Borrowing Constraints, Markups, and Misallocation*

Huiyu Li **Chen Lian** **Yueran Ma** **Emily Martell**

March 12, 2025

Abstract

We document new facts that link firms' markups to borrowing constraints: (1) firms with looser constraints have higher markups, especially in industries where assets are difficult to borrow against and firms rely more on earnings to borrow; (2) markup dispersion is also higher in industries where firms rely more on earnings to borrow. We explain these relationships using a standard Kimball demand model augmented with borrowing against assets and earnings. The key mechanism is a two-way feedback between markups and borrowing constraints. First, firms with looser constraints charge higher markups, as looser constraints allow them to attain larger market shares. Second, higher markups relax borrowing constraints when firms rely on earnings to borrow, as those with higher markups have higher earnings. The interaction between markups and borrowing constraints has important allocative efficiency implications. High-markup firms can be too large because they face looser borrowing constraints. Introducing borrowing constraints also lowers the overall TFP losses from markup dispersion.

*Li: Federal Reserve Bank of San Francisco, Huiyu.Li@sf.frb.org; Lian: UC Berkeley and NBER, chen_lian@berkeley.edu; Ma: Uchicago and NBER, yueran.ma@chicagobooth.edu; Martell: UC Berkeley, emily_martell@berkeley.edu. We thank Corina Boar, Joel David, Chad Jones, Pete Klenow, Virgiliu Midrigan, Michael Peters, Kunal Sangani, Yongseok Shin, Venky Venkateswaran, Tom Winberry, and seminar participants at the ASSA, Federal Reserve Bank of Richmond, Federal Reserve Bank of San Francisco, Barcelona Summer Forum, Northwestern University, Swiss Annual Macro Conference, UC Berkeley, and VII MadMac Annual Conference. The views in this paper are solely the responsibility of the authors and should not be interpreted as reflecting the views of the Federal Reserve Bank of San Francisco or the Board of Governors of the Federal Reserve System or the U.S. Census Bureau. The Census Bureau has reviewed this data product to ensure appropriate access, use, and disclosure avoidance protection of the confidential source data used to produce this product. This research was performed at a Federal Statistical Research Data Center under FSRDC Project Number 2441 (CBDRB-FY25-P2441-R11891). Yueran Ma acknowledges that this material is based upon work supported by the NSF under Grant # 2144769.

1 Introduction

In a number of economic models (e.g., Edmond, Midrigan, and Xu, 2023; Baqaee, Farhi, and Sangani, 2024a,b), variations in firms' markups are driven by differences in their productivity. This is because, under the standard Kimball demand aggregator, high-productivity firms have larger market shares and hence charge higher markups. In this case, markup dispersion leads to TFP loss, as high markup firms are always too small.

In this paper, we present empirical and theoretical analyses showing that firms' markups can also reflect the tightness of their borrowing constraints. Firms with looser borrowing constraints have larger market shares and hence charge higher markups. However, the allocative efficiency implications are different. High markup firms that face looser constraints, rather than having high productivity, may become too large. Moreover, introducing borrowing constraints lowers the overall TFP losses from markup dispersion.

We first provide new empirical evidence that links firms' markups with their borrowing constraints, using both US Compustat data for public firms and the U.S. Census of Manufacturing data. We present the following findings. We start by documenting that less constrained firms in an industry have higher markups (Fact 1a). For measures of markups, we use several estimates from De Loecker, Eeckhout, and Unger (2020) as well as translog markups following De Ridder (2024). For measures of constraint tightness, our baseline analysis uses cash holdings (normalized by book assets), which is commonly used in the literature (for example, in Gilchrist et al. (2017)): firms with abundant cash holdings can use internal funds, whereas firms with little cash need to rely on external borrowing. We also consider alternative measures, such as log capital wedge (production worker hours over capital) which aligns with our model subsequently.

Furthermore, we find that the relationship between markups and borrowing constraint tightness is stronger among industries where firms have low liquidation values of pledgeable assets and therefore rely more on earnings to borrow (Fact 1b). Markup dispersion is also higher in industries that rely more on earnings to borrow (Fact 2). We verify that these industry-level relationships hold for both US Compustat data and the Census of manufacturing data.

We then show how these firm-level and industry-level relationships between firms' markups and their borrowing constraints can be quantitatively explained by a standard Kimball demand model augmented with borrowing against assets and earnings. In this model, the competitive final good producer aggregates differentiated intermediate goods through a Kimball aggregator. Monopolistically competitive intermediate input producers use variable inputs and capital and are subject to asset- and earnings-based borrowing constraints. Intermediate input producers not only optimally choose the amount of variable inputs and capital used for production but also endogenously choose their saving.

The key mechanism in the model is a two-way feedback between markups and constraint tightness. The first step of this feedback applies in general to both asset-based and earnings-based constraints. Firms with looser borrowing constraints have lower marginal costs and, hence, higher market shares and markups. The second step of this feedback operates only under earnings-based borrowing constraints. In that case, firms with higher markups have higher earnings, which loosen their earnings-based borrowing constraints. This step is absent under asset-based borrowing constraints. Both steps of the two-way feedback explain Fact 1a. The observation that the second step of the two-way feedback operates only under earnings-based borrowing constraints explains Fact 1b and Fact 2.

The interaction between markups and borrowing constraints has important allocative efficiency implications. High-markup firms can be too large if they are not more productive, but face looser constraints. Introducing borrowing constraints also lowers the overall TFP losses from markup dispersion. For this purpose, we rewrite the Hsieh and Klenow (2009) formula for TFP loss (and extend it to the case with Kimball demand) and express the TFP loss as a function of the dispersion in log markup and constraint tightness (measured in log capital wedge) and the covariance between log markup and constraint tightness.

$$\begin{aligned}\text{TFP loss} \quad &\approx \quad \underbrace{\text{log markup dispersion}}_{>0} \quad + \quad \text{dispersion in constraint tightness (measured in log capital wedge)} \\ &\quad + \underbrace{\text{covariance of log markup \& constraint tightness}}_{<0 \text{ if higher markup firms less constrained}} \end{aligned} \tag{1}$$

This formula illustrates how introducing borrowing constraints mitigates the TFP loss from markup dispersion. Without borrowing constraints, markup dispersion leads to TFP loss as high-markup firms are too small, as summarized by the first term in (1). With borrowing constraints, the last term in the TFP formula (1) mitigates the TFP loss from markup dispersion because high-markup firms are less constrained (Fact 1a) as a result of the two-way feedback, and this term is negative. Intuitively, as looser borrowing constraints help originally too-small high-markup firms grow, they can mitigate the TFP loss from markup dispersion.

Furthermore, the cross-industry differences documented in Fact 1b and Fact 2 also matter for how much the borrowing constraint channel mitigates the TFP loss from markup dispersion. In industries that rely more on earnings to borrow, markup dispersion is higher, but markups are also more negatively correlated with constraint tightness. As a result, all else equal, the net TFP loss from markup dispersion, defined as the sum of the first and third terms in (1) can be smaller in those industries.

Several papers find that financially constrained firms raise prices in response to temporary negative shocks (Chevalier and Scharfstein, 1995; Gilchrist et al., 2017; Meinen and Soares, 2022; Minseog and Park, 2024; Renkin and Züllig, 2024). The interpretation is that firms intertemporally substitute their investment in building the customer base, and they cut back on such investment when their financial constraints tighten. Meanwhile, Kim (2021) finds that firms cut prices to liquidate their inventory in response to negative credit supply shocks due to their banks' exposure to the Lehman crisis. Our analyses instead focus on long-run steady-state variations in markups across firms. For example, constrained firms cannot always underinvest in customer capital in the steady state (even though they can cut back in bad times and engage in intertemporal substitution). Eventually, they will have low market share (Renkin and Züllig, 2024), and correspondingly lower long-term markups.

Related literature Our paper is related to two streams of literature. One stream is the work on misallocation due to markup dispersion. Recent research finds significant productivity loss due to misallocation from markup dispersion. Baqaee and Farhi (2020) estimate that eliminating all markup dispersion across firms could raise aggregate TFP by as much as 15%. Edmond, Midrigan, and Xu (2023) calculates smaller but economically significant TFP losses of 2 to 6% due to misallocation from markup dispersion. In both papers, high markup firms are too small. Meanwhile, Aghion, Bergeaud, Boppart, Klenow, and Li (2022) feature multiple sources of price-cost dispersion, and emphasize that whether high markup firms should be bigger depends on the source of their markups. We complement these studies by analyzing borrowing constraints as another source of markup dispersion. Our analysis suggests that some high markup firms could be too large because their cost advantage arises from looser borrowing constraints (rather than higher productivity).

Another stream of work that we connect to is misallocation due to financial frictions, such as Buera, Kaboski, and Shin (2011), Khan and Thomas (2013), Moll (2014), Midrigan and Xu (2014), David and Venkateswaran (2019), Ottonello and Winberry (2020), and corporate finance studies reviewed by Eisfeldt and Shi (2018). A number of papers in this literature examine borrowing against physical assets. Our work studies borrowing against operating earnings, which is prevalent among U.S. nonfinanial firms (Lian and Ma, 2021; Drechsel, 2023; Adler, 2024). Li (2022) shows that borrowing against earnings generates less misallocation than borrowing against physical assets. We present new evidence on the relationship between markups and corporate borrowing, and highlight that a two-way feedback mechanism between variable markups and earning-based borrowing aligns with the evidence.

Overall, the literature on misallocation in the sense of Hsieh and Klenow (2009) and Restuccia and Rogerson (2008) tends to study the role of a single distortion individually. Our analysis suggests that product market and financial market distortions may offset each other, which provides new perspectives for understanding the productivity loss from different distortions.

2 Motivating Facts

In this section, we present new empirical facts about the relationship between firms' markups and borrowing constraints. In future sections, we show how a two-way feedback between markups and constraint tightness can explain all the empirical facts documented here.

Fact 1a. Less constrained firms have higher markups We start with the firm-level relationship between borrowing constraint tightness and markups in Panel A of Figure 1. This firm-year level binscatter plot uses the log firm-level markup from De Loecker, Eeckhout, and Unger (2020) on the y-axis, and we confirm that the results are similar using a number of other markup estimates in Table 1. The x-axis uses a common proxy for firms' borrowing constraint tightness, namely firms' cash holdings (normalized by book assets), following Gilchrist et al. (2017), for example. This measure reflects the abundance of firms' internal funds and is inversely related to their borrowing constraint tightness. The sample covers US Compustat firms annually, so we can observe their cash holdings, and we use industry (3-digit NAICS code) by year fixed effects to pinpoint the heterogeneity among firms in the same industry at a given point in time.

