NBER WORKING PAPER SERIES

SORTING OUT THE REAL EFFECTS OF CREDIT SUPPLY

Briana Chang
Matthieu Gomez
Harrison Hong

Working Paper 28842
<http://www.nber.org/papers/w28842>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
May 2021

We thank Samuel Hanson, David Scharfstein, Raj Iyer, Jose-Luis Peydro, Rafael Repullo, Daniel Paravisini, Amir Sufi, Michael Schwert, Oliver Giesecke, Shangjin Wei, and seminar participants at American Finance Association Meetings, NBER Corporate Finance Meetings, Imperial College Business School, Columbia University, Seoul National University, and Fudan University for helpful comments. Address correspondence to Briana Chang, Department of Finance, Wisconsin School of Business, UW-Madison, 5274 E Grainger Hall, 975 University Avenue, Madison, WI 53703, Tel: 608-265-9171. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2021 by Briana Chang, Matthieu Gomez, and Harrison Hong. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Sorting Out the Real Effects of Credit Supply
Briana Chang, Matthieu Gomez, and Harrison Hong
NBER Working Paper No. 28842
May 2021
JEL No. G0,G01,G2,G23

ABSTRACT

We document that banks which cut lending more during the Great Recession were lending to riskier firms. To explain this evidence, we build a competitive matching model of bank-firm relationships in which risky firms borrow from banks with low holding costs. Based on default probabilities and equilibrium loan rates, we use our sorting model to recover the latent bank holding cost distribution. The measure of banks with low holding costs dropped during the Great Recession. This credit supply shift conservatively accounted for around 50% of the decline in corporate loans over this period. Our attribution cannot be captured by panel regression estimates from the bank lending channel literature.

Briana Chang
Department of Finance, Investment
and Banking
University of Wisconsin - Madison
bchang@bus.wisc.edu

Matthieu Gomez
Columbia University
435 Riverside Drive, Apt 41
10025 New York
gomez.matthieu@gmail.com

Harrison Hong
Department of Economics
Columbia University
1022 International Affairs Building
Mail Code 3308
420 West 118th Street
New York, NY 10027
and NBER
hh2679@columbia.edu

1. Introduction

A large macro-finance literature seeks to quantify the effects of bank lending for business cycle outcomes (Bernanke and Blinder (1988), Bernanke and Gertler (1989), Kashyap, Stein, and Wilcox (1993)). One particularly striking episode is the Great Recession, during which aggregate loan volume dropped by half (Ivashina and Scharfstein (2010)). How much of this drop was due to changes in financial markets (shift in credit supply) as opposed to changes in firm riskiness (shift in credit risk) is still an open question. We provide an answer using a matching model of bank lending.

First, we provide strong evidence of sorting between banks and firms. Banks which cut lending the most during the Great Recession were lending to the riskiest firms. This correlation is robust across all these measures of firm downside risk used in the banking literature: (1) borrower loan spread or interest rates, (2) the yields of the public debt for firms that issue in both public and syndicated private corporate loan markets, and (3) borrower leverage. The economic magnitudes are large. For instance, a 1 p.p. higher borrower loan or bond spread is associated with a decrease of bank lending growth of 0.70 standard deviation, with a t -statistic of around 4.

We then examine the ex-ante characteristics of the banks that were lending to the riskiest firms. These banks typically had a low deposit-to-asset ratio and a close proximity to Lehman in terms of co-syndication of corporate loans. In short, they were investment banks rather than commercial banks. This suggests that this sorting pattern is driven by banks' ability to take on risk, presumably due to their talent to sell securitized loans. For instance, marketing documents for Lehman during this period pointed to their ability to package the riskiest loans and to sell them to institutional investors compared to other banks.¹ While firms in theory could have borrowed from banks with differential exposure

¹See <https://bit.ly/3j0rz9H>. Banks that securitized had a competitive advantage over traditional banks during the years before the Great Recession of 2008 (Nadauld and Weisbach (2012)).

to securitization so as to diversify their funding risks during the crisis period, they did not in practice.

Second, we develop a competitive one-to-one matching model of a credit market where banks differ in their holding costs (per unit of risk). For the Great Recession episode, we associate lower holding costs with bank talent to sell securitized loans. But bank holding costs can in principle be related to different bank characteristics emphasized in the literature:² our approach does not require us to take a stand on relevant underlying bank characteristics.

The joint surplus of a given match is simply the NPV of the project minus the total cost of the loan. When there is complementarity between firm riskiness and bank talent, firms with greater downside risk borrow from banks with lower holding costs in equilibrium, consistent with the sorting in the data. The interest rate of the loan determines the split of the surplus. The riskiest firms benefit the most from the best banks (i.e. those with lowest per unit holding costs) and thus are willing to pay higher interests to these banks. In contrast, the safest firms with the best credit ratings benefit the least from the best banks and so they match with the worst banks and get all the surplus in the relationship. The loan rates are dependent on the distribution of bank holding costs (or equivalently talent to sell securitized loans) and the distribution of firm credit risk across the entire credit market.

A market clearing condition characterizes the type of the marginal lender, and therefore the mass of firms that obtain loans. The marginal type of firm that gets a loan can be understood from a social planner's perspective: the benefit of lending to the marginal firm must equal its marginal cost. The marginal benefit is the NPV of the project net of the holding costs of the best bank. The marginal cost is moving other firms to worse

²These include size (Chen and Song (2013), Berger, Miller, Petersen, Rajan, and Stein (2005)), bank capitalization (Schwert (2018)), proximity (Petersen and Rajan (2002)), specialized knowledge (Paravisini, Rappoport, Schnabl, and Wolfenzon (2014)) or networks (Engelberg, Gao, and Parsons (2012)).

banks. Hence, an outward shift in the distribution of firm risk or an outward shift in the bank holding cost distribution implies that fewer firms can borrow or lower aggregate lending because this marginal cost increases.

Third, we then show how to use our model to attribute time-series changes in aggregate lending (i.e. between a pre-crisis and a crisis period) to shifts of these distributions. In general, default rates of loans are observable while bank holding cost per unit of credit risk is not. One method to estimate matching models when underlying agent distributions are homogeneous over time is from [Gabaix and Landier \(2008\)](#), who consider CEO-firm matching. However, the distribution of bank holding costs is likely to change across non-crisis and crisis periods.

In the Great Recession, the fragility of the securitized loan market ([Gennaioli, Shleifer, and Vishny \(2012\)](#), [Gorton and Metrick \(2012\)](#)) led to the financial distress of investment banks. This suggests that it is the measure of low cost banks (reliant on the securitized loan market) that shrinks. But in other periods, it might be the measure of banks with other characteristics that declines. For instance, time-varying risk-appetite of a representative investor for credit risk as in [Greenwood and Hanson \(2013\)](#) or [López-Salido, Stein, and Zakrajšek \(2017\)](#) can lead to time-varying cost of capital for different types of banks in different periods depending on their exposure to credit sentiment.³

Hence, we provide an alternative approach to account for time-varying bank holding cost distributions. We show how to recover bank holding costs by using default rates and equilibrium loan payoffs or interest rates. We then take these estimates and use our model to calculate a counterfactual — holding fixed the pre-crisis firm default rates, how much can the change in distribution of bank holding costs account for the drop in the loans supplied to risky firms?

Fourth, we then apply our method to corporate loan data from the Great Recession.

³Alternatively, low interest rates can lead to search for yield behavior that affects credit supply ([Martinez-Miera and Repullo \(2017\)](#)).

For our empirical analysis, we focus on our subsample of firms that issue both public debt and that borrow from banks. There was roughly 4000 firm loans in the pre-crisis period (2005-2007) and but only 2500 firm loans during the Great Recession or crisis period (2008-2010). We use the Moody's credit rating of the firm's last issued debt as our proxy for firm credit risk and then calculate our measure of bank holding cost using the historical default probabilities associated with these credit ratings. These cross-sectional distributions are estimated separately for the 2005-2007 pre-crisis period and the 2008-2010 crisis period.

Our novel method to recover bank holding cost using default rates and equilibrium interest rates yields sensible results. Consistent with the financial distress of investment banks such as Lehman, we find that there was a smaller measure of low cost banks (i.e. those best at selling securitized loans) in the crisis period—or equivalently by the market clearing condition, a smaller measure of firms receiving loans from low cost banks. We provide bootstrap standard errors and discuss some robustness calculations regarding these estimates of credit supply. Firm credit risk also went up significantly even among the firms that received credit during the crisis (i.e. an outward shift in credit risk). And of course far fewer high credit risk firms received any loans, consistent with a potentially important role of credit risk in the observed decline in aggregate lending during the crisis.

We then attribute how much of the decline in the extensive margin (from 4000 firm loans to 2500 firm loans) is due to the increase in bank holding costs versus firm default probabilities. Our estimates point to a sizable effect due to an outward shift in bank holding costs. We find that this credit supply shift accounts for conservatively 49.3% of the observed drop in aggregate number of loans during the Great Recession.⁴ Using our most aggressive estimates for the change in the bank holding cost distribution, the credit supply effect is as large as 71.9%. An equivalent calculation to identify a credit

⁴The remaining 50% is attributed to changes in credit risk of firms.

risk channel is to hold fixed the bank holding cost distribution and ask how much can the change in firm credit risk explain of the drop in aggregate loans.

Our calculation of the aggregate lending decline weighs all loans equally but a more economically sensible measure would be to risk-weight the loans. In this instance, the role of bank holding cost heterogeneity is even pronounced because it was the banks best at securitization and that could absorb the riskiest loans pre-crisis that could no longer perform this role during the crisis.

Finally, we connect our model to the reduced form approach in the bank lending channel literature which regresses firms loans on the health of the banks they are borrowing from (Khwaja and Mian (2008), Amiti and Weinstein (2018)). Because this panel regression approach does not explicitly model the reallocation of firms to banks (extensive margins), it is generally silent on how to relate cross-sectional effects of bank health to their effects on aggregate lending (see also Herreno (2019)). Nonetheless, it is worthwhile to consider how panel regression estimates behave in our setting.

