INSEAD The Business School for the World®

Working Paper

2022/57/FIN

(Revised version of 2022/45/FIN)

Intermediary-Based Loan Pricing

Pierre Mabille
INSEAD, pierre.mabille@insead.edu

Olivier Wang
New York University, olivier.wang@nyu.edu

November 18, 2022

How do shocks to banks transmit to loan terms faced by borrowers on different loan markets? In our model of multidimensional contracting between heterogeneous risky borrowers and intermediaries with limited lending capacity, loan terms depend on loan demand elasticities and default elasticities. These two sufficient statistics predict how the cross-section of loan terms and bank risk react to changes in capital and funding costs. Using empirical estimates, they explain the heterogeneous transmission of shocks across loan markets and borrower risk categories. Accounting for non-price loan terms is important for dynamics because their endogenous response can increase the persistence of credit crises.

History: First draft: November 2021.
JEL classification: G12; G21; G28; G33; G51.
Keywords: Banks; Credit Supply; Credit Risk; Covenants; Financial Crisis; Intermediary Asset Pricing; Sufficient Statistics.

Acknowledgments
We thank Viral Acharya, Arpit Gupta, John Kuong, Alexi Savov, Philipp Schnabl, Alp Simsek, Venky Venkateswaran, and our discussants Laurent Calvet, William Diamond, Sebastian Doerr, and Fabrice Tourre for comments. This paper bene ted from participants at the WFA Meeting, the SGF Conference, the CEBRA Annual Meeting, the GdRE Symposium on Money Banking and Finance, the INSEAD Finance Symposium, and from seminars at NYU Stern, INSEAD, BI Oslo, the Federal Reserve Board, Banque de France, and the Junior European Finance Seminar. We thank the Alfred P. Sloan Foundation and the Macro Financial Modeling (MFM) project for supporting this research.

1 Introduction

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

Variations in loan terms are well documented in the cross section and over time. Within a loan market, they differ between borrowers and they depend on the financial health of banks.[1] However, it is still unclear why changes in banks' balance sheets are transmitted differently to the loan terms faced by borrowers *across* loan markets. While credit supply goes hand in hand with banks' financial health across loan markets, shocks to banks have heterogeneous effects on borrowers with different credit risks in each market. In particular, during the credit boom that preceded the Great Recession in the United States, growing 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). Understanding the heterogeneous transmission of bank shocks to borrowers across loan markets is important not only because banks lend on multiple markets, but also because it is crucial for the allocation of credit supply between agents with different needs to borrow. Despite the central role of credit conditions in macro-finance, what is missing is a model that can simultaneously explain the transmission of bank shocks to the price and non-price terms of loans faced by borrowers across loan markets.

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. Their balance sheets affect credit supply and asset prices as in specialized asset markets (He and Krishnamurthy 2013, Brunnermeier and Sannikov 2014). The novelty of our approach, which is specific to credit markets, is to allow banks to use both interest rates and non-price terms to control borrowers' default risk when this risk is endogenous to the terms of the contract. We show that two sufficient statistics, the interest-rate elasticity of borrowers' loan demand and the elasticity of the default rate to the size of the repayment, 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 is heterogeneous across loan markets (e.g., mortgages vs. credit cards), across borrower

[1] See, e.g., Agarwal et al. (2010), Ramcharan et al. (2016), Adelino et al. (2016), Chakraborty et al. (2018), Chakraborty et al. (2020),Jorda et al. (2021),Chodorow-Reich and Falato (2021), and Diamond and Landvoigt (2021).

risk categories within a market (e.g., safe vs. risky), and for banks with different balance sheets (e.g., well vs. poorly capitalized). Moreover, we highlight an important implication of accounting for changes in non-price terms for the dynamics of credit crises. The persistence of credit crises is endogenous, and it is higher on loan markets where non-price terms are tightened more than interest rates, since lower rates slow down the recapitalization of banks after negative shocks.

Our analysis is motivated by key facts from the empirical literature on bank shocks for which a unified explanation is missing. First, banks face financial constraints, and the more losses (gains) on their assets they make, the less (more) they lend. Second, banks jointly set interest rates and non-price terms to control credit supply. Changes in bank heath are associated with changes in loan-to-value limits for mortgages, lending standards for credit cards and C&I loans, and interest rates on all three markets. 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 for C&I loans for which non-price terms such as covenants also adjust. Despite separate interest in these markets, there is a limited understanding of what explains their differences and commonalities.

Our model of multidimensional loan contracting leads to 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: either in economies with endogenous credit rationing and interest rates but fixed loan quantities (Stiglitz and Weiss 1981), or in Walrasian economies where interest rates adjust to clear loan markets but borrowing constraints are exogenous (Holmstrom and Tirole 1997). Similarly, borrowing constraints are crucial in many macro-finance models, but their changes are usually assumed to be exogenous (e.g., Bernanke et al. 1999, Jermann and Quadrini 2012, Favilukis et al. 2017, Guerrieri and Lorenzoni 2017).

In the model, the 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. Loan markets differ from generic asset markets such as stocks in two ways. 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 their loan repayment value is too high. The elasticity α of borrowers' default probability to the repayment value captures this channel. Our model accommodates special cases with moral hazard, adverse selection, liquidity or strategic default, and unsecured or collateralized lending. Second, we depart from a Walrasian setting where the loan repayment value is linear in a single interest rate that clears loan markets. Rather, banks offer multidimensional non-linear contracts with price and non-price terms which capture their effort to screen and monitor borrowers. Non-price terms arise from the feedback between

loan prices and loan payoffs. Since increasing interest rates also increases default risk, non-price terms allow banks to manage this risk while holding rates fixed, and α determines their relative adjustment. The interest rate on a given loan compensates banks for their cost of funds (risk-free rate), borrower default risk (credit risk premium), and the tightness of their capacity constraint (excess loan premium). The problem of the bank generates a loan contract curve, which relates borrower and lender characteristics to loan terms. It provides a comprehensive measure of credit tightness based on all loan terms and lender health, which generalizes credit surfaces focusing on borrower risk (e.g., Geanakoplos 2010). It is described by the elasticity ϵ of loan demand when borrowers are endogenously constrained by multiple loan terms.

We use the model to understand how shocks to banks' lending capacity and cost of funds affect equilibrium loan terms. When their balance sheets deteriorate, banks trade off tightening price and non-price terms. Our second contribution is a set of pass-through formulas highlighting the role of the demand elasticity ϵ and the default elasticity α as two sufficient statistics which determine how bank shocks are transmitted to the cross-section of loan terms. The benefit of our approach is that even though they are endogenous, it suffices to compute these two moments identified in borrower- and loan-level data to analyze the effect of credit crises.

We begin by analyzing the impact of credit supply shocks, modeled as changes in banks' lending capacity. Our results explain why credit expansions are transmitted to risky borrowers on loan markets with a low elasticity of loan demand $\epsilon < 1$ such as mortgages, and to safe borrowers on markets with a high elasticity $\epsilon > 1$ such as credit cards.[2] A positive credit supply shock increases loan quantities and lowers interest rates on both markets, but markets with a low ϵ see a smaller increase in quantities and a larger decrease in rates, which lead to a decrease in loan values to be repaid. The default elasticity α determines the resulting change in credit risk. In markets with a low ϵ, the decrease in repayment values induces a small decrease in credit risk for safe borrowers with a low α, but a large decrease for risky borrowers with a high α. The opposite happens in markets with a high ϵ where repayment values increase. As a result, the returns from lending to risky borrowers increase on markets with a low ϵ, while they increase for safe borrowers on markets with a high ϵ. These changes create an opportunity for banks to increase their total profits, which leads them to increase lending to risky borrowers on inelastic markets and to safe borrowers on elastic markets.

Next, we study the transmission of shocks to banks' cost of funds due to changes in monetary policy. Our results explain why the transmission of monetary policy is dampened when bank balance sheets are impaired, and highlight that the strength of the bank lending channel is het-

[2]Estimates for the interest rate-elasticity of mortgage debt are generally lower than 1 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). In contrast, estimates for the elasticity of credit card debt are always greater than 1 (e.g., 1.3 in Gross and Souleles 2002).

erogeneous across loan markets. The pass-through of the policy rate to interest rates and loan sizes depends on the elasticity of loan demand ϵ and on banks' lending capacity. Unconstrained banks transmit changes in the policy rate more than one-for-one to low-elasticity borrowers such as mortgages, but this transmission is weaker for high-elasticity borrowers such as credit cards, resulting in sticky interest rates. Low policy rates also give rise to covenant-lite lending for low-elasticity borrowers, but to tighter covenants for high-elasticity borrowers. 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 of a policy rate hike, but also from the positive effects of a rate cut.

What are the implications for the credit risk of the entire portfolio of bank loans? First, a shock to credit supply or banks' costs of funds changes both price and non-price terms within a given pool of borrowers. Second, it reallocates bank lending towards specific borrowers. Shocks such as policy rate cuts increase the credit risk of borrowers with a high loan demand elasticity ϵ, and decrease it for borrowers with a low ϵ. When high-elasticity borrowers also have a high risk α, the net effect is an increase in the total credit risk of banks.

We conclude by highlighting the importance of accounting for changes in non-price terms when analyzing how credit crises play out over time. In a dynamic version of the model, changes in interest rates due to bank shocks feed back into banks' capital through retained earnings, and determine the tightness of their lending capacity constraints in future periods. Our third contribution is to show that the persistence of credit crises is endogenous, and that it is larger on loan markets where non-price terms adjust more than interest rates in response to shocks. The two elasticities ϵ and α drive the dynamics of crises. Negative bank shocks result in higher interest rates and tighter non-price terms. The more non-price terms are tightened, the less interest rates need to increase to compensate banks for default risk. In turn, lower rates decrease the profits earned by constrained banks, which slows down their recapitalization and makes the crisis more persistent. These results shed light on the slow recovery of credit markets after the Great Recession. We illustrate them in a numerical example based on the U.S. mortgage and credit card markets. Interestingly, diversifying lending towards high-elasticity borrowers such as credit cards increases the persistence of crises for banks which specialize in low-elasticity borrowers such as mortgages. Conversely, diversifying lending towards low-elasticity borrowers allows banks which specialize in high-elasticity borrowers to implement cross-subsidization between the two markets. In that case, the larger initial increase in interest rates on low-elasticity borrowers allows banks specialized in high-elasticity borrowers to recapitalize themselves more quickly after a negative shock.

Related literature. Our work contributes to the microeconomic literature on credit markets and the macro-finance literature on the dynamics of credit crises. We connect two separate approaches in a model of multidimensional loan contracting which jointly endogenizes the price and non-price terms of loans: models with endogenous credit rationing but fixed loan quantities (Stiglitz and Weiss 1981); and models with endogenous interest rates but fixed credit limits (Holmstrom and Tirole 1997). We focus on the intensive margin of rationing whereby risky borrowers face tighter non-price terms rather than being excluded from credit markets altogether, and generalize existing results which are surveyed in Jaffee and Stiglitz (1990).

As in the data, loan terms endogenously depend on both borrower and lender characteristics (Section 2). Our analysis therefore complements models with heterogeneous households where the interest rate on unsecured consumer loans depends on borrower characteristics (Chatterjee et al. 2007, Livshits et al. 2007), and models of the credit surface where an increase in borrower credit risk leads lenders to tightening collateral requirements (Geanakoplos 2010). We add to these settings by showing that lenders' financial conditions also affect the terms of lending, as in Diamond and Landvoigt (2021) who study the credit surface of borrowers in a rich quantitative model of the mortgage market. Our contribution is theoretical: we analytically explain how multiple loan terms react to banks shocks as a function of sufficient statistics on the underlying loan markets. This general formulation delivers a new analysis of the transmission of bank shocks *across* different loan markets in a unified setting, and allows us to study the implications for the persistence of credit crises on these markets. Estimates for the elasticities of loan demand and default rates to interest rates are well identified in empirical settings which include survey data (Fuster and Zafar 2021), regression discontinuity designs (Fuster and Willen 2017, Best et al. 2019), and structural models (Buchak et al. 2020, Robles-Garcia 2020, Benetton 2021). In our model, we show that they depend on structural parameters which affect borrowers' demand and default. Across markets, their variation arises from moral hazard, adverse selection, and borrowers' liquidity constraints (Adams, Einav and Levin 2009, Einav, Jenkins and Levin 2012).

