The Pricing and Economic Impact of Legal Risk *

Dean Ryu†

June, 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 using earnings call transcripts. Firms exposed to class action lawsuits, regulatory scrutiny related to climate issues and banking, and cybersecurity threats tend to exhibit higher risk scores. Legal risk has real and financial consequences: a one standard deviation increase predicts a 3–8% decline in future investment and reduces firms' reliance on debt when facing financing needs, consistent with a precautionary motive. A long-short portfolio that buys high-risk firms and sells zero-risk firms earns an annualized return of 4.7% (t-stat: 3.6), increasing to 7.1% (t-stat: 4.9) post-2010 when regulatory oversight intensified, and remains robust to factor zoo concerns. 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, Thummim Cho, Robin Greenwood, Umit Gurun, Tarek Hassan, Jonas Heese, Rustam Jamilov, Mathias Kruttli, Max Miller, Ji Yeol Jimmy Oh, Andrea Pagliuca (discussant), Ludovic Phalippou, David Schoenherr, Markus Schwedeler, Vitalii Tubdenov, Jaesun Yun (discussant), and seminar participants at the Allied Korean Finance Associations Joint Conference, FMA 2024 New Ideas Session, FMA 2025 European Doctoral Consortium, International Accounting & Finance Doctoral Symposium, LAPE Spring 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

> *"[...] [U]nder the Rule of Law, the government is prevented from stultifying individual efforts by ad hoc action. Within the known rules of the game, the individual is free to pursue his personal ends and desires, certain that the powers of government will not be used deliberately to frustrate his efforts."*
> – Friedrich A. Hayek, *The Road to Serfdom* (1944, p. 76)

A well-functioning economy relies on the rule of law. Legal and regulatory institutions such as courts, government agencies, and financial regulators enable economic activity by enforcing contracts and protecting property rights, thereby fostering investment and facilitating transactions. Yet these same institutions also introduce legal risk. As legal systems expand and regulatory oversight intensifies, firms face increasing complexity in compliance and heightened exposure to litigation. Legal risk, in this sense, is not merely incidental but a fundamental feature of modern markets.[1]

[INSERT FIGURE 1 HERE]

In the U.S., legal risk has become increasingly salient due to the growing complexity of legal proceedings, regulatory scrutiny, and contractual disputes. Litigation-prone industries[2] have expanded substantially in size and market capitalization since the early 2000s, as shown in Figure 1. This trend is mirrored by a surge in securities class action lawsuits, which have more than doubled since the 1990s despite a shrinking number of publicly listed firms,[3] and by a tripling of patent disputes since 2005 (Cohen et al. 2019). Regulatory enforcement actions have also accelerated in the post–Global Financial Crisis era. As a result, firms face mounting legal exposure with growing consequences for valuation and real decisions.

To examine these dynamics, this paper studies how legal risk affects publicly traded firms through two key channels: market valuation and firm behavior. Legal risk is embedded

[1] A notable counterexample is Venezuela under President Hugo Chávez in the late 1990s and early 2000s, where the shift from rule *of* law to rule *by* law eroded judicial independence and enabled widespread expropriation. The Chávez administration bypassed legal norms, nationalized key industries without consistent compensation, and concentrated power in the executive. This breakdown in institutional checks heightened investor uncertainty, triggered capital flight, and collapsed private investment—illustrating how the absence of credible legal frameworks can destabilize markets.

[2] These include firms in the pharmaceutical and biotechnology sectors (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).

[3] See https://corpgov.law.harvard.edu/2021/03/11/recent-trends-in-securities-class-action-litigation/.

in prices as investors update expectations about cash flows, litigation outcomes, and regulatory exposure. At the same time, firms respond by adjusting investment, governance, and financing choices.[4] These channels motivate the core questions of this paper: To what extent is legal risk systematically priced by investors, and how does it shape real economic decisions?

To address these questions, I construct a firm-level measure of legal risk using textual analysis of earnings call transcripts. While prior work emphasizes the macroeconomic role of institutions (Acemoglu et al. 2001, Acemoglu and Johnson 2005), little is known about how legal risk is perceived and managed at the firm level. My approach captures firms' attention to legal issues as expressed in real-time communication with investors, offering a dynamic lens into evolving legal exposure. The resulting measure, **LRISK**, tracks concerns over class actions, climate regulation, cybersecurity threats, and financial oversight (see Sections 3.2 and 3.3). Firms with elevated legal risk tend to reduce future investment and shy away from debt financing when externally constrained—consistent with legal uncertainty increasing the expected cost of financial distress.

In addition to firm-level analysis, I 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).

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

Taken together, this paper highlights the growing importance of legal risk in both finance and economics. While legal and regulatory institutions underpin market stability, they

[4]See Online Appendix A for case evidence on how legal risk materialization impacts firm valuations.

also introduce constraints 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.

A second strand in comparative economics explores how differences in legal origins and institutional quality shape financial development (e.g., La Porta et al. 1998; Glaeser and Shleifer 2002; Rajan and Zingales 2003). These studies emphasize how judicial efficiency and regulatory enforcement affect corporate behavior and market outcomes. While this paper focuses on the U.S., where legal enforcement is relatively strong, preliminary cross-country evidence suggests that the pricing of legal risk varies by legal origin—being more prominent in common law systems.

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.

This paper also contributes 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 5.1, 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 describes the construction of the firm-level legal risk measure using textual analysis of earnings call transcripts. Section 3 examines cross-sectional variation in legal risk across firms and industries, supported by detailed case studies. Section 4 presents the asset pricing results, demonstrating that legal risk is a priced factor in the cross-section of stock returns, and addresses potential concerns related to data mining and the broader "factor zoo" critique. This section also discusses potential pricing mechanism. Section 5 evaluates the robustness and distinctiveness of the legal risk factor relative to other text-based measures, rules out behavioral explanations, and considers variation across legal origins. Section 6 explores the real effects of legal risk on corporate investment and financing decisions. Section 7 concludes.

2 Data and Methods

2.1 Extracting Legal Risk from Corporate Earnings Conference Calls

[INSERT TABLE 1 HERE]

I construct a text-based measure of ex-ante legal risk at the firm-level using quarterly earnings conference calls transcripts of U.S. firms.[5] A key advantage of this approach is that it captures forward-looking legal concerns, rather than ex-post litigation outcomes. In that

[5]Typically, firms hold quarterly earnings calls, each aligned with the release of their financial results. These calls begin with a structured management presentation, where senior executives provide an overview of the company's recent performance, financial health, and key operational developments. Following the presentation, the call opens to a Q&A session, during which analysts and, occasionally, investors engage management directly, probing into areas such as strategic initiatives, market challenges, and specific financial metrics. This interactive portion of the call is critical, as it allows stakeholders to gauge management's preparedness, transparency, and responsiveness to both expected and unforeseen circumstances.

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). Table 1 highlights instances where firms explicitly discuss legal risks in earnings calls, demonstrating how legal uncertainty is a relevant dimension of firm valuation.

For instance, in an October 2024 earnings call, Alphabet Inc. (Google) 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 statement refers to the U.S. Department of Justice's (DOJ) high-profile antitrust lawsuit against Google, which alleges that the company has abused its dominant position in online search by securing exclusive agreements—such as the Apple ISA (Information Services Agreement)—to pre-install its search engine on mobile devices. The potential voiding of these contracts could significantly disrupt Google's traffic acquisition strategy and revenue model, particularly its lucrative mobile search business. The legal uncertainty surrounding the case has not only weighed on Google's stock price but also influenced broader market expectations about future regulatory interventions in the technology sector.

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.[6] 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.

A key aspect of this analysis is the construction of a training library of legal bigrams. I initiate this process using the Loughran-McDonald (LM) Dictionary (Loughran and McDonald 2011), particularly its "litigious (Fin-Lit)" category, which encompasses legal jargon and terms related to litigation. The 2023 edition of the LM Dictionary contains 86,553 keywords, 903 of which fall under the "litigious" category. However, rather than relying on unigrams of legal terms as listed in the LM dictionary, the current paper's approach focuses on developing a dynamic bigram (or trigram) library tailored to context-specific

[6]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).

legal terminology to minimize misclassification and reduce measurement error.[7]

[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 Stock Market and Accounting Data

To examine the financial and real effects of legal risk, I match the firm-level legal risk exposure—constructed from quarterly earnings call transcripts—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

[7] Refer to Section 5.1 for a comparison highlighting how my approach differs from that of Loughran and McDonald (2011).

[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 1 illustrates the rising number of litigation-prone industries around the turn of the century, underscoring the growing relevance of legal risk since 2000.

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 Firm-level Variation in Legal Risk

This section examines firm-level variation in legal risk exposure, as constructed in the previous section. Through detailed case studies, it explores how well-known firms and industries respond to emerging legal risks, demonstrating that the legal risk exposure measure introduced in this paper effectively captures firms' concerns about legal uncertainty and helps assess both its internal consistency and external validity.

3.1 Legal Risk: Descriptive Statistics

[INSERT FIGURE 2 HERE]

Figure 2 demonstrates that the sample of earnings call-releasing firms analyzed in this study is representative of the overall U.S. stock market. The top panel shows the number of firms releasing earnings calls from 2002 to 2024. The sample consists of standard U.S. stocks listed on the three major exchanges—NYSE, NASDAQ, and NYSE American (formerly AMEX)—and includes only ordinary shares, identified by CRSP share codes 10 or 11. Around 55% of firms began publishing earnings calls in 2002, and this percentage steadily increased, surpassing 80% by 2010. By the late 2010s, the number had reached 85%, though there was a slight decline following the onset of COVID-19. A similar pattern is observed in how these firms represent the market. The middle panel shows that, after 2010, these firms account for approximately 85% of the entire market capitalization. The bottom panel of the figure presents the ratio of the top 100, 500, and 1500 firms by market capitalization at the end of each June. The data indicates that nearly all of the top 100 and 500 firms consistently publish earnings calls each year, suggesting that the sample considered in this paper is highly representative of the U.S. stock market.