Using an extension of our model to allow a firm to borrow from two banks, we show that larger aggregate credit supply effects in our model correspond to smaller firm fixed-effects panel regression estimates.⁵ We then additionally introduce a simple of form stickiness in bank-firm relationships. We show that our credit supply effect attribution is robust to this stickiness, but this stickiness does lead to larger firm fixed effects panel regression estimates.⁶

Our paper makes several contributions. In terms of empirical findings, we focus on an under-appreciated dimension of sorting: we show that, before the Great Recession, investment banks tended to lend to riskier firms. In terms of methodology, we use assortative matching models developed for CEO labor, underwriters, and venture capitalists

⁵See also Jiménez, Mian, Peydró, and Saurina (2019) for another rationale for why firm fixed effects estimates need to be adjusted to be interpreted as capturing credit supply effects.

⁶Moreover, we show that the absence of difference between OLS and firm fixed effects estimates cannot be interpreted as evidence of random sorting (Chodorow-Reich (2013)).

(Terviö (2008), Gabaix and Landier (2008), Sørensen (2007), Chang and Hong (2019)) to shed light on banking relationships. We show how the model can be used to infer the bank holding cost distribution and to separate credit risk versus supply effects. Future work can consider sorting on multiple dimensions (Chiappori and Salanié (2016)). Our approach to disentangle credit supply from credit risk shocks can be applied to other settings, including in the context of mortgages since Justiniano, Primiceri, and Tambalotti (2019) point to the importance of lending standards for the housing boom.

2. Empirical Evidence

In this section, we document that banks whose health suffered more during the Great Recession were lending to riskier firms. Moreover, we show that these banks have low deposit-to-asset ratios and closer relationships with Lehman. Taken together, our results suggest that sorting is a first-order concern when addressing bank capital shocks for aggregate loans and firm output during the Great recession.

2.1. Data

Dealscan. Our principal source of data is the Thomson Reuters Dealscan database. Dealscan collects loan-level information on syndicated loans from Securities and Exchange Commission filings, company statements, and media reports. The data include the names of the borrowers and lenders, the purpose of the loan, the role of each bank in the agreement (underwriter, agent, adviser, etc...), and the interest rate of the loan. Following the literature, we exclude loans to financial companies (SIC between 6000 and 6999) from the sample. Because we focus on the Great Recession period around 2008-2009, we only keep borrowers that obtained a loan between 2004 and August 2008 or that obtained a loan prior to 2004 which matured after October 2007.

Measures of Firm Risk. We are interested in borrower characteristics that are associated with downside or credit risk. We propose three different measures. The first measure is the average of all-in-drawn *loan spread*, which is the borrower’s last loan credit spread over LIBOR plus annual fees to the lenders.

The second measure is the *bond spread*, which is the average spread of public bonds issued by the borrower measured in January 2007 before the crisis. To obtain this borrower level characteristic, we first look for bonds issued by the borrower in the Fixed Income Securities Database (FISD).⁷ We then obtain the spread of each bond in January 2007 from the Lehman Corporate Bond Data. We obtain a borrower level spread by averaging the spread of all outstanding bonds in January 2007, weighted by their face value. Because this measure is only available for firms that issue public bonds, this only covers 30% of the initial sample of all borrowers.

The last measure is the *market leverage* of each firm in January 2007. This measure is only available for public firms. To obtain it, we match borrowers to Compustat using the DealScan-Compustat Link from [Chava and Roberts \(2008\)](#). We define leverage as the ratio of market value of asset to market value of equity.⁸ While our first measure (loan spread) is available for all firms in our sample, the two other measures (bond spread and market leverage) are only available for public firms.

In [Appendix A](#), we show that these three measures of downside risk are negatively related to firm performance during the financial crisis—that is our measures of ex-ante downside risk are correlated with ex-post performances during economic downturns. As such, the investment opportunities of firms with greater downside risk are more sensitive to the business cycle (see also [Philippon \(2009\)](#)).

⁷We match Dealscan with FISD by using the first 6 numbers of the CUSIP, keeping only matches with similar names.

⁸The market value of asset is constructed as the market value of equity plus debt in current liabilities and long term debt.

2.2. Bank Health and Firm Risk

We are interested in measuring the systematic relationship between a bank health or performance during the financial crisis and the risk of the firms it was lending to. To do this, we estimate by OLS the following model:

$$\Delta L_{it} = \alpha + \beta \bar{X}_i + \epsilon_{it},$$

where i denotes a bank, ΔL_{it} is the bank lending growth during the crisis, and $\bar{X}_i$ reports the average of the firms' characteristics for all syndicated loans in which the bank is a lead lender. Bank lending growth is defined as

$$\Delta L_{it} = \frac{L_{\text{crisis}}}{L_{\text{normal}}},$$

where L_{crisis} denotes the total amount of loans originated in the 9-month period from October 2008 to June 2009, and L_{normal} denotes the average of the the total amount of loans originated from October 2005 to June 2006 and October 2006 to June 2007.⁹

Panel A of Table 1 reports the result of regressing bank lending growth during the crisis on our three measures of firm downside risk. We find that all three measures of firm downside risk negatively predict corresponding bank lending growth: the banks that did the worst during the crisis were systematically lending to firms with higher loan spread, higher bond spread, and higher leverage.

Using borrower loan spread, we find that a one p.p. increase in the average borrower loan spread is associated with a decrease of bank lending growth by 0.65 standard deviation. This is not only a large economic effect but also highly statistically significant with a t -statistic of 3.25. The R^2 from the regression is 19%. When we use borrower bond

⁹That is, we are measuring the abnormal loan growth over the financial crisis period by normalizing with loan growth over the prior years during the same October to June window.

Table 1: Sorting on Observables

	β	t -stat	R^2	N
	(1)	(2)	(3)	(4)
<i>Panel A: Bank Lending Growth 06-09</i>				
Borrower Loan Spread	−0.65***	3.25	0.19	43
Borrower Bond Spread	−0.70***	4.21	0.21	38
Borrower Leverage	−3.65***	3.32	0.24	43
<i>Panel B: Bank Lehman Distance</i>				
Borrower Loan Spread	−1.37***	3.39	0.37	42
Borrower Bond Spread	−1.18***	3.34	0.24	37
Borrower Leverage	−8.33***	3.99	0.52	42
<i>Panel C: Bank Deposit</i>				
Borrower Loan Spread	−1.24***	5.30	0.48	43
Borrower Bond Spread	−1.06***	4.07	0.32	38
Borrower Leverage	−6.15***	6.55	0.47	43

Notes: This table estimates the model

$$Y_i = \alpha + \beta \bar{X}_i + \epsilon_i,$$

where i denotes a bank, Y_i is alternatively the bank lending growth from 2006-2009 (Panel A), bank Lehman distance (Panel B), bank deposit (Panel C). $\bar{X}_i$ denotes the average observable of borrowers from bank i in 2004-2006. It is alternatively the borrower loan spread, bond spread, and market leverage. Estimation with WLS — estimates in column (1) and t -statistics in column (2).

spread as our measure of firm downside risk, we get a decrease of bank lending growth by 0.70 standard deviation with a t -statistic of 4.21. Furthermore, a one unit increase in the average borrower leverage is associated with a decrease of bank lending growth by 3.7 standard deviation and t -statistic of 3.32. Again, regardless of the measure of borrower downside risk we use, we find significant sorting.

To visualize this sorting result, Figure 1 displays a scatter-plot between bank lending growth from 2006 to 2009 on the y-axis and borrower loan spread on the x-axis. The larger circles indicate the number of loans from the bank and hence reflect the size of the bank. Investment banks such as Lehman Brothers, Bear Stearns, and Merrill Lynch all drastically decreased lending during the period and were also matched with borrowers with high spreads, i.e. high downside risk firms.

Overall, our results suggest a strong positive assortative matching between banks and borrowers. Banks that were particularly hit during the financial crisis were lending to the

Figure 1: Bank Lending Growth 2006-2009 and Corporate Borrower Loan Spread

Notes: This figure plots (standardized) bank lending growth over the period of 2006-2009 on the corporate borrower loan spread in the bank's portfolio of loans.

riskiest firms. In this world, it is hard to differentiate between the “bank” or credit supply effect and the “firm” or credit risk effect in explaining the negative real effects during the Great Recession associated with risky firms borrowing from hit banks.

2.3. Bank Characteristics and Firm Risk

We now explore the characteristics of these banks: what were the ex-ante characteristics of the banks lending to the riskier firms in the economy before the financial crisis? We find that these banks tend to have low bank deposit to asset ratio, as well as a high proximity to Lehman. These findings have direct implications for the existing literature on the financial crisis, which has used these characteristics as instruments for bank performance following [Ivashina and Scharfstein \(2010\)](#). Our results show that these bank characteristics are systematically correlated with the riskiness of the firms they are lending to.

We first study the correlation between firm risk and the distance to Lehman, defined as the fraction of a bank’s syndication portfolio where Lehman Brothers had no lead role. As above, the dependent variable is standardized to have a standard deviation of one. Panel B of Table 1 reports the result of regressing bank Lehman distance on the three measures of borrower characteristics. We find that all three measures of borrower downside risk are negatively related to Lehman distance. The correlation is actually even stronger quantitatively than when it comes to predicting bank lending growth in Panel A. Using borrower loan spread, a one p.p. increase in the average borrower loan spread is associated with a decrease of Lehman distance by more than 1.37 standard deviations. The t -statistic is 3.39. The R^2 is now 37%. Using borrower bond spread, the economic magnitude with minus 1.18 standard deviations with a t -statistic of 3.34. A one unit increase in the average borrower leverage is associated with a decrease of bank lending growth by 8.33 standard deviation. The t -statistic is 3.99 and the R^2 is 52%.

We then study the correlation between firm risk and the ratio of *bank deposit to asset* (Chodorow-Reich (2013)). Panel C of Table 1 reports the result of regressing bank deposit on the three measures of firm downside risk. Quantitatively, a one p.p. increase in the average borrower loan spread is associated with a decrease of bank deposits by 1.24 standard deviations with a t -statistic of 5.3 and an R^2 of 48%. Using the borrower loan spread, the coefficient is -1.06 with a t -statistic of 4.07. A one unit increase in the average borrower leverage is associated with a decrease of bank lending growth by 6.15 standard deviations. The t -statistic is 6.55 with an R^2 of 47%.

Our findings suggest that, before the crisis, the banks lending to the riskiest firms looked like investment banks, rather than commercial banks. This may be explained by their comparative ability to take risk, as well as their easier access the securitization markets. In fact, many investment banks such as Lehman, Goldman or Bear Stearns were leaders in the securitization of risky corporate loans (Shivdasani and Wang (2011)). As

mentioned above, marketing documents for Lehman during this period pointed to their special ability to assess and package the riskiest loans compared to other banks.

