NBER WORKING PAPER SERIES

THE AGGREGATE DEMAND FOR BANK CAPITAL

Milton Harris
Christian Opp
Marcus Opp

Working Paper 27858
http://www.nber.org/papers/w27858

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
September 2020

We are grateful for feedback from our discussants Bengt Holmstrom, Guido Lorenzoni, Balint Horvath and Frederic Boissay, as well as comments by Juliane Begenau, Peter DeMarzo, Jason Donaldson, Arvind Krishnamurthy, and Ernst-Ludwig von Thadden. In addition, we thank participants at the NBER Summer Institute 2019 (Risks of Financial Institutions; Capital Markets and the Economy), the EFA 2019, the Chicago Booth Asset Pricing Conference, Bocconi BAFFI CAREFIN, European Central Bank, HKU, HKUST, University of Oregon, University of Mannheim, University of Rotterdam, NUS, and Tilburg University. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

The Aggregate Demand for Bank Capital
Milton Harris, Christian Opp, and Marcus Opp
NBER Working Paper No. 27858
September 2020
JEL No. G12,G21,G23,G28

ABSTRACT

We propose a novel conceptual approach to transparently characterizing credit market outcomes in economies with multi-dimensional borrower heterogeneity. Based on characterizations of securities' implicit demand for bank equity capital, we obtain closed-form expressions for the composition of credit, including a sufficient statistic for the provision of bank loans, and a novel cross-sectional asset pricing relation for securities held by regulated levered institutions. Our framework sheds light on the compositional shifts in credit prior to the 07/08 financial crisis and the European debt crisis, and can provide guidance on the allocative effects of shocks affecting both banks and the cross-sectional distribution of borrowers.

Milton Harris
Graduate School of Business
University of Chicago
5807 South Woodlawn Avenue
Chicago, IL 60637
milt@chicagobooth.edu

Christian Opp
Simon Business School
University of Rochester
3-110N Carol Simon Hall
Rochester, NY 14627
and NBER
opp@rochester.edu

Marcus Opp
Stockholm School of Economics
Drottninggatan 98
111 60 Stockholm
Stockholm
Sweden
marcus.opp@hhs.se

1. Introduction

Recent empirical evidence highlights the diverse cross-sectional implications of shocks affecting bank capital. In particular, the empirical literature has documented how banks affect real activity not only by alleviating credit rationing but also by reaching for yield in risky asset classes.[1] This rich evidence reveals that the allocative effects of various shocks affecting banks and the real economy depend on which types of lending to the cross-section of borrowers are affected. That is, *compositional* effects are of first-order importance, not just aggregate quantities.

Yet understanding how the composition of credit is determined is generally complex, since firms in practice exhibit heterogeneity along multiple relevant dimensions and simultaneously demand credit from heterogeneous banks. Correspondingly, each bank-borrower pair has effectively a separate demand and supply curve for credit, and this multitude of demand and supply curves are interdependent. For instance, one borrower experiencing a positive shock to her investment opportunities — a canonical credit demand shock — may imply a negative supply shock for another borrower, since banks might reallocate funds to the former borrower.[2] Moreover, banks face regulatory constraints and deposit subsidies, which appear to be of first-order importance for their decisions (see, e.g., Jiang, Matvos, Piskorski, and Seru (2020) for related evidence). These frictions introduce non-linearities implying that banks optimally account for the correlation structure of investments in their portfolio, in particular in tail risk states. Given the associated complexity of equilibrium aggregation, existing theoretical frameworks feature only limited notions of borrower heterogeneity. Yet heterogeneity is key to understanding the cross-sectional dispersion in firms' credit elasticities with respect to financial and real shocks.