[INSERT FIGURE 3 HERE]

Using earnings call transcript data, Figure 3 illustrates the changes in legal risk exposure for prominent corporate entities, most of which are constituents of the well-known S&P 500 index. In each firm's figure, the y-axis categorizes legal risk exposure into three levels: 1 (no legal risk), 2 (moderate legal risk), and 3 (high legal risk), as defined in the following section.

The figure reveals that firms frequently transition between different levels of legal risk exposure, highlighting that legal risk is dynamic and not confined to a specific subset of firms. For instance, Ford Motor Company faced multiple lawsuits in the mid-2000s, including Employment Discrimination related to Apprenticeship Testing (2004–2005), discrimination in loan rates (2005), and Product Liability cases involving sudden acceleration and defective vehicle components (2006). During this period, legal risk exposure, as indicated in the 2006 earnings call transcripts, escalated from a moderate level (2) to the highest level (3). Similarly, Coca-Cola experienced heightened mentions of legal risk during 2020 and 2021, coinciding with significant legal challenges. These included Environmental Greenwashing Allegations (*Earth Island Institute v. Coca-Cola Co.*) where the company was accused of overstating its sustainability efforts while contributing substantially to plastic pollution. Additionally, Coca-Cola faced the 100% Recyclable Labeling Lawsuit (*Swartz v. Coca-Cola Co.*), which challenged the company's claims about the recyclability of its plastic bottles.

[INSERT FIGURE 4 HERE]

I next explore the industry distribution of legal risk measure. Figure 4 illustrates the distribution of firms across three legal risk categories—No Risk, Medium Risk, and High Risk—represented by green, yellow, and red bars, respectively, conditional on their industry. The industry definitions follow the Fama-French 17 industry classification. On average, firms' mentions of legal risk appear balanced across industries, as indicated by the relatively even distribution of firm-month observations on the y-axis. However, certain industry categories exhibit higher levels of legal risk, as documented in the transcripts of earnings calls, compared to others.

For instance, the utilities sector (denoted as "Utils (14)") shows a significant concentration of firm-month observations in the high-risk portfolio. This concentration aligns with the unique characteristics of the utilities sector, including the nature of its services, the regulatory framework within which it operates, and its distinct business models. Similarly, the finance and accounting literature identifies several sub-industries with elevated exposure

to legal risk: pharmaceuticals and biotech products (SIC 2833–2836; industry category 7), computer and office equipment manufacturing (SIC 3570–3577; industry category 11), electronic equipment and components (SIC 3600–3674; industry category 11), and retail industries (SIC 5200–5962, primarily categorized under industry label 15, with others under label 8).[9] The figure highlights that these industry categories tend to exhibit higher levels of legal risk exposure, as evidenced by the prevalence of yellow (Medium Risk) and red (High Risk) bars compared to green (No Risk). Overall, the industry decomposition presented in Figure 4 confirms that the portfolio analysis is not disproportionately influenced by any single industry or a small subset of industries. Instead, it provides a comprehensive view of the legal risks faced by firms across various industries.

3.2 Case Studies: Are Brown Firms More Exposed to Legal Risk?

[INSERT FIGURE 5 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.[10] 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 5 illustrates firm-level legal risk exposure 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

[9] Computer programming and IT services (SIC 7370–7374), although known for legal risk, are omitted from the analysis as they are classified under the "Others" category in the Fama-French 17 industry system.

[10] 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.

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

Overall, the spikes in legal risk exposure in Figure 5 align with major legal disputes, validating the effectiveness of the firm-level legal risk measure in capturing periods of heightened regulatory and litigation risk.

3.3 Case Studies: Are Banks More Exposed to Legal Risk?

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 6 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 mis-

leading 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 6 (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.

4 The Pricing Impact of Legal Risk

4.1 Portfolio Construction and Asset Pricing Results

I assess the economic value 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 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.[11]

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. 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 (though unreported results show that the pricing effect remains robust even when utilities and financials are included).

[INSERT TABLE 3 HERE]

Table 3 presents the results. In the baseline analysis, average excess returns increase from 0.638% to 1.035% per month between no and high legal risk stocks in value-weighted portfolios. The corresponding 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.[12] This strong cross-sectional relationship between legal risk and future stock returns is notable, especially considering that existing literature has not documented such a connection. 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 from July 2004 to June 2024. 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).

Next, I examine whether existing asset pricing models account for the legal risk factor. The

[11] 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. I focus on larger firms for consistency with the broader asset pricing literature, and because legal risk is more consequential for these firms, both in exposure and economic impact.

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

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.

4.2 Addressing the Factor Zoo Critique

Although this paper does not primarily aim to introduce a new risk factor, it inevitably engages with the broader "factor zoo" critique (Cochrane 2011). 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). In addition, the **LAW** factor is supported by strong statistical evidence: its full-sample t-statistic of 3.66 (2004–2024), and 4.91 in the post-2010 subsample (see the next section), exceed the 3.0 threshold recommended by Harvey et al. (2016) for identifying robust asset pricing anomalies.

Second, 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.1). Figure A.2 in the Online Appendix further shows that the **LAW** factor is largely uncorrelated with 193 known anomalies, with no absolute Pearson correlation exceeding 0.4, highlighting its distinctiveness in the factor zoo.

Third, the text-based construction of the **LAW** factor offers a unique advantage: it enables decomposition of legal risk into distinct thematic categories based on keyword clusters. As detailed in Table 2, legal content is classified into three subcomponents (**LRISK_COURT**, **LRISK_TRIAL**, and **LRISK_TERMS**), which are then aggregated to form the composite **LAW** factor. To assess the economic relevance of each category, I construct long-short portfolios separately for each subcomponent.[13] The results show that both macro-level legal risk (**LAW_COURT**) and micro-level trial-related exposure (**LAW_TRIAL**) are individually priced in the cross-section of returns. This flexible structure underscores the strength of a text-based approach in uncovering economically meaningful components of legal risk.

4.3 Time-variation of Aggregate Legal Risk

[INSERT FIGURE 7 HERE]

How does legal risk change over time? This section tracks fluctuations in aggregate legal risk and discusses their broader implications. The measure is constructed each quarter by classifying firms into two groups: zero firms, which did not mention legal risk in any of the past four quarterly earnings call transcripts, and non-zero firms, which did. The coverage ratio, defined as the proportion of non-zero legal risk firms relative to all firms, captures the extensive margin of legal risk in the corporate sector.

The aggregate measure of legal risk, presented in Figure 7, exhibits notable time variation that aligns with major regulatory developments. Between 2003 and 2007, aggregate

[13] For instance, I construct a **LAW_COURT** factor by going long on firms with high legal risk based solely on keywords in **LRISK_COURT**, and short on firms with no legal risk. The L1 group (firms with no legal risk) is held constant across both the **LAW** and **LAW_COURT** factor constructions, allowing for a cleaner comparison of pricing power that stems entirely from differences in the high-risk group (L3).

legal risk trended upward, as reflected in the standardized coverage ratio. This period was characterized by both regulatory tightening and deregulatory pressures. On balance, however, legal risk awareness increased.[14]

Following the Global Financial Crisis (GFC), the legal landscape shifted decisively toward greater regulatory scrutiny. Key reforms—including the Dodd-Frank Act's whistleblower provisions and the SEC's intensified enforcement activities—increased firms' exposure to legal oversight. As a result, aggregate legal risk remained elevated throughout the immediate post-crisis years.

By 2012–2013, however, the measure briefly subsided—possibly reflecting regulatory adjustment. However, this dip proved short-lived. From 2018 onward, legal risk began to rise again, reflecting structural developments such as the Supreme Court's *Cyan* (*2018*) ruling, which facilitated forum shopping in securities class actions, and the growing legal salience of cybersecurity threats. As Jamilov et al. (2025) note, evolving SEC disclosure rules and a wave of high-profile cyber incidents after 2015 have also contributed to heightened legal exposure—particularly in data-sensitive sectors. The COVID-19 shock further amplified aggregate legal risk. A multiple structural change procedure of Bai and Perron (2003) identify four statistically significant breakpoints: December 2007, March 2013, June 2017, and June 2020, which closely aligns with the major regulatory, legal, and macroeconomic shifts described above. In Online Appendix D, I show that the aggregate measure of legal risk developed here does not coincide with well-known uncertainty measures such as Economic Policy Uncertainty (Baker et al. 2016) and Macroeconomic Uncertainty (Jurado et al. 2015), highlighting that LRISK represents a unique source of variance.

[INSERT TABLE 4 HERE]

Table 4 reports the portfolio results for the 2010–2024 subsample, a period marked by post-GFC reforms that elevated aggregate legal risk and intensified regulatory scrutiny across firms. Reflecting this shift, the long-short portfolio strategy yields a higher annualized return of 7.1% (t-stat: 4.91), driven by the strong performance of firms with high legal risk. This result highlights the evolving legal environment in the U.S. following the financial crisis. Notably, the enhanced performance of the strategy is largely attributable to the

[14]On one hand, the Sarbanes-Oxley Act (SOX) of 2002 introduced stricter compliance requirements, heightened board accountability, and imposed heavier penalties for misconduct, encouraging more proactive legal disclosures. On the other hand, deregulatory measures such as the Gramm-Leach-Bliley Act (GLBA) of 1999 and the adoption of Universal Demand (UD) laws reduced firms' exposure to shareholder litigation. Nonetheless, the post-SOX compliance shift appears to have had the stronger effect.

short-leg of the portfolio. Firms with no disclosed legal risk (L1) exhibit significantly negative alphas, suggesting that investors are willing to accept lower returns from firms perceived to face minimal legal exposure. This pricing dynamic reinforces the idea that legal risk has become a salient dimension of cross-sectional return variation in the post-GFC era.