At the end of this section, we also check that the main empirical results hold using capital wedge in U.S. Census of Manufacturing data (production worker hours relative to capital) as an alternative proxy for financial constraints, which maps to the model in Section 3. The finance literature has constructed other proxies for borrowing constraint tightness (Kaplan and Zingales, 1997; Whited and Wu, 2006), which have been subject to much debate (Farre-Mensa and Ljungqvist, 2016). Furthermore, these proxies contain net income in their construction, which relate to firm profits and can be affected by markups. Accordingly, they can be mechanically correlated with markups measures, so we do not use them.

Figure 1, Panel A, shows a strong positive relationship between firms' cash holdings and markups. We confirm the statistical significance of this relationship in column (1) of Table 1, Panel A. This means that less constrained firms have higher markups. The remaining columns perform robustness checks using several other markup proxies. Column (2) uses the markup from De Loecker, Eeckhout, and Unger (2020) where the production function includes overhead as a factor of production. Column (3) uses the simple accounting markup from De Loecker, Eeckhout, and Unger (2020), which is the ratio of sales to cost of good sold scaled by the industry cost share. Column (4) uses translog markups following De Ridder (2024). We use industry (3-digit NAICS code) by year fixed effects, and double cluster standard errors by both industry and year. Appendix Table IA1 shows that the results are robust to further controls such as firm age.[1]

[1] Compustat does not have a direct measure of firm age. We use the older one between year since incorporation (we collect incorporation year data from Datastream) and year since IPO (IPO year is from Compustat).

One concern is that the revenue-based markup estimates in De Loecker, Eeckhout, and Unger (2020) are not informative of the level of markups (Bond et al., 2021). As De Ridder, Grassi, and Morzenti (2024) point out, though biases in these types of estimates do affect the measurement of the average level of the markup, they measure the dispersion of markups across firms reasonably well, which is our sole focus. Also, the relationship between firms' borrowing constraint tightness and markups does not need to be interpreted causally in either direction to be of interest. In our subsequent analysis, we explain how this relationship arises from a two-way feedback loop between borrowing constraint tightness and markups. Moreover, their univariate relationship (without controls) is of particular interest, as the univariate covariance between markup and constraint tightness directly enters TFP loss (20).

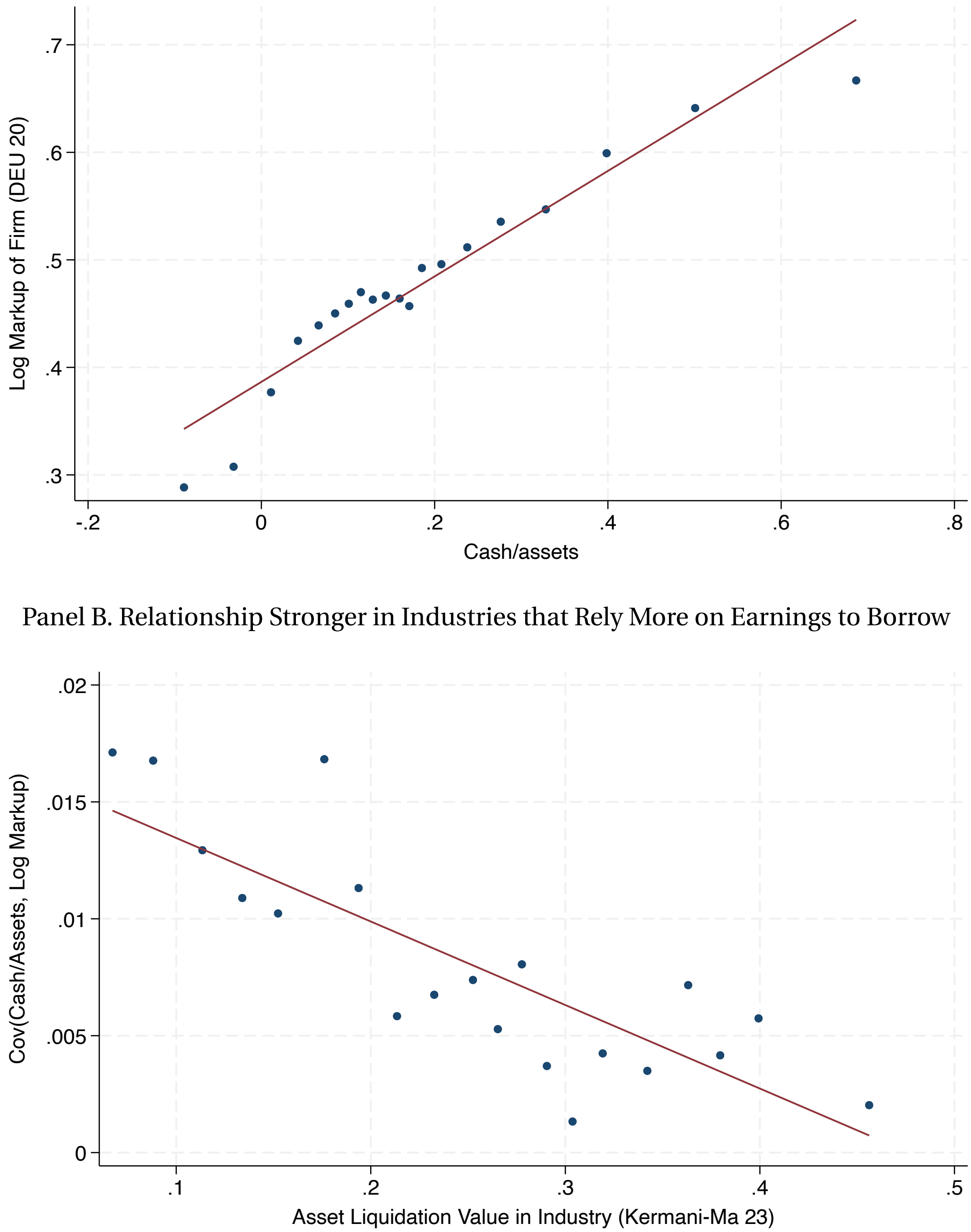
Fact 1b. The relationship between borrowing constraint tightness and markups is stronger in industries that rely more on earnings to borrow Furthermore, we find that the relationship between borrowing constraint tightness and markups is stronger among industries where firms have low liquidation value of pledgeable assets and therefore rely more on earnings to borrow (Lian and Ma, 2021; Kermani and Ma, 2023). In the binscatter plot in Figure 1, Panel B, we study the covariance between firm-level cash holdings and log markups (i.e., the covariance of the two variables on the two axes of Panel A) in each industry-year on the y-axis. We use the industry-level liquidation value of pledgeable assets (normalized by book assets) from Kermani and Ma (2023) on the x-axis, which are driven by physical characteristics of assets used in each industry and shape the extent to which firms need to borrow based on earnings, When the liquidation value of pledgeable assets is high, firms in that industry can directly pledge them for asset-based debt; when the liquidation value of pledgeable assets is low, firms in that industry resort to cash flow-based debt and rely more on earnings to borrow. We verify that firms and industries with low liquidation value have more cash flow-based debt subject to earnings-based borrowing constraints in Appendix Figure C.1. Figure 1, Panel B, shows that the relationship between cash holdings and markups in Panel A is especially strong in low liquidation value industries that rely more on earnings to borrow. We confirm the statistical significance of this relationship in column (1) of Table 1, Panel B. Columns (2) to (4) use other markup measures (for the covariance between firms' cash holdings and log markups on the left hand side) parallel to those in Panel A. We use year fixed effects, and double cluster standard errors by both industry and year.

Fact 2. Markup dispersion is higher in industries that rely more on earnings to borrow Next, we show that markup dispersion is higher in industries with low liquidation value of pledgeable assets that rely more on earnings to borrow. The binscatter plot in Figure 2 shows the variance of firm-level markup within an industry-year on the y-axis, and the industry's liquidation value of pledgeable assets on the x-axis (same x-axis as Figure 1, Panel B). We confirm the statistical significance of this

Figure 1: Financial Constraints and Firm Markups

Panel A. Less Constrained Firms Have Higher Markups

Panel B. Relationship Stronger in Industries that Rely More on Earnings to Borrow



Notes. Panel A shows a binscatter plot of the relationship between firms' cash holdings (normalized by book assets) and log markups, using US Compustat firms annually. The x-axis represents firms' cash holdings, and the y-axis represents the log firm-level markup from De Loecker, Eeckhout, and Unger (2020). Industry (3-digit NAICS code) by year fixed effects are included. Panel B shows the covariance between firm-level cash holdings and log markups in each industry-year on the y-axis, and the industry's liquidation value of fixed assets plus working capital (normalized by book assets) on the x-axis. The liquidation value data are from Kermani and Ma (2023). Year fixed effects are included.