In short, our results show that certain banks (e.g. investment banks) were lending to the riskiest firms, and that they particularly suffered during the crisis. This pattern suggests that understanding the matching between banks and firms is key to understanding the distribution of credit during the Great Recession. Firms in theory could have borrowed from banks with differential exposure to securitization so as to diversify their funding risks during the crisis period. But they did not in practice. During this period, a firm's first and second lenders' health tended to be highly correlated, around 0.9 and highly statistically significant.

3. Matching in Credit Markets

In this section, we develop a frictionless and competitive matching model of a credit market. Firms differ in their credit riskiness and banks differ with respect to holding costs per unit of credit risk. We use the sorting model to establish the connection between interest rates and the extensive margin of loan supply to the underlying distribution of bank holding costs, which thus allows us to examine the aggregate effect of credit supply shocks.

3.1. Model

Credit Market. Firms own projects but do not have capital. Banks are endowed with capital, and they can either offer funding to a firm or earn a risk-free gross return $1 + r_f$.

There is a continuum of firms. Each firm has one project that requires 1 unit of investment. The project of firm $i \in [0, N]$ succeeds with probability $1 - \delta[i]$ and yields return $y_H[i]$. It fails with probability $\delta[i]$ and yields return $y_L[i]$. A high i denotes a firm

with a higher probability of default. That is, firms are ranked by their default probability $\delta'[i] \geq 0$. We make the following two assumptions:

Assumption A1. $y_H[i] \geq 1 + r_f > y_L[i]$, i.e. the payoff is large enough to pay back the loan at the risk free rate only if the project succeeds.

Assumption A2. The NPV of the project, $NPV[i] \equiv (1 - \delta[i])y_H[i] + \delta[i]y_L[i] - (1 + r_f) \equiv y$, is constant across i .

There is also a continuum of heterogeneous risk-neutral banks, indexed by $j \in [0, N]$, who differ in their ability of handling risks, i.e holding costs. In the Great Recession period that is the motivation for our study, we associate this lower holding cost with better talent in selling securitized loans to institutional investors.

Specifically, let $C(i, j)$ denote the total cost of holding risk for bank j when lending to firm i . We assume that $C_1(i, j) \geq 0$: fixing any bank j , the cost is higher if lending to riskier firms (i.e., the ones with higher default probability). We also assume that $C_2(i, j) \leq 0$: a bank with a higher j is better at holding risk in the sense that it has a lower holding cost given any firm i .

For simplicity, we assume that each firm can only at most borrow from one bank. Any individual firm i can be matched with a bank j or remains autarky. That is, we use one-to-one matching to model the relationship between firms and banks. The one-to-one relationship can also be interpreted as matching between a single project and a bank manager in an environment where firms have multiple projects and banks have multiple managers, where each project needs one dollar and each manager can lend one dollar (see Appendix B).

Within a match, bank and firm agree on a debt contract, which specifies a promised repayment d . We assume that the project is pledgeable and thus when the firm defaults, the bank obtains the value of the project. Banks understand that they will receive

$\min\{d, y_L[i]\}$. A bank's payoff when lending to firm i with repayment d thus yields,

$$w(i, j|d) = (1 - \delta[i])d + \delta[i]y_L[i] - C(i, j) - (1 + r_f).$$

The firm's payoff, on the other hand, yields

$$u(i, j|d) = (1 - \delta[i])(y_H[i] - d).$$

While a firm's payoff does not directly depend on its matching bank b , different banks can charge different interest rates in equilibrium, and thus banks' abilities can indirectly affect firms' payoffs.

The joint surplus between a matching pair is given by:

$$s(i, j) \equiv w(i, j|d) + u(i, j|d) = y - C(i, j),$$

which is simply the NPV of the project minus its holding cost and it is independent of the repayment d . In other words, the repayment d only affects how these two agents split the surplus, where a higher repayment d means a lower surplus to the firm.

Sorting in Equilibrium. Our equilibrium concept follows the standard assignment framework. The bank's decision can be rewritten as choosing the firm optimally while taking the equilibrium utility $U[i]$ as given:

$$W(j) = \max_i \{y - C(i, j) - U[i]\}. \quad (1)$$

Fixing any given level of firm payoff, lending to a riskier firm means a higher holding cost. Hence all banks prefer to match with safer firms. As all banks compete for those firms, these firms must obtain higher payoffs. That is, safer firms must receive a better loan

term in the sense that $U[i]$ must decrease in i .

All banks thus trade off between a riskier firm vs. a higher rate. Banks lending to the riskier firms are compensated by getting a higher payment. The matching outcome is then determined by which bank is more willing to absorb firm risk.

The following lemma gives the condition for sorting.

Lemma 1. *If $C_{12}(i, j) < 0$, a more talented bank is matched with a riskier firm.*

The matching in general depends on the complementarity between firm riskiness and bank ability in the surplus function, where $s_{12}(i, j) = -C_{12}(i, j)$. When $C_{12}(i, j) < 0$, it means that a bank's ability matters more for a riskier firm in the sense that it can reduce the firm's borrowing cost more. Thus, it is relatively cheap for banks with higher ability to take on the riskier loans and $s_{12}(i, j) > 0$. In case when such cost is additive, for example, it then implies no sorting.

Given the sorting outcome i, j the firms' utility $U[i]$ is pinned down competitively in equilibrium so that all banks are matched with the correct firms. That is,

$$U'[i] = -C_1(i, j^*(i)). \quad (2)$$

As in the standard assignment framework, the marginal payoff increase for firm i is given by his contribution to the surplus within the match given his optimal assignment $j^*(i)$.

Since we observe that riskier firms tend to borrow from banks that are better at securitization (the ones with lower holding costs), we then focus on the characterization where $C_{12}(i, j) < 0$. In this case, we can prove the following result:

Lemma 2. *When $C_{12}(i, j) < 0$, the equilibrium consists of a marginal rank $i^* \in [0, N]$ such that (1) for all $i \leq i^*$, his matching bank is given by $j^*(i) = N - (i^* - i)$ and (2) $U'[i]$ solves Equation (2) with $U[i^*] = 0$.*

Notice that the marginal risky firm i^* is matched with the best bank $j^*(i^*) = N$, while the safest firm $i = 0$ is matched with the marginal bank $N - i^*$.

3.2. Credit Rationing

The aggregate loan supply in our model is described by the cutoff type i^* , the rank of the last firm getting credit (i.e. the marginal firm). As shown below, our model naturally gives rise to credit rationing in the sense that some positive NPV projects are excluded.

Derivation. According to Lemma 2, given $U'[i]$ and the boundary condition $U[i^*] = 0$, we have

$$0 = U[0] - \int_0^{i^*} C_1(i, j^*(i)) di. \quad (3)$$

Given that the safest type receives all the surplus $U[0] = y - C(0, N - i^*)$. Equation (3) can then be rewritten as

$$y - C(0, N - i^*) = \int_0^{i^*} C_1(i, j^*(i)) di. \quad (4)$$

This equation pins down i^* , the rank of the marginal firm getting credit.

Alternatively, we can derive the same condition from the view point of banks in a similar way. Applying the envelope theorem to Equation (1) and using the fact that the best bank receives all the surplus, where $W[N^*] = y - C(i^*, N)$, the condition yields

$$y - C(i^*, N) = - \int_{N-i^*}^N C_2(i^*(j'), j') dj', \quad (5)$$

which uniquely pins down the marginal type i^* . By integration by parts, Equations (4)

and (5) are equivalent.

Interpretation. Both equations highlight that the compensation to the best bank and the utility to the safest firm are determined by their competing banks and firms. Intuitively, if other banks become better at taking risks and/or other firms become safer, then the marginal type i^* goes up, which lowers $W[N]$ and $U[0]$.

Because the equilibrium is efficient, Equation (5) can be interpreted from the social planner's view point. The LHS of Equation (5) can be understood as the benefit of adding the marginal firm, which is the net NPV of that firm, taking into account the holding cost valued at the best bank. The RHS of Equation (5), on the other hand, captures the distributional effect that the holding costs have increased for the existing firms' loans, as each of them is now borrowing from a slightly worse bank. The marginal type is thus determined so that the benefit equals the cost.

Because of the negative sorting, some firms cannot borrow even though they have positive surplus. To see this, given that Equation (5) is strictly larger than zero, it means that, by continuity, firms that are slightly above the marginal type have positive surplus ($y - C(i^* + \epsilon, N) > 0$). They are nevertheless excluded. This is in contrast to the positive sorting case, where the marginal type is pinned down so that their joint surplus is zero. In this sense, our model captures credit rationing.

Intensive Margin. Our model focuses on the extensive margin. In Appendix B, we consider the case with endogenous loan size (where intensive margin is allowed) and show that similar methods can be applied as long as the loan size does not change much across firm risk ranking, which is what we observe in the data.

3.3. Comparative Statics

Given that the aggregate distribution matters for the aggregate loan supply, the following lemma below formalizes how it is affected by bank holding cost (credit supply effect) and by firm default probability (credit risk effect).

Bank Holding Cost Distribution. Let $C(i, j) = c(\delta[i], \kappa[j])$, where $c_\delta > 0$, $c_\kappa > 0$, $c_{\delta\kappa} > 0$, and $\kappa'[j] < 0$. That is, the holding cost $c(\delta, \kappa)$ increases with firms' default probability. Moreover, a more talented bank (banks with higher j) has a lower κ and thus lower holding costs, given $\delta[i]$.

We define talent as becoming more scarce when $\kappa'[j]$ becomes steeper (less talented banks) while the best bank's ability $\kappa[N]$ remains the same.¹⁰ And we define firms as becoming riskier if the default probability $\delta[i]$ is weakly higher under the new distribution for all firms with ranking i .

Lemma 3. *The total loan supply decreases when bank holding costs are higher (i.e. bank talent becomes more scarce), when firms are riskier, and when y is lower.*

4. Credit Supply versus Credit Risk Effects

During a crisis, typically both the distribution of firm riskiness and the distribution of bank holding cost shift. As shown in Lemma 3, the quantity of aggregate lending depends on both firm riskiness and bank holding cost. We now establish how to decompose these two effects with the existence of sorting.