4.4 Channels of Systematic Legal Risk

While firm-specific legal events may appear idiosyncratic, they can exert systematic force through multiple (not mutually exclusive) channels. This section outlines three distinct but complementary mechanisms: economic spillovers, belief updating about aggregate legal uncertainty, and granularity. Together, these perspectives help explain why legal risk, though firm-specific in origin, may be priced in aggregate asset markets.

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 sense, 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.

5 Validation

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 meaningful differences between common law and civil law systems.

[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.

5.1 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.[16] 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.[17]

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

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

[INSERT TABLE 5 HERE]

How the legal risk measure relates to the Political Risk 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 Table 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, 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 6 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 6. 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 4, and find that both **LRISK** (replicating the results from Table 3) 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.

5.2 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 B.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.[18] However, my empirical results in Table B.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.

5.3 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

[18] 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.

6 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.[19] 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

[19] 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 7 HERE]

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

$$\text{INV}_{i,t+1} = \alpha + \beta\ \text{LRISK}_{i,t} + \gamma\ \mathbf{X}_{i,t} + \theta_i + \eta_t + \delta_{s(i)} + \varepsilon_{i,t+1} \tag{2}$$

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) as well as industry fixed effects $(\delta_{s(i)})$ where $s(i)$ maps firm i to its corresponding two-digit SIC industry. This procedure ensures that the results are not attributed to time-invariant firm characteristics, common macroeconomic shocks, or persistent industry-specific factors. Also, independent variables are lagged one year to limit concerns of reverse causality.

For the dependent variable, INV, I use four different measures of investment, each scaled by beginning-of-year total assets. CAPX + AQC captures tangible investment, including capital expenditures and acquisitions, reflecting the firm's spending on physical assets and M&A activity. R&D represents intangible investment, measuring expenditures related to innovation, product development, and technological advancement. To capture both tangible and intangible investment in a single measure, I include Δ AT, defined as the annual change in total assets, and total investment, which aggregates capital expenditures, acquisitions, and R&D (See Table 7 for details). These broader measures provide a more comprehensive view of the firm's overall investment activity. For ease of interpretation, investment variables are expressed in percentage terms, and legal risk is standardized.

[INSERT TABLE 8 HERE]

Table 8 presents the results. Across four distinct measures of firm-level investment (including both tangible and intangible assets), all regressions reveal a consistent pattern:

higher legal risk significantly predicts lower future investment. For economic interpretation, consider column (2), where the dependent variable is tangible investment, defined as the sum of capital expenditures (CAPX) and acquisitions (AQC). A one standard deviation increase in legal risk is associated with an approximate 3% decline in tangible investment in the following year.[20]

The effect is even more pronounced when focusing on intangible investment. Column (8) reports results using R&D expenditures as the dependent variable, showing that a one standard deviation increase in legal risk is associated with an estimated 7.8% decline in R&D investment in the subsequent year.[21] Finally, when aggregating across all investment types—including CAPX, acquisitions, R&D, and the change in intangible assets—the results remain consistent. Column (11) shows that a one standard deviation increase in legal risk predicts a 7.0% reduction ($-1.233/17.75 \approx -6.94\%$) in total firm-level investment, underscoring the broader economic impact of legal uncertainty on corporate investment behavior. These empirical results are consistent with the notion that legal risk introduces uncertainty about the future operating environment, prompting firms to adopt more cautious investment behavior. Faced with elevated legal uncertainty, firms appear to scale back or delay investment to avoid committing capital under conditions of heightened risk: an effect theoretically grounded in models of irreversible investment under uncertainty (Bernanke 1983).

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:

$$\text{Finance Mode}_{i,t} = \alpha + \beta_1 \, \text{FD}_{i,t} + \beta_2 \, \text{LRISK}_{i,t} + \beta_3 \, (\text{FD}_{i,t} \times \text{LRISK}_{i,t}) + \gamma \, \mathbf{X}_{i,t} + \theta_i + \eta_t + \varepsilon_{i,t} \quad (3)$$

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 7 for variable definitions). The coefficient of interest is β_3, which captures the interaction between the financing deficit and legal risk (FD $\times$ LRISK),

[20] The sample mean of tangible investment (CAPX + AQC) is 8.45, implying an economic effect of $-0.267/8.45 \approx -3.16\%$.

[21] The sample mean of R&D investment (xrd) is 5.92, implying an economic effect of $-0.463/5.92 \approx -7.82\%$.

and tests whether legal uncertainty moderates the firm's reliance on debt financing when internal resources are insufficient.

[INSERT TABLE 9 HERE]

Panel A in Table 9 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 9 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.

This behavior 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.

7 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

Acemoglu, D. and Johnson, S.: 2005, Unbundling Institutions, *Journal of Political Economy* **113**(5), 949–995.

Acemoglu, D., Johnson, S. and Robinson, J. A.: 2001, The Colonial Origins of Comparative Development: An Empirical Investigation, *American Economic Review* **91**(5), 1369–1401.

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.

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.

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.

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. 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.

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, 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.

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

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.

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.

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.

Rajan, R. G. and Zingales, L.: 2003, The great reversals: the politics of financial development in the twentieth century, *Journal of Financial Economics* **69**(1), 5–50.

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.

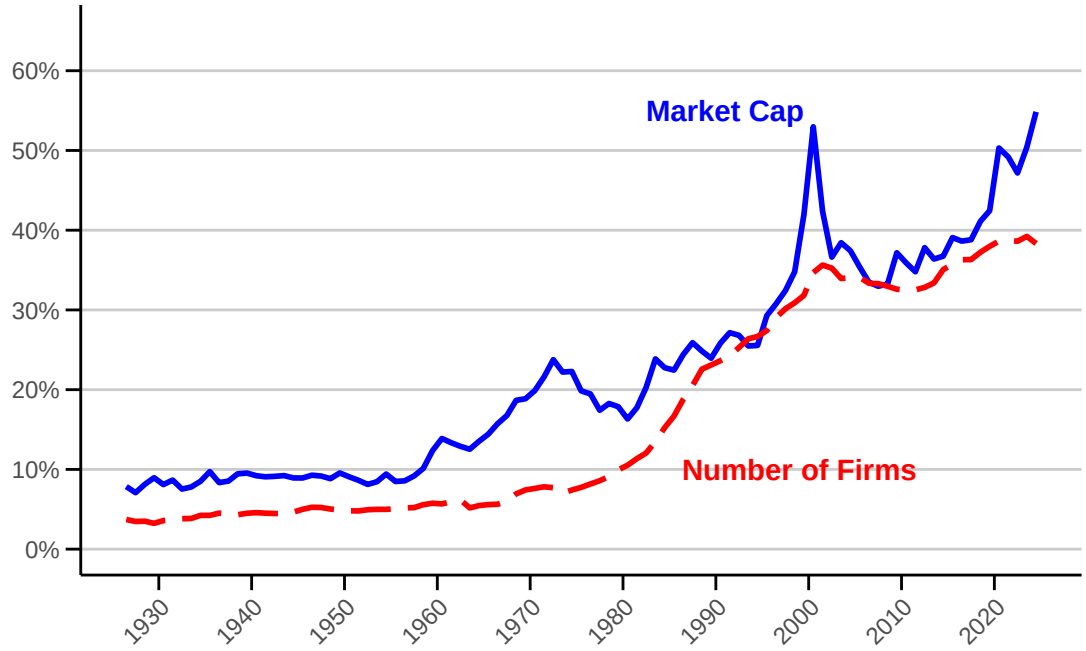
Stambaugh, R. F.: 2017, Mispricing Factors, *Review of Financial Studies* **30**(4), 1270–1315.

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.

