to these deregulatory trends, the Sarbanes-Oxley Act (SOX) of 2002 marked a sharp shift toward regulatory tightening in response to high-profile accounting scandals such as Enron and WorldCom. SOX introduced stringent rules on internal controls, mandated executive certification of financial reports, and imposed heavier disclosure burdens on public firms. While SOX sought to restore investor confidence and strengthen oversight, its implementation introduced considerable legal and compliance risk, particularly for firms with opaque reporting practices or weak governance structures.

Taken together, these concurrent shifts created a complex legal landscape. On one side, deregulatory measures such as the GLBA and the enduring impact of UD laws reduced the external constraints on managerial behavior, encouraging risk-taking and innovation. On the other, SOX raised the cost of non-compliance and increased the legal consequences of disclosure failures. The aggregate effect was an elevation in firm-level legal risk, as companies navigated a regulatory environment that was both permissive and punitive. This asymmetry—lax enforcement in some domains and heightened scrutiny in others—contributed to mounting systemic fragilities that were ultimately exposed during the GFC.

F.2 Stringent Legal Shifts After the Global Financial Crisis

Background: The 2008 GFC exposed systemic weaknesses in the U.S. financial system, including excessive risk-taking, regulatory gaps, and corporate misconduct. The resulting economic downturn and public outrage over corporate malfeasance prompted a wave of legislative and regulatory reforms designed to prevent similar crises in the future. The Dodd-Frank Wall Street Reform and Consumer Protection Act of 2010 was a cornerstone of these reforms, introducing comprehensive measures to enhance market oversight, improve transparency, and promote corporate accountability.

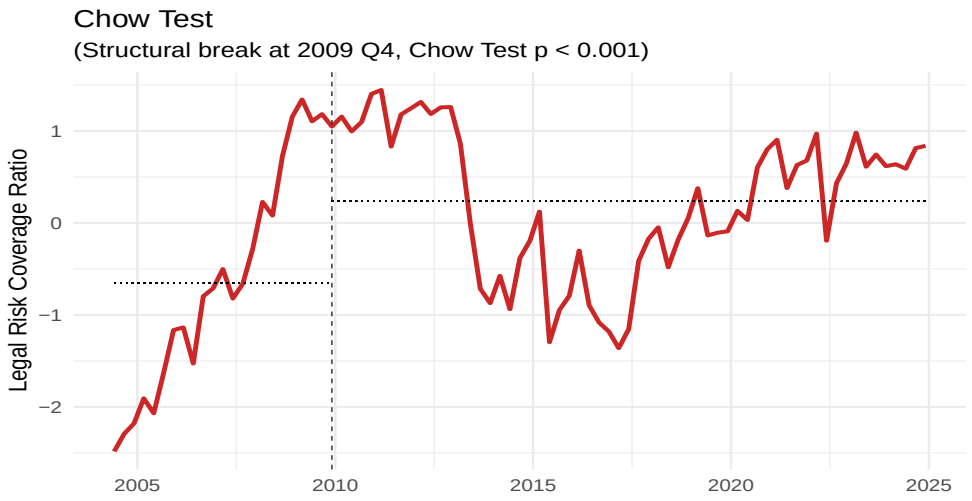
The Securities and Exchange Commission (SEC) played a pivotal role in these reforms by implementing stricter enforcement measures. In 2010, the SEC established specialized units within its Enforcement Division, such as the Market Abuse Unit and the Asset Management Unit, to address complex financial misconduct more effectively. Additionally, the SEC's cooperation program encouraged firms and individuals to assist in investigations through tools like cooperation agreements, deferred prosecution agreements, and non-prosecution agreements. According to Leone et al. (2021), the program's early years were marked by higher enforcement probabilities and increased penalties. However, from 2011 to 2014, enhanced cooperation led to reduced enforcement probabilities and lower penalties, reflecting the program's evolving effectiveness.

A landmark provision in securities law enforcement came with the whistleblower program established under Section 922 of the Dodd-Frank Act. This program offered monetary rewards ranging from 10% to 30% of sanctions collected from SEC enforcement actions exceeding $1 million, alongside robust anti-retaliation protections for whistleblowers. These incentives encouraged individuals to report misconduct early, thereby strengthening the SEC's ability to detect and deter securities violations. Heese and Perez-Vacazos (2021) examine how the costs of retaliation influence employees' decisions to whistleblow, underscoring the program's effectiveness. Beyond the SEC's whistleblower program, the *Cyan, Inc. v. Beaver County Employees Retirement Fund* (*2018*) decision further heightened legal risk in the U.S. by allowing class-action lawsuits under the Securities Act of 1933 to proceed in state courts, increasing the risk of forum shopping. This exposed public companies, particularly those undergoing IPOs or secondary offerings, to costly and unpredictable litigation across multiple venues. Beyond direct legal costs, Cyan raised uncertainty around liability, complicated securities defense strategies, and increased the legal risk premium investors demanded. The prospect of parallel lawsuits in different states strained corporate legal teams, drove up settlement costs, and deterred some firms from

