Granular Firms and Endogenous Risk Premia

Hyein Han*

This version: November 2025

Abstract

This paper studies how financial risks shape the liquidity and leverage choices of the largest and most market-exposed U.S. public firms. These granular firms hold substantially more cash and use less leverage than other firms, especially when aggregate volatility rises. They also face disproportionately large increases in credit spreads when the excess bond premium (EBP) spikes, driven almost entirely by its non-default componentevidence that the interaction of size and market exposure generates undiversifiable risk priced by investors. To explain these patterns, I develop a structural model in which investors are imperfectly diversified across a finite set of dominant firms. Idiosyncratic shocks to granular firms load into the stochastic discount factor, creating a firm-specific price-impact wedge that raises marginal borrowing costs through a steeper, downward-sloping bond pricing schedule. The model matches key cross-sectional patterns in cash, leverage, and credit spreads, and implies that granular risk premia account for roughly one-fifth to one-quarter of the post-2000 increase in cash holdings among top-1% firms by market capitalization

Keywords: Granular firms; Undiversifiable risk; Endogenous risk premia; Corporate liquidity; Asset pricing
JEL Codes: G12, G32, E44, D53

*Department of Economics, New York University. E-mail: hh2969@nyu.edu. I am especially grateful to Ricardo Lagos and Simon Gilchrist for their invaluable guidance and support. I also thank Mark Gertler, Jaroslav Borovika, Virgiliu Midrigan, Viral Acharya, and Thomas Philippon for many helpful suggestions in shaping this paper. I am grateful to Keshav Dogra, Marco Del Negro, Giorgio Topa, and seminar participants at the NY Fed Junior Macro Workshop, the Midwest Macroeconomics Meeting (Cleveland Fed), and the NYU Student Macro Lunch Seminar for valuable comments. All remaining errors are my own.

1 Introduction

What financial risks do the largest firms face, and how do these risks shape their liquidity and leverage decisions? This question has become increasingly salient as economic concentration amplifies the macroeconomic footprint of individual firms. Over the past two decades, a striking pattern has emerged: firms at the very top of the size distribution–the top 1% of U.S. public companies by market capitalization–have adopted markedly more conservative financial policies than the rest of the corporate sector. These "granular" firms hold historically high levels of cash, maintain unusually low leverage, and exhibit behavior that becomes increasingly pronounced as their systematic risk exposure rises.

What makes this behavior puzzling is that, by every conventional metric, these firms appear the least financially constrained. They enjoy strong fundamentals, low default probabilities, and unparalleled access to capital markets–conditions under which standard corporate finance theories would predict lower cash holdings and higher reliance on cheap external finance. Yet granular firms behave in the opposite manner. This divergence raises a fundamental question: why do the largest and most creditworthy firms in the economy choose to underborrow and oversave?

In this paper, I argue that understanding this puzzle requires distinguishing between size alone and the interaction of size with systematic risk exposure. While the top 1% of the market-cap distribution contains many large firms, the conservative financial behavior I document is concentrated among those whose valuations co-move strongly with the aggregate market. Throughout the analysis, I therefore refer to granular firms as those that are not only extremely large but also have high equity betas–the subset for whom firm-specific shocks cannot be diversified away by investors. This classification aligns with the granular hypothesis in macroeconomics, but with an important refinement: granularity becomes financially consequential only when size is paired with high market exposure.

Using matched data from Compustat, CRSP, and TRACE, I document three empirical facts about these large, market-sensitive firms. First, they have accumulated disproportion-

ately large cash buffers since the early 2000s, diverging sharply from both mid-sized firms and other top-1% firms with lower betas. Second, their cash holdings are higher and their leverage is lower, and both margins respond much more strongly to aggregate volatility: granular firms increase cash and reduce leverage as market risk rises, in stark contrast to the predictions of standard tradeoff or pecking-order theories.

Third, and most importantly, granular firms experience disproportionately large increases in corporate bond spreads when investors aggregate risk premia rise, as captured by movements in the aggregate excess bond premium (EBP). Using the Gilchrist and Zakrajšek (2012) decomposition of credit spreads into expected default losses and the excess bond premium, I show that the spread increases for granular firms are driven almost entirely by the excess-premium component rather than by changes in expected default risk. This pattern indicates that when intermediary risk-bearing capacity tightens and aggregate EBP spikes, investors demand additional compensation for absorbing undiversifiable shocks from large, market-exposed issuers, generating a granular risk premium. In response, granular firms sharply curtail net debt issuance and accumulate cash, providing direct evidence that rising aggregate risk premia have systematic consequences for their financing and liquidity policies.

To explain these patterns, I develop a structural model in which a risk-averse representative investor holds an imperfectly diversified portfolio consisting of a continuum of small firms and a finite set of granular firms. In contrast to standard representative-agent models with perfectly diversified investors, idiosyncratic shocks to granular firms enter directly into the investors marginal utility and thus into the stochastic discount factor. This generates a firm-specific, state-contingent price wedge: the required return on a granular firm's debt rises with its exposure to undiversifiable risk, particularly in bad states. Anticipating this price impact, granular firms optimally delever and accumulate liquidity, effectively self-insuring against the heightened risk premium embedded in their own debt.

I incorporate this perspective into a corporate debt financing framework that builds on recent structural models of long-term debt (e.g., Gomes and Schmid, 2021) but departs from

them by endogenizing the investors pricing kernel. Rather than taking discount rates as exogenous or fully diversified, the representative investor in the model prices bonds in a way that directly reflects exposure to each granular firm. This feature allows a firms own idiosyncratic risk–when sufficiently large–to feed into the pricing of its debt. In contrast to environments with atomistic firms, the model generates a feedback loop from a granular firm's risk to its cost of capital, inducing financial conservatism even in the absence of traditional frictions.

The mechanism operates through an endogenous price-impact wedge: bond prices incorporate both expected discounted repayment and a risk adjustment reflecting the covariance between the firms repayment and the investors pricing kernel. When a granular firm issues more debt, its bond price declines not only because expected default rises but also because the investor demands additional compensation for absorbing the firms undiversifiable, issuer-specific risk. For firms that are both large and highly market-exposed, this wedge becomes quantitatively meaningful even without aggregate shocks. Consequently, granular firms behave as if financially constrained–not because they lack access to credit, but because their own financing decisions move market prices in ways that feed back into their cost of capital. This mechanism rationalizes the empirical behavior of granular firms, including their countercyclical accumulation of cash and sharp deleveraging in periods of elevated macro-financial uncertainty.

Quantitatively, the model replicates key cross-sectional patterns in leverage, credit spreads, and cash holdings among granular and non-granular firms. It also clarifies why even extremely safe, investment-grade issuerssuch as the granular firms in the top 1%–face credit spreads far exceeding expected default losses. A variance decomposition of log bond prices shows that only 20-30% of price dispersion is explained by the risk-neutral price component, while the remaining variation is accounted for by the risk-adjustment component–the price-impact wedge–together with the covariance between the wedge and the risk-neutral price. This decomposition highlights that the well-known "credit-spread puzzle" for high-grade is-

suers is particularly relevant for granular firms. Because they are both large and highly exposed to the market, investors require substantial additional compensation for bearing their undiversifiable, issuer-specific risk, even when their default probabilities are low.

Motivated by the rise in idiosyncratic volatility since the early 2000s–documented in Section 4.3, I feed an increasing volatility sequence into the model and compute equilibrium liquidity policies under two environments: the baseline economy with the granular risk-premium channel active, and a counterfactual economy in which large firms behave as price takers. As volatility rises, the gap between baseline and counterfactual cash holdings widens, reflecting the growing importance of the price-impact wedge. By 2020, this widening differential accounts for roughly one-fifth to one-quarter of the observed increase in cash holdings among top-1% firms, quantifying how heightened firm-level uncertainty interacts with the granular risk-premium channel to elevate liquidity demand.

Contributions This paper makes both empirical and theoretical contributions. Empirically, I document a new set of facts about the financial behavior of firms that are simultaneously extremely large and highly exposed to aggregate risk. Prior work has examined cash policies, leverage choices, and precautionary savings across the full firm-size distribution, but has not isolated the subset of top-1% firms whose shocks carry macroeconomic significance. I show that these granular firms account for a disproportionate share of the post-2000 rise in corporate cash holdings, and that their cash and leverage choices respond far more strongly to fluctuations in aggregate risk premia than those of smaller firms or even other large, low-beta firms. Using secondary bond-market data combined with the Gilchrist and Zakrajšek (2012) decomposition, I further show that granular firms face systematically higher excess bond premia during episodes of heightened investor risk aversionevidence that markets assign a distinct, non-default risk premium to their debt. Together, these findings reveal an empirically meaningful channel through which size and systematic exposure interact to influence corporate financing behavior.

Theoretically, I develop a tractable model in which a representative investor prices the debt of a finite set of large firms under incomplete diversification. Unlike standard corporate-finance or asset-pricing frameworks that treat discount rates as exogenous or fully diversified, the model features a firm-specific pricing kernel: idiosyncratic shocks to sufficiently large and high-beta firms load directly into the investors marginal utility and thus into bond prices. This generates a downward-sloping demand curve for granular firm's debt, an endogenous price-impact wedge, and a feedback loop from a firms own risk to its cost of capital. The mechanism induces granular firms to deleverage and accumulate liquidity–even in the absence of traditional frictions–matching the financial conservatism observed in the data.

In doing so, the paper provides a unified explanation for the rise in cash holdings among dominant U.S. public firms and for the persistent credit-spread wedge faced by even investment-grade issuers. The model quantifies how the combination of heightened post-2000 idiosyncratic volatility and the granular risk-premium channel contributes significantly to the increase in liquidity buffers. More broadly, the analysis highlights how incomplete diversification in credit markets creates endogenous risk premia that shape the balance-sheet policies of the largest, most systemically relevant firms. This mechanism offers a new rationale for the emergence of large cash hoards and low leverage among granular firmsone that differs from traditional explanations based on agency frictions, pecking-order motives, or asymmetric information, and instead arises from endogenous risk premia embedded in their own debt.

Related Literature

Drivers of corporate financial policies This paper contributes to the literature on the determinants of corporate cash holdings and financing policies. Understanding firms' incentives to accumulate liquid assets is a core question in corporate finance, extensively examined both in cross-section and over time. The literature has documented classic motives such as agency problems, tax considerations, transaction costs, and precautionary savings

(e.g. Opler et al., 1999, Foley et al., 2007,Riddick and Whited, 2009). Almeida, Campello and Weisbach (2004) empirically estimate that financially constrained firms save a significantly higher fraction of incremental cash flows (positive cash flow sensitivity), directly linking cash holdings to precautionary liquidity management. Bates, Kahle and Stulz (2009) document a pronounced secular rise in U.S. cash ratios and trace it to firms with high cash-flow volatility, intensive R &D, and low dividend payouts–attributes consistent with precautionary liquidity demand.

At an aggregate level, recent studies identify structural transformations that have further amplified the secular upward trend in corporate cash holdings. Graham and Leary (2018) document a pronounced "Nasdaq effect," arguing the rise in economy-wide cash ratios since the 1980s was driven largely by a new wave of technology-intensive firms entering public markets with structurally higher liquidity needs. Begenau and Palazzo (2021) show that this shift in firm composition towards younger, innovation-oriented firms can fully account for the observed rise in average cash ratios. On the other hand, Falato et al. (2022) highlight the increasing reliance on intangible assets, which lowers collateral values and reduces debt capacity, further incentivizing firms to maintain elevated liquidity buffers. And macro-financial forces, from declining real interest rates (Azar, Kagy and Schmalz, 2016) to repatriation-tax incentives (Faulkender, Hankins and Petersen, 2019), have lowered the cost of carrying cash.

My empirical findings complement and extend this literature by documenting a previously overlooked dimension: the role of very large, market-sensitive firms in driving corporate liquidity patterns. Whereas prior studies often emphasize precautionary saving motives among smaller or financially constrained firms, I document that even the largest and traditionally unconstrained firms accumulate cash for precautionary reasons when they face undiversifiable granular risk–risk that arises from the interaction of their scale and their exposure to aggregate conditions. This insight adds a new facet to the corporate liquidity puzzle: large, market-sensitive firms self-insure not because of agency conflicts or financing frictions, but because their own shocks carry risk premia in credit markets. This perspective complements

the secular trends noted by Bates, Kahle and Stulz (2009), but refines the narrative by revealing important heterogeneity within the firm-size distribution. By linking large-firm cash and debt policies directly to measures of macro-financial stress, I identify a new channel in the corporate liquidity puzzleone in which size and systematic risk, rather than financing frictions alone, drive firms buffer-stock behavior.

This paper also relates to studies of corporate debt policy and maturity structure under risk. Firms can respond to heightened uncertainty not only by hoarding cash but also by adjusting leverage and debt maturity to mitigate refinancing risk. For example, firms facing uncertainty may reduce short-term debt or raise long-term financing as a precaution (Jungherr and Schott, 2021). Empirically, the 2007-09 crisis offers a clear example: many firms drew down revolving credit lines and issued long-dated bonds–so-called "borrowing to save–to lock in liquidity before market conditions worsened. Recent evidence by Xiao (2022) further shows that, in that episode, the firms that ramped up leverage the most simultaneously hoarded safe assets, preserving liquidity as a self-insurance strategy.

In contrast, I find that the very largest, systemically important firms adopt the opposite adjustment: when aggregate volatility or excess bond premia spike, they tighten issuance (delever) and accumulate cash rather than increase debt. This pattern holds not only in crisis episodes but persistently over the past two decades. By linking the debt-to-asset and cash-to-asset responses of the top 1 % firms directly to VIX and EBP fluctuations, we demonstrate that undiversified risk exposure drives a structural, precautionary rebalancing of the balance sheet–deleveraging and cash hoarding–instead of the borrowing expansions emphasized in prior work. My results thus enrich the precautionary saving literature by showing that, for the largest market-sensitive firms, heightened uncertainty triggers self-insurance through both reduced external financing and increased internal liquidity, rather than through simulataneous borrowing.

Granular Macroeconomics: Large Firms and Macro-Financial Risk My work is also connected to the growing literature on granular hypothesis and its idea that shocks to individual large firms can influence aggregate outcomes. Gabaix (2011) showed that idiosyncratic firm-level shocks do not wash out in the aggregate if the firm size distribution is fat-tailed, invalidating the diversification argument of Lucas (1995). This insight has been corroborated: a handful of large firms explain a significant fraction of fluctuations in output and other aggregates (e.g., Carvalho and Grassi, 2019). Moreover, Malevergne, Santa-Clara and Sornette (2009) document that the granular residual–the excess return of an individual large firm over the market average–is a priced factor in U.S. equities, implying investors demand compensation for undiversifiable large-firm risk, a point directly relevant to our theory.

Building on these ideas, this paper investigates how the granular nature of the economy affects risk transmission through financial markets. I provide evidence that when aggregate volatility rises, it disproportionately impacts firms at the top of the size distribution. These dominant firms have equity betas that make them highly sensitive to market-wide shocks, and in crises they experience especially pronounced increases in credit spreads (long-term bond yield spreads). This aligns with the notion that their idiosyncratic distress is not entirely idiosyncratic from the investors' perspective–it has systematic consequences.

The theoretical framework developed in this paper links these observations to research on asset pricing under incomplete risk sharing. In the classical Arrow-Debreu world, idiosyncratic firm risk is fully diversifiable and does not command a risk premium (e.g., Lucas, 1978). But if some risks cannot be fully shared or arbitraged away, the representative investor's stochastic discount factor will reflect those uninsurable risks. I posit that the investors' pricing kernel does not entirely wash out granular firm risk, so that a large firm's idiosyncratic shocks carry a risk premium in equilibrium. This setup is in line with empirical anomalies where idiosyncratic volatility is priced or where large firms' stock returns have a common component that investors cannot diversify away (e.g., Gao, 2022). By linking a

firm's cash-and-debt policy to the state-contingent price of its own idiosyncratic risk, the model integrates corporate-finance policies with incomplete-markets asset pricing.

Roadmap

The remainder of the paper is organized as follows. Section 2 documents the secular trends in the U.S. corporate sector, estimates firm-level leverage and cash policies, and provides suggestive evidence for the existence of granular risk premiums via the cross-section of corporate bond spreads. Section 3 introduces the structural model that rationalizes these empirical patterns, emphasizing the role of a pricing kernel wedge that leads large firms to deleverage and accumulate liquidity. Section 4 presents the model calibration and uses it to quantify the effects of the price-impact wedge on firm balance sheets, validating the model's implications through counterfactual exercises. Section 5 concludes.

2 Empirical Evidence

This section documents how the largest and most market-sensitive U.S. public firms manage liquidity and leverage in response to fluctuations in macro-financial risk, as captured by movements in aggregate volatility and investor risk premia. Following the terminology of the granular-hypothesis literature (e.g., Gabaix, 2011), I begin by identifying firms in the top 1 percent of the market-capitalization distribution in each quarter.[1] To align the empirical analysis with the mechanism developed in the model, I treat systematic risk exposure as a continuous dimension of granularity: within the top-1% size group, firms with higher equity

[1]While the 1 percent threshold is admittedly arbitrary, the results are robust to alternative cutoffs, such as the top 3 or 5 percent. The 1 percent criterion is chosen as it yields the most pronounced heterogeneity relative to other firms within the same industry. Similar patterns hold when using total assets instead of market capitalization to rank firms. This definition allows for time-varying composition of the granular group, which is especially relevant given the shift in firm prominence over time–for instance, the post-2000 rise of technology firms displacing traditional manufacturing leaders. Since the objective is to assess the aggregate implications of a corporate saving glut, the analysis focuses on firms that occupy the largest positions in the economy at each point in time. Accordingly, the main specification relies on the top 1 percent by market capitalization, with robustness checks using alternative definitions.

betas are those whose shocks are least diversifiable for investors and therefore most sensitive to changes in aggregate volatility (VIX) and credit-market risk premia (EBP).

Using this framework, I show that large, market-sensitive firms differ systematically from both smaller firms and large low-beta firms. They have accumulated disproportionately high cash buffers, maintain lower leverage, and display financial policies that react strongly to changes in VIX and to movements in the excess bond premium (EBP). These empirical patterns motivate the model developed in the next section, where the interaction of size and market sensitivity generates an endogenous risk premium that shapes firms financing and liquidity decisions.

The empirical analysis draws on quarterly data from CRSP/Compustat for firm-level financials and stock market information, and from the Trade Reporting and Compliance Engine (TRACE) for secondary-market bond prices and credit spreads. Details on sample construction, variable definitions, and data sources are provided in the Appendix.

2.1 The Evolution of Corporate Cash

The substantial rise in corporate cash holdings in recent decades has received considerable attention in the literature. However, much of the existing discourse has focused on the financing policies of the average public firm or on firms that are plausibly credit constrained-typically younger, smaller firms with limited tangible assets (e.g., Bates, Kahle and Stulz, 2009; Graham and Leary, 2018). In contrast, less is known about the behavior of the very largest firms, which account for a growing share of aggregate cash balances.

The left panel of Figure 1 plots the time series of aggregate cash-to-assets ratios for three firm-size groups. For the bottom 90 percent of firms by market capitalization, the ratio has increased steadily since the early 1980s, a trend often attributed to compositional changes in the public-firm universe, particularly the entry of high-cash NASDAQ firms and successive IPO cohorts with structurally elevated liquidity.

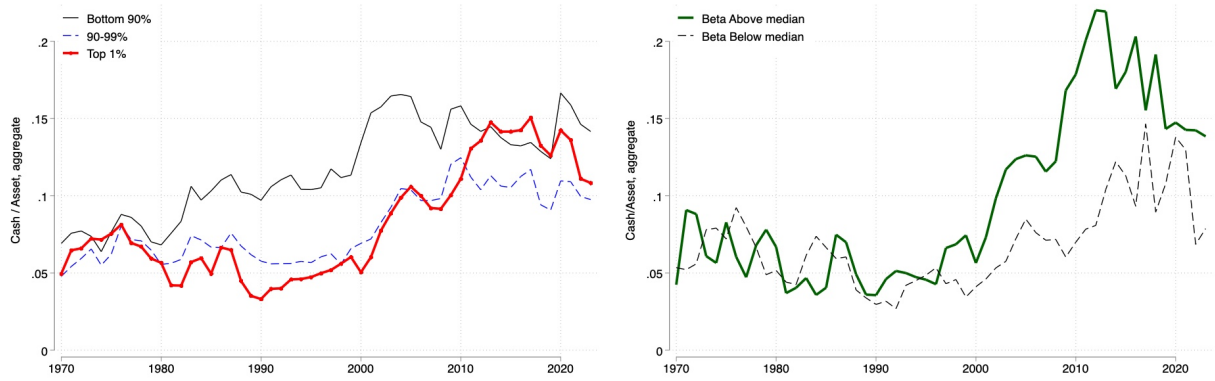
What is more striking, however, is the behavior of the largest firms. Beginning in the

early 2000s, the top 1 percent of firms experienced a rapid acceleration in cash accumulation, surpassing even the 90th–99th percentile groupfirms that are themselves large and systemically relevant. By the late 2010s–after the Global Financial Crisis (GFC) and prior to the COVID-19 pandemic–the top 1 percent had overtaken the bottom 90 percent in terms of aggregate cash-to-assets ratios.