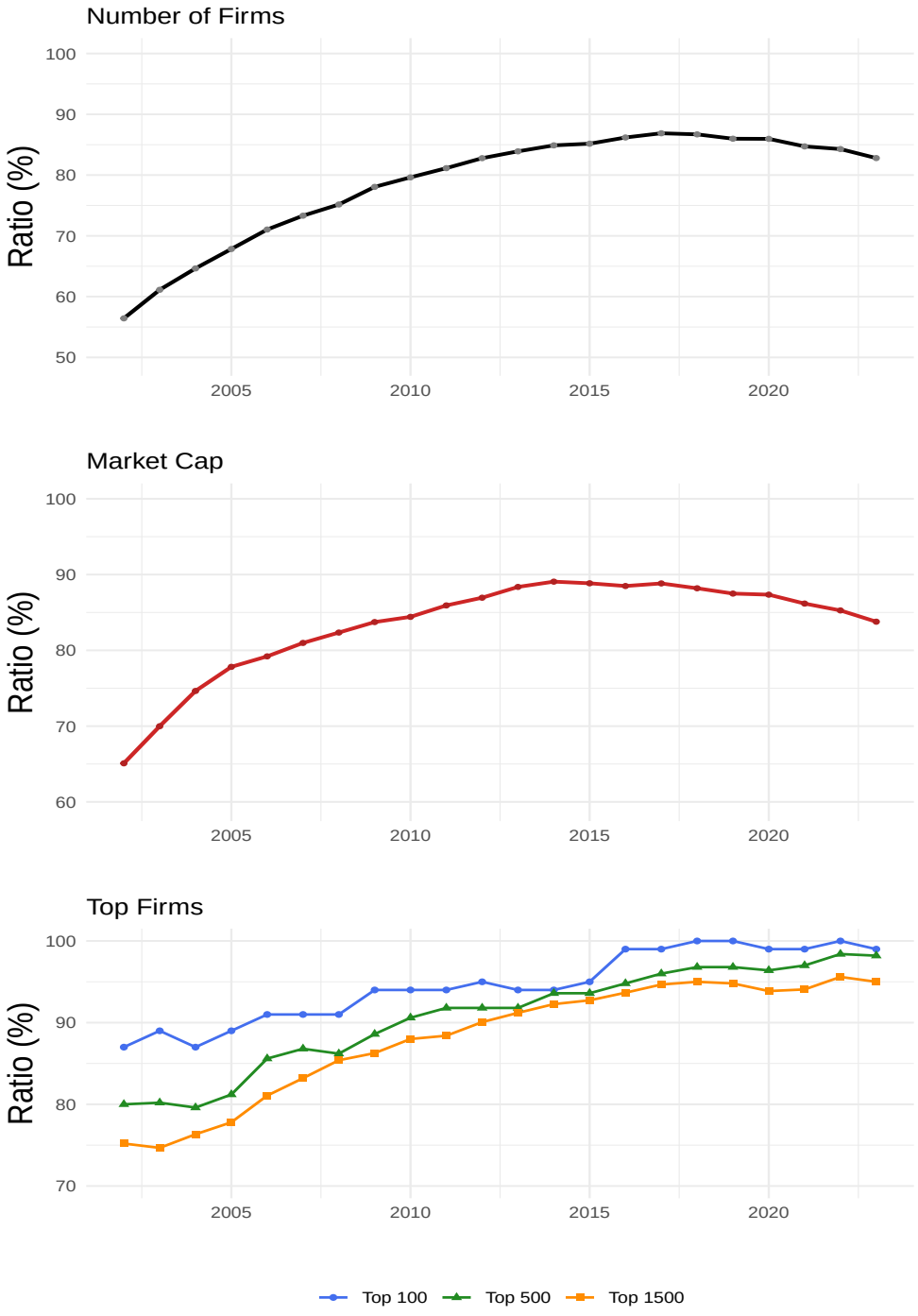
List of Figures

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



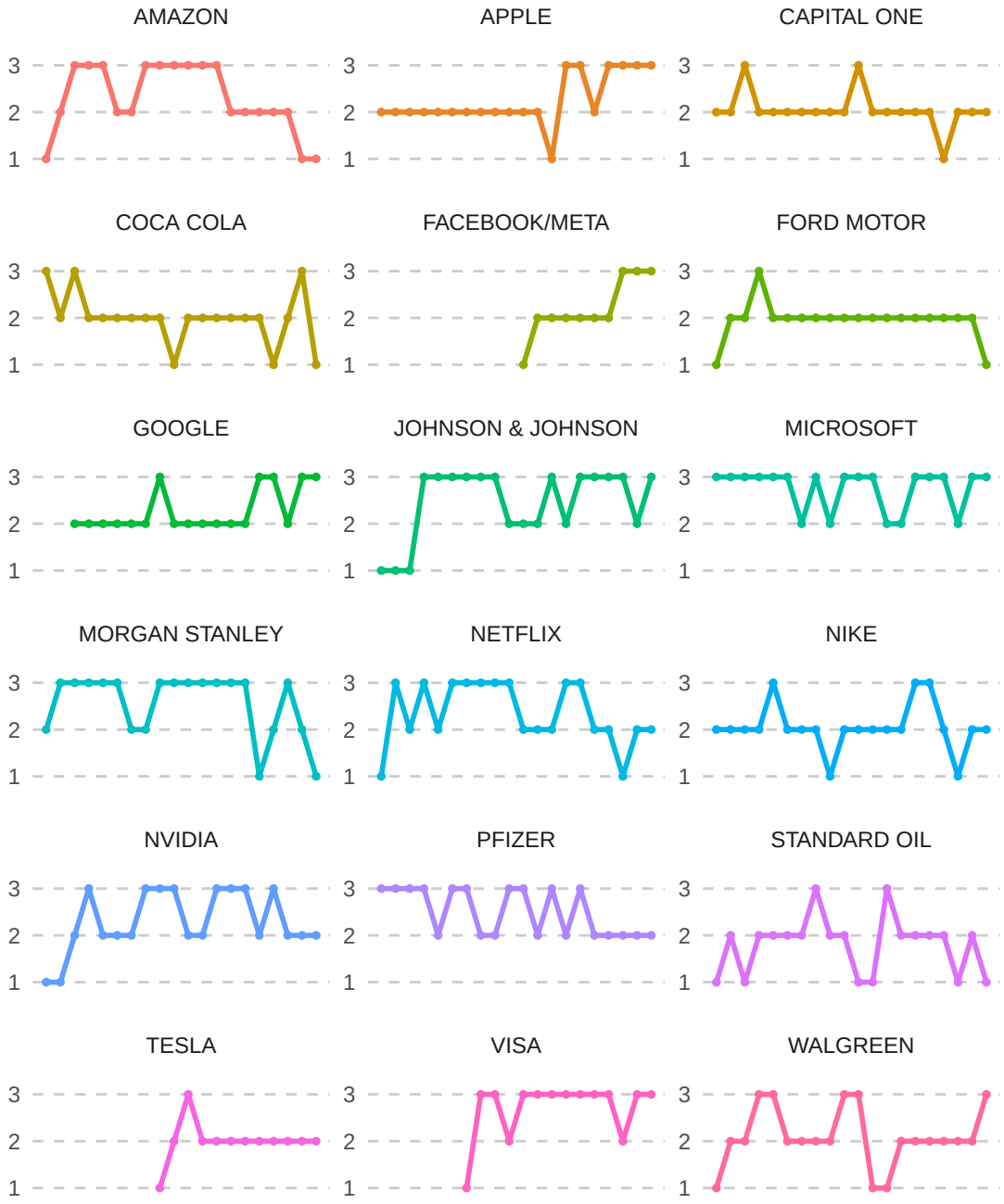
Note: This figure illustrates the number of firms (red dotted line) and the market capitalization ratio (blue solid line) of companies operating in high-litigation industries, relative to all publicly traded firms on the NYSE, NASDAQ, and NYSE American (formerly AMEX) from 1926 to 2024. The industries 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.

Figure 2: **Summary Statistic on U.S. Firms Publishing Earnings Calls Transcripts**



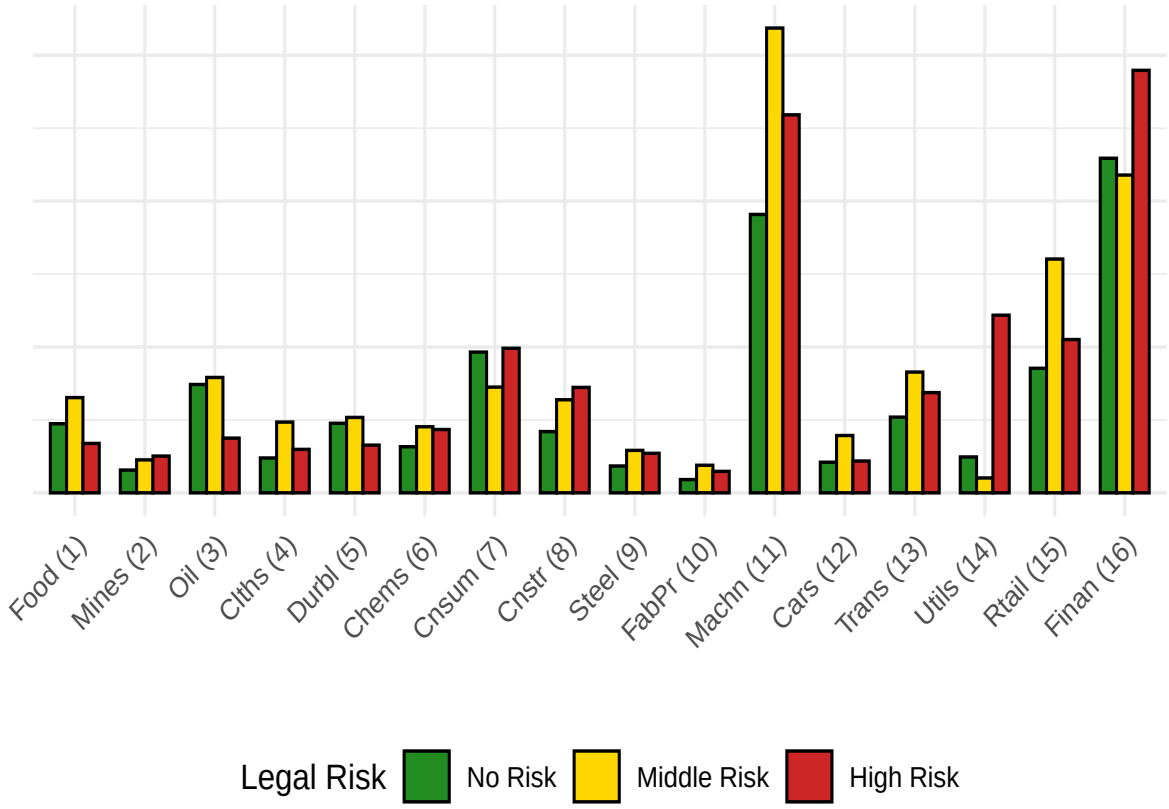
Note: The figures display the following: the proportion of firms publishing earnings call transcripts (top), the proportion of market capitalization of those firms that publish earnings calls (middle), and the proportion of the top 100, top 500, and top 1500 firms by market capitalization that publish earnings call transcripts (bottom). 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 3: **Legal Exposure of Well-known Firms**



