Intermediary-Based Loan Pricing*

Pierre Mabille† Olivier Wang‡
INSEAD NYU Stern

June 2022

How do shocks to banks transmit to the price and non-price terms of loans? We build a model of multidimensional contracting between heterogeneous risky borrowers and intermediaries with limited lending capacity and show that loan terms depend on two moments: the elasticities of loan demand and default rates to interest rates. These two sufficient statistics predict how the cross-section of loan terms and banks' portfolio risk react to changes in lenders' capital, regulation, and funding costs. Using empirical estimates, we show that they explain the heterogeneous transmission of shocks across loan markets and borrower risk categories. Accounting for non-price terms is important because they endogenously increase the persistence of credit crises on loan markets where they adjust more than interest rates.

JEL classification: G12, G21, G28, G33, G51.

Keywords: bank loans, credit supply, credit risk, covenants, financial crisis, intermediary asset pricing, sufficient statistics.

*We thank Viral Acharya, Simon Gilchrist, Arpit Gupta, John Kuong, Virgiliu Midrigan, Alexi Savov, Philipp Schnabl, Alp Simsek, Venky Venkateswaran, and Laurent Calvet, Sebastian Doerr, Fabrice Tourre, and Laurent Weill (discussants) for helpful comments. This paper benefited from seminar participants at NYU Stern, INSEAD, the Federal Reserve Board, the Junior European Finance Seminar, Banque de France, BI Oslo, and from participants at the CEBRA Annual Meeting, the 37th GdRE Symposium on Money Banking and Finance, the SGF Conference, and the INSEAD Finance Symposium. We thank the Alfred P. Sloan Foundation and the Macro Financial Modeling (MFM) project for supporting this research.

†Email: pierre.mabille@insead.edu. Website: https://www.pierremabille.com.

‡Email: olivier.wang@nyu.edu. Website: https://www.olivierwang.com.

1 Introduction

The terms of loans offered by banks determine borrowers' access to credit and banks' exposure to credit risk, making them crucial for the economy and financial stability. They make up rich contracts which typically consist of prices (interest rates) and non-price terms such as quantity limits (e.g., collateral) and other restrictions (e.g., covenants, maturity), whose variations across borrowers and over time have attracted widespread attention. Recently, non-price terms have come under renewed scrutiny for households during the housing cycle (e.g., Acharya et al. 2020) and for firms during the Covid-19 recession (e.g., Chodorow-Reich et al. 2021).

However, it is still unclear how loan terms are set in the cross-section of borrowers and how they are affected by the financial health of the very lenders which offer them. Especially intriguing is that, while credit supply generally goes hand in hand with bank health (e.g., Chodorow-Reich and Falato 2021), credit contractions and expansions affect various borrowers differently and through different loan terms. The credit boom which preceded the Great Recession forcefully illustrates it: thriving U.S. banks expanded credit to risky borrowers in the mortgage market (Mian and Sufi 2009), but to safe borrowers in the credit card market (Agarwal et al. 2018). The mortgage expansion hinged on both relaxed leverage constraints and lower rates, but the increase in credit card debt was mostly driven by heightened credit limits. Despite the outsized role of credit conditions in macro-finance (e.g., Stiglitz and Weiss 1981, Holmstrom and Tirole 1997, Bernanke et al. 1999), what is missing is a model that can jointly explain how shocks to banks are transmitted to the price and non-price terms of loans faced by different borrowers.

We fill this gap in an equilibrium model of multidimensional loan contracting between heterogeneous risky borrowers and financial intermediaries with limited lending capacity. Banks' capacity constraints and costs of funds are subject to shocks which arise from changes in asset prices, financial regulation, and monetary policy. We show that two sufficient statistics, the interest rate elasticities of borrowers' loan demand and default probability, can explain key features of credit markets: (i) Banks jointly use interest rates and non-price terms to control the supply of credit. (ii) The transmission of bank shocks to loan terms varies across markets (e.g., mortgages vs. credit cards), borrowers within a market (e.g., safe vs. risky), and bank health (e.g., well vs. poorly capitalized). (iii) The responses of the different loan terms govern the persistence of credit contractions and expansions. We emphasize the role of banks' balance sheets for credit supply in specialized asset markets as in intermediary asset pricing (He and Krishnamurthy 2013, Brunnermeier and Sannikov 2014). But we highlight a specificity of credit markets: banks use non-price terms to control borrowers' default risk when it is endogenous due to information frictions.

We motivate our analysis by documenting heterogeneity in loan terms and the transmission of bank shocks for the main classes of loans. First, banks lend less when they make more losses on their assets. Over time, total loan growth is strongly negatively correlated with bank charge-off rates across mortgage, credit card, and commercial and industrial (C&I) loans. Second, banks jointly set interest rates, quantity limits, and other non-price terms to control default risk and credit supply. Changes in bank heath are associated with changes in loan-to-value (LTV) limits for mortgages, credit standards for credit cards and C&I loans, and interest rate spreads for all three markets. However, interest rates are stickier for credit cards and therefore changes in credit supply mostly happen through quantity limits. Third, the transmission of bank shocks to different borrowers is heterogeneous across markets. Credit supply shocks are mostly transmitted to risky borrowers on mortgage markets, but to safe borrowers on credit card markets, as well as C&I loans for which non-price terms such as covenants also adjust. Despite interest in these markets taken separately, there is a limited understanding of what explains their differences and commonalities.

We build and solve a model of multidimensional loan contracting. We make three contributions. Our first contribution is to jointly endogenize the price and non-price terms of loan contracts. While these features are well-known in the data, they have been studied separately in theoretical settings so far: either in economies with endogenous credit rationing but fixed loan rates (e.g., Stiglitz and Weiss 1981); or in Walrasian economies where rates adjust to clear loan markets but borrowing constraints are exogenous (e.g., Holmstrom and Tirole 1997).[1] In our model, realistic price and non-price terms of loans are the optimal outcomes of a contracting problem between heterogeneous borrowers and banks. Banks compete for borrowers subject to capacity constraints on lending, which depend on their balance sheets. Two features distinguish loan markets from generic asset markets (such as stock markets). First, asset payoffs are endogenous to asset prices. Interest rates affect default probabilities and losses given default because of microeconomic frictions which induce borrowers to default when the loan repayment value is too high. Our model accommodates special cases with moral hazard, adverse selection, and a "double trigger" default motive. The elasticity of borrowers' repayment probability $\epsilon_{1-\mu}$ to interest rates captures this channel. Second, we depart from a Walrasian framework where the loan repayment value is linear in a single interest rate clearing credit markets. Instead, banks can offer multidimensional non-linear contracts with quantity limits and other non-price terms which capture lender screening and monitoring effort. Non-price terms arise from the feedback between loan prices and loan payoffs for banks. Quantity limits and screening or monitoring allow banks to manage borrowers' default risk while holding rates fixed. The interest rate on a given loan compensates banks for their cost of funds (risk-free rate), the borrower credit risk (credit risk premium), and their capacity constraint tightness (excess loan premium). This problem gener-

[1] In many macro-finance models, borrowing constraints are key for households, firms, and asset prices but their variations are also usually assumed to be exogenous (e.g., Bernanke et al. (1999), Jermann and Quadrini (2012), Favilukis et al. (2017), Guerrieri and Lorenzoni (2017)).

ates a multidimensional contract curve, which relates both borrower and lender characteristics to loan terms in a similar way to reduced-form credit surfaces which focus on borrowers (e.g., Geanakoplos 2010). The contract curve provides a comprehensive measure of credit tightness based on all the terms of loan contracts. It is described by its interest rate elasticity ϵ_ℓ, which captures how demand changes in response to rates when borrowers are endogenously constrained by multiple loan terms. In turn, ϵ_ℓ depends on two objects. The first is the unconstrained interest rate elasticity of loan demand ϵ_ℓ^u when borrowers borrow as much as they want at the going interest rate, which depends on preferences and technology. The second is the elasticity of borrowers' repayment probability $\epsilon_{1-\mu}$, which determines whether banks adjust interest rates or non-price terms to control borrowers' default probability and depends on preferences, technology, and microeconomic frictions.

We use our framework to understand how shocks to banks' lending capacity and cost of funds affect the price and non-price terms of loan in equilibrium. When banks' balance sheets deteriorate, they trade off tightening the various terms. Our second contribution is a set of pass-through formulas which highlight the role of the interest rate elasticities of borrowers' loan demand and repayment probabilities, ϵ_ℓ and $\epsilon_{1-\mu}$, as two sufficient statistics which determine how bank shocks are transmitted to the cross-section of loan terms. Our approach has the benefit of relying on two objects which are identified in borrower- and loan-level data.

We begin by studying the impact of credit supply shocks, modeled as changes in banks' lending capacity. Our results explain why credit contractions and expansions affect safe and risky borrowers differently and through various loan terms across markets (e.g., credit cards vs. mortgages). The higher the elasticity of borrowers' loan demands ϵ_ℓ on a given market, the more their loan sizes vary in response to a given interest rate change. Positive credit supply shocks lower interest rates across markets, but markets with more elastic borrowers see larger increases in sizes and smaller decreases in rates. As a result, loan repayment values increase in markets with more elastic borrowers, whereas they decrease in markets with inelastic borrowers due to the dominating fall in quantities. The repayment elasticity $\epsilon_{1-\mu}$ determines how credit risk reacts to these changes. In more elastic markets, the increase in loan repayment values generates a small increase in credit risk for safe borrowers, but a large increase for risky ones. To maximize their total profits, banks adjust loan terms to equate the returns from lending to safe and risky borrowers – if it were not the case, they could increase their total profits by lending more to borrowers with a higher return. This leads them to lending more to safe borrowers on markets with more elastic borrowers (e.g., credit cards), and more to risky borrowers on markets with less elastic borrowers (e.g., mortgages).[2]

[2]Estimates for the interest rate elasticity of mortgage debt tend to be lower than one and range between 0.07 and 0.5 (e.g., Best et al. 2019, Fuster and Zafar 2021, Benetton 2021), with the exception of DeFusco and Paciorek (2017).

We next study how shocks to banks' cost of funds which arise from monetary policy shocks transmit to the cross-section of loan terms. Our results explain why the transmission of monetary policy is weakened when bank balance sheets are impaired, and why the strength of the bank lending channel is heterogeneous across loan markets. The pass-through of the policy rate to loan rates and sizes depends on the elasticity ϵ_ℓ and on banks' lending capacity. Unconstrained banks transmit changes in the policy rate more than one-for-one to low-elasticity borrowers (e.g., mortgages), but this transmission is dampened for higher-elasticity ones, resulting in stickier rates (e.g., credit cards). When banks are capacity-constrained but lending capacity is not sensitive to the policy rate, the transmission of monetary policy is further dampened. This feature tends to insulate borrowers from the negative effects on credit supply of a policy rate hike, but also from the positive effects of a rate cut.

We explore the implications for the credit risk of the entire portfolio of bank loans. A monetary policy shock has two effects. First, a change in both price and non-price terms within a given pool of borrowers; second, a reallocation of bank loans towards specific borrowers. We find that policy rate cuts tend to increase the credit risk of borrowers with a high loan demand elasticity ϵ_ℓ and to decrease it for others. If high-elasticity borrowers are also riskier (high $\epsilon_{1-\mu}$), then policy rate cuts increase credit risk for the entire portfolio of bank loans. We also show how a low policy rate environment gives rise to covenant-lite lending for low elasticity borrowers but to tighter covenants for high elasticity ones, a novel finding in line with the data.

We conclude by showing that the two elasticities ϵ_ℓ and $\epsilon_{1-\mu}$ also drive the dynamics of banking and credit supply crises. We build a dynamic version of the model where bank shocks and borrower credit risk affect loan spreads which feed back into bank capital, which affects the tightness of lending capacity constraints over time. Our third contribution is to show that the adjustment of non-price terms in response to bank shocks increase the persistence of financial crises – a finding which speaks to the post-2008 experience. Negative bank shocks result in higher spreads and tighter non-price terms. The more they are tightened, the less spreads need to increase to control credit risk. The lesser increase in spreads in turn lowers excess returns earned by constrained banks. It slows down their recapitalization and makes the credit crunch more persistent.

We illustrate the quantitative implications of our results in a version of the model calibrated to the U.S. mortgage market. The model matches empirical estimates for ϵ_ℓ and $\epsilon_{1-\mu}$. In response to a contraction in bank lending capacity, the excess loan premium on mortgages increases, reflecting the tightness of banks' constraints. Loan sizes fall and mortgage spreads increase for all households, lowering LTV ratios. As in the data, the responses of loan terms differ across

In contrast, estimates for the elasticity of credit card debt tend to be greater than one (e.g., Gross and Souleles 2002 estimate an elasticity of 1.3).

borrowers. Loan size falls twice more for borrowers with high ϵ_ℓ and $\epsilon_{1-\mu}$ while their mortgage spread increases by less. As a result, credit risk falls sharply for high-elasticity borrowers and increases for low-elasticity ones. Two complementary policies partly mitigate the credit crunch. First, direct household debt relief reduces default risk by making them effectively richer, and allows relatively more borrowing during the credit crunch, mitigating the decrease in LTV ratios. Second, a recapitalization of banks relaxes their lending capacity constraint, and achieves similar results by increasing the total volume of credit available.

Related literature Our results contribute to the microeconomic literature on credit markets and to the macro-finance literature on the dynamics of credit crises.

First, our model helps explain the transmission mechanism from losses in banks' portfolios to reductions in credit documented in empirical analyses (e.g., Peek and Rosengren 1997, Murfin 2012, Chodorow-Reich 2014, Huber 2018, Greenstone et al. 2020). Reductions in new credit take the form of tighter lending standards and higher loan spreads. We endogenize this mechanism and study its implications. Our model explains why banks use non-price terms to control the volume of credit for a given rate, such as debt covenants for firms (Chodorow-Reich and Falato 2021) and credit card limits for households (Agarwal et al. 2018). It generates heterogeneous responses of loan terms in the cross-section of borrowers and credit markets, as in the data when banks' balance sheets change (Khwaja and Mian 2008, Chakraborty et al. 2018, Ivashina et al. 2020) and when their cost of fund changes because of monetary policy (Jimenez et al. 2012, Chakraborty et al. 2020) or access to external finance (Paravisini 2008, Ivashina and Scharfstein 2010). We trace these differences back to two sufficient statistics with well-identified empirical estimates: the interest rate elasticities of borrowers' loan demand and repayment probability. Estimates for these elasticities come from survey data (Fuster and Zafar 2021), regression discontinuity analyses (Best et al. 2019, Fuster and Willen 2017), and structural models (Buchak et al. 2020, Benetton 2021, Robles-Garcia 2020). We show how they depend on structural parameters in our model which affect borrowers' loan demand and default. Their variation across markets may arise from moral hazard, adverse selection, or borrowers' liquidity constraints (Adams, Einav and Levin 2009, Einav, Jenkins and Levin 2012).

Second, we add to canonical macro-finance models by introducing multidimensional credit contracts with both price and non-price terms. We build on the theoretical literature on credit rationing (Stiglitz and Weiss 1981, Jaffee and Russell 1976) surveyed in Jaffee and Stiglitz (1990), and focus on rationing at the intensive margin: risky borrowers face tighter non-price terms rather than being excluded from credit markets. Our analysis complements models of the credit surface (e.g., Geanakoplos 2010) where an increase in borrower credit risk in bad times leads more pessimistic lenders to requiring more collateral, and heterogeneous household models where the price of unsecured consumer loans depends on borrower characteristics (e.g., Chatterjee et al. 2007, Livshits et al. 2007). We show how lenders' financial conditions affect the terms of lending, as in recent work by Diamond and Landvoigt (forthcoming) who focus on the credit surface in a rich quantitative model of the mortgage market. Our contribution is more theoretical: we analytically isolate key elasticities which predict how credit surfaces react to various shocks. This general formulation allows us to compare different kinds of loan markets, in particular mortgages, credit cards and corporate loans.

We extend the intermediary asset pricing framework (He and Krishnamurthy 2013, Brunnermeier and Sannikov 2014, Gromb and Vayanos 2018) to credit markets by modeling the joint effect of banks' capacity constraints and borrower default risk on interest rates, quantity limits, and other non-price terms of loans. Endogenizing these terms generates the excess bond premium found in the data (Gilchrist and Zakrajsek 2012). It also generates a new transmission mechanism of shocks as non-price adjustments affect the dynamics of credit crises. As in canonical macro-finance models (e.g., Gertler and Kiyotaki 2010, Rampini and Viswanathan 2019), credit crises occur when banks' net worth falls. However, bank losses need not generate a sharp increase in spreads which would quickly recapitalize banks and make the credit crunch short-lived. Instead, the endogenous tightening of non-price terms by banks results in spreads increasing by less, which slows down banks' recapitalization and makes the credit crisis more persistent. This finding helps explain the persistence of the post-2008 U.S. mortgage crisis (e.g., Justiniano et al. 2019).