Table 1: Financial Constraint and Firm Level Markups

Panel A. Less Constrained Firms Have Higher Markups

	Firm Log Markup			
	(1)	(2)	(3)	(4)
Cash/Assets	0.45***	0.41***	0.26***	0.34***
	(0.04)	(0.04)	(0.02)	(0.05)
Fixed Effects	Industry × Year			
Observations	129,178	86,598	64,893	99,887
R^2	0.30	0.26	0.20	0.30

Panel B. Relationship Stronger in Industries that Rely More on Earnings to Borrow

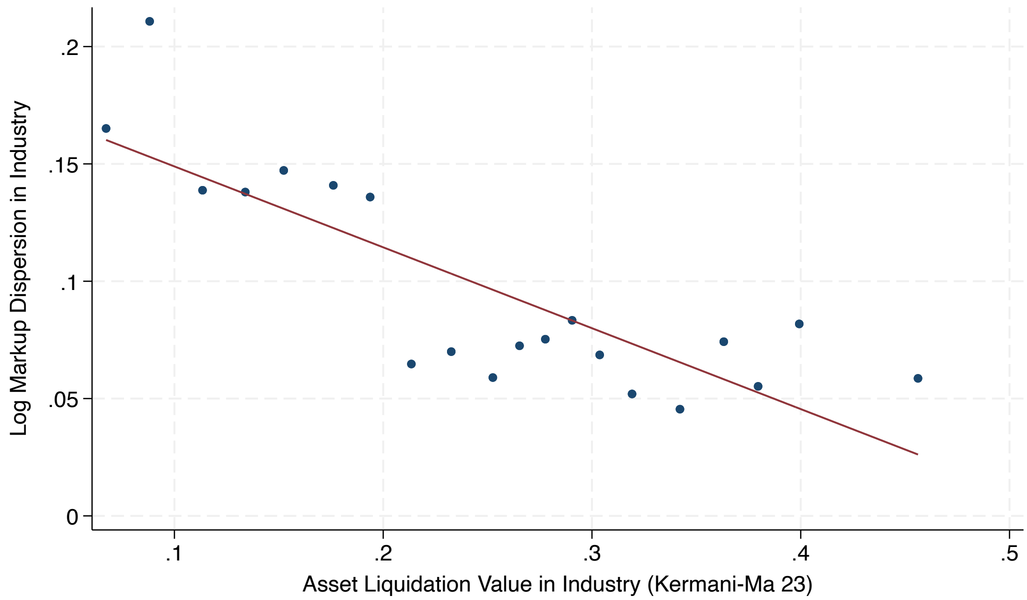
	Industry Cov(Log Markup, Cash/Assets)			
	(1)	(2)	(3)	(4)
Asset Liquidation Value	-0.03***	-0.02*	-0.01*	-0.02***
	(0.01)	(0.01)	(0.01)	(0.01)
Fixed Effects	Year			
Observations	2,453	2,329	2,354	2,358
R^2	0.06	0.03	0.02	0.04

Notes. Panel A shows firm-year level regressions $\text{Log Markup}_{it} = \alpha_{ind(i)t} + \beta \text{Cash/Assets}_{it} + \varepsilon_{it}$, among US Compustat firms. Log Markup_{it} for firm i in year t uses the baseline markup from De Loecker, Eeckhout, and Unger (2020) in column (1), the markup with overhead in production input from De Loecker, Eeckhout, and Unger (2020) in column (2), the markup using accounting ratio (sales/cost of goods) sold from De Loecker, Eeckhout, and Unger (2020) in column (3), and translog markup from De Ridder (2024) in column (4). $\alpha_{ind(i)t}$ is industry (3-digit NAICS code) by year fixed effects. Standard errors are double clustered by industry and year. Panel B shows industry-year level regressions $\text{Cov(Log Markup, Cash/Assets)}_{kt} = \alpha_t + \beta \text{Liqval}_{kt} + \varepsilon_{kt}$, among US Compustat firms. $\text{Cov(Log Markup, Cash/Assets)}_{kt}$ is the covariance of firm-level log markup and firm cash/assets for firms in industry k in year t. The set of markups is the same as the columns in Panel A. The variable $Liqval_{kt}$ is the industry's liquidation value from fixed assets and working capital normalized by the industry's total book assets. The liquidation value is the book value of property, plant, and equipment, inventory, and receivables multiplied by their respective industry-level liquidation recovery rates from Kermani and Ma (2023), normalized by total book assets. α_t represents year fixed effects. Standard errors are double clustered by industry and year.

relationship in column (1) of Table 2. Columns (2) to (4) use other markup measures parallel to those in Table 1.

For industry-level markup dispersion, we supplement markups measures using Compustat data with markups measured using Census of manufacturing in column (5) of Table 2. The benefit of census data is that we can directly observe variable inputs, as well as more comprehensive coverage of a wide range of firms. However, we are limited to Economic Census years (years ending with 2 and 7)

Figure 2: Industry Markup Dispersion



Notes. This figure shows a binscatter plot of the relationship between the variance firm-level log markups in each industry-year on the y-axis, and the industry's liquidation value of fixed assets plus working capital (normalized by book assets) on the x-axis. The liquidation value data are from Kermani and Ma (2023), normalized by total book assets. Year fixed effects are included.

Table 2: Industry Markup Dispersion

	Variance of Firm-Level Log Markup in Industry				
	(1)	(2)	(3)	(4)	(5)
Asset Liquidation Value	-0.29***	-0.26***	-0.18***	-0.12***	-0.06***
	(0.07)	(0.06)	(0.04)	(0.03)	(0.02)
Fixed Effects			Year		
Observations	2,454	2,330	2,354	2,358	1,100
R^2	0.13	0.10	0.09	0.10	0.10

Notes. This table shows industry-year level regressions $\text{Var(Log Markup)}_{kt} = \alpha_t + \beta \text{Liqval}_{kt} + \varepsilon_{kt}$, among US Compustat firms. $\text{Var(Log Markup)}_{kt}$ is the variance of firm-level log markup for firms in industry k in year t. The set of markups is the same as the columns in Table 1. The variable $Liqval_{kt}$ is the industry's liquidation value from fixed assets and working capital normalized by the industry's total book assets. The liquidation value is the book value of property, plant, and equipment, inventory, and receivables multiplied by their respective industry-level liquidation recovery rates from Kermani and Ma (2023). Year fixed effects are included. Standard errors are double clustered by industry and year.

and manufacturing.

Additional robustness checks In our model in Section 3, the tightness of financial constraints will be reflected by the capital wedge, so we also measure the capital wedge in census data (the log of production worker hours relative to capital) and repeat the main analysis above using the capital wedge as a measure of financial constraint tightness. We utilize the Census of Manufacturing data for this check because production hours are not available in Compustat data. High capital wedge indicates tighter financial constraint as constrained firms cannot acquire enough capital (so the results using capital wedge have the opposite sign as those using cash holdings above because higher capital wedge indicates tighter constraints whereas higher cash holdings indicate looser constraints). For Fact 1a, we verify that firms with higher capital wedge have significantly lower markups. When we regress log markups on the capital wedge in the Census data, the regression coefficient is -0.0116** (s.e. 0.01). For Fact 1b, we also find that the above relationship is stronger for industries with low asset liquidation value. When we regress the covariance of log markup and the capital wedge on the industry asset liquidation value, we obtain a regression coefficient of 0.088* (s.e. 0.026), i.e., the negative relation between markup and capital wedge is stronger in industries that rely more on earnings to borrow.

Relevance of earnings-based borrowing constraints In the empirical analyses and in the model, we highlight that borrowing against earnings can shape the relationship between markups and financial constraints. A natural question is whether large public firms in Compustat are financially constrained. Lian and Ma (2021) provide an extensive study of the relevance of earnings-based borrowing constraints among large US nonfinancial firms. They find that almost 70% of large Compustat nonfinancial firms have earnings-based borrowing constraints written explicitly in their debt contracts. In a given year, about 10% of such firms violate the constraints and 25% are within one standard deviation to violation. This importance of earnings-based borrowing constraints is frequently discussed by firms' financial reports. For example, Starwood Hotel's 2011 annual report states: "A substantial decrease in EBITDA (as defined in our credit agreements) ... could make it difficult for us to meet our debt service requirements and restrictive covenants and force us to sell assets and/or modify our operations." GNC's 2012 annual report states: "These restrictions may prevent us from taking actions that we believe would be in the best interest of our business and may make it difficult for us to successfully execute our business strategy or effectively compete with companies that are not similarly restricted."

3 Model

We set up an industry equilibrium model to explain the above firm- and industry-level evidence and study the allocative implications. The industry's final good is produced by competitive firms that aggregate differentiated intermediate input through a Kimball aggregator. Monopolistically competitive intermediate input producers use variable inputs and capital and are subject to asset- and earnings-based borrowing constraints.

Industry final good producers. The industry's final good is produced by perfectly competitive producers that aggregate a continuum of intermediate inputs $y(i)$ into the final output Y via a Kimball aggregator $\Upsilon(\cdot)$:

$$\max_{\{y(i)\}_i} PY - \int p(i)\, y(i)\, di \quad \text{s.t.} \quad \int_0^1 \Upsilon\left(\frac{y(i)}{Y}\right) di \;=\; 1, \tag{2}$$

where $\Upsilon(1) = 0$, $\Gamma'(\cdot) > 0$, $\Gamma''(\cdot) < 0$, and P captures the price of industry final good. Optimality by the final good producers implies that the market share of each intermediate good i, $\tilde{y}(i) = y(i)/Y$, is related to its price $p(\tilde{y}(i))$ by

$$p(\tilde{y}(i)) = \frac{\lambda}{Y}\Upsilon'(\tilde{y}(i)), \tag{3}$$

where λ captures the Lagrange multiplier on the constraint in (2). Relationship (3) can be interpreted as the demand curve faced by the intermediate good producer i.

When illustrating key mechanisms of the model and how it quantitatively explains the empirical evidence in the previous section, we use a popular specification of $\Upsilon(\cdot)$ based on Dotsey and King (2005) for the Kimball aggregator,

$$\Upsilon(\tilde{y}(i)) = \frac{\sigma}{\sigma(1-\eta)-1}\left(\left[(1-\eta)\,\tilde{y}(i)+\eta\right]^{1-\frac{1}{\sigma(1-\eta)}} - 1\right) + 1, \tag{4}$$

which implies the demand elasticity

$$\sigma(\tilde{y}(i)) \equiv -\frac{\partial \ln \tilde{y}(i)}{\partial \ln p(i)} = \sigma\left(1 + \eta\frac{1-\tilde{y}(i)}{\tilde{y}(i)}\right), \tag{5}$$

where $\sigma > 0$ governs the average demand elasticity and $\eta > 0$ governs the "superelasticity," i.e., how the demand elasticity of good i decreases with its market share $\tilde{y}(i)$. When $\eta = 0$, the demand system corresponds to the standard Constant Elasticity of Substitution (CES) case. Note that this specification is closely related to the specification in Klenow and Willis (2016) and, in fact, shares the exact same first-order approximation with Klenow and Willis (2016). Due to differences in higher-order terms, Dotsey and King (2005)'s specification generates greater markup dispersion across varieties and allows the model to better match certain empirical features of the data, facilitating our quantitative analysis. We henceforth use Dotsey and King (2005)'s specification throughout the paper.

Industry intermediate goods producers. Each industry is populated by a unit measure of intermediate goods producers $i \in [0,1]$ operating under monopolistic competition. We write the intermediate goods producer's optimization problem recursively and omit the index i. The relevant state variable for is its pre-production net worth a and productivity z. Each intermediate good producer produces a differentiated good using a Cobb-Douglas technology in capital k and variable inputs v, which include labor and materials:[2]

$$y = zk^{\alpha} v^{1-\alpha}, \tag{6}$$

where $\alpha \in (0,1)$. Each producer faces idiosyncratic productivity shocks. Its productivity, denoted by z, evolves exogenously according to an AR(1) process:

$$\log z' = \rho \log z + \epsilon_z, \tag{7}$$

with $\rho \in [0,1)$ and $\epsilon_z \sim \mathcal{N}\left(0, \sigma_e^2\right)$ independent across firms.