To understand this heterogeneity within the top of the distribution, the right panel of Figure 1 decomposes cash ratios among the largest firms by equity beta. A clear pattern emerges: the rise in cash holdings is concentrated almost entirely among the largest firms with above-median market betas. In contrast, large firms with lower market exposure exhibit substantially more modest changes in liquidity over time. This visual evidence suggests that the interaction between scale and market sensitivity–rather than size alone–is central to explaining why some dominant firms accumulate significantly larger liquidity buffers.

Appendix Figure A2 further illustrates this divergence by showing the cross-sectional distribution of cash-to-assets ratios across size percentiles in selected years. While the relationship between firm size and cash holdings was roughly monotonic in earlier periods, a distinct U-shape emerges in the post-GFC years, with the top 1 percent sharply deviating from mid-sized firms. This reversal underscores the importance of examining not only aggregate trends but also heterogeneity at the top of the size distribution. Despite substantial cross-industry variation in riskiness and liquidity needs, this U-shaped pattern is remarkably consistent across major sectorsincluding manufacturing, services, and retailas documented in Appendix Figure A2.

Figure 1: Cash Holdings by Size and Market Sensitivity



Notes: The left panel plots the annual aggregate cash-to-assets ratio for three firm groups based on market capitalization: bottom 90% (black solid line), 90–99% (blue dashed line), and top 1% (red dotted line). Aggregate cash-to-assets is calculated as the sum of cash and short-term investments divided by the sum of book assets within each group. The right panel shows, among the top 1% largest firms, cash-to-assets ratios for firms with above-median versus below-median equity betas. Firm-level equity beta is estimated using a 60-month (5-year) rolling window of monthly excess stock returns regressed on excess market returns under a CAPM specification, requiring at least 24 non-missing observations. Financial firms, utilities, and quasi-government entities are excluded from the sample.

2.2 Systematic and Undiversifiable Risk, and Liquidity Demand

A natural starting point for understanding firms liquidity policies is the classical liquidity-management framework of Holmström and Tirole (1998), in which idiosyncratic shocks are diversifiable through intermediaries, while aggregate shocks—such as market-wide volatility spikes or funding disruptions—cannot be insured. In such environments, as emphasized by Acharya, Almeida and Campello (2013), firms with higher exposure to aggregate risk face a greater likelihood of encountering financing frictions precisely when liquidity is most valuable. This creates an endogenous motive for self-insurance: the higher a firms beta, the more strongly it should value internal liquidity.

For most firms in the economy, this logic is rooted entirely in exposure to *aggregate* shocks. However, for the largest firms—those at the top of the size distribution—beta takes on an additional interpretation. Because these firms account for a nontrivial share of

aggregate activity and asset-market movements, their own idiosyncratic shocks become less diversifiable for investors. In this sense, systematic exposure and firm size interact: higher beta not only reflects sensitivity to the market, but also proxies for the extent to which a large firms shocks themselves enter the market. This distinction is central to the granular mechanism developed in the model, where idiosyncratic shocks scaled by firm size become imperfectly diversifiable and therefore command risk premia in credit markets.

With this interpretation in mind, I examine empirically whether firms with greater systematic exposure hold more cash. As illustrated in Figure 1, the rise in cash among the top 1% of firms is concentrated among those with higher equity betas, suggesting that market sensitivity—not size alone—is a key driver of heterogeneity in liquidity behavior. Firm-level beta is estimated using a 60-month rolling window of monthly excess stock returns regressed on excess market returns under a CAPM specification, requiring at least 24 non-missing observations.[2]

Before turning to formal regressions, Table 1 presents unconditional summary statistics documenting how cash holdings vary jointly with firm size and market beta. Among firms in the top 1 percent of the market- capitalization distribution, there is a clear positive association between beta and liquidity: the cash-to-assets ratio increases monotonically across beta quartiles. This pattern is consistent with a precautionary motive, but one that appears especially salient for *granular* firmsthose whose large scale combined with greater market sensitivity makes their idiosyncratic shocks less diversifiable for investors. In this sense, the beta gradient in cash holdings among the largest firms is suggestive of an undiversifiable risk channel rather than a size effect alone.

A similar, though less pronounced, pattern appears for firms in the 90th–99th percentile of the size distribution. In contrast, no systematic relationship between beta and cash

[2]Throughout the main analysis, beta refers to the firm's equity (levered) beta, estimated from a regression of monthly excess stock returns—measured relative to the 1-month Treasury bill rate—on excess market returns. The regression uses a 60-month rolling window, requiring a minimum of 24 months of non-missing return observations. Since equity beta is mechanically influenced by financial leverage, I also conduct robustness checks using asset (i.e., unlevered) betas. The results are qualitatively unchanged.

holdings is observed among firms in the bottom 90 percent. These differences underscore that the connection between market sensitivity and liquidity is strongest for firms at the top of the size distribution whose shocks are most likely to be non-diversifiableprecisely the firms that exhibit granularity in the sense developed later in the model. The patterns also hold within industries: restricting the sample to manufacturing yields the same monotonic relationship between beta and liquidity (Table A2 in Appendix A.4).

Table 1: Firm Size and Market Beta, All sectors pooled

Size	Beta Quartile	Market Equity ($ M)	Beta	Cash-to-Assets	Leverage	Equity issuance	Debt issuance
	Q1	363	0.28	17.21	24.12	1.78	0.97
$< 90\%$	Q2	596	0.75	16.55	23.71	1.21	1.07
	Q3	712	1.08	17.43	23.78	1.23	1.19
	Q4	740	1.57	20.86	23.40	1.67	1.24
	Q1	18417	0.43	9.61	28.12	-0.79	1.51
$90\% - 99\%$	Q2	15337	0.77	9.80	25.89	-0.61	1.36
	Q3	12684	1.07	10.35	24.76	-0.44	1.33
	Q4	9050	1.51	15.31	23.32	0.01	1.64
	Q1	135235	0.46	10.31	25.15	-0.70	0.80
$> 99\%$	Q2	104470	0.74	10.24	23.19	-0.74	0.59
	Q3	125607	1.05	13.89	21.32	-0.95	0.90
	Q4	128254	1.44	18.85	14.73	-0.96	0.90

Notes: This table reports summary statistics for three firm size groups over the sample period 1970Q1-2019Q4. Within each size group, firms are sorted into quartiles based on their firm-level equity beta, reported in the second column. Column (3) presents the average market equity (in millions of dollars) for each size-beta subgroup. Columns (5) through (8) report average values of financial variables, expressed as percentages.

Together, these descriptive facts motivate the panel regression analysis that follows, which examines whether market beta predicts cash holdings after controlling for firm fundamentals and whether the strength of this relationship varies systematically with firm size.

2.2.1 Panel Regression Estimates

The unconditional patterns in Table 1 show that cash holdings increase with market beta primarily among the largest firms, consistent with the notion that scale and market sensitivity jointly give rise to undiversifiable risk exposures. To formally assess whether these relationships hold after accounting for firm characteristics and industry-level shocks, I now estimate panel regressions of cash holdings on market beta. I begin with the baseline specification:

$$\text{Cash/Asset}_{i,t} = \alpha_i + \alpha_{\text{industry},t} + \gamma \text{Beta}_{i,t-1} + \Gamma' Z_{i,t-1} + \varepsilon_{i,t} \tag{1}$$

where α_i denotes firm fixed effects and $\alpha_{\text{industry},t}$ captures industry-by-time fixed effects at the 2-digit SIC level. The vector $Z_{i,t-1}$ includes standard firm-level controls cited in the corporate finance literature as determinants of cash holdings.

Table 2 reports the regression results under three specifications. Column (1) presents a baseline specification without fixed effects, capturing the average effect of market beta on cash holdings. Column (2) adds industry-by-time fixed effects and firm controls, allowing for comparisons within industries over time while adjusting for firm-level characteristics. Column (3) further includes firm fixed effects, identifying the effect of time variation in beta within the same firm

Across all specifications, the coefficient on lagged market beta is positive and statistically significant, indicating a robust relationship between aggregate risk exposure and liquidity demand. In Column (2), the coefficient implies that a one standard deviation increase in market beta is associated with a 1.8 percentage point increase in the cash-to-assets ratio, holding other factors constant. This magnitude is economically meaningful: for a firm with an average cash-to-assets ratio of 20%, a 1.8 percentage point increase represents around 10% relative rise in cash holdings. This supports the hypothesis that aggregate risk exposure (as measured by market beta) is an important driver of corporate liquidity demand, above and beyond firm size, leverage, investment policy, and other conventional controls.

Given the evidence that market beta is a robust predictor of corporate cash holdings, I next examine whether this relationship is stronger among the largest firms–and whether it varies systematically with firm size and exposure to aggregate risk. To test this, I estimate the following panel regression, which includes interaction terms between market beta and

Table 2: Determinants of cash holdings

	(1)	(2)	(3)
L.Beta	2.267***	1.789***	0.735***
	(0.26)	(0.16)	(0.13)
L.Size		-0.158	-1.655***
		(0.10)	(0.28)
L.Tobin's Q		0.413***	0.163***
		(0.05)	(0.05)
L.Leverage		-0.123***	-0.091***
		(0.01)	(0.01)
L.Acquisition		-0.070***	-0.039***
		(0.01)	(0.01)
L.R&D Intenstiy		0.506***	-0.159***
		(0.02)	(0.03)
L.Distance to Default		0.635***	0.456***
		(0.04)	(0.03)
N	177519	177519	177519
R^2	0.095	0.378	0.723
Firm controls	no	yes	yes
Firm FE	no	no	yes
Time X Sector FE	yes	yes	yes

Notes: This table reports panel regression results examining the determinants of corporate cash holdings over the period 1980-2019. The dependent variable is the ratio of cash and short-term investments to total assets. The key independent variable is lagged firm-level market beta, standardized over the full sample. Firm-level controls include size (log of market equity), Tobin's Q, leverage (book debt over total assets), acquisition activity, R&D intensity, distance to default, cash flow ratio, cash flow volatility, sales growth (year-over-year), capital expenditures ratio (as a share of assets, asset tangibility), and firm age. Standard errors are two-way clustered at the firm and industry-quarter level. Industries are based on 2-digt SIC level.The full table results are shown in Appendix A3.

.

firm size group indicators:

$$\begin{aligned}\text{Cash/Asset}_{i,t} = \alpha_{\text{industry},t} + \sum_{g=1}^{5} \delta_g \, \mathbf{1}(\text{SizeGroup}_i = g) + \sum_{g=1}^{5} \gamma_g \left[\mathbf{1}(\text{SizeGroup}_i = g) \times \text{Beta}_{i,t-1}\right] \\ + \beta_2' Z_{i,t-1} + \beta_3 \left(Z_{i,t-1} \times \text{Beta}_{i,t-1}\right) + \varepsilon_{i,t}\end{aligned} \tag{2}$$

where firms are grouped into five size categories based on market capitalization percentiles at period t: bottom 50%, 50-75%, 75-90%, 90-99%, and top 1%. The regression includes industry-by-time fixed effects at the 2-digit SIC level to control for sector-specific macro shocks and business cycle exposure. This is important because certain sectors–such as

durable goods manufacturing, construction, or consumer discretionary–tend to have higher average market betas due to procyclical demand patterns, while others (e.g., utilities or consumer staples) exhibit more countercyclical behavior. By conditioning on these fixed effects, the analysis isolates variation within industries, allowing a more transparent comparison of firm behavior across the size distribution while controlling for industry-level cyclicality.

In this specification, the base group is firms in the bottom 50% of the size distribution. The coefficients δ_g captures level differences in average cash holdings across size groups when market beta is zero (i.e., at the standardized mean, which is around 1 in the sample). The interaction terms γ_g capture how the sensitivity of cash holdings to market beta varies across size groups. Specifically, it reflects how much more (or less) responsive firms in size group g are to market beta, relative to the base group. This interaction is economically meaningful: it allows for a test of whether exposure to aggregate risk has different implications for firms of different sizes, even when beta is held constant. For example, smaller firms may exhibit high beta due to volatile cash flows and substantial idiosyncratic risks, while large firms may have high beta because they are heavily integrated into macroeconomic dynamics. As such, the same level of beta may have different implications for financing constraints and liquidity policy across the firm size distribution in a given major sector[3].

The results, presented in Table 3, show that the responsiveness of cash holdings to market beta increases monotonically with firm size. The interaction coefficients γ_g are positive and statistically significant for all size groups above the median, with the effect most pronounced for the top 1% of firms. In Column (2), for example, the interaction term for the top 1% is 5.47, implying that a one standard deviation increase in beta is associated with a 5.47 percentage point increase in the cash-to-assets ratio–substantially higher than the baseline group. This supports that the precautionary savings motive in response to aggregate risk is

[3]From the specification, particularly in Column (2), I further interact each firm control with beta to ensure that the beta effect is not simply proxying for leverage, growth opportunities, payout policy, R&D intensity, or other covariates. If beta merely reflected one of these channels (for example, high-growth firms), then allowing, say, sales-growth × beta to soak up that channel would attenuate the top-1 percent interaction coefficient–but it does not.

Table 3: Heterogeneity in the Relationship between Cash Holdings and Market Beta

	(1)	(2)
L.Beta	1.901***	0.273
	(0.24)	(0.39)
Size $(50-75\%)$	-0.064	5.038***
	(0.47)	(0.43)
Size $(75-90\%)$	-3.798***	4.770***
	(0.59)	(0.66)
Size $(90-99\%)$	-4.876***	6.717***
	(0.63)	(0.86)
Size $(>99\%)$	-2.589*	10.397***
	(1.54)	(1.57)
$(50-75\%)\times$ L.Beta	0.535	1.675***
	(0.36)	(0.31)
$(75-90\%)\times$ L.Beta	1.623***	3.217***
	(0.45)	(0.45)
$(90-99\%)\times$ L.Beta	2.679***	4.793***
	(0.53)	(0.61)
$(>99\%)\times$ L.Beta	3.507**	5.467***
	(1.67)	(1.45)
L.Size		-1.266***
		(0.15)
L.Distance to Default (D2D)		0.525***
		(0.04)
L.Beta $\times$ L.D2D		0.061**
N	177738	177738
R^2	0.105	0.391
Firm controls	no	yes
Firm controls X Beta	no	yes
Time X Sector FE	yes	yes
Firm and Industry X Time Two-way clustering	yes	yes

Notes: This table reports panel regressions of the cash-to-assets ratio on market beta, size dummies, and their interactions. Column (1) includes only size dummies, beta, and size $\times$ beta interactions. Column (2) adds firm controls–log size, Tobin's Q, leverage, acquisition intensity, R&D intensity, distance-to-default, cash-flow ratio, cash-flow volatility, sales growth, capital expenditures (asset tangibility), and firm age–and interactions between each control and beta. All continuous variables are winsorized at the 1st and 99th percentiles; beta is standardized to mean 0, variance 1. Industry-by-quarter fixed effects are at the 2-digit SIC level; standard errors are two-way clustered at the firm and industry-quarter levels. The full coefficient matrix is reported in the Appendix A4.

disproportionately concentrated among the largest firms.

Before delving deeper into the precautionary-saving motives of granular firms, it is instructive to examine how aggregate-risk exposure shapes other dimensions of capital structure. To that end, I re-estimate the size-beta interaction specification using (i) total leverage and (ii) long-term debt-to-assets as the dependent variable (Table 4). Following Friewald,

Nagler and Wagner (2022), I decompose leverage into short-maturity (debt due within three years) and long-maturity (debt due beyond three years) components to capture potential differences in how creditors price firms' risk over different horizons.

Table 4: Capital structure: leverage and maturity

	(1) Leverage	(2) Leverage	(3) Long-term debt	(4) Long-term debt
L.Beta	-0.671***	-2.171***	0.131	-1.059***
	(0.24)	(0.47)	(0.21)	(0.40)
Size $(50-75\%)$	-0.200	-1.693***	4.409***	0.899**
	(0.50)	(0.49)	(0.42)	(0.43)
Size $(75-90\%)$	0.827	-2.293***	6.436***	0.165
	(0.65)	(0.77)	(0.54)	(0.69)
Size $(90-99)\%$	0.874	-5.276***	6.186***	-3.786***
	(0.74)	(1.10)	(0.57)	(0.93)
Size $(>99\%)$	-4.547***	-13.885***	1.682*	-12.953***
	(1.23)	(1.81)	(0.98)	(1.58)
$(50-75\%) \times$ L.Beta	0.473	0.046	-0.107	-0.124
	(0.39)	(0.38)	(0.34)	(0.34)
$(75-90\%) \times$ L.Beta	-0.580	-1.299**	-1.397***	-1.371***
	(0.48)	(0.54)	(0.43)	(0.50)
$(90-99)\%\times$ L.Beta	-1.978***	-3.035***	-1.846***	-2.041***
	(0.67)	(0.84)	(0.53)	(0.72)
$(>99)\%\times$ L.Beta	-2.635**	-3.040**	-2.494***	-2.151*
	(1.15)	(1.40)	(0.87)	(1.18)
L.Size		3.977***		3.949***
		(0.20)		(0.18)
L.Distance to Default		-1.906***		-1.252***
		(0.05)		(0.04)
N	180445	180445	170577	170577
R^2	0.042	0.282	0.061	0.232
Firm controls	no	yes	no	yes
Firm x Beta	no	yes	no	yes
Time X Sector FE	yes	yes	yes	yes
Firm FE	no	no	no	no
Firm and Industry x Time Two-way clustering	yes	yes	yes	

Notes: See Table 3 for detailed variable definitions. Columns (1) and (2) use the total debt-to-assets ratio; Columns (3) and (4) use the long-term debt-to-assets ratio (debt maturing after three years). Column (2) and (4) add the full set of firm controls and their interactions with beta. All continuous variables are winsorized at the 1st and 99th percentiles; beta is standardized to mean 0, variance 1. Industry-by-quarter fixed effects are at the 2-digit SIC level; standard errors are two-way clustered at the firm and industry-quarter levels. The complete results appear in Appendix A5.

The estimates in Table 4 confirm that, as with cash holdings, the sensitivity of both total leverage and long-term leverage to market beta declines monotonically with firm size. In the fully controlled specification (Columns 2 and 4), the interaction term for the top 1 percent is -3.04 in total leverage (Column 2) and -2.15 in long-term debt (Column 4), each significant at the 5 percent level. At first glance, this result is counterintuitive. Because equity beta is mechanically linked to leverage–higher debt amplifies equity volatility–one would normally

expect high-beta firms to carry more debt. However, I find the opposite among the largest firms: a one-standard-deviation increase in beta is associated with roughly a 3 percentage-point decline in total leverage and a 2.2 point decline in long-term debt for the top 1 percent. This negative beta-leverage relationship indicates that, rather than levering up to ride their high-beta status, the very largest firms de-lever, especially at the long end of the maturity spectrum, as their exposure to systematic risk rises.

Combined with our earlier finding that high-beta large firms also accumulate more cash, these results paint a consistent picture: granular firms self-insure by reducing external debt and building internal liquidity when their exposure to systematic risk rises. This behavior is surprising, since large, high-beta firms would ostensibly face lower borrowing costs and easier credit access. Instead, it suggests that these firms confront unique risk factors–perhaps related to their systemic importance or tail-risk exposure–that induce a stronger precautionary-saving motive.

In the next section, I investigate this hypothesis more directly by examining how granular firms adjust their cash buffers in response to fluctuations in the financial market, particularly during episodes of heightened market volatility and investor risk aversion.

2.2.2 Evidence on self-insurance motives among granular firms

If firms actively respond to heightened borrowing costs in volatile markets by strengthening their balance sheets–accumulating cash and curtailing external debt–this behavior would provide direct evidence of a precautionary self-insurance motive. Although balance-sheet adjustments are reported only quarterly and thus may appear sluggish, the cumulative changes from one quarter to several ahead still reflect firms' strategic financial policies during episodes of market distress. To test this channel, I estimate the following panel regression:

$$\begin{aligned}\Delta \log(\text{Debt or Cash})_{t-1,t+h} = \; & \alpha_{\text{industry},t} \; + \; \sum_{g=1}^{3} \delta_g \, \mathbf{1}(\text{Size}_i = g) \; + \; \sum_{g=1}^{3} \gamma_g \, \mathbf{1}(\text{Size}_i = g) \; \times \; \text{Beta}_{i,t-1} \\ & + \; \sum_{g=1}^{3} \beta_g \, \mathbf{1}(\text{Size}_i = g) \; \times \; \text{Beta}_{i,t-1} \; \times \; \Delta \log \text{VIX}_{t-1} \\ & + \; \Gamma_1' Z_{i,t-1} \; + \; \Gamma_2' \Big(Z_{i,t-1} \times \Delta \log \text{VIX}_{t-1} \Big) \; + \; \varepsilon_{i,t} \, . \end{aligned} \tag{3}$$

where the dependent variable is the cumulative, industry-deflated growth rate of cash (or debt) from quarter $t-1$ to $t+h$, i.e, $\Delta \log(\text{Cash})_{t-1,t+h}$ is defined as $\log(\text{Cash})_{t+h} - \log(\text{Cash})_{t-1}$ and analogously for debt.[4] Size groups $g = 1, 2, 3$ partition the sample into bottom 90 percent, 90-99 percent, and top 1 percent of market capitalization, respectively. The key interaction β_g measures whether, conditional on size group and beta, an increase in market volatility (proxied by $\Delta \log VIX_{t-1}$) amplifies or dampens subsequent cash or debt growth. γ_g captures how the baseline beta-cash (or beta-debt) sensitivity differs across size groups. β_g tests whether this sensitivity is further modulated by contemporaneous spikes in aggregate volatility. $Z_{i,t-1}$ includes the usual firm-level controls (size, Tobin's Q, leverage, R%D intensity, tangibility, age) and their interactions with $\Delta \log VIX_t$.