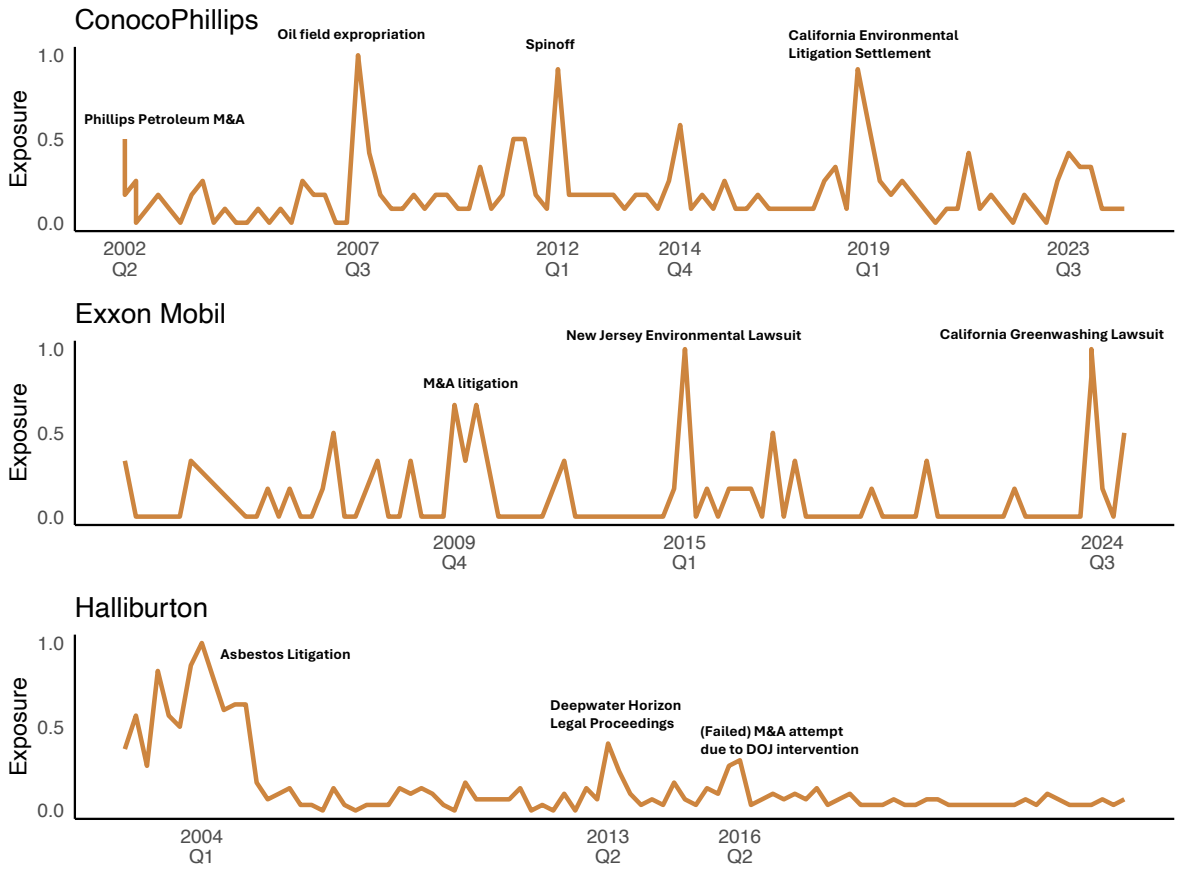
Note: This figure illustrates the legal risk exposure of prominent corporate entities analyzed in this study. The x-axis shows the years over which portfolio sorting based on firm-level legal risk exposure was conducted, spanning from 2003 to 2023. The y-axis categorizes the legal risk into three levels: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

Figure 4: **Industry Distribution of Legal Risk**



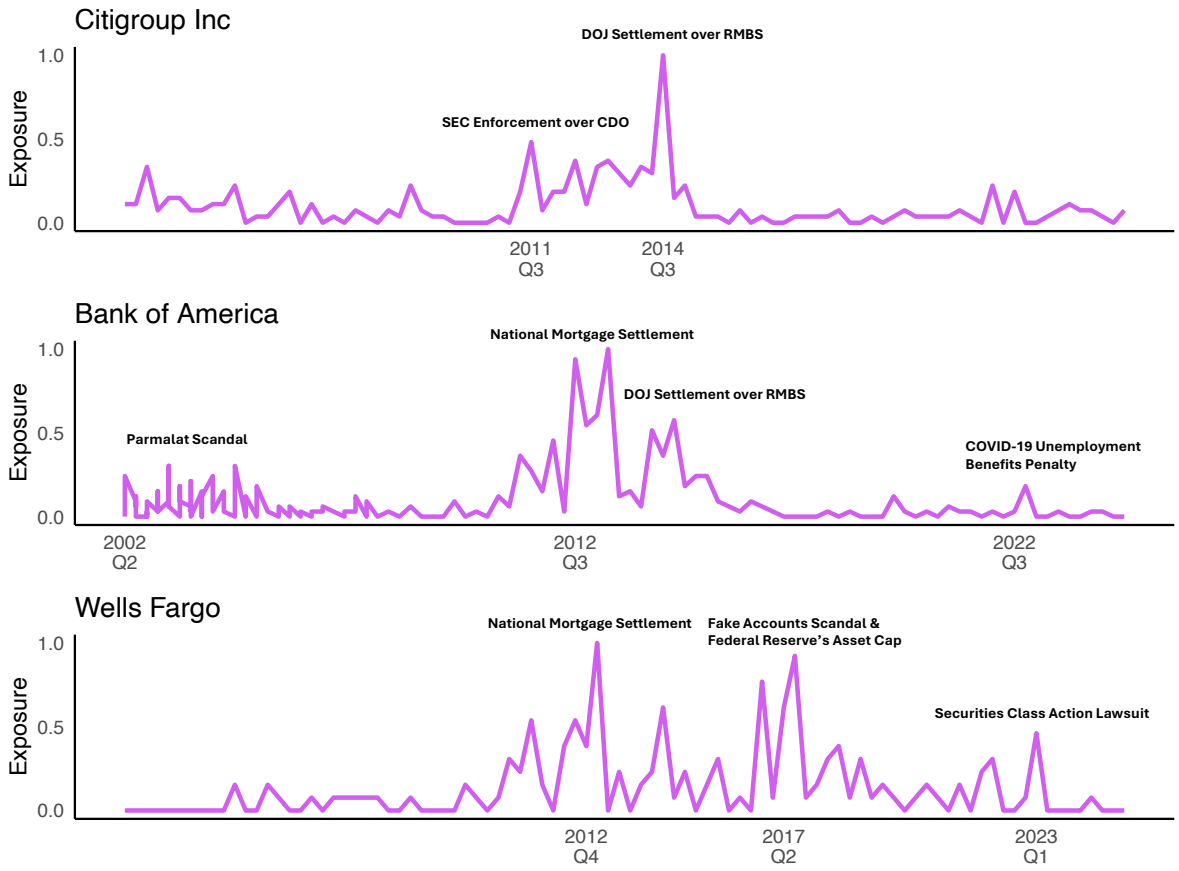
Note: The figures illustrate the industry contributions to the tercile portfolios described in this paper. The x-axis represents the Fama-French 17 industry categories: Food (FF industry category 1), Mines (2), Oil (3), Clothes (4), Consumer Durables (5), Chemicals (6), Consumer goods (7), Construction (8), Steel (9), Fabricated Products (10), Machinery and Business Equipment (11), Cars (12), Transportation (13), Utilities (14), Retails (15), Finance (16) and Others (omitted). The y-axis displays the frequency of firm-month observations categorized into one of the three portfolio sorts based on legal risk exposure: green (No legal risk), yellow (Medium legal risk), and red (High legal risk).

Figure 5: **Case Studies: Brown Firms**



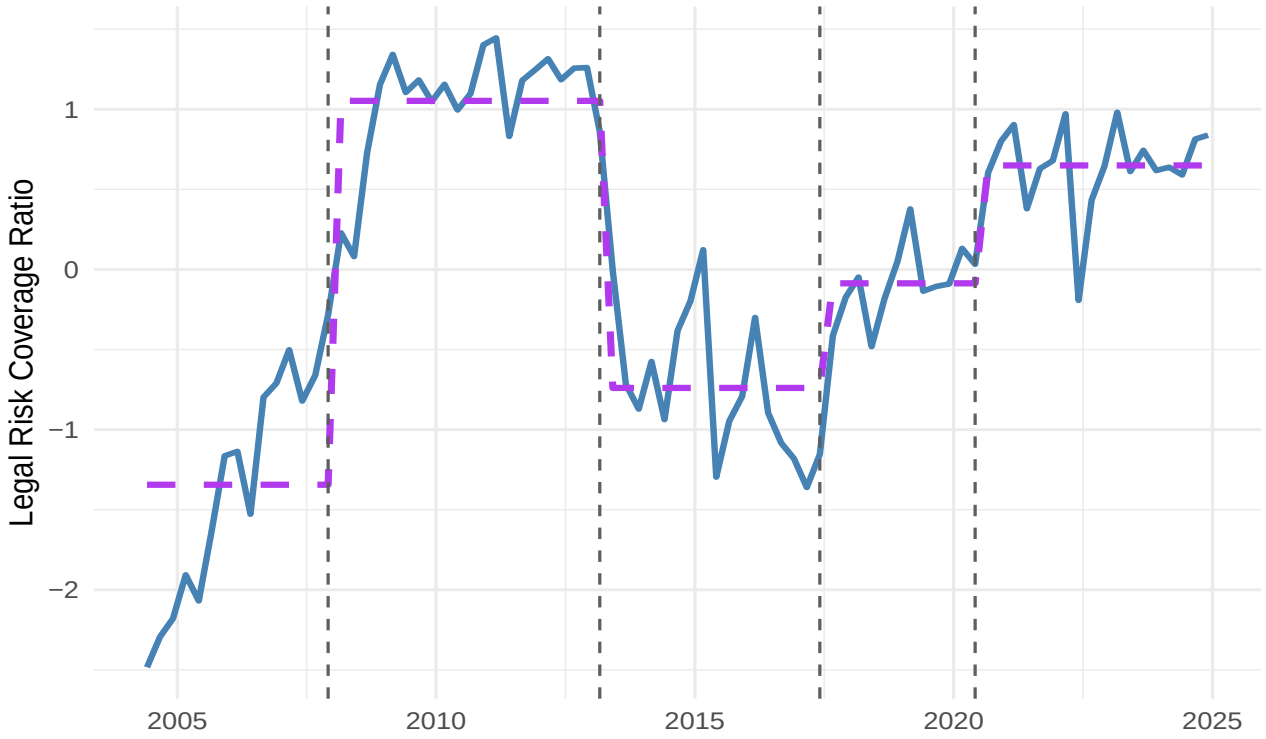
Note: This figure illustrates the firm-level legal risk exposure of three representative companies in the oil industry: ConocoPhillips, ExxonMobil, and Halliburton. All three are constituents of the S&P 500. The brown solid line in each figure represents the standardized firm-level legal risk exposure,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

Figure 6: **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 7: **Aggregate Measure of Legal Risk**



Note: This figure presents an aggregate measure of legal risk, captured by the coverage ratio of firms referencing legal risk. Each quarter, firms are classified into two groups: zero firms, defined as those that did not mention legal risk in their earnings call transcripts over the past four quarters, and non-zero firms, which did include such mentions. The coverage ratio is then calculated as the proportion of non-zero firms to the total number of firms. To facilitate comparison over time, this ratio is standardized using a z-score transformation, highlighting fluctuations relative to the historical average. The time series is further segmented using the Bai and Perron (2003) multiple structural break estimation procedure. Segment means between these breaks are plotted to visualize regime shifts in the underlying trend of legal risk coverage over time.