Following Moll (2014), the producer can borrow $b \equiv k - a$ to acquire capital beyond its net worth for production, and the price of capital goods is normalized to one. Specifically, the producer chooses its capital k and variable inputs v to maximize its post-production net worth

$$n(a,z) \equiv \max_{k,v} \quad p\left(y/Y\right) y - p_v v - c - (k-a)(1+r) + k(1-\delta), \tag{8}$$

subject to the demand of its good in (3), the production technology (6), and asset- and earnings-based borrowing constraints:

$$b \equiv k - a \leq \max\left\{\phi^A k, \phi^E \max\left\{0, p\left(y/Y\right) \cdot y - p_v v - c\right\}\right\}, \tag{9}$$

where p_v is the price of the variable input, c is the fixed operating cost, r is the interest rate, δ is the depreciation rate of capital, ϕ^A is the capacity of asset-based borrowing, ϕ^E is the capacity of earnings-based borrowing constraints, and $p\left(y/Y\right) \cdot y - p_v v - c$ is the producer's earnings.

The producer's borrowing capacity in (9) is determined by the maximum implied by asset-based borrowing, as in Moll (2014), and earnings-based borrowing, Lian and Ma (2021). This specification of borrowing capacity in (9) follows the contractual evidence in Lian and Ma (2021).[3]

After the production, the producer decides how much dividend d to pay and how much to save a'. Specifically, the producer chooses a' to maximize its value

$$V(a,z) = \max_{a' \geq 0} \quad d + \beta(1-\pi_d)\mathbb{E}_z\left[V\left(a', z'\right) \mid z\right] + \beta\pi_d a' \tag{10}$$

subject to budget constraint

$$\psi_d d^2 \mathbb{I}(d<0) + d = n(a,z) - a', \tag{11}$$

where $n(a,z)$ is its post-production net worth in (8), π_d is the exogenous exit rate, and ψ_d is costly equity issuance (reflected as a negative dividend payment). Exiting firms are replaced by entrants

[2] Our model is equivalent to a model in which labor and materials inputs, l and m, are explicitly modeled. In that model, the firm's production is given by $y = zk^{\alpha}\left(l^{\gamma} m^{1-\gamma}\right)^{1-\alpha}$ and the price of labor and material inputs is given by p_l and p_m. It is equivalent to our current model with $v = l^{\gamma} m^{1-\gamma}$ and $p_v = \frac{p_l^{\gamma} p_m^{1-\gamma}}{(\gamma)^{\gamma}(1-\gamma)^{1-\gamma}}$.

[3] If a firm borrows earnings-based debt, the earnings-based constraints always restrict a firm's total debt (including both earnings-based and asset-based debt) as a function of its total earnings. This explains why the borrowing capacity in (9) is not given by a sum of the firm's asset-based and earnings-based borrowing capacities.

with an initial net worth a^{ent} and the average productivity $\mathbb{E}[z]$.

The model closely follows Moll (2014), particularly its discrete time version detailed in Appendix G, with only two key differences. First, the producer's borrowing capacity in (9) is determined by the maximum implied by both asset-based and earnings-based borrowing, rather than by asset-based borrowing alone. Second, instead of modeling the producer as an entrepreneur with log utility, the producer's value is linear in its dividend payment and subject to costly equity issuance and an exogenous probability of exit. The latter assumptions closely follow Ottonello and Winberry (2020, 2024) and better capture large firms subject to earnings-based constraint, as studied in Lian and Ma (2021) and the previous Section.

Given the exogenous interest rate r, the price of variable inputs p_v, and the demand for the industry final good D, an industry equilibrium consists of a collection of intermediate goods producers' decision rules, prices, and quantities $\{a'(a,z), b(a,z), d(a,z), y(a,z), v(a,z), k(a,z), p(a,z)\}$, the industry final good price and quantity P and Y, and the cumulative distribution function of intermediate goods producers' states $G(a,z)$ such that: (i) Intermediate goods producers and final good producers optimize; (ii) The industry final good price and quantity satisfy industry demand $PY = D$, which follows naturally if the technology aggregates goods produced by different industries in a Cobb-Douglas production form, as in Hsieh and Klenow (2009); iii) The distribution of intermediate goods producers' states $G(a,z)$ is stationary given its saving policy $a'(a,z)$ and the process of z in (7).

4 Borrowing Constraints and Markup Dispersion

In this section, we study how borrowing constraints and markups are related in our model through the two-way feedback between markups and constraint tightness. We then show how this two-way feedback helps explain the empirical evidence in Section 2.

Markup dispersion. In our model, markup dispersion depends on both productivity dispersion, as in Edmond, Midrigan, and Xu (2023) and Baqaee, Farhi, and Sangani (2024a,b), and the dispersion of borrowing constraint tightness. To see this, first note that each intermediate goods producer's optimal choice of variable inputs v imply that

$$p(\tilde{y}) = \frac{\sigma(\tilde{y})}{\sigma(\tilde{y})-1}\frac{p_v}{(1-\alpha)z(k/v)^{\alpha}} = \mu(\tilde{y})MC \tag{12}$$

where $MC = \frac{p_v}{(1-\alpha)z(k/v)^{\alpha}}$ is the marginal cost of producing one additional unit of the intermediate good using additional variable inputs, and $\mu(\tilde{y}) = \frac{\sigma(\tilde{y})}{\sigma(\tilde{y})-1}$ is the markup implied by the demand elasticity in (5). Note that (12) holds regardless of whether the borrowing constraint (9) binds, as it restricts capital choice but not the choice of variable inputs, following Moll (2014).

We then write the intermediate goods producer's marginal cost as a function of its productivity z and the capital wedge $1+\tau_k = \frac{\alpha}{1-\alpha}\frac{p_v}{r+\delta}\frac{v}{k}$, defined as in Hsieh and Klenow (2009):

$$MC = \frac{(r+\delta)^{\alpha}\, p_v^{1-\alpha}\,(1+\tau_k)^{\alpha}}{\alpha^{\alpha}\,(1-\alpha)^{1-\alpha}\, z}, \tag{13}$$

which decreases with the producer's productivity z and increases with its capital wedge τ_k. The capital wedge τ_k is high if the producer uses "too little" capital relative to variable inputs (so $\frac{v}{k}$ is high), driving the marginal product of capital too high relative to the cost of capital. In our model, a firm has a high capital wedge because of binding financial constraints, so the capital wedge can be viewed as a measure of borrowing constraint tightness.

Based on the demand curve in (3), the demand elasticity in (5), and (12), we can link the intermediate goods producer's market share, and hence markup, to its marginal cost. We can then express markup dispersion as the dispersion of marginal cost, and further as productivity dispersion and the dispersion of borrowing constraint tightness based on (13).

The expressions take a particularly simple form if we take an approximation. In particular, we approximate around the point where all intermediate goods producers operate efficiently (so their capital wedges $\tau_k = 0$ and markups $\mu = 1$) and have the same TFP z. In fact, this approximation generalizes the Hsieh and Klenow (2009) approach based on log-normal distributions and arrives at the exact same formula for TFP loss in (5). Under this approximation, we can express markup dispersion as follows.

Proposition 1. *Consider two intermediate goods producers $i, j \in [0,1]$. Under the specification of $\Upsilon(\cdot)$ based on Dotsey and King (2005) in (4), to first order, the differences in their markups are given by*

$$\underbrace{\log\mu(i) - \log\mu(j)}_{\text{markup dispersion}} \approx \frac{\eta}{(\sigma-1)}\underbrace{\left(\log\tilde{y}(i) - \log\tilde{y}(j)\right)}_{\text{market share dispersion}} \approx -\frac{\sigma\eta}{\sigma-1+\sigma\eta}\underbrace{\left(\log MC(i) - \log MC(j)\right)}_{\text{marginal cost dispersion}} \tag{14}$$

$$\approx \frac{\sigma\eta}{\sigma-1+\sigma\eta}\underbrace{\left(\log z(i) - \log z(j)\right)}_{\text{productivity dispersion}} - \frac{\alpha\sigma\eta}{\sigma-1+\sigma\eta}\underbrace{\left(\log(1+\tau_k(i)) - \log\left(1+\tau_k(j)\right)\right)}_{\text{dispersion in constraint tightness}}. \tag{15}$$

(14) captures the standard properties of the Kimball aggregator. Intermediate goods producers with lower marginal costs, and hence higher market shares, face less elastic demand (e.g., in (5)) and have higher markups. (15) shows that markup dispersion in our model depends on both productivity dispersion and the dispersion of borrowing constraint tightness. As standard in the literature (e.g., Edmond, Midrigan, and Xu, 2023; Baqaee, Farhi, and Sangani, 2024a,b), higher markups can arise because higher-productivity firms have lower marginal costs and larger market shares. Unique to our model, higher markups can also arise because firms with looser borrowing constraints (i.e., lower capital wedges τ_k) have lower marginal costs and higher market shares.

Two-way feedback between markups and constraint tightness. Our model embeds a two-way feedback between markups and constraint tightness, which helps explain the empirical facts in Section 2. The first step of this feedback applies in general to both asset-based and earnings-based constraints. As discussed earlier, firms with looser borrowing constraints have lower marginal costs and, hence, higher market shares and markups. The second step of this feedback operates only under earnings-based borrowing constraints. In that case, firms with higher markups have higher earnings, which loosen their earnings-based borrowing constraints. This step is absent under asset-based borrowing constraints. Both steps of the two-way feedback explain Fact 1a in Section 2 that less constrained firms have higher markups. The observation that the second step of the two-way feedback operates only under earnings-based borrowing constraints explains Fact 1b in Section 2 that the relationship between financial constraints and markups is stronger in industries that rely more on earnings to borrow. As further explained shortly, this observation also explains Fact 2 in Section 2 that markup dispersion is higher in industries that rely more on earnings to borrow.

We now provide simple illustrations of this two-way feedback.[4] Figure 3 plots the markups (x-axis) and constraint tightness measured in capital wedges (y-axis) of firms that share the same level of productivity, z, but differ in their net worth, a. We observe that firms with higher a—hence, lower capital wedges and looser borrowing constraints—have lower markups, as described by the first step of the two-way feedback.