capital-raising activities. The ruling also impacted D&O insurance markets, as insurers adjusted premiums to reflect heightened litigation exposure. Alongside other post-GFC regulatory shifts, Cyan played a key role in shaping the modern U.S. legal risk landscape and its asset pricing implications.

Impact on Firms Listed on U.S. Public Markets: The post-GFC reforms significantly elevated the legal risk exposure for publicly listed firms. Companies faced greater regulatory scrutiny and higher compliance costs as they bolstered internal controls and governance frameworks to mitigate legal risks. Financial institutions like JPMorgan Chase and Goldman Sachs incurred substantial penalties for past misconduct and adopted stringent compliance measures in response. The whistleblower provisions of the Dodd-Frank Act further amplified firms' legal risk, evidenced by high-profile cases like the $104 million reward to UBS whistleblower Bradley Birkenfeld, whose disclosures led to significant penalties for the firm. These reforms marked a transformative shift in the U.S. legal landscape, with firms now encountering intensified accountability for both historical and ongoing practices.

Figure A.7: **Chow Test**



Note: This figure plots the Chow test around 2010 using the aggregate legal risk measure introduced in Section 3.5. The measure is based on the coverage ratio—the proportion of firms referencing legal risk in their earnings call transcripts. Each quarter, firms are classified into two groups: zero firms, which made no mention of legal risk in the past four quarters, and non-zero firms, which did. The coverage ratio is computed as the share of non-zero firms relative to the total number of firms. To facilitate interpretation over time, the series is standardized using a z-score transformation, highlighting deviations from the historical average.

The findings in this paper, based on post-2010 data, are consistent with the results in Figure A.7, where the Chow test clearly indicates a structural break around 2010. This break marks a significant shift in the legal environment, corresponding to the implementation of major post-GFC reforms. Together, these results underscore the substantial impact of legal regime changes on U.S. public markets.

G Modeling Legal Risk in a Real Options Framework

This section presents two complementary real options models that incorporate legal risk: a McDonald and Siegel (1986)–type model in which investment yields a one-time payoff, and a Dixit and Pindyck (1994)–type model in which investment generates an ongoing stream of cash flows. The purpose is not to provide full mathematical derivations, but rather to offer intuition for how legal risk alters investment incentives, specifically, by increasing the cost of commitment or amplifying uncertainty. In both frameworks, legal risk modifies the firm's investment behavior by raising the option value of waiting, thereby leading firms to scale back or delay irreversible investment.

G.1 McDonald–Siegel Framework with Legal Risk as an Investment Cost

Following McDonald and Siegel (1986), consider a firm that holds an irreversible option to invest in a project with stochastic value V_t. Investment requires a fixed cost F, and the firm chooses the optimal stopping time τ to maximize the expected discounted payoff:

$$\underbrace{X(V)}_{\text{Value of waiting}} = \sup_{\tau} \mathbb{E}_0 \left[e^{-r\tau} (V_\tau - F) \right], \tag{A.2}$$

where r is the discount rate.

The project value V_t evolves according to geometric Brownian motion:

$$\frac{dV_t}{V_t} = \mu \, dt + \sigma \, dZ_t, \tag{A.3}$$

with constant drift μ, volatility σ, and Z_t a standard Brownian motion.

The firm invests only when $V_t \geq V^*$, where V^* is the optimal investment threshold:

$$\beta = \frac{1}{2} - \frac{\mu}{\sigma^2} + \sqrt{\left(\frac{\mu}{\sigma^2} - \frac{1}{2}\right)^2 + \frac{2r}{\sigma^2}}, \tag{A.4}$$

$$V^* = \frac{\beta}{\beta - 1} F. \tag{A.5}$$

Equations (A.4) and (A.5) can be derived using dynamic programming and Itô's lemma.[6]