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: **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. Investment is defined as CAPX / PPENT; Asset tangibility is defined as PPENT / AT; Long-term Debt / Total Liabilities is defined as DLTT/ LT. Panel C shows the average transition probabilities. ***/**/* denote the statistical significance at 1%/5%/10% level.

Table 4: **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 5: **Correlation between LRISK and other Text-based Firm-level Exposure Measures**

	LRISK	LRISK_COURT	LRISK_TRIAL	LRISK_TERMS	PRISK	NPRISK	RISK	PSENT	NPSENT	SENT
LRISK	1.00									
LRISK_COURT	0.80	1.00								
LRISK_TRIAL	0.88	0.48	1.00							
LRISK_TERMS	0.44	0.20	0.29	1.00						
PRISK	0.09	0.09	0.07	0.03	1.00					
NPRISK	0.05	0.04	0.03	0.03	0.19	1.00				
RISK	0.08	0.08	0.06	0.03	0.47	0.49	1.00			
PSENT	-0.15	-0.13	-0.14	-0.09	-0.07	-0.09	-0.14	1.00		
NPSENT	-0.10	-0.08	-0.07	-0.08	-0.06	-0.04	-0.09	0.30	1.00	
SENT	-0.18	-0.15	-0.14	-0.10	-0.12	-0.12	-0.21	0.68	0.56	1.00

Note: This table presents the Pearson correlation between **LRISK** and its three components—**LRISK_COURT**, **LRISK_TRIAL**, and **LRISK_TERMS**—alongside relevant text-based measures. Political and non-political risk, as well as sentiment measures, are sourced from Hassan et al. (2019). Correlations are computed at the aggregate level, though first grouping by firm or year before calculating correlations does not affect the main implications.

Table 6: **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 7: **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: $(\texttt{ceq} + \texttt{dltt})/\texttt{at}$.
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: $\texttt{ppent}/\texttt{at}$.
Intangibility	Ratio of intangible assets to total assets: $\texttt{intan}/\texttt{at}$.
Z-score	Altman Z-score (1968)
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	Capital investment (%): 100 × [capx + aqc] / lagged total assets. Missing values in capx and aqc are imputed with zeros.
Δ AT	Change in total assets (%): 100 × (at - L1.at) / lagged total assets.
R&D	R&D intensity (%): 100 × xrd / lagged total assets. Missing values in xrd are imputed with zeros.
Total	Total investment (%): 100 × [capx + aqc + xrd + Δintan] / lagged total assets. Δintan is the annual change in intangible 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
Intangibility	59,163	0.18	0.21	0.00	0.01	0.10	0.30	0.90
Z-score	52,710	0.94	6.28	-81.73	0.73	2.02	3.23	12.43
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	60,149	7.24	14.48	-14.11	0.88	3.22	7.84	204.84
Δ AT	57,324	11.93	49.31	-82.90	-3.89	4.55	15.42	750.22
R&D	60,149	6.25	15.26	0.00	0.00	0.00	5.89	194.45
Total	60,149	16.57	36.44	-44.36	1.32	7.19	18.43	523.37

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 8: **Legal Risk and Firm Investment**

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	CAPX + AQC	CAPX + AQC	CAPX + AQC	Δ AT	Δ AT	Δ AT	R&D	R&D	R&D	Total	Total	Total
LRISK	**-0.223****	**-0.267****	**-0.330*****	**-1.086*****	**-1.022*****	**-1.467*****	**-0.387*****	**-0.463*****	**-0.489*****	**-0.867*****	**-1.233*****	**-1.478*****
	(-2.569)	**(-2.389)**	**(-2.910)**	**(-3.470)**	**(-3.055)**	**(-4.092)**	**(-3.998)**	**(-3.374)**	**(-3.659)**	**(-3.270)**	**(-4.284)**	**(-5.164)**
Size	-3.632***	-3.374***	-2.330***	-24.254***	-23.291***	-16.693***	-3.304***	-2.255***	-1.768***	-13.687***	-10.501***	-6.792***
	(-9.890)	(-9.763)	(-6.716)	(-14.236)	(-13.368)	(-12.172)	(-12.659)	(-8.262)	(-7.020)	(-13.911)	(-10.842)	(-7.460)
Tobin's Q	3.302***	2.524**	2.202**	-6.727**	-10.014**	-11.015**	-8.865***	-5.507***	-5.776***	-4.536	-1.049	-2.429
	(5.319)	(2.596)	(2.228)	(-2.113)	(-2.569)	(-2.718)	(-5.826)	(-3.652)	(-3.883)	(-1.423)	(-0.277)	(-0.640)
Leverage	-9.714***	-8.160***	-7.369***	-7.537*	-7.392	-0.109	-0.360	-3.090***	-2.730***	-14.728***	-13.266***	-10.048***
	(-8.026)	(-5.926)	(-5.554)	(-1.826)	(-1.417)	(-0.019)	(-0.541)	(-3.417)	(-2.994)	(-6.231)	(-4.748)	(-3.594)
Sales Growth		0.124***	0.107**		0.458	0.301		0.060	0.053		0.320**	0.255*
		(2.946)	(2.309)		(1.329)	(0.891)		(0.907)	(0.800)		(2.537)	(2.064)
Tangibility		-9.961***	-10.623***		-8.575	-2.618		3.478**	3.353**		-19.370***	-19.925***
		(-4.699)	(-5.021)		(-1.583)	(-0.425)		(2.304)	(2.299)		(-5.299)	(-5.591)
Intangibility		-16.011***	-16.358***		-10.981***	-14.270***		-0.417	-0.455		-60.601***	-61.851***
		(-8.233)	(-8.405)		(-3.284)	(-3.936)		(-0.427)	(-0.465)		(-12.362)	(-12.276)
Z-score		0.311***	0.278***		0.094	-0.220		-0.623***	-0.645***		-0.100	-0.252
		(4.460)	(4.201)		(0.245)	(-0.590)		(-5.398)	(-5.719)		(-0.427)	(-1.155)
Constant	33.882***	37.603***	30.650***	190.269***	185.018***	137.771***	35.050***	26.311***	23.087***	120.797***	112.670***	87.651***
	(12.614)	(14.302)	(11.319)	(15.263)	(14.577)	(14.083)	(15.230)	(14.410)	(13.033)	(16.101)	(15.152)	(11.790)
Firm Fixed Effects?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects?	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No
Industry Fixed Effects?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	46,357	36,546	36,546	46,343	36,531	36,531	46,357	36,546	36,546	46,357	36,546	36,546
R^2	0.263	0.254	0.246	0.229	0.239	0.215	0.799	0.814	0.812	0.265	0.261	0.250

Note: This table reports panel regressions of alternative investment measures on lagged firm-level legal risk exposure (LRISK). The data spans 2002 - 2024. The dependent variables include capital expenditures and acquisitions (denoted as CAPX), changes in total assets (ΔAT), R&D expenditures (R&D), and a comprehensive investment measure combining capital expenditures, acquisitions, R&D, and changes in intangible assets (denoted as Total). All investment measures are scaled by beginning-of-year total assets. LRISK is a firm-level measure of legal risk constructed from textual analysis of corporate disclosures. Control variables include firm size, Tobin's Q, leverage, sales growth, asset tangibility, intangible intensity, and Altman's Z-score. Standard errors are two-way clustered at the firm and year levels. Coefficient estimates are reported with *t*-statistics in parentheses. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

Table 9: **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**: Market Responses to Legal Risk: Case Evidence
- **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 Market Responses to Legal Risk: Case Evidence

Legal risk is not merely an abstract concern for firms—it has tangible financial consequences that investors and corporate decision-makers systematically monitor. High-profile lawsuits, regulatory actions, and emerging legal challenges frequently trigger significant market reactions, demonstrating that legal uncertainty is a key determinant of firm value. Boards of directors are also highly responsive to heightened litigation risk, adjusting their oversight and decision-making accordingly (Laux 2010).[1] The following cases illustrate how legal risk translates into substantial stock price movements and strategic corporate responses.