The rest of the paper is organized as follows. Section 2 documents stylized facts on heterogeneity in loan terms. Section 3 describes our model of multidimensional loan contracting and the credit market equilibrium. Section 4 studies the transmission of credit supply shocks and monetary policy shocks to the cross-section of loan terms. Section 5 analyzes the dynamics of credit crises, which depends on the interaction between the cross-section of loan terms and banks' balance sheets. We illustrate our findings in a version of the model calibrated to the U.S. mortgage market, where we compare the effectiveness of debt relief and bank recapitalization policies. Section 6 concludes.

2 Evidence on Heterogeneity in Loan Terms

This section documents motivating facts on heterogeneity in loan terms and the transmission of shocks for the main classes of bank loans: mortgages, credit cards, and C&I loans. The goal of the model in the next section is to explain these facts and understand their implications for the persistence of banking and credit crises.

Data sources. We combine data on bank losses and the volume and characteristics of mortage, credit card, and commercial and industrial (C&I) loans.

Data on charge-off rates comes from the Consolidated Reports of Condition and Income of the Federal Financial Institutions Examination Council at the U.S. Federal Reserve Board.

For mortgages, data on loan sizes, loan rates, and maximum LTV and DTI ratios comes from Fannie Mae and Freddie Mac. For credit cards and C&I loans, data on the net percentage of banks tightening lending standards and spreads comes from the Senior Loan Officer Opinion Survey (SLOOS) of the U.S. Federal Reserve Board.

Bank health and loan growth. First, banks lend less when they make more losses on their assets. The upper panel of Appendix Figure 12 plots year-to-year percentage changes in the quantity of mortgage (solid line, left axis) and the associated charge-off rates by banks (dashed line, right axis). There is a strongly negative correlation at the aggregate level between loan growth and bank charge-off rates. As the upper panels of Appendix Figures 13 and 14 show, this relationship holds for credit card and C&I loans too.

Interest rates and non-price terms. Second, banks simultaneously set interest rates, quantity limits, and other non-price terms to control default risk and credit supply. This is shown in the lower panel of Appendix Figure 12, which plots the maximum LTV ratios on mortgages issued by banks (blue line, left axis), reflecting the looseness of borrowing constraints, and the associated interest rate spread over the Federal Funds Rate (red line, right axis). The lower panels of Appendix Figures 13 and 14 respectively plot the net percentages of banks tightening credit standards on credit card and C&I loans (blue line, left axis), which reflects the tightness of borrowing constraints, and the associated interest rate spreads (red line, right axis). Comparing Appendix Figure 12 and Appendix Figure 13 suggests that interest rates are stickier for credit cards than for mortgages, and therefore that credit card expansions and contractions mostly occur through changes in credit limits.

Heterogeneous transmission of bank shocks. Third, the transmission of bank shocks to different borrowers is heterogeneous across markets. Credit supply shocks are mostly transmitted to risky borrowers on mortgage markets, and to safe borrowers on credit card markets. The upper left panel of Appendix Figure 15 plots the average loan sizes in dollars of high-FICO score, safe mortgage borrowers (blue line) and low-FICO score, risky mortgage borrowers (red line). The upper right panel plots the associated loan rates, the lower left panel the associated maximum DTI ratios, and the lower right panel the associated maximum LTV ratios. Mortgage sizes grew for both risky and safe borrowers in the 2000s, but then fell more deeply and persistently for risky borrowers despite similar interest rate changes. This reflects the relaxation and subsequent tightening of mortgage standards on subprime borrowers around the Great Recession (Mian and Sufi 2009, Justiniano et al. 2019).

In contrast, credit card expansions and contractions are mostly transmitted to safe borrowers. Appendix Figure 16 shows it by plotting changes in borrower credit card limits by FICO score bins in response to a decrease in banks' cost of funds (from Agarwal et al. 2018). Risky households with a higher marginal propensity to borrow face smaller changes in credit limits than safe households.

Corporate loan sizes are also more volatile for safe borrowers, despite similar changes in interest rates across risk categories. Appendix Figure 17 shows it by plotting the average sizes in dollars of C&I loans (blue line, left axis) and the corresponding interest rates (red line, right axis), for safe firms in the left panel and risky firms in the right panel. This reflects the recent evidence that most of the adjustment in the volume of corporate loans occurs through non-price terms. In particular, covenants are stricter after banks suffer defaults (Murfin, 2012), while abundant liquidity spurs covenant-lite lending (Diamond, Hu and Rajan, 2020).

3 A Model of Multidimensional Loan Contracting

This section describes a general model of multidimensional loan contracting with financial frictions affecting lenders. We add two key features to the intermediary asset pricing framework. First, asset prices in credit markets, i.e., interest rates, affect default probabilities and losses given default. Therefore asset payoffs are endogenous to asset prices. Second, endogenous credit risk makes it optimal for banks to impose quantity limits and other restricting non-price terms on borrowers' loans, in addition to adjusting rates. Banks offer non-linear contracts in the interest rate, and the model departs from the standard Walrasian setting where only rates adjust.

3.1 Environment

We consider a unit continuum of identical lenders indexed by b, "banks", and a unit mass of heterogeneous borrowers indexed by i.

Banks. Banks have funding cost R^f. The expected profit on a loan contract to borrower i, which consists of an interest rate R^i (price term), a loan amount l^i (quantity limit), and a vector of non-price terms $\boldsymbol{z}^i = \left(z_k^i\right)_k$, is

$$\pi^i\left(R^i, l^i, \boldsymbol{z}^i\right) = \left(R^i - R^f\right) l^i - \lambda^i\left(R^i l^i, \boldsymbol{z}^i\right) - c(\boldsymbol{z}^i) l. \tag{1}$$

λ^i is the bank's expected loss on its loan to borrower i. Absent default, bank profits would be $\left(R^i - R^f - c(z^i)\right) l^i$. It reflects the channels through which endogenous default risk affect the bank's expected profit, such as liquidity default, adverse selection, or debt overhang. In most cases, R and l only affect the expected loss through the face value of the debt to be repaid Rl. $c(z^i)$ is the cost of control for lenders due to monitoring and screening efforts when borrowers can default. $c(\cdot)$ is an increasing and convex function of the non-price terms $z^i \geq 0$ such that $c(0) = 0$. It captures other non-price terms such as covenants and loan document requirements which allow the bank to reduce default risk and potentially improve the recovery value in case of default, but which are costly to implement.

Our results derive from the fact that we can always write an isomorphic model with zero recovery rate and an "effective default probability" μ defined as

Definition 1. $\mu^i\left(Rl, z\right) = \frac{\lambda^i(Rl,z)}{Rl}$ *is the effective default probability.*

Expected profit can then be written as $\pi = \left[R\left(1-\mu^i\right) - R^f - c(z^i)\right] l$. With a positive recovery rate, μ will be higher than the actual default probability. We work with μ as a primitive, and define the elasticity of the repayment rate, one of the two sufficient statistics from which our results derive.

Definition 2. *The elasticity of the loan repayment probability* $1-\mu$ *to its face value* Rl *is*

$$\epsilon_{1-\mu}\left(Rl, z\right) = \frac{Rl\mu'\left(Rl, z\right)}{1-\mu\left(Rl, z\right)}$$

We make the following assumption.

Assumption 1. *The elasticity of the repayment probability satisfies* $0 \leq \epsilon_{1-\mu} < 1$ *everywhere.*

Assumption 1 ensures that the zero profit curve $\pi\left(R, l, z\right) = 0$, or more generally iso-return curves (defined as constant $\frac{\pi(R,l,z)}{l}$), are always upward sloping, since they have a slope $\frac{dR/R}{dl/l} = \frac{\epsilon_{1-\mu}}{1-\epsilon_{1-\mu}}$. We rule out standard credit rationing of the type studied by Williamson (1987), which takes place when borrowers' loan demand curve is always above lenders' backward-bending supply curve. Since its implications are well-known, we focus on settings where borrowers have access to credit but at a smaller scale than they would like, because banks set binding non-price terms to limit their credit risk.

To analyze the transmission of credit supply shocks, we study banks with a capacity constraint on total lending

$$\int \rho^i l^i di \leq \overline{L} \tag{2}$$

where l^i is the dollar amount lent to borrower i and $\rho^i \in [0, 1]$ is a risk weight which measures how much balance sheet space a loan to borrower i requires. The constraint penalizes borrowers with high ρ. Heterogeneity in ρ can arise from regulatory risk weights. It can also arise from banks' ability to securitize a particular type of loan and take it off their balance sheets. For instance, conforming mortgages have a low weight ρ and non-conforming mortgages have a high weight ρ, which can further depend on liquidity in the private label securitization market.

Banks' lending capacity $\overline{L}$ can arise from regulatory constraints (e.g., the Basel regulation) and market-based constraints imposed by bank creditors due to informational issues such as those affecting the bank-borrower relationship. A large literature (e.g., Holmstrom and Tirole 1997, Gertler and Kiyotaki 2010) provides microfoundations for the moral hazard or limited commitment problems that lead banks themselves to be credit constrained. We focus on how credit supply shocks to banks' lending capacity are transmitted to different borrowers through the multiple terms of loans.

Special cases. The reduced-form model nests rich environments with ex-ante or ex-post asymmetric information which makes the expected loss endogenous to loan terms through the default probability, the loss given default, or both. It has the benefit that our results do not depend on the underlying microfoundation which determines the dependence of λ on Rl. Before focusing on the general case, we detail special cases of the model to illustrate how various microeconomic frictions map to λ: liquidity default, collateralized loans, and adverse selection.

Liquidity default. A higher repayment value Rl makes it harder to repay, hence $\lambda = \mu\left(Rl, z\right) \times Rl$ where $\mu\left(\cdot\right)$ is increasing. For instance, borrowers have stochastic income y at the date of repayment and default if and only if $Rl \geq y$, so $\mu\left(Rl\right) = \mathbf{P}\left(y \leq Rl\right)$. More generally, with a standard debt contract the bank recovers $\rho\left(y, z\right)$ in case of default, which happens when y falls below a threshold $\hat{y}\left(Rl, z\right)$. In that case,

$$\underbrace{\mu\left(Rl, z\right)}_{\text{effective default prob.}} = \underbrace{F\left(\hat{y}\left(Rl, z\right)\right)}_{\text{actual default prob.}} \left(1 - \mathbf{E}\left[\frac{\rho\left(y, z\right)}{Rl} | y \leq \hat{y}\left(Rl, z\right)\right]\right)$$

Collateralized loans. We consider a simple example of the mortgage market. At date $t = 0$, households can borrow l from the bank and buy a house of price P_0, contributing a downpayment d such that $P_0 = d + l$. Then they consume

$$c_0 = y_0 - d = y_0 - P_0 + l$$

At date $t = 1$, income y_1 and house price P_1 are realized. Households have utility

$$\begin{cases} u\left(y_1 - Rl + \chi P_1\right) & \text{if they repay the mortgage} \\ u\left(\kappa y_1 + \underline{c}\right) & \text{if they default} \end{cases}$$

$(1-\kappa)\, y_1$ captures the disutility of renting as well as the costs of exclusion from financial markets. χP_1 captures the pecuniary value of owning. Hence households default if and only if

$$z_1 \equiv y_1 + \frac{\chi}{1-\kappa} P_1 \leq \frac{\underline{c} + Rl}{1-\kappa}$$

We can define their $t = 0$ value function as

$$V(l, R) = u(y_0 - P_0 + l) + \beta \left[\int_{z_1 \leq \frac{\underline{c}+Rl}{1-\kappa}} u(\kappa y_1 + \underline{c})\, dF(y_1, P_1) + \int_{z_1 > \frac{\underline{c}+Rl}{1-\kappa}} u(y_1 - Rl + \chi P_1)\, dF(y_1, P_1) \right]$$

If upon default the bank recovers ζP_1, then

$$\mu(Rl) = \int \int_{z_1 \leq \frac{\underline{c}+Rl}{1-\kappa}} \left(1 - \frac{\zeta P_1}{Rl}\right) dF(y_1, P_1)\,.$$

Adverse selection. When borrowers' types are not observable, a higher repayment Rl attracts a worse distribution of borrowers. It endogenously increases the risk of default, with $\lambda = \mu(Rl, z) \times Rl$, where $\mu(\cdot)$ is increasing. For instance, suppose borrowers have an unobservable propensity to default u_i and an unobservable component in the utility from loans v_i. Holding rates fixed, if $\mathbf{Cov}(u_i, v_i) > 0$, then borrowers who demand a higher loan size are more likely to default.

Borrowers. Borrowers are characterized by their indirect utility over loan contracts $V^i(l, R, z)$. We make the following standard assumption:

Assumption 2. *For each i, $V^i_R < 0$, and the marginal utility of additional borrowing is lower at higher interest rates: $V^i_{lR} < 0$.*

Assumption 2 holds in most settings, such as those described earlier. The second part of the assumption implies that the unconstrained loan demand curve, defined as the solution l to $V^i_l(l, R, z) = 0$, is decreasing in the interest rate R.

The second sufficient statistics from which our results derive is the elasticity of borrowers' unconstrained loan demand to the interest rate.

Definition 3. *The elasticity of the unconstrained loan demand to the interest rate R is*

$$\epsilon^u_\ell = -\frac{R}{l}\frac{dl}{dR}\Big|_{V_l=0}$$

Finally, tighter non-price terms $z^i \geq 0$ leading borrowers to renouncing control is costly for them, i.e. $V^i_z \leq 0$.

3.2 Equilibrium: Bertrand-Nash with Capacity Constraints

Banks are perfectly competitive and subject to capacity constraints on lending. For each borrower i, they post contracts $C^i = \left(R^i, l^i, z^i\right)$ with commitment, which consist of an interest rate, a loan size which limits the quantity borrowed at the interest rate, and a vector of non-price terms which can further limit borrowing. Contracts are exclusive, so borrowers cannot borrow from multiple banks. They optimally choose which bank they apply to. If rejected, we assume that they can reapply for a loan at the same and at other banks.

A bank with lending capacity $\bar{L}$ choose contracts C^i and the probabilities x^i to grant loans to borrowers i to solve

$$\begin{aligned} \max_{\{x^i,R^i,l^i,z^i\}} & \int x^i \pi^i \left(l^i, R^i, z^i\right) di \\ \text{s.t.} & \int x^i \rho^i l^i di \leq \bar{L} \\ & V^i \left(l^i, R^i, z^i\right) \geq \overline{V}^i \end{aligned}$$

l^i captures the intensive margin of credit, and x^i captures the extensive margin. Banks are subject to a participation constraint from borrowers and $\pi^i \left(l^i, R^i, z^i\right)$ is the profit per loan conditional on take-up for borrower i.

Definition 4. *An equilibrium is an optimal strategy $i \mapsto \left\{x^i_b, C^i_b\right\}$ for each bank b such that borrowers optimize:*

$$\overline{V}^i = \max_{b'} V^i \left(C^i_{b'}\right)$$

and markets clear:

$$1 = \int_{\arg\max_{b'} V^i\left(C^i_{b'}\right)} x^i_b db.$$

We focus on equilibria with symmetric banks such that all banks offer the same contract and choose $x^i = 1$ for all borrowers.

Proposition 1. *In a symmetric equilibrium,*

i. l^i, R^i, and z^i satisfy for each i

$$\tau^i \left(l^i, R^i, z^i\right) = \frac{\epsilon^i_{1-\mu} \left(R^i l^i, z^i\right)}{1 - \epsilon^i_{1-\mu} \left(R^i l^i, z^i\right)} \tag{3}$$

where $\tau^i \left(l, R, z\right) = -\frac{l V^i_l (l,R,z)}{R V^i_R (l,R,z)}$.

ii. Banks make the same profit per risk-weighted dollar for each borrower, i.e.,

$$\frac{\pi^i\left(l^i,R^i,z^i\right)}{\rho^i l^i}=\frac{\pi^j\left(l^j,R^j,z^j\right)}{\rho^j l^j}\equiv\nu \tag{4}$$

where ν is the multiplier on the bank's capacity constraint.

iii. The marginal cost of tightening non-price terms for banks is weakly lower than their marginal benefit, i.e.,

$$\frac{\partial c}{\partial z_k}\left(z^i\right)=-R^i l^i\frac{\partial\mu^i}{\partial z_k}-\frac{V^i_{z_k}}{V^i_R}l^i\left(1-\mu^i\right)\left(1-\epsilon^i_{1-\mu}\right) \tag{5}$$

Proof. See Appendix B.1. □

τ can be interpreted as an intertemporal wedge measuring how constrained borrowers are. If $\tau^i=0$, then borrowers i are on their unconstrained demand curve defined by $V^i_l=0$. If $V^i_l>0$, then borrowers could increase their utility by borrowing more. They are even more borrowing-constrained when the additional cost of borrowing $V^i_R<0$ is low. Totally differentiating τ at $\tau=0$ when borrowers are unconstrained provides an expression for the elasticity of loan demand $\epsilon^u_\ell=\frac{R\tau_R}{\ell\tau_\ell}$.