Incorporating Legal Risk Legal risk may increase the effective cost of exercising the investment option. This cost could reflect anticipated litigation, regulatory compliance expenses, or other legal frictions. Departing from the baseline assumption of a fixed investment cost F, I allow the cost to vary with a firm's legal risk exposure. Specifically, let the effective investment cost be:

$$F^*(\ell) = F + \lambda \ell, \tag{A.6}$$

where ℓ denotes firm-level legal risk and $\lambda > 0$ is the marginal legal cost per unit of exposure. For example, consider a firm planning to build a new production facility. While the construction cost is nominally fixed, high legal risk—such as exposure to environmental litigation or regulatory review—effectively inflates that cost by requiring additional legal

[6]The value function $X(V)$ satisfies the Hamilton–Jacobi–Bellman (HJB) equation:

$$rX(V) = \mathbb{E}\left[\frac{dX(V_t)}{dt}\right].$$

Applying Itô's lemma to a twice-differentiable function $X(V_t)$, where V_t follows geometric Brownian motion:

$$dX(V_t) = X'(V_t)\, dV_t + \frac{1}{2} X''(V_t)(dV_t)^2.$$

Substituting $dV_t = \mu V_t\, dt + \sigma V_t\, dZ_t$ and noting that $(dZ_t)^2 = dt$, we obtain:

$$\mathbb{E}[dX(V_t)] = \left(\mu V X'(V) + \frac{1}{2}\sigma^2 V^2 X''(V)\right) dt.$$

Thus, the HJB equation becomes:

$$rX(V) = \mu V X'(V) + \frac{1}{2}\sigma^2 V^2 X''(V).$$

Guessing a solution of the form $X(V) = AV^\beta$ and substituting into this equation leads to the quadratic characteristic equation for β. The economically meaningful root (i.e., $\beta > 1$) yields the optimal threshold in Equation (A.5).

counsel, permitting delays, insurance premiums, or contingency reserves. In this sense, legal risk operates like a surcharge: the firm must invest not only in bricks and mortar, but also in navigating the legal landscape that surrounds the project.

The new investment threshold becomes:

$$V^*(\ell) = \frac{\beta}{\beta - 1}(F + \lambda \ell), \tag{A.7}$$

which increases linearly in legal risk. Thus, legal uncertainty raises the break-even point required to justify investment and makes investment less likely, even if the expected project value is unchanged.

G.2 Dixit–Pindyck Framework with Legal Risk as Uncertainty

I now consider a related model of irreversible investment under uncertainty, following Dixit and Pindyck (1994). The key difference is that investment yields a continuous stream of payoff flows rather than a one-time gain.

Assume a firm can invest at cost F to receive perpetual payoff flow $\pi(V_t) = aV_t$, where $a > 0$ and V_t evolves as in Equation (A.3). The present value of investing at time t is:

$$\Pi(V) = \frac{aV}{r - \mu}, \tag{A.8}$$

assuming $r > \mu$.[7]

Let $X(V)$ denote the value of waiting to invest. The firm invests when $V_t \geq V^*$, and otherwise waits. The option value for $V < V^*$ is:

$$X(V) = AV^{\beta}, \tag{A.9}$$

where β is given by Equation (A.4) and arises from solving the associated HJB equation.

[7]The equation (A.8) follows from taking the expected present value of the perpetual cash flow stream $\pi(V_t) = aV_t$, where V_t follows geometric Brownian motion. Since $\mathbb{E}[V_t] = Ve^{\mu t}$, we have:

$$\Pi(V) = \mathbb{E}_0\left[\int_0^\infty e^{-rt} aV_t\, dt\right] = a\int_0^\infty e^{-rt}\mathbb{E}[V_t]\, dt = aV\int_0^\infty e^{-(r-\mu)t}\, dt = \frac{aV}{r-\mu},$$

which converges only if $r > \mu$.

Using standard value matching and smooth pasting conditions,[8] the investment trigger becomes:

$$V^* = \frac{\beta}{\beta - 1} \cdot \frac{F(r - \mu)}{a}. \tag{A.10}$$

Incorporating Legal Risk Legal risk increases uncertainty about future cash flows, which I model as stochastic volatility:

$$\sigma(\ell) = \sigma_0 + \kappa\ell, \quad \kappa > 0. \tag{A.11}$$

This raises the option value of waiting and leads to a higher threshold:

$$V^*(\ell) = \frac{\beta(\ell)}{\beta(\ell) - 1} \cdot \frac{F(r - \mu)}{a}, \tag{A.12}$$

with

$$\beta(\ell) = \frac{1}{2} - \frac{\mu}{\sigma(\ell)^2} + \sqrt{\left(\frac{\mu}{\sigma(\ell)^2} - \frac{1}{2}\right)^2 + \frac{2r}{\sigma(\ell)^2}}. \tag{A.13}$$

As legal risk rises, the volatility of future payoffs increases, the firm becomes more selective, and investment is delayed despite positive expected returns. This aligns with the empirical finding that firms with greater legal exposure tend to invest less in physical and intangible assets.