Figure 4 plots the markups (x-axis) and constraint tightness measured in capital wedges (y-axis) of firms that share the same level of net worth, a, but differ in their productivity, z. Panel A studies the case in which earnings-based constraints are binding. We observe that firms with higher z—hence, higher markups—have lower capital wedges and looser borrowing constraints because they have higher earnings to borrow against. Panel B instead plots the case without earning-based constraints ($\phi^E = 0$). We observe that firms with higher z—hence, higher markups—have higher capital wedges and tighter borrowing constraints because more productive firms seek to borrow more but have the same borrowing capacity as other firms with the same a but lower z under asset-based constraints.

Quantitative evaluation.

We calibrate the model in two steps. First, we exogenously assign a subset of parameters. Second, we choose the remaining parameters in order to match moments in the data. Table 3 lists the parameters that we assign. The model period is one year, so we set discount rate $\beta = 0.97$. We set ϕ^A, the

[4]We consider the following parametrization for the illustration: $\phi^E = 3.5$, $\phi^A = 0.25$, $\delta = 0.1$, $r = 0.02$, $p_v = 1$, $\alpha = 0.15$, $\sigma = 4.8$, $\eta = 0.99$, and a specification of the Kimball aggregator based on Dotsey and King (2005) in (4). The distribution of intermediate goods producers' states, $G(a, z)$, is calculated to match the empirical distribution of lagged book equity and sales share in an industry with a median level of liquidation value of pledgeable assets, ϕ^A, Food Manufacturing with the NAICS code 311.

Figure 3: Markups and Capital Wedges of Firms with Different Net Worth a but the Same Productivity z

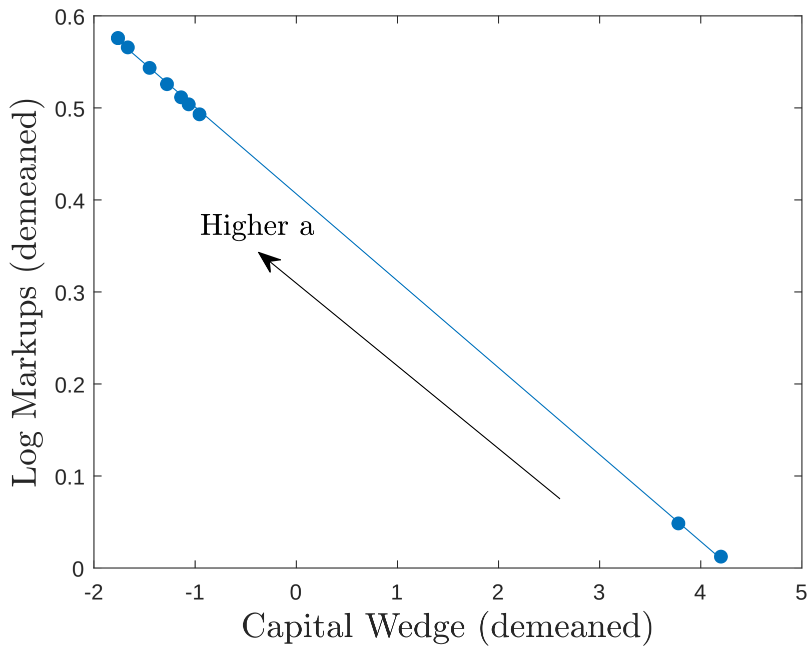
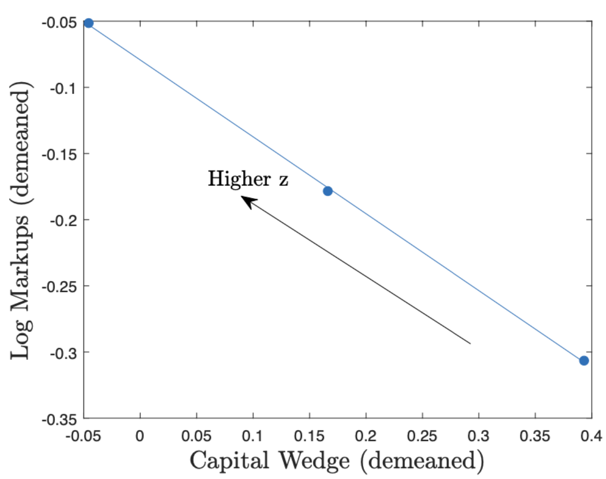
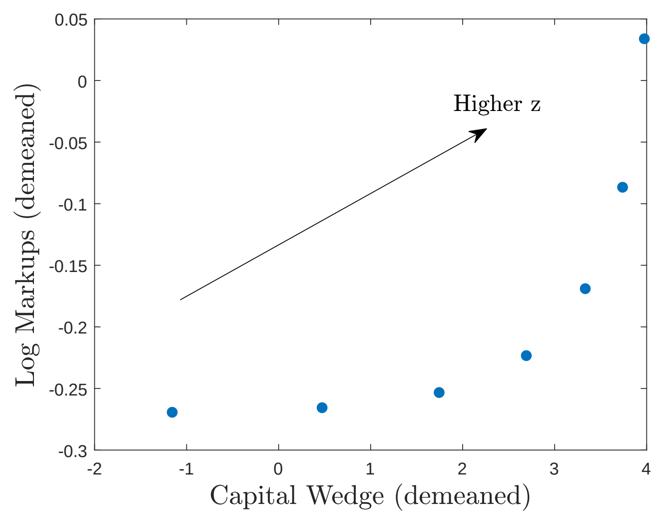


Figure 4: Markups and Capital Wedges of Firms with the Same Net Worth a but Different Productivity z



Panel A. Binding Earnings-based Constraints



Panel B. No Earnings-based Constraints

capacity of asset-based borrowing, to 0.2, representing the aggregate liquidation value of fixed assets and working capital for Compustat manufacturing firms, normalized by their aggregate book assets. We set ϕ^E, the capacity of earnings-based borrowing, to 3.5, following the contractual evidence in Lian and Ma (2021). Capital depreciates at rate $\delta = 0.1$ annually, following David and Venkateswaran (2019). The interest rate is $r = 2\%$, the historical average of the Fed Cleveland 10-Year Real Interest Rate. The capital share $\alpha = 0.15$, calculated from the capital cost share of the BLS-BEA KLEMS manufacturing data. The persistence of idiosyncratic productivity is $\rho = 0.93$, again from David and Venkateswaran (2019). Its average $\mathbb{E}[z]$ is normalized to be 1. The annual exogenous exit rate is 10%, calculated based on the Census BDS data.

We then choose the parameters listed in Table 4 to match the empirical moments reported in Table 5. We target the ratio of aggregate earnings to aggregate debt (0.38), calculated based on Quarterly Financial Report (QFR) manufacturing firms. It is informative about the fixed operating cost c. We target the median firm's debt-to-asset ratio (0.16), calculated based on Compustat manufacturing firms; and the elasticity of book equity with sales share, corresponding to the firm-level regression coefficient of $\ln a$ on $\ln py$ in our model, calculated based on Compustat manufacturing firms. Those two moments are informative about the distribution of intermediate goods producers' states $G(a,z)$ and hence the standard deviation of the idiosyncratic shock σ_e and the equity issuance cost ψ_d. We target the aggregate markup (1.45) from De Ridder, Grassi, and Morzenti (2024) and the elasticity of markups with sales share, corresponding to the firm-level regression coefficient of $\ln\mu$ on $\ln py$ in our model (0.16). Those two moments are informative about the two parameters governing the elasticity of substitution and the superelasticity, σ and η. Table 5 shows that our model matches the targeted moments reasonably well. The calibrated parameters in Table 4 are broadly comparable to existing estimates in the literature.

Our model can explain Fact 1a in Section 2 that less constrained firms have higher markups because of the two-way feedback between markups and constraint tightness. In the benchmark calibration with $\phi^A = 0.2$, we find that firms' markups and their constraint tightness (measured using capital wedge) are negatively correlated, specifically, $Cov\left(\log\mu, \log(1+\tau_k)\right) \approx -0.09$. This is similar to the covariance between log markup and log capital wedge (measured using ln (production worker hours per unit of capital) in the Census), which is not explicitly targeted in our calibration.

Our model can also explain Fact 1b in Section 2 that the negative relationship between financial constraints and markups is stronger in industries that rely more on earnings to borrow. We solve the industry equilibrium for different values of the capacity for asset-based borrowing, ϕ^A, while keeping other parameters fixed. This approach is motivated by the evidence in Lian and Ma (2021) and Kermani and Ma (2023) that the capacity for asset-based borrowing, ϕ^A, differs significantly across industries due to cross-industry variations in asset specificity, whereas the capacity for earnings-based borrowing, ϕ^E, remains roughly constant across industries.

The blue line in Figure 5 plots the covariance between log markup and constraint tightness (measured in log capital wedge), $Cov\left(\log\mu, \log(1+\tau_k)\right)$, in the model under different values of ϕ^A (y-axis) against ϕ^A (x-axis). In the model, we find that firms' markups and their constraint tightness are more negatively correlated in industries with low ϕ^A, which rely more on earnings to borrow, as the second step of the two-way feedback operates in those industries. The slope of Figure 5, i.e., the industry-level

Table 3: Assigned Parameters

Parameter	Description	Source	Value
ϕ^E	EBC constraint	Lian and Ma (2021)	3.5
ϕ^A	ABL constraint	Compustat manufacturing	0.2
δ	depreciation rate	David and Venkateswaran (2019)	0.1
r	interest rate	Fed Cleveland 10-Year Real Interest Rate	2%
α	capital share	BLS-BEA KLEMS manufacturing	0.15
ρ	persistence of idiosyncratic productivity	David and Venkateswaran (2019)	0.93
x	exogenous exit rate	Firm exit rate Census BDS	10%
β	discount factor	standard value	0.97
$\mathbb{E}[z]$	average productivity	normalize	1