The novelty of our approach is to extend the intermediary asset pricing framework (He and Krishnamurthy 2013, Brunnermeier and Sannikov 2014, Gromb and Vayanos 2018) to credit markets, by modeling the effect of banks' capacity constraints and borrower default risk on interest rates and the non-price terms of loans. Jointly endogenizing loan terms gives rise to an excess loan premium as in recent data (Gilchrist and Zakrajsek 2012), which compensates banks when their capacity constraints bind. It also generates a new transmission mechanism of shocks as non-price terms affect the dynamics of credit crises. As in canonical macro-finance models, a credit crunch arises from a decrease in banks' net worth (Gertler and Kiyotaki 2010, Rampini and Viswanathan 2019). In contrast with these models, bank losses do not lead to a large increase in interest rates which would quickly recapitalize banks when non-price terms are also tightened. Instead, their

tightening leads to a lower increase in in interest rates, which slows down the recapitalization of banks and makes the credit crisis endogenously more persistent. This novel mechanism sheds light on the slow recovery of credit markets after the Great Recession (Justiniano et al. 2019).

Outline. The rest of the paper is organized as follows. Section 2 discusses motivating evidence on heterogeneity in loan terms. Section 3 presents the model of multidimensional loan contracting. Section 4 analyzes the transmission of credit supply and monetary policy shocks to the cross-section of loan terms. Section 5 studies the dynamics of credit crises, which depends on the interaction between the cross-section of loan terms and banks' balance sheets. We describe the empirical implications of our results in both sections. Using a calibrated version of the model for the U.S. mortgage and credit card markets, we show that they can explain key features of these markets: the heterogeneous transmission of bank shocks across borrowers and the persistence of credit crises. Section 6 concludes.

2 Evidence on Heterogeneity in Loan Terms

The goal of our model is to explain key facts on heterogeneity in loan terms and the transmission of bank shocks to the main classes of loans – mortgages, credit cards, and C&I loans — as well as to understand their implications for the persistence of credit crises. In this section, we briefly discuss the main patterns documented in the empirical literature.

Bank health and loan growth. Banks face financial constraints, hence the health of their balance sheet matters for how much they can lend. As shown by the extensive literature on bank shocks, banks tighten credit supply following losses on their assets (e.g., Peek and Rosengren 1997, Khwaja and Mian 2008, Schnabl 2012, Chodorow-Reich 2014, Huber 2018, Amiti and Weinstein 2018, Greenstone et al. 2020). Conversely, they expand lending following positive shocks (e.g., Gilje et al. 2016).

Interest rates and non-price terms. Credit supply shocks affect the interest rates, quantity limits, and other non-price terms which are simultaneously set by banks to control default risk and the supply of credit. Reductions in lending due to bank shocks take the form of both higher interest rate spreads and tighter lending standards. Our model explains why banks use non-price terms to limit lending for a given interest rate. Such terms include debt covenants for firms (Murfin, 2012; Chodorow-Reich and Falato, 2021), LTV and credit card limits (Agarwal et al., 2018), and auto loans terms for households (Benetton et al., 2022). Shocks to banks generate heterogeneous responses for these loan terms across different credit markets. This is the case

when their balance sheets deteriorate or improve (Khwaja and Mian 2008, Chakraborty et al. 2018, Ivashina et al. 2020), and when their cost of funds changes because of monetary policy (Jimenez et al. 2012, Chakraborty et al. 2020), or access to external finance (Paravisini 2008, Ivashina and Scharfstein 2010).

Heterogeneous transmission of shocks to borrowers. Our model seeks to explain a key feature of the data: the transmission of bank shocks to risk and safe borrowers is heterogeneous across loan markets. Credit supply shocks are transmitted to risky borrowers in mortgage markets, and to safe borrowers in credit card markets. In particular, Mian and Sufi (2009) and Justiniano et al. (2019) emphasize the expansion and subsequent contraction of credit supply to subprime borrowers before the Great Recession. While mortgage sizes grew for both risky and safe borrowers, they fell more deeply and persistently for risky borrowers despite similar changes in interest rates.

In contrast, expansions in credit card lending are mostly transmitted to safe borrowers. Agarwal et al. (2018) show that in response to a decrease in banks' cost of funds, households with a low FICO score who are risky but have a high propensity to borrow face a smaller increase in credit limits than households with a high FICO score who are safe. Corporate loan sizes are also more volatile for safe borrowers, despite similar changes in interest rates across risk categories. This is illustrated in Appendix Figure 11, which plots 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. These changes reflect that most of the adjustment in corporate lending occurs through non-price terms. 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. We consider a static setting with one-period loans characterized by an interest rate R^i (price term), a loan amount l^i (quantity limit), and a vector of non-price terms $z^i = \left(z_k^i\right)_k$.[3] The expected profit on a loan contract to borrower i is

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

The term $\mu^i\left(R^i l^i, z^i\right) R^i l^i$ is the bank's expected loss on its loan to borrower i, as a function of the loan terms, where μ^i denotes the product of the default probability with one minus the recovery rate on the loan to borrower i. We call μ^i the *effective default probability* because expected profits can be rewritten as

$$\pi = \left[R\left(1 - \mu^i\right) - R^f\right] l^i. \tag{2}$$

With a positive recovery rate upon default, μ^i will be lower than the actual default probability. We work with μ^i as a primitive as it reflects the multiple channels through which endogenous default risk affects the bank's expected profit, such as liquidity default, adverse selection, or debt overhang.

We make the simplifying assumption that μ depends on the product Rl instead of R and l separately. For the one-period loan contracts that we consider, this means that R and l only affect the expected loss through the face value of the debt to be repaid Rl; we give several examples below in which this is indeed the case. The vector z^i captures other non-price terms, such as covenants and loan documentation requirements, which allow the bank to reduce default risk and potentially improve the recovery value in case of default.

Effective default probability. The benefit of our general specification using the effective default probability μ is that it can nest several environments with ex-ante or ex-post asymmetric information that makes the expected loss endogenous to loan terms through the default probability, the loss given default, or both. Therefore, our results do not depend on the underlying microfoundation that determines the dependence of μ on Rl. Appendix B details three examples of the model to illustrate how microeconomic frictions map to μ: liquidity default for unsecured loans, liquidity and strategic defaults for collateralized loans, and adverse selection.

Capacity constraint. To analyze the transmission of credit supply shocks, we study banks with

[3] Allowing for longer loan maturities that also respond endogenously to credit supply conditions would be an interesting extension that we leave for future work.

a capacity constraint on total lending

$$\int \rho^i l^i di \leq \overline{L} \tag{3}$$

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 specific 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.

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

Assumption 1. *For each i, $V^i_R < 0$, $V^i_z \leq 0$, and the marginal utility of additional borrowing is lower at higher interest rates: $V^i_{lR} < 0$.*

Assumption 1 holds in most settings, such as with three types of microeconomic frictions in Appendix B. Tighter non-price terms (higher z^i) improve the recovery value for lenders by lowering μ^i, but they imply that borrowers must give up control rights more often (for instance in the case of covenants), which is costly for them ($V_z \leq 0$). The last part of the assumption is standard: it simply states in our context 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.

3.2 Two sufficient statistics: α and ϵ

We first define the default elasticity α, one of the two sufficient statistics from which our results derive.

Definition 1. *The default elasticity is*

$$\alpha(Rl, z) = \frac{Rl\mu'(Rl, z)}{1 - \mu(Rl, z)}. \tag{4}$$

The higher α, the more credit risk depends on the terms of the loan contract. It is more convenient to use α defined in (4) than the elasticity of μ, but they capture the same idea. We make the following assumption.

Assumption 2. *The default elasticity satisfies* $0 \leq \alpha < 1$ *everywhere.*

The first inequality in Assumption 2 implies that the effective default probability μ is increasing in the loan repayment value Rl. The second inequality ensures that bank revenues are increasing in the loan's interest rate. Banks' zero profit curves $\pi(R, l, z) = 0$, and their iso-return curves defined as a constant $\frac{\pi(R,l,z)}{l}$, are always upward sloping since they have a slope $\frac{dR/R}{dl/l} = \frac{\alpha}{1-\alpha}$.

We rule out credit rationing at the extensive margin as in 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 credit rationing at the intensive margin, 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.

The solution to equation (12) defines a "loan contract curve" $\ell^i\left(R^i, \boldsymbol{z}^i\right)$ which is a generalization of a simple loan demand curve.

Then, we define the elasticity of the loan contract curve with respect to the interest rate, which is the second sufficient statistic from which our results derive.

Definition 2. *The interest-rate elasticity of loan demand* ℓ *is*

$$\epsilon^i = -\frac{R^i}{\ell^i}\frac{d\ell^i}{dR^i}. \tag{5}$$

It is a central object that will determine the response of the cross-section of loan terms to bank shocks. By analogy, we call it a demand elasticity, as it denotes the change in quantity borrowed for a given change in the loan rate.

We can relate ϵ to the more familiar elasticity ϵ_u of the unconstrained loan demand function $\ell_u(R, \boldsymbol{z})$ that solves $V_l^i(l, R, \boldsymbol{z}) = 0$. It is more common to think of the unconstrained loan demand elasticity as a primitive, but in many settings ϵ is indeed what can be estimated using exogenous variation in loan rates, and we will treat ϵ as the primitive sufficient statistic. Nevertheless it is useful to understand the relationship between ϵ and ϵ_u:

Proposition 1. *The interest rate elasticity* ϵ *can be decomposed as a weighted average*

$$\epsilon = \omega \cdot \epsilon_u + (1 - \omega) \cdot 1 \tag{6}$$

where $\omega = \frac{1}{1+\frac{R\alpha'}{(1-\alpha)^2|\tau_l|}}$ *and* $\epsilon_u = \frac{R\tau_R}{l\tau_l}$ *is the interest elasticity of the unconstrained loan demand.*

Proof. See Appendix C.1. □

The elasticity ϵ is a weighted average of 1 and the unconstrained demand elasticity ϵ_u. If the effective default probability μ is constant such that $\alpha = 0$ (i.e., exogenous default risk), then $\omega = 1$ and ϵ is simply the unconstrained elasticity ϵ_u. If $\alpha' > 0$ (i.e., the default elasticity increases with debt), then $\epsilon > \epsilon_u$ if $\epsilon_u < 1$ and conversely $\epsilon < \epsilon_u$ if $\epsilon_u > 1$.

On loan markets where the elasticity of loan demand is low ($\epsilon < 1$), such as mortgages, endogenous default risk α makes the contract curve ℓ more elastic than the unconstrained demand ℓ_u, bringing the elasticity ϵ 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, hence the loan size adjustment, is larger for riskier borrowers with a high α. In response to a small increase $dx = \frac{Rl\alpha'}{(1-\alpha)^2}$, the elasticity ϵ 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 ($\epsilon > 1$). 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 closer to one. Hence a positive term $\frac{Rl\alpha'}{(1-\alpha)^2}$ due to endogenous default risk $\alpha > 0$ acts as an elasticity dampener which brings back ϵ closer to one.

3.3 Equilibrium: Bertrand-Nash with capacity constraints

Banks are risk-neutral, perfectly competitive, and subject to capacity constraints on lending.[4] For each borrower i, banks 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

[4] We abstract from bank risk aversion as studied by, e.g., Diamond and Landvoigt (2021).