In this paper, we aim to bridge this gap by offering a novel approach to transparently characterizing the credit market equilibrium in economies with rich borrower heterogeneity. Our main conceptual contribution is to depart from the conventional focus on demand and supply curves for credit, and to instead determine loans' implicit reservation prices for *bank*

[1]For evidence consistent with a bank lending channel, see, e.g., Kashyap et al. (1993, 1994), Gertler and Gilchrist (1994), Peek and Rosengren (2000), Khwaja and Mian (2008), Jiménez et al. (2012), Iyer et al. (2014), Chodorow-Reich (2014). For evidence on risk taking, see, e.g., Acharya et al. (2014), Jiménez et al. (2014), Iannotta et al. (2018).

[2]The empirical literature has recognized the complexity associated with separate credit demand and supply curves for each borrower-bank pair. As a result, the literature now typically relies on identification approaches that use variation within borrowers (see, e.g., Khwaja and Mian, 2008), which provides limited insights on the determinants of compositional changes at the aggregate level.

capital. This metric is essential for the composition of credit, since banks' lending decisions are governed by the objective to maximize shareholder value, and because banks' equity capital is a key scarce resource in the presence of financial frictions. We show that applying this approach delivers precise predictions for the pricing and the composition of credit, even in the presence of multi-dimensional borrower heterogeneity.

The presented framework can serve as a laboratory to examine the cross-sectional allocative effects of policy interventions and shocks affecting the economy. It reveals how the borrower ranking based on reservation prices identifies the group of banks' *marginal borrowers*, the set borrowers whose access to bank credit responds most elastically to small shocks. As a result, even focused policy interventions that exclusively increase the risk weights of (infra-marginal) high-risk borrowers may locally only cause credit rationing of (marginal) safe borrowers. More generally, our formula for borrowers' reservation prices provides sharp predictions for the implications of large shocks affecting both the financial sector and the real economy. As example applications, we discuss how our framework can coherently integrate various stylized facts that have been linked to the 07/08 financial crisis and the European debt crisis. Moreover, we highlight potential unintended consequences on bank risk-taking of policies aiming to improve public market access for firms, such as the EU's MIFID II initiative.

An essential feature of our analysis is that asset prices are determined in equilibrium by marginal investors that endogenously differ across assets. Despite this complexity, we obtain a unifying asset pricing relation for the cross-section of securities held by regulated levered institutions. This relation reveals how securities' expected returns reflect premia that are not only *increasing* in assets' implicit demand for scarce regulatory capital, but also *decreasing* in their contribution to levered institutions' portfolio tail risk. As a result, risky securities with low regulatory "risk weights" may command *negative* premia, consistent with empirical evidence on inflated prices in asset classes typically held by levered financial institutions, including insurance companies.[3] Counter to conventional wisdom, this asset pricing relation also unveils the logic for why certain borrowers' cost of capital may decrease in response to system-wide increases of capital requirements and how borrower heterogeneity is essential for this result. Intuitively, relatively safe firms may face lower financing costs if increases in capital requirements render the funding of risky borrowers unprofitable for banks, thereby freeing up capital and decreasing the equilibrium scarcity premium of bank capital.

[3] See, e.g., Greenwood and Hanson (2013).

We obtain these results in a flexible general equilibrium model of the cross-section of credit that can accommodate any finite number of borrower types and aggregate states. Borrower types differ in terms of investment opportunities, public market access, and regulatory risk classifications. Relative to public markets, banks differ in their credit supply due to a socially beneficial monitoring advantage, and by virtue of having access to implicitly subsidized debt financing via the anticipation of taxpayer bailouts or deposit insurance.[4] As in practice, banks are subject to Basel I-III bank capital requirements. Our model can accommodate any statistical relation between securities' actual riskiness and regulatory capital charges, allowing it to account for imperfections of regulations prevailing in practice (such as, for example, the zero capital charges on Greek sovereign debt that applied prior to the European debt crisis). Overall, our framework thus features two central frictions that the recent empirical literature has highlighted as influencing banks' leverage choices, deposit subsidies and capital regulations (see Jiang, Matvos, Piskorski, and Seru (2020)).