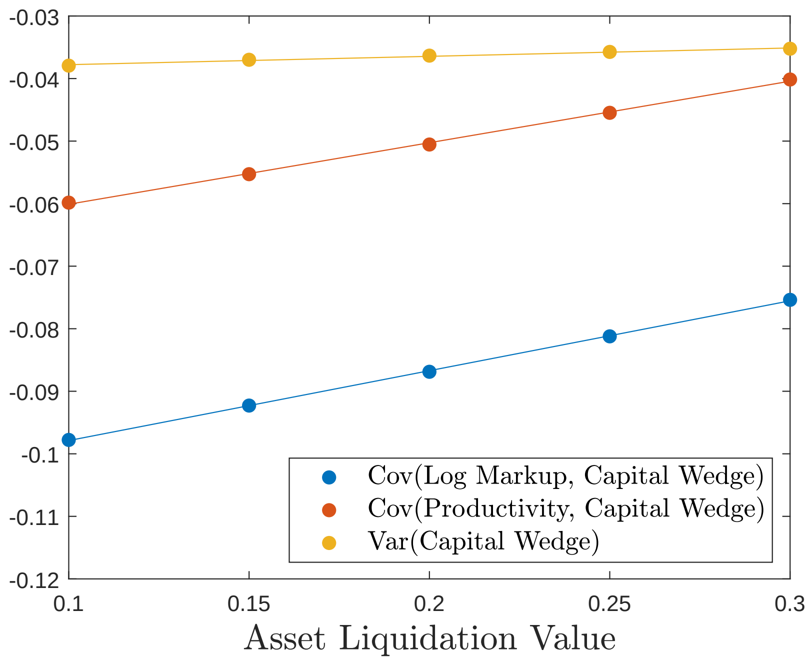
Table 4: Calibrated Parameter

Parameter	Description	Value
c	fixed operating cost	0.37
σ	elasticity of substitution	6
η	governs superelasticity	0.4
σ_e	sd of idiosyncratic productivity shock	0.29
ϕ^d	equity issuance cost	0.01

Table 5: Calibration Target

Calibration Target	Source	Data	Model
Aggregate Earnings / Debt	QFR manufacturing	0.38	0.38
Median debt/asset	Compustat manufacturing	0.16	0.20
Aggregate markup	De Ridder, Grassi, and Morzenti (2024)	1.45	1.49
Elasticity of markups wrt sales share	Edmond, Midrigan, and Xu (2023)	0.16	0.17
Elasticity of book equity wrt sales share	Compustat manufacturing	0.88	0.90

Figure 5: Less constrained firms have higher markups; esp. in industries relying on earnings to borrow



regression coefficient of $Cov\left(\log\mu, \log(1+\tau_k)\right)$ on ϕ^A is 0.125. This is similar to its counterpart 0.0878 in the Census, where we measure $\log(1+\tau_k)$ using ln (production worker hours per unit of capital) in the Census and is not explicitly targeted in our calibration.

To further under how the the two-way feedback helps us explain Fact 1b, we further decompose $Cov\left(\log\mu, \log(1+\tau_k)\right)$.

Proposition 2. *Under the specification of* $\Upsilon(\cdot)$ *based on Dotsey and King (2005) in (4), to second order, the covariance between log markup and constraint tightness (measured in log capital wedge) is given by*

$$Cov\left(\log\mu, \log(1+\tau_k)\right) \approx \frac{\sigma\eta}{\sigma-1+\sigma\eta} Cov\left(\log z, \log(1+\tau_k)\right) - \frac{\alpha\sigma\eta}{\sigma-1+\sigma\eta} Var\left(\log(1+\tau_k)\right) \tag{16}$$

This decomposition follows directly from (15) in Proposition 1. The first term in (16), captured by the red line in Figure 5, arises because high productivity firms have high markups. As a result, a higher $Cov\left(\log z, \log(1+\tau_k)\right)$ leads to a higher $Cov\left(\log\mu, \log(1+\tau_k)\right)$. In industries with lower ϕ^A and greater reliance on earnings for borrowing, the second step of the two-way feedback is operative and this term is negative. As illustrated in Panel A of Figure 3, firms with higher productivity have higher earnings and face looser borrowing constraints. In industries with higher ϕ^A and greater reliance on assets for borrowing, the second step of the two-way feedback is not operative and this term is positive. As illustrated in Panel B of Figure 3, firms with higher productivity seek to borrow more and face tighter borrowing constraints. Together, $Cov\left(\log z, \log(1+\tau_k)\right)$ increases with ϕ^A and helps explain Fact 1b that $Cov\left(\log\mu, \log(1+\tau_k)\right)$ is more negative in industries with lower ϕ^A, The second term in (16), captured by the yellow line in Figure 5, arises because firms with looser borrowing constraints also have high markups. As a result, a lower $Var(\log(1+\tau_k))$ leads to a higher $Cov(\log\mu, \log(1+\tau_k))$. In industries with lower ϕ^A, firms are more constrained overall and have a higher $Var(\log(1+\tau_k))$. As a result, this term also helps explain why $Cov(\log\mu, \log(1+\tau_k))$ is more negative in industries with lower ϕ^A.

Our model can also explain Fact 2 in Section 2 that markup dispersion is higher in industries that rely more on earnings to borrow. To build intuition, we link the variance of log markups in an industry with the covariance between log markup and constraint tightness (measured in log capital wedge) studied in Proposition 2.

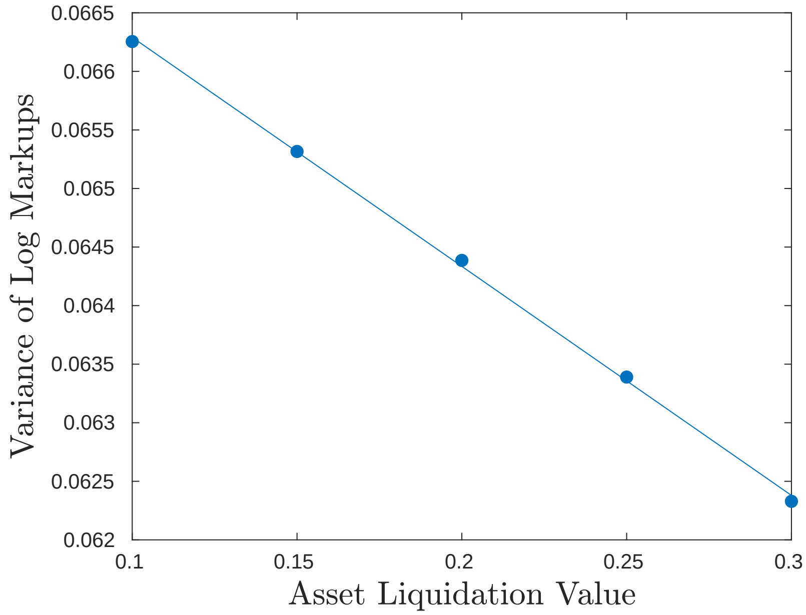
Proposition 3. *Under the specification of $\Upsilon(\cdot)$ based on based on Dotsey and King (2005) in (4), to second order, the variance of log markups is given by*

$$
\begin{aligned}
Var(\log\mu) &\approx \left[\frac{\sigma\eta}{\sigma-1+\sigma\eta}\right]^2 \left[Var(\log z) + \alpha^2 Var(\log(1+\tau_k)) - 2\alpha Cov(\log z, \log(1+\tau_k))\right] \\
&\approx \left[\frac{\sigma\eta}{\sigma-1+\sigma\eta}\right]^2 Var(\log z) - \left[\frac{\alpha\sigma\eta}{\sigma-1+\sigma\eta}\right]^2 Var(\log(1+\tau_k)) \\
&\quad \underbrace{-\frac{2\alpha\sigma\eta}{\sigma-1+\sigma\eta} Cov(\log\mu, \log(1+\tau_k))}_{>0 \text{ if higher markup firms less constrained}} \qquad (17)
\end{aligned}
$$

The first equation again directly uses how intermediate producers' productivity and constraint tightness impact their markups in (15). The second equation use (16) to replace $Cov(\log z, \log(1+\tau_k))$ and show that $Var(\log\mu)$ decreases with $Cov(\log\mu, \log(1+\tau_k))$, the object of interest in Fact 1b and Figure 5. Intuitively, if larger firms with higher markups are also less constrained, they become even larger through borrowing and hence achieve even higher markups. This leads to greater markup dispersion. Because $Cov(\log\mu, \log(1+\tau_k))$ is more negative in industries with lower ϕ^A and that rely more on earnings to borrow, as shown in Figure 5, we know that markup dispersion is also higher in those industries.

In Figure 6, we plot the variance of log markup $Var(\log\mu)$ in the model under different values of ϕ^A (y-axis) against ϕ^A (x-axis). We find that markup dispersion is higher in industries with low ϕ^A, which rely more on earnings to borrow. The slope of Figure 6, i.e., the industry-level regression coefficient of $Var(\log\mu)$ on ϕ^A is -0.023. This is in the ballpark but smaller than the magnitude of its counterpart, -0.062, in the Census, as shown in column (5) of Table 2. The smaller magnitude of the slope in the model could arise because of the well-known fact that model markups are often much smaller than their empirical counterpart due to measurement error (Bils, Klenow, and Ruane, 2021).

Figure 6: Higher Markup Dispersion in Industries that Rely More on Earnings to Borrow



5 Allocative Efficiency

We now study our model's implications for allocative efficiency. The planner allocates the industry's total variable inputs and capital, V and K, across intermediate goods producers to maximize industry final output, subject to the same technology of the final and intermediate goods producers, as given in (2) and (6), but not the borrowing constraints.

$$\max_{\{v^P(i),k^P(i)\}} Y^P \quad s.t. \quad \int_0^1 \Upsilon\left(\frac{y^P(i)}{Y^P}\right) di = 1, \quad y^P(i) = z(i)\left(k^P(i)\right)^{\alpha}\left(v^P(i)\right)^{1-\alpha}, \tag{18}$$

$$V = \int_i v^P(i)\, di, \quad K = \int_i k^P(i)\, di.$$

As we focus on allocative efficiency, the planner takes the industry's total variable inputs and capital as given, $V = \int_i v(i)\, di$ and $K = \int_i k(i)\, di$, from the industry equilibrium studied in Section 3 and 4.[5]

Consider two intermediate goods producers $i, j \in [0,1]$. Under the specification of $\Upsilon(\cdot)$ based on Dotsey and King (2005) in (4), the differences in their output shares in the planner's problem, $\tilde{y}^P(i) = y^P(i)/Y^P$ and $\tilde{y}^P(j) = y^P(j)/Y^P$, are then linked by their relative productivity,

$$\log\left(\tilde{y}^P(i) + \frac{\eta}{1-\eta}\right) - \log\left(\tilde{y}^P(j) + \frac{\eta}{1-\eta}\right) = \sigma(1-\eta)\left(\log(z(i)) - \log(z(j))\right).$$

This differs from the equilibrium differences in their relative market shares in two ways, as the latter

[5] Inefficiency in the aggregate level of variable inputs and capital is a separate issue outside the scope of the paper.

also depend on their relative markups and relative constraint tightness.

$$\log\left(\tilde{y}(i)+\frac{\eta}{1-\eta}\right)-\log\left(\tilde{y}(j)+\frac{\eta}{1-\eta}\right)=\sigma(1-\eta)(\log z(i)-\log z(j))$$
$$-(\log\mu(i)-\log\mu(j))-\alpha(\log(1+\tau_k(i))-\log(1+\tau_k(j))).$$

The following proposition summarizes the difference between the intermediate goods producer's equilibrium market share and its output share in the planner's problem.