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 $\bar{L}$ takes as given the set of banks posting the best contracts $\Omega^i = \arg\max_{b' \neq b} V^i(C^i_{b'})$ for each borrower i, and chooses contracts C^i and the probabilities x^i to grant loans to borrowers i to solve

$$\max_{\{x^i, R^i, l^i, z^i\}} \int x^i \pi^i \left(l^i, R^i, z^i\right) di \tag{7}$$

$$\text{s.t.} \int x^i \rho^i l^i di \leq \bar{L} \tag{8}$$

$$V^i \left(l^i, R^i, z^i\right) \geq \overline{V}^i \tag{9}$$

$\pi^i \left(l^i, R^i, z^i\right)$ is the profit per loan conditional on take-up for borrower i. l^i captures the intensive margin of credit, and x^i captures the extensive margin. $\overline{V}^i = \max_{b' \neq b} V^i(C^i_{b'})$ is the outside option of borrower i.

In our formulation we assume that the bank can control the probability of lending through x^i. This is true conditional on the participation constraint $V^i \left(l^i, R^i, z^i\right) \geq \overline{V}^i$ holding. Borrower i would not apply if the bank offered a contract with a strictly lower utility than the outside option $\overline{V}^i$, and the resulting probability of lending would be zero irrespective of x^i. Since a bank can always reach the same outcome by setting $x^i = 0$ and deciding to deny credit to borrower i, it is without loss of generality to restrict attention to contracts that respect the participation constraint.

Definition 3. *An equilibrium is an optimal strategy $i \mapsto \left\{x^i_b, C^i_b\right\}$ for each bank b such that borrowers optimize:*

$$\overline{V}^i = \max_{b'} V^i \left(C^i_{b'}\right) \tag{10}$$

and markets clear:

$$1 = \int_{\arg\max_{b'} V^i\left(C^i_{b'}\right)} x^i_b db. \tag{11}$$

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

Proposition 2. *In a symmetric equilibrium,*

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

$$\tau^i \left(l^i, R^i, z^i\right) = \frac{\alpha^i \left(R^i l^i, z^i\right)}{1 - \alpha^i \left(R^i l^i, z^i\right)} \tag{12}$$

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.,

$$\frac{\pi^i\left(l^i, R^i, \boldsymbol{z}^i\right)}{\rho^i l^i} = \frac{\pi^j\left(l^j, R^j, \boldsymbol{z}^j\right)}{\rho^j l^j} \equiv \nu \tag{13}$$

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

iii. Other optimal non-price terms $\boldsymbol{z}$ satisfy

$$-R^i l^i \frac{\partial \mu^i}{\partial z_k} = \frac{V^i_{z_k}}{V^i_R} l^i \left(1-\mu^i\right)\left(1-\alpha^i\right) \tag{14}$$

Proof. See Appendix C.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.

Part (i) of Proposition 2 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 their default elasticity α. If there is no endogenous default risk and $\alpha = 0$, then borrowers are not constrained and the interest rate does not depend on their loan size. The riskier borrowers are (the higher $\alpha > 0$), 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., $V^i_z = 0$, then banks minimize the effective default probability by setting $-R^i l^i \frac{\partial \mu^i}{\partial z_k} = 0$. This defines an optimal tightness $\hat{z}^i$, where R^i and l^i are endogenously determined by (12) and

(13). 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 $\boldsymbol{z}^i$ as given for now, the unconstrained loan quantity $l^{i,*}$ is defined as solving (12) together with the zero-profit conditions $\pi^i\left(l^i, R^i, \boldsymbol{z}^i\right) = 0$. Banks are said to be *unconstrained* if given unconstrained loan quantities $\left\{l^{i^*}\right\}$, their constraint (3) is slack, and in that case $\nu = 0$. Otherwise, banks are *constrained*, and must make a positive profit per dollar $\nu > 0$.

3.4 Implications

We derive four implications of our setting where endogenous default risk affects interest rates and non-price terms, which are key for the transmission of bank shocks.

Implication 1: When do banks constrain borrowers? Equation (12) 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.

Implication 2: Excess loan premium. Interest rates can be decomposed into borrower-level and bank-level factors. A first-order approximation of equilibrium interest rates satisfies

$$\begin{aligned} & \log R_i = \log\left(R^f + \rho^i \nu\right) - \log\left(1 - \mu_i\right) \\ \Leftrightarrow \qquad & r^i \approx r^f + \mu^i + \rho^i \nu \end{aligned} \tag{15}$$

Equation (15) shows that the loan rate is increasing in banks' cost of funds, the effective default probability μ^i, and the tightness of banks' capacity constraints weighted by the borrower's risk-weight. The latter generates an excess loan premium. Loan rates net of the premium $r^i - \rho^i \nu$ are

actuarially fair and exactly compensate banks for individual default risk. Importantly, condition (15) is an equilibrium outcome 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 quantities l^i by more for riskier borrowers i, as measured by their default elasticity α^i.

Implication 3: Interest rates do not fully capture credit conditions. With endogenous default risk, the equilibrium interest rate can be a non-monotonic function of the borrower's income risk, so that rates alone do not capture credit conditions. The full loan contract needs to be specified to measure how tight credit is. For instance, 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.

Implication 4: Bank-dependent credit surface The previous implication itself implies that a credit surface can be interpreted as a special case of the multidimensional loan contract curve ℓ, which depends on banks' capacity constraint and cost of funds.

Credit surfaces 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 loan contract curve ℓ generalizes credit surfaces to include multiple non-price terms (l, z) in addition to interest rates R. Furthermore, ℓ highlights that in addition to borrower characteristics, these terms depend on *lender* characteristics, captured by banks' capacity constraint $\overline{L}$ and the excess loan premium v. The effect of shocks on the resulting credit surfaces can be characterized with the two sufficient statistics ϵ and α.

Borrowers are constrained at R if they would like to borrow more at the prevailing interest rate, such that $V_l\left(\ell\left(R\right), R\right) > 0$. This may be the case even if lenders are unconstrained, with $\ell\left(R\right) = L^* < \overline{L}$. Equilibrium condition (12) implies that borrowers are unconstrained only if $\alpha\left(\ell\left(R\right) R\right) = 0$. If V is separable, $V\left(l, R\right) = u\left(l\right) - w\left(Rl\right)$, then we can interpret the unconstrained condition $V_l = 0$ as a standard Euler equation

$$\frac{u'\left(l\right)}{w'\left(Rl\right)} = R \tag{16}$$

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 2021). 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.[5]

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

$$\max_{l} u\left(l\right) - w\left(R\left(l|\bar{L}\right) l\right) \tag{17}$$

hence

$$\frac{u'\left(l\right)}{w'\left(R\left(l|\bar{L}\right) l\right)} = R\left(l|\bar{L}\right)\left[1 + \frac{lR'\left(l|\bar{L}\right)}{R\left(l|\bar{L}\right)}\right] \tag{18}$$

Combining with (12), the function $R\left(l|\bar{L}\right)$ solves the differential equation (for each type)

$$\frac{d\log R\left(l|\bar{L}\right)}{d\log l} = \frac{\alpha\left(lR\left(l|\bar{L}\right)\right)}{1-\alpha\left(lR\left(l|\bar{L}\right)\right)} \tag{19}$$

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.

4 Transmission of Bank Shocks to Loan Terms

This section studies the heterogeneous transmission of shocks to banks' balance sheets to the loan terms of borrowers with different credit risks across loan markets with different loan demand elasticities. We analyze credit supply shocks that affect banks' lending capacity and monetary policy shocks that affect their cost of funds. We focus on the responses of interest rates R, quantity limits l, non-price terms z, and the resulting default risk born by banks. We show that a calibration of our model can explain the heterogeneous transmission of the U.S. credit boom of the 2000s to the mortgage and the credit card markets.

4.1 Credit supply shocks

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-

[5]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).

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 this heterogeneous transmission with a formula that governs how loan quantity limits and interest rates vary across borrowers in response to a shock to bank lending capacity $\overline{L}$. The formula relies on the following elasticity, which can be constructed for each borrower from the previous two sufficient statistics ϵ and α:

Definition 4. *Let the risk-adjusted elasticity of loan demand be*

$$\widetilde{\epsilon}^i = \frac{\epsilon^i}{(1-\alpha^i)+\epsilon^i\alpha^i}. \tag{20}$$

A key property is that the risk-adjusted elasticity is closer to 1 than ϵ:

Lemma 1. *$\widetilde{\epsilon}^i$ lies between 1 and ϵ^i.*

4.1.1 Bank to borrower transmission

The next proposition is one of the main results of the paper. It shows that the risk-adjusted elasticity of loan demand governs the responses of loan terms to changes in bank lending capacity:

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

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

$$\begin{aligned}
\frac{d\log l^i}{d\log \overline{L}} &= \frac{\widetilde{\epsilon}^i \frac{\rho^i}{R^i(1-\mu^i)}}{\sum \omega^j \widetilde{\epsilon}^j \frac{\rho^j}{R^j(1-\mu^j)}} \approx \frac{\widetilde{\epsilon}^i \rho^i}{\sum \omega^j \widetilde{\epsilon}^j \rho^j} \\
\frac{d\log R^i}{d\log \overline{L}} &= -\frac{1}{\epsilon^i} \times \frac{d\log l^i}{d\log \overline{L}} \approx -\frac{1}{(1-\alpha^i)+\widetilde{\epsilon}^i\alpha^i} \times \frac{\rho^i}{\sum \omega^j \widetilde{\epsilon}^j \rho^j}
\end{aligned} \tag{21}$$

Proof. See Appendix C.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 $\widetilde{\epsilon}^i$ can be constructed from demand elasticities ϵ^i and default elasticities α^i.

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 $\widetilde{\epsilon}^i$ is high. On loan markets with a low elasticity of loan demand $\epsilon < 1$ such as mortgages, risky borrowers with a high default elasticity α have a higher risk-adjusted loan demand elasticity $\widetilde{\epsilon}^i$, and safe borrowers with a low α have a

lower $\widetilde{\epsilon}^i$. 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 $\epsilon > 1$ 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.

4.1.2 Effect on default risk

The elasticity ϵ determines how changes in total credit supply affect the default risk μ 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. Combining the responses of rates and quantities we see that the credit risk of borrower i responds as

$$\frac{d\mu^i}{d \log \overline{L}} = (1 - \mu^i)\alpha^i \left(1 - \frac{1}{\epsilon^i}\right) \frac{d \log l^i}{d \log \overline{L}}. \tag{22}$$

In response to a negative credit supply shock, credit risk decreases for elastic borrowers ($\epsilon^i > 1$) and increases for inelastic ones ($\epsilon^i < 1$).

4.2 Monetary policy shocks

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 (Geanako-

plos 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 to borrower transmission

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

$$\frac{d\log R^i}{d\log R^f} = \frac{1}{1-\alpha^i+\alpha^i\epsilon^i}. \tag{23}$$

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 $\alpha^i\left(\epsilon^i-1\right) > 0$. In contrast, it is stronger if $\alpha^i\left(\epsilon^i-1\right) < 0$. Therefore, interest rates are sticky in response to monetary policy shocks on loan markets where the elasticity of loan demand ϵ is high such as credit cards. They are more volatile on markets where ϵ 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 equal risk-weights $\rho^i = 1$ 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:*

$$\frac{d\log R^i}{d\log R^f} = \frac{\widetilde{\epsilon}^i}{\epsilon^i}, \tag{24}$$

$$\frac{d\log l^i}{d\log R^f} = -\widetilde{\epsilon}^i. \tag{25}$$

- *When banks are constrained:*

$$\frac{d\log R^i}{d\log R^f} = \frac{\widetilde{\epsilon}^i}{\epsilon^i} \times \left(\frac{-1}{\widetilde{\epsilon}} \frac{d\log \bar{L}}{d\log R^f}\right), \tag{26}$$

$$\frac{d\log l^i}{d\log R^f} = -\widetilde{\epsilon}^i \times \left(\frac{-1}{\widetilde{\epsilon}} \frac{d\log \bar{L}}{d\log R^f}\right). \tag{27}$$

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

Proof. See Appendix C.1. □

4.2.2 Effect on default risk

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 (intensive margin); second, a reallocation of bank loans towards specific borrowers (extensive margin).