¹⁰By definition, $\kappa[N] = \kappa[j] + \int_j^N \kappa'[j']dj'$. Hence, more scarce talent means a distribution under which $\kappa[j]$ is weakly higher.

4.1. Counterfactual

To this end, we need to formalize what we mean by credit risk and credit supply effects. Let $t \in \{0, 1\}$ denote the pre-crisis and crisis periods and $i^*(\delta_t[i], y_t, \kappa_t[j])$ denote the marginal type given the characteristics of banks $\kappa_t[j]$ and firms $(\delta_t[i], y_t)$. Define the counterfactual marginal type $i^*(\delta_0[i], y_0, \kappa_1[i])$ as the marginal firm type assuming the firm distribution remains the same between the two periods while the bank holding cost distribution changes.

According to Equation 5, $i^*(\delta_0[i], y_0, \kappa_1[i])$ is then given by the variable i^* that solves the following equation:

$$y_0 = c(\delta_0[0], \kappa_1[N - i^*]) + \int_0^{i^*} c_\delta(\delta_0[i], \kappa_1[N - (i^* - i)]) \delta'_0[i] di. \quad (6)$$

We can then define the credit supply effect ϕ_S as:

$$\phi^S \equiv \frac{i^*(\delta_0[i], y_0, \kappa_0[i]) - i^*(\delta_0[i], y_0, \kappa_1[i])}{i^*(\delta_0[i], y_0, \kappa_0[i]) - i^*(\delta_1[i], y_1, \kappa_1[i])}.$$

The numerator represents the change in the marginal lender type if firms' characteristics remain the same. The denominator represents the change in the loan volume during the crisis, where both firm characteristics and bank holding costs have changed. Therefore, the credit supply effect represents the contribution of changes in banks' characteristics to the change in credit volume. In particular, if all banks' characteristics remain the same (i.e., $\kappa_0[i] = \kappa_1[i]$), then $\phi^S = 0$, i.e. there is no credit supply effect.

Similarly, one can define the credit risk effect as the change in the loan supply when considering the counterfactual where banks' characteristics remain the same (taking the distribution at time $t = 1$):

$$\phi^D \equiv \frac{i^*(\delta_0[i], y_0, \kappa_1[i]) - i^*(\delta_1[i], y_1, \kappa_1[i])}{i^*(\delta_0[i], y_0, \kappa_0[i]) - i^*(\delta_1[i], y_1, \kappa_1[i])},$$

The credit risk effect represents the contribution of changes in firms' characteristics to the change in credit volume. As above, if all firms' characteristics remain the same (i.e., $y_0 = y_1$ and $\delta_0[i] = \delta_1[i]$), then $\phi^D = 0$, i.e. there is not credit risk effect. Note that, by definition, $\phi^S + \phi^D = 1$, i.e. the credit supply and credit risk effects sum up to one.

4.2. Estimating Bank Holding Cost Distribution

As our definition highlights, in order to quantify the credit supply effect, one must estimate the change in the bank holding cost distribution, which generally is not observable. This is particularly challenging when firm characteristics are also changing over time. In particular, the method used in the literature by [Gabaix and Landier \(2008\)](#) cannot be directly applied here, as it assumes that underlying talent distribution remains constant over time. Rather, we adapt their approach in the following manner.

First, we assume as they do that the loan cost function is separable in firm and bank types.

Assumption A3. Loan cost function is separable $c(\delta, \kappa) = v(\delta)\kappa$ and $v(\delta)$ is linear in δ .

Since the unobserved talents are ordinal, it is WLOG to assume that the loan cost function is linear in bank holding cost κ , which should be interpreted as the effective holding cost across rankings. This holding cost, in theory, can depend on varied banks' characteristics. For example, if the holding cost depends on the deposit-to-asset ratio, the effective holding cost is thus given by $\kappa(d)$. The function form of $\kappa(d)$ does not matter as our approach allows us to cover the relevant holding cost κ directly.

Then, using the equilibrium condition in the sorting model, one can infer a bank's per unit risk holding cost by using the information from equilibrium interest rates and firms' default probabilities. Let $L_t[i] \equiv y_t - U_t[i]$ denote the loan payoff of the bank that lends to firm i . Such a payoff is observable and can be measured by $L_t[i] = (1 - \delta_t[i])D_t^*[i] +$

$\delta_t[i]y_{t,L}[i]$, where $D_t^*[i]$ is the repayment and $y_{t,L}[i]$ is the recover value. Equation (2) can thus be rewritten as

$$\frac{L'_t[i]}{\delta'_t[i]} = c_\delta(\delta_t[i], \kappa_t[j^*(i)]), \quad (7)$$

which depends on the matching bank's holding cost $\kappa_t[j^*(i)]$, conditional on a firm's characteristic $\delta_t[i]$.

Baseline: Linear in Firm Default Risk. When $v(\delta)$ is linear in δ with $v'(\delta) = a_0$, then Equation (7) is reduced to

$$\frac{L'_t[i]}{\delta'_t[i]} = a_0\kappa_t[j^*(i)] = a_0\kappa_t([N - (i^* - i)]). \quad (8)$$

One can thus infer the matching bank's holding cost $\kappa_t[j^*(i)]$ by looking at the ratio of the change in loan payoff to the change in default probability of firm i . Intuitively, when the default probability of a riskier firm increases more relative to the interest rate the firm paid, it must be because the matching bank can absorb the risk better (i.e., a relative low $\kappa_t[j^*(i)]$).

Recall that i^* is the type of the marginal firm obtaining credit. So $i^* - i$ close to zero corresponds to risky firms close to the cutoff, while large values of $i^* - i$ corresponds to safer firms. The sorting implies that the most risky firm (i.e., $(i^* - i) \rightarrow 0$) is always matched with the best bank, who has the lowest cost $\kappa[N]$; the safest firm (i.e., $(i^* - i) \rightarrow i^*$), on the other hand, is matched to the worst bank conditioned on that bank being active with $\kappa[N - i^*]$. That is, $\kappa[N - (i^* - i)]$ is increasing in $i^* - i$ when there is heterogeneity in bank ability or holding cost. Only when banks are homogeneous, we would then expect the ratio is constant across firm risk rankings.

Alternative: Convexity in Firm Default Risk. In our empirical work, we will use the baseline linear specification. More generally, the cost function can be non-linear

in firms' default probability. In this case, the right hand side depends on the bank's characteristic $\kappa[j]$ as well as firm's characteristic $\delta[i]$. Naturally, if the cost is given by the variance of the project, then $v(\delta)$ will be convex in the firms' default probability.¹¹ Such an effect will increase $\frac{L'_t[i]}{\delta'_t[i]}$ and thus our approach would *underestimate* bank's ability of holding risks.¹² In other words, using a linear specification in empirical work will give a conservative estimate for bank holding cost heterogeneity, which is our preference in this paper. Nevertheless, since default probability is observable, one can still back out $\kappa[j]$ given any parametric assumption on $v(\delta)$.

4.3. Additional Implementation Issues

Note that, to calculate the counterfactual $i^*(\delta_0[i], y_0, \kappa_1[i])$ that solves Equation (6), one would need to know (1) the NPV of projects y_0 and (2) the whole distribution of $\kappa_1[j]$ and $\delta_0[i]$, both of which might not be observable from the data.

The first difficulty, i.e. knowing the NPV of projects, can be resolved by using the fact that the loan supply under the non-crisis period, denoted by $i_0^* \equiv i^*(\delta_0[i], y_0, \kappa_0[i])$, solves

$$y_0 = c(\delta_0[0], \kappa_0[N - i^*]) + \int_0^{i_0^*} c_\delta(\delta_0[i], \kappa_0[N - (i^* - i)]) \delta'_0[i] di,$$

where RHS can be estimated from the data. To address the second difficulty that we only observe the δ 's and κ 's of active banks and firms, we can *extrapolate* these values estimated from the sample and use them to calculate $i^*(\delta_0[i], y_0, \kappa_1[i])$.

¹¹For example, set $y_L = 0$ and given that all projects have the same NPV, the variance of the project is thus increasing and convex in the default probability given by $v(\delta) = (1 - \delta) \left(\frac{y}{1-x} - y \right)^2 + \delta y^2$ with $v''(\delta) > 0$.

¹²The opposite is true when $v(\delta)$ is concave in δ .

5. Great Recession Decline in Aggregate Lending

In this section, we now apply the methods from Section 4 to account for the effect of credit supply on the observed drop in aggregate lending during the Great Recession.

5.1. Estimates of $\delta[i]$ and $\kappa[i]$

We estimate banks' holding cost distribution $\kappa[i]$ for two time periods: a pre-crisis span from 2005-2007 and a crisis span of 2008-2010. Given that we are assuming that the cost function is given by $c(\delta[i], \kappa[j]) = \delta[i]\kappa[j]$, i.e. that Equation 8 holds, $\kappa[j]$ is the per unit cost of bank j handling risk. The more talented the bank, the lower is the bank's per unit cost. We will equate a more talented bank with a bank having a lower per unit cost in this section. A decrease in credit supply is then an outward shift in the per unit cost of holding risk for banks.

In the pre-crisis period, we have roughly 4000 firm loans from Dealscan. In the crisis period, we have roughly 2500 firm loans. (In our model, each firm is a loan so we will use firm and firm loan interchangeably.) We take each firm's Moody's credit rating and use the average historical default rate associated with that rating as our proxy for $\delta[i]$. As our proxy for $L[i]$, we take this to be $(1 - \delta[i])$ times the credit spread over LIBOR for the firm loan. We assume that $y_L[i] = 0$ in this calculation. That is, we reasonably assume that heterogeneity in recovery rate is small across firms relative to the credit spread.

Figure 2: Firm Probability of Default by Credit Rating Rankings

Notes: This figure plots for the pre-crisis (2005-2007) and the crisis (2008-2010) periods historical firm probability of default δ based on Moody's credit ratings from the firm's most recent public debt issuance. The pre-crisis period has roughly 4000 firm loans placed into four bins based on their credit rating rankings. The crisis-period has around 2500 firm loans placed into four bins based on their credit rating rankings. The ranks are all normalized by 4000 the number of firm loans in the pre-crisis period. Observations near zero in the x-axis are the risky firms, while observations near 1 are the safe firms. Dotted lines represent bootstrap standard error bands.