The triple-interaction for the top 1% in the cash-growth regression is large, positive, and significant at all horizons, indicating that high-beta granular firms boost cash balances sharply when volatility spikes. The corresponding triple-interaction in the debt-growth regressions is large and negative for the top 1%, showing that these firms concurrently curtail debt issuance (or paydown debt) in response to the same shocks. The magnitude of both effects grows with horizon, illustrating that precautionary self-insurance is not just a one-quarter reaction but a sustained balance-sheet adjustment. Neither the 90-99% nor the bottom 90% groups exhibit a comparable pattern: their triple-interaction coefficients are small, statistically indistinguishable from zero, or even opposite in sign.

[4]All dollar-value series are deflated by using the Implied Price Index of Gross Value Added in the U.S. Nonfarm Business Sector

Table 5: Precautionary saving and deleveraging

	h=0		h=4		h=9	
$\Delta \log(y_{i,t+h})$	Cash	Total Debt	Cash	Total Debt	Cash	Total Debt
(90-99)% $\times\Delta$ LVIX	0.928	6.213*	-5.512	4.499	-10.227	4.733
	(4.94)	(3.58)	(5.89)	(6.38)	(7.98)	(7.64)
$>$ 99% $\times$ Δ LVIX	-9.342	13.748**	-33.601***	20.110**	-36.943**	19.846
	(6.54)	(5.68)	(10.30)	(9.60)	(17.10)	(13.37)
L.Beta $\times\Delta LVIX$	-3.368**	-0.463	-5.032**	2.650	-1.581	0.710
	(1.50)	(1.42)	(2.30)	(2.52)	(3.80)	(3.67)
(90-99)% $\times$ L.Beta $\times\Delta$ LVIX	2.875	-3.806	6.933	1.476	10.873*	1.606
	(4.08)	(3.46)	(4.84)	(6.00)	(6.05)	(6.87)
$>$ 99%$\times$ L.Beta $\times\Delta$ LVIX	12.665**	-18.119***	28.685***	-23.486***	31.671*	-26.146**
	(5.83)	(5.44)	(9.10)	(8.32)	(18.06)	(12.91)
N	300841	244927	263610	217492	228651	192403
R^2	0.016	0.018	0.023	0.053	0.032	0.087
Firm controls	yes	yes	yes	yes	yes	yes
Firm $\times\Delta LVIX$	yes	yes	yes	yes	yes	yes
Time X Sector FE	yes	yes	yes	yes	yes	yes

Notes: Table 5 reports estimates of firms' cumulative cash- and debt-growth responses to a spike in market volatility (Δlog VIX), interacted with market beta and size group. Columns (1)-(2) show the horizon $h = 0$ results, (3)-(4) the $h = 4$ horizon, and (5)-(6) the $h = 9$ (nine-quarter) horizon. In each pair, the first column is the cash-growth regression and the second is total debt-growth. Other controls include size, age, Tobin's Q, leverage, tangibility, cash flow volatility, sales growth (year-over-year) and their interaction with $\Delta \log VIX_{t-1}$.Industry-by-quarter fixed effects are at the 2-digit SIC level; standard errors are two-way clustered at the firm and industry-quarter levels. The complete results appear in Appendix need to add.

Economically, a positive β_g in the cash equation means that high-beta firms in size group g increase cash growth more when volatility rises, over and above the baseline effect in the smallest firms. Conversely, a negative β_g in the debt equation implies that these same firms reduce debt growth. The fact that these patterns are concentrated among the top 1%- and grow over multi-quarter horizons-provides strong evidence that granular firms engage in precautionary self-insurance: they build internal liquidity and deleverage when external financing becomes riskier and more expensive.

In the next section, I provide more direct evidence that the largest firms indeed face higher borrowing costs in the corporate secondary bond market, as measured by credit spreads. To identify the drivers of these costs, I follow the decomposition of Gilchrist and Zakrajšek (2012), which separates the raw credit spread into expected default loss, capturing the firm's fundamental probability of default; and excess bond premium, reflecting the additional compensation financial intermediaries demand due to their limited risk-bearing capacity. This

decomposition allows us to determine whether high spreads for granular firms arise from their intrinsic default risk or from an extra risk premium required by investors. Applying this methodology across our firm-size groups thus reveals how investors differentially price the borrowing costs of small, mid-sized, and the very largest firms.

2.3 Inspecting the mechanism: Corporate Bonds Market

In order to examine the heterogeneity in borrowing costs across firms, I construct daily credit spreads on senior, unsecured, fixed-coupon U.S. corporate bonds from TRACE over August 2002-December 2019.[5] Following Gilchrist and Zakrajšek (2012) (GZ), I then construct a credit spread for each bond on each trading day as the difference between the bond's yield-to-maturity implied by its daily price and the yield-to-maturity of a synthetic risk-free security obtained by discounting the bond's contractual cash flows at the zero-coupon Treasury term structure (interpolated from Gürkaynak, Sack and Swanson, 2007). In the Appendix A.3, I report summary statistics of bond characteristics in our sample and compare the resulting credit-spread series to the excess bond premium of Gilchrist and Zakrajšek (2012)–constructed from the publicly available BofA-ML index on FRED–to confirm that our TRACE-based measure replicates the key dynamics of the canonical GZ series.

Summary statistics by firm-size group are reported in Table 6. As expected, the largest firms issue the greatest volume of debt, enjoy the lowest yields and credit spreads, and tap the market for the longest maturities: The average number of distinct bonds traded per firm-month rises from 1.44 in the bottom 50% to 12.39 in the top 1%. Regarding cost of funds, effective yields and credit spreads both decline monotonically with size, from 10.8% and 8.93% in the smallest firms to 3.55% and 1.30% in the largest. Large issuer bonds trade more often, averaging 15.9 trading days per month in the top 1% versus 12.0 days for the bottom 50%.

[5]Bond data are cleaned following Dickerson, Mueller and Robotti (2023) and matched to Compustat/CRSP via https://openbondassetpricing.com/. I restrict the sample to only U.S.-domiciled, public, fixed-coupon, non-convertible, non-asset-backed issues (ex. 144A, equity-linked, agency-backed, MBS)

Table 6: Summary statistics by size group

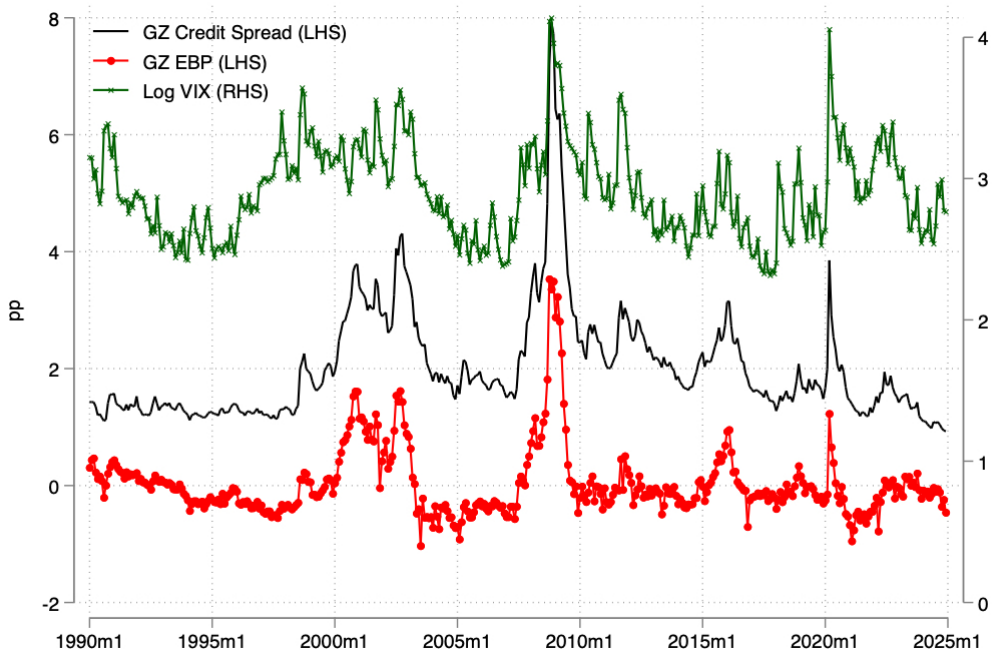
Mean	Size < 50%	50-75	75-90	90-99	>99
No. of Bonds per Firm/Month	1.44	1.88	2.65	5.92	12.39
Effective Yield (%)	10.80	7.09	5.35	4.25	3.55
Spread (%)	8.93	5.11	3.27	1.94	1.30
Coupon (%)	7.90	6.93	6.06	5.33	4.68
Amount Issued ($ Mills)	342.43	386.32	426.50	619.48	1082.32
Maturity At Issue (Years)	9.19	10.12	11.82	14.96	16.18
Time to Maturity (Years)	5.62	6.49	7.80	10.55	11.76
Duration	4.02	4.67	5.42	6.82	7.58
Credit Rating	BBB3	BBB1	A2	A2	A2
Callable (%)	100	99	98	97	95
No. trading days/Month	11.96	12.41	12.21	12.83	15.86

Notes: Sample covers August 2002-December 2019; 377,880 firm-month observations from 1,029 non-financial, non-utility, non-quasi-government issuers and 7,376 bonds. Size groups are defined by market-capitalization percentiles each quarter: bottom 50%, 50-75%, 75-90%, 90-99%, and top 1%. Credit spreads are yield-to-maturity over the Treasury curve; effective yields and coupons are annual rates. Credit ratings convert S&P' s letter grades to numeric scale (AAA=1, C=21); Median credit rating is reported within each group. Trading days count distinct TRACE reporting days within each month.

Figure 2 plots the GZ credit spread and its excess-bond-premium (EBP) component on the left axis, together with the log of the VIX index on the right axis, over January 1990-January 2025 (including our TRACE sample period). Peaks in VIX coincide with large jumps in both the aggregate credit spread and the EBP component. Even outside major crises such as the 2008-2009 Global Financial Crisis and the Covid-19, the three series exhibit pronounced comovement: when implied equity volatility rises, credit spreads and the EBP tend to increase as well. The EBP series (residual spread) often spikes proportionally in the most acute stress episodes, underscoring that limited financial-intermediary risk-bearing capacity amplifies corporate borrowing costs beyond what default probabilities alone would predict. These regularities confirm that periods of heightened market volatility are precisely the times when firms face materially higher external borrowing costs and greater difficulty rolling or extending debt.

In the next section, I formally test whether high-beta granular firms–those in the top 1 percent of the size distribution–face disproportionately steeper borrowing costs in the

Figure 2: Credit Spreads and Excess Bond premiums



Notes: The panel plots GZ spread and its residual component, excess bond preimum, on the left axis. For comparison, log VIX is plotted on the right axis. GZ spread is the average (equal-weighted) credit spread on senior unsecured bonds issued by nonfinancial firms while GZ EBP measures the component of corporate bond credit spreads not explained by expected default risk, capturing shifts in investor risk appetite.

secondary bond market. By quantifying the sensitivity of credit spreads and excess bond premia to firm-level beta and size, I assess whether the elevated cost of debt for these firms provides the economic incentive behind their pronounced cash accumulation. Combined with the preceding evidence on cash holdings and deleveraging, this analysis will clarify whether the highest-beta, largest firms are indeed self-insuring against the outsized risk premiums they incur when tapping external finance.

2.3.1 Evidence on higher borrowing cost for high beta granular firms

I begin by showing that firm-level equity beta predicts corporate bond spreads and that this sensitivity is amplified when the aggregate bond market becomes more risk-averse, as captured by the excess bond premium (EBP). To that end, I decompose the monthly change in bond i's spread for issuer j as

$$\Delta cs_{ij,t} = \Delta \widehat{cs}_{ij,t} + \Delta ebp_{ij,t} \tag{4}$$

where $\Delta cs_{ij,t} = cs_{ij,t} - cs_{ij,t-1}$ is the observed change in the bond's credit spread; $\widehat{cs}_{ij,t}$ is the fitted spread from a structural "distance-to-default" model (á la Merton-KMV), controlling for bond characteristics such as callable, duration, maturity, age, issuance, and coupon rate; $ebp_{ij,t}$ is the residual component-i.e, the excess bond premium-which captures the portion of spread movements not explained by fundamentals and reflects shifts in intermediary risk-bearing capacity.

I run the following panel regression:

$$\Delta cs_{ij,t} = \alpha_i + \beta_{\text{industry}_j,t} + \gamma(\text{Beta}_{j,t-1} \times \Delta EBP_t) + \delta \text{Beta}_{j,t-1} + \varepsilon_{ij,t} \tag{5}$$

Table 7 presents evidence that firms with higher market beta face steeper borrowing costs precisely when aggregate risk premia rise. In Column (1), the coefficient on the interaction between lagged beta and contemporaneous Δ EBP is 0.2545, implying that a one-percentage-

Table 7: Sensitivity of Bond Spreads to Aggregate Risk Exposure

	(1) $\Delta cs_{ij,t}$	(2) $\Delta ebp_{ij,t}$	(3) $\Delta \widehat{cs}_{ij,t}$	(4) $\Delta cs_{ij,t}$	(5) $\Delta ebp_{ij,t}$	(6) $\Delta \widehat{cs}_{ij,t}$
L.Beta	0.0030	0.0028	-0.0012***	0.0031	0.0027	-0.0013***
	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
L.Beta $\times \Delta EBP_t$	0.2545***	0.2575***	-0.0301***	0.2517***	0.2543***	-0.0300***
	(0.04)	(0.03)	(0.00)	(0.04)	(0.03)	(0.00)
L.Size				-0.0065	-0.0063*	-0.0001
				(0.00)	(0.00)	(0.00)
L.Leverage				0.0631**	0.0576**	-0.0117***
				(0.03)	(0.03)	(0.00)
N	331756	327125	327125	330277	325670	325670
R^2	0.373	0.379	0.885	0.373	0.380	0.886
Firm controls	no	no	no	yes	yes	yes
Time X Sector FE	yes	yes	yes	yes	yes	yes
Bond FE	yes	yes	yes	yes	yes	yes

Notes: Each column reports estimates of $\Delta cs_{ij,t} = \alpha_i + \beta_{\text{industry}_j,t} + \gamma(\text{Beta}_{j,t-1} \times \Delta EBP_t) + \delta \text{Beta}_{j,t-1} + \Gamma_1' Z_{i,t-1} + \varepsilon_{ij,t}$ where $\Delta cs_{ij,t}$ is the change in credit spread for bond j of firm i, and $Z_{i,t-1}$ includes log size, leverage, debt-refinancing intensity, firm age, and year-over-year sales growth and its interaction with EBP. All specifications include bond fixed effects, credit-rating fixed effects, and time x sector dummies at the 3-digit SIC level. Standard errors (in parentheses) are two-way clustered at the firm and industry-time level.

point increase in the aggregate excess bond premium is associated with an additional 25 basis-point widening of a high-beta firm's credit spread. Columns (2)-(3) confirm that this effect is almost entirely driven by the EBP component of spreads (rather than expected default losses).

Columns (4)-(6) show that these results are robust to adding standard firm-level controls (size, leverage, debt-refinancing intensity, age, and sales growth) and their own interactions with Δ EBP. The interaction term coefficient remains essentially unchanged (0.2517 in the full-spread regression), and continues to load almost exclusively on the excess-premium margin. Together, these findings demonstrate a tight link between aggregate intermediary stress and firm-specific borrowing costs–one that is amplified for firms whose equity valuations are most sensitive to the market.

Crisis Episodes: Heterogeneous transmission of investor risk aversion The preceding results indicate that, on average, firms whose equity valuations are more sensitive

to market movements pay higher borrowing costs–driven primarily by the excess-premium component of credit spreads. I now ask whether this effect is especially pronounced for the largest, granular firms during periods of extreme market distress. To test this, I re-estimate the bond-spread regression with additional interactions for size groups, focusing exclusively on the height of the Global Financial Crisis (2007 - 2009), when aggregate EBP surged by more than two percentage points:

$$\begin{aligned} cs_{ij,t+h} - cs_{ij,t-1} = \alpha_i + \beta_{\text{industry}_{j,t}} + \sum_{g=1}^{3} \delta_g \mathbf{1}(\text{Size Group}_j = g) \times \Delta EBP_t \\ + \sum_{g=1}^{3} \gamma_g \big(\mathbf{1}(\text{Size Group}_j = g) \times \text{Beta}_{j,t-1} \times \Delta EBP_t \big) \\ + Z'_{i,t-1}\beta_2 + (Z'_{j,t-1} \times \Delta EBP_t)\beta_3 + \varepsilon_{ij,t} \end{aligned} \tag{6}$$

where size groups $g = 1, 2, 3$ index the bottom 90 %, 90-99 %, and top 1 % of firms by market capitalization, respectively. By conditioning on the GFC window, I isolate episodes when financial-intermediary risk bearing was severely impaired.[6]

It is well known that bonds of different maturities expose the issuer to distinct risks and serve different financing purposes. Short-term instruments (< 3 years) primarily carry rollover risk–firms must refinance or repay shortly after issuance–whereas long-term bonds (≥ 10 years) fund longer-horizon investments and lock in funding costs over an extended period (e.g., He and Xiong, 2012; Friewald, Nagler and Wagner, 2022). Because our core question concerns why firms would substitute away from cash and short-term liquidity into marketable assets, the most relevant margin is the trade-off between holding cash and issuing long-dated debt. Accordingly, in what follows, I restrict the sample to senior unsecured bonds with original maturities of ten years or more. Focusing on this long-term segment isolates the strategic decision to raise durable external funds in lieu of internal cash buffers.

[6] I focus on the 2007-2009 window because our TRACE-based sample begins in mid-2002, missing the full dot-com collapse, and because aggregate EBP only spiked meaningfully during the GFC. It rose again in early 2020, and extending the sample through COVID-19 downturn (ending just before March 2020 liquidity facilities) yields the similar results.

Table 8: Granular Firm's Bond-Market Risk Premia

Long-term maturity bonds (≥10yrs)	h=3			h=6		
	$\Delta cs_{ij,t+h}$	$\Delta ebp_{ij,t+h}$	$\Delta \widehat{cs}_{ij,t+h}$	$\Delta cs_{ij,t+h}$	$\Delta ebp_{ij,t+h}$	$\Delta \widehat{cs}_{ij,t+h}$
(90-99)% $\times\Delta$EBP$_t$	-0.517**	-0.548*	0.034	-0.557	-0.667	0.133
	(0.25)	(0.29)	(0.06)	(0.55)	(0.63)	(0.10)
$>$ 99% $\times$ ΔEBP$_t$	-0.831***	-0.858**	0.035	-0.735	-0.866	0.136
	(0.30)	(0.34)	(0.07)	(0.64)	(0.72)	(0.12)
L.Beta $\times\Delta$EBP$_t$	-0.317	-0.409	0.093*	-1.035**	-1.227***	0.185***
	(0.21)	(0.26)	(0.05)	(0.40)	(0.45)	(0.08)
(90-99)% $\times$L.Beta $\times$ ΔEBP$_t$	0.587**	0.691**	-0.102*	1.236***	1.435***	-0.188***
	(0.26)	(0.31)	(0.05)	(0.44)	(0.48)	(0.04)
$>$ 99% $\times$ L.Beta $\times$ ΔEBP$_t$	0.438**	0.536**	-0.097	1.446***	1.759***	-0.327***
	(0.22)	(0.27)	(0.06)	(0.45)	(0.50)	(0.07)
N	3375	3356	3356	3322	3300	3300
R^2	0.816	0.745	0.985	0.855	0.720	0.980
Firm controls	yes	yes	yes	yes	yes	yes
Firm controls $\times\Delta$EBP$_t$	yes	yes	yes	yes	yes	yes
Bond FE	yes	yes	yes	yes	yes	yes
Time X Sector FE	yes	yes	yes	yes	yes	yes

Notes: Table 8 shows the estimates for long-term ($\geq$ 10 years) senior unsecured bonds during the GFC window (2007m9-2009m6). $Z_{i,t-1}$ includes log size, leverage, debt-refinancing intensity, firm age, and year-over-year sales growth and its interaction with EBP. All specifications include bond fixed effects, credit-rating fixed effects, and time x sector dummies at the 3-digit SIC level. Standard errors (in parentheses) are two-way clustered at the firm and industry-time level. Firm-level market beta is standardized (mean 0, standard deviation 1) over the 2002-2019 sample period. The results are presented for selected horizons $h = 3$ and $h = 6$ for readability. Full results are reported in the Appendix. need to add

Table 8 reports the estimated response of long-term bond spreads to a one-percentage-point increase in the aggregate excess bond premium (EBP) during the Global Financial Crisis. In the three-quarter horizon (columns 1-3), high-beta firms in the top 1 percent of the size distribution experience an additional 44 basis point widening of their credit spread, of which 54 basis points materialize in the excess-premium component and only 10 basis points dissipate from the expected default-risk margin. By the six-quarter horizon (columns 4-6), these effects amplify: the total spread increase reaches 1.45 points, with 1.76 points in the excess-premium residual. At longer horizons, these coefficients substantially exceed those estimated for the 90-99 percentile firms, underscoring that the "granular firms bear a disproportionately large risk premium when market-wide risk aversion peaks. This persistent, outsized sensitivity of borrowing costs to aggregate EBP helps explain why the largest, most market-exposed firms accumulate unusually large cash buffers in volatile periods.