Part (i) of Proposition 1 highlights that endogenous default risk is the reason why banks use quantity limits and other non-price terms in addition to interest rates to control lending. Banks optimally constrain borrowers depending on the elasticity of their repayment rate to debt $\epsilon_{1-\mu}$. If there is no endogenous default risk and $\epsilon_{1-\mu}=0$, then borrowers are not constrained and the interest rate does not depend on their loan size. The riskier borrowers are when $\epsilon_{1-\mu}>0$, i.e., the more likely they are to default when the face value of their debt increases, then the more banks will restrict the size of their loans for a given interest rate. An excessively high interest would further increase borrowers' default risk and banks' expected losses, therefore banks turn to quantity limits and other non-price terms which restrict borrowing.

Part (ii) describes the optimal capital allocation across different classes of borrowers. Banks use both price (R) and non-price terms (l and z) to equalize the profit per risk-weighted dollar ν across borrowers. If it were not the case, banks could increase their total profits by lending more to borrowers with a higher profit per dollar. When $\nu=0$, the relationship between interest rates, loan sizes, and non-price terms is determined by a zero profit condition for banks. This is because unconstrained banks compete for borrowers up to the point where they make zero profits. When $\nu>0$ and banks' capacity constraints are binding, profit per risk-weighted dollar are higher. Total credit supply $\overline{L}$ is lower and borrowers pay higher rates which lower their total demand accordingly.

Part (iii) describes how banks set non-price terms z. If borrowers have no preferences over z, i.e., $V_z^i = 0$, then banks equalize the marginal cost of non-price terms to their marginal benefit of lowering the effective default probability. This defines an optimal tightness $\hat{z}^i$, where R^i and l^i are endogenously determined by (3) and (4). In general, non-price terms are costly for borrowers due to for instance a loss of control rights ($V_z^i < 0$). This lowers the marginal benefit from tightening z because banks need to keep attracting borrowers, hence banks optimally relax non-price terms to $z^i < \hat{z}^i$.

Unconstrained vs. constrained banks. For each borrower i and taking z^i as given for now, the unconstrained loan quantity $l^{i,*}$ is defined as solving (3) together with the zero-profit conditions $\pi^i\left(l^i, R^i, z^i\right) = 0$. Banks are said to be *unconstrained* if given unconstrained loan quantities $\left\{l^{i^*}\right\}$, their constraint (2) is satisfied, that is

$$\int \rho^i l^{i^*} \leq \overline{L}.$$

Otherwise, banks are *constrained*, and must make a positive profit per dollar $\nu > 0$.

3.3 Implications

We derive three implications of our setting for the effect of endogenous default risk on interest rates and non-price terms, which are key for the transmission of bank shocks. We use a special case of the model to illustrate them and refer to it as our workhorse model. The rest of the paper presents the results for a generic income process that satisfies the main assumptions.

Special case. Borrower income is i.i.d. and follows a Pareto distribution with shape parameter α, where a higher value corresponds to a riskier distribution. The cumulative distribution function of income y is $F_y(y) = 1 - (\frac{y_{min}}{y})^{\alpha}$ if $y_{min} \leq y \leq y_{max}$ and $F_y(y) = 0$ if $y < y_{min}$ or $y > y_{max}$, where $\alpha, y_{min}, y_{max} > 0$. We assume Pareto income risk in all examples and figures for consistency.

Excess loan premium. We decompose interest rates into borrower-level and bank-level factors. A first-order approximation of equilibrium interest rates satisfies

$$\begin{aligned} & \log R_i = \log\left(R^f + \rho^i \nu\right) - \log\left(1 - \mu_i\right) \\ \Leftrightarrow \qquad & r^i \approx r^f + \mu^i + \rho^i \nu \end{aligned}$$

Interest rates paid by borrowers are increasing in banks' time value of money captured by the risk-free rate, individual default risks, and the tightness of banks' capacity constraints weighted by borrowers' importance in the constraint. The latter generates an excess loan premium. If risk weights $\rho^i = 1$ for all i , then interest rates net of the common premium $r^i - \nu$ are actuarially fair and exactly compensate banks for individual default risk. Importantly, this can only be achieved thanks to the endogenous non-price terms l^i, which give banks an additional instrument to control credit risk and thereby offset the common increase in interest rates that arises from credit supply shocks or monetary shocks. Banks optimally tighten l_i by more for riskier borrowers i, as measured by their repayment elasticity $\epsilon^i_{1-\mu}$.

When do banks constrain borrowers? Equation (3) implies that banks only impose a binding borrowing constraint when credit risk is endogenous:

Corollary 1. *Suppose that μ^i is independent of C^i. Then borrower i's allocation can be implemented with a price-posting mechanism where banks only quote an interest rate R^i and borrowers borrow as much as they want given R^i.*

In the case of exogenous default risk where μ is constant, a credit supply shock translates into a higher rate. It is enough for banks to charge a higher rate to compensate for higher default risk and control credit supply. Loans can still have time-varying risk through changes in the effective default probability μ (similar to equity with risky dividends). This is the same case as other asset markets where there is no feedback loop between asset prices and asset payoffs. Menus of loan contracts and non-price terms (e.g., covenants), which are key features of credit markets, only arise from endogenous default risk.

Interest rates do not fully capture credit conditions. With endogenous default risk, the equilibrium interest rate is a non-monotonic function of the borrower's income risk, so that rates alone do not capture credit conditions. Riskier borrowers may be charged a lower interest rate and still be more credit-constrained than safer borrowers, because banks tighten their quantity limits by more. On Figure 1, the black curve depicts the rate R charged to borrowers by banks, the red curve depicts the shadow rate $R(1 + \tau)$, which accounts for the credit rationing wedge τ. The rate R is increasing in risk α at low risk levels. Then it can decrease from the point where an increase in R would decrease bank's expected profits because it would increase the borrower's default probability. At that point, the associated loan size decreases more to compensate for the weaker increase, or even for the decrease in the rate. It translates into a shadow loan rate $R(1+\tau)$ which monotonically increases in borrower's risk. The wedge acts like a tax imposed by banks on borrowers, such that they borrow less for a given rate.

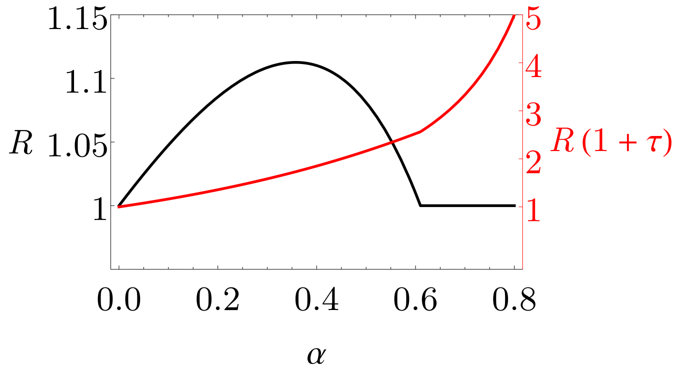

Figure 1: Equilibrium interest rate as a function of borrower risk α, where α is the shape parameter of a Pareto distribution (a higher α corresponds to a riskier distribution). The black curve depicts the contractual rate charged to borrowers by banks. The red curve depicts the shadow rate, which accounts for the rationing wedge.

3.4 Multidimensional Loan Contract Curve ℓ

Banks subject to capacity constraints offer a multidimensional loan contract curve, which is defined by (3). The elasticity of borrowers' loan demand with capacity constraints to the interest rate determines the response of the cross-section of loan terms to bank shocks. The two elasticities that we have defined – for borrowers' repayment probability and their unconstrained loan demand – are sufficient statistics for this term.

Definition 5. *The multidimensional loan contract curve $\ell^i\left(R^i, \boldsymbol{z}^i\right)$ is the solution to (3).*

Definition 6. *The interest rate elasticity of the contract curve is*

$$\epsilon_\ell = -\frac{R}{\ell}\frac{d\ell}{dR},$$

where ℓ satisfies (3) and (4).

The elasticity of the contract curve ℓ^i depends on the unconstrained loan demand elasticity ϵ^u_ℓ, which reflects borrower preferences V^i, and on the repayment elasticity $\epsilon^i_{1-\mu}$, which captures the strength of microeconomic frictions affecting repayment risk. Proposition 2 decomposes the elasticity of ℓ between these two terms and characterizes the effect of endogenous default risk on the contract curve (omitting the i superscript for simplicity).

Proposition 2. *The interest rate elasticity of the contract curve can be decomposed as*

$$\epsilon_\ell = \frac{-R\tau_R + \frac{Rl\epsilon'_{1-\mu}}{\left(1-\epsilon_{1-\mu}\right)^2}}{-l\tau_l + \frac{Rl\epsilon'_{1-\mu}}{\left(1-\epsilon_{1-\mu}\right)^2}} \tag{6}$$

If $\epsilon'_{1-\mu} > 0$ (i.e., the elasticity of repayment rates is higher at higher interest rates), then

$$\epsilon_\ell > \epsilon_\ell^u = \frac{R\tau_R}{l\tau_l} \Leftrightarrow \epsilon_\ell^u < 1$$

Hence if the equilibrium effective default probability μ is constant such that $\epsilon_{1-\mu} = 0$ (exogenous default risk), then $\epsilon_\ell = \epsilon_\ell^u$.

Proof. See Appendix B.1.

□

On loan markets where the elasticity of borrowers' unconstrained loan demand is low ($\epsilon_\ell^u < 1$), such as mortgages, endogenous default risk $\epsilon_{1-\mu}$ makes the contract curve ℓ more elastic than the unconstrained demand ℓ^u, bringing the elasticity ϵ_ℓ closer to one. Hence it generates a larger adjustment in loan sizes than in interest rates in response to credit supply shocks.

The increase in the elasticity of loan demand with constrained banks ϵ_ℓ, hence the loan size adjustment, is larger for riskier borrowers with a high $\epsilon_{1-\mu}$. In response to a small increase $dx = \frac{Rl\epsilon'_{1-\mu}}{\left(1-\epsilon_{1-\mu}\right)^2}$, the elasticity ϵ_ℓ increases by $\frac{d}{dx}\left(\frac{-R\tau_R+x}{-l\tau_l+x}\right) = \frac{R\tau_R-l\tau_l}{(l\tau_l+x)^2}$. Therefore, the quantity limits of risky borrowers vary more on loan markets where the demand for credit is less elastic.

What is the intuition for this result? Suppose that banks did not impose binding quantity limits. Then, for a given reduction in loan size l (e.g., due to a negative credit supply shock or a monetary contraction), the interest rate faced by borrowers on less elastic markets would have to increase by more than one-for-one to induce them to reduce their loan demand. This would result in an increase in the total face value of the loan Rl. Hence it would increase default risk μ and lower bank's expected profits π. Instead, when quantity limits are endogenous, the optimal response of banks is to offer a contract with a lower interest rate but with a binding borrowing constraint. The quantity limit forces borrowers to adjust the loan size demanded, which effectively translates into a more elastic contract curve.

The opposite is true for elastic loan markets such as credit cards ($\epsilon_\ell^u > 1$). Endogenous default risk instead decreases the elasticity of loan demand with constrained banks, which generates a larger adjustment in interest rates than in loan sizes. The contract curve is less elastic for risky borrowers, which brings the elasticity ϵ_ℓ closer to one. Hence a positive term $\frac{Rl\epsilon'_{1-\mu}}{\left(1-\epsilon_{1-\mu}\right)^2}$ due to endogenous default risk $\epsilon_{1-\mu} > 0$ acts as an elasticity dampener which brings back ϵ_ℓ closer to one.

Sufficient statistics for ϵ_ℓ. Proposition 2 highlights that the two elasticities $\epsilon_{1-\mu}$ and ϵ_ℓ^u are sufficient statistics for the elasticity of loan demand ϵ_ℓ when banks are constrained. Even though these two objects are endogenous, it suffices to measure them for each loan market irrespective of their determinants to compute the responses of loan terms to bank shocks in the cross section of borrowers.

The elasticity of borrowers' repayment rates $\epsilon_{1-\mu}$ is a measure of borrower risk as it determines the sensitivity of the bank's recovery value to interest rate changes. Safe borrowers have a low value of $\epsilon_{1-\mu}$ while risky borrowers have a high value (Di Maggio et al. (2017), Fuster and Willen (2017)). The elasticity of borrowers' unconstrained loan demand ϵ_ℓ differs across loan markets. It is lower than one for mortgages and ranges between 0.07 and 0.50 (Best, Cloyne, Ilzetzki and Kleven 2019, Fuster and Zafar (2021), Benetton (2021)). The typical loan size is equal to the difference between a borrower's targeted house size and its down payment, which vary very little with interest rates. It is higher than one and around 1.30 for credit cards, which tend to be used for consumption smoothing instead (Gross and Souleles (2002)).

Figure 2 illustrates the determinants of the contract curve elasticity ϵ_ℓ in our workhorse model. We reproduce key features of the data. First, the contract curve elasticity increases in the elasticity of intertemporal substitution, but it increases less for riskier borrower types (Best, Cloyne, Ilzetzki and Kleven 2019). Second, ϵ_ℓ increases in borrowers' cash-on-hand and income, and it increases by less at higher levels (Buchak, Matvos, Piskorski and Seru 2020).

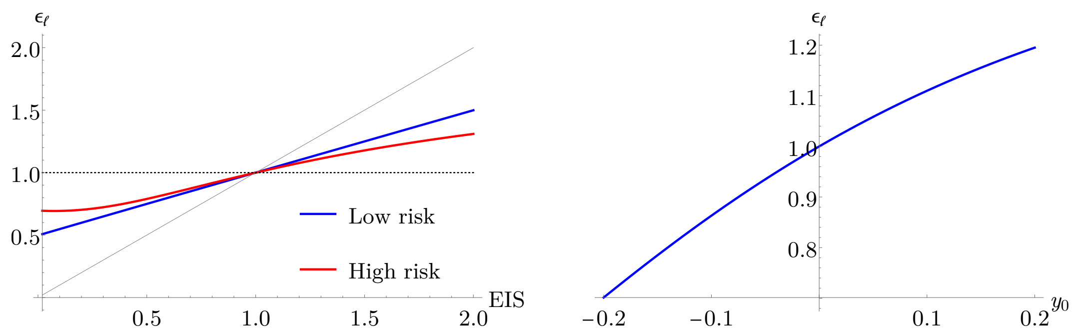

Figure 2: Sensitivity of the interest rate-elasticity of the contract curve curve to borrowers' elasticity of intertemporal substitution and initial endowment y_0.

Effect on default risk μ. The elasticity ϵ_ℓ determines how changes in total credit supply affect the default risk borne by banks. When $\overline{L}$ falls, loan rates increase because of the higher excess loan premium ν associated with a tighter capacity constraints for banks. However, despite the decrease in loan volume, the resulting change in default risk μ is ambiguous. The higher loan rate R leads to more default for fixed loan sizes, but the reduction in loan sizes also reduces the likelihood of default. The balance between these two forces depends on the elasticity ϵ_ℓ as

$$\frac{d}{d\overline{L}}\mu\left(R\left(l\right)l\right) = R\mu'\left(1 - \frac{1}{\epsilon_\ell}\right)$$

Since $\mu' > 0$, a reduction in lending volume due to a negative credit supply shock increases default risk μ if and only if the elasticity ϵ_ℓ is below one. Following our previous discussion, this is the case if and only if the unconstrained loan demand elasticity $\epsilon_\ell^u = \frac{R\tau_R}{l\tau_l}$ is itself lower than one, i.e., on less elastic loan markets.

3.5 Lender-Dependent Credit Surface

Credit surfaces usually map loan sizes and *borrower* characteristics to interest rates (e.g., Geanakoplos 2010, Geanakoplos and Rappoport 2019). They can be estimated with loan- and borrower-level data in multiple settings including household and sovereign debt. The multidimensional loan contract curve ℓ generalizes them to multiple non-price terms (l, z) in addition to interest rates R. Importantly, it highlights how they depend on *lender* health, which is captured by banks' capacity constraint $\overline{L}$ or the excess loan premium ν. The resulting credit surfaces can be characterized by the two sufficient statistics $\epsilon_{1-\mu}$ and ϵ_ℓ^u, instead of relying on nonparametric estimation.