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

$$d\log\left(\mathbf{E}\left[1-\mu^i\right]\right) \approx \mathbf{Cov}\left(\mu^i, \widetilde{\epsilon}^i\right) d\log R^f - \frac{\mathbf{E}\left[d\mu^i\right]}{\mathbf{E}\left[1-\mu^i\right]} \tag{28}$$

4.3 Empirical Implications

Building on the general results above, we now illustrate them with a special case of the model. Our results explain a key feature of mortgage and credit card markets: the heterogeneous transmission of bank shocks across borrowers.

4.3.1 Calibration

Technology and preferences. Borrowers have preferences with constant relative risk aversion. 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 $\alpha, y_{\min}, y_{\max} > 0$. Loans are unsecured. In that case, the default elasticity α defined in (4) is constant, and equal to the Pareto parameter α.

Sufficient statistics: Empirical estimates. The default elasticity α 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 α while risky borrowers have a high value (Di Maggio et al. 2017, Fuster and Willen 2017).

The elasticity of borrowers' loan demand ϵ 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).

Sufficient statistics: Model determinants. Figure 1 illustrates the determinants of the demand elasticity ϵ in our workhorse model. We reproduce key features of the data. First, ϵ increases in the elasticity of intertemporal substitution, but it increases less for riskier borrower types (Best, Cloyne, Ilzetzki and Kleven 2019). Second, ϵ increases in borrowers' cash-on-hand and income, and it increases by less at higher levels (Buchak, Matvos, Piskorski and Seru 2020).

Interest rates do not fully capture credit conditions. On Figure 2, 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. At high risk levels α, a lower R becomes optimal because a higher rate would only increase the borrower's default probability. High risk borrowers do borrow less, however, but the reduction in lending is achieved through a tightening of their quantity limit. It translates into a *shadow*

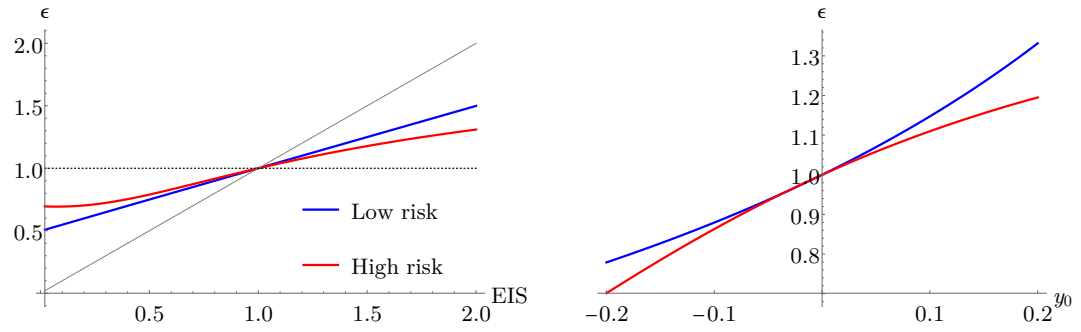

Figure 1: Sensitivity of the interest rate-elasticity ϵ to the elasticity of intertemporal substitution (EIS) and the initial endowment y_0 for low risk ($\alpha = 0$, in blue) and high risk ($\alpha = 0.5$, in red) borrowers.

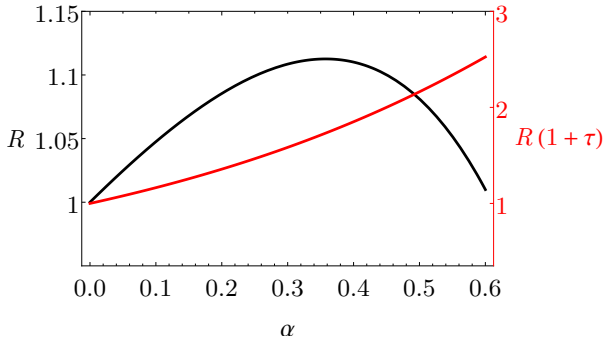

Figure 2: Equilibrium interest rate as a function of borrower risk α, where α is the shape parameter of a Pareto distribution (a higher α corresponds to a riskier distribution). The black curve depicts the contractual rate R (left scale) charged to borrowers by banks. The red curve depicts the shadow rate $R(1+\tau)$ (right scale), which accounts for the rationing wedge.

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.

Bank-dependent credit surface. 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 equals the default elasticity. 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.

In the case of a constant default elasticity α, we can solve the differential equation (19) in closed form:

$$R(l) = \bar{R} \times \left(\frac{l}{\bar{L}}\right)^{\frac{\alpha}{1-\alpha}} \tag{29}$$

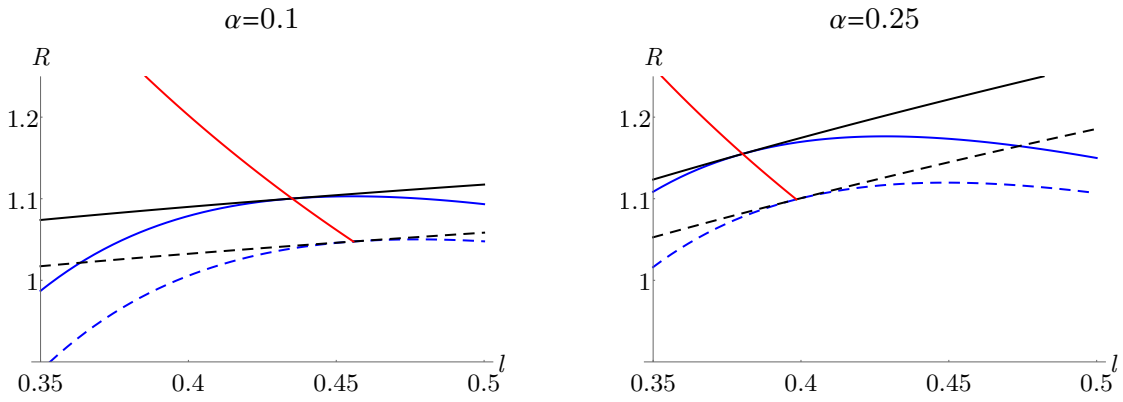

Figure 3: Each borrower α faces an increasing interest rate schedule $R_\alpha(l)$ (black lines). Borrowers choose the point at which their indifference curve (blue lines) is tangent to the rate schedule. The red line is the contract curve $\ell_\alpha(R)$ obtained by varying total lending capacity $\bar{L}$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

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 (29) 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

$$\left\{ \frac{R_i}{\ell_i\left(R_i\right)^{\frac{\alpha^i}{1-\alpha^i}}} \right\}_i \tag{30}$$

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 $\ell_i(R_i)$ drops) and less elastic markets (where R_i jumps while $\ell_i(R_i)$ is sticky) shift upwards. Second, interestingly, these changes are heterogeneous across borrower types. Differences in borrower risk α^i 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.

4.3.2 Credit supply shocks

Proposition 3 describes 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 $d\log\bar{L} > 0$ 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 $\epsilon < 1$ resulted in a decrease in the face value Rl to be repaid, hence a decrease in default risk μ for all borrowers. For

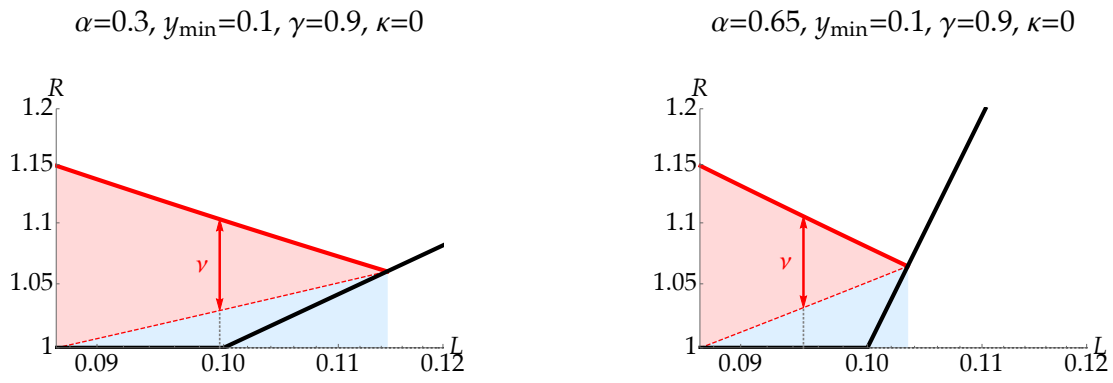

Figure 4: Borrowers with high $\gamma = 0.9$, log-log scale. Left: safer borrowers (low α), right: riskier borrowers (high α). The vertical red segment has length $\nu \approx 7\%$: the bank equalizes ν across types, which tells us how quantities and rates react for each type. Both types have the same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

risky borrowers with a high default elasticity α, 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 $\epsilon > 1$ 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 default elasticity α, 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.

The reason for this result is that endogenous default risk affects the elasticity of borrowers' demand for loan. Figure 4 shows that on markets with a high elasticity of loan demand such as credit cards, endogenous default risk makes risky borrowers effectively less elastic ($1 < \widetilde{\epsilon}^b < \widetilde{\epsilon}^a$). Their loan quantities increases by less and their loan rates decrease by more in response to a positive credit supply shock $d \log \overline{L} > 0$. As Figure 5 shows, on inelastic markets such as mortgages, risky borrowers are effectively more elastic ($\widetilde{\epsilon}^a < \widetilde{\epsilon}^b < 1$). Their loan quantities increase by more and their loan rates decrease by less when $d \log \overline{L} > 0$.

4.3.3 Monetary policy shocks

Bank lending channels. 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

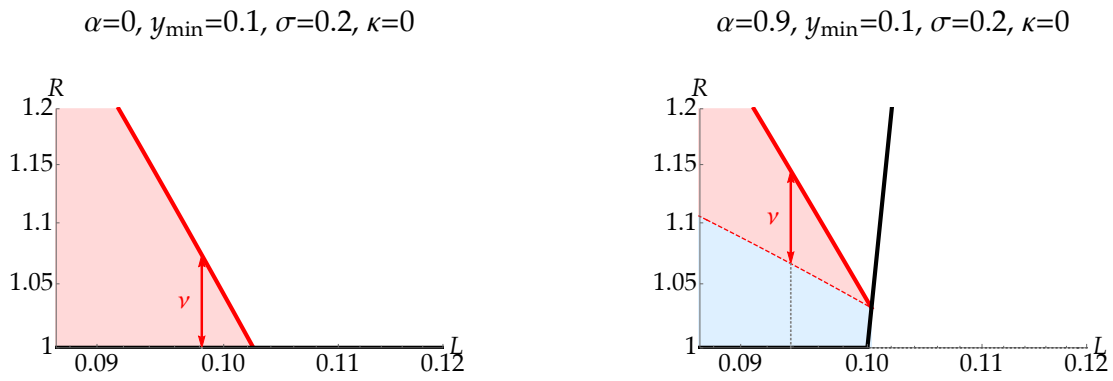

Figure 5: Borrowers with low EIS $\sigma = 0.2$, log-log scale. Left: safe borrowers ($\alpha = 0$), right: risky borrowers (high α). The vertical red segment has length $\nu \approx 7\%$: the bank equalizes ν across types, which tells us how quantities and rates react for each type. The contract curve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

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.

Bank portfolio risk. Figure 7 shows how individual default probabilities react to changes in banks' cost of funds relative to a baseline $R^f = 5\%$, 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 start with 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.