In Figure 2, we plot on the x-axis $i^* - i$, firm ranks based on default probability. We have $i_{2005}^* = 4000$ firms loans in the pre-crisis period. We use 4000 to normalize the x-axis. On the y-axis is the default probability. Firms $i^* - i$ near zero are riskier, while firms with $i^* - i$ closer to one are safer. To address noisiness in the data, we construct four bins for the 2005-2007 with around 1000 observations per bin. We can see that default probabilities are near zero for firms closer to 1 but rise considerably as we get closer to 0. For firms near the cutoff i^* , the default probability is around 0.18.

For the 2008-2010 sample, we also create four bins similar to the pre-crisis period.¹³ There are roughly 600 observations per bin. The default probabilities for the safest firms (i.e. those close to 1) are still close to zero. But default probabilities rise quickly and are higher than pre-crisis levels. This reflects downgrades in credit ratings during this period for a large set of firms. Moreover, there are fewer risky firm loans in the crisis period as witnessed by the red line's observations being concentrated at lower values of i .

In Figure 3, we again plot on the x-axis $i^* - i$, firm ranks based on default probability. On the y-axis is the expected payoff of the loan $L[i]$. We can see that banks charge much higher interest rates during the crisis period compared to the pre-crisis period.

In Figure 4, we plot bank per unit cost $\kappa[j]$ as a function of j/j_{2005}^* . For both the pre-crisis and post-crisis, we see that the curves are upward sloping consistent with heterogeneity in bank holding cost or talent and assortative matching. The banks that make loans to the riskiest firms, i.e. those near 0 in the x-axis, have the lowest cost. The banks that make loans to the safer firms, i.e. those near one in the x-axis, have higher cost. Since fewer loans are made, it also means, in our matching framework, that fewer banks are making loans.

Consider again the blue pre-crisis curve. The blue dot at 0.125 represents the mass of firms (i.e. 12.5% of the 4K firms in 2005) that are being lent to by banks with close

¹³Whether it is 2005-2007 or 2008-2010, we are still normalizing the x-axis by 4000.

Figure 3: Expected Loan Repayment by Firm Credit Rating Rankings

Notes: This figure plots for the pre-crisis (2005-2007) and the crisis (2008-2010) periods the expected loan repayment L by firm credit rating rankings. The pre-crisis period has roughly 4000 firm loans placed into eight bins based on their credit rating rankings. The crisis-period has around 2500 firm loans placed into six bins based on their credit rating rankings. The ranks are normalized by 4000 the number of firm loans in the pre-crisis period. Observations near zero in the x-axis are the risky firms, while observations near 1 are the safe firms. Dotted lines represent bootstrap standard error bands

Figure 4: Bank Holding Cost by Firm Credit Rating Rankings

Notes: This figure plots for the pre-crisis (2005-2007) and the crisis (2008-2010) periods the bank holding cost ($\kappa = \frac{L'}{\delta'}$) by bank rankings. On the axis is ranking of the banks j normalized by the 2005 j^* . Banks closer to 0 on the x-axis are banks making loans to riskiest firms and are more talented at securitization.

to zero per unit holding cost. There is a mass of firms running from 0.125 to 0.375 (i.e. 25% of the 4K firms in 2005) that are being lent to by banks with slightly higher holding costs. These costs are rising slowly over this segment. In other words, 37.5% of firms in 2005 were given loans by banks with low per unit holding costs. Finally the mass from 0.375 onwards then represents the firms that are being matched with progressively higher cost banks. The slope of the blue curve is much steeper over this interval.

Now consider the red crisis curve. The main difference between the red and blue curves is that there is a smaller mass of firms receiving loans from banks with low holding costs, i.e. 0.19 versus 0.375 relative to the number of firm loans in 2005. This key empirical finding — that the mass of the low cost banks shrank considerably in the crisis period — is consistent with the financial stress of investment banks during the Great Recession and disruptions in securitization market. It reassuring that our methodology yields sensible inferences.

Robustness. In the model, the interest rate is a monotonic function of a firm’s probability of default. That is, whether one ranks firms by their probability of defaults or by their interest rates should give exactly the same result. To quantify how close the data is to our stylized model, we report the rank-correlation (Spearman correlation) between a firm’s probability of default and a firm’s interest rate in the period 2005-2007 in Table 2. The stylized model gives a correlation of one. In the data, however, we obtain a correlation around 60%, which reflects the fact that the distributions are close, but not exactly similar.

This result suggests that there are other determinants of the interest rate beyond the probability of default. To handle this, we average the bank-firm matches into five groups determined by the quantiles of the probability of default. After doing this step, there is a perfect rank-correlation between the average probability of default within a group and the average interest rate.

Table 2: Spearman Correlation Probability of Default and Interest Rate in 2005-2007

	Rank Probability of Default
	(1)
Rank Interest Rate	0.60*** (0.01)

5.2. Calculating Counterfactuals

Two group specification. To calculate our counterfactual, we first fit the data in Figure 4 to distributions of banks and firms characterized by (1) a measure of talented banks (safe firms), with holding costs (default probability) that are approximated to be zero, denoted by n_t^b and n_t^f (where $t = 0$ denotes the pre-crisis period and $t = 1$ is the crisis period, respectively, and (2) banks that are lower ranked have positive holding cost and is linear decreasing in j where $\kappa'_t[j] = -h_t^b$ for $j \leq N - n_t^b$. Similarly, default probability of non-safe firms is linear increasing in its ranking (i.e., $\delta'_t[i] = h_t^f > 0$ for $i \geq n_t^f$).

We refer banks with close to zero holding costs as talented banks. These banks, according to Figure 4, are lending to firms that represented by the first two dots. That is, the total measure of talented banks drops from 0.375*4K before crisis to 0.192*4K, which gives $n_0^b = 1.5K$ and $n_1^b = 0.76K$. One can thus interpret that the observed supply shocks as if the very talented banks (the top investment banks) are hit in the sense that their holding costs, which was zero before crisis, go up during the crisis to the level that they're no longer active.

Figure 5 illustrates the sorting outcome in this simple set-up and the essence of our counterfactual. The counterfactual loan amount $i^*(\delta_0[i], y_0, \kappa_1[i])$ is calculated using the estimated banks' holding cost distribution during the crisis period and firms' riskiness during the pre-crisis period. These two distributions are represented by the green and blue line, respectively, in Figure 5.

Figure 5: The blue line represents the riskiness $\delta[i]$ of firm i and the green line represents the holding cost of his matching bank $\kappa[N - (i - i^*)]$.

Specifically, according to Equation (5), given any distribution, the marginal firm is given by

$$\int_{n_t^f}^{i^* - n_t^b} h_t^f \kappa_t[N - (i^* - i)] di = \frac{h_t^f h_t^b}{2} (i^* - n_t^b - n_t^f)^2 = y_t. \quad (9)$$

That is, in this simple example, the grey area under $\kappa_1[i]$ (the green line) represents the aggregate risk absorbing capacity, which is higher for a higher NPV y_t and/or a lower h_t^f (i.e., firms are relative safe). The counterfactual loan amount $i^*(\delta_0[i], y_0, \kappa_1[i])$ thus solves

$$h_1^b (i^* - n_1^b - n_0^f)^2 = y_0 = h_0^b (i_0^* - n_0^b - n_0^f)^2. \quad (10)$$

RHS of the equation represents the aggregate risk absorbing capacity before crisis, where $i_0 = 4000$ is the credit volume before crisis. The difference between $i_0^* - i^*$ thus represents the credit supply effect on the credit volume, which depends on the change in the banks' ability to hold risk, captured by the value of n_t^b and h_t^b .

Since the estimated value of κ is more noisy for the less talented banks, we first consider a more conservative estimate by setting $h_0^b = h_1^b$. In this case, the change in the credit volume is then reduced to the change of measure of talented banks. Hence,

	j	κ	$\frac{\Delta\kappa}{\Delta j}$
2005-07	0.125	0	0
	0.375	0.1671	$h_0^{bm} = 0.668$
	0.625	1.1796	$h_0^b = 4.05$
2008-10	0.064	0	0
	0.192	0.221	$h_1^{bm} = 1.723$
	0.320	1.552	$h_1^b = 10.40$

Table 3: Banks' holding costs

$i_0^* - i^* = n_1^b - n_0^b = 1.5K - 0.76K = 740$. In other words, the credit supply effect is around $\frac{740}{(4000-2500)} = 0.493$

Three group specification. We now further divide the talented banks in the 2-group specification into two sub groups. Only the first subgroup has zero holding costs, while the second subgroup now also has a positive linear slope, denoted by h_t^{bm} , which is smaller than the one of last group $h_t^{bm} < h_t^b$. The size of the first and second subgroup is denoted by n_t^b and n_t^{bm} , respectively. The equivalent of Equation 9 under 3-group specification yields

$$\frac{h_t^{bm} (n_t^{bm})^2}{2} + \frac{h_t^b}{2} \left\{ \left(i^* - (n_t^f + n_t^b + n_t^{bm}) \right)^2 \right\} + h_t^{bm} n_t^{bm} (i^* - (n_t^f + n_t^b + n_t^{bm})) = y_t.$$

The size of these banks and the corresponding slopes can be obtained from Figure 4. The first two columns of Table 3 provides the values of $(j, \kappa_t[j])$. The measure of the second group is given by $n_0^{bm} = 0.375 - 0.125$ and $n_1^{bm} = 0.192 - 0.0638$, respectively. The third column shows the corresponding slope for the last two groups. The measure of safe firms can be obtained from Figure 2, which yields $n_0^f = 0.375$.¹⁴

The estimated credit supply effect is summarized in Table 4. Given that the estimated κ is more noisy for the group of high per unit cost banks, we provide two additional values of $\frac{h_1^b}{h_0^b}$. That is, $\frac{h_1^b}{h_0^b} = 1$ assumes that the slope remains the same for the non-talented banks,

¹⁴We refer that the last two groups of firms as safe firms, as their default probability are close to zero.

ϕ^S	$\frac{h_1^b}{h_0^b} = \frac{10.4}{4.05}$	$\frac{h_1^b}{h_0^b} = 1.5$	$\frac{h_1^b}{h_0^b} = 1$
3-group	0.719	0.604	0.502

Table 4: Credit supply effect under 3-Group Specification

which thus gives the most conservative estimate for the credit supply effect. We find that the credit supply effect ranges from 0.502 to 0.719 across these specifications. That is, a conservative estimate of the credit supply effect from our two-group and three-group specifications is around 50%.