Proposition 4. *Consider the intermediate goods producer $i\in[0,1]$. Under the specification of $\Upsilon(\cdot)$ based on Dotsey and King (2005) in (4), to first order, the difference between its equilibrium market share and market share in the planner's problem is given by*

$$\log\tilde{y}(i)-\log\tilde{y}^P(i) = -\sigma\underbrace{\left(\log\mu(i)-\int\log\mu(j)dj\right)}_{\textit{markup relative to avg}}-\sigma\alpha\underbrace{\left(\log(1+\tau_k(i))-\int\log(1+\tau_k(j))dj\right)}_{\textit{constraint tightness relative to avg}} \quad (19)$$

As standard in the literature (e.g., Edmond, Midrigan, and Xu, 2023; Baqaee, Farhi, and Sangani, 2024b,a), firms with above-average markups are too small ($\tilde{y}(i)<\tilde{y}^P(i)$) because these more productive firms charge higher markups. Unique to our model, firms with above-average capital wedge are also too small because of the borrowing constraints. As high-markup firms face looser borrowing constraints (Fact 1a), they can become too large relative to the planner when the second channel dominates.

We can then calculate the equilibrium TFP relative to the planner's problem. For this purpose, we define the (gross) equilibrium TFP and its counterpart in the planner's problem as

$$Z=\frac{Y}{K^\alpha V^{1-\alpha}} \quad \text{and} \quad Z^P=\frac{Y^P}{K^\alpha V^{1-\alpha}}.$$

The TFP loss $\log Z^P-\log Z$ in our model captures the misallocation of variable inputs and capital across intermediate goods producers due to markups and financial constraints. We now express this TFP loss as a function of the covariance between the log markup and constraint tightness (measured in the log capital wedge) and their variances, the key objects we studied in Section 4.

Proposition 5. *Under the specification of $\Upsilon(\cdot)$ based on Dotsey and King (2005) in (4), to second order, the equilibrium TFP loss given by*

$$\underbrace{\log Z^P-\log Z}_{\textit{TFP loss}} \approx \underbrace{\frac{\sigma}{2}Var(\log\mu)}_{\textit{markup dispersion}} + \underbrace{\frac{\sigma\alpha^2+\alpha(1-\alpha)}{2}Var(\log(1+\tau_k))}_{\textit{dispersion in constraint tightness}} \quad (20)$$
$$+ \underbrace{\sigma\alpha Cov(\log\mu,\log(1+\tau_k))}_{<0\ \textit{if higher markup firms less constrained}}$$

The TFP loss formula (20) is essentially the same formula for TFP loss in Hsieh and Klenow (2009), re-expressed for our purposes. Compared to Hsieh and Klenow (2009), we extend in two dimensions.

First, we generalize the formula to cases with Kimball demand, beyond the Cobb-Douglas demand case in the original derivation. Second, we consider a second-order approximation, so we do not need the underlying variables to be log-normally distributed.

(20) illustrates how introducing borrowing constraints mitigates the TFP loss from markup dispersion. Without borrowing constraints, markup dispersion leads to TFP loss as high-markup firms are too small, as summarized by the first term in (19). With borrowing constraints, the last term in the TFP formula (20) mitigates the TFP loss from markup dispersion because high-markup firms are less constrained (Fact 1a) as a result of the two-way feedback, and this term is negative. Intuitively, as looser borrowing constraints help originally too-small high-markup firms grow, they can mitigate the TFP loss from markup dispersion.

In the benchmark calibration with the capacity for asset-based borrowing $\phi^A = 0.2$, we can decompose the TFP loss in (20) as:

$$\underbrace{17\%}_{\text{TFP loss}} \approx \underbrace{19\%}_{\text{markup dispersion}} + \underbrace{6\%}_{\text{dispersion in constraint tightness}} - \underbrace{8\%}_{\text{higher markup firms less constrained}} .$$

In other words, the fact that higher-markup firms are less constrained mitigates the TFP loss from markup dispersion by $8\%/19\% \approx 42\%$.

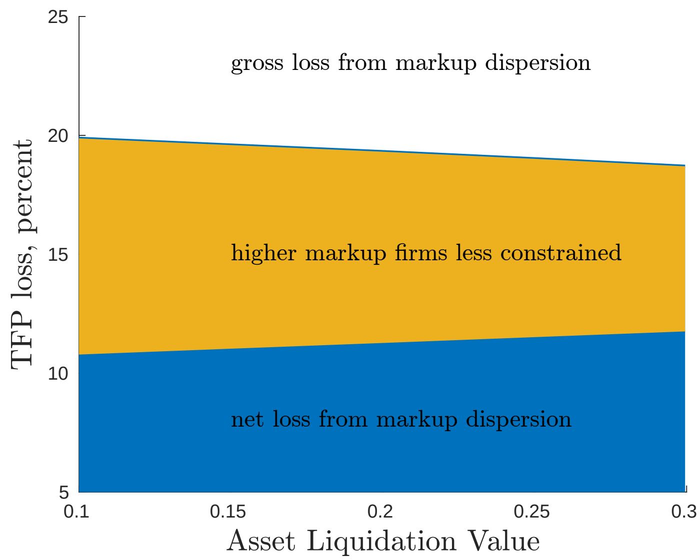
Furthermore, the cross-industry differences documented in Fact 1b and Fact 2 and studied in Figures 5 - 6 also matter for how much the borrowing constraint channel mitigates the TFP loss from markup dispersion. In industries with low ϕ^A, which rely more on earnings to borrow, markup dispersion $Var\left(\log \mu\right)$ is higher, but markups are also more negatively correlated with constraint tightness ($Cov\left(\log \mu, \log\left(1+\tau_k\right)\right)$ is more negative). As a result, the net TFP loss from markup dispersion, defined as the sum of the first and third terms in (20), $\frac{\sigma}{2} Var\left(\log \mu\right) + \sigma\alpha Cov\left(\log \mu, \log\left(1+\tau_k\right)\right)$, can be smaller in those industries.

In Figure 5, we verify this intuition. As in Figures 5 - 6, we solve the industry equilibrium for different values ϕ^A. The top line plots the gross TFP loss from markup dispersion, $\frac{\sigma}{2} Var\left(\log \mu\right)$, which is larger in industries with low ϕ^A, as in Figure (6). The bottom line plots the net TFP loss from markup dispersion, $\frac{\sigma}{2} Var\left(\log \mu\right) + \sigma\alpha Cov\left(\log \mu, \log\left(1+\tau_k\right)\right)$, which is instead smaller in industries with low ϕ^A, exactly because markups are more negatively correlated with constraint tightness in industries with low ϕ^A.

6 Conclusion

In this paper, we document new facts that link firms' markups to borrowing constraints: (1) firms with looser constraints have higher markups, especially in industries where assets are difficult to borrow

Figure 7: TFP Loss From Markup Dispersion



against and firms rely more on earnings to borrow; (2) markup dispersion is also higher in industries where firms rely more on earnings to borrow. We explain these relationships using a standard Kimball demand model augmented with borrowing against assets and earnings. The key mechanism is a two-way feedback between markups and borrowing constraints. First, firms with looser constraints charge higher markups, as these constraints allow them to attain larger market shares. Second, higher markups relax borrowing constraints when firms rely on earnings to borrow, as those with higher markups have higher earnings. The interaction between markups and borrowing constraints has important allocative efficiency implications. High-markup firms can be too larger because they face looser constraints. Introducing borrowing constraints also lowers the overall TFP losses from markup dispersion.

References

Adler, Konrad. Financial covenants, firm financing, and investment. Working paper, 2024.

Aghion, Philippe, Antonin Bergeaud, Timo Boppart, Peter Klenow, and Huiyu Li. Good rents versus bad rents: R&d misallocation and growth. Technical report, Working Paper, 2022.

Baqaee, David R, Emmanuel Farhi, and Kunal Sangani. The supply-side effects of monetary policy. *Journal of Political Economy*, 2024a. 132(4):1065–1112.

Baqaee, David Rezza and Emmanuel Farhi. Productivity and misallocation in general equilibrium. *The Quarterly Journal of Economics*, 2020. 135(1):105–163.

Baqaee, David Rezza, Emmanuel Farhi, and Kunal Sangani. The darwinian returns to scale. *Review of Economic Studies*, 2024b. 91(3):1373–1405.

Bils, Mark, Peter J Klenow, and Cian Ruane. Misallocation or mismeasurement? *Journal of Monetary Economics*, 2021. 124:S39–S56.

Bond, Steve, Arshia Hashemi, Greg Kaplan, and Piotr Zoch. Some unpleasant markup arithmetic: Production function elasticities and their estimation from production data. *Journal of Monetary Economics*, 2021. 121:1–14.

Buera, Francisco J., Joseph P. Kaboski, and Yongseok Shin. Finance and development: A tale of two sectors. *American Economic Review*, 2011. 101(5):1964–2002.

Chevalier, Judith A and David S Scharfstein. Liquidity constraints and the cyclical behavior of markups. *American Economic Review*, 1995. 85(2):390–396.

David, Joel M and Venky Venkateswaran. The sources of capital misallocation. *American Economic Review*, 2019. 109(7):2531–2567.

De Loecker, Jan, Jan Eeckhout, and Gabriel Unger. The rise of market power and the macroeconomic implications. *Quarterly Journal of Economics*, 2020. 135(2):561–644.

De Ridder, Maarten. Market power and innovation in the intangible economy. *American Economic Review*, 2024. 114(1):199–251.

De Ridder, Maarten, Basile Grassi, and Giovanni Morzenti. The hitchhiker guide to markup estimation. 2024. Working paper.