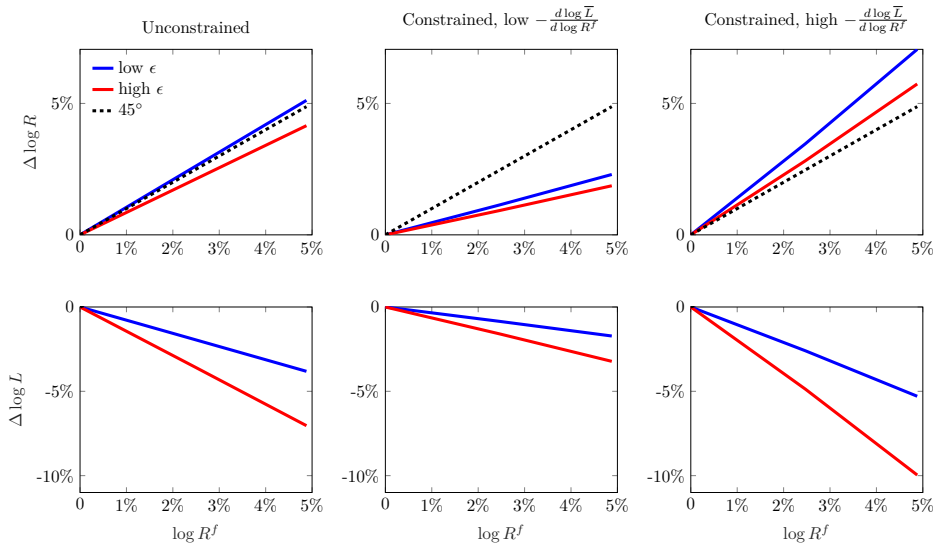

Figure 6: Percentage changes in loan rates and sizes as a function of banks' funding cost. Changes are plotted in cases where banks are unconstrained (left panels), constrained with inelastic lending capacity (middle panels), and constrained with elastic lending capacity (right). For each case, they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

The increase in the bank's default risk due to the composition effect also increases when R^f decreases.

Other non-price rerms: 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 8 shows that this trade off depends on the elasticity of loan demand ϵ. 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.

5 Endogenous Persistence of Credit Crises

After having analyzed how loan terms reacts to bank shocks, we now turn to the transition dynamics of credit crises. We extend the model dynamically, and show that the impact of a deterioration in banks' balance sheets is endogenously more persistent on loan markets where non-price terms adjust more than interest rates. We illustrate these findings by calibrating the model to the U.S. mortgage and credit card markets.

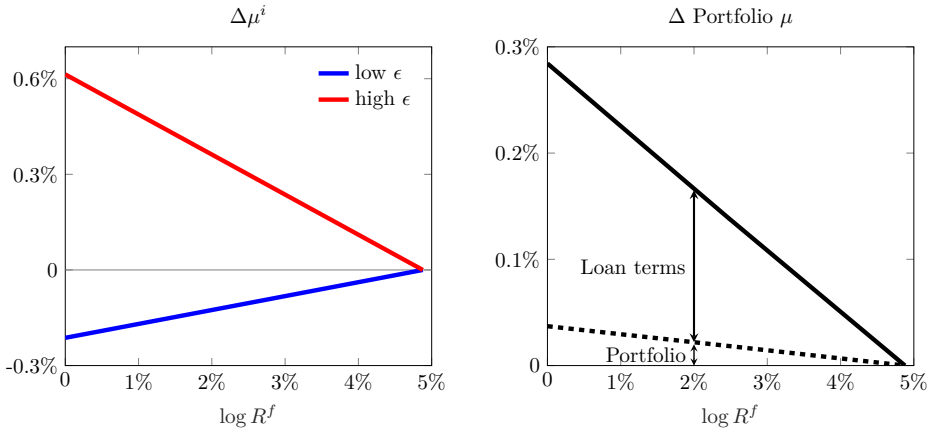

Figure 7: Effect of the decline in the risk-free rate, starting from a baseline $R^f = 5\%$, on individual default probabilities for borrowers with low vs. high elasticity ϵ (left panel), and on the total effective default probability of the bank's loan portfolio (right panel).

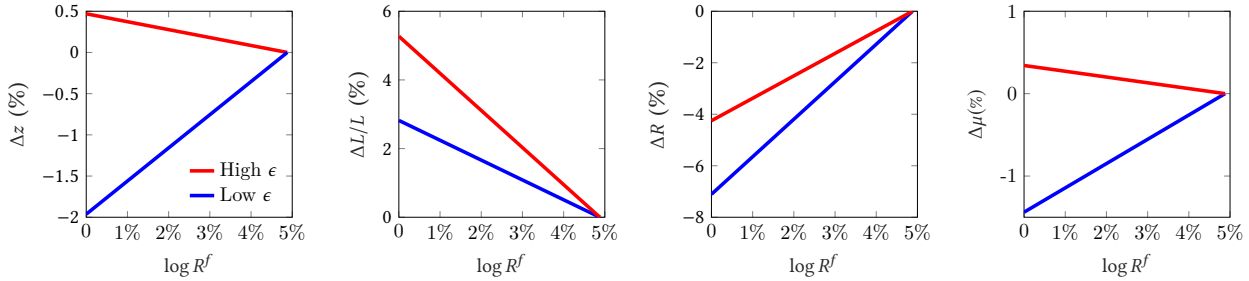

Figure 8: Covenant-lite loans: effect of bank's cost of fund on loan terms starting from a baseline $\log R^f = 5\%$, contrasting low elasticity ϵ (blue) and high ϵ (red) borrowers.

5.1 Dynamic model

We first analyze banks that are active on a single loan market, and then turn to banks that simultaneously lend on multiple markets. Time is discrete and the economy has an infinite horizon. Banks face overlapping generations of borrowers, and their capital varies over time as a function of their profits and losses on loan markets. Starting from equilibrium lending l^* in a steady state where banks are marginally unconstrained and thus $\nu = 0$, a negative credit supply shock lowers banks' capacity to $\bar{l}_0 < L^*$, i.e., by a proportional amount

$$\delta = \frac{L^* - \bar{l}_0}{L^*} \tag{31}$$

that increases the excess loan premium to $\nu_0 > 0$. After the initial shock the economy evolves deterministically under perfect foresight.

As in the large literature on the dynamics of banking crises (Gertler and Kiyotaki, 2010), we assume that banks are unable to raise capital in the short run hence equity, which determines

their lending capacity, recovers slowly through retained earnings. This is the only source of intertemporal linkages in the model: a high excess loan premium at date t increases retained earnings and thus equity growth from t to $t+1$, which increases lending capacity at $t+1$. We capture these dynamics by assuming that banks' lending capacity, and therefore the total supply of loans, evolves according to

$$l_{t+1} = (1 + \phi \nu_t)\, l_t \tag{32}$$

for some parameter $\phi > 0$. Thus high spreads ν_t recapitalize the banks and increase their lending capacity in the next period.[6]

Substituting for the loan contract curve with capacity constraints which arises from the bank's problem, we obtain that at each date

$$(1 + \phi \nu_t)\, \ell\left(R\left(l\left(\nu_t\right), \nu_t\right)\right) = \ell\left(R\left(l\left(\nu_{t+1}\right), \nu_{t+1}\right)\right). \tag{33}$$

As previously, $R(l, \nu)$ solves $R\left(1 - \mu\left(Rl\right)\right) = R^f + \nu$ holding R^f and z fixed, and $l(\nu)$ solves the static loan market clearing condition $l = \ell\left(R\left(l, \nu\right)\right)$. Linearizing the law of motion for the total supply of loans around its steady state l^* and using our previous expressions for $\frac{\partial R}{\partial \nu}$, $\frac{\partial R}{\partial l}$, the implicit function theorem gives the law of motion for the excess loan premium ν_t:

$$\nu_{t+1} = \left(1 - \frac{\phi R^f}{\widetilde{\epsilon}}\right) \nu_t \tag{34}$$

where $\widetilde{\epsilon} = \frac{\epsilon}{(1-\alpha)+\epsilon\alpha}$ is the risk-adjusted elasticity of loan demand defined previously.

5.2 Impact and persistence of bank shocks

Impact. The initial jump in the excess loan premium, or profits per risk-weighted dollar, is

$$\nu_0 = \frac{R^f}{\widetilde{\epsilon}} \delta. \tag{35}$$

The larger it is, the more quickly the resulting increase in banks' profits allows the total supply of loans tot recover back to the unconstrained steady state l^*. Conversely, a low excess loan premium ν_0 recapitalizes the bank more slowly and makes the decrease in the supply of loans l_t more persistent. Therefore, accounting for non-price terms is key for the persistence of credit crises. On loan markets where they are tightened more, interest rates and the excess loan premium increase by less, making the recapitalization of the bank slower and the credit crisis more persistent. The persistence of the credit crisis endogenously depends on the risk-adjusted elasticity of loan de-

[6] The law of motion (32) is standard. For instance, it obtains when banks face a leverage constraint $l_t \leq k \cdot n_t$, where n_t denotes the net worth at t, and banks retain a fraction $b \in (0, 1)$ of their earnings in every period. In that case $l_{t+1} = l_t(1 + b\nu_t)$.

mand $\widetilde{\epsilon}$. On markets with a low average demand elasticity such as mortgages, endogenous default risk $\alpha > 0$ increases the elasticity $\widetilde{\epsilon}$ which determines the response of non-price terms, including quantity limits, to credit supply shocks. A higher $\widetilde{\epsilon}$ leads to a larger tightening of non-price terms and a lower increase in the interest rate on such markets.

Proposition 5 describes the impact of the risk-adjusted elasticity of loan demand on the full dynamics of lending.

Proposition 5. *Denote* $\varphi = \frac{\phi R^f}{\widetilde{\epsilon}}$ *and the size of the initial credit supply shock* $\delta = \frac{l^* - \bar{L}_0}{l^*}$. *In a first-order approximation, the dynamics of the excess loan premium* v_t *and bank lending* L_t *follow*

$$v_t = \frac{R^f}{\widetilde{\epsilon}} \delta \left(1 - \varphi\right)^t, \tag{36}$$

$$l_t = l^* \left[1 - \delta \left(1 - \varphi\right)^t\right]. \tag{37}$$

Persistence. The persistence of the credit crisis is measured by the half-life of the excess loan premium v, defined as the time T that it takes for v to revert back to half of its initial value $v_T = v_0/2$, before it ultimately goes back to zero. With Proposition 5, we can compute the half-life of the credit crisis

$$T = \frac{\log 2}{-\log\left(1 - \varphi\right)}. \tag{38}$$

The persistence of the credit crisis over the full transition dynamics also depends on the adjustment of non-price terms through the risk-weighted elasticity $\widetilde{\epsilon}$ in φ, similar to the initial jump in the excess loan premium. The more non-price terms are tightened in response to the credit supply shock, the less interest rates increase. Therefore, the longer it will take for the excess loan premium v_t to go back to zero and for credit supply l_t to go back to its unconstrained value l^*.

This key result distinguishes intermediary-based loan pricing with multidimensional contracts from canonical macro-finance models (e.g. Kiyotaki and Moore 1997, Gertler and Kiyotaki 2010). In these models, a special case of Proposition 5 holds. Credit risk μ is exogenous, so $\alpha = 0$. As a result, the risk-adjusted elasticity of loan demand $\widetilde{\epsilon}$ is lower for instance on mortgage markets, and equal to the unconstrained loan demand elasticity ϵ_u. Therefore, quantity limits and the non-price terms of loans are tightened by less after a negative credit supply shock. In contrast, interest rates R increase more in order to lower borrower demand for credit and equate it with the bank's lending capacity. This leads to a larger increase in the excess loan premium, or profits per dollar v, which quickly recapitalizes the bank. Credit supply l_t quickly goes back to its unconstrained value l^*, and the credit crisis is short-lived. In contrast, in our model default risk is endogenous with $\alpha > 0$, which increases the risk-adjusted elasticity $\widetilde{\epsilon}$. The larger elasticity leads to a larger tightening in quantity limits and non-price terms after a negative credit supply shock, and a lower increase in interest rates. The resulting excess loan premium stays persistently lower,

slowing down the recapitalization of the bank and the recovery of credit supply. In this case, the credit crisis is endogenously more persistent.