Investors systematically care about legal risk as they actively monitor and react to such risks. For instance, shares of GSK dropped over 9% after a Delaware judge allowed 70,000 lawsuits alleging that its discontinued heartburn drug, Zantac, caused cancer to proceed, wiping out nearly $8.9 billion in market value. These lawsuits have posed a significant concern, erasing approximately $40 billion in combined market value from GSK, Sanofi, Pfizer, and Haleon during a similar legal scare in August 2022.[2]

Recent evidence further highlights the significant financial repercussions of legal risk. Climate change litigation, for example, has emerged as a material financial risk for firms in carbon-intensive industries. A study examining lawsuits against major polluters found that litigation announcements lead to statistically significant declines in share prices, reflecting investor concerns over regulatory penalties, reputational damage, and future compliance costs.[3] As climate-related litigation becomes more prevalent, firms exposed to environmental liabilities face heightened legal uncertainty, reinforcing the notion that legal risk is a systematic force shaping investment decisions.

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 ad-

[1] See also https://fticommunications.com/the-decade-of-disputes-litigation-is-on-the-boardroom-agenda-for-the-year-ahead/

[2] https://www.reuters.com/business/healthcare-pharmaceuticals/gsk-shares-tumble-9-after-70000-zantac-lawsuits-allowed-move-forward-2024-06-03/

[3] https://www.theguardian.com/environment/2023/may/22/big-polluters-share-prices-fall-climate-lawsuits-fossil-fuels-study

ditional 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 a series of additional analyses to evaluate the robustness and distinctiveness of the **LAW** factor. I begin with alternative calendar-time portfolio analyses, using different formation windows and a zero–nonzero legal risk sorting approach. I then address data mining concerns to directly confront the "factor zoo" critique. Importantly, the subcomponents of legal risk are individually priced, underscoring the unique text-based construction and economic richness of the **LAW** factor.

Next, I test whether the **LAW** factor is subsumed by broader behavioral explanations, including market sentiment, mispricing-based factors, and belief disagreement. I also investigate whether short-term mood fluctuations—proxied by day-of-the-week return patterns—can explain the legal risk premium. Finally, I explore the relationship between the **LAW** factor and ESG-based risk factors. Collectively, these exercises reinforce the robustness and unique economic interpretation of the **LAW** factor, allowing me to rule out potential data mining concerns.

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

This section expands the empirical analysis by introducing additional portfolio strategies to test the robustness of the main findings. The primary result employs a (12-12) portfolio strategy, where portfolios are formed at the end of June each year using the past 12 months (i.e., 4 quarters) of earnings calls transcripts information, and stocks are weighted by their

market capitalization as of that date. The formed portfolio lasts for the next 12 months, ending in June of the following year. This approach aligns with the standard methodology for constructing risk factor anomalies.

To leverage the higher frequency of quarterly earnings conference calls, Table B.1 reports three additional portfolio strategies: (12-3), (12-6), and (6-6). Moreover, I present an alternative "zero-nonzero" approach under the (12-12) strategy, where portfolios are sorted by stocks with zero versus positive exposure to legal risk. To ensure that any performance differences arise solely from the portfolio formation periods, I preserve the original setting by using market capitalization weights as of the end of June each year.

The results from all four strategies, summarized in Table B.1, demonstrate strong statistical significance. These findings affirm that variations in portfolio formation periods do not affect the core conclusion. In additional unreported analyses, I confirm that including both utility and financial firms (which were previously excluded in the construction of the LAW factor in the main text) and capping the market capitalization of my earnings call firm sample at the 95th percentile (which approximately corresponds to the 80th percentile NYSE breakpoint[4] as used in Jensen et al. (2023) yields qualitatively similar results.

C.2 Data Mining Concerns

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.

[4]This discrepancy appears to stem from the inclusion of NASDAQ and NYSE American (formerly known as AMEX) firms in my sample, in addition to NYSE firms.

Table B.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	0.950	0.970	**0.347*** **
	(2.02)	(3.26)	(3.29)	**(3.34)**
FF3 alpha	-0.185	0.107	0.154	**0.339*** **
	(-2.26)	(2.03)	(2.81)	**(3.32)**
FF5 alpha	-0.156	0.062	0.133	**0.289*** **
	(-2.13)	(1.25)	(2.97)	**(3.09)**

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

	L1 No legal risk	L2 Middle legal risk	L3 High legal risk	L3 - L1 Long - Short
Excess return	0.693	0.946	0.979	**0.286*** **
	(2.21)	(3.29)	(3.29)	**(3.03)**
FF3 alpha	-0.163	0.109	0.153	**0.316*** **
	(-2.07)	(2.04)	(2.87)	**(3.33)**
FF5 alpha	-0.163	0.042	0.120	**0.283*** **
	(-2.24)	(0.87)	(2.32)	**(3.08)**

Panel C: (6-6) strategy (%)

	L1 No legal risk	L2 Middle legal risk	L3 High legal risk	L3 - L1 Long - Short
Excess return	0.746	1.003	0.953	**0.207** **
	(2.54)	(3.41)	(3.22)	**(2.40)**
FF3 alpha	-0.085	0.171	0.112	**0.198** **
	(-1.32)	(3.28)	(1.94)	**(2.31)**
FF5 alpha	-0.117	0.122	0.063	**0.180*** **
	(-1.88)	(2.45)	(1.22)	**(2.27)**

Panel D: (12-12) strategy using Zero-Nonzero legal risk approach (%)

	L1 No legal risk	L2 Any legal risk	L2 - L1 Long - Short
Excess return	0.645	0.969	**0.324*** **
	(2.03)	(3.37)	**(3.44)**
FF3 alpha	-0.209	0.139	**0.348*** **
	(-2.17)	(3.66)	**(3.64)**
FF5 + mom alpha	-0.205	0.089	**0.294*** **
	(-2.10)	(2.88)	**(3.05)**

Note: In both panels, the average future excess returns (2004:07 - 2024:06) alongside the alpha results derived from portfolios constructed based on ex-ante legal risk measures are presented. Panel A presents a portfolio strategy that uses (12-3) strategy, and Panel B presents (3-3) strategy. Firms in P1 (Low litigation risk) are those with no exposure to litigation risk. Firms with litigation 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.

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.1 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.[5]

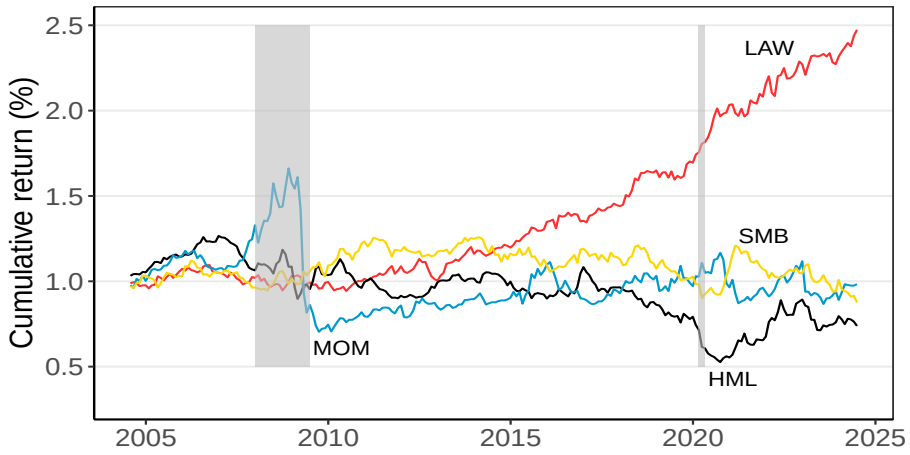
Importantly, the **LAW** factor demonstrates resilience during economic downturns, including the 2008 Global Financial Crisis and the 2020 COVID shock. Legal risks become particularly salient in these periods, as firms and investors are already under financial stress. Investors tend to value firms whose 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).

Figure A.2 further underscores the uniqueness of the legal risk factor (**LAW**) within the broader factor zoo. Specifically, I examine the correlation between **LAW** and the full set of 193 anomalies with complete data that align with the legal risk factor's sample period (July 2004 to December 2023; the anomaly data is sourced from Chen and Zimmermann 2022). The results indicate that none of these anomalies exhibit an absolute Pearson correlation greater than 0.4, underscoring that the legal risk factor constructed in this paper is a distinct and significant driver of cross-sectional variation in stock returns.[6]

[5] 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.

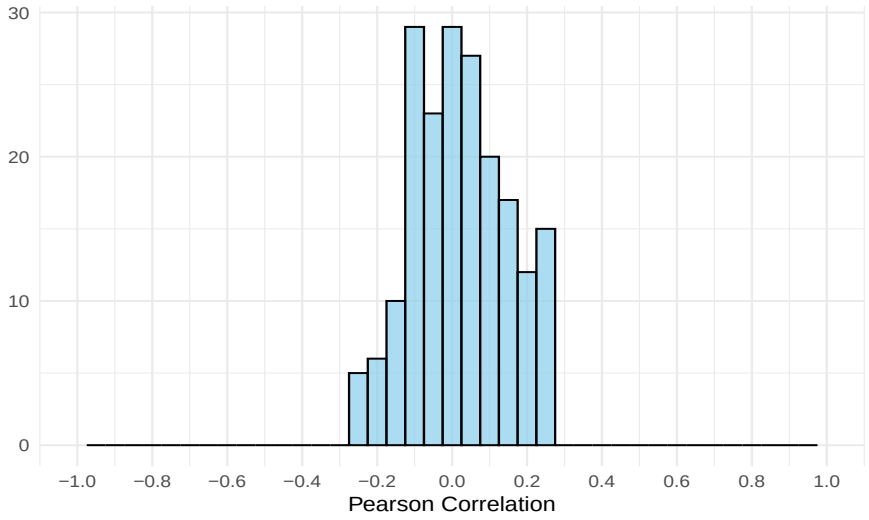
[6] A similar judgment rule based on Pearson correlation is used in Birru et al. (2024). Additionally, I find that the correlations between **LAW** and other well-known factors are as follows: **MKT** (-0.14), **SMB** (-0.26), **HML** (-0.19), **MOM** (0.25), **RMW** (0.10), **CMA** (0.04), and **LIQ** (-0.02).