Dotsey, Michael and Robert G King. Implications of state-dependent pricing for dynamic macroeconomic models. *Journal of Monetary Economics*, 2005. 52(1):213–242.

Drechsel, Thomas. Earnings-based borrowing constraints and macroeconomic fluctuations. *American Economic Journal: Macroeconomics*, 2023. 15(2):1–34.

Edmond, Chris, Virgiliu Midrigan, and Daniel Yi Xu. How costly are markups? *Journal of Political Economy*, 2023. 131(7):1619–1675.

Eisfeldt, Andrea L and Yu Shi. Capital reallocation. *Annual Review of Financial Economics*, 2018. 10(1):361–386.

Farre-Mensa, Joan and Alexander Ljungqvist. Do measures of financial constraints measure financial constraints? *Review of Financial Studies*, 2016. 29(2):271–308.

Gilchrist, Simon, Raphael Schoenle, Jae Sim, and Egon Zakrajšek. Inflation dynamics during the financial crisis. *American Economic Review*, 2017. 107(3):785–823.

Hsieh, Chang-Tai and Peter Klenow. Misallocation and manufacturing tfp in china and india. *The Quarterly journal of economics*, 2009. 124(4):1403–1448.

Kaplan, Steven N and Luigi Zingales. Do investment-cash flow sensitivities provide useful measures of financing constraints? *Quarterly Journal of Economics*, 1997. 112(1):169–215.

Kermani, Amir and Yueran Ma. Asset specificity of nonfinancial firms. *The Quarterly Journal of Economics*, 2023. 138(1):205–264.

Khan, Aubhik and Julia K Thomas. Credit shocks and aggregate fluctuations in an economy with production heterogeneity. *Journal of Political Economy*, 2013. 121(6):1055–1107.

Kim, Ryan. The effect of the credit crunch on output price dynamics: The corporate inventory and liquidity management channel. *Quarterly Journal of Economics*, 2021. 136(1):563–619.

Klenow, Peter J and Jonathan L Willis. Real rigidities and nominal price changes. *Economica*, 2016. 83(331):443–472.

Li, Huiyu. Leverage and productivity. *Journal of Development Economics*, 2022. 154:102752.

Lian, Chen and Yueran Ma. Anatomy of corporate borrowing constraints. *The Quarterly Journal of Economics*, 2021. 136(1):229–291.

Meinen, Philipp and Ana Cristina Soares. Markups and financial shocks. *Economic Journal*, 2022. 132(647):2471–2499.

Midrigan, Virgiliu and Daniel Yi Xu. Finance and misallocation: Evidence from plant-level data. *American economic review*, 2014. 104(2):422–458.

Minseog, Kim. and Geunyong Park. Corporate debt maturity and output price dynamics. Technical report, Working Paper, 2024.

Moll, Benjamin. Productivity losses from financial frictions: Can self-financing undo capital misallocation? *The American Economic Review*, 2014. 104(10):3186–3221. ISSN 00028282. URL http://www.jstor.org/stable/43495317.

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

Ottonello, Pablo and Thomas Winberry. Capital, ideas, and the costs of financial frictions. *NBER Working Paper*, 2024.

Renkin, Tobias and Gabriel Züllig. Credit supply shocks and prices: Evidence from danish firms. *American Economic Journal: Macroeconomics*, 2024. 16(2):1–28.

Restuccia, Diego and Richard Rogerson. Policy distortions and aggregate productivity with heterogeneous establishments. *Review of Economic Dynamics*, 2008. 11(4):707–720. ISSN 1094-2025. URL https://www.sciencedirect.com/science/article/pii/S1094202508000203.

Whited, Toni M and Guojun Wu. Financial constraints risk. *Review of Financial Studies*, 2006. 19(2):531–559.

Appendices for:
Borrowing Constraints, Markup Dispersion, and Misallocation

This Appendix contains further material for the article “Borrowing Constraints, Markup Dispersion, and Misallocation.” Any references to equations, figures, tables, assumptions, propositions, lemmas, or sections that are not preceded by “A.”—“C.” refer to the main article.

A Supplementary theoretical details

[Under Construction]

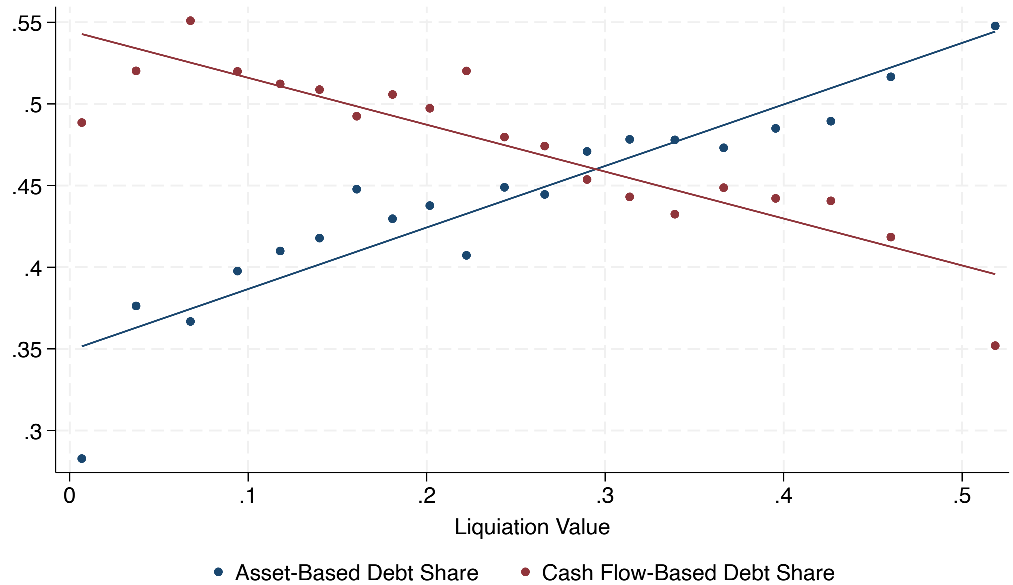
B Additional results for quantitative analysis

[Under Construction]

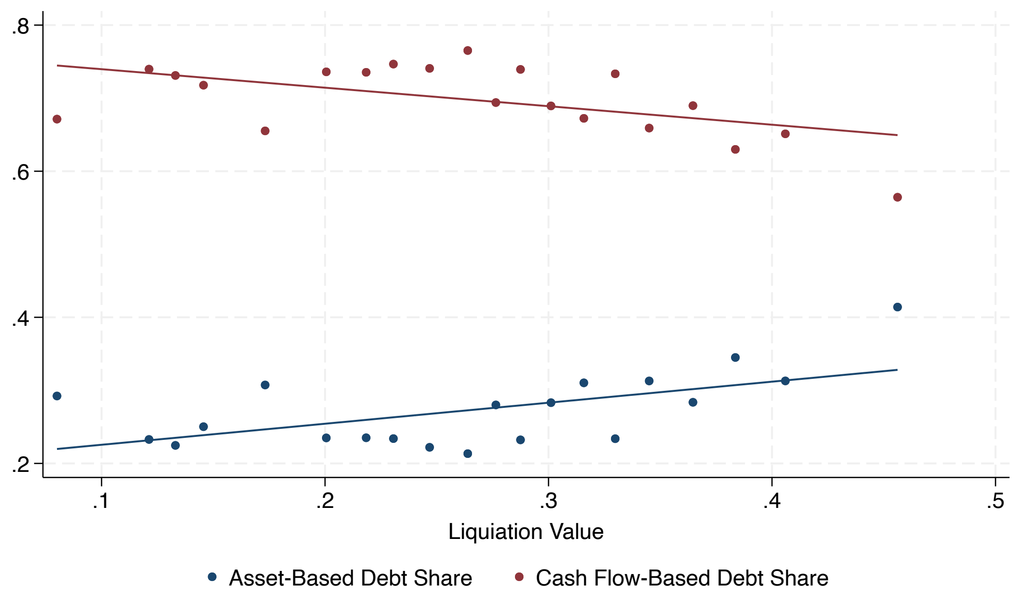
C Additional empirical results

Figure C.1: Liquidation Value and Firm Borrowing

Panel A. Firm Level



Panel B. Industry Level



Notes. Panel A shows binscatter plot of firm-year level share of asset-based debt and cash flow-based debt in total debt against the firm's liquidation value of fixed assets and working capital. Firm-level liquidation value is the book value of property, plant, and equipment, inventory, and receivables multiplied by their respective industry-level liquidation recovery rates from Kermani and Ma (2023), normalized by book assets. Panel B shows binscatter plot of industry-year level share of asset-based debt and cash flow-based debt in total debt against the industry's liquidation value of fixed assets and working capital. Industry-level debt share is calculated using total asset based debt and cash flow-based debt in the industry divided by total debt in the industry. Industry-level liquidation value is the book value of property, plant, and equipment, inventory, and receivables multiplied by their respective industry-level liquidation recovery rates from Kermani and Ma (2023), normalized by total book assets.

Table IA1: Financial Constraint and Firm Level Markups, Controlling for Firm Age

	Firm Log Markup			
	(1)	(2)	(3)	(4)
Cash/Assets	0.45***	0.40***	0.27***	0.35***
	(0.04)	(0.03)	(0.02)	(0.05)
Log Firm Age	-0.01	-0.01	-0.01**	-0.01
	(0.01)	(0.01)	(0.01)	(0.00)
Fixed Effects	Industry × Year			
Observations	54,150	38,621	29,927	44,090
R^2	0.33	0.29	0.23	0.34

Notes. This table shows firm-year level regressions Log Markup$_{it} = \alpha_{ind(i)t} + \beta$Cash/Assets$_{it} + \gamma$Log Age$_{it} + \varepsilon_{it}$, among US Compustat firms. Log Markup$_{it}$ for firm i in year t uses the baseline markup from De Loecker, Eeckhout, and Unger (2020) in column (1), the markup with overhead in production input from De Loecker, Eeckhout, and Unger (2020) in column (2), the markup using accounting ratio (sales/cost of goods) sold from De Loecker, Eeckhout, and Unger (2020) in column (3), and translog markup from De Ridder (2024) in column (4). Compustat does not have a direct measure of firm age. We use the older one between year since incorporation (we collect incorporation year data from Datastream) and year since IPO (IPO year is from Compustat). Industry (3-digit NAICS code) by year fixed effects are included. Standard errors are double clustered by industry and year.