Taking Stock of the Empirical Evidence

1. A pronounced run-up in cash holdings among the largest U.S. public firms.

2. Within this group, higher-beta firms hold more cash and use less leverage, reflecting greater exposure to undiversifiable risk.

3. Granular firms dynamically increase cash and reduce debt issuance when aggregate volatility (VIX) rises.

4. Their credit spreads are disproportionately sensitive to credit-market stress, particularly through the excess bond premium in crisis episodes.

These facts motivate the structural framework that follows. While the empirical analysis identifies the interaction of size and market sensitivity (beta) as the key predictor of liquidity and leverage behavior, this reduced form reflects a deeper mechanism: for sufficiently large firms, idiosyncratic shocks are no longer fully diversifiable. In the model, the

interaction of firm size and idiosyncratic risk creates undiversifiable exposures that investors price, generating a firm-specific price-impact wedge in corporate bond markets. This wedge–an endogenous risk premium–steepens the bond pricing schedule and shapes granular firms liquidity and leverage choices.

3 Model

To rationalize the empirical evidence on why large ("granular") issuers simultaneously stockpile cash and avoid long-maturity leverage, I build a tractable corporate financing model. Firms choose between (i) issuing defaultable long-term debt and (ii) saving in a risk-free liquid asset ("cash") to smooth payouts in the face of idiosyncratic cash-flow risk. On the liability side, default follows the standard structural approach used in corporate credit models (e.g., Kuehn and Schmid, 2014; Gomes and Schmid, 2021), abstracting from physical capital and investment adjustment. On the asset-pricing side, a risk-averse investor prices bonds competitively but is imperfectly diversified because a finite set of large issuers accounts for a non-negligible share of outstanding corporate debt. This limited risk-bearing capacity makes the bond demand curve for granular issuers much steeper, as their idiosyncratic shocks impose undiversifiable risk on the investor and therefore require additional compensation.[7] When a granular firm issues more debt, the additional risk premium embedded in its bond spread rises steeply, imposing a marginal cost that is substantially higher than that faced by smaller, effectively price-taking issuers.

[7]When corporate debt supply is concentrated in a few large issuers, issuer-specific supply shocks (through issuance and default risk) cannot be fully diversified by the marginal bondholder. As a result, the covariance between a granular issuers bond payoff and the investors stochastic discount factor becomes endogenous to that issuers own supply, generating a wedge between risk-neutral and actual borrowing costs. This mechanism parallels the intuition behind the macro granularity literature (e.g., Gabaix, 2011), in that firm size makes idiosyncratic shocks partly non-diversifiable. However, for tractability, the model abstracts from general-equilibrium spillovers or aggregate feedback effects and focuses solely on the issuer-level price impact created by imperfect investor diversification.

3.1 Setup

Time is discrete and indexed by $t \in \{0, 1, ...\}$, with each period representing one quarter. There is no aggregate uncertainty; all risk is idiosyncratic. At the beginning of the period t, each firm i observes its realized cash flow Y_{it} and then decides whether or not to default. If it defaults, equity is wiped out, creditors liquidate, and the firm exits. If it continues, then the firm services its outstanding long-term debt (coupon and amortization) and pays a fixed operating cost. Next, the firm chooses (i) net issuance of defaultable long-term debt I_{it} in a competitive bond market, (ii) next-period cash $A_{it+1} \geq 0$ and (iii) dividends $D_{it} \geq 0$ (no external equity issuance).

3.1.1 Firm problem

The economy's production sector comprises two ex-ante heterogeneous types of firms:

1. A finite number of *granular* firms $i \in N << \infty$
2. A continuum of *competitive fringe* firms $j \in [0, 1]$

By nature, the fringe firms are atomistic and thus price takers, while the granular firms are monopsonists in the bond market. I characterize price impact for granular issuers in Section 3.1.2.

Preferences The firm is modeled as having risk-averse preferences over intertemporal cash flows, represented by a strictly concave utility function $u(\cdot)$ and a discount factor $\beta \in (0, 1)$. This reduced-form preference captures the idea that the firm values a smoother stream of dividend payments over time and is averse to fluctuations in cash flows.[8]

[8]While it is common to assume that firms are owned by fully diversified shareholders and thus behave as risk-neutral, empirical evidence suggests that many large firms act as if they are risk averse. This assumption can be interpreted as a reduced-form way to capture several institutional features: (i) managerial risk aversion due to compensation contracts or career concerns (e.g., Panousi and Papanikolaou, 2012); (ii) a desire to smooth cash flows over time; or (iii) ownership concentration or under-diversification among large shareholders (e.g., Greenwald and Stiglitz, 1990). Amel-Zadeh, Kasperk and Schmalz (2022), for example, document that nearly one-fifth of the largest U.S. firms have a nonfinancial blockholder or insider as their largest shareholder. These features provide empirical motivation for modeling corporate financial decisions with risk-averse preferences.

Limited liability implies that at the beginning of each period t, shareholders can choose to default if the firm's continuation value is negative. Let $D_{i,t}$ denote dividends paid to equity holders in period t. The equity value of firm i at time t is given by the Bellman equation

$$V_{i,t} = \max\Big\{0,\ \max\big[u(D_{i,t}) + \beta\, \mathbf{E}_t\big[V_{i,t+1}\big]\big]\Big\}, \tag{7}$$

where $\mathbf{E}_t[\cdot]$ denotes the conditional expectation given information at time t. The outer maximum captures the option to default: if the optimal continuation value inside the brackets is negative, shareholders optimally default and receive zero.

Endowment process Each firm i receives an exogenous cash-flow endowment Y_{it} which evolves according to:

$$Y_{it} = Y_{it}^P Y_{it}^T \tag{8}$$

$$\ln Y_{it}^P = \ln Y_{it-1}^P + \eta_{it} \quad \text{, with } \eta_{it} \sim i.i.d\, N(-0.5\sigma_\eta^2, \sigma_\eta) \tag{9}$$

$$Y_{it}^T = 1 + \varepsilon_{it} \quad \text{, with } \varepsilon_{it} \sim i.i.d\, N(\mu_\varepsilon, \sigma_\varepsilon) \tag{10}$$

where Y_{it}^P is the persistent component, Y_{it}^T the transitory component[9].

Securities. Firms may hold a one-period risk-free liquid asset (cash) priced at q_a^τ, implying a gross return of $r_a^\tau \equiv 1/q_a^\tau$.[10] Firms face a non-negativity constraint on cash holdings: $Ai, t+1 \geq 0$; secured borrowing against cash is not allowed.

Debt Structure Long-term corporate debt follows the exponential maturity formulation commonly used in sovereign-debt and corporate-finance models (e.g., Hatchondo and Mar-

[9] A multiplicative permanent-transitory decomposition with lognormal components in the cash flow process follows the recent corporate-finance literature that emphasizes the role of both temporary and permanent shocks to firms' cash flows. See, e.g., Poeschl (2023) for a recent application.

[10] The return on cash is net of corporate taxes, reflecting the assumption that interest income is taxable. This creates a tax-induced incentive for firms to borrow due to the associated interest deductibility (i.e., the debt tax shield).

tinez, 2009; Bianchi, Hatchondo and Martinez, 2018; Gomes and Schmid, 2021). Let B_{it} denote face value outstanding at the start of t. In each period, a fraction $\delta \in (0, 1]$ of principal amortizes and a per-unit coupon $c \geq 0$ is due. The law of motion with net issuance I_{it} (chosen after servicing) is

$$B_{i,t+1} = (1 - \delta)B_{it} + I_{it}. \tag{11}$$

Dividends and Cash Flows At the end of period t, dividends D_{it} are paid out of after-tax operating cash flows, net of fixed costs, cash accumulation, and debt service. The firms flow-of-funds constraint is

$$D_{it} = (1 - \tau)\big(Y_{it} - \Psi_{it}\big) + A_{it} - q_a^{\tau} A_{it+1} - \big[(1 - \tau)c + \delta\big]B_{it} + q_t I_{it} - \Phi(I_{it}). \tag{1}$$

Here, Y_{it} is operating cash flow before taxes and fixed costs, and Ψ_{it} is a fixed operating cost. The variable A_{it} denotes the beginning-of-period stock of risk-free liquid assets (cash), and q_a^{τ} is the after-tax price of next periods cash holdings; the firm faces the constraint $A_{it+1} \geq 0$ and cannot borrow against cash. The firm pays corporate income tax at rate $\tau \in [0, 1)$ on operating income net of fixed costs. Consistent with the U.S. tax code, interest payments receive a tax deduction through the term $(1 - \tau)cB_{it}$, while δB_{it} reflects the amortization of principal.[11]

The fixed operating cost Ψ_{it} is assumed proportional to the permanent component of income, $\Psi_{it} = \psi Y_{it}^{P}$ with $\psi > 0$, which ensures stationarity of the detrended problem. This fixed cost can induce firms to exit endogenously, even when unlevered, and helps match the average level of cash holdings in the data.

Long-term debt issuance is denoted I_{it}, with adjustment costs $\Phi(I_{it}) = \psi_1 |I_{it}| + \psi_2 I_{it}^2$ applied symmetrically to issuance and retirement. These costs capture frictions in bond markets and help replicate empirical patterns in firms' issuance behavior.

[11] Interest deductibility is a standard feature of corporate-debt models and generates the classic tax advantage of debt.

3.1.2 Investor Block: Downward-sloping Demand

To build intuition for why granular firms face steeper borrowing costs, I begin by writing the standard asset-pricing equation for a long-term corporate bond. For any bond issued by firm i, competitive investors require a price that satisfies

$$q_{it} = \mathbf{E}_t\big[M_{t,t+1} R_{i,t+1}\big], \tag{12}$$

$$R_{i,t+1} \equiv \big[1 - \mathbf{1}(\text{default}_{i,t+1})\big]\big[c + \delta + (1-\delta) q_{i,t+1}\big], \tag{13}$$

where $R_{i,t+1}$ is the state-contingent payoff and $M_{t,t+1}$ is the investors stochastic discount factor (SDF).

The Euler equation can be decomposed into a risk-neutral component and a risk-adjustment wedge:

$$q_{it} = \frac{\mathbf{E}_t[R_{i,t+1}]}{r_{f,t}} + \underbrace{\text{Cov}_t\big(M_{t,t+1}, R_{i,t+1}\big)}_{\text{endogenous risk premium}}, \tag{14}$$

where $r_{f,t}^{-1} = \mathbf{E}_t[M_{t,t+1}]$. The first term is the risk-neutral price. The second captures the extra compensation investors require for holding payoffs that co-move with marginal utility.

For typical "fringe" firms, whose idiosyncratic shocks are fully diversifiable, $\text{Cov}_t(M_{t,t+1}, R_{i,t+1})$ is close to zero (apart from any exposure to aggregate shocks). By contrast, for a granular firm that accounts for a non-negligible share of investors portfolios, idiosyncratic shocks affect the investors wealth and thus load into the SDF. When bond payoffs tend to be low precisely in high-$M_{t,t+1}$ states, $\text{Cov}_t(M_{t,t+1}, R_{i,t+1})$ becomes more negative, so that the bond trades further below its risk-neutral price and the required spread rises. Moreover, as a granular firm issues additional debt, the investors exposure to its undiversifiable risk increases, making this covariance term even more negative in magnitude. This steepens the issuer-specific pricing schedule: each extra unit of issuance pushes the price further below the risk-neutral benchmark. I refer to this endogenous gap between the actual bond price

and its risk-neutral counterpart as the *price-impact wedge*.

In the next subsection, I formalize the investors portfolio problem and introduce an approximation that preserves this endogenous price-impact channel while keeping the model tractable.

Investors Problem A representative risk-averse investor with wealth W_t allocates across the N granular firms long-term bonds and a diversified *fringe index* representing the continuum of atomistic issuers.[12] Let α_{it} denote the share of wealth invested in granular firm i's bond, and let

$$\alpha_{Ft} = 1 - \sum_{i=1}^{N} \alpha_{it}$$

be the share invested in the fringe index.

Each asset delivers a gross return

$$r_{i,t+1} = \frac{R_{i,t+1}}{q_{it}}, \qquad r_{F,t+1} = \frac{R_{F,t+1}}{q_{Ft}},$$

where

$$R_{i,t+1} = \big[1 - \mathbf{1}\{\text{default}_{i,t+1}\}\big]\big(c + \delta + (1-\delta)q_{i,t+1}\big), \qquad R_{F,t+1} = c + \delta + (1-\delta)q_{F,t+1}.$$

The investor chooses consumption c_t and portfolio weights $\{\alpha_{it}\}_{i=1}^{N}$ to solve

$$V(W_t) = \max_{c_t, \{\alpha_{it}\}} \frac{c_t^{1-\gamma} - 1}{1-\gamma} + \beta_L \, \mathbb{E}_t V(W_{t+1}), \tag{15}$$

$$\text{s.t.} \qquad W_{t+1} = r_{p,t+1}\big(W_t - c_t\big), \tag{16}$$

$$r_{p,t+1} = r_{F,t+1} + \sum_{i=1}^{N} \alpha_{it}\big(r_{i,t+1} - r_{F,t+1}\big). \tag{17}$$

[12]The fringe index is the value-weighted portfolio of fringe firms. Because the fringe is a unit mass of atomistic issuers whose idiosyncratic shocks wash out, the index is well diversified and serves as the investors outside asset in the corporate bond market.

The first-order conditions and envelope condition imply the familiar Euler equations:

$$1 = \mathbb{E}_t\big[M_{t,t+1}\, r_{i,t+1}\big], \qquad i = 1, \ldots, N, \tag{18}$$

$$1 = \mathbb{E}_t\big[M_{t,t+1}\, r_{F,t+1}\big], \tag{19}$$

where the investors stochastic discount factor (SDF) is

$$M_{t,t+1} = \frac{\beta_L u'(c_{t+1})}{u'(c_t)} = \beta_L \left(\frac{c_{t+1}}{c_t}\right)^{-\gamma}.$$

Equivalently, bond prices satisfy the asset-pricing relations

$$q_{it} = \mathbb{E}_t\big[M_{t,t+1} R_{i,t+1}\big], \qquad q_{Ft} = \mathbb{E}_t\big[M_{t,t+1} R_{F,t+1}\big]. \tag{20}$$

For granular firms–that is, firms large enough that their idiosyncratic shocks affect the investors wealth–the covariance $\mathrm{Cov}_t(M_{t,t+1}, R_{i,t+1})$ is negative and its magnitude grows with issuance. This is the origin of the *issuer-specific risk premium* and motivates the tractable approximation developed below.

Samuelson-Merton Approximation and the Price-Impact Wedge A direct solution to the investors problem is computationally infeasible. The SDF $M_{t,t+1}$ depends on the investors portfolio return $r_{p,t+1}$, which is itself a function of all granular firms bond returns and thus of their issuance and default decisions. Because issuance today affects next periods bond price and default probability, and because long-term debt links these choices recursively into the future, computing $M_{t,t+1}$ requires solving a large fixed-point problem over policies, prices, and continuation values. Solving the full equilibrium would require tracking the entire joint distribution of firms balance-sheet positions and the evolving covariance between each firms repayment $R_{i,t+k}$ and the SDF–an intractable state space even for moderate values of N.

To retain the key economic channel while avoiding this curse of dimensionality, I adopt a standard quadratic (small-risk) approximation à la Campbell and Viceira (2002). This approach linearizes the investors optimality conditions around a constant consumptionwealth ratio and yields a meanvariance portfolio rule. Crucially, it delivers a closed-form expression for bond prices in which the issuer-specific risk premium appears as a multiplicative "price-impact" wedge on top of the risk-neutral price, while preserving the dependence of this wedge on idiosyncratic risk and firm size.

Lemma 1 (Mean-Variance Optimal Portfolio Rule)**.** *Under the quadratic approximation and a constant consumption-wealth ratio, the investors optimal share in bond i satisfies*

$$\alpha_{it}^{*} \approx \frac{1}{\gamma} \frac{\mu_{it}}{\sigma_{it}^{2}}, \qquad i = 1, \ldots, N.$$

Proof. See Appendix B.

To focus on issuer-level price impact, I assume excess-return covariances across granular firms are negligible (investors specialize in a single issuer). Then $\Sigma_t = \operatorname{diag}(\sigma_{1t}^2, \ldots, \sigma_{Nt}^2)$. Market clearing requires

$$\alpha_{it} W_t = q_{it} B_{i,t+1}.$$

Substituting Lemma 1 and using $r_{i,t+1} = R_{i,t+1}/q_{it}$ yields the closed-form pricing equation

$$q_{it} = \underbrace{\frac{\mathbb{E}_t(R_{i,t+1})}{r_{F,t}}}_{\text{Risk-neutral price}} \times \underbrace{\Phi_{it}}_{\text{Price-impact wedge}} \tag{21}$$

$$\Phi_{it} \equiv 1 - \underbrace{\gamma \frac{\mathbb{V}_t(R_{i,t+1})}{\mathbb{E}_t(R_{i,t+1})}}_{\text{idiosyncratic risk}} \times \underbrace{\frac{B_{i,t+1}}{W_t}}_{\text{granularity}} \tag{22}$$

Interpretation The first term $\frac{\mathbb{E}_t(R_{i,t+1})}{r_{F,t}}$ is the risk-neutral price, which would prevail if investors were perfectly diversified. In that benchmark, the bond is valued by discounting its expected repayment at the return on the diversified fringe index, and all issuers—large

or small—would face identical, perfectly elastic demand curves.

By contrast, the second term $\Phi_{it} < 1$ captures the price-impact wedge faced by a granular borrower. Because the firm's idiosyncratic shocks cannot be fully diversified, the investors required yield rises whenever (i) the bonds payoff variance increases, (ii) the firm issues more debt—raising its share of investor wealth $\frac{B_{i,t+1}}{W_t}$, (iii) investor risk aversion γ increases. In each case, the covariance between the bond payoff and the investors SDF becomes more negative, the wedge Φ_{it} shrinks, and the bond trades further below its risk-neutral price. This reflects the additional compensation investors demand to bear the undiversifiable risk created by the interaction of idiosyncratic uncertainty and firm size.

3.2 Recursive Formulation

I focus on a Recursive Markov Equilibrium in detrended variables. Let X_{it} denote the level of any firm variable X, and define its detrended counterpart by

$$\widehat{X}_{it} \equiv \frac{X_{it}}{Y_{it}^P}, \qquad \widehat{X}_{i,t+1} = \frac{X_{i,t+1}}{Y_{i,t+1}^P} = \widehat{X}'_{it}\, e^{-\eta_{t+1}},$$

where Y_{it}^P is the permanent component of income and η_{t+1} is the permanent productivity shock. Because the model is homogeneous in Y_{it}^P, detrending renders the firms problem stationary. The state for a firm is the tuple $s = (b, a, \eta, \varepsilon)$, where b and a are outstanding debt and cash, respectively, and (η, ε) collects permanent and transitory shocks, respectively.

If the firm does not default, its value satisfies

$$v(s) = \max_{\widehat{b}', \widehat{a}', d} \left\{ u(d) + \beta \, \mathbb{E} \left[e^{\eta'} v(s') \mid \text{no default} \right] \right\} \tag{23}$$

$$\begin{aligned} \text{s.t.} \quad d &= (1-\tau)\big[(1+\varepsilon) - \psi\big] + a - q_a^{\tau} \widehat{a}' - \big[(1-\tau)c + \delta\big] b \\ &\quad + q(\widehat{b}', \widehat{a}', s) \left[\widehat{b}' - (1-\delta) b \right] - (\xi_1 |i| + \xi_2 i^2) \end{aligned} \tag{24}$$

$$i = \widehat{b}' - (1-\delta) b, \tag{24a}$$

$$\widehat{b}' = b' e^{-\eta'}, \qquad \widehat{a}' = a' e^{-\eta'}. \tag{24b}$$

The continuation value incorporates future default risk. A firm defaults if its continuation value is non-positive. Define the default threshold $\varepsilon^*(b, a)$ by

$$v(b, a, \varepsilon^*(b, a), \eta) = 0.$$

The firm repays and continues if $\varepsilon > \varepsilon^*(b, a)$; otherwise it exits, bondholders recover $\lambda(\cdot)$, and equity value is zero.