Figure A.1: **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.

Figure A.2: **Risk Factor Correlation**



Note: The figure displays a histogram of Pearson correlation coefficients between the legal risk factor ("LAW" factor) and 193 risk anomalies identified in the finance literature. The sample period spans from July 2004 to December 2023, based on the availability of the factor dataset. The x-axis represents the correlation coefficients (ranging from -1 to 1), while the y-axis shows the number of risk anomalies within each correlation bin.

The text-based construction of the **LAW** factor offers a unique advantage, as it enables a decomposition of legal risk into distinct thematic categories based on keyword clusters. As detailed in Table 2 in the main text, the analysis begins by classifying legal content into three subcomponents—**LRISK_COURT**, **LRISK_TRIAL**, and **LRISK_TERMS**. While the main Table 3 reports results for a long-short portfolio where the long leg contains firms with above-median mentions across all three legal risk categories, I present complementary results that isolate each component. For instance, for the **LAW_COURT** factor, I long firms with above-median mentions of court-related legal terms (capturing macro-level legal risk) and short firms with no legal risk mentions—identical to the L1 group used in the main analysis. This approach leverages the extensive-margin sorting method, which offers flexibility and enables clean comparisons between the bundled **LAW** factor and its individual subcomponents. I conduct the same exercise for **LRISK_TRIAL** (micro-level legal risk) and **LRISK_TERMS** (generic legal terminology).

The results are as follows: for the **LAW_COURT** factor, over the full 2004–2024 sample (240 months), the FF3 (FF5) alpha is 0.291% (0.274%) per month, with Newey–West adjusted t-statistics of 2.42 (2.09). Similarly, for the **LAW_TRIAL** factor, the FF3 (FF5) alpha is 0.299% (0.252%) with t-statistics of 2.76 (2.26). In contrast, the **LAW_TERMS** factor is not priced. Taken together, this subcomponent analysis suggests that the pricing power of the **LAW** factor is driven by economically meaningful legal exposures, both from macro-level (**LAW_COURT**) and micro-level (**LAW_TRIAL**) legal risks.

Finally, as demonstrated in the case studies, legal risk is closely linked to both environmental concerns and the banking sector. Moreover, I confirm that the LAW factor is not explained by ESG factors (Pastor et al. (2022) and Pedersen et al. (2021); Table B.4 in the Online Appendix) and subsumes the bank size factor (Gandhi and Lustig 2015; unreported), reinforcing the argument that the LAW factor has a broader scope.

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. 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.

Table B.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 B.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.

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 B.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.

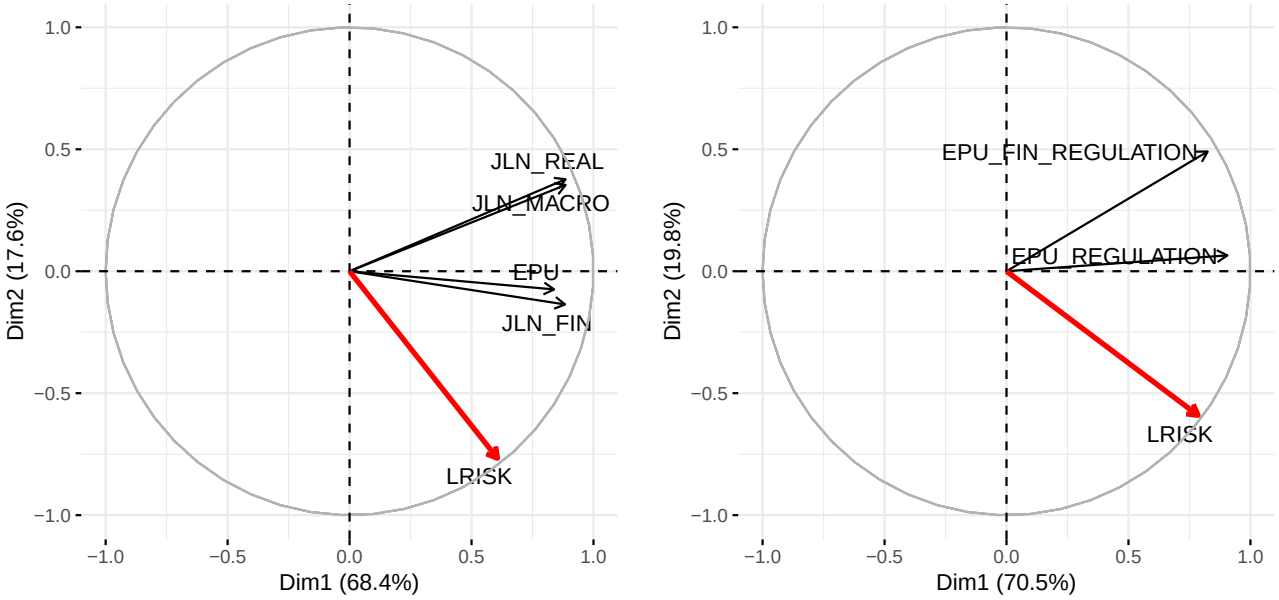
C.4 LAW Factor vs. ESG Factors

In Table B.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.3: **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 4.3—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 observation. 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.3 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 B.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

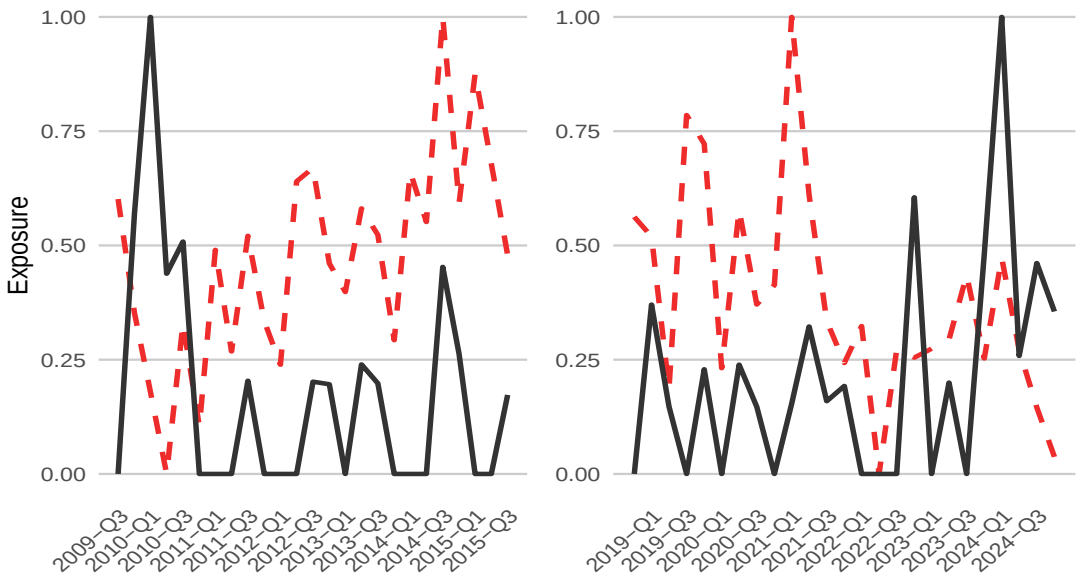
Panel B: LAW factor is dependent variable

	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.4: **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[7] 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

[7]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.4). 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.

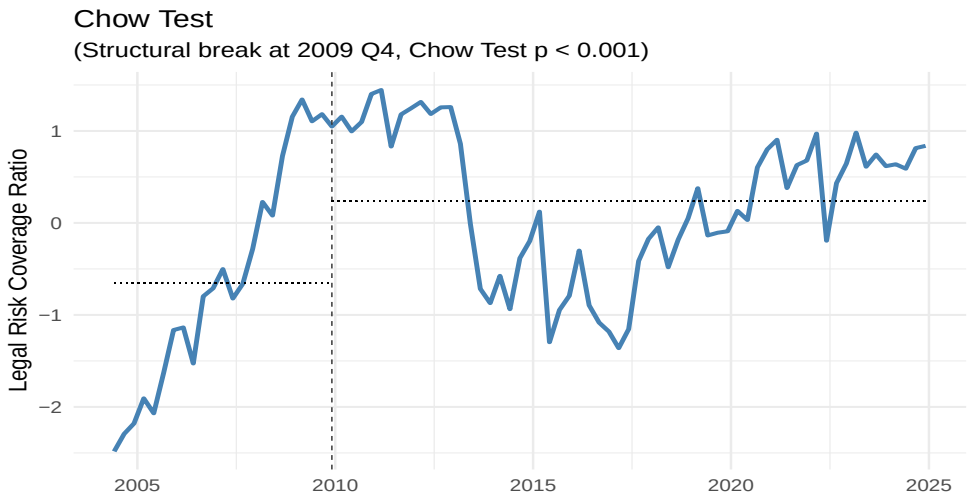
The Securities and Exchange Commission (SEC) played a pivotal role in these reforms by implementing stricter enforcement measures. In 2010, the SEC established specialized units within its Enforcement Division, such as the Market Abuse Unit and the Asset Management Unit, to address complex financial 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.5: **Chow Test**



Note: This figure plots the Chow test around 2010 using the aggregate legal risk measure introduced in Section 4.3. 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.5, 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.[8]

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

[8]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 $r > \mu$.[9]

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.

[9]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,[10] 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.

[10] 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.

Birru, J., Gokkaya, S., Liu, X. and Markov, S.: 2024, Quants and market anomalies, *Journal of Accounting and Economics* **78**, 101688.

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.

Gandhi, P. and Lustig, H.: 2015, Size Anomalies in U.S. Bank Stock Returns, *Journal of Finance* **70**(2), 733–768.

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.

Laux, V.: 2010, Effects of Litigation Risk on Board Oversight and CEO Incentive Pay, *Management Science* **56**(6), 938–948.

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.