Borrowers are constrained at R if they would like to borrow more at the prevailing interest rate, such that $V_l(\ell(R), R) > 0$. This may be the case even if lenders are unconstrained, with $\ell(R) = L^* < \overline{L}$. Equilibrium condition (3) implies that borrowers are unconstrained only if $\epsilon_{1-\mu}(\ell(R)R) = 0$. If V is separable, $V(l, R) = u(l) - w(Rl)$, then we can interpret the unconstrained condition $V_l = 0$ as a standard Euler equation

$$\frac{u'(l)}{w'(Rl)} = R$$

When borrowers are constrained, the equilibrium with endogenous borrowing constraints—which arises from the optimal contract between banks and borrowers—can also be implemented as a competitive equilibrium in which borrowers choose l subject to a non-linear interest rate schedule $R\left(l|\bar{L}\right)$ (e.g., Livshits, MacGee and Tertilt 2007, Chatterjee, Corbae, Nakajima and Rios-Rull 2007, Diamond and Landvoigt forthcoming). The collection of schedules faced by different borrower types traces out the credit surface. Then, bank lending capacity $\bar{L}$ acts as a credit supply shifter of the surface, which captures the effect of bank health on loam terms.[3]

Given the function $R\left(\cdot|\bar{L}\right)$ associated with her type, a borrower solves

$$\max_l u(l) - w\left(R\left(l|\bar{L}\right) l\right)$$

[3]We only make the dependence on $\bar{L}$ explicit below, but other shocks can also shift the credit surface (e.g., changes in R^f and to the distribution of default risk).

hence

$$\frac{u'(l)}{w'\left(R\left(l|\bar{L}\right)l\right)}=R\left(l|\bar{L}\right)\left[1+\frac{lR'\left(l|\bar{L}\right)}{R\left(l|\bar{L}\right)}\right]$$

Combining with (3), the function $R\left(l|\bar{L}\right)$ solves the differential equation (for each type)

$$\frac{d\log R\left(l|\bar{L}\right)}{d\log l}=\frac{\epsilon_{1-\mu}\left(lR\left(l|\bar{L}\right)\right)}{1-\epsilon_{1-\mu}\left(lR\left(l|\bar{L}\right)\right)} \tag{7}$$

Equivalently, $R\left(l|\bar{L}\right)$ gives rise to the locus $R\left(1-\mu\left(Rl\right)\right)=R^f+\nu\left(\bar{L}\right)$, where $\nu\left(\bar{L}\right)$ is the excess loan premium. Note that the function $l\mapsto R\left(l|\bar{L}\right)$ is conceptually different from the inverse contract curve that we have described earlier, which is the equilibrium outcome as we vary $\bar{L}$ instead. To pin down the exact level of $R\left(l|\bar{L}\right)$, we use as boundary condition the fact that $R\left(l|\bar{L}\right)$ must go through the actual equilibrium contract $\left(\bar{R},\bar{L}\right)$. The interest rate schedules capture the supply side of credit. Borrower preferences V then determine which contract (R,l) is chosen.

Special case. As an illustration, Figure 3 plots the interest rate schedules for two types of households indexed by the Pareto parameter α of their income process. Borrower credit risk is increasing in α, which controls the repayment elasticity $\epsilon_{1-\mu}$. A negative shock to total lending capacity $\bar{L}$ corresponding to an excess loan premium of $\nu=5\%$ shifts the interest rate schedule upwards for both borrowers.

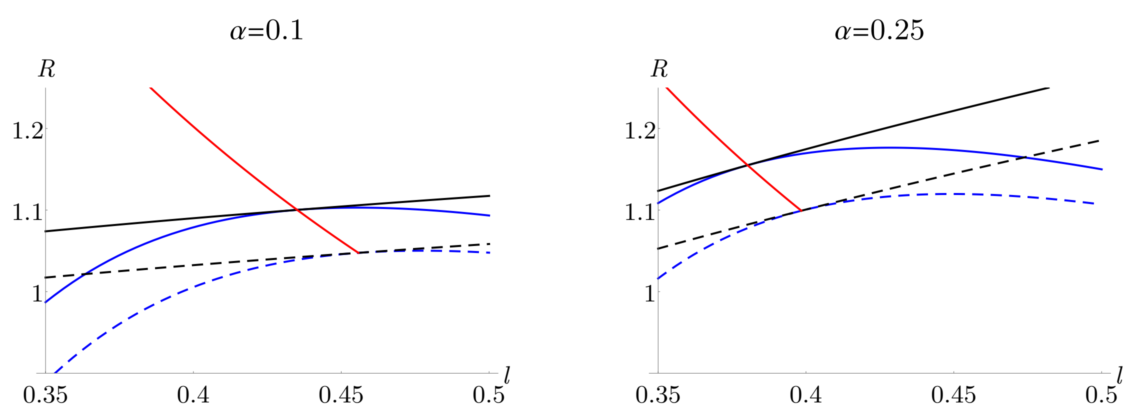

Figure 3: Each borrower α faces an increasing interest rate schedule $R_\alpha(l)$ (black lines). Borrowers choose the point at which their indifference curve (blue lines) is tangent to the rate schedule. The red line is the contract curve $\ell_\alpha(R)$ obtained by varying total lending capacity $\bar{L}$ or equivalently the excess loan premium ν.

In the case of a constant repayment elasticity $\epsilon_{1-\mu}$, we can solve the differential equation (7) in closed form:

$$R(l) = \bar{R} \times \left(\frac{l}{\bar{L}}\right)^{\frac{\epsilon_{1-\mu}}{1-\epsilon_{1-\mu}}} \tag{8}$$

where $\bar{R}$ is the equilibrium rate given lending capacity $\bar{L}$, that is $\ell\left(\bar{R}\right) = \bar{L}$. With several borrower types (e.g., FICO scores for households or credit ratings for firms), we obtain one curve (8) per borrower. The credit surface consists of the collection of these curves.

The effect of shocks on the credit surface is captured by the sufficient statistics

$$\left\{\frac{R_i}{\ell_i\left(R_i\right)^{\frac{\epsilon^i_{1-\mu}}{1-\epsilon^i_{1-\mu}}}}\right\}_i$$

where $\{R_i\}_i$ are the equilibrium interest rates. First, a negative shock to total lending capacity $\bar{L}$ induces a general upward shift in the credit surface. The interest rate schedules faced on both more elastic loan markets (where R_i is sticky while $\ell_i\left(R_i\right)$ drops) and less elastic markets (where R_i jumps while $\ell_i\left(R_i\right)$ is sticky) shift upwards. Second, interestingly, these changes are heterogeneous across borrower types. Differences in borrower risk $\epsilon^i_{1-\mu}$ determine which parts of the surface increase by more, as we show below.

These results are key to explain how the cross-section of loan terms respond to bank shocks, and how these responses differ across loan markets. In the next section, we start by analyzing the impact of credit supply shocks, then we turn to monetary policy shocks. We conclude by studying their implications for the full dynamics of banking and credit crises.

4 Transmission of Bank Shocks to Loan Terms

This section studies how bank shocks transmit to the loan terms of borrowers with different credit risks across loan markets with different elasticities of loan demand. We start by analyzing credit supply shocks which affect banks' lending capacity, and then turn to monetary policy shocks which affect their cost of funds. We focus on interest rates R and quantity limits l for given non-price terms z, and then we analyze their endogenous response.

4.1 Credit Supply

The transmission of credit supply shocks to borrowers with different credit risks is heterogeneous across loan markets in the data. For credit cards, credit supply expansions tend to benefit low-risk (high FICO score) borrowers, and to be not passed through to high-risk borrowers with a higher propensity to consume (Agarwal et al. 2018). For mortgages, credit supply expansions tend to benefit higher-risk borrowers as illustrated by the subprime boom of the 2000s ((Mian and Sufi, 2009)). These differences are a puzzle for existing macro-finance models. Our main result explains them with a formula which governs how loan quantity limits and interest rates vary across borrowers in response to a shock to bank lending capacity $\overline{L}$.

Definition 7. *Let the risk-adjusted elasticity of loan demand be*

$$\widetilde{\epsilon}_\ell^i = \frac{\epsilon_\ell^i}{\left(1 - \epsilon_{1-\mu}^i\right) + \epsilon_\ell^i \epsilon_{1-\mu}^i},$$

which lies between 1 and ϵ_ℓ^i.

The next proposition shows that the risk-adjusted elasticity of loan demand governs the responses of loan terms to changes in bank lending capacity. This elasticity depends on two sufficient statistics, the repayment elasticity and the elasticity of loan demand.

Proposition 3. *Denote the risk-weighted loan share of borrower i as $\omega^i = \frac{\rho^i \ell^i(R^i)}{\sum \rho^i \ell^i(R^i)}$.*

A change in $\overline{L}$ affect borrowers i's loan quantities and rates as follows:

$$\frac{d\log l^i}{d\log \overline{L}} = \frac{\widetilde{\epsilon}_\ell^i \frac{\rho^i}{R^i(1-\mu^i)}}{\sum \omega^j \widetilde{\epsilon}_\ell^j \frac{\rho^j}{R^j(1-\mu^j)}} \approx \frac{\widetilde{\epsilon}_\ell^i \rho^i}{\sum \omega^j \widetilde{\epsilon}_\ell^j \rho^j}$$

$$\frac{d\log R^i}{d\log \overline{L}} = -\frac{1}{\widetilde{\epsilon}_\ell^i} \times \frac{d\log l^i}{d\log \overline{L}} \approx -\frac{1}{\left(1 - \epsilon_{1-\mu}^i\right) + \widetilde{\epsilon}_\ell^i \epsilon_{1-\mu}^i} \times \frac{\rho^i}{\sum \omega^j \widetilde{\epsilon}_\ell^j \rho^j}$$

Proof. See Appendix B.1. □

Proposition 3 provides closed-form formulas for the endogenous responses of loan contracts in the cross-section of borrowers with different risks and on loan markets with different demand elasticities, in terms of two sufficient statistics. The risk-adjusted elasticities $\widetilde{\epsilon}_\ell^i$ can be constructed from contract curve elasticities ϵ_ℓ^i and repayment elasticities $\epsilon_{1-\mu}^i$. Ultimately, the elasticities ϵ_ℓ^i are obtained from $\epsilon_{1-\mu}^i$ and unconstrained loan demand elasticities $\epsilon_\ell^{u,i}$ using equation (6).

All else equal, the transmission of credit supply shocks $d\log \overline{L}$ to the loan size l^i of borrower i is larger if the risk-adjusted elasticity of loan demand $\widetilde{\epsilon}_\ell^i$ is high. On loan markets with a low elasticity of loan demand $\epsilon^\ell < 1$ such as mortgages, risky borrowers with a high repayment rate elasticity $\epsilon_{1-\mu}$ have a higher risk-adjusted loan demand elasticity $\widetilde{\epsilon}_\ell^i$, and safe borrowers with a low $\epsilon_{1-\mu}$ have a lower $\widetilde{\epsilon}_\ell^i$. On these markets, endogenous default risk increases the elasticity of loan demand for risky borrowers and decreases it for safe borrowers. In response to a credit supply shock, risky borrowers have more volatile loan sizes and less volatile interest rates than safe borrowers. The opposite is true on loan markets with a high elasticity of loan demand $\epsilon^\ell > 1$ such as credit cards. In response to a credit supply shock, safe borrowers have more volatile loan sizes and less volatile interest rates than risky borrowers.

The transmission of shocks $d\log \overline{L}$ to the loan size l^i is larger if borrower i has a high risk-weight ρ^i in the bank's capacity constraint. However, this effect is asymmetric across loan markets. On markets with a high elasticity of loan demand, the quantity limits and the interest rates of borrowers with low risk-weights are insulated from credit supply shocks as the changes $\frac{d\log l^i}{d\log \overline{L}}$ and $\frac{d\log R^i}{d\log \overline{L}}$ are small. On markets with a low elasticity of loan demand, the the quantity limits of borrowers with low risk-weights are also insulated, but their interest rates can experience a sharp increase.

Credit crunch. As an illustration, consider an economy with two borrower types a and b in equal mass, where b borrowers are risky. Both types of borrowers have the same preferences and technology, and they only differ in their risk. As Figure 4 shows, on markets with a high elasticity of loan demand such as credit cards, endogenous default risk makes risky borrowers effectively less elastic ($1 < \widetilde{\epsilon}_\ell^b < \widetilde{\epsilon}_\ell^a$). Their loan quantities contract by less and their loan rates increase by more in response to a negative credit supply shock $d\log \overline{L} < 0$. As Figure 5 shows, on inelastic markets such as mortgages, risky borrowers are effectively more elastic ($\widetilde{\epsilon}_\ell^a < \widetilde{\epsilon}_\ell^b < 1$). Their loan quantities contract by more and their loan rates increase by less when $d\log \overline{L} < 0$.

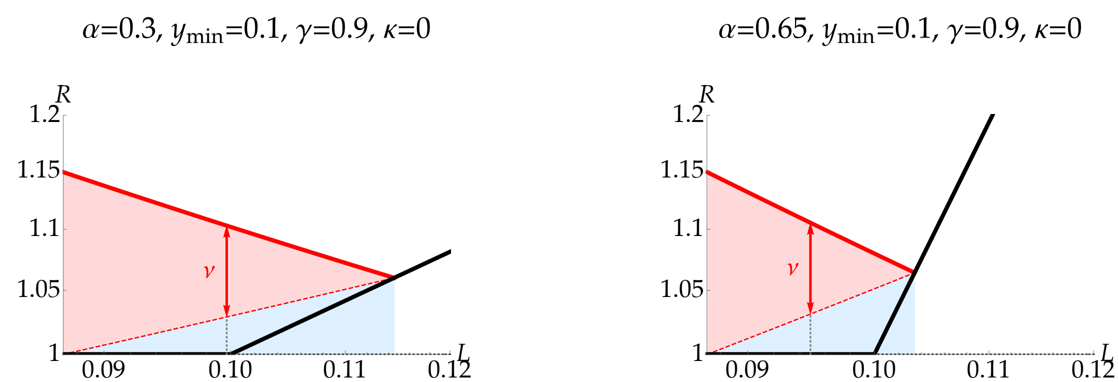

Figure 4: Firms with high $\gamma = 0.9$, log-log scale. Left: safer borrowers (low α), right: riskier borrowers (high α). The vertical red segment has length $\nu \approx 7\%$: the bank equalizes ν across types, which tells us how quantities and rates react for each type. Both types have the same actual loan demand elasticity (the orange line with slope $-(1-\gamma)$) but the contract curve is less elastic for riskier borrowers ($1 < \epsilon_\ell^b < \epsilon_\ell^a$) hence credit contracts more for safer borrowers.

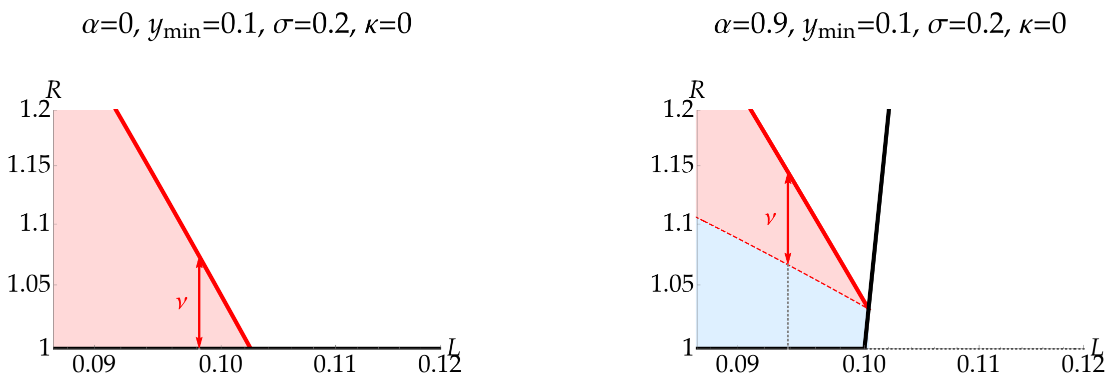

Figure 5: Households with low EIS $\sigma = 0.2$, log-log scale. Left: safe borrowers ($\alpha = 0$), right: risky borrowers (high α). The vertical red segment has length $\nu \approx 7\%$: the bank equalizes ν across types, which tells us how quantities and rates react for each type. The contract curve is more elastic for risky borrowers ($\epsilon_\ell^a < \epsilon_\ell^b < 1$) hence credit contracts more for them.

Credit boom. Proposition 3 explains the heterogeneous transmission of the credit boom of the 2000s to risky borrowers on mortgage markets and to safe borrowers on credit markets. The positive credit supply shock $d \log \overline{L} > 0$ generated an increase in loan quantities l and a decrease in interest rates R on both markets. For mortgages, the low elasticity of loan demand $\epsilon_\ell < 1$ resulted in a decrease in the face value Rl to be repaid, hence a decrease in default risk μ for all borrowers. For risky borrowers with a high repayment rate elasticity $\epsilon_{1-\mu}$, the decrease in default risk was larger, hence the increase in expected profits per risk-weighted dollars $\frac{\pi}{\rho l}$ was also larger for these borrowers. In order to increase their total profits, banks optimally increased lending l to them. Therefore, loan quantities increased more for risky than for safe mortgage borrowers during the credit boom. The opposite happened for credit cards. The high elasticity of loan demand $\epsilon_\ell > 1$ resulted in an increase in the face value Rl to be repaid, hence an increase in default risk μ. For safe borrowers with a low repayment rate elasticity $\epsilon_{1-\mu}$, the increase in default risk was smaller, hence the decrease in expected profits per risk-weighted dollars $\frac{\pi}{\rho l}$ was also smaller. In order to increase their total profits, banks optimally increased lending l to them. Therefore, loan quantities increased more for safe than for risky credit card borrowers.