Bond prices incorporate the price-impact wedge derived in Section 3.1.2. Let $\mathcal{R}(\widehat{b}', \widehat{a}', s')$ denote next-period repayment. Under the mean-variance approximation and log utility for the investor ($\gamma = 1$), and normalizing detrended investor wealth $\widehat{W}_t \equiv W_t / Y_t^P$ so that $\gamma / \widehat{W} = 1$[13], the issuer-specific price-impact wedge depends only on repayment risk and debt issuance. The recursive bond price can be written as

$$q(\widehat{b}', \widehat{a}', s) = \frac{\mathbb{E}_{s'|s} \big[\mathcal{R}(\widehat{b}', \widehat{a}', s') \big]}{r_F} \, \Phi(\widehat{b}', s), \tag{25}$$

[13]This normalization exploits the fact that, under the quadratic (small-risk) approximation, the issuer-specific price-impact wedge depends on the product $(\gamma/\widehat{W})\widehat{b}'$, where $\widehat{b}'$ is detrended debt. Because the firm's problem is homogeneous in the permanent component of income and only the ratio $B_{i,t+1}/W_t$ matters for pricing, we can measure debt in units of investor risk-bearing capacity and set $\gamma/\widehat{W} \equiv 1$ without loss of generality. Economically, this corresponds to a partial-equilibrium environment in which the representative investor is large and holds a diversified portfolio of outside assets, so that detrended wealth is approximately constant and unaffected by any single issuers decision.

where r_F^{-1} is the discount factor implied by returns on the diversified fringe index, and

$$\mathcal{R}(\widehat{b}', \widehat{a}', s') = \mathbf{1}\{v(s') > 0\}\big[c + \delta + (1-\delta)q(\widehat{b}'', \widehat{a}'', s')\big] + \mathbf{1}\{v(s') \leq 0\}\,\lambda(\widehat{b}'', \widehat{a}'', s'). \quad (26)$$

The price-impact wedge is given by

$$\Phi(\widehat{b}', s) = 1 - \frac{\mathbb{V}_{s'|s}\big(\mathcal{R}(\widehat{b}', \widehat{a}', s')\big)}{\mathbb{E}_{s'|s}\big(\mathcal{R}(\widehat{b}', \widehat{a}', s')\big)}\,\widehat{b}',$$

so that the wedge shrinks and the bond price moves further below its risk-neutral benchmark as the firms detrended debt $\widehat{b}'$ and repayment risk increase.

3.3 Equilibrium

Definition. A *Recursive Markov Equilibrium* (log-utility baseline) for this economy consists of (i) a set of firms' policy functions for each granular firm $i = 1, \ldots, N$ and representative fringe firms including borrowing, saving, and dividend $\{\widehat{b}', \widehat{a}', d\}$; (ii) value functions $\{v\}$ and implied default thresholds $\{\varepsilon^*\}$; (iii) bond price schedules $\{q(\widehat{b}', \hat{a}', s)\}$ for granular firms and a constant price q_F for the diversified fringe index; such that the following conditions hold:

1. Given the bond price functions $q(\cdot)$, each firm solves the dynamic programming problem defined in Equation (23), subject to constraints (24). The optimal policy functions $(\widehat{b}', \widehat{a}', d)$ and associated value function $v(s)$ satisfy the Bellman equation, and the default rule is defined by:

$$v(s) \leq 0 \quad \Leftrightarrow \quad \varepsilon \leq \varepsilon^*(b, a).$$

2. Taking as given firms' policy and value functions, bond price q satisfies Equation 25.

3.4 Characterization of optimal policy under the price-impact wedge

Granular firms, unlike atomistic firms in the competitive fringe, internalize the effect of their financial decisions on the pricing of their own debt. Issuing more debt raises their exposure to undiversifiable risk, which is priced by the representative investor. This feedback is embedded in the bond price function $q(\cdot)$ via the price-impact wedge $\Phi(\cdot) \in (0, 1)$, a distortion relative to the frictionless, risk-neutral valuation. The wedge increases the marginal cost of external finance with firm size, repayment risk, and leverage.

The price-impact wedge also gives saving a special, strategic role that goes beyond standard precautionary motives. In a frictionless environment, saving in a risk-free asset smooths dividends and reduces default risk; however, for a granular firm, saving has an additional benefit: it directly attenuates the price-impact wedge. Higher cash holdings reduce the variance of next-period repayment and improve the firms repayment profile from the investors perspective. Because the wedge is proportional to the ratio of repayment variance to expected repayment, a reduction in this variance shrinks the wedge and raises the bond price toward its risk-neutral benchmark. Thus, saving does not simply buffer cash-flow risk–it strategically improves credit conditions by weakening the sensitivity of bond prices to undiversifiable issuer-specific risk. This mechanism enhances the marginal value of liquidity and leads granular firms to accumulate more cash than fringe firms even when they face identical fundamentals and default probabilities.

Figure 3: Impact of borrowing $\widehat{b}'$ on bond pricing schedules

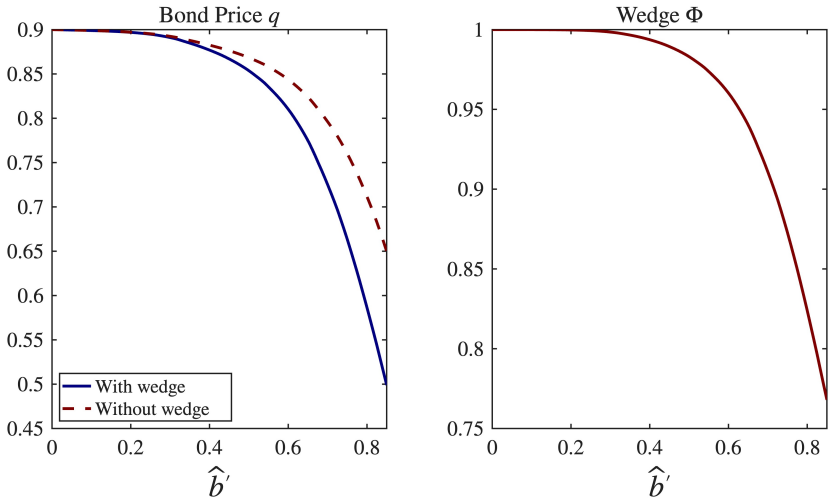
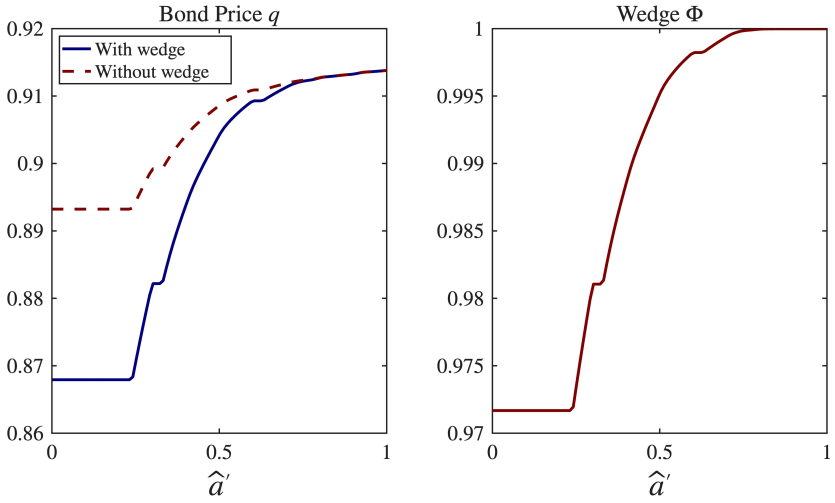


Figure 4: Impact of saving $\widehat{a}'$ on bond pricing schedules



Notes: Figures 3 and 4 plot bond pricing schedules as functions of next-period debt $\widehat{b}'$ and next-period cash $\widehat{a}'$, respectively, holding current balance-sheet variables at their medians and ε at its unconditional mean. In each figure, the solid line shows the baseline model with the price-impact wedge; the dashed line shows a counterfactual in which the firm is treated as a competitive-fringe issuer. Borrowing steepens the decline in bond prices when the wedge is active, while saving attenuates the wedge and raises bond prices more strongly than in the fringe benchmark.

In equilibrium, these pricing schedules lead granular firms to adopt more conservative financial policies: they borrow less and save more than otherwise identical fringe firms. The next figures show optimal policies as functions of current debt and cash, conditional on the

median realization of idiosyncratic shocks.

Figure 5: Optimal borrowing policy w.r.t. outstanding debt b

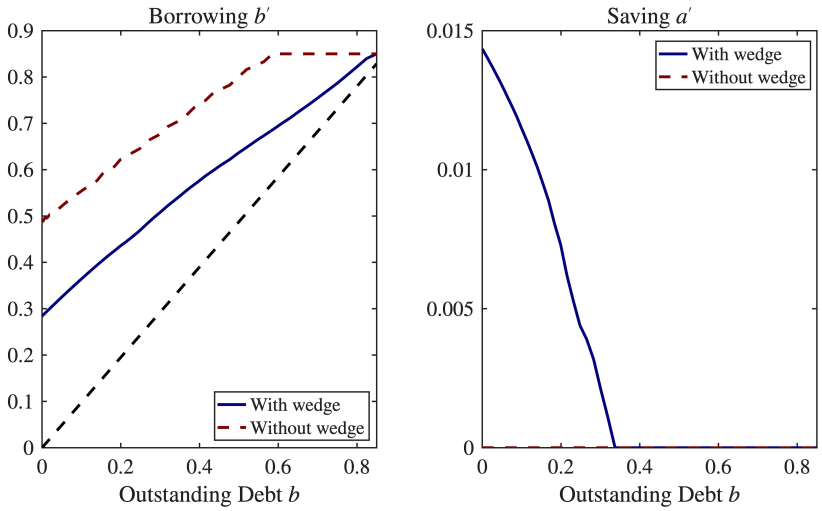
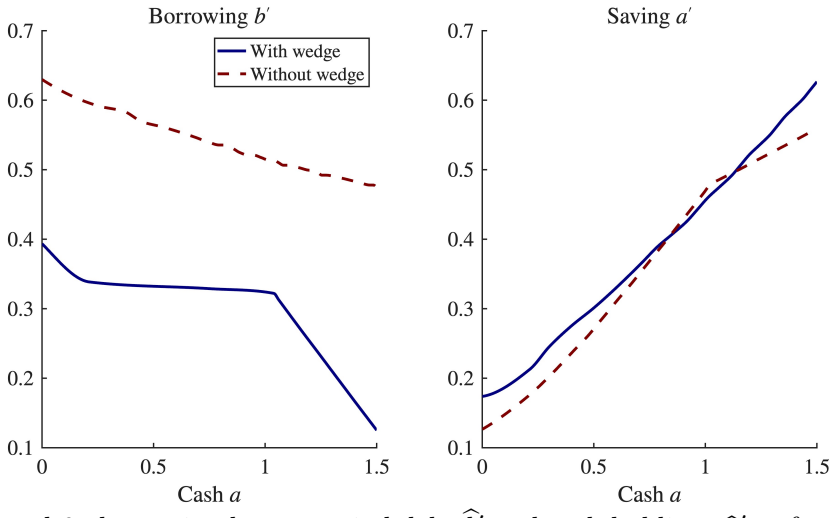


Figure 6: Optimal saving policy w.r.t. cash balances a



Notes: Figures 5 and 6 plot optimal next-period debt $\widehat{b}'$ and cash holdings $\widehat{a}'$ as functions of current debt b and current cash a, respectively, holding the other balance-sheet variable at its median and ε at its unconditional mean. The solid line corresponds to the baseline model with the price-impact wedge; the dashed line shows the counterfactual without the wedge. Granular firms choose lower $\widehat{b}'$ for any given level of debt and higher $\widehat{a}'$ for any given level of cash, reflecting the steeper marginal cost of borrowing and higher marginal value of liquidity created by the price-impact wedge.

These comparative statics highlight the core mechanism: the price-impact wedge increases the sensitivity of borrowing costs to debt issuance and amplifies the benefits of

liquidity. Granular firms therefore optimally underborrow and oversavenot due to tighter constraints or weaker fundamentals, but because their own financing decisions affect the pricing kernel and thus their cost of capital. The next proposition formalizes these implications, comparing the optimal choices of a granular issuer to those of an otherwise identical fringe firm that takes prices as given.

These comparative statics highlight the core mechanism: the price-impact wedge increases the sensitivity of borrowing costs to debt issuance and amplifies the marginal value of liquidity. Granular firms therefore optimally underborrow and oversavenot because they face tighter constraints or weaker fundamentals, but because their own financing choices affect the pricing kernel and thus their cost of capital. The next proposition formalizes these implications by comparing the optimal policies of a granular issuer to those of an otherwise identical fringe firm that takes prices as given.

Proposition 1 (**Oversaving and Underborrowing due to the Price-Impact Wedge**)**.** *Consider two firms with identical fundamentals and default probabilities. If one firm is granular and internalizes its effect on the investors SDF (i.e., faces the price-impact wedge* $\Phi(\cdot) < 1$*), while the other is a competitive-fringe firm that takes prices as given (i.e.,* $\Phi(\cdot) = 1$*), then in equilibrium:*

1. The granular firm chooses strictly lower next-period debt b' than the fringe firm, all else equal;

2. The granular firm chooses strictly higher next-period cash a' than the fringe firm.

Proof. See Appendix B.

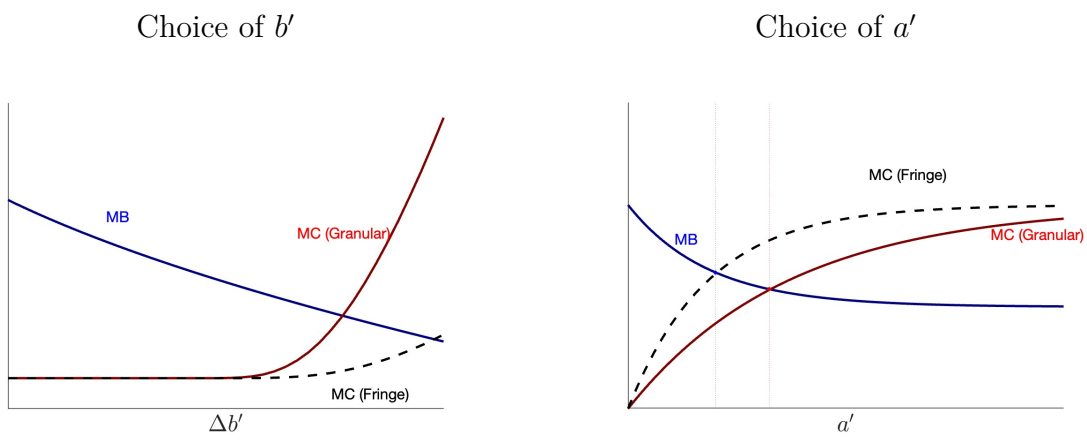
Granular firms face endogenously higher borrowing costs because the bond price they receive declines with issuance both through standard default-risk effects and through the additional price-impact wedge arising from undiversifiable issuer-specific risk. Conversely, greater cash holdings reduce repayment variance and expand the wedge, raising the bond price. Competitive-fringe firms also face a downward-sloping pricing schedule due to default

risk, but their prices do not respond to their own size in investors portfolios and therefore lack the additional wedge component. Accordingly, the granular firms bond pricing function $q(\widehat{b}', \widehat{a}', s)$ is more sensitive to $\widehat{b}'$ and $\widehat{a}'$ than that of an otherwise identical fringe firm, reflecting the extra price impact created by granularity.

This asymmetry alters perceived marginal trade-offs. For granular firms, the marginal cost of borrowing rises quickly with issuance, and the marginal benefit of liquidity is high because saving attenuates the price-impact wedge. As a result, granular firms issue less debt and accumulate more cash than an otherwise identical fringe firm. This behavior constitutes a form of strategic self-insurance against steep endogenous borrowing costs, arising purely from firms' influence on investor risk premia.

Holding fixed the default threshold $\varepsilon^*(b, a)$, a granular firm therefore selects a lower $\widehat{b}'$ and a higher $\widehat{a}'$ relative to a fringe firm. Figure 7 illustrates these comparative statics: for the same marginal benefit of external funds, the granular firm encounters a steeper marginal borrowing cost schedule and a flatter marginal saving cost schedule. Differences in financial policy thus arise not from differences in fundamentals, but from the endogenous pricing feedback associated with granularity.

Figure 7: Underborrowing and Oversaving



Note: The left panel illustrates that the marginal cost of issuing $\widehat{b}'$ is higher for granular firms because their bond price schedule is steeper. The right panel shows that the marginal cost of saving $\widehat{a}'$ is lower, since additional liquidity reduces repayment variance and attenuates the price-impact wedge.

4 Quantitative Analysis

This section evaluates the models ability to account for the financial behavior of large, market-exposed U.S. public firms and the cross-sectional patterns in their credit spreads. The quantitative analysis proceeds in three steps.

(i) *Baseline fit.* I first assess how much of the observed leverage and cash-to-asset ratios of granular firms can be rationalized by the price-impact wedge derived in the model. The wedge is calibrated to be consistent with empirically observed credit spreads and generates heterogeneous borrowing costs across firms that differ in size and idiosyncratic risk exposure.

(ii) *Cyclical response to credit tightening.* I then study the firm-level effects of a credit-tightening shock and compare the model-implied adjustments in saving and leverage to the empirical responses documented in Section 2. This exercise serves as a validation of the core mechanism: the model should reproduce the sharper deleveraging and greater liquidity accumulation exhibited by granular firms when credit conditions tighten.

(iii) *Contribution to the secular rise in corporate cash.* Finally, I quantify the extent to which the granular risk-premium channel explains the secular increase in cash holdings among the largest U.S. firms since the early 2000s. As documented in Section 4.3, this period features a persistent rise in undiversifiable idiosyncratic volatility. Through the endogenous credit-spread mechanism, higher volatility strengthens the strategic saving motive for granular firms. In a counterfactual economy where this mechanism is shut down (i.e., granular firms are treated as price-taking borrowers), model-implied cash holdings fall by roughly one-quarter to one-third. This decline implies that the interaction between rising idiosyncratic volatility and the granular risk-premium channel can account for a substantial fraction of the observed increase in cash holdings between 2000 and 2020.

4.1 Calibration and Estimation

The model is calibrated at a quarterly frequency using firm-level data from Compustat, CRSP, and TRACE. The calibration strategy targets moments that capture key differences in financial policies and credit pricing between granular firms and the competitive fringe. Details of the numerical solution algorithm are provided in Appendix C.

Parameter values are chosen using a combination of external calibrationdrawing on standard values in the literatureand internal calibration to match a set of empirical moments for each firm type. Table 9 summarizes the parameter choices, and Table 10 compares the corresponding long-run moments in the data and in model simulations.

Table 9: Parameter choices

Parameter	**Value**	**Description**	**Target / Data source**
(a) Externally calibrated parameters			
β	0.9902	Discount factor of equity holders	Annual risk-free rate (4%)
δ	1/40	Amortization rate on long-term debt	TRACE 10-year bond maturity
τ	0.3	Corporate income tax rate	Graham (2002)
χ	0.4	Recovery rate	Gomes and Schmid (2020)
(b) Internally calibrated parameters			
$(\sigma_\eta^G, \sigma_\eta^F)$	(0.15, 0.47)	Std. dev. of permanent income shock	Estimated from earnings process
$(\sigma_\varepsilon^G, \sigma_\varepsilon^F)$	(0.15, 0.38)	Std. dev. of transitory income shock	Estimated from earnings process
ξ_1	0.05	Linear debt-issuance cost	Quarterly issuance frequency
ξ_2	0.3	Quadratic debt-issuance cost	Conditional issuance size
(ψ^G, ψ^F)	(0.0, 0.2)	Fixed production cost	Annual default rate

Note: This table reports parameter values used to calibrate the model. $(\sigma_\eta^G, \sigma_\eta^F)$ denotes the standard deviation of permanent cash-flow shocks for granular and fringe firms, respectively; $(\sigma_\varepsilon^G, \sigma_\varepsilon^F)$ denotes the corresponding transitory components. Granular firms are defined as the top 1% of firms by market capitalization, with the remaining firms classified as fringe, over 1980Q1-2019Q4. The earnings process is estimated separately for each group using EBITDA-to-assets ratios, as described in Appendix D.1. Debt adjustment costs follow $\Phi(i) = \xi_1|i| + \xi_2 i^2$ for $i \neq 0$. (ψ^G, ψ^F) represents fixed operating costs for granular and fringe firms.

Table 10: Statistics - Data vs. Model

Description	Type	Data	Model
Leverage (%)	Granular	20.54	27.44
	Fringe	23.09	39.52
Cash-to-Assets (%)	Granular	13.40	14.80
	Fringe	16.86	26.06
Cash Flow-to-Assets (%)	Granular	2.95	2.56
	Fringe	2.67	2.22
Debt issuance, frequency	Granular	33.50	27.70
	Fringe	29.28	59.79
Debt issuance, size	Granular	0.63	0.855
	Fringe	1.46	3.42
Annual default rate (%)	Granular	0.07	3.77
	Fringe	7.43	9.26
Credit spreads (bp)	Granular	130	84
	Fringe	284	161

Note: This table compares data moments for top 1% largest firms (granular) and remaining firms (fringe) over the sample period to model moments from a simulation with (300,4000) respective firms over 2,000 quarters; the first 1,000 periods are discarded. For calculating the annual default rate for data moments, due to data limitation, I instead use distance to default (dd) measure estimated from Merton model, and compute implied default probability $N(-dd)$ where $N(\cdot)$ is standard normal distribution.

While the model captures several important cross-sectional patterns, a few limitations in the moment fit are worth noting.

Credit spreads. Model-implied credit spreads are lower than in the data for both granular and fringe firms. Because the model abstracts from aggregate uncertainty and from

intermediary-based or liquidity-driven risk premia, investors price credit risk purely through undiversifiable idiosyncratic shocks. This creates the familiar tension in structural credit models: increasing spreads requires raising default risk, but doing so pushes default rates far above empirically plausible levels. As a result, the model can match the dispersion in spreads across firm types more easily than their levels.