Analyzing a general equilibrium environment of this type is in principle technically challenging and distinguishes our paper from the existing literature. In particular, there is an important interaction between the general joint distribution of borrower characteristics (i.e., state-contingent cash flow risk, bank dependence, and regulatory classifications) and the presence of non-linearities in banks' objectives (due to the aforementioned deposit subsidies and regulations emphasized by the empirical literature). Banks account for the joint distribution of borrower characteristics and cash flows in their portfolio and leverage decisions and generally choose heterogeneous strategies, leading to differences in banks' default risk and default states. At the same time, banks' decisions jointly determine yields and allocations in general equilibrium. In this setting, we show that our novel approach to characterizing the credit market equilibrium is particularly useful in maintaining tractability and transparency.

The key measure derived in our analysis is the implicit *price* of bank capital that is associated with any given bank loan. This price is defined as the present value of a loan to bank equity holders, per unit of equity capital needed to fund the loan. This metric differs from the contractually specified interest rate, although it is an increasing function of the latter. Contrary to the interest rate, it is the relevant metric for banks' investment decisions when capital is scarce,[5] consistent with banks' focus on the return on equity in practice. As standard in price theory, the aggregate demand curve is then based on the corresponding *reservation prices*. Reservation prices are those prices that encode the maximum interest rate

[4]See Atkeson et al. (2018) and Duffie (2018) for evidence on this distortion.

[5]It is related to profitability indices used in capital budgeting contexts (Berk and DeMarzo, 2014).

a borrower would be willing to accept from a bank if only non-bank funding was available as an alternative. Importantly, these prices have to be determined conditional on a bank holding an optimal portfolio *and* a portfolio that is the best match for the particular loan among all optimal portfolios held by heterogeneous banks in equilibrium—portfolios matter since banks face a non-linear objective and thus have to account for the co-movement among all securities held. Despite these complexities, we obtain closed-form expressions for securities' reservation prices.

These closed-form solutions reveal how various dimensions of borrower heterogeneity affect a loan's position in the demand curve for bank capital. The reservation price of a loan exceeds a value of one by a premium that is given by the following ratio of borrower-specific quantities: (1) banks' and borrowers' joint incremental private surplus from bank funding relative to that obtainable under non-bank funding, and (2) the effective amount of bank capital used to fund a borrowers' loan. Incremental private surplus (the numerator) emerges from banks' comparative advantages in *both* monitoring and in funding investments with implicitly subsidized debt, and is computed conditional on the best feasible match among the various optimal portfolios held by banks in equilibrium. The second quantity (the denominator) maps units of credit into corresponding units of bank capital. A \$100 loan provided by a bank that funds its investment with 8% equity effectively "consumes" only \$8 of the bank's capital. As the incremental private surplus reflects the "put" value obtained from banks' ability to fund risky loans at subsidized rates, a wedge emerges; that is, the ranking of borrowers based on these reservation prices is generally not aligned with the ranking that would maximize allocative efficiency. The severity of this distortion, in turn, depends on securities' regulatory capital charges, which are determined by so-called "risk weights" in practice.

The *credit* market equilibrium is then pinned down by the intersection of demand and supply for *bank capital*. As our paper's contribution lies in micro-founding the aggregate *demand* curve for bank capital, we keep the modeling of the supply side parsimonious. In particular, we allow for flexible specifications for the costs of raising additional bank capital via issuances of outside equity (as in Decamps et al., 2011, Bolton et al., 2013), deviating from a common assumption in the financial accelerator literature that equity issuances are infeasible (e.g., Bernanke and Gertler, 1989). In equilibrium, bank credit is extended to all borrowers with reservation prices for bank capital above the marginal borrower type's reservation price, which is also the equilibrium price of bank capital. Borrowers with reservation prices below this equilibrium price issue bonds in public markets, if feasible. Our approach thus yields an intuitive *sufficient statistic* characterizing bank funding in the cross-section; a borrower obtains bank credit if the difference between a her reservation price and the equilibrium price of bank capital is weakly positive. Moreover, the equilibrium price for bank capital is key in determining the