4.2 Monetary Policy

Monetary policy shocks to the risk-free rate R^f change banks' cost of funds, and the transmission of these changes is heterogeneous across loan markets in the data. Interest rates on new loans are sticky for credit cards, but they vary significantly over time for mortgages. Furthermore, the transmission of monetary policy depends on bank health. It can be weaker when bank balance sheets are impaired (e.g., Jimenez et al. 2012, Acharya et al. 2019), or it can be stronger (Geanakoplos and Rappoport (2019)). The goal of this section is to explain these features by extending our previous results to monetary policy shocks.

4.2.1 Bank Lending Channels: Loan Markets

The bank lending channels of monetary policy depend on borrower credit risk and the elasticity of loan demand on various markets. The transmission of shocks to banks' funding cost R^f into loan rates and quantities also depends on whether banks' lending capacity constraints are binding.

Unconstrained banks. If banks' loan supply $\bar{L}$ is sufficiently high (i.e., larger than the total unconstrained loan demand L^*), then the interest rate pass-through works through banks' zero profit condition for each borrower $R_i\left(1-\mu_i\right)=R^f$. The interest rate transmission after accounting for changes in loan quantities is

$$\frac{d\log R_i}{d\log R^f}=\frac{1}{1-\epsilon^i_{1-\mu}+\epsilon^i_{1-\mu}\epsilon^i_{\ell}}$$

The interest rate R increases less than one for one with banks' funding cost R^f and the transmission of monetary policy is weaker if $\epsilon^i_{1-\mu}\left(\epsilon^i_{\ell}-1\right)>0$. In contrast, it is stronger if $\epsilon^i_{1-\mu}\left(\epsilon^i_{\ell}-1\right)<0$. Therefore, interest rates are sticky in response to monetary policy shocks on loan markets where the elasticity of loan demand ϵ^{ℓ} is high such as credit cards. They are more volatile on markets where ϵ^{ℓ} is low such as mortgages.

Constrained banks. The change in bank lending capacity following a monetary policy shock $-\frac{d\log\bar{L}}{d\log R^f}$ is key for the resulting changes in loan terms. One interpretation of $\bar{L}$ is that it arises from a regulatory constraint faced by the bank which limits its total lending depending on its capital. In that case, monetary shocks affect $\bar{L}$ in a way that depends on the duration of equity. Another interpretation is that $\bar{L}$ is determined by banks' market power on deposits, together with constraints on wholesale funding. While our loan market is perfectly competitive, one possibility is to interpret the "deposit channel of monetary policy" in Drechsler, Savov and Schnabl 2017 as an effect of R^f on banks' total loan supply $\bar{L}$, inclusive of how banks use their market power on deposits.

If $\bar{L}$ does not react to R^f, then steady state interest rates and quantities will not change and R^f only affects banks' static profit per dollar. As we show in the next section, monetary policy can still have a dynamic effect on total lending $\bar{L}_t$ over the transition. If $\bar{L}$ reacts to R^f, then the transmission of monetary policy to loan terms on a given market can be either dampened or amplified depending on the elasticity of loan demand. The next proposition formalizes these results (assuming $\rho^i=1$ for simplicity).

Proposition 4. *Suppose $\rho^i = 1$. The pass-through of banks' funding cost R^f to loan rate R^i is*

- *When banks are unconstrained:*

$$\frac{d\log R^i}{d\log R^f} = \frac{1}{1-\epsilon^i_{1-\mu}+\epsilon^i_{1-\mu}\epsilon^i_\ell} = \frac{\widetilde{\epsilon}^i_\ell}{\epsilon^i_\ell}$$
$$\frac{d\log l^i}{d\log R^f} = -\widetilde{\epsilon}^i_\ell$$

- *When banks are constrained*

$$\frac{d\log R^i}{d\log R^f} = -\frac{1}{\epsilon^i_\ell}\times\frac{\widetilde{\epsilon}^i_\ell}{\widetilde{\epsilon}_\ell}\frac{d\log \bar{L}}{d\log R^f},$$
$$\frac{d\log l^i}{d\log R^f} = -\widetilde{\epsilon}^i_\ell\times\left(-\frac{1}{\widetilde{\epsilon}_\ell}\frac{d\log\bar{L}}{d\log R^f}\right)$$

If $\frac{d\log\bar{L}}{d\log R^f} = -\widetilde{\epsilon}_\ell$, monetary policy transmission to loan terms does not depend on the bank's capacity constraint.

Proof. See Appendix B.1. □

Figure 6 describes the percentage changes in interest rates and loan sizes in response to changes in deposit rates, as a function of the elasticity of loan demand and banks' capacity constraint. First, when faced with a tightening of monetary policy, unconstrained banks (left panels) pass through the increase in the policy rate more than one-for-one on loan markets with a low demand elasticity, and less than one-for-one on markets with a high elasticity, where the bank lending channel is dampened. However, the decrease in loan sizes is steeper on markets on these markets, so that borrowers are eventually relatively more credit-rationed. Second, when banks are constrained but their capacity constraints are not very sensitive to the policy rate (middle panels), the transmission of monetary policy is further dampened. Loan terms are largely determined by bank's lending capacity, so the relative insensitivity of the latter to the policy rate partly insulates borrowers from a credit tightening. Conversely, an insensitive bank lending capacity reduces the transmission of policy cuts to the cross-section of loan terms. Third, a high sensitivity of banks' capacity constraints to policy rates (right panels) amplifies the transmission of monetary policy to loan terms, since interest rates react more than one-for-one on all markets.

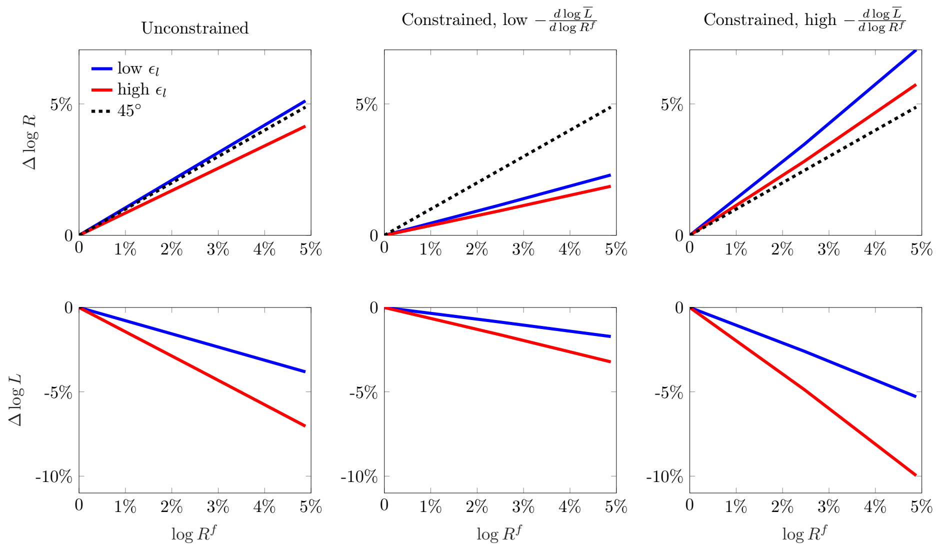

Figure 6: Percentage changes in loan rates and sizes as a function of banks' funding cost. Changes are plotted in cases where banks are unconstrained (left panels), constrained with inelastic lending capacity (middle panels), and constrained with elastic lending capacity (right). For each case, they are plotted for borrowers with a low (blue) vs. high (red) elasticity of loan demand.

4.2.2 Bank Lending Channels: Borrowers

Proposition 4 explains why a decrease in banks' cost of funds is transmitted differently to safe and risky borrowers across loan markets with different elasticities of loan demand, in a similar way to credit supply booms.

Elastic loan markets. On markets with a high demand elasticity such as credit cards, banks do not pass-through lower funding costs to risky borrowers (Agarwal et al. 2018). This is because, as we have shown, safe borrowers with a low endogenous default risk $\epsilon_{1-\mu}$ have a higher effective elasticity of loan demand $\widetilde{\epsilon}_\ell^i$ on these markets. Therefore banks transmit reductions in R^f relatively more to low-risk borrowers, both when they are constrained and unconstrained.

Inelastic loan markets. On markets with a low demand elasticity such as mortgages, policy rates can affect loan terms through multiple channels, which include adjustable-rate mortgage payments, interest rates on newly originated fixed-rate mortgages, payment-to-income constraints, and refinancing rates. We illustrate these effects using the special case of our model with collateralized loans described in Section 3.1. We extend the model to include mortgages, and compute the loan-to-value (LTV) and payment-to-income (PTI) ratios associated with a given loan. Using the same notation, the LTV is $\frac{l}{P_0}$ and the PTI is $(R-1)l/y_0$.

Figure 7 depicts their responses to changes in banks' cost of fund. Because the elasticity of loan demand is low for mortgages, risky borrowers have a higher effective elasticity of demand on this market, and safe borrowers have a lower elasticity. Three features arise. First, borrowers with a high demand elasticity have higher LTV and PTI ratios than borrowers with a low elasticity. Second, their LTV decreases more than for low elasticity borrowers when banks' cost of funds increases, but their PTI increases less. For a given change in loan rates, they adjust their optimal loan size by more than low elasticity borrowers. This results in strongly heterogeneous responses in loan term changes between risky and safe borrowers. Third, these changes are amplified by a greater sensitivity of bank lending capacity to their cost of fund, as implied by Proposition 4.

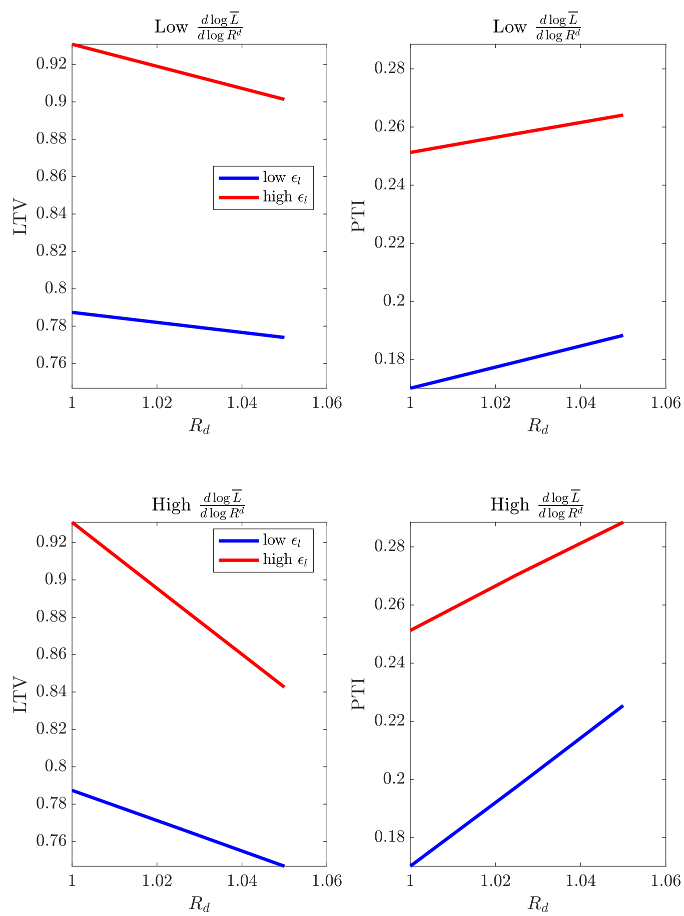

Figure 7: Effect of the bank deposit rate on mortgage terms, for low (blue) and high (red) elasticity borrowers, and as a function of the elasticity of banks' lending capacity to banks' cost of fund (low in upper panels, high in lower panels).

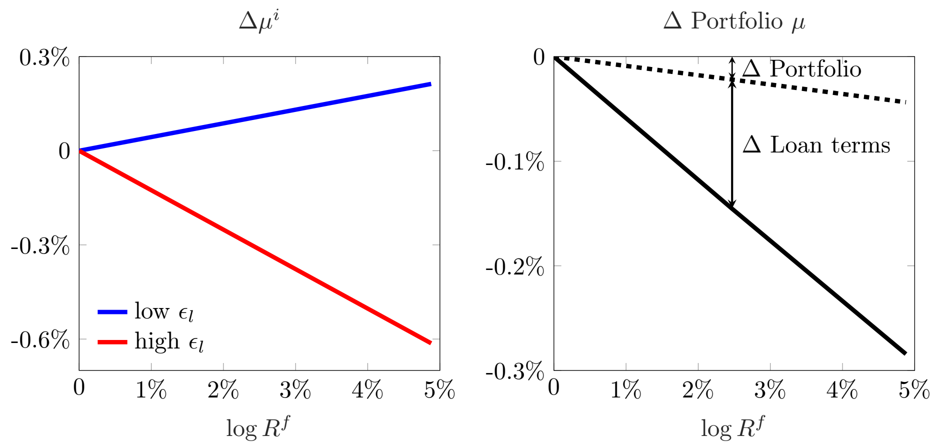

Figure 8: Effect of the risk-free rate on individual default probabilities for borrowers with low vs. high loan demand elasticity (left panel), and on the total default probability of the bank's loan portfolio (right panel).

4.2.3 Banks' Portfolio Risk: Intensive vs. Extensive Margin

With endogenous credit risk, changes in loan terms due to variations in banks' cost of fund will affect the total credit risk borne by banks. We now analyze how banks' portfolio risk reacts to a decrease in R^f. The total effect can be decomposed into two parts: first, a change in the price and non-price terms of loans for a given set of borrowers; second, a reallocation of bank loans towards specific borrowers.

Within borrower groups, lower interest rates lead to less credit risk on loan markets with a low demand elasticity as Rl decreases, and more credit risk on markets with a high elasticity as Rl increases. In addition, there is a composition effect towards high elasticity borrowers because their loan sizes increase more after a decrease in R^f. The effect on the bank's portfolio risk depends on the credit risk of these borrowers. On markets with a low average demand elasticity such as mortgages, the borrowers with a higher elasticity have a higher risk of default. On these markets, a credit boom increases the weight of risky borrowers in banks' portfolios. The opposite is true for markets with a high demand elasticity such as credit cards, where the borrowers with a higher elasticity have a lower risk. The total effect is

$$d \log \left(\mathbf{E}\left[1-\mu^i\right]\right) \approx \mathbf{Cov}\left(\mu^i, \tilde{\epsilon}^i_\ell\right) d \log R^f - \frac{\mathbf{E}\left[d\mu^i\right]}{\mathbf{E}\left[1-\mu^i\right]}$$

Figure 8 shows how individual default probabilities react to changes in banks' cost of fund, and the resulting change in the total credit risk of the bank's portfolio of loans. First, the decrease in R^f results in a smaller decrease in the loan rate of high elasticity borrowers, who face a larger increase in loan size. As a result, their credit risk increases sharply when rates decrease. In contrast, the risk of low elasticity borrowers decreases, even though their interest rate decreases more than one-for-one with R^f. Interestingly, endogenous default risk leads to opposite results for borrowers with different demand elasticities. Second, even when these borrowers have identical weights in bank's loan portfolio, the total effect is an increase in the total credit risk borne by the bank. Quantitatively, most of it arises from changes in loan terms, which induce a large increase in credit risk for high elasticity borrowers.

4.2.4 Other Non-Price Terms: Reaching for Yield and Covenant-Lite Loans

The issuance of loans with weak covenants has been linked to historically low risk-free rates (e.g., Roberts and Schwert 2020). In the model, lenders trade off the price and non-price terms as their cost of funds decreases. Figure 9 shows that this trade off depends on the elasticity of loan demand ϵ_ℓ. Low interest rates are associated with looser covenants z^* for low elasticity borrowers, whose default risk falls when rates are low because Rl decreases. However, they are associated with tighter covenants for high elasticity borrowers for which Rl increases.

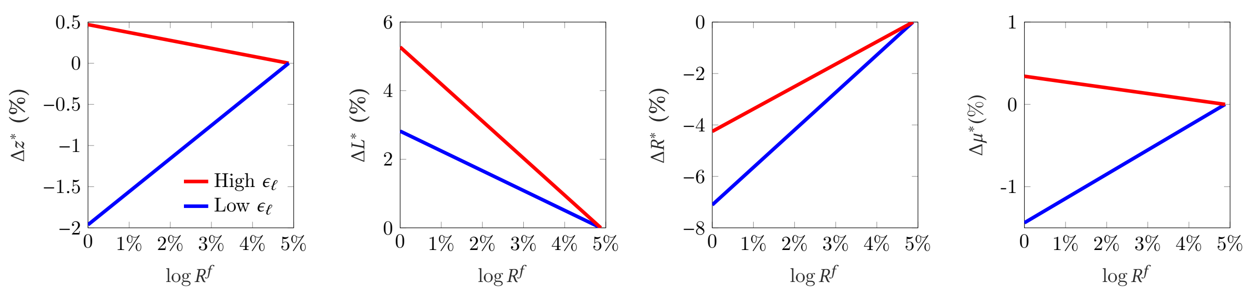

Figure 9: Covenant-lite loans: effect of bank's cost of fund on loan terms starting from a baseline $\log R^f = 5\%$, contrasting low elasticity (blue) and high elasticity (red) borrowers.