Impact vs. persistence. How are the loan payments of overlapping generations of borrowers affected in response to a negative credit supply shock? On loan markets where the elasticity of loan demand $\widetilde{\epsilon}$ is high, the initial increase in the excess loan premium ν_0 is smaller. Current borrowers are hurt less as the crisis is initially milder with a low excess loan premium ν. However, the high elasticity $\widetilde{\epsilon}$ also implies that the crisis is more persistent, so that future borrowers are more affected. The net effect gives rise to an intertemporal trade-off between generations of borrowers.

With non-linear contracts and endogenous credit risk, the speed of bank recapitalization changes relative to standard models with linear contracts where no quantity limits are imposed by banks. The direction of the effect depends on the elasticity of loan demand:

- If $\epsilon < 1$, then $\frac{\log 2}{-\log\left(1-\frac{\phi R^f}{\widetilde{\epsilon}}\right)} \leq T \leq \frac{\log 2}{-\log\left(1-\phi R^f\right)}$: relative to a model with linear contracts, crises are milder on impact, but more persistent;
- If $\epsilon > 1$, then $\frac{\log 2}{-\log\left(1-\phi R^f\right)} \leq T \leq \frac{\log 2}{-\log\left(1-\frac{\phi R^f}{\widetilde{\epsilon}}\right)}$: relative to a model with linear contracts, crises are stronger on impact, but less persistent.

In total, does intermediary-based loan pricing with multidimensional contracts lead to a worse effect of credit contractions over the full transition dynamics? Interestingly, we find that the effects of endogenous impact and persistence of the crisis exactly compensate each other in present value terms relative to standard models. Proposition 6 describes this neutrality result.

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

$$\sum_{t=0}^{\infty} \nu_t = \frac{\delta}{\phi} \tag{39}$$

and is therefore independent of the risk-adjusted elasticity of loan demand $\widetilde{\epsilon}$.

Proposition 6 highlights the benefit from using sufficient statistics. The cumulative impact of a credit supply shock measured by the cumulative excess loan premium depends on the initial shock δ and the bank's balance sheet through the product of its leverage and earnings retention ratio ϕ. The cumulative impact of the shock is independent of the specific features of credit markets, such as borrower preferences, the microeconomic and information frictions which generate endogenous default risk $\alpha > 0$, and the feasible contract space with linear vs. multidimensional non-linear contracts. This result allows to compare credit crises in the cross-section and the time

series. As long as the bank parameter ϕ remains the same or can be controlled for, the cumulative spread relative to the initial shock δ should be unchanged.

Banks simultaneously active on each loan market. We now combine the transition dynamic analysis with our results on the incidence of credit supply shocks in the cross-section of loan markets. For simplicity, assume identical risk-weights $\rho^i = 1$. In a first-order approximation, the aggregate supply of loans in each period is

$$\begin{aligned}\sum_i l^i_{t+1}(\nu_{t+1}) &= (1+\phi\nu_t)\sum_i l^i_t(\nu_t) \\ &= (1+\phi\nu_t)\sum_i l^{i,*}\left(1-\epsilon^i\nu_t\right)\end{aligned} \tag{40}$$

Therefore, the excess loan premium ν_t follows a similar transition dynamics as in Equation (36) when banks are active on a single market at a time. The difference is that the risk-weighted elasticity of loan demand which governs the persistence of credit crises is a weighted average $\sum \omega^i \widetilde{\epsilon}^i$ of these elasticities on the various markets on which the bank is active, where more weight is given to markets with higher steady state loan shares $\omega^i = \frac{l^{i,*}}{\sum l^{i,*}}$. The speed at which the bank is recapitalized by an increase in interest rate spreads and the persistence of the credit crisis are governed by the average risk-adjusted elasticity. As a result, credit crises will be more persistent when banks lend to more elastic borrowers.

How are the different generations of borrowers affected over time in this case? The evolution of loan quantities on each market is given by

$$l^i_t = l^{i,*} - \widetilde{\epsilon}^i \nu_t. \tag{41}$$

Therefore, all loan quantities recover at the same speed across markets, which is determined by their common excess loan premium ν_t. Loan markets where borrowers i have a higher elasticity $\widetilde{\epsilon}^i$ suffer a larger initial tightening, and the differences in loan quantities with borrowers j on less elastic markets, who are less rationed, is determined by the ratio of their elasticities of loan demand. When they are different, the difference in loan quantities remains large and constant over time:

$$\frac{l^{j,*} - l^j_t}{l^{i,*} - l^i_t} = \frac{\widetilde{\epsilon}^j}{\widetilde{\epsilon}^i} \quad \forall t \tag{42}$$

How do these changes affect banks' balance sheets? Profit per dollar ν_t is common across borrowers, so the bank earns expected profits equal to $\nu_t l^i_t$ from type i borrowers. As a result, elastic borrowers for which l^i_t decreases more (and R^i_t increases less) induce lower expected profits for the banks that lend to them. Inelastic borrowers for which R^i_t increases more (l^i_t decreases less)

are hurt less by the credit crunch in terms of the loan amount borrowed, but they are also paying for the excess interest rate spread that allows the bank to recapitalize itself.

5.3 Empirical Implications

After having described the general results, we now use a special case of the model to illustrate them. Our findings explain a key feature of the U.S. mortgage and credit card markets: the persistence of credit crises.

We calibrate a simple model of U.S. credit markets where loan contracts have multiple price and non-price terms. The economy is populated by overlapping generations of borrowers with different elasticities of loan demand, to which banks allocate the total supply of credit. Banks lend to borrowers on the markets for mortgages and credit cards. First, we analyze the transition dynamics of a credit crisis when banks are active on each loan market separately. Then, we turn to the case where each bank is active on the two markets simultaneously.

5.3.1 Calibration

The two markets have different elasticities of loan demand, and for simplicity the same number of borrowers. We calibrate the model using data on U.S. mortgage and credit card markets before the Great Recession to match six moments, which are reported in Table 1.

Borrowers on the mortgage market have a low elasticity of loan demand ϵ. Empirical estimates for the low ϵ are generally below 1, and range from 0.05 in Benetton (2021) (in a structural model of the U.S. banking system), to 0.11 in Fuster and Zafar (2021) (in survey data), and 0.5 in Best et al. (2019) (using bunching at LTV thresholds). Borrowers on the credit card market have a high elasticity ϵ. Estimates for the high ϵ are generally above 1; we target 1.3 as in Gross and Souleles (2002). We use the EIS of the two borrower types to match the low and the high ϵ, and obtain 0.6 and 1.4.

We consider a single risk category for each market. To calibrate the default elasticity α, we first compute the elasticity $\frac{Rl\mu'}{\mu}$ which can be more easily estimated. We obtain an elasticity of default rates to interest rates of $\frac{Rl\mu'}{\mu} = 1.45$, within the range of empirical estimates of 1.1 in Fuster and Willen (2017) and 2 in Di Maggio et al. (2017). With an effective default probability $\mu = 0.1$ this corresponds to $\alpha = 0.16$.

For mortgages, we target average interest rates and LTV limits before the recession, in 2000-2007. We match values of respectively 15% and 0.82. The interest rate of 15% is also in line with the average annual percentage rate on credit cards, which ranges between 14% and 17% (Federal Reserve Board, G.19 Consumer Credit).

Variable	Description	Value	Target	Source
low ϵ	Elasticity of loan demand: mortgages	0.6	$[0.11, 0.5]$	See text
high ϵ	Elasticity of loan demand: credit cards	1.4	1.3	See text
α	Default elasticity	0.16	$[0.12, 0.22]$	See text
$R-1$	Interest rate	15%	[3%, 18%]	Primary Mortgage Survey
LTV	Loan-to-value limit	0.82	$[0.8, 1]$	Urban Institute

Table 1: U.S. mortgage and credit card markets: targeted moments

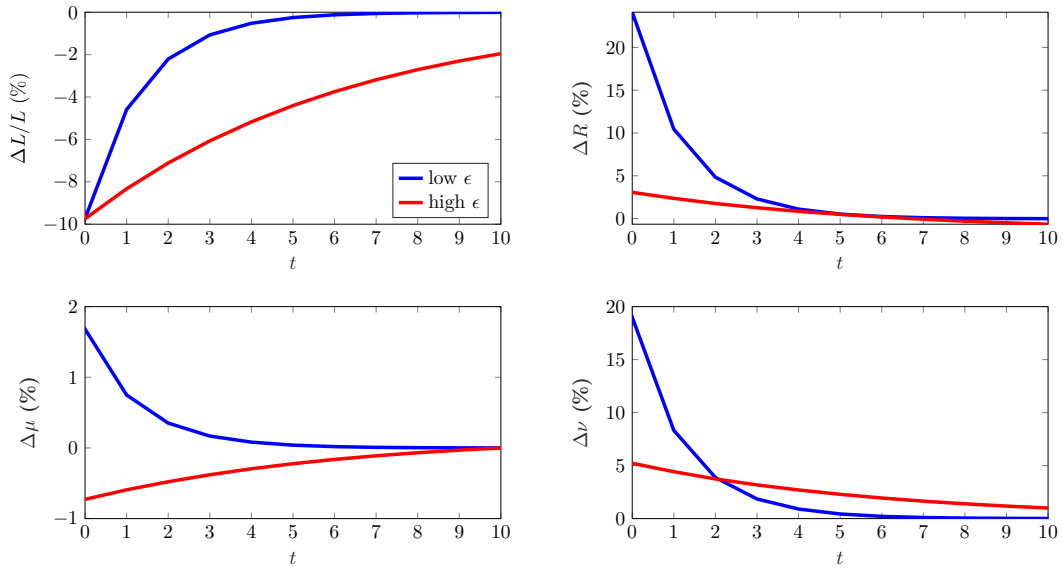

Figure 9: Dynamics of the cross-section of loans terms in the U.S. mortgage (blue) and credit card (red) markets in response to a tightening in banks' lending capacity. Impulse response functions for loan sizes, loan-to-value and payment-to-income ratios, excess loan premium, mortgage spreads, and default risk are plotted for each loan market.

5.3.2 Transition dynamics

We study the transitions dynamics of a credit crisis on the two loan markets. The economy starts from a steady state with $\nu = 0$ where banks are unconstrained and $L^* \leq \overline{L}$. At $t = 0$, bank lending capacity $\overline{L}$ contracts 10% below its steady state value L^*.

Banks separately active on each loan market. Figure 9 plots the transition dynamics of the economy when separate banks lend on each market. By construction, loan sizes L decrease by 10% upon impact on each loan market. The excess loan premia ν increase differently on each market in response to the negative credit supply shock. The increases in ν reflect the tightness of the banks' capacity constraints for mortgages with a low elasticity of loan demand $\epsilon < 1$ (blue) and for credit cards with a high $\epsilon > 1$ (red).

The dynamic responses of the two loan markets are consistent with our theoretical results. In response to the negative credit supply shock, interest rates R increase and loan sizes L decrease on both markets. The credit crisis has a larger impact on mortgage markets with a low elasticity of demand ϵ. The interest rate R and the resulting excess loan premium ν increase between three and four times as much as on credit card markets with a high ϵ, for the same initial decrease in L of 10%. As a result, the face value Rl to be repaid by borrowers increases, which leads to an increase in their endogenous default risk μ and a decrease in banks' expected profits. Over time, interest rates R are increased by more and non-price terms L are tightened by less than on markets with a high elasticity of demand ϵ. The larger increase in R and the higher profits per dollar ν recapitalize the banks faster by feeding back into total profits. The resulting credit crisis is less persistent, with the loan size L going back faster to its steady state value L^* and profits per dollar ν going back faster to zero.