[8] The value matching condition requires that the value of waiting and the value of investing coincide at the investment threshold V^*:

$$X(V^*) = \Pi(V^*) - F.$$

This ensures there is no value lost or gained from exercising the option at the optimal point. Smooth pasting requires that the derivatives also match at V^*:

$$X'(V^*) = \Pi'(V^*).$$

This condition ensures optimality by preventing kinks in the value function—i.e., if the slopes did not match, the firm could increase its value by slightly delaying or advancing the investment decision. Together, these conditions pin down the unique investment threshold V^* where the firm is indifferent between waiting and investing.

References for Online Appendix

Baker, S. R., Bloom, N. and Davis, S. J.: 2016, Measuring Economic Policy Uncertainty, *Quarterly Journal of Economics* **131**(4), 1593–1636.

Birru, J.: 2018, Day of the week and the cross-section of returns, *Journal of Financial Economics* **130**(1), 182–214.

Chen, A. Y. and Dim, C.: 2024, High-Throughput Asset Pricing, *Working Paper* .

Chen, A. Y. and Zimmermann, T.: 2022, Open source cross-sectional asset pricing, *Critical Finance Review* **27**(2), 207–264.

Chordia, T., Subrahmanyam, A. and Tong, Q.: 2014, Have capital market anomalies attenuated in the recent era of high liquidity and trading activity?, *Journal of Accounting and Economics* **58**(1), 41–58.

Daniel, K., Hirshleifer, D. and Sun, L.: 2020, Short- and Long-Horizon Behavioral Factors, *Review of Financial Studies* **33**(4), 1673–1736.

Daniel, K. and Moskowitz, T. J.: 2016, Momentum crashes, *Journal of Financial Economics* **122**(2), 221–247.

Dixit, A. K. and Pindyck, R. S.: 1994, *Investment Under Uncertainty*, Princeton University Press, Princeton, NJ.

Florackis, C., Louca, C., Michaely, R. and Weber, M.: 2023, Cybersecurity Risk, *Review of Financial Studies* **36**(1), 351–407.

Harvey, C. R., Liu, Y. and Zhu, H.: 2016, ... and the Cross-Section of Expected Returns, *Review of Financial Studies* **29**(1), 5–68.

Heese, J. and Perez-Vacazos, G.: 2021, The effect of retaliation costs on employee whistleblowing, *Journal of Accounting and Economics* **71**(2-3), 101385.

Houston, J. F., Lin, C. and Xie, W.: 2018, Shareholder Protection and the Cost of Capital, *Journal of Law and Economics* **61**(4), 677–710.

Huang, D., Jiang, F., Tu, J. and Zhou, G.: 2015, Investor Sentiment Aligned: A Powerful Predictor of Stock Returns, *Review of Financial Studies* **28**(3), 791–837.

Huang, D., Li, J. and Wang, L.: 2021, Are disagreements agreeable? Evidence from information aggregation, *Journal of Financial Economics* **141**(1), 83–101.

Jamilov, R., Rey, H. and Tahoun, A.: 2025, The Anatomy of Cyber Risk, *Working Paper* .

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

Jurado, K., Ludvigson, S. C. and Ng, S.: 2015, Measuring Uncertainty, *American Economic Review* **105**(3), 1177–1216.

Leone, A. J., Li, E. X. and Liu, M.: 2021, On the SEC's 2010 enforcement cooperation program, *Journal of Accounting and Economics* **71**(1), 101355.

McDonald, R. and Siegel, D.: 1986, The Value of Waiting to Invest, *Quarterly Journal of Economics* **101**(4), 707–728.

Novy-Marx, R. and Velikov, M.: 2025, AI-Powered (Finance) Scholarship, *Working Paper* .

Pastor, L., Stambaugh, R. F. and Taylor, L. A.: 2022, Dissecting green returns, *Journal of Financial Economics* **146**(2), 403–424.

Pedersen, L. H., Fitzgibbons, S. and Pomorski, L.: 2021, Responsible investing: The ESG-efficient frontier, *Journal of Financial Economics* **142**(2), 572–597.

Stambaugh, R. F.: 2017, Mispricing Factors, *Review of Financial Studies* **30**(4), 1270–1315.

Stambaugh, R. F., Yu, J. and Yuan, Y.: 2012, The short of it: Investor sentiment and anomalies, *Journal of Financial Economics* **104**(2), 288–302.