5 Endogenous Persistence of Credit Crises

After analyzing how the cross-section of loan terms reacts to changes in credit supply and monetary policy in the steady state of the model, we turn to the transition dynamics of credit crises. We first study the impact and the persistence of a deterioration in banks' balance sheets. Then we illustrate our results with a calibrated model of the U.S. mortgage market. We conclude by discussing how policy interventions can improve the slow recovery from a credit crisis.

5.1 Dynamic Model

For simplicity, we start by assuming that there is a single type of borrower. Starting from an unconstrained steady state loan demand l^* with $\nu^* = 0$, we consider a negative credit supply shock which lowers banks' capacity $\bar{L}$ to $\bar{L}_0$ and increases the excess loan premium (the associated Lagrange multiplier) to $\nu_0 > 0$. The law of motion for the total supply of loans is

$$l_{t+1} = (1 + \phi \nu_t)\, l_t$$

where ϕ is equal to the earnings retention ratio times the leverage.

At each date, we have

$$(1 + \phi \nu_t)\, \ell \left(R \left(l \left(\nu_t \right), \nu_t \right) \right) = \ell \left(R \left(l \left(\nu_{t+1} \right), \nu_{t+1} \right) \right),$$

where as earlier $R(l, \nu)$ solves $R(1 - \mu(Rl)) = R^f + \nu$ (holding R^f and z fixed), and where $l(\nu)$ solves the static loan market clearing condition $l = \ell(R(l, \nu))$. Linearizing around the steady state l^* and using our previous expressions for $\frac{\partial R}{\partial \nu}, \frac{\partial R}{\partial l}$, we obtain the law of motion for the excess loan premium ν_t:

$$\nu_{t+1} = \left(1 - \frac{\phi R^f}{\widetilde{\epsilon_\ell}} \right) \nu_t$$

where as earlier $\widetilde{\epsilon_\ell} = \frac{1}{\left(1 - \epsilon_{1-\mu}\right) \frac{1}{\epsilon_\ell} + \epsilon_{1-\mu}}$. The initial jump is $\nu_0 = \frac{R^f}{\widetilde{\epsilon_\ell}} \times \frac{l^* - \bar{L}_0}{l^*}$. Therefore:

Proposition 5. *Let $\varphi = \frac{\phi R^f}{\widetilde{\epsilon_\ell}}$. To first-order in the size of the initial credit supply shock $\delta = \frac{l^* - \bar{L}_0}{l^*}$, the excess loan premium ν_t and bank lending L_t follow*

$$\nu_t = \frac{R^f}{\widetilde{\epsilon_\ell}} \delta \left(1 - \varphi \right)^t, \tag{9}$$

$$l_t = l^* \left[1 - \delta \left(1 - \varphi \right)^t \right]. \tag{10}$$

We can measure the persistence of the crisis through the half-life of ν, defined as the time T such that the excess loan premium has reverted to one-half of its initial value $\nu_T = \nu_0/2$:

$$T = \frac{\log 2}{-\log \left(1 - \varphi \right)}.$$

A special case of Proposition 5 holds in canonical macro-finance models (e.g. Kiyotaki and Moore 1997, Gertler and Kiyotaki 2010): credit risk μ is exogenous, hence $\widetilde{\epsilon_\ell}$ and ϵ_ℓ are both equal to the unconstrained loan demand elasticity ϵ_ℓ^u. Accounting for non-linear contracts and endogenous default risk brings the elasticity $\widetilde{\epsilon_\ell}$ closer to 1, hence the half-life gets closer to $\frac{\log 2}{-\log\left(1 - \phi R^f\right)}$, relative to the exogenous default (constant μ) case. In particular, in the limit of infinitely elastic unconstrained loan demand, the half-life increases without bounds when credit risk μ is constant, while it remains bounded as long as $\epsilon_{1-\mu} > 0$. This is because, as shown earlier, default risk decreases endogenously during credit crises if $\widetilde{\epsilon_\ell} > 1$. In very elastic markets, interest rates do not decrease to offset the lower default risk. As a result, banks earn excess returns v_t that recapitalize them back to the steady state in finite time.

Impact vs. persistence. For a given initial shock, the crisis is more persistent if loan demand is very elastic since the loan rate R and hence profits cannot jump by a large amount. How are borrowers' loan payments affected? At a higher elasticity ϵ_l, the initial jump v_0 is also smaller. Overall, this gives rise to an intertemporal trade-off between successive generations of borrowers. Current borrowers are hurt more (as measured by the spread v they face) with a lower elasticity $\widetilde{\epsilon_\ell}$, as the crisis is initially very sharp. But a low elasticity $\widetilde{\epsilon_\ell}$ also implies that the crisis is short-lived hence future borrowers are barely affected. By contrast, the crisis is more persistent and future borrowers are more affected with a higher elasticity $\widetilde{\epsilon_\ell}$.

An important consequence of accounting for non-linear contracts and endogenous credit risk is that the speed of bank recapitalization changes relative to the benchmark with linear contracts (i.e., no quantity limit imposed by the lenders). The direction of this effect depends on the elasticity:

- if $\epsilon_l < 1$, then $\frac{\log 2}{-\log\left(1-\frac{\phi R^f}{\widetilde{\epsilon_l}}\right)} \leq T \leq \frac{\log 2}{-\log\left(1-\phi R^f\right)}$: relative to a benchmark with linear contracts, crises are milder on impact, but more persistent;
- if $\epsilon_l > 1$, then $\frac{\log 2}{-\log\left(1-\phi R^f\right)} \leq T \leq \frac{\log 2}{-\log\left(1-\frac{\phi R^f}{\widetilde{\epsilon_l}}\right)}$: relative to a benchmark with linear contracts, crises are sharper on impact, but more short-lived.

Interestingly, we find that impact and persistence balance each other exactly in the following present value sense:

Proposition 6. *For an initial credit supply shock $\delta = \frac{l^*-\bar{L}_0}{l^*}$, the cumulative excess loan premium is given by*

$$\sum_{t=0}^{\infty} v_t = \frac{\delta}{\phi}$$

and is therefore independent of $\widetilde{\epsilon_\ell}$.

This result highlights the benefit from using sufficient statistics in our analysis. In the relatively wide range of settings that we consider, the cumulative impact of the shock, as measured by the cumulative excess loan premium, does not depend on the details of the environment, such as borrower preferences, the particular information frictions that affect $\epsilon_{1-\mu}$, and the feasible contract space (i.e., linear vs. non-linear contracts). Proposition 6 provides a testable prediction for different crises across time and space: as long as the parameter ϕ remains the same (or can be controlled for), the cumulative spread relative to the percent impact effect on quantities δ should be unchanged.

Dynamics with heterogeneous borrowers. We now combine the dynamic analysis with our previous results on the incidence of credit supply shocks in the cross-section of borrowers. Assume equal risk-weights $\rho^i = 1$ for simplicity. To first order in each period, the aggregate loan supply is

$$\begin{aligned}\sum_i l^i_{t+1}\left(\nu_{t+1}\right) &= \left(1+\phi\nu_t\right)\sum_i l^i_t\left(\nu_t\right) \\ &= \left(1+\phi\nu_t\right)\sum_i l^{i,*}\left(1-\epsilon^i_\ell \nu_t\right)\end{aligned}$$

hence the excess loan premium ν_t follows the same dynamics (9) as with homogeneous borrowers, except that $\widetilde{\epsilon}_\ell = \sum \omega^i \widetilde{\epsilon}^i_\ell$ is now a weighted average of individual elasticities with weights equal to the steady state loan shares $\omega^i = \frac{l^{i,*}}{\sum l^{i,*}}$. Thus the speed of recapitalization is now governed by the average (risk-adjusted) elasticity: credit crunches will be more persistent if on average banks lend to more elastic borrowers.

How are different borrowers affected over time? Each loan quantity evolves as

$$l^i_t = l^{i,*} - \widetilde{\epsilon}^i_\ell \nu_t.$$

Therefore all loan quantities recover at the same speed, since they all depend on the common excess loan premium ν_t. High $\widetilde{\epsilon}^i_\ell$ borrowers suffer a larger initial tightening, and for two types i, j the relative tightening is constant over time:

$$\frac{l^{j,*}-l^j_t}{l^{i,*}-l^i_t} = \frac{\widetilde{\epsilon}^j_\ell}{\widetilde{\epsilon}^i_\ell} \quad \forall t$$

Finally, the bank earns expected profits $\nu_t l^i_t$ from type i borrowers, and profit per dollar ν_t is common to all borrowers. As a result, elastic borrowers suffer a large contraction in credit. Inelastic borrowers are hurt less by the credit crunch in terms of quantity borrowed, but they are also the ones paying for the excess spreads that allow the bank to recapitalize.

5.2 Numerical Example: Mortgage Market

To illustrate our findings, we conclude by analyzing a credit crisis in a simple calibrated model of the U.S. mortgage market, where loan contracts with many price and non-price terms are traded. We study the transition dynamics of mortgage markets in response to a contraction of banks' balance sheets. We introduce overlapping generations of households in the two-period model of Appendix 3.1, and let banks allocate credit to multiple borrowers with different loan supply elasticities. We calibrate the model to recent U.S. data and discuss estimates for our two sufficient statistics: the interest-rate elasticities of loan supply, and of borrowers' repayment probability.

We then present impulse response functions, and study two policy interventions to mitigate the shortgage of credit: direct household debt relief and bank recapitalization. In addition to mortgage spreads, we focus on two non-price dimensions of loans: maximum loan-to-value (LTV) and payment-to-income (PTI) ratios, which limit mortgage sizes for a given interest rate.

5.2.1 Calibration

There are two equal measures of types of borrowers, with respectively high and low elasticities of loan demand. We calibrate the model pre-crisis steady state to match six key moments for the U.S. mortgage market in 2000-2007, shown in the table below.

We target a low interval for the interest rate-elasticity of borrowers' loan demand, which reflects the range of estimates available in the literature. We obtain 0.6 for low-elasticity (low EIS) borrowers, and 1.4 for high-elasticity (high EIS) borrowers. In survey data, Fuster and Zafar 2021 estimate an elasticity of 0.11. Relying on the discontinuity of interest rates at various loan sizes, Best et al. 2019 estimate the elasticity of LTV to be 0.5 for the U.K. nortgage market. In a structural model of the U.S. banking system, Benetton 2021 estimates an elasticity of 0.07.[4]

To calibrate the interest-rate elasticity of borrowers' repayment probability, we compute the elasticity of the complement event: the elasticity of borrowers' default probability, which can be more easily estimated. We also target an interval which reflects the range of estimates available in the literature. We obtain a default rate elasticity of 1.45. Fuster and Willen 2017 estimate that the elasticity of mortgage default to monthly payment size is 1.1, and Di Maggio et al. 2017 that it is 2.

Finally, we target values for mortgage rates, maximum LTV and PTI ratios to reflect averages for those variables over the period 2000-2007. We obtain an average mortgage rate of 15%, a LTV ratio of 0.82, and a PTI ratio of 0.15, with the first two moments in the corresponding intervals of $[3\%, 18\%]$ and $[0.8, 1]$ for new originated mortgages (source: Black Knight, eMBS, HMDA, SIFMA, CoreLogic and Urban Institute), and the third moment below the $[0.3, 0.5]$ interval.

Variable	Description	Value	Target	Source
ϵ_l	Int. rate elasticity of loan demand	0.6, 1.4	$[0.11, 5]$	See text
ϵ_μ	Int. rate elasticity of default prob.	1.45	$[0.15, 2]$	See text
R	Mortgage rate	1.15	$[1.03, 1.18]$	Primary Mortgage Survey (30-Year FRM)
LTV	Max. loan-to-value	0.82	$[0.8, 1]$	Urban Institute
PTI	Max. payment-to-income	0.15	$[0.3, 0.5]$	Urban Institute

[4]The interest rate elasticity of mortgage demand can be decomposed into a within-loan elasticity for households who borrow using the same loan, and a between-loan elasticity for households who switch between loan products altogether. those who change loan altogether. We focus on the within-loan elasticity because our model does not feature a discrete loan product choice.

5.2.2 Results

Figure 10 describes the dynamics of a credit crisis with multidimensional mortgage contracts. At $t = 0$, the mortgage market is in steady state. At $t = 0^+$, banks' lending capacity $\overline{L}$ unexpectedly contracts, resulting in an increase in the excess loan premium ν, which reflects the tightness of the banks' constraint (black line). Loan sizes fall and mortgage spreads increase for both types of households, as a result of the negative shock to credit supply. Because the dynamics of house prices remains unchanged, LTV ratios fall in response to the decrease in loan sizes. Interestingly, the response of PTI ratios depends on the relative changes in interest rates and loan sizes. Because the increase in rates dominates the decrease in loan sizes, PTI ratios increase for both borrower types.

The responses of loan terms are heterogeneous when disaggregated across borrowers. Loan size falls twice as much for high elasticity borrowers as for low elasticity ones, while the mortgage spread increases by less. As a result, default risk falls sharply for high elasticity borrowers, while it increases by about the same amount for low elasticity ones.

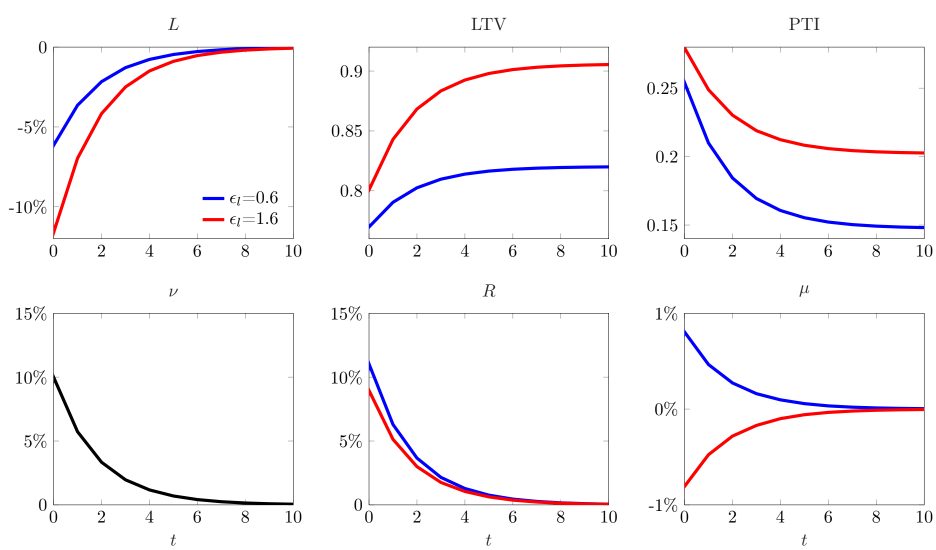

Figure 10: Dynamics of the cross-section of loans terms in the U.S. mortgage market in response to a tightening in banks' lending capacity. Impulse response functions for loan sizes, loan-to-value and payment-to-income ratios, excess loan premium, mortgage spreads, and default risk are plotted for low (blue lines) and high interest-rate elasticity borrowers (red lines).

5.2.3 Credit-Ameliorating Policies: Debt Relief and Bank Recapitalization

We conclude by studying the effectiveness of two policies designed to ameliorate the shortgage of credit in response to a credit supply shock. Direct borrower debt relief is modeled as a lump-sum transfer when households initially borrow from banks, which effectively reduces their indebtedness level. Bank recapitalization is modeled as a relaxation of banks' lending constraints, which can be implemented by direct equity injections. Both policies have been advocated during the U.S. mortgage market crisis, yet without clear guidance on the best way to direct credit to the borrowers who need it most.

Figure 11 illustrates how these policies work. Thick lines show laissez-faire outcomes under the same credit crisis as Figure 10, and dotted lines show the equilibrium path under the two policies, for the two types of borrowers. In our calibration, the two policies have very similar effects in mitigating the fall in LTV ratios due to a decrease in loan sizes. In the case of debt relief, households internalize that they are effectively richer when the loan is originated, and can borrow relatively more without excessively increasing their credit risk. In the case of bank recapitalization, banks' lending capacity does not fall as much, and therefore more credit is available to households, who decide to use it because they are credit-constrained. While the two policies have similar effects across borrowers, debt relief increases the speed at which loan size recovers for high elasticity borrowers slightly more than bank recapitalization.