The credit crisis is endogenously more persistent on loan markets with a high elasticity ϵ where non-price terms L are tightened by more and interest rates R increase by less over time. The higher persistence of the crisis illustrates the importance of accounting for changes in non-price terms when analyzing credit crunches, though it is initially less stark. The lower increase in R and the lower profits per dollar ν recapitalize the banks more slowly. The loan size L goes back much more slowly to its steady state value L^* and profits per dollar ν remain positive much longer. The face value Rl to be repaid decreases, which leads to a decrease in the endogenous default risk μ of borrowers. Overall, the net effect is an increase in the total default risk faced by the banking sector, since lending to low ϵ borrowers increases more quickly than to high ϵ borrowers.

Banks simultaneously active on two loan markets. Figure 10 plots the transition dynamics when each bank simultaneously lends to mortgage (blue) and credit card borrowers (red). The average loan size between the two markets decreases by 10% upon impact. As each bank lends to both groups of borrowers, it equalizes profits per dollar for each group to the single excess loan premium ν (black). The negative credit supply shock makes banks' capacity constraints on all loans more binding, which leads to an increase in ν. Upon impact, mortgage borrowers with a low elasticity of loan demand $\epsilon < 1$ face a larger increase in interest rates R and smaller decrease in loan sizes L, and credit card borrowers with a high $\epsilon > 1$ face a lower increase in R and a larger tightening in L. Therefore, the face values of loans to be repaid Rl increase for low-elasticity borrowers and decrease for high-elasticity borrowers, leading respectively to an increase and a decrease in their endogenous default risks μ. Since the impact of the shock is larger on markets with a high elasticity of loan demand ϵ, the composition of total lending shifts to borrowers with a low elasticity ϵ over time. The net effect is an increase in the total default risk of each bank.

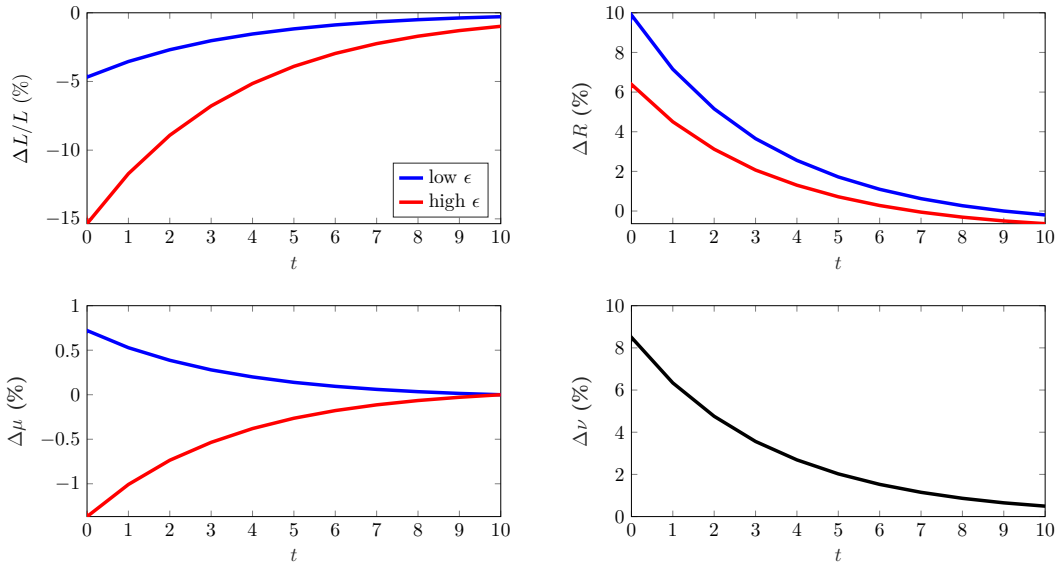

Figure 10: Dynamics of the cross-section of loans terms in the U.S. mortgage (blue) and credit card (red) markets in response to a tightening in banks' lending capacity. Impulse response functions for loan sizes, loan-to-value and payment-to-income ratios, excess loan premium, mortgage spreads, and default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

As implied by our theoretical results, the persistence of the credit crisis over time is similar across the two loan markets and governed by the dynamics of the excess loan premium ν. Therefore, diversifying lending towards high-elasticity borrowers increases the persistence of credit crises for banks which specialize in low-elasticity borrowers. Conversely, diversifying lending towards low-elasticity borrowers allows banks which specialize in high-elasticity borrowers to implement *cross-subsidization* between the two markets. In that case, the larger initial increase in interest rates on low-elasticity borrowers allows banks specialized in high-elasticity borrowers to recapitalize themselves more quickly after a negative credit supply shock. This also highlights the importance of comparing changes in interest rates to changes in non-price terms when analyzing the effect of credit crises on banks that lend to multiple borrowers.

6 Conclusion

We build a model of multidimensional contracting between heterogeneous borrowers and intermediaries with limited lending capacity. We show that two sufficient statistics, the interest-elasticity of borrowers' loan demand ϵ and the elasticity of default rates to repayment value α, 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 for which a unified explanation has been missing so far, especially the heterogeneous transmission of bank shocks

across loan markets and borrower risk categories, as well as the rise of covenant-lite lending in low risk-free rate environments. We highlight an important implication of accounting for non-price terms. Credit crises are endogenously more persistent on inelastic loan markets, where non-price terms adjust more than interest rates in response to bank shocks. More generally, the two sufficient statistics that we emphasize can provide useful guidance for understanding and comparing the results of quantitative structural models applied to different loan markets. Beyond our results in this paper, our approach allows to analyze more complex market structures for loans such as imperfect bank competition and borrowers' search for loan terms, which are important extensions for future research.

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.

Adelino, Manuel, Antoinette Schoar, and Felipe Severino (2016) "Loan Originations and Defaults in the Mortgage Crisis: The Role of the Middle Class," *The Review of Financial Studies*, Vol. 29, pp. 1635–1670.

Agarwal, Sumit, Souphala Chomsisengphet, and Chunlin Liu (2010) "The Importance of Adverse Selection in the Credit Card Market: Evidence from Randomized Trials of Credit Card Solicitations," *Journal of Money, Credit and Banking*, Vol. 42, pp. 743–754.

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.

Amiti, Mary and David E. Weinstein (2018) "How Much Do Idiosyncratic Bank Shocks Affect Investment? Evidence from Matched Bank-Firm Loan Data," *The Journal of Political Economy*, Vol. 126.

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

Benetton, Matteo, Sergio Mayordomo, and Daniel Paravisini (2022) "Credit Fire Sales: Captive Lending as Liquidity in Distress," March.

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 (2021) "Credit Cycles with Market Based Household Leverage," *The Journal of Financial Economics*, Vol. forthcoming.

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.

Gilje, Erik P., Elena Loutskina, and Philip E. Strahan (2016) "Exporting Liquidity: Branch Banking and Financial Integration," *The Journal of Finance*, Vol. 71, pp. 1159–1184.

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 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.

Jorda, Oscar, Bjorn Richter, Moritz Schularick, and Alan M. Taylor (2021) "Bank Capital Redux: Solvency, Liquidity, and Crisis," *The Review of Economic Studies*, Vol. 88, pp. 260–286.

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.

Khwaja, Asim and Atif Mian (2008) "Tracing the Impact of Bank Liquidity Shocks: Evidence from an Emerging Market," *American Economic Review*, Vol. 98, pp. 1413–1442.

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.

Ramcharan, Rodney, Stephane Verani, and Skander J. Van Den Heuvel (2016) "From Wall Street to Main Street: The Impact of the Financial Crisis on Consumer Credit Supply," *The Journal of Finance*, Vol. 71, pp. 1323–1356.

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.

Schnabl, Philipp (2012) "The International Transmission of Bank Liquidity Shocks: Evidence from an Emerging Market," *The Journal of Finance*, Vol. 67, pp. 897–932.

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.

Internet Appendix

A Data

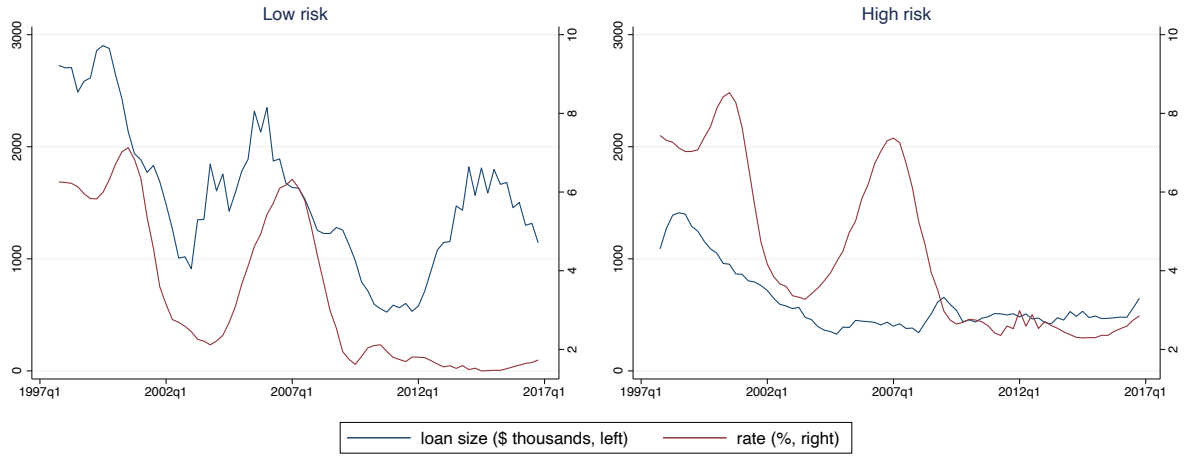

Figure 11: Loan size and loan rates for low and high risk short-term commercial and industrial loans. Source: Board of Governors of the Federal Reserve System (US).

B Effective Default Probability: Special Cases

Unsecured loans: pure liquidity default. The simplest case is an unsecured loan. The borrower (a firm or a household) is subject to an income shock y at the date of repayment, and a "liquidity default" happens if the realization of y is too low. A higher repayment value Rl makes it harder to repay. With zero recovery value, borrowers default if and only if $Rl \geq y$ and the lender gets nothing, so $\mu(Rl) = \mathbf{P}(y \leq Rl)$. More generally, the bank recovers $\rho(y, z)$ in case of default which happens when y falls below a threshold $\hat{y}(Rl, z)$. In that case,

$$\underbrace{\mu(Rl, z)}_{\text{effective default prob.}} = \underbrace{F(\hat{y}(Rl, z))}_{\text{actual default prob.}} \left(1 - \mathbf{E}\left[\frac{\rho(y, z)}{Rl} | y \leq \hat{y}(Rl, z)\right]\right) \tag{43}$$

Collateralized loans: liquidity and strategic default. Next, we consider collateralized loans, as in the case of a mortgage. 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

$$c_0 = y_0 - d = y_0 - P_0 + l. \tag{44}$$

At the date of repayment $t = 1$, income y_1 and house price P_1 are realized. Borrowers have utility

$$\begin{cases} u\left(y_1 - Rl + \chi P_1\right) & \text{if they repay} \\ u\left(\kappa y_1 + \underline{c}\right) & \text{if they default} \end{cases} \tag{45}$$

$(1-\kappa)\, y_1$ 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

$$w_1 \equiv y_1 + \frac{\chi}{1-\kappa} P_1 \leq \frac{\underline{c} + Rl}{1-\kappa} \tag{46}$$

This condition captures both liquidity defaults, stemming from a low realization of income y_1, as well as "strategic defaults", driven by a low realization of P_1. We can define the borrower's $t = 0$ value function as

$$V(l, R) = u\left(y_0 - P_0 + l\right) + \beta \left[\int_{w_1 \leq \frac{\underline{c}+Rl}{1-\kappa}} u\left(\kappa y_1 + \underline{c}\right) dF\left(y_1, P_1\right) + \int_{w_1 > \frac{\underline{c}+Rl}{1-\kappa}} u\left(y_1 - Rl + \chi P_1\right) dF\left(y_1, P_1\right) \right]$$

If upon default the bank recovers ζP_1, then the effective default probability is given by

$$\mu(Rl) = \int \int_{w_1 \leq \frac{\underline{c}+Rl}{1-\kappa}} \left(1 - \frac{\zeta P_1}{Rl}\right) dF\left(y_1, P_1\right). \tag{47}$$

Adverse selection. When borrowers' types are not observable, a higher repayment Rl attracts a worse distribution of borrowers. This is the classic problem analyzed by Stiglitz and Weiss (1981). For instance, suppose in the previous examples that the distribution of income F varies across types: safe borrowers have a better distribution (in the sense of stochastic dominance) than risky borrowers. In that case a higher face value Rl has the additional effect of attracting relatively more risky borrowers and thus increasing the effective default probability μ.