6. Relating to Firm-Bank Panel Regression Estimates

Finally, we connect our model to the reduced form approach in the bank lending channel literature which regresses firm loans on the health of the banks they are borrowing from. To explicitly map our results to the firm fixed-effect estimator from the literature ([Khwaja and Mian \(2008\)](#)), we propose an extended environment where each firm needs two dollars, and each bank can lend 1 dollar to a firm.

In this simple set-up, a firm can obtain two dollars of loans (one dollar from each bank) or zero loan otherwise if and only if $i \leq i_t^*$. Consider the same credit supply shock as in our counterfactual exercise in Section 5.2 (two group specification) where some talented banks are hit and thus the measure of talented banks is now lower $n_1^b < n_0^b$.

Given that the bank holding costs in the economy increase, the marginal firm must decrease to $i_1^* < i_0^*$. The model predicts that the riskier firms with ranking $i \in [i_1^*, i_0^*]$ will be driven out of the market, independent of whether the firm was borrowing from a hit bank or not. For these firms, the estimated firm fixed-effect estimator is zero; that is, $\beta_{out}^{FE} = 0$.¹⁵

¹⁵Specifically, for those firms, the change in the loan between firm i and bank j yields $\Delta L_{i,j} = -1 \forall j \in \{H, N\}$, where $j = H$ iff the bank is hit. The estimator solves $(\Delta L_{i,j} - \Delta \bar{L}_i) = \beta^{FE}(e_{i,j} - \bar{e}_i)$, where $e_{i,j} = -1$ iff the lending bank is hit $j \in H$; $\Delta \bar{L}_i$ and $\bar{e}_i$ denote the average change in the loan supply and exposure for firm i . For those firms $\beta^{FE} = 0$, as $\Delta L_{i,j} = -1$ and $\Delta \bar{L}_i = -1$.

On the other hand, firms above the cutoff can secure funding by reallocating to other banks. The estimated fixed effect for firms that are able to refinance is thus higher due to the substitution effect. Specifically, in this set-up, a firm with ranking $i \in [i_0^* - n_0^b, i_1^*]$ loses one dollar if its original bank is hit, but it can refinance through another bank that is not hit; hence, $\beta_{ref}^{FE} = 2$.

Fixed Effect Estimator Does Not Capture Aggregate Credit Supply Effect.

The firm fixed-effect estimator in the literature can be understood as the average of β_{out}^{FE} and β_{rf}^{FE} ,

$$\beta^{FE}(\mu) = (1 - \mu)\beta_{ref}^{FE} + \mu\beta_{out}^{FE},$$

where μ denotes the weight on firms that are driven out of the market, conditional on firms having one hit bank.¹⁶

A larger drop in the credit supply thus means a higher measure of firms that would be driven out of the market, which implies a higher weight on β_{out}^{FE} and a lower β^{FE} . Indeed, consider an extreme case where all low cost banks that are hit are replaced by new banks that enter the economy. Since there is effectively no change in the aggregate distribution of risk capacity, the aggregate supply effect is zero. However, the fixed effect estimator is highest, as it purely captures the refinance/substitution effects.

The Effect of Sticky Relationships. We next evaluate the robustness of attributions in the presence of a form of bank-firm sticky relationship. When their banks are hit, we assume that firms can only rematch with probability $\lambda \in [0, 1]$. With probability $1 - \lambda$, these firms lose access to refinancing. In general, we refer to hit banks as the ones having extremely high holding costs and thus no longer active. This specification captures that $1 - \lambda$ of these firms lose their loans whenever their borrowing banks are hit. In other

¹⁶In this simple example, $\mu = \frac{n_0^b - n_1^b}{n_0^b}$ is the percentage of talented or low cost banks that are hit.

words, some safer firms that could have refinanced are now exogenously removed out of the market due to the sticky relationship.¹⁷

This notion of sticky relationship effectively makes surviving firms riskier, relative to the perfect reallocation or matching setting. Nevertheless, conditional on firms that can reallocate or rematch, the equilibrium characterization and thus our methodology remains the same. In particular, given that firms' riskiness is observable, our estimate of the bank holding cost distribution remains the same. In other words, the effect of sticky relationship will show up as a credit risk effect in our estimate.

The sticky relationship, however, will increase the firm fixed-effect panel regression estimate. As before, the sensitivity of firm loan to bank health is heterogeneous across three groups of firms and can now be ranked in a descending order: (i) firms that are able to refinance through other banks to maintain the same amount of lending (ii) firms that lose one dollar when their banks are hit due to sticky relationships (iii) firms that lose 2 dollars and are driven out of the market because of competition. That is, $\beta_{ref}^{FE} > \beta_0^{FE} > \beta_{out}^{FE}$.

The average fixed effect estimator is then:

$$\beta^{FE}(\lambda, \mu) = (1 - \eta(\mu, \lambda))(\lambda\beta_{ref}^{FE} + (1 - \lambda)\beta_0^{FE}) + \eta(\mu, \lambda)\beta_{out}^{FE},$$

where $\eta(\mu, \lambda)$ denotes the measure of firms that lose 2 dollars whether or not their banks are hit. As before, the measure $\eta(\mu, \lambda)$ increases with the drop in the aggregate credit supply μ . Moreover, it also increases with the reallocation probability λ , with $\eta(\mu, 1) = \mu$. This comes from the fact that, when relationships are more sticky (lower λ), more safe firms are removed from the market. As a result, riskier firms can survive as there is less

¹⁷For simplicity, we further assume that all the shocks here are unexpected. One could extend the model for expected shocks. In the case that the types and shocks are i.i.d, the matching outcome remains the same. More generally, the specification of shocks and persistence of types however could then affect the matching outcomes, as what matters is the expected surplus of the relationship.

competition. Hence, for a given credit supply shock μ , more sticky relationships lead to a higher β^{FE} .

No-Sorting Heuristic. The bank lending channel literature loosely uses a no-sorting heuristic based on the consistency of panel regressions estimates (OLS and firm fixed effects), which also applies during the Great Recession (Chodorow-Reich (2013)).¹⁸ However, in our framework, consistency of OLS and firm fixed effects estimates are not informative about sorting, as our empirical findings and analysis undoubtedly indicate. In the Appendix C, we provide additional reasons outside our model for why general equilibrium effects complicate conclusions regarding sorting based on panel regression estimates.

7. Conclusion

We show in this paper that the bank loans for risky firms are more likely to be syndicated by banks with a greater propensity to securitize. This sorting affects causal inference of bank credit supply effects for firm output which is not easily addressed with standard reduced-form approaches for identification. We show that this sorting arises naturally in a competitive matching model of a credit market where banks with low holding costs (e.g. investment banks better at selling securitized loans) lend to risky firms. We then use our model to assess the importance of credit supply (heterogeneity in bank holding costs) from credit risk (heterogeneity in firm default rates) effects. We implement our approach using available data on loan interest rates and historical default rates from credit ratings.

¹⁸For instance, Khwaja and Mian (2008) write “Thus the difference between the OLS estimate $\hat{\beta}^{OLS}$ and the FE estimate $\hat{\beta}^{FE}$ provide a direct test of how (...) the bank liquidity shock is correlated with (...) the firm productivity shock.”. Similarly, Chodorow-Reich (2013) write “Under certain assumptions, the difference in the point estimates between regressions including and excluding the fixed effects captures the amount of bias induced by not-as-good-as random matching of borrowers and lenders. Specifically, the true model of bank lending must be additively separable over bank health and firm characteristics.”. See also Jiménez, Ongena, Peydró, and Saurina (2017), Jiménez, Mian, Peydró, and Saurina (2019), Schnabl (2012) for similar reasonings.

References

- Amiti, M., and D. E. Weinstein, 2018, “How much do idiosyncratic bank shocks affect investment? Evidence from matched bank-firm loan data,” *Journal of Political Economy*, 126(2), 525–587.
- Berger, A. N., N. H. Miller, M. A. Petersen, R. G. Rajan, and J. C. Stein, 2005, “Does function follow organizational form? Evidence from the lending practices of large and small banks,” *Journal of Financial Economics*, 76(2), 237–269.
- Bernanke, B., and A. Blinder, 1988, “Credit, money and aggregate demand,” *American Economic Review*, 78(2), 435–439.
- Bernanke, B. S., and M. Gertler, 1989, “Agency costs, collateral, and business fluctuations,” *American Economic Review*, 79(1), 14–31.
- Chang, B., and H. Hong, 2019, “Selection versus talent effects on firm value,” *Journal of Financial Economics*, 133(3), 751–763.
- Chava, S., and M. R. Roberts, 2008, “How does financing impact investment? The role of debt covenants,” *Journal of Finance*, 63(5), 2085–2121.
- Chen, J., and K. Song, 2013, “Two-sided matching in the loan market,” *International Journal of Industrial Organization*, 31(2), 145–152.
- Chiappori, P.-A., and B. Salanié, 2016, “The econometrics of matching models,” *Journal of Economic Literature*, 54(3), 832–61.
- Chodorow-Reich, G., 2013, “The employment effects of credit market disruptions: Firm-level evidence from the 2008–9 financial crisis,” *Quarterly Journal of Economics*, 129(1), 1–59.
- Engelberg, J., P. Gao, and C. A. Parsons, 2012, “Friends with money,” *Journal of Financial Economics*, 103(1), 169–188.
- Gabaix, X., and A. Landier, 2008, “Why has CEO pay increased so much?,” *Quarterly Journal of Economics*, 123, 49–100.
- Gennaioli, N., A. Shleifer, and R. Vishny, 2012, “Neglected risks, financial innovation, and financial fragility,” *Journal of Financial Economics*, 104(3), 452–468.
- Gorton, G., and A. Metrick, 2012, “Securitized banking and the run on repo,” *Journal of Financial Economics*, 104(3), 425–451.
- Greenwood, R., and S. G. Hanson, 2013, “Issuer quality and corporate bond returns,” *Review of Financial Studies*, 26(6), 1483–1525.
- Herreno, J., 2019, “The Aggregate Effect of Bank Lending Cuts,” working paper, Columbia University mimeo.
- Ivashina, V., and D. Scharfstein, 2010, “Bank lending during the financial crisis of 2008,” *Journal of Financial Economics*, 97(3), 319–338.
- Jiménez, G., A. Mian, J.-L. Peydró, and J. Saurina, 2019, “The real effects of the bank lending channel,” *Journal of Monetary Economics*.