Issuance dynamics. The linear-quadratic debt-adjustment cost is assumed common across granular and fringe firms. This parsimonious specification helps discipline leverage dynamics, but limits the models flexibility in jointly matching issuance frequency and issuance size. In particular, the model overpredicts issuance frequency for fringe firms while matching granular firms more closely. Allowing for heterogeneity in adjustment costs, or incorporating additional financing frictions such as credit lines or short-term rollover risk, could help bridge this gap.

Overall, despite these limitations, the calibrated model reproduces several salient differences between granular and fringe firms in leverage, cash policies, issuance behavior, and the relative pricing of credit risk. These features are essential for the quantitative exercises that follow.

Before turning to the counterfactual exercises, it is useful to examine the sources of variation in credit spreads faced by granular firms in the model. In particular, the model implies that most of the dispersion in credit pricing arises not from the risk-neutral component associated with expected default, but from the endogenous risk-premium channel generated by the price-impact wedge. To see this more clearly, consider the decomposition

$$\log q = \underbrace{\log q_{RF}}_{\text{Risk-neutral}} + \underbrace{\log \Phi}_{\text{Price-impact wedge}}, \tag{27}$$

$$Var(\log q) = Var(\log q_{RF}) + Var(\log \Phi) + 2\, Cov(\log q_{RF}, \log \Phi). \tag{28}$$

Because expected repayment and $\Phi(\cdot)$ are jointly determined, the covariance term is generally nonzero. Nonetheless, this variance decomposition is informative: it allows us to quantify

how much of the observed variation in credit spreads is attributable to default risk alone, to the price-impact wedge, and to the interaction of the two.

Table 11: Variance Decomposition and counterfactuals

	$Var(\log q_{RF})$	$Var(\log \Phi)$	Covariance	Cash ratio	Leverage
Baseline	0.25	0.43	0.32	13.27	26.36
Counterfactuals	1	-	-	9.94	26.50

Notes: This table decomposes granular firms' bond prices into a risk-neutral component q_{RF} and an endogenous price-impact wedge $\Phi(\cdot)$, satisfying $\log q = \log q_{RF} + \log \Phi$. Because expected default and the wedge are jointly determined, the variance decomposition includes the covariance term. The baseline row reports the variance shares from the full model; the counterfactual sets $\Phi \equiv 1$, treating granular firms as price-takers. The last two columns show the implied cash and debt ratios. Shutting down the wedge reduces the cash ratio from 13.27% to 9.94%, indicating that the price-impact channel induces substantial precautionary liquidity, while leverage remains nearly unchanged.

Table 11 shows that, in the baseline economy, only about 25% of the variation in granular firms' bond prices is explained by the risk-neutral component q_{RF}. By contrast, the price-impact wedge alone accounts for roughly 44% of the variance, with the remainder captured by the covariance between the two terms. This pattern highlights a central implication of the model: even for firms that are fundamentally safeprecisely those that resemble investment-grade issuers in the data a large portion of the observed credit-spread variation reflects an *endogenous risk premium*, not expected default losses.

This provides a natural explanation for the well-documented "credit spread puzzle" (e.g., Huang and Huang, 2012; Gilchrist and Zakrajšek, 2012), in which investment-grade firms exhibit credit spreads that are far too wide to be justified by default probabilities alone. In the present framework, these excess spreads arise because granular firms' own financing decisions shift investor risk exposure and therefore move market prices. Even holding expected default fixed, granular firms face systematically tighter credit terms than counterfactual fringe firms facing the same fundamentals. In this sense, granularity endogenously

generates a non-default premium that resembles the "excess bond premium" in the data.

The final two columns of Table 11 show the implications for firms' balance-sheet choices. In a counterfactual economy where the endogenous pricing wedge is shut down (i.e., $\Phi \equiv 1$ so granular firms behave as price-takers), the cash ratio falls from 13.27% to 9.94%. This sizable decline demonstrates that the price-impact channel induces granular firms to engage in additional precautionary saving: liquidity serves to shrink repayment variance and thereby relax the wedge. By contrast, leverage remains nearly unchanged across the two environments, consistent with the models mechanism operating primarily through liquidity buffers rather than through the scale of debt issuance.

Taken together, the variance decomposition and the counterfactual illustrate a unified mechanism: large, investment-grade firms face higher-than-expected credit spreads because their own idiosyncratic shocks carry undiversifiable consequences for the investor, and this endogenous pricing feedback induces them to operate with elevated liquidity buffers. Granular firms thus appear "safe" in fundamentals yet face elevated credit spreads–a pattern that the model generates without appealing to exogenous intermediary frictions or latent risk factors.

4.2 Counterfactual exercise: Impulse response to VIX-type shock [In Progress]

As a test of model fit, I examine whether it can replicate the dynamic response of firms leverage and cash holdings to a tightening of credit conditions, in the spirit of the panel regressions documented in Section 2. To parallel the empirical specification in which the VIX enters as an aggregate uncertainty measure, I extend the volatility process so that the transitory cash-flow shock for each firm contains both an idiosyncratic and an aggregate component. Formally, let

$$\sigma_{\varepsilon,t} = \sigma_{\varepsilon,t}^{\text{agg}} + \sigma_{\varepsilon}^{\text{idiosyn}},$$

where $\sigma_{\varepsilon}^{\text{idiosyn}}$ is constant and firm-specific, while $\sigma_{\varepsilon,t}^{\text{agg}}$ is an aggregate state variable common to all firms. This extension preserves tractability while allowing aggregate uncertainty to shift firms perceived risk in a way comparable to movements in the VIX.

The aggregate volatility state evolves according to

$$\sigma_{\varepsilon,t+1}^{\text{agg}} = \sigma_{\varepsilon,0}^{\text{agg}} + \rho\big(\sigma_{\varepsilon,t}^{\text{agg}} - \sigma_{\varepsilon,0}^{\text{agg}}\big) + \varepsilon_{t+1}^{\sigma}, \tag{29}$$

where $\sigma_{\varepsilon,0}^{\text{agg}}$ is the steady-state volatility reported in Table 9, the persistence parameter $\rho = 0.8$ matches the empirical persistence of the VIX, and $\varepsilon_{t+1}^{\sigma}$ is an unexpected aggregate volatility innovation. A VIX-type shock is implemented as a one-time surprise jump $\varepsilon_{120}^{\sigma} > 0$ at period $t = 120$, after which aggregate volatility gradually reverts according to (29). This specification captures both the sudden onset and slow unwinding of a high-uncertainty regime.

Impulse response functions (IRFs) are computed as differences in means between two sets of simulated economies: one subjected to the aggregate volatility shock and one under the baseline model without shocks. Each set consists of 100 independently simulated economies, each run for 200 periods with the first 100 periods discarded as burn-in. In the treated economies, the shock occurs at $t = 120$, which is normalized to period 0 in Figure 8.

To isolate the role of the granular risk-premium channel, I also simulate a counterfactual economy in which the price-impact wedge is shut downso granular firms behave as competitive price-takers in credit markets. Comparing IRFs across the two model variants reveals how endogenous credit pricing amplifies the precautionary saving and deleveraging responses of granular firms following a VIX-type shock. In particular, the wedge-based model is evaluated on its ability to reproduce the stronger beta- and size-dependent adjustments in cash and leverage documented in the empirical panel regressions.

Figure 8: Impulse responses to VIX-type shock: data vs. model

[Insert here]

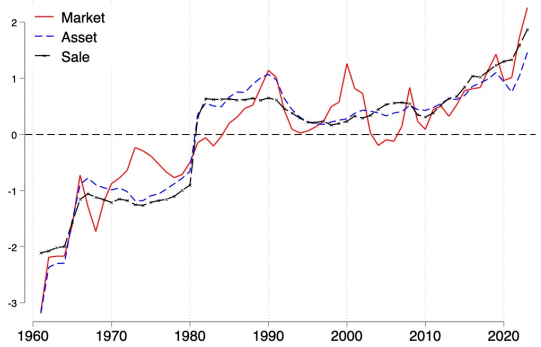
4.3 Explaining the Secular Rise in Corporate Saving Since the 2000s

The final objective of the quantitative analysis is to assess whether the granular risk-premium channel can account for the secular increase in corporate cash holdings observed since the early 2000s. As documented in the empirical section, cash balances have risen sharply across the economy, with the increase particularly pronounced among the largest, most market-exposed firms. While standard precautionary motives predict greater saving when uncertainty rises, the key question here is whether the endogenous tightening of credit conditions associated with granularity amplifies this trend beyond what a frictionless model would imply.

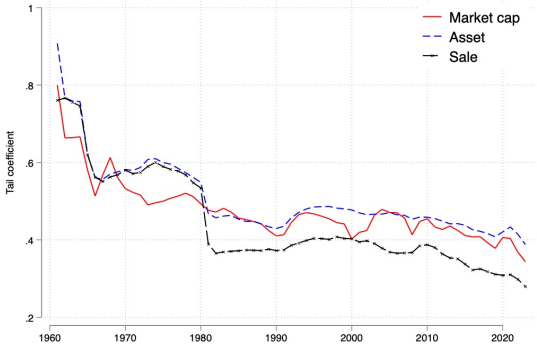
Rising granularity and undiversifiable volatility. A natural starting point is the evolving distribution of firm size. Figure 9 shows that firm size dispersion has increased markedly over the past two decades, with the upper tail of the distribution thickening as measured by declining Pareto exponents. The rise in concentration implies that shocks to dominant firms now account for a larger share of aggregate activity and are therefore less diversifiable for investors.

Figure 9: Increasing size dispersion

(a) Size dispersion



(b) Pareto tail coefficient

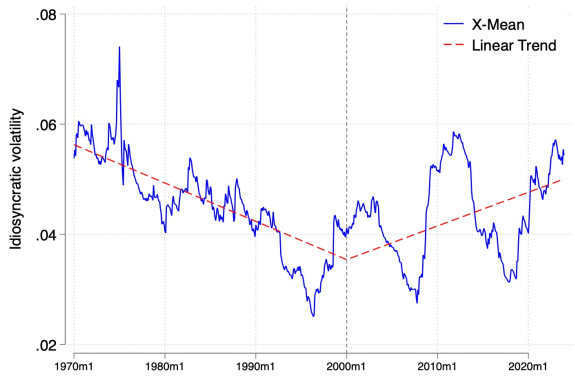


Notes: The left panel plots the cross-sectional dispersion of firm size in each year using Compustat data. Dispersion is measured as the standard deviation of log(Sales), log(Total Assets), and log(Market Equity), computed at the end of each year. Each series is standardized to have zero mean and unit standard deviation over the sample period to facilitate comparison. The right panel shows the corresponding Pareto tail exponents for each size measure, estimated annually via the rank-size regression $\log(\text{Rank} - 1/2) = a + b \log(\text{Size})$, following Gabaix and Ibragimov (2011). A lower exponent indicates a thicker upper tail, reflecting greater concentration among the largest firms.

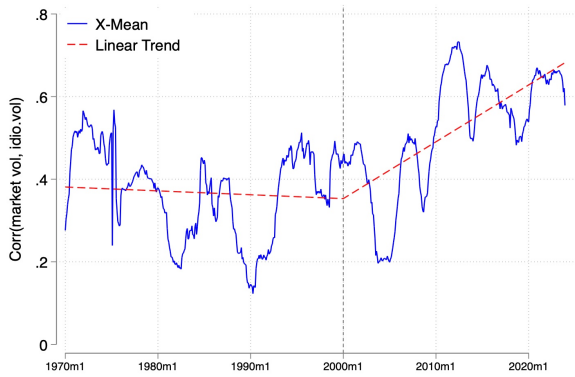
At the same time, firm-level idiosyncratic volatility has also increased. Figure 10 documents two key trends in firm-level risk. The left panel shows a rising trend in the cross-sectional mean of idiosyncratic volatility, particularly after 2000. The right panel highlights a growing correlation between idiosyncratic volatility and aggregate market volatility, suggesting increased integration and reduced scope for diversification. These patterns are consistent with heightened firm-level uncertainty and a stronger link between firm-specific and systematic risk in recent decades.

Figure 10: Idiosyncratic volatility

(a) Cross-sectional mean



(b) Correlation with market volatility



Notes: Idiosyncratic volatility is estimated from firm-level factor regressions: monthly stock returns are regressed over a 60-month rolling window using models such as CAPM, Fama-French 3- and 5-factor, and q-factor models: $r_{it} = \beta_{it}' F_t + \varepsilon_{it}$. The standard deviation of the residual term $\widehat{\varepsilon}_{it}$ serves as the idiosyncratic volatility for firm i in month t. The left panel plots the cross-sectional mean of this measure over time. Linear trends are estimated separately for the pre-2000 and post-2000 periods. The right panel shows the rolling 60-month correlation between the cross-sectional average of idiosyncratic volatility and market volatility, defined as the standard deviation of the monthly value-weighted return.

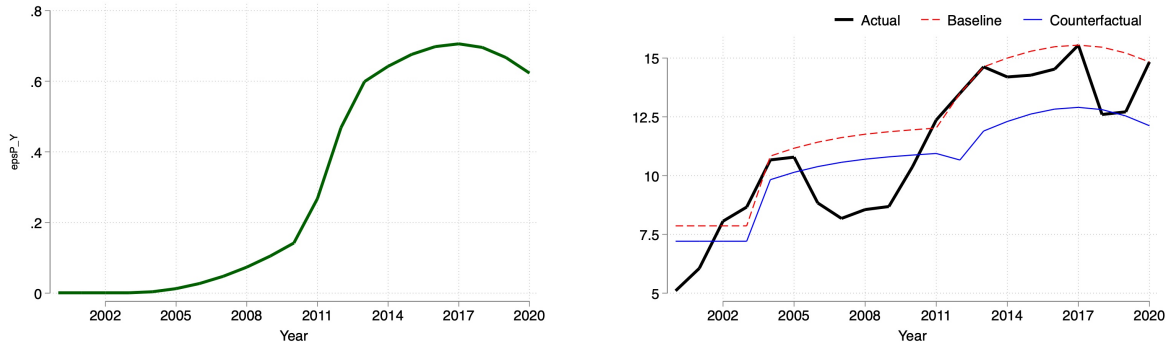
To account for the upward trend in corporate saving and to quantify the contribution of the granular risk-premium channel to precautionary cash accumulation, this final exercise estimates a time path of idiosyncratic volatility that allows the model to match the observed increase in the aggregate cash-to-asset ratio between 2000 and 2020. Holding all other parameters fixed, the permanent volatility parameter is calibrated to match the cash ratio in 2000, 2012, 2017, and 2020, and interpolated between these years using cubic splines. Using this constructed sequence, the model is then simulated under two environments: the baseline case with endogenous credit pricing and the price-impact wedge $\Phi(\cdot)$, and a counterfactual economy in which granular firms are price-takers and face no wedge.

In the counterfactual scenario where the granular risk-premium channel is shut down, the resulting cash ratios are substantially lower. The difference between the baseline and counterfactual paths–especially in the terminal year 2020–provides a direct quantitative measure of how much of the secular rise in corporate cash holdings is attributable to the interac-

tion between rising undiversifiable idiosyncratic volatility and the endogenous tightening of credit spreads faced by granular firms. This comparison isolates the share of the upward trend driven specifically by the endogenous credit-pricing mechanism, as opposed to purely precautionary considerations.

Figure 11: Model-Implied Time Series of Cash Ratios

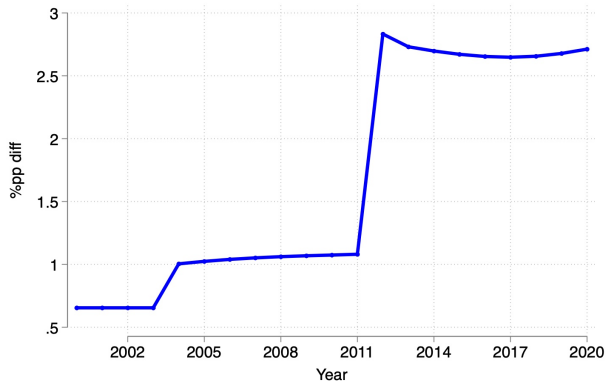
(a) Estimated σ_P^G **(b)** Actual vs. Model-Implied Cash Ratio



Notes: The left panel plots the estimated time series of permanent idiosyncratic cash-flow volatility for granular firms. The parameter σ_P^G is calibrated to match the observed cash-to-asset ratio among the top 1% of firms in 2000, 2012, 2017, and 2020, and interpolated via cubic splines in the intervening years. The right panel compares the actual cash ratio of large firms with the model-implied series under the estimated volatility sequence (red dashed). The blue line shows the corresponding counterfactual in which granular firms face no price-impact wedge $\Phi(\cdot)$, using the same volatility sequence.

The widening gap between the baseline and counterfactual series is plotted in Figure 12. The difference between the two paths grows steadily as idiosyncratic volatility rises, reflecting the strengthening of the endogenous credit channel. When volatility is high, the price-impact wedge significantly erodes credit prices, magnifying the benefit of additional liquidity accumulation.

Figure 12: Gaps between Baseline vs. Counterfactual cash-to-asset ratio



In years where the model closely matches observed data, shutting down the granular risk-premium channel reduces model-implied cash ratios by 2.112.83 percentage points, with an average decline of 2.56 p.p. To interpret this magnitude, note that the cash ratio of top-1% firms rises from roughly 5% in 2000 to about 15% in 2020a 10 percentage point increase. Thus, a reduction of 2.12.8 p.p. implies that the granular risk-premium mechanism accounts for approximately 2028% of the secular rise in corporate cash holdings among dominant firms.

In other words, about one-fifth to one-quarter of the post-2000 increase in liquidity buffers can be attributed to the interaction between (i) rising undiversifiable idiosyncratic volatility and (ii) tighter credit pricing induced by the price-impact wedge. The remainder reflects standard precautionary motives and broader structural forces affecting cash demand.

Taken together, these findings highlight that the post-2000 rise in cash holdings among top-1% firms cannot be explained solely by higher volatility or technological change. Instead, the endogenous tightening of credit spreads for granular firms plays an important role in amplifying the corporate saving response, deepening the wedge between granular and fringe firms, and producing the secular patterns observed in the data.

5 Conclusion

This paper provides a new explanation for why the largest and most creditworthy firms in the economy choose to underborrow and oversave. Using detailed bond and balance sheet data, I show that dominant firms internalize their impact on equilibrium asset prices: their debt issuance raises undiversifiable risk, which investors price into credit spreads through a firm-specific pricing-kernel wedge. This endogenous risk premium induces these firms to self-insure by holding more cash and issuing less debt, particularly in periods of heightened uncertainty.

Embedding this mechanism into a structural model of corporate financing with incomplete risk-sharing, I demonstrate that the precautionary behavior of granular firms is not a sign of binding constraints or agency frictions, but a rational response to the equilibrium pricing environment they face. The model provides a micro-founded link between firm-level risk, credit-market pricing, and balance sheet choices, thereby bridging corporate finance and asset pricing perspectives.

Quantitatively, the model reproduces key cross-sectional patterns in leverage, liquidity, and credit spreads across large and smaller firms. The endogenous pricing wedge explains a substantial portion of the excess credit spreads observed among investment-grade issuers and amplifies their precautionary saving response to volatility shocks. Moreover, the interaction between rising firm-level idiosyncratic volatility since the early 2000s and the granular risk-premium channel accounts for roughly 20–28% of the secular increase in cash holdings among the largest U.S. firms.

More broadly, the analysis highlights the macro-financial consequences of granularity. When firm-specific shocks generate undiversifiable risk for investors, dominant firms become both sources and absorbers of systemic risk. Their balance sheets respond not only to internal fundamentals, but also to their endogenous role in shaping risk premia in credit markets. These findings underscore the importance of integrating corporate-finance models with asset-pricing considerations in environments with large, market-exposed firms.

From a policy perspective, the results suggest that observed credit spreads on large, highly rated firms embed equilibrium forces beyond expected default risk. If spreads partly reflect undiversifiable granular risk, then market-based measures of firm risk may not fully capture fundamentals, and macro-financial stability may be influenced by the behavior of a small set of dominant firms. Understanding the dual role of these firmsas both contributors to and mitigators of systemic riskis an important direction for future research.

References

Acharya, Viral, Heitor Almeida, and Murillo Campello. 2013. "Aggregate risk and the choice between cash and lines of credit." *The Journal of Finance*, 68(5): 2059–2116.

Almeida, Heitor, Murillo Campello, and Michael S Weisbach. 2004. "The cash flow sensitivity of cash." *The journal of finance*, 59(4): 1777–1804.

Amel-Zadeh, Amir, Fiona Kasperk, and Martin C Schmalz. 2022. "Mavericks, universal, and common owners-The largest shareholders of US public firms." *European Corporate Governance Institute–Finance Working Paper*, , (838).

Azar, José A, Jean-Francois Kagy, and Martin C Schmalz. 2016. "Can changes in the cost of carry explain the dynamics of corporate cash holdings?" *The Review of Financial Studies*, 29(8): 2194–2240.

Bates, Thomas W, Kathleen M Kahle, and René M Stulz. 2009. "Why do US firms hold so much more cash than they used to?" *The journal of finance*, 64(5): 1985–2021.