C Proofs and derivations

C.1 Main Propositions

Proof of Proposition 2. Each bank solves

$$\max_{\{x^i, R^i, l^i, \boldsymbol{z}^i\}} \int x^i \pi^i\left(l^i, R^i, \boldsymbol{z}^i\right) di \tag{48}$$

$$\text{s.t.} \int x^i \rho^i l^i di \leq \bar{L} \tag{49}$$

$$V^i\left(l^i, R^i, \boldsymbol{z}^i\right) \geq \overline{V}^i \tag{50}$$

Denote ν the multiplier on the bank lending constraint (49) and λ_i the one on borrower i's participation constraint (50). The first-order conditions with respect to l^i, R^i and x^i are respectively

$$x^i \pi_R^i + \lambda_i V_R^i = 0 \tag{51}$$

$$x^i \pi_l^i + \lambda_i V_l^i - \nu \rho^i = 0 \tag{52}$$

$$\pi^i - \nu \rho^i l^i = 0 \tag{53}$$

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

$$\frac{\pi^i}{\rho^i l_i} = \nu \tag{54}$$

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

$$-\frac{V_l^i}{V_R^i} = \frac{\frac{\pi^i}{l^i} - \pi_l^i}{\pi_R^i}. \tag{55}$$

Using

$$\pi^i = \left[R^i \left(1 - \mu^i\right) - R^f\right] l^i \tag{56}$$

we have

$$\begin{aligned} \frac{\frac{\pi^i}{l^i} - \pi_l^i}{\pi_R^i} &= \frac{R^i}{l^i} \frac{l^i \mu_l^i / \left(1 - \mu^i\right)}{1 - R^i \mu_R^i / (1 - \mu^i)} \\ &= \frac{R^i}{l^i} \frac{\alpha^i}{1 - \alpha^i} \end{aligned}$$

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

Proof of Proposition 1. We fix one borrower type i and omit the superscripts i. Differentiating (12) yields

$$-\frac{l\tau_l}{\tau}\frac{dl}{l} - \frac{R\tau_R}{\tau}\frac{dR}{R} = \theta\left(\frac{dl}{l} + \frac{dR}{R}\right)(1+\tau) \tag{57}$$

where $\theta = \frac{Rl\alpha'}{\alpha}$. Using $1 + \tau = \frac{1}{\alpha - 1}$ hence $\tau\,(1+\tau) = \frac{-\alpha}{(1-\alpha)^2}$ we get

$$\epsilon = \frac{-R\tau_R + \frac{Rl\alpha'}{(1-\alpha)^2}}{-l\tau_l + \frac{Rl\alpha'}{(1-\alpha)^2}}. \tag{58}$$

Letting $x = \frac{Rl\alpha'}{(1-\alpha)^2}$, 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 $\epsilon > \epsilon_u = \frac{R\tau_R}{l\tau_l}$ 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

$$\sum_i dl^i = d\bar{L} \tag{59}$$

$$-\sum_i \frac{l^i}{\bar{L}}\epsilon^i d\log R^i = d\log \bar{L} \tag{60}$$

To obtain $d\log R^i$, rewrite (13) as

$$\frac{R^i\left[1-\mu^i\left(R^i\ell^i\left(R^i\right),z^i\right)\right]-R^f}{\rho^i} = v \tag{61}$$

and differentiate to get for i, j

$$\frac{d\log R^i}{\rho^i}R^i\left(1-\mu^i\right)\left[1-\alpha^i\left(1-\epsilon^i\right)\right] = \frac{d\log R^j}{\rho^j}R^j\left(1-\mu^j\right)\left[1-\alpha^j\left(1-\epsilon^j\right)\right] \tag{62}$$

Therefore

$$\begin{aligned} -1 &= \sum_i \frac{l^i}{\bar{L}}\epsilon^i\frac{d\log R^i}{d\log \bar{L}} \\ &= \frac{l^i}{\bar{L}}\epsilon^i\frac{d\log R^i}{d\log \bar{L}} + \sum_{j\neq i}\frac{l^j}{\bar{L}}\epsilon^j\frac{d\log R^j}{d\log \bar{L}} \\ &= \frac{d\log R^i}{d\log \bar{L}}\frac{R^i\left(1-\mu^i\right)\left[1-\alpha^i\left(1-\epsilon^i\right)\right]}{\rho^i}\left\{\sum_j \omega^j\widetilde{\epsilon}^j\frac{\rho^j}{R^j\left(1-\mu^j\right)}\right\} \end{aligned} \tag{63}$$

where $\omega^j = \frac{l^j}{\bar{L}}$ are loan weights and $\widetilde{\epsilon}^j = \frac{\epsilon^j}{1-\alpha^j(1-\epsilon^j)}$ is the risk-adjusted elasticity. This rewrites

$$\frac{d\log R^i}{d\log \bar{L}} = -\frac{\rho^i}{R^i\left(1-\mu^i\right)}\frac{\widetilde{\epsilon}^i}{\epsilon^i}\times\frac{1}{\sum_j \omega^j\widetilde{\epsilon}^j\frac{\rho^j}{R^j(1-\mu^j)}} \tag{64}$$

which implies

$$\frac{d\log l^i}{d\log \bar{L}} = -\epsilon^i\frac{d\log R^i}{d\log \bar{L}} \tag{65}$$

and

$$\frac{d\log R^i}{d\log \overline{L}} = -\frac{1}{\epsilon^i}\frac{d\log l^i}{d\log \overline{L}}. \tag{66}$$

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

$$\frac{d\log l^i}{d\log \overline{L}} \approx \frac{\rho^i \widetilde{\epsilon}^i}{\sum_j \omega^j \rho^j \widetilde{\epsilon}^j}. \tag{67}$$

C.2 Other Calculations

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

$$\mu'(Rl) = \kappa f\left(\hat{y}(Rl)\right) + \frac{1-\kappa}{(Rl)^2}\int_{y_{\min}}^{\hat{y}(Rl)} y dF(y) \tag{68}$$

while

$$\begin{aligned}\alpha(Rl) &= \frac{Rl\mu'(Rl)}{1-\mu(Rl)} \\ &= \frac{Rl\kappa\frac{f(Rl)}{F(Rl)} + (1-\kappa)\,\mathbf{E}\left[\frac{y}{Rl}|y \le Rl\right]}{\frac{1-F(Rl)}{F(Rl)} + (1-\kappa)\,\mathbf{E}\left[\frac{y}{Rl}|y \le Rl\right]}\end{aligned} \tag{69}$$

If $\kappa = 0$ then $\alpha \in [0,1]$.

Pareto distribution.

- Suppose $\kappa = 0$ and

$$F(y) = 1 - \left(\frac{y_{\min}}{y}\right)^{\alpha} \tag{70}$$

for $\alpha > 0$ and $y > y_{\min},$, then $f(y) = \alpha\frac{1-F(y)}{y}$ and

$$\alpha(Rl) = \alpha \times \frac{1-\left(\frac{y_{\min}}{Rl}\right)^{1-\alpha}}{1-\alpha\left(\frac{y_{\min}}{Rl}\right)^{1-\alpha}} \in [0,1] \tag{71}$$

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 $\alpha(Rl) \in [0,1]$. The general formula is

$$\begin{aligned}
\alpha(Rl) &= \alpha \times \frac{1-\alpha\kappa-(1-\kappa)\left(\frac{y_{\min}}{Rl}\right)^{1-\alpha}}{1-\alpha\kappa-\alpha(1-\kappa)\left(\frac{y_{\min}}{Rl}\right)^{1-\alpha}} \\
\alpha'(Rl) &= \frac{(1-\alpha)^2\alpha(1-\kappa)y_{\min}(1-\alpha\kappa)\left(\frac{y_{\min}}{Rl}\right)^{\alpha}}{\left(\alpha(1-\kappa)y_{\min}-Rl(1-\alpha\kappa)\left(\frac{y_{\min}}{Rl}\right)^{\alpha}\right)^2} \\
\theta(Rl) = \frac{Rl\alpha'(Rl)}{\alpha(Rl)} &= \frac{(1-\alpha)^2(1-\kappa)y_{\min}}{\left(\alpha\frac{1-\kappa}{1-\alpha\kappa}\left(\frac{y_{\min}}{Rl}\right)^{1-\alpha}-1\right)\left(\frac{1-\kappa}{1-\alpha\kappa}\left(\frac{y_{\min}}{Rl}\right)^{1-\alpha}-1\right)}
\end{aligned} \tag{72}$$

- With the power law example,

$$\theta(Rl) = \frac{(1-\alpha)^2 Rl y_{\min}\left(\frac{y_{\min}}{Rl}\right)^{\alpha}}{\left(y_{\min}-Rl\left(\frac{y_{\min}}{Rl}\right)^{\alpha}\right)\left(\alpha y_{\min}-Rl\left(\frac{y_{\min}}{Rl}\right)^{\alpha}\right)} \tag{73}$$

  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

$$V(l,R) = u(y_0+l) + \beta u(y_1 - Rl) \tag{74}$$

    we see that

$$\tau(l,R) = \frac{u'(y_0+l)}{\beta R u'(y_1-Rl)} - 1 \tag{75}$$

    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

$$\frac{R\tau_R}{l\tau_l} = \frac{(y_0+l)(\sigma(y_1-Rl)+Rl)}{l(Ry_0+y_1)} \tag{76}$$

    If $y_0 = 0$ then this simplifies to

$$\frac{R\tau_R}{l\tau_l} = \sigma \times \left(1-\frac{Rl}{y_1}\right) + 1 \times \frac{Rl}{y_1} \tag{77}$$

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

  - Firms

$$V(l,R) = f(k_0+l) - Rl \tag{78}$$

then

$$\tau(l, R) = \frac{f'(k_0 + l)}{R} - 1 \tag{79}$$

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

$$\frac{R\tau_R}{l\tau_l} = \frac{f'(k_0 + l)}{-lf''(k_0 + l)} \tag{80}$$

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 $k_0 = 0$,

$$\frac{R\tau_R}{l\tau_l} = \frac{1}{1-\gamma} \geq 1 \tag{81}$$

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:

$$V(l, R) = u(y_0 + l) + \beta \left[\int_{Rl}^{\infty} u(y_1 - Rl)\, dF(y_1) + \int_0^{Rl} u(\kappa y_1)\, dF(y_1) \right] \tag{82}$$

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

$$V(l, R) = \int_{\frac{Rl}{(y_0+l)^\gamma}}^{\infty} \left[A(y_0 + l)^\gamma - Rl \right] dF(A) + \kappa (y_0 + l)^\gamma \int_0^{\frac{Rl}{(y_0+l)^\gamma}} A dF(A) \tag{83}$$