- Jiménez, G., S. Ongena, J.-L. Peydró, and J. Saurina, 2017, “Macroprudential policy, countercyclical bank capital buffers, and credit supply: evidence from the Spanish dynamic provisioning experiments,” *Journal of Political Economy*, 125(6), 2126–2177.
- Justiniano, A., G. E. Primiceri, and A. Tambalotti, 2019, “Credit supply and the housing boom,” *Journal of Political Economy*, 127(3), 1317–1350.
- Kashyap, A., J. Stein, and D. Wilcox, 1993, “Monetary policy and credit conditions: Evidence from the composition of external finance,” *American Economic Review*, 83(1), 78–98.
- Khwaja, A. I., and A. Mian, 2008, “Tracing the impact of bank liquidity shocks: Evidence from an emerging market,” *American Economic Review*, 98(4), 1413–42.
- López-Salido, D., J. C. Stein, and E. Zakrajšek, 2017, “Credit-market sentiment and the business cycle,” *Quarterly Journal of Economics*, 132(3), 1373–1426.
- Martinez-Miera, D., and R. Repullo, 2017, “Search for yield,” *Econometrica*, 85(2), 351–378.
- Nadauld, T. D., and M. S. Weisbach, 2012, “Did securitization affect the cost of corporate debt?,” *Journal of Financial Economics*, 105(2), 332–352.
- Paravisini, D., V. Rappoport, P. Schnabl, and D. Wolfenzon, 2014, “Dissecting the effect of credit supply on trade: Evidence from matched credit-export data,” *Review of Economic Studies*, 82(1), 333–359.
- Petersen, M. A., and R. G. Rajan, 2002, “Does distance still matter? The information revolution in small business lending,” *Journal of Finance*, 57(6), 2533–2570.
- Philippon, T., 2009, “The bond market’s q,” *Quarterly Journal of Economics*, 124(3), 1011–1056.
- Schnabl, P., 2012, “The international transmission of bank liquidity shocks: Evidence from an emerging market,” *Journal of Finance*, 67(3), 897–932.
- Schwert, M., 2018, “Bank capital and lending relationships,” *Journal of Finance*, 73(2), 787–830.
- Shivdasani, A., and Y. Wang, 2011, “Did structured credit fuel the LBO boom?,” *Journal of Finance*, 66(4), 1291–1328.
- Sørensen, M., 2007, “How smart is smart money? A two-sided matching model of venture capital,” *The Journal of Finance*, 62(6), 2725–2762.
- Terviö, M., 2008, “The difference that CEOs make: An assignment model approach,” *American Economic Review*, 98, 642–668.

Appendices

A. Downside Risk and Economic Performances

In this Appendix section, we show that our three firm downside risk measures are negatively related to firm performance during the financial crisis—that is our measures of downside risk are correlated with cyclicalities or sensitivity to economic downturns. To show this, we run regressions of the form

$$\Delta CAPX_i = \alpha + \beta X_i + \epsilon_i,$$

where CAPX from Compustat measures the growth of investment between 2006 and 2009, and X_i denotes one of the three borrower characteristics. We run these specifications with and without fixed effects.

Table [A1](#) reports the results. In column (1), we find that borrower loan spread from the Dealscan universe is negatively correlated with borrower CAPX growth from 2006 to 2009. A one percentage point (p.p) increase in loan spread (or in bond spread for public firms) is associated with a 5% decrease in the CAPX growth. In column (2), we add bank fixed effects. That is for the firms in the portfolio of the same bank, we find that borrower loan spread predicts a lower CAPX growth over the financial crisis period. Interestingly, the coefficients of interest are virtually identical in columns (1) and (2) and both have t -statistics around 3.

In columns (3) and (4), our independent variable of interest is borrow bond spread from their public debt. Notice that the number of observations is around 500 in columns (3) and (4) compared to columns (1) and (2). As we mentioned earlier, the reason for difference in sample sizes across the two measures is that not all firms issue public debt, which is a prerequisite for having our second measure. Nonetheless, this subsample is interesting

since it provides a different measure of borrower downside risk. We get virtually identical point estimates in column (3) without bank fixed effects and in column (4) with bank fixed effects as in columns (1) and (2) when we had the entire Dealscan sample. In column (3), the coefficient of interest is -0.049 with a t -statistic of -2.8. In column (4), we have a coefficient of -0.048 with a t -statistic of -2.6 when we introduce bank fixed effects. Not only are these two coefficient similar with each other, they are also similar to those obtained in columns (1) and (2).

In columns (5) and (6), we find that a one standard deviation increase in market leverage is associated with a 1.3% decrease in CAPX growth when there are no bank fixed effects and 1.8% decrease in CAPX growth when there are bank fixed effects. The sample here is almost similar to Dealscan absent data lost in the merger between Dealscan and Compustat. The t -statistics in columns (5) and (6) are comparable to those in the earlier columns. Hence, regardless of our measure of borrower downside risk, we get robustly similar results that risky borrowers are more procyclical than less risky ones.

The findings in Table A1 can to a large extent be anticipated from findings in earlier papers. Risky firms should be more procyclical than less risky firms as in a Q-theory set-up proposed in Philippon (2009). Moreover, Greenwood and Hanson (2013) find that aggregate corporate debt issuance is positively correlated with aggregate credit risk appetite. They find that credit risk appetite was extremely high before the financial crisis and fell substantially during the financial crisis. While these findings might to a large extent be anticipated, they are nonetheless important for the macro-finance literature on credit supply effects to the extent borrower risk characteristics are correlated with whether or not a bank is hit during the financial crisis.

Table A1: Characteristics Related to Downside Risk Predict CAPX Growth during the Financial Crisis

	Borrower CAPX Growth 06-09					
	(1)	(2)	(3)	(4)	(5)	(6)
Borrower Loan Spread	-.05*** (-3)	-.05*** (-2.7)				
Borrower Bond Spread			-.049*** (-2.8)	-.048*** (-2.6)		
Borrower Leverage					-.13** (-2.3)	-.18*** (-3.1)
Bank FE	No	Yes	No	Yes	No	Yes
R^2	.0087	.04	.024	.098	.01	.042
N	1913	1912	599	592	1709	1708

Notes: This table estimates the model

$$\Delta CAPX_i = \alpha + \beta X_i + \epsilon_i,$$

where i denotes a firm, X_i is alternatively its loan spread in columns (1) and (2), bond spread in columns (3) and (4), and market leverage in columns (5) and (6). The specification is estimated using bank fixed effects in columns (2), (4), and (6). Estimation with OLS. t -statistics in parenthesis.

B. Extensions

B.1. Many-to-Many Matching

We now explain how our baseline one-to-one matching model can be extended to the environment with multiple relationships.

Assume each bank j has $m(j)$ measure of managers, where each manager can lend 1 dollar to a firm. Each manager makes their loan decisions separately but share the same bank-specific risk capacity $\kappa[j]$. A firm i , on other hand, has $l(i)$ projects and each needs one dollar. Each project gives the same independent payoff as our baseline model, with the firm-specific default probability $\delta[i]$.

The baseline model can then be reinterpreted as matching between each manager and each project, where the only difference is the underlying distribution of κ and δ takes into account the capacity distribution $m(j)$ and $l(i)$. Specifically, the measure of projects that has default probability lower than $\delta[i]$ is then given by $\int_0^i l(\tilde{i}) d\tilde{i}$ and the measure of managers with holding cost lower than $\kappa[j]$ is then given by $\int_0^j m(\tilde{j}) d\tilde{j}$.

For instance, consider the distribution used for the counterfactual exercise and suppose that $\kappa_t[j] = 0$ for any $j \geq j_t^*$, with $t \in \{0, 1\}$. One can interpret that there are effectively $n_t^b = \int_{j_t^*}^N m_t(j) dj$ measures of talented managers, who thus lend to firms with relative risky projects with ranking from $i \in [\hat{i}_t, i_t^*]$, where the market clearing implies that $\int_{\hat{i}_t}^{i_t^*} l_t(i) di = \int_{j_t^*}^N m_t(j) dj$.

B.2. Loan Size

In our baseline model, we normalize the loan size to be one and are interested in whether a firm will obtain the loan (the extensive margin). We now extend our environment to allow agents to optimally choose the loan size (the intensive margin).

To do so, we assume that there is diminishing return in the investment. The payoff for each project in each state $s \in \{H, L\}$ is given by $R_s^i(q) = qy_s^i - \frac{\gamma}{2}q^2$, where $\gamma > 0$. Thus, conditional on the investment q , the expected payoff the project yields

$$Y(q) = q \left(\underbrace{\delta^i y_L^i + (1 - \delta^i) y_H^i}_{\equiv y} - (1 + r_f) \right) - \frac{\gamma}{2} q^2.$$

Same as in the baseline model, we assume that, conditional on the investment, (1) the expected value of the project is the same, while projects differ in their riskness; (2) the holding cost depends on the riskness of the project as well as banks' ability.

Additionally, we assume that the holding cost is linear increasing in the loan size. Given any match (i, j) , bank and firm then choose the optimal loan to maximize the joint surplus, which yields

$$s(i, j) = \max_q Y(q) - qC(\delta[i], \kappa[j]).$$

The optimal loan size between (i, j) is given by the FOC condition

$$Y'(q^*(i, j)) = C(\delta[i], \kappa[j]).$$

One can thus see that, conditional on banks' ability, loan size increases with firm's riskiness. In equilibrium, however, the relationship depends on the sorting outcome.

Lemma 4. *A more talented bank is matched with a riskier firm, if*

$$C_{12}(i, j) < \left(\frac{C_1(i, j)C_2(i, j)}{y - C(\delta, \kappa)} \right) < 0.$$

To understand this, the complementarity in the surplus functions with intensive margins, is given by

$$s_{12}(i, j) = -q^*(i, j)C_{12}(i, j) - q_j^*(i, j)C_i(i, j).$$

where the first term is the same as before which is positive as long as $C_{12} < 0$. The second term, however, captures the fact that the better bank can give out more loan and it is more costly to have more loan to the risky bank, which is captured by $q_j^*(i, j) > 0$ and $C_i(i, j) > 0$.

Lemma above thus implies that, compared to the baseline model, a higher degree of complementarity in the cost function is needed in order to achieve the same sorting outcome. If $C_{12} \simeq 0$, the sorting outcome reverses (as $s_{12} < 0$).