Begenau, Juliane, and Berardino Palazzo. 2021. "Firm selection and corporate cash holdings." *Journal of Financial Economics*, 139(3): 697–718.

Bianchi, Javier, Juan Carlos Hatchondo, and Leonardo Martinez. 2018. "International reserves and rollover risk." *American Economic Review*, 108(9): 2629–2670.

Campbell, John Y, and Luis M Viceira. 2002. *Strategic asset allocation: portfolio choice for long-term investors.* Clarendon Lectures in Economic.

Carvalho, Vasco M, and Basile Grassi. 2019. "Large firm dynamics and the business cycle." *American Economic Review*, 109(4): 1375–1425.

Dickerson, Alexander, Philippe Mueller, and Cesare Robotti. 2023. "Priced Risk in Corporate Bonds." *Journal of Financial Economics*, 150: article 103707.

Falato, Antonio, Dalida Kadyrzhanova, Jae Sim, and Roberto Steri. 2022. "Rising intangible capital, shrinking debt capacity, and the US corporate savings glut." *The Journal of Finance*, 77(5): 2799–2852.

Faulkender, Michael W, Kristine W Hankins, and Mitchell A Petersen. 2019. "Understanding the rise in corporate cash: Precautionary savings or foreign taxes." *The Review of Financial Studies*, 32(9): 3299–3334.

Foley, C Fritz, Jay C Hartzell, Sheridan Titman, and Garry Twite. 2007. "Why do firms hold so much cash? A tax-based explanation." *Journal of financial economics*, 86(3): 579–607.

Friewald, Nils, Florian Nagler, and Christian Wagner. 2022. "Debt refinancing and equity returns." *The Journal of Finance*, 77(4): 2287–2329.

Gabaix, Xavier. 2011. "The granular origins of aggregate fluctuations." *Econometrica*, 79(3): 733–772.

Gabaix, Xavier, and Rustam Ibragimov. 2011. "Rank- 1/2: a simple way to improve the OLS estimation of tail exponents." *Journal of Business & Economic Statistics*, 29(1): 24–39.

Gao, Junxiong. 2022. "Granular Asset Pricing."

Gilchrist, Simon, and Egon Zakrajšek. 2012. "Credit spreads and business cycle fluctuations." *American economic review*, 102(4): 1692–1720.

Gomes, Joao, and Lukas Schmid. 2021. "Equilibrium asset pricing with leverage and default." *The Journal of Finance*, 76(2): 977–1018.

Graham, John R, and Mark T Leary. 2018. "The evolution of corporate cash." *The Review of Financial Studies*, 31(11): 4288–4344.

Greenwald, Bruce C, and Joseph E Stiglitz. 1990. "Asymmetric information and the new theory of the firm: Financial constraints and risk behavior."

Gürkaynak, Refet S, Brian P Sack, and Eric T Swanson. 2007. "Market-based measures of monetary policy expectations." *Journal of Business & Economic Statistics*, 25(2): 201–212.

Hatchondo, Juan Carlos, and Leonardo Martinez. 2009. "Long-duration bonds and sovereign defaults." *Journal of international Economics*, 79(1): 117–125.

Hatchondo, Juan Carlos, Leonardo Martinez, and Cesar Sosa-Padilla. 2016. "Debt dilution and sovereign default risk." *Journal of Political Economy*, 124(5): 1383–1422.

He, Zhiguo, and Wei Xiong. 2012. "Rollover risk and credit risk." *The Journal of Finance*, 67(2): 391–430.

Holmström, Bengt, and Jean Tirole. 1998. "Private and public supply of liquidity." *Journal of political Economy*, 106(1): 1–40.

Huang, Jing-Zhi, and Ming Huang. 2012. "How much of the corporate-treasury yield spread is due to credit risk?" *The Review of Asset Pricing Studies*, 2(2): 153–202.

Jungherr, Joachim, and Immo Schott. 2021. "Optimal debt maturity and firm investment." *Review of Economic Dynamics*, 42: 110–132.

Koijen, Ralph SJ, and Motohiro Yogo. 2019. "A demand system approach to asset pricing." *Journal of Political Economy*, 127(4): 1475–1515.

Kuehn, Lars-Alexander, and Lukas Schmid. 2014. "Investment-based corporate bond pricing." *The Journal of Finance*, 69(6): 2741–2776.

Lucas, Robert E. 1978. "Asset prices in an exchange economy." *Econometrica: journal of the Econometric Society*, 1429–1445.

Lucas, Robert E. 1995. "Understanding business cycles." In *Essential readings in economics*. 306–327. Springer.

Malevergne, Yannick, Pedro Santa-Clara, and Didier Sornette. 2009. "Professor zipf goes to wall street." National Bureau of Economic Research.

Merton, Robert C. 1974. "On the pricing of corporate debt: The risk structure of interest rates." *The Journal of finance*, 29(2): 449–470.

Opler, Tim, Lee Pinkowitz, René Stulz, and Rohan Williamson. 1999. "The determinants and implications of corporate cash holdings." *Journal of financial economics*, 52(1): 3–46.

Ottonello, Pablo, and Thomas Winberry. 2020. "Financial heterogeneity and the investment channel of monetary policy." *Econometrica*, 88(6): 2473–2502.

Panousi, Vasia, and Dimitris Papanikolaou. 2012. "Investment, idiosyncratic risk, and ownership." *The Journal of finance*, 67(3): 1113–1148.

Poeschl, Johannes. 2023. "Corporate debt maturity and investment over the business cycle." *European Economic Review*, 152: 104348.

Riddick, Leigh A, and Toni M Whited. 2009. "The corporate propensity to save." *The Journal of Finance*, 64(4): 1729–1766.

Xiao, Jasmine. 2022. "Borrowing to save and investment dynamics." *Available at SSRN 3478294*.

Appendix A Data

A.1 Firm Balance Sheet Data

- Cash-to-Assets Ratio: Cash/Short-term investment / Total Asset ($CHEQ/ATQ$)
- Market beta for each stock is obtained using monthly data over five-year horizons as in Koijen and Yogo (2019) and run regression of monthly excess returns (over the 1-month Treasury bill rate) onto excess market returns using a 60-month moving window (with at least 24 months of non-missing returns)
- Size: log of Total Assets (ATQ)
- Leverage: Total Debt / Total Asset; Total Debt = sum of Debt in Current Liabilities and Total Long-Term Debt ($DLCQ + DLTTQ$)
- Return on assets (ROA): operating income before depreciation ($OIBDP$) scaled by beginning-of-year total assets ATQ

A.2 Sample Selection

I follow Ottonello and Winberry (2020) for constructing our sample: exclude all firm-quarters for which

- Firm not incorporated in the US
- Firms in finance, insurance, and real estate sectors (SIC codes 6000-6799), utilities (SIC codes 4900-4999), and nonoperating establishments (SIC = 9995), Industrial Conglomerates (SIC= 9997)
- Negative capital or assets
- Investment rate is in the top or bottom 0.5% of the cross-sectional distribution
- Investment spell is shorter than 40 quarters
- Net current assets as a share of total assets higher than 10 or below -10
- Leverage higher than 10 or negative
- Quarterly real sales growth above 1 or below -1
- Negative sales or liquidity

A.3 Corporate Bond Data

Table A1: Summary Statistics

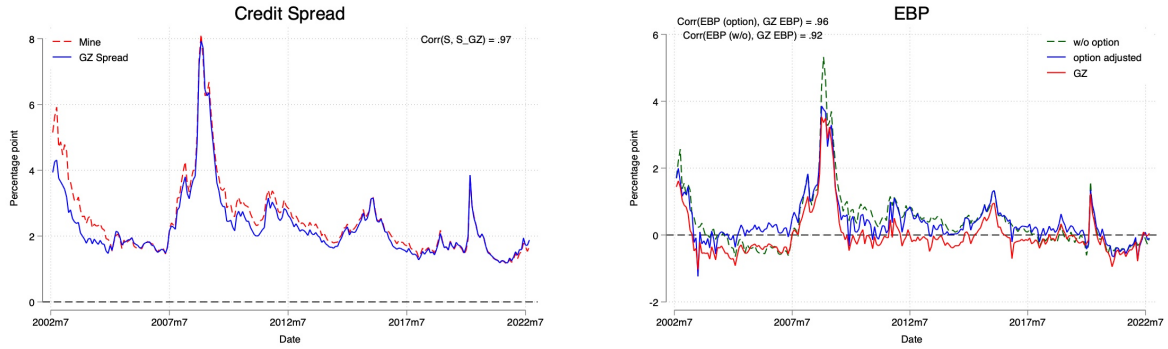
	Mean	SD	Min	p50	Max
# of Bonds per Firm/Month	4.12	5.05	1.00	2.00	89.00
Effective Yield (%)	4.79	2.87	0.23	4.42	38.96
Spread(%)	2.58	2.71	0.05	1.81	34.98
Coupon(%)	5.58	1.96	0.00	5.72	15.50
Amount Issued (M)	625.44	556.07	1.00	500.00	15000.00
Maturity At Issue (Years)	13.94	9.33	1.32	10.02	50.02
Time to Maturity (Years)	9.70	8.20	1.00	6.68	30.00
Duration (Macaulay, Years)	6.39	3.98	0.50	5.34	19.90
Credit Rating	.	.	D	BBB2	AAA
Callable (% of sample)	97.11	.	.	.	.

Note: Sample period: August 2002-December 2019. Credit spread is the yield-to-maturity over the synthetic Treasury curve. Credit rating" reports the 25th percentile, median, and 75th percentile Moody's rating category (numeric scale: Aaa = 1, C = 21). Callable bonds constitute 97.1 % of observations. Summary statistics for the 377,421 observations; Number of firms= 1,029 Number of bonds = 7,376. Exclude finance, utilities, and quasi-government firms.

Sanity Check Figure A1 validates our TRACE-based measures against the benchmark BofA-ML indices from Gilchrist and Zakrajšek (2012). In the left panel, the dashed red line plots the equal-weighted average TRACE credit spread (yield-to-maturity over the Treasury curve), while the solid blue line shows the GZ option-adjusted spread (OAS) from the BofA-ML index. The two series track one another closely–with a sample correlation of 0.97–demonstrating that our TRACE cleanup and aggregation reproduce the canonical credit-spread dynamics.

The right panel compares the excess-bond-premium (EBP) components constructed via the Merton (1974) distance-to-default framework. Again, the TRACE-based EBP (solid green) co-moves tightly with the publicly available GZ OAS-based EBP (dashed purple), both in level and through cycle. This close correspondence confirms that our TRACE-derived spreads and their decomposition faithfully replicate the GZ measures.

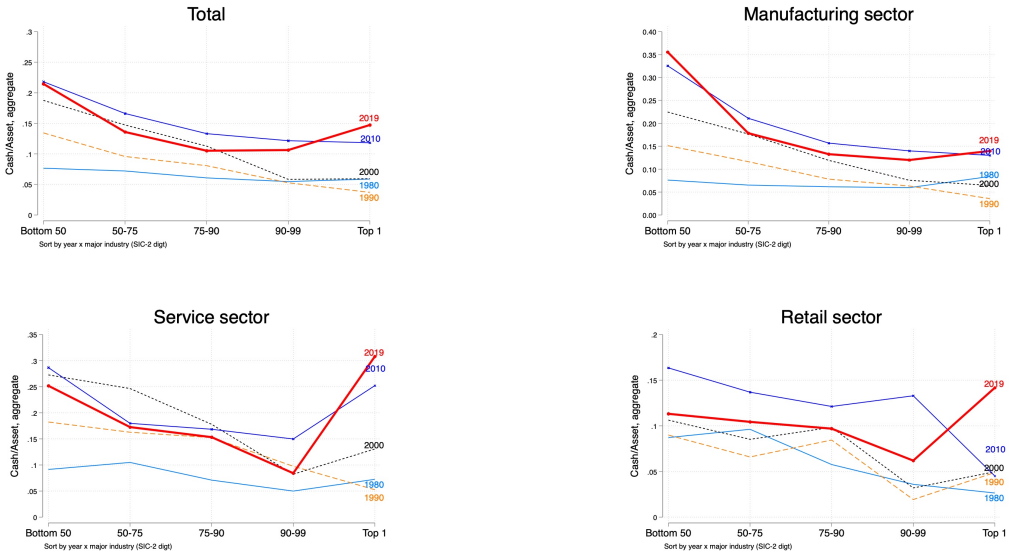
Figure A1: Comparison with GZ spreads



Notes: The figure plots aggregate credit spreads and their decomposition into excess bond premium (EBP), comparing our TRACE-based series with the canonical GZ indices. In the left panel, the blue solid line shows the option-adjusted credit spread, which strips out the value of embedded call provisions, while the green solid line shows the non-adjusted spread (yield-to-maturity over the Treasury curve). The dashed red line is the GZ OAS series derived from the BofA-ML index. In the right panel, we decompose each series into expected-default and EBP components using the Merton (1974) distance to default framework: the blue solid line is our TRACE-based option-adjusted EBP, the green solid line is the non-adjusted EBP, and the dashed red line is the GZ EBP. Since most bonds in our sample are callable–and therefore sensitive to risk-free rate movements–option adjustment is crucial for isolating pure credit risk; the close co-movement across all measures confirms that this adjustment has minimal impact on cyclical dynamics.

A.4 Additional empirics

Figure A2: Distribution of Corporate Cash: Cash-to-Assets ratio



Notes: Each panel plots the cross-sectional distribution of cash-to-assets ratios across firm size percentiles (bottom 50%, 50-75%, 75-90%, 90-99%, and top 1%) in selected years. The top-left panel reports the distribution for the full non-financial sample. The remaining panels restrict attention to specific sectors–manufacturing, services, and retail–based on 2-digit SIC classifications. Across sectors, the distribution becomes increasingly U-shaped over time, with both the smallest and the largest firms holding more cash relative to mid-sized firms. Financial firms, utilities, and quasi-government entities are excluded from the sample.

Table A2: Firm Size and Market Beta, Manufacturing sector only

Market Equity	Beta Quartile	Market Equity ($ M)	Beta	Cash-to-Assets	Leverage	Eq. issuance	Debt issuance
$< 90\%$	Q1	289	0.30	19.22	22.44	2.25	0.85
	Q2	506	0.77	18.71	21.89	1.66	1.05
	Q3	631	1.10	20.03	21.73	1.73	1.09
	Q4	677	1.58	24.00	21.25	2.14	1.06
$90\% - 99\%$	Q1	19435	0.44	10.79	28.42	-0.75	1.74
	Q2	15711	0.80	10.34	25.24	-0.64	1.22
	Q3	11501	1.09	10.55	24.21	-0.48	1.16
	Q4	8939	1.51	16.20	20.54	-0.03	1.38
$> 99\%$	Q1	139915	0.46	11.70	23.15	-0.73	0.89
	Q2	93470	0.76	10.89	19.79	-0.81	0.69
	Q3	121761	1.06	14.83	19.24	-1.00	0.71
	Q4	139972	1.48	20.78	11.34	-1.04	0.68

Notes: This table reports summary statistics for three firm size groups within the manufacturing sector over the sample period 1970Q1-2019Q4. The sample is restricted to firms with primary 2-digit SIC codes between 20 and 39. Within each size group, firms are sorted into quartiles based on their firm-level equity beta, reported in the second column. Column (3) presents the average market equity (in millions of dollars) for each size-beta subgroup. Columns (5) through (8) report average values of key financial variables, expressed as percentages.

Table A3: Full regression results for Table 2

	(1) cheat_rat	(2) cheat_rat	(3) cheat_rat
L.Market Beta	2.267***	1.789***	0.735***
	(0.26)	(0.16)	(0.13)
L.Size		-0.158	-1.655***
		(0.10)	(0.28)
L.Tobin's Q		0.413***	0.163***
		(0.05)	(0.05)
L.CF volatility		0.557***	0.492***
		(0.08)	(0.08)
L.CF ratio		-0.476***	-0.121***
		(0.04)	(0.03)
L.Sale growth		0.000	-0.000**
		(0.00)	(0.00)
L.Firm age		-0.085***	
		(0.01)	
L.Dividend payer dummy		-4.123***	0.237
		(0.37)	(0.35)
L.Leverage		-0.123***	-0.091***
		(0.01)	(0.01)
L.Tangibility		-7.578***	-3.572**
		(1.89)	(1.51)
L.Acquisition		-0.070***	-0.039***
		(0.01)	(0.01)
L.R&D Intensity		0.506***	-0.159***
		(0.02)	(0.03)
L.Distance to Default		0.635***	0.456***
		(0.04)	(0.03)
N	177519	177519	177519
R^2	0.095	0.378	0.723
Firm controls	no	yes	yes
Firm FE	no	no	yes
Time X Sector FE	yes	yes	yes
Firm and Industry x Time Two-way clustering	yes	yes	yes

Table A4: Full regression results for Table 3

	(1) cheat_rat	(2) cheat_rat
L.Market Beta	1.901***	0.273
	(0.24)	(0.39)
Size $(50-75\%)$	-0.064	5.038***
	(0.47)	(0.43)
Size $(75-90\%)$	-3.798***	4.770***
	(0.59)	(0.66)
Size $(90-99\%)$	-4.876***	6.717***
	(0.63)	(0.86)
Size $(>99\%)$	-2.589*	10.397***
	(1.54)	(1.57)
$(50-75\%)$ × L.Market Beta	0.535	1.675***
	(0.36)	(0.31)
$(75-90\%)$ × L.Market Beta	1.623***	3.217***
	(0.45)	(0.45)
$(90-99\%)$ × L.Market Beta	2.679***	4.793***
	(0.53)	(0.61)
$(>99\%)$× L.Market Beta	3.507**	5.467***
	(1.67)	(1.45)
L.Size		-1.266***
		(0.15)
L.Tobin's Q		0.379***
		(0.05)
L.Leverage		-0.121***
		(0.01)
L.CF volatility		0.523***
		(0.07)
L.Sale growth		0.000*
		(0.00)
L.CF ratio		-0.497***
		(0.04)
L.Distance to Default		0.525***
		(0.04)
L.Tangibility		-9.834***
		(1.82)
L.Acquisition		-0.073***
		(0.01)
L.R&D Intensity		0.466***
		(0.02)
L.Dividend payer dummy		-4.086***
		(0.37)
L.Firm age		-0.082***
		(0.01)
L.Market Beta × L.Size		-0.442***
		(0.09)
L.Market Beta × L.Tobin's Q		-0.104***
		(0.04)
L.Market Beta × L.Leverage		0.017**
		(0.01)
L.Market Beta × L.CF Vol		0.076
		(0.05)
L.Market Beta × L.Sale Growth		-0.000*
		(0.00)
L.Market Beta × L.CF ratio		0.006
		(0.03)
L.Market Beta × L.D2D		0.061**
		(0.03)
L.Market Beta × L.Tangibility		-5.491***
		(1.36)
L.Market Beta × L.Acquisition		-0.018**
		(0.01)
L.Market Beta × L.R&D Intensity		0.067***
		(0.01)
L.Market Beta × L.Dividend Payer Dummy		-1.138***
		(0.28)
L.Market Beta × L.Age		-0.004
		(0.00)
N	177738	177738
R^2	0.105	0.391
Firm controls	no	yes
Firm x Beta	no	yes
Time X Sector FE	yes	yes
Firm and Industry X Time Two-way clustering	yes	yes

Table A5: Full regression results for Table 4

	(1) Leverage	(2) Leverage	(3) Long-term debt	(4) Long-term debt
L.Market Beta	-0.671***	-2.171***	0.131	-1.059***
	(0.24)	(0.47)	(0.21)	(0.40)
Size $(50-75\%)$	-0.200	-1.693***	4.409***	0.899**
	(0.50)	(0.49)	(0.42)	(0.43)
Size $(75-90\%)$	0.827	-2.293***	6.436***	0.165
	(0.65)	(0.77)	(0.54)	(0.69)
Size $(90-99\%)$	0.874	-5.276***	6.186***	-3.786***
	(0.74)	(1.10)	(0.57)	(0.93)
Size $(>99\%)$	-4.547***	-13.885***	1.682*	-12.953***
	(1.23)	(1.81)	(0.98)	(1.58)
$(50-75\%)$ × L.Market Beta	0.473	0.046	-0.107	-0.124
	(0.39)	(0.38)	(0.34)	(0.34)
$(75-90\%)$ × L.Market Beta	-0.580	-1.299**	-1.397***	-1.371***
	(0.48)	(0.54)	(0.43)	(0.50)
$(90-99\%)$ × L.Market Beta	-1.978***	-3.035***	-1.846***	-2.041***
	(0.67)	(0.84)	(0.53)	(0.72)
$(>99\%)$× L.Market Beta	-2.635**	-3.040**	-2.494***	-2.151*
	(1.15)	(1.40)	(0.87)	(1.18)
L.Size		3.977***		3.949***
		(0.20)		(0.18)
L.Sale growth		-0.000		0.000
		(0.00)		(0.00)
L.Firm age		-0.013		-0.019**
		(0.01)		(0.01)
L.Tobin's Q		0.197***		0.154***
		(0.05)		(0.04)
L.Distance to Default		-1.906***		-1.252***
		(0.05)		(0.04)
L.CF volatility		0.580***		0.339***
		(0.08)		(0.07)
L.Industry Sigma		0.081		0.139
		(0.12)		(0.10)
L.Acquisition		0.122***		0.122***
		(0.01)		(0.01)
L.R&D Intensity		-0.077***		-0.054***
		(0.02)		(0.02)
L.Dividend payer dummy		-0.191		-0.331
		(0.44)		(0.40)
L.Market Beta × L.Size		0.198		0.174
		(0.14)		(0.12)
L.Market Beta × L.Sale Growth		0.000**		-0.000
		(0.00)		(0.00)
L.Market Beta × L.Age		-0.008		-0.006
		(0.01)		(0.01)
L.Market Beta × L.Tobin's Q		-0.058*		-0.032
		(0.03)		(0.03)
L.Market Beta × L.D2D		-0.117***		-0.150***
		(0.04)		(0.03)
L.Market Beta × L.CF Vol		0.026		0.023
		(0.06)		(0.06)
L.Market Beta × L.Industry Sigma		0.176**		0.152**
		(0.08)		(0.07)
L.Market Beta × L.Acquisition		0.006		0.011
		(0.01)		(0.01)
L.Market Beta × L.R&D Intensity		0.044***		0.025**
		(0.01)		(0.01)
L.Market Beta × L.Dividend Payer Dummy		0.741**		0.721**
		(0.37)		(0.34)
N	180445	180445	170577	170577
R^2	0.042	0.282	0.061	0.232
Firm controls	no	yes	no	yes
Firm x Beta	no	yes	no	yes
Time X Sector FE	yes	yes	yes	yes
Firm FE	no	no	no	no
Firm and Industry x Time Two-way clustering	yes	yes	yes	yes

Appendix B Proofs

Proof of Lemma 1 We establish Lemma 1 in two steps. First, under a one-period horizon–with the investor maximizing next-period wealth W_{t+1} and a second-order approximation of utility-the optimal portfolio weights coincide with the familiar Markowitz solution. Second, when excess returns are i.i.d. or the investor has log utility–so that the consumption-to-wealth ratio remains essentially constant–intertemporal hedging demands drop out and the infinite-horizon Bellman equation produces the same first-order condition as the one-period case (see Campbell and Viceira (2002), Sec. 2). Hence, the long-term optimal weights coincide with the short-term Markowitz rule. Accordingly, it suffices to derive the one-period result, which immediately extends to the infinite-horizon setting.