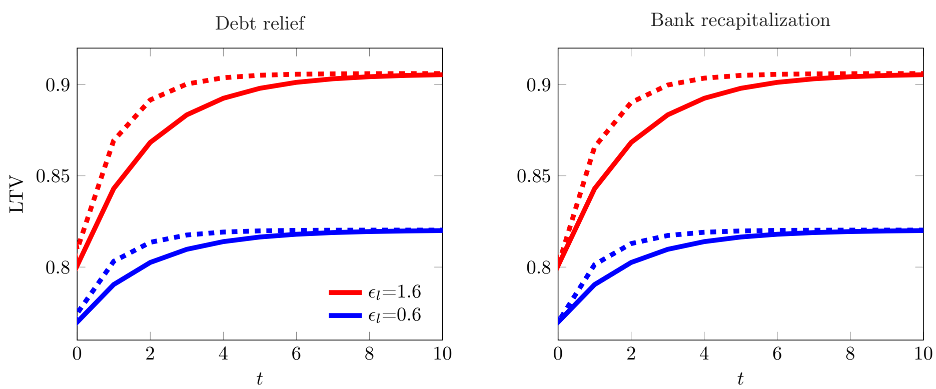

Figure 11: Impulse responses under two credit-ameliorating policies: debt relief and bank recapitalization. Thick lines show laissez-faire outcomes, and dotted lines show the equilibrium path under the two policies. Red (resp. blue): high (resp. low) elasticity borrowers.

6 Conclusion

We propose a model of multidimensional contracting between heterogeneous borrowers and intermediaries with limited lending capacity. We show that two sufficient statistics, the elasticities of borrowers' loan demand and default rates to interest rates, predict how the cross-section of loan terms and banks' portfolio risk react to changes in bank capital and funding costs. Our results help explain key features of loan markets, especially the heterogeneous transmission of shocks across borrowers and loan products, as well as the rise of covenant-lite lending in low risk-free rate environments. They provide useful guidance for comparing applications to specific loan markets in structural models. Our results also have normative implications. The two elasticities drive the dynamic effect of credit crises through the combination of impact and persistence. They help understand how policies can best direct credit to borrowers who need it the most during downturns.

References

Acharya, Viral, Katherina Bergant, Matteo Crosignani, Tim Eisert, and Fergal McCann (2020) "The Anatomy of the Transmission of Macroprudential Policies," *The Journal of Finance*.

Acharya, Viral V, Björn Imbierowicz, Sascha Steffen, and Daniel Teichmann (2019) "Does the Lack of Financial Stability Impair the Transmission of Monetary Policy?," *Journal of Financial Economics*, Vol. forthcoming.

Adams, William, Liran Einav, and Jonathan Levin (2009) "Liquidity Constraints and Imperfect Information in Subprime Lending," *American Economic Review*, Vol. 99, pp. 49–84.

Agarwal, Sumit, Souphala Chomsisengphet, Neale Mahoney, and Johannes Stroebel (2018) "Do Banks Pass through Credit Expansions to Consumers Who want to Borrow?" *The Quarterly Journal of Economics*, Vol. 133, pp. 129–190.

Benetton, Matteo (2021) "Leverage Regulation and Market Structure: A Structural Model of the UK Mortgage Market," *The Journal of Finance*.

Bernanke, Ben S., Mark Gertler, and Simon Gilchrist (1999) "Chapter 21 The financial accelerator in a quantitative business cycle framework," Vol. 1 of Handbook of Macroeconomics: Elsevier, pp. 1341–1393.

Best, Michael Carlos, James S Cloyne, Ethan Ilzetzki, and Henrik J Kleven (2019) "Estimating the Elasticity of Intertemporal Substitution Using Mortgage Notches," *The Review of Economic Studies*, Vol. 87, pp. 656–690.

Brunnermeier, Markus K. and Yuliy Sannikov (2014) "A Macroeconomic Model with a Financial Sector," *American Economic Review*, Vol. 104, pp. 379–421.

Buchak, Greg, Gregor Matvos, Tomasz Piskorski, and Amit Seru (2020) "Beyond the Balance Sheet Model of Banking: Implications for Bank Regulation and Monetary Policy," working paper.

Chakraborty, Indraneel, Itay Goldstein, and Andrew MacKinlay (2018) "Housing Price Booms and Crowding-Out Effects in Bank Lending," *The Review of Financial Studies*, Vol. 31, pp. 2806–2853.

——— (2020) "Monetary stimulus and bank lending," *Journal of Financial Economics*, Vol. 136, pp. 189–218.

Chatterjee, Satyajit, Dean Corbae, Makoto Nakajima, and Jose Victor Rios-Rull (2007) "A Quantitative Theory of Unsecured Consumer Credit with Risk of Default," *Econometrica*, Vol. 75, pp. 1525–1589.

Chodorow-Reich, Gabriel (2014) "The Employment Effects of Credit Market Disruptions: Firm-level Evidence from the 2008-09 Financial Crisis," *Quarterly Journal of Economics*, Vol. 129, pp. 1–59, Lead article.

Chodorow-Reich, Gabriel, Olivier Darmouni, Stephan Luck, and Matthew Plosser (2021) "Bank liquidity provision across the firm size distribution," *Journal of Financial Economics*.

Chodorow-Reich, Gabriel and Antonio Falato (2021) "The Loan Covenant Channel: How Bank Health Transmits to the Real Economy," *The Journal of Finance*.

DeFusco, Anthony A. and Andrew Paciorek (2017) "The Interest Rate Elasticity of Mortgage Demand: Evidence from Bunching at the Conforming Loan Limit," *American Economic Journal: Economic Policy*, Vol. 9, pp. 210–240.

Di Maggio, Marco, Amir Kermani, Benjamin J. Keys, Tomasz Piskorski, Rodney Ramcharan, Amit Seru, and Vincent Yao (2017) "Interest Rate Pass-Through: Mortgage Rates, Household Consumption, and Voluntary Deleveraging," *American Economic Review*, Vol. 107, pp. 3550–3588.

Diamond, Douglas W, Yunzhi Hu, and Raghuram G Rajan (2020) "Pledgeability, industry liquidity, and financing cycles," *The Journal of Finance*, Vol. 75, pp. 419–461.

Diamond, William and Tim Landvoigt (forthcoming) "Credit Cycles with Market Based Household Leverage."

Drechsler, Itamar, Alexi Savov, and Philipp Schnabl (2017) "The Deposits Channel of Monetary Policy*," *The Quarterly Journal of Economics*, Vol. 132, pp. 1819–1876.

Einav, Liran, Mark Jenkins, and Jonathan Levin (2012) "Contract Pricing in Consumer Credit Markets," *Econometrica*, Vol. 80, pp. 1387–1432.

Favilukis, Jack, Sydney C. Ludvigson, and Stijn Van Nieuwerburgh (2017) "The Macroeconomic Effects of Housing Wealth, Housing Finance, and Limited Risk Sharing in General Equilibrium," *The Journal of Political Economy*, Vol. 125, pp. 140–223.

Fuster, Andreas and Paul S. Willen (2017) "Payment Size, Negative Equity, and Mortgage Default," *American Economic Journal: Economic Policy*, Vol. 9, pp. 167–191.

Fuster, Andreas and Basit Zafar (2021) "The Sensitivity of Housing Demand to Financing Conditions: Evidence from a Survey," *American Economic Journal: Economic Policy*, Vol. 131.

Geanakoplos, John (2010) *The Leverage Cycle*, pp. 1–65: University of Chicago Press.

Geanakoplos, John and David Rappoport (2019) “Credit Surfaces, Economic Activity, and Monetary Policy,” working paper.

Gertler, Mark and Nobuhiro Kiyotaki (2010) “Financial Intermediation and Credit Policy in Business Cycle Analysis,” in *Handbook of Monetary Economics*, Vol. 3, Chap. 11.

Gilchrist, Simon and Egon Zakrajsek (2012) “Credit Spreads and Business Cycle Fluctuations,” *American Economic Review*, Vol. 102, pp. 1692–1720.

Greenstone, Michael, Alexandre Mas, and Hoai-Luu Nugyen (2020) “Do Credit Market Shocks Affect the Real Economy? Quasi-experimental Evidence from the Great Recession and "Normal" Economic Times,” *American Economic Journal: Economic Policy*, Vol. 12, pp. 200–225.

Gromb, Denis and Dimitri Vayanos (2018) “The Dynamics of Financially Constrained Arbitrage,” *Journal of Finance, Forthcoming*, Vol. 73, pp. 1713–1750.

Gross, David B. and Nicholas S. Souleles (2002) “Do Liquidity Constraints and Interest Rates Matter for Consumer Behavior? Evidence from Credit Card Data,” *The Quarterly Journal of Economics*, Vol. 117, pp. 149–185.

Guerrieri, Veronica and Guido Lorenzoni (2017) “Credit Crises, Precautionary Savings, and the Liquidity Trap,” *The Quarterly Journal of Economics*, Vol. 132, pp. 1427–1467.

He, Zhiguo and Arvind Krishnamurthy (2013) “Intermediary Asset Pricing,” *American Economic Review*, Vol. 103, pp. 732–70.

Holmstrom, Bengt and Jean Tirole (1997) “Financial Intermediation, Loanable Funds, and The Real Sector,” *The Quarterly Journal of Economics*, Vol. 112, pp. 663–691.

Huber, Kilian (2018) “Disentangling the Effects of a Banking Crisis: Evidence from German Firms and Counties,” *American Economic Review*, Vol. 108, pp. 868–98.

Ivashina, Victoria, Luc Laeven, and Enrique Moral-Benito (2020) “Loan Types and the Bank Lending Channel,” April.

Ivashina, Victoria and David Scharfstein (2010) “Bank lending during the financial crisis of 2008,” *Journal of Financial Economics*, Vol. 97, pp. 319–338.

Jaffee, Dwight M. and Thomas Russell (1976) “Imperfect Information, Uncertainty, and Credit Rationing,” *The Quarterly Journal of Economics*, Vol. 90, pp. 651–666.

Jaffee, Dwight and Joseph Stiglitz (1990) "Credit rationing," in B. M. Friedman and F. H. Hahn eds. *Handbook of Monetary Economics*, Vol. 2 of Handbook of Monetary Economics: Elsevier, Chap. 16, pp. 837–888.

Jermann, Urban and Vincenzo Quadrini (2012) "Macroeconomic Effects of Financial Shocks," *American Economic Review*, Vol. 102, pp. 238–271.

Jimenez, Gabriel, Steven Ongena, Jose-Luis Peydro, and Jesus Saurina (2012) "Credit Supply and Monetary Policy: Identifying the Bank Balance-Sheet Channel with Loan Applications," *American Economic Review*, Vol. 102, pp. 2301–26.

Justiniano, Alejandro, Giorgio Primiceri, and Andrea Tambalotti (2019) "Credit Supply and the Housing Boom," *Journal of Political Economy*, Vol. 127, pp. 1317–1350.

Khwaja, Asim Ijaz and Atif Mian (2008) "Tracing the Impact of Bank Liquidity Shocks: Evidence from an Emerging Market," *American Economic Review*, Vol. 98, pp. 1413–42.

Kiyotaki, Nobuhiro and John Moore (1997) "Credit Cycles," *The Journal of Political Economy*, Vol. 105, pp. 211–248.

Livshits, Igor, James MacGee, and Michele Tertilt (2007) "Consumer Bankruptcy: A Fresh Start," *American Economic Review*, Vol. 97, pp. 402–418.

Mian, Atif and Amir Sufi (2009) "The Consequences of Mortgage Credit Expansion: Evidence from the U.S. Mortgage Default Crisis*," *The Quarterly Journal of Economics*, Vol. 124, pp. 1449–1496.

Murfin, Justin (2012) "The Supply-Side Determinants of Loan Contract Strictness," *The Journal of Finance*, Vol. 67, pp. 1565–1601.

Paravisini, Daniel (2008) "Local Bank Financial Constraints and Firm Access to External Finance," *The Journal of Finance*, Vol. 63, pp. 2161–2193.

Peek, Joe and Eric S. Rosengren (1997) "The International Transmission of Financial Shocks: The Case of Japan," *American Economic Review*, Vol. 87, pp. 495–505.

Rampini, Adriano and S. Viswanathan (2019) "Financial Intermediary Capital," *Review of Economic Studies*, Vol. 86, pp. 413–455.

Roberts, Michael and Michael Schwert (2020) "Interest Rates and the Design of Financial Contracts," working paper.

Robles-Garcia, Claudia (2020) "Competition and Incentives in Mortgage Markets: The Role of Brokers," February.

Stiglitz, Joseph E and Andrew Weiss (1981) "Credit Rationing in Markets with Imperfect Information," *American Economic Review*, Vol. 71, pp. 393–410.

Williamson, Stephen D. (1987) "Costly Monitoring, Loan Contracts, and Equilibrium Credit Rationing," *The Quarterly Journal of Economics*, Vol. 102, pp. 135–145.

Appendix

A Figures

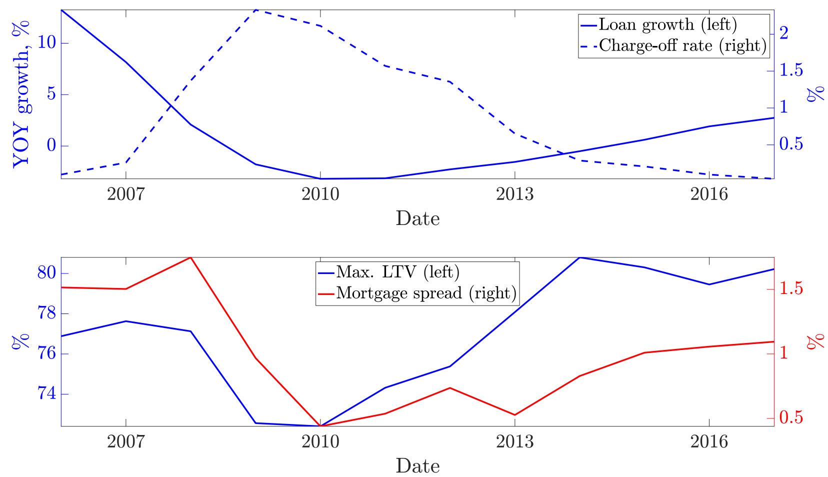

Figure 12: Bank balance sheets and mortgage loans. Sources: Federal Reserve Board, Fannie Mae, Freddie Mac.

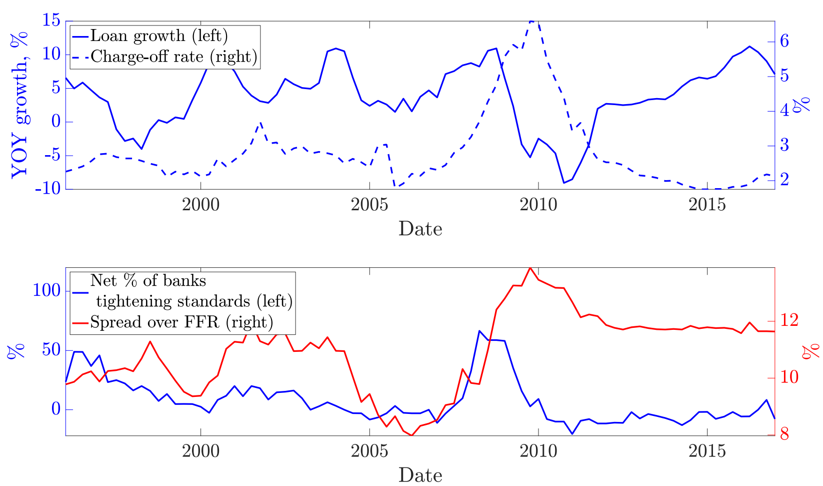

Figure 13: Bank balance sheets and credit card loans. Sources: Federal Reserve Board, SLOOS.

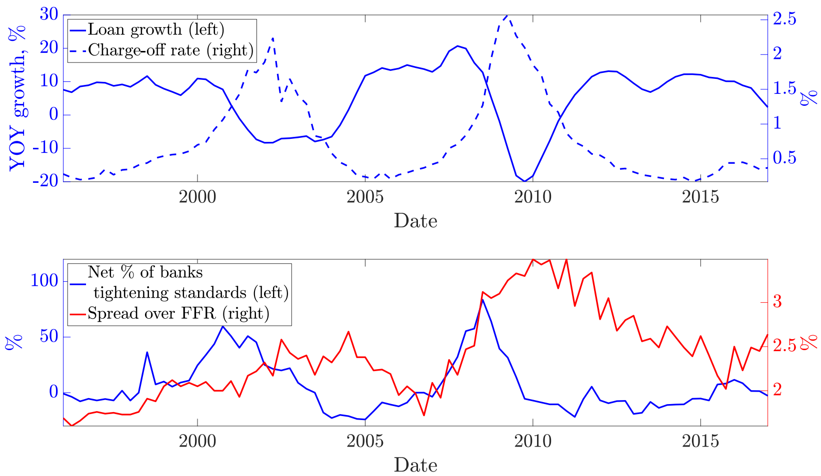

Figure 14: Bank balance sheets and commercial and industrial loans. Sources: Federal Reserve Board, SLOOS.