Given that we observe a more talented bank is matched with a riskier firm empirically, we thus assume that such condition holds. The equilibrium can then be characterized similarly as before, where the equivalent of Equation (2) becomes

$$U'[i] = s_i(i, j^*(i)) = -q^*(i, j^*(i))C_1(i, j^*(i)).$$

The loan size for firm i , which is given by $Q[i] \equiv q^*(i, j^*(i))$. Note that, because of the sorting, the loan size can be non-monotonic in firms' default probability, as the riskier firm is hold by the more talented bank.

Let $L[i]$ denote the total loan payoff to the bank, where $L[i] = Q[i] \{(1 - \delta[i])D^*[i] + \delta[i]y_L[i]\}$, which now takes into account the loan size. Given that the payoff of firm is given by $U[i] = Y(q^*(i, j(i))) - L[i]$, the equivalence of Equation (7) then becomes

$$\frac{1}{Q[i]} \frac{L'[i]}{\delta'[i]} = \frac{1}{\delta'[i]} \left(\frac{y - \gamma Q[i]}{Q[i]} \right) \frac{dQ[i]}{di} + C_i(i, j). \quad (\text{B.11})$$

Thus, Equation (7) is the special case of Equation (B.11) when $Q'[i] = 0$. In case that $Q[i]$ does not change much across ranking, which is also what we observed empirically, our results thus remain robust.

C. No-Sorting Heuristic

In this section, we discuss a sorting heuristic often used in the literature, which was introduced in Khwaja and Mian (2008). Researchers often interpret the difference between the OLS and FE estimate of a regression on lending growth on bank change in health as a measure of sorting. However, we now argue that there can be sorting even when the OLS and FE estimates are similar..

To clarify the discussion, it is useful to remember that the OLS estimate can always be written as the weighted average of the AGG estimate (i.e. the estimate in a regression of the average lending growth on the average change in bank health) and the FE estimate. Even in absence of sorting, we can expect the AGG and the FE estimate to be different. This is because the AGG estimate measures an income effect (faced with a change in average lending cost, by how much does a firm's demand for loan decrease?), while the

FE estimate measures a substitution effect (faced with a relative differential in lending cost between two banks, by how much does a firm reallocate between two banks?). For instance, in our model, the FE estimate would be infinite (since firms can freely reallocate between banks) while the AGG estimate could be zero, or even negative, depending on the structure of shocks hitting the economy.

This general argument even holds in the exact framework explored by [Khwaja and Mian \(2008\)](#). Consider a firm j that borrows from multiple bank $i \in \mathbb{I}_j$, where $\mathbb{I}_j$ is exogeneous. The firm combines loans from multiple banks to produce an output Y_j with a CES production function:

$$Y_j \equiv A_j \bar{L}_j^{1-\frac{1}{\alpha}} \quad (\text{C.12})$$

where A_j is the productivity and $\bar{L}_j = \left(\sum_{i \in \mathbb{I}_j} L_{ij}^{\frac{\chi-1}{\chi}} \right)^{\frac{\chi}{\chi-1}}$ is the total lending. The parameter $\chi > 0$ is the elasticity of substitution between loans of different banks. The higher χ , the easiest it is to substitute loans across different banks. The parameter α controls the return to scale of the firm.¹⁹

The firm takes the interest rate of the banks r_i as given, and chooses its bank loans $(L_{ij})_{i \in \mathbb{I}_j}$ to maximize total profit — revenue minus costs of repaying loans. Firm opti-

¹⁹Alternatively, we could assume that production has linear return to scale and that firms face an isoelastic demand with elasticity α

$$Y_j = B_j p^{-\alpha}$$

where B_j is a demand shifter. The revenue function becomes $B_j^{\frac{1}{\alpha}} A_j^{1-\frac{1}{\alpha}} \bar{L}_j^{1-\frac{1}{\alpha}}$. Denoting $\tilde{A}_j = B_j^{\frac{1}{\alpha}} A_j^{1-\frac{1}{\alpha}}$ the “business condition”, which is a combination of firm specific productivity and demand shifter, we can write the revenue function as $\tilde{A}_j \bar{L}_j^{1-\frac{1}{\alpha}}$.

mization gives:

$$\ln \bar{L}_j = -\alpha \ln \bar{r}_j + \alpha \ln A_j - \alpha \ln \left(1 - \frac{1}{\alpha}\right) \quad (\text{C.13})$$

$$\ln L_{ij} = -\chi (\ln r_i - \ln \bar{r}_j) + \ln \bar{L}_j \quad (\text{C.14})$$

where $\bar{r}_j = \left(\sum_{i \in \mathbb{I}_j} r_i^{1-\chi}\right)^{\frac{1}{1-\chi}}$ is the average interest rate proposed to the firm j .

After a change in interest rates, the change in lending from bank i to firm j can be written as:

$$\Delta \ln L_{ij} = \underbrace{-\chi \Delta \ln r_i}_{\text{Bank Component}} + \underbrace{\alpha \Delta \ln A_j + (\chi - \alpha) \Delta \ln \bar{r}_j}_{\text{Firm Component}} \quad (\text{C.15})$$

In particular the lending growth from bank i to firm j is additively separable into a bank component and a firm component. This equation is consistent with the reduced form equations used in the empirical banking literature. The firm component reflects two pieces, only one of which has been recognized in the literature. The first is the change in TFP or risk piece denoted by $\alpha \Delta \ln A_j$, which has been well-studied by researchers. However, it also depends on the change in the firm cost of capital which has been ignored in the literature (i.e. the $\Delta \ln \bar{r}_j$ term). Formally, the difference between the OLS and the

FE estimate in a regression of lending growth on bank interest rate is:²⁰

$$\beta^{\text{OLS}} = \beta^{\text{FE}} + \underbrace{\alpha \frac{\text{Cov}(\Delta \ln A_j, \Delta \ln r_i)}{\text{Var}(\Delta \ln r_i)}}_{\text{sorting term}} + \underbrace{(\chi - \alpha) \frac{\text{Cov}(-\Delta \ln \bar{r}_j, \Delta \ln r_i)}{\text{Var}(\Delta \ln r_i)}}_{\text{cross elasticity term}}. \quad (\text{C.16})$$

It is the sum of two distinct terms. As stressed by the existing literature, there is a *sorting term*, depends on the covariance between a bank liquidity shock and a firm productivity shock. It measures the extent to which the matching between firms and banks is non-random. The term is positive if banks with higher health are matched with firms with higher productivity growth. This term appears in empirical banking literature.

However, there is also a *cross-elasticity term*. Because we think of loans as easily substitutable (at least among the set of banks from which the firm borrows from), we expect $\chi > \alpha$, and, therefore, that this term is negative.²¹ In conclusion, in the framework of [Khwaja and Mian \(2008\)](#), the difference between the FE and OLS estimate measures sorting only when $\chi = \alpha$. This is a strong assumption: in words, it assumes that the interest rate charged by a bank does not effect the firm's demand from other banks (i.e. the income effect exactly compensates the substitution effect).

²⁰There are two additional terms in this derivation that roughly cancel out and which we omit for simplicity:

$$\begin{aligned} \beta^{\text{OLS}} = & \beta^{\text{FE}} \underbrace{\alpha \frac{\text{Cov}(\Delta \ln A_j, \Delta \ln r_i)}{\text{Var}(\Delta \ln r_i)}}_{\text{sorting term}} + \underbrace{-\mu \frac{\text{Cov}(-\Delta \ln \bar{r}_j, \Delta \ln r_i)}{\text{Var}(\Delta \ln r_i)}}_{\text{cross elasticity term}} + \\ & + \frac{\text{Cov}(-\chi \Delta \ln r_i, \delta_i)}{\text{Var}(\delta_i)} - \frac{\text{Cov}(-\chi(\Delta \ln r_i - \Delta \ln \bar{r}_j), \delta_i - \bar{\delta}_j)}{\text{Var}(\delta_i - \bar{\delta}_j)} \end{aligned}$$

²¹The presence of this cross elasticity term is related to Marshall-Hicks formula, which relates the price-elasticity of input i to the elasticity of substitution between different factors:

$$\frac{\partial \ln L_{ij}}{\partial \ln r_i} = \chi + (\alpha - \chi)s_i$$

where s_i denotes the cost share of input i .

D. Omitted Proofs

D.1. Proofs for Lemmas 1 and 2

From Equation (1), by monotone comparative statics, if $C_{12}(i, j) < 0$, then $i^*(j)$ increases with j . That is, a higher j (more talented bank) will choose a higher i (more riskier firms). Or, equivalently, $j^*(i)$ increases with i . Hence, given any marginal type i^* , the market-clearing condition thus imply that $j^*(i) = N - (i^* - i)$, where the most risky firm is matched with the best bank $j^*(i^*) = N$.

Lastly, given any equilibrium payoff $U(i)$ and matching outcome $j^*(i)$, it must satisfy Equation (2), which is the FOC from banks' optimization problem.

D.2. Proof for Lemma 3

Let $F(i^*) \equiv c(\delta[0], \kappa[N - i^*]) + \int_0^{i^*} c_\delta(\delta[i], \kappa[j^*(i)])\delta'[i]di$, where $\frac{dF(i^*)}{di^*} > 0$. From Equation (4), we thus have higher y , higher i^* .

We define banks talent are more scarce under $\kappa_1[j]$, where $\kappa'_1[j] \leq \kappa'_2[j] \leq 0$ and $\kappa_1[N] = \kappa_2[N]$. This implies that $\kappa_1[j] \geq \kappa_2[j]$. Let i^* be the loan supply under $\kappa_2[j]$, we thus have

$$\begin{aligned} \int_0^{i^*} c_\delta(\delta[i], \kappa_1[j^*(i)])\delta'[i]di &> \int_0^{i^*} c_\delta(\delta[i], \kappa_2[j^*(i)])\delta'[i]di \\ &= y - c(\delta[0], \kappa_2[N - i^*]) > y - c(\delta[0], \kappa_1[N - i^*]). \end{aligned}$$

hence the loan supply i' under $\kappa_1[j]$ must be smaller than i^* . Same logic can be derived when firms become riskier in the sense that $\delta_1[i] \geq \delta_2[i]$. This thus increases RHS of Equation (5), which thus implies that the loan supply must be lower.