Suppose the investor cares only about next-period wealth W_{t+1} and chooses portfolio weights α_t to maximize expected utility of next period wealth, $\mathbb{E}_t u(W_{t+1})$. The optimization problem is to maximize $\mathbb{E}_t[u(W_t - c_t)(r_{Ft+1} + \alpha_t'(r_{t+1} - r_{Ft+1})]$ with respect to α_t. The first-order condition for each bond i is:

$$\mathbb{E}_t[u'(W_{t+1}) \cdot (r_{it+1} - r_{Ft+1})] = 0 \quad \forall\, i = 1, 2, ..., N \tag{B.1}$$

The FOC is difficult to solve analytically because the optimal weight α_t is inside the non-linear function $u'(W_{t+1})$. To make progress, we perform a second-order Taylor expansion of the marginal utility function $u'(W_{t+1})$ around the wealth level that would result from investing fully in the fringe index, $W_F = (W_t - c_t)R_{Ft+1}$.

The wealth can be expressed as $W_{t+1} = W_F + (W_t - c_t)\alpha_t'\tilde{r}_{t+1}$ where $\tilde{r}_{t+1} = R_{t+1} - R_{Ft+1}\iota$ is excess return over fringe index. The expansion of u' around the benchmark wealth gives

$$u'(W_{t+1}) \approx u'(W_F) + u''(W_F) \cdot [(W_t - c_t)\alpha_t'\tilde{r}_{t+1}] \tag{B.2}$$

Substituting B.2 back into Equation B.1 and using the linearity of expectations yields:

$$\mathbb{E}_t[u'(W_F)\tilde{r}_{t+1}] + \mathbb{E}_t[u''(W_F)(W_t - c_t)\tilde{r}_{t+1}\tilde{r}_{t+1}']\alpha_t \approx 0 \tag{B.3}$$

$$\mathbb{E}_t[u'(W_F)]\mathbb{E}_t\tilde{r}_{t+1} + (W_t - c_t)\mathbb{E}_t[u''(W_F)]\mathbb{E}_t[\tilde{r}_{t+1}\tilde{r}_{t+1}']\alpha_t \approx 0 \tag{B.4}$$

where the second equation is expanded under the assumption that the fringe return r_{Ft+1} is uncorrelated with the excess return $\tilde{r}_{t+1}$, because the fringe index is a fully diversified portfolio of small (atomistic) firms, its return is uncorrelated with the excess returns of the granular bonds.

Rearranging the equation to solve for α_t gives:

$$\alpha_t \approx \underbrace{\Big(- \frac{\mathbb{E}_t u'(W_F)}{(W_t - c_t)\mathbb{E}_t u''(W_F)} \Big)}_{\approx 1/\gamma} \Sigma_t^{-1} \mu_t \tag{B.5}$$

where $\Sigma_t = \mathbb{E}_t \tilde{r}_{t+1} \tilde{r}'_{t+1}$, $\mu_t = \mathbb{E}_t \tilde{r}_{t+1}$. Under CRRA utility $u(W) = \frac{W^{1-\gamma}}{1-\gamma}$ and assuming that fringe index return volatility is negligible (by full diversification), we arrive at the final expression as in the main text.

Proof of Proposition 1 We derive the firm's Euler conditions for borrowing and saving under a bond pricing schedule that embeds endogenous price impact. Let $i \equiv \widehat{b}' - (1-\delta)b$ denote net new bond issuance. The firm chooses $\widehat{a}'$ and $\widehat{b}'$ to maximize continuation value, taking into account its influence on the bond price $q(\widehat{b}', \widehat{a}', s)$.

- **FOC with respect to $\widehat{a}'$:**

$$q_a - \underbrace{\frac{\partial q(\widehat{b}', \widehat{a}', s)}{\partial \widehat{a}'}}_{>0} \cdot i \geq \beta \left[\int_{\eta'} \int_{\varepsilon^{*\prime}(b', a')}^{\infty} \left(\frac{u(d')}{u(d)} \right) dG(\varepsilon') dF(\eta') \right] \qquad \text{(equality if } a' > 0)$$

- **FOC with respect to $\widehat{b}'$:**

$$q(\widehat{b}', \widehat{a}', s) + \underbrace{\frac{\partial q(\widehat{b}', \widehat{a}', s)}{\partial b'}}_{<0} \cdot i = \beta \Bigg[\int_{\eta'} \int_{\varepsilon^{*\prime}(b', a')}^{\infty} \frac{u(d')}{u(d)} \times \{c + \delta + (1-\delta) q(b'', a'', s')\} \, dG(\varepsilon') dF(\eta') \Bigg]$$

where

$$b'' = \widehat{b}(\widehat{b}', \widehat{a}', s'), \quad a'' = \widehat{a}(\widehat{b}', \widehat{a}', s'),$$

The left-hand side of the FOC for $\widehat{a}'$ captures the marginal cost of saving: purchasing one more unit of the risk-free asset at price q_a, net of its effect on the firm's bond price via $\frac{\partial q}{\partial \widehat{a}'} > 0$. Accumulating more liquidity improves the bond price by reducing both the expected default probability and its variance. Hence, this indirect improvement in pricing incentivizes firms to save more than they would if they took $q(\cdot)$ as given.

The left-hand side of the FOC for $\widehat{b}'$ captures the marginal benefit of borrowing: the proceeds from issuing one unit of debt at price $q(\cdot)$, net of the negative price impact $\frac{\partial q}{\partial \widehat{b}'} < 0$, which reflects that higher issuance lowers the value of all outstanding debt. Thus, the

granular firm internalizes that borrowing more depresses the price of its bonds due to investor risk aversion.

Together, these two FOCs formalize the price-kernel mechanism: relative to an atomistic firm facing constant prices, a granular firm perceives higher marginal cost of debt and greater marginal benefit of liquidity. As a result, it chooses lower $\widehat{b}'$ (underborrowing) and higher $\widehat{a}'$ (oversaving), conditional on default probability.

To derive further the sign of $\frac{dq}{da'}$ and $\frac{dq}{db'}$, we examine the pricing schedules from Equation 25

$$q(\widehat{b}', \widehat{a}', s) = \frac{1}{\overline{R}_F}\, \mathbb{E}_{s'|s}\Big[\, \mathcal{R}(\widehat{b}', \widehat{a}', s')\; \Phi(\widehat{b}', \widehat{a}', s)\Big],$$

where

$$\Phi(\widehat{b}', \widehat{a}', s) = 1 - \gamma\, \frac{\mathrm{Var}_{s'|s}[\mathcal{R}]}{\mathbb{E}_{s'|s}[\mathcal{R}]}\, \frac{\widehat{b}'}{\widehat{W}}, \qquad 0 < \Phi < 1. \tag{$*$}$$

Holding the exogenous state s fixed:

$$\frac{\partial q}{\partial \widehat{a}'} = \frac{1}{R_F}\Big\{ \underbrace{\mathbb{E}\big[\tfrac{\partial \mathcal{R}}{\partial \widehat{a}'}\, \Phi\big]}_{\text{(A)}} + \underbrace{\mathbb{E}\big[\mathcal{R}\big]\, \frac{\partial \Phi}{\partial \widehat{a}'}}_{\text{(B)}} \Big\}, \tag{B.6}$$

$$\frac{\partial q}{\partial \widehat{b}'} = \frac{1}{R_F}\Big\{ \underbrace{\mathbb{E}\big[\tfrac{\partial \mathcal{R}}{\partial \widehat{b}'}\, \Phi\big]}_{\text{(C)}} + \underbrace{\mathbb{E}\big[\mathcal{R}\big]\, \frac{\partial \Phi}{\partial \widehat{b}'}}_{\text{(D)}} \Big\}. \tag{B.7}$$

Sign of $\partial q/\partial \widehat{a}'$. Increasing cash raises the continuation value, lowering the default threshold: $\partial \varepsilon^{*\prime}/\partial \widehat{a}' < 0$. Hence the survival probability and expected repayment increase, so termă(A)ă> 0. From $(*)$, Φ does not depend directly on $\widehat{a}'$ (only through moments of $\mathcal{R}$); incorporating the covariance term shows (B) ≥ 0. Therefore $\partial q/\partial \widehat{a}' > 0$.

Sign of $\partial q/\partial \widehat{b}'$. Higher debt raises the default threshold ($\partial \varepsilon^{*\prime}/\partial \widehat{b}' > 0$), so expected repayment falls: term (C) < 0. Moreover, from $(*)$,

$$\frac{\partial \Phi}{\partial \widehat{b}'} = -\gamma\, \frac{\mathrm{Var}_{s'|s}[\mathcal{R}]}{\mathbb{E}_{s'|s}[\mathcal{R}]}\, \frac{1}{\widehat{W}} \; < 0,$$

so term (D)< 0 [14] Thus $\partial q/\partial \widehat{b}' < 0$.

Comparison to a price-taker. For a competitive-fringe (price-taking) firm, the red derivative terms $\partial q/\partial \widehat{a}'$ and $\partial q/\partial \widehat{b}'$ vanish. Because $\partial q/\partial \widehat{b}' < 0$ and $\partial q/\partial \widehat{a}' > 0$, the granular firm perceives a higher marginal cost of debt and a higher marginal benefit of liquidity than the

[14] Strictly speaking, $\widehat{b}'$ also affects the conditional moments $\mathbb{E}_{s'|s}[\mathcal{R}]$ and $\mathrm{Var}_{s'|s}[\mathcal{R}]$; these additional derivative terms reinforce the same sign, so we omit them for brevity.

fringe firm. Holding fixed the same default probability, the inequalities force the granular firm to choose

$$\widehat{b}'_{\text{granular}} \leq \widehat{b}'_{\text{fringe}}, \qquad \widehat{a}'_{\text{granular}} \geq \widehat{a}'_{\text{fringe}},$$

establishing *under-borrowing* and *over-saving* relative to a price-taking counterpart. □

Appendix C Numerical Algorithms

The model is solved using a value function iteration scheme, building on techniques developed in Hatchondo, Martinez and Sosa-Padilla (2016) and Poeschl (2023). Because the equilibrium of the infinite-horizon model may not be unique–due to endogenous default and bond pricing feedback–we approximate the solution by backward induction from a finite terminal horizon. That is, we compute the value and policy functions recursively by using a fixed terminal value function and a corresponding terminal bond price as initial guesses for the backward pass.

Baseline case ($\gamma = 1$)

Since there is no aggregate risk and for fringe firms their idiosyncratic risks are washed out in the index, q_F is deterministic in a stationary equilibrium. Hence, we only need to solve for granular firm's problem given q_F, thus $\overline{R}_F$.

General Case ($\gamma \neq 1$)

Since equilibrium bond prices depend on the representative investor's wealth W, which in turn influences the gross return on the fringe index (the benchmark return used to price all debt), we adopt a *Krusell-Smith*-type algorithm. This involves positing and updating a forecasting rule for $\ln R_{F,t+1}$ as a function of $\ln W_t$ and solving the model jointly with this approximate law of motion.

The overall procedure is composed of an outer fixed-point iteration on the return forecasting coefficients (Krusell-Smith loop) and an inner value function iteration conditional on those coefficients.

Model Solution

1. **Aggregate Forecasting Rule.** The aggregate return on the fringe index $R_{F,t+1}$ is approximated by a log-linear forecasting rule:

$$\ln R_{F,t+1} = \theta_0 + \theta_1 \ln W_t, \tag{C.1}$$

   where W_t denotes aggregate investor wealth, and $\Theta = (\theta_0, \theta_1)^\top$ is estimated using an outer Krusell-Smith fixed-point iteration. In implementation, (C.1) is expressed as a degree-1 Chebyshev polynomial in $\ln W$ via `fundef`.

2. **Outer Loop (Krusell-Smith Fixed Point).**

(a) **Initialization.** Set initial guess $\Theta^{(0)} = (\ln 1.01, 0)^\top$.

(b) **Iterate for** $k = 0, 1, 2, \ldots$ **until convergence:**

i. *Solve Firm Problems.* For each firm, solve the value and policy functions on a discrete (b, a, s) grid taking $R_F^{(k)}(W)$ from (C.1) as given. The bond price function is computed as

$$q^{(k)}(s, \widehat{b}') = \frac{\mathbb{E}_t[R_{t+1}]}{R_F^{(k)}(W)} \left[1 - \frac{\mathbb{V}_t[R_{t+1}]}{\mathbb{E}_t[R_{t+1}]} \cdot \frac{\gamma}{W} \cdot \widehat{b}'\right],$$

where R_{t+1} is the firm's state-contingent repayment and γ denotes investor risk aversion.

ii. *Simulate Economy.* Simulate P independent panels of T periods. At each t:

- Update firm states using policy rules;
- Solve the market-clearing condition $g(W) = q_F(W)B_F + \sum_i q_i(W)B_i - W = 0$ for W;
- Record $(\ln R_{F,t+1}, \ln W_t)$, with $R_{F,t+1}^{-1} = \mathbb{E}_t[R_{F,t+1}]$ and $R_{F,t+1} = c + \delta + (1 - \delta)q_{F,t+1}$.

iii. *Update Coefficients.* Run pooled OLS: $\ln R_{F,t} = \theta_0 + \theta_1 \ln W_t + \varepsilon_t$ to obtain $\Theta^{(k+1)}$.

iv. *Convergence Test.* If $\|\Theta^{(k+1)} - \Theta^{(k)}\| < 10^{-4}$, stop; else, set $k \leftarrow k + 1$ and repeat.

Typically, convergence is achieved within 3-6 iterations, with final $R^2 > 0.999$ in the forecasting rule.

3. **Inner Loop (Value Function Iteration).** Fix Θ and solve for firm policies:

(a) Start with initial guesses for expected value $\mathbb{E}V(\widehat{b}', \widehat{a}')$, $\mathbb{E}[R]$, and $\mathrm{Var}[R]$.

(b) Solve for optimal policies $(\widehat{b}', \widehat{a}', e)$ and value function $v(b, a, s)$.

(c) Compute bond price $q(b, a, s; \widehat{b}', \widehat{a}')$ from the policy rules and pricing formula.

(d) Evaluate $\mathbb{E}V$ and $\mathbb{E}[R]$, $\mathrm{Var}[R]$ using:

- Gauss-Legendre quadrature (30 nodes) for continuous ε;
- Tauchen-discretized grid (5 points) for permanent shock η;
- Default cutoffs ε^* for piecewise integration.

(e) Iterate until the changes in expectations and bond prices fall below 10^{-4}, or a cap of 300 iterations is reached.

Simulation Procedure We simulate a panel of $N_G = 40$ granular firms and 1000 fringe firms for $T = 2000$ periods, discarding the first 1000 periods as burn-in. Defaulted firms are replaced with entrants initialized at $Y^P = 1$, $b = a = 0$.

- **Wealth dynamics.** Aggregate investor wealth evolves according to:

$$W_{t+1} = (1 - \kappa) R_{p,t+1} W_t, \quad \text{with } \kappa = 0.02,$$

  where $R_{p,t+1}$ is the investor portfolio return (weighted average across bonds and fringe index).

- **VIX Shock Experiment.** At period t^*, impose a volatility-induced wealth contraction:

$$W_{t^*} \leftarrow (1 + \Delta) W_{t^*}, \quad \Delta = -0.15,$$

  and trace the impulse responses of credit spreads, cash ratios, and leverage differentials across granular and fringe firms.

Appendix D Calibration / Estimation

D.1 Estimation of Permanent and Transitory Shock Variances

We model firm earnings Y_t as the product of a permanent component $Y_{P,t}$ and a transitory component $Y_{T,t}$, so that

$$Y_t = Y_{P,t} \times Y_{T,t}.$$

The permanent component follows a random walk in logs,

$$\log Y_{P,t} = \log Y_{P,t-1} + \eta_t, \qquad \eta_t \sim \mathcal{N}(0, \sigma_\eta^2),$$

while the transitory component is a one-period multiplicative shock,

$$Y_{T,t} = 1 + \epsilon_t, \qquad \epsilon_t \sim \mathcal{N}(0, \sigma_\epsilon^2).$$

The innovations η_t and ϵ_t are mutually independent and independent over time.

Taking logarithms and first differences of Y_t yields the transformation

$$\Delta \log Y_t \equiv \log Y_t - \log Y_{t-1} = \eta_t + \epsilon_t - \epsilon_{t-1}.$$

Because η_t and ϵ_t are uncorrelated white-noise processes, this differenced series has zero mean

and a moving-average (MA(1)) covariance structure. In particular, the variance of $\Delta \log Y_t$ and its lag-one autocovariance satisfy

$$\gamma(0) = \mathrm{Var}(\Delta \log Y_t) = \sigma_\eta^2 + 2\,\sigma_\epsilon^2, \qquad \gamma(1) = \mathrm{Cov}(\Delta \log Y_t, \Delta \log Y_{t-1}) = -\sigma_\epsilon^2 \tag{C.2}$$

These expressions follow from the fact that $\Delta \log Y_t = \eta_t + (\epsilon_t - \epsilon_{t-1})$; the cross-products between contemporaneous innovations vanish because η_t and ϵ_t are independent, while the negative autocovariance arises from the term $-\epsilon_{t-1}$. Higher-order autocovariances are zero, implying that $\Delta \log Y_t$ is an MA(1) process.

Given observed quarterly earnings data, we estimate σ_η and σ_ϵ by matching these theoretical moments to sample moments for each group of firms. For a given group (that is, the top 1% granular firms or the remaining fringe firms), let $\widehat{\gamma}_0$ denote the sample variance of $\Delta \log Y_t$ and $\widehat{\gamma}_1$ denote the sample autocovariance at lag one. Under the model, these statistics satisfy the system $\widehat{\gamma}_0 \approx \sigma_\eta^2 + 2\sigma_\epsilon^2$ and $\widehat{\gamma}_1 \approx -\sigma_\epsilon^2$ according to Equation (C.2). Solving for σ_η and σ_ϵ yields the closed-form estimators

$$\hat{\sigma}_\epsilon = \sqrt{-\widehat{\gamma}_1}, \qquad \hat{\sigma}_\eta = \sqrt{\max\{\widehat{\gamma}_0 - 2\,\hat{\sigma}_\epsilon^2,\, 0\}}.$$

In practice, we compute quarterly earnings (EBITDA) for each firm, take the log difference by firm, and then calculate $\widehat{\gamma}_0$ and $\widehat{\gamma}_1$ for each group over the 1980q1-2019q4 sample. Substituting these sample moments into the formulas above yields group-specific estimates of the standard deviation of permanent earnings innovations ($\hat{\sigma}_\eta$) and transitory earnings shocks ($\hat{\sigma}_\epsilon$). The grouping ensures that the top 1% of firms by market capitalization (granular firms) and the remaining firms (fringe firms) can exhibit different income-volatility profiles in line with theoretical expectations. The estimated parameter set $(\hat{\sigma}_\epsilon^G, \hat{\sigma}_\eta^G) = (0.1516,\, 0.1469)$ and $(\hat{\sigma}_\epsilon^F, \hat{\sigma}_\eta^F) = (0.4164,\, 0.3736)$ for granular and fringe firms, respectively.