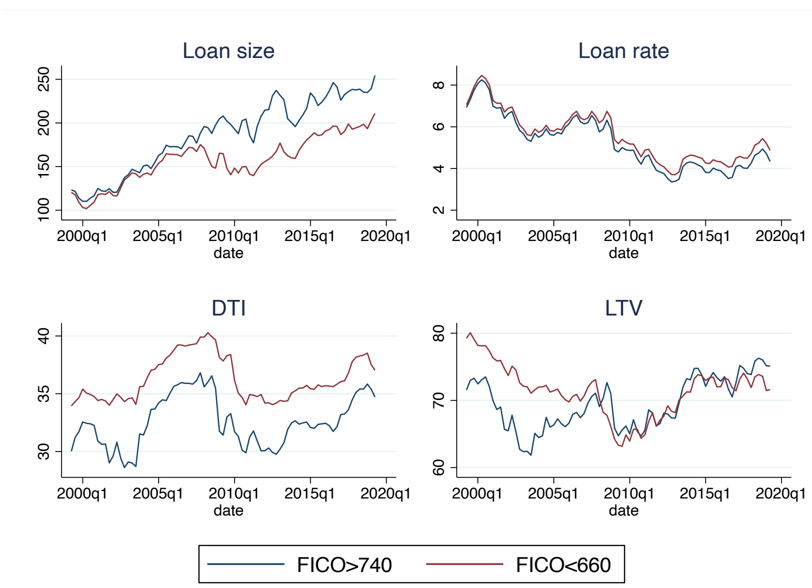

Figure 15: Conforming mortgage loan terms for low-risk (FICO above 750) and high-risk (FICO below 650) borrowers. Sources: Fannie Mac, Freddie Mae.

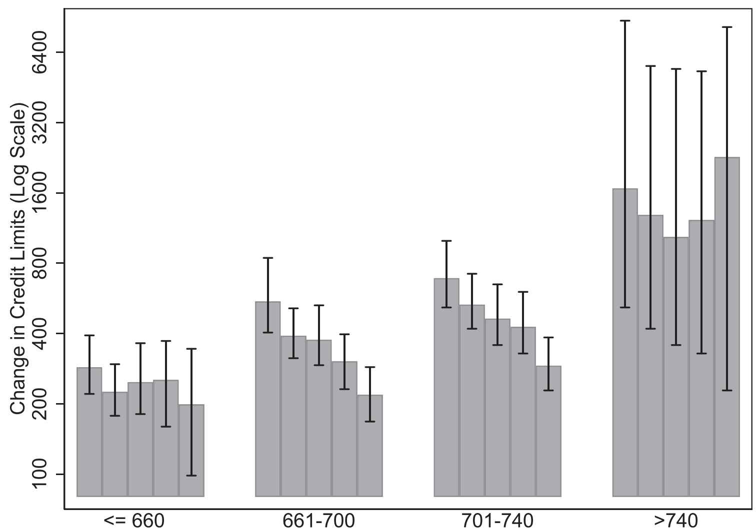

Figure 16: Source: Agarwal et al. 2018.

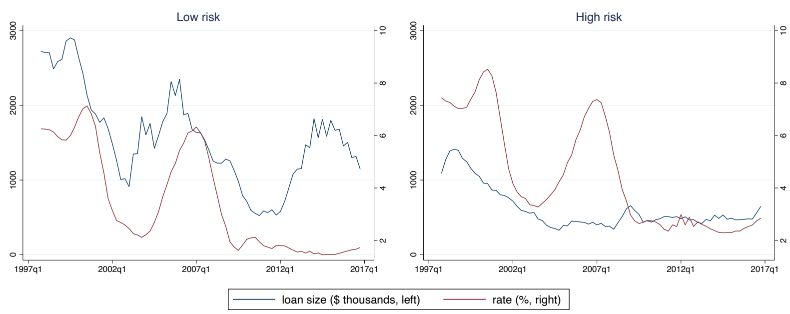

Figure 17: Loan size and loan rates for low and high risk short-term commercial and industrial loans. Source: Board of Governors of the Federal Reserve System (US).

B Proofs and derivations

B.1 Main Propositions

Proof of Proposition 1. Each bank solves

$$\max_{\{x^i,R^i,l^i,z^i\}} \int x^i \pi^i \left(l^i, R^i, z^i\right) di$$

$$\text{s.t.} \int x^i \rho^i l^i di \leq \bar{L} \tag{11}$$

$$V^i \left(l^i, R^i, z^i\right) \geq \overline{V}^i \tag{12}$$

Denote ν the multiplier on the bank lending constraint (11) and λ_i the one on borrower i's participation constraint (12). The first-order conditions with respect to l^i, R^i and x^i are respectively

$$x^i \pi^i_R + \lambda_i V^i_R = 0$$

$$x^i \pi^i_l + \lambda_i V^i_l - \nu \rho^i = 0$$

$$\pi^i - \nu \rho^i l^i = 0$$

Therefore banks must equalize the profit per risk-weighted dollar across loans

$$\frac{\pi^i}{\rho^i l_i} = \nu$$

Note that this nests the case in which the lending constraint is not binding and thus $\nu = 0$ and banks make zero profits.

In a symmetric equilibrium with $x^i = 1$ for all i, the price and quantity of each loan must solve

$$-\frac{V^i_l}{V^i_R} = \frac{\frac{\pi^i}{l^i} - \pi^i_l}{\pi^i_R}.$$

Using

$$\pi^i = \left[R^i \left(1 - \mu^i\right) - R^d\right] l^i$$

we have

$$\frac{\frac{\pi^i}{l^i} - \pi^i_l}{\pi^i_R} = \frac{R^i}{l^i} \frac{l^i \mu^i_l / \left(1 - \mu^i\right)}{1 - R^i \mu^i_R / (1 - \mu^i)}$$

$$= \frac{R^i}{l^i} \frac{\epsilon^i_{1-\mu}}{1 - \epsilon^i_{1-\mu}}$$

where the second line uses $\mu^i = \mu^i \left(R^i l^i, z^i\right)$.

Proof of Proposition 2. We fix one borrower type i and omit the superscripts i. Differentiating (3) yields

$$-\frac{l\tau_l}{\tau}\frac{dl}{l}-\frac{R\tau_R}{\tau}\frac{dR}{R}=\theta\left(\frac{dl}{l}+\frac{dR}{R}\right)(1+\tau)$$

where $\theta=\frac{Rl\epsilon'_{1-\mu}}{\epsilon_{1-\mu}}$. Using $1+\tau=\frac{1}{\epsilon_{1-\mu}-1}$ hence $\tau(1+\tau)=\frac{-\epsilon_{1-\mu}}{\left(1-\epsilon_{1-\mu}\right)^2}$ we get

$$\epsilon_l=\frac{-R\tau_R+\frac{Rl\epsilon'_{1-\mu}}{\left(1-\epsilon_{1-\mu}\right)^2}}{-l\tau_l+\frac{Rl\epsilon'_{1-\mu}}{\left(1-\epsilon_{1-\mu}\right)^2}}.$$

Letting $x=\frac{Rl\epsilon'_{1-\mu}}{\left(1-\epsilon_{1-\mu}\right)^2}$, we have $\frac{d}{dx}\left(\frac{-R\tau_R+x}{-l\tau_l+x}\right)=\frac{l\tau_l-R\tau_R}{(b+x)^2}$ hence if $\theta>0$ then $\epsilon_l>\frac{R\tau_R}{l\tau_l}$ if and only if $l\tau_l-R\tau_R>0$.

Proof of Proposition 3 and Proposition 4. We detail the case where R^f is fixed and $\overline{L}$ is shocked; the converse case follows exactly the same steps. First, the bank lending constraint implies

$$\sum_i dl^i=d\bar{L}$$

$$-\sum_i\frac{l^i}{\bar{L}}\epsilon_l^i d\log R^i=d\log\bar{L}$$

To obtain $d\log R^i$, rewrite (4) as

$$\frac{R^i\left[1-\mu^i\left(R^i\ell^i\left(R^i\right),z^i\right)\right]-R^f}{\rho^i}=\nu$$

and differentiate to get for i,j

$$\frac{d\log R^i}{\rho^i}R^i\left(1-\mu^i\right)\left[1-\epsilon^i_{1-\mu}\left(1-\epsilon^i_l\right)\right]=\frac{d\log R^j}{\rho^j}R^j\left(1-\mu^j\right)\left[1-\epsilon^j_{1-\mu}\left(1-\epsilon^j_l\right)\right]$$

Therefore

$$\begin{aligned}
-1 &= \sum_i \frac{l^i}{\bar{L}} \epsilon_l^i \frac{d\log R^i}{d\log \bar{L}} \\
&= \frac{l^i}{\bar{L}} \epsilon_l^i \frac{d\log R^i}{d\log \bar{L}} + \sum_{j\neq i} \frac{l^j}{\bar{L}} \epsilon_l^j \frac{d\log R^j}{d\log \bar{L}} \\
&= \frac{d\log R^i}{d\log \bar{L}} \frac{R^i\left(1-\mu^i\right)\left[1-\epsilon_{1-\mu}^i\left(1-\epsilon_l^i\right)\right]}{\rho^i} \left\{\sum_j \omega^j \epsilon_l^j \frac{\rho^j}{R^j\left(1-\mu^j\right)}\right\}
\end{aligned}$$

where $\omega^j = \frac{l^j}{\bar{L}}$ are loan weights and $\epsilon_l^j = \frac{\epsilon_l^j}{1-\epsilon_{1-\mu}^j\left(1-\epsilon_l^j\right)} = \frac{\epsilon_l^j}{\left(1-\epsilon_{1-\mu}^j\right)+\epsilon_l^j\epsilon_{1-\mu}^j}$ is the risk-adjusted elasticity. This rewrites

$$\frac{d\log R^i}{d\log \bar{L}} = -\frac{\rho^i}{R^i\left(1-\mu^i\right)} \frac{\epsilon_\ell^i}{\epsilon_l^i} \times \frac{1}{\sum_j \omega^j \epsilon_l^j \frac{\rho^j}{R^j\left(1-\mu^j\right)}}$$

which implies

$$\begin{aligned}
\frac{d\log l^i}{d\log \bar{L}} &= -\epsilon_l^i \frac{d\log R^i}{d\log \bar{L}} \\
&= \frac{\frac{\rho^i}{R^i\left(1-\mu^i\right)}\epsilon_\ell^i}{\sum_j \omega^j \epsilon_l^j \frac{\rho^j}{R^j\left(1-\mu^j\right)}}
\end{aligned}$$

and $\frac{d\log R^i}{d\log \bar{L}} = -\frac{1}{\epsilon_l^i}\frac{d\log l^i}{d\log \bar{L}}$.

Since $R^i\left(1-\mu^i\right) = R^f + \rho^i \nu$, for small $\left(\rho^i - \rho^j\right)\nu$ we have $R^i\left(1-\mu^i\right) \approx R^j\left(1-\mu^j\right)$ hence

$$\frac{d\log l^i}{d\log \bar{L}} \approx \frac{\rho^i \epsilon_\ell^i}{\sum_j \omega^j \rho^j \epsilon_l^j}.$$

B.2 Other Calculations

The effective default probability is lower than the actual one thanks to the positive recovery rate. Then

$$\mu'\left(Rl\right) = \kappa f\left(\hat{y}\left(Rl\right)\right) + \frac{1-\kappa}{\left(Rl\right)^2}\int_{y_{\min}}^{\hat{y}(Rl)} y dF\left(y\right)$$

while

$$\begin{aligned}
\epsilon_{1-\mu}\left(Rl\right) &= \frac{Rl\mu'\left(Rl\right)}{1-\mu\left(Rl\right)} \\
&= \frac{Rl\kappa\frac{f(Rl)}{F(Rl)} + \left(1-\kappa\right)\mathbf{E}\left[\frac{y}{Rl}|y\leq Rl\right]}{\frac{1-F(Rl)}{F(Rl)} + \left(1-\kappa\right)\mathbf{E}\left[\frac{y}{Rl}|y\leq Rl\right]}
\end{aligned}$$

If $\kappa = 0$ then $\epsilon_{1-\mu} \in [0, 1]$.

Pareto distribution.

- Suppose $\kappa = 0$ and
$$F(y) = 1 - \left(\frac{y_{\min}}{y}\right)^{\alpha}$$
for $\alpha > 0$ and $y > y_{\min}$,, then $f(y) = \alpha \frac{1-F(y)}{y}$ and
$$\epsilon_{1-\mu}(Rl) = \alpha \times \frac{1 - \left(\frac{y_{\min}}{Rl}\right)^{1-\alpha}}{1 - \alpha\left(\frac{y_{\min}}{Rl}\right)^{1-\alpha}} \in [0, 1]$$
When $y_{\min}$ is very small this is approximately α. When $\alpha = 1$ this is 0. More generally if $\kappa \leq \frac{1}{\alpha}$ then $\epsilon_{1-\mu} \in [0, 1]$. The general formula is
$$\begin{aligned}
\epsilon_{1-\mu} &= \alpha \times \frac{1 - \alpha\kappa - (1-\kappa)\left(\frac{y_{\min}}{Rl}\right)^{1-\alpha}}{1 - \alpha\kappa - \alpha(1-\kappa)\left(\frac{y_{\min}}{Rl}\right)^{1-\alpha}} \\
\epsilon'_{1-\mu} &= \frac{(1-\alpha)^2 \alpha(1-\kappa) y_{\min}(1-\alpha\kappa)\left(\frac{y_{\min}}{Rl}\right)^{\alpha}}{\left(\alpha(1-\kappa)y_{\min} - Rl(1-\alpha\kappa)\left(\frac{y_{\min}}{Rl}\right)^{\alpha}\right)^2} \\
\theta_{1-\mu} &= \frac{(1-\alpha)^2(1-\kappa)\, y_{\min}}{\left(\alpha\frac{1-\kappa}{1-\alpha\kappa}\left(\frac{y_{\min}}{Rl}\right)^{1-\alpha} - 1\right)\left(\frac{1-\kappa}{1-\alpha\kappa}\left(\frac{y_{\min}}{Rl}\right)^{1-\alpha} - 1\right)}
\end{aligned}$$

- With the power law example,
$$\theta_{1-\mu} = \frac{(1-\alpha)^2 Rl y_{\min}\left(\frac{y_{\min}}{Rl}\right)^{\alpha}}{\left(y_{\min} - Rl\left(\frac{y_{\min}}{Rl}\right)^{\alpha}\right)\left(\alpha y_{\min} - Rl\left(\frac{y_{\min}}{Rl}\right)^{\alpha}\right)}$$
is always positive. If $\alpha > 1$ the denominator is the product of two positive terms. If $\alpha < 1$ it's the product of two negative terms.

Examples of borrower utility $V(l, R)$.

- Starting with no risk hence no default:
  - Households
$$V(l, R) = u(y_0 + l) + \beta u(y_1 - Rl)$$
we see that
$$\tau(l, R) = \frac{u'(y_0 + l)}{\beta R u'(y_1 - Rl)} - 1$$

Consistent with the intertemporal wedge interpretation, $\tau \geq 0$ measures how constrained the household ends up since $u_0' = \beta R (1+\tau) u_1'$. Suppose CRRA utility with EIS σ, $u(c) = c^{1-1/\sigma}$. Then

$$\frac{R\tau_R}{l\tau_l} = \frac{(y_0 + l)(\sigma(y_1 - Rl) + Rl)}{l(Ry_0 + y_1)}$$

If $y_0 = 0$ then this simplifies to

$$\frac{R\tau_R}{l\tau_l} = \sigma \times \left(1 - \frac{Rl}{y_1}\right) + 1 \times \frac{Rl}{y_1}$$

This is a weighted average of σ and 1, so it's above 1 if and only if $\sigma \geq 1$.

– Firms

$$V(l, R) = f(k_0 + l) - Rl$$

then

$$\tau(l, R) = \frac{f'(k_0 + l)}{R} - 1$$

we have the same wedge interpretation: $f'(k_0 + l) = (1+\tau) R$. Then

$$\frac{R\tau_R}{l\tau_l} = \frac{f'(k_0 + l)}{-lf''(k_0 + l)}$$

the inverse curvature of the production function. So for $f(k) = Ak^\gamma$ we have $\frac{R\tau_R}{l\tau_l} = \frac{(k_0+l)}{l(1-\gamma)}$. With $k_0 = 0$,

$$\frac{R\tau_R}{l\tau_l} = \frac{1}{1-\gamma} \geq 1$$

A higher γ leads to a higher interest rate elasticity of the unconstrained loan demand.

- Once we add risk and default we need to compute V numerically:

  – Households with income shocks y_1:

$$V(l, R) = u(y_0 + l) + \beta \left[\int_{Rl}^{\infty} u(y_1 - Rl)\, dF(y_1) + \int_0^{Rl} u(\kappa y_1)\, dF(y_1) \right]$$

  – Firms with stochastic TFP shocks A, so that firm repays if and only if $Af(k_0 + l) \geq Rl$:

$$V(l, R) = \int_{\frac{Rl}{(y_0+l)^\gamma}}^{\infty} \left[A(y_0 + l)^\gamma - Rl\right] dF(A) + \kappa (y_0 + l)^\gamma \int_0^{\frac{Rl}{(y_0+l)^\gamma}} A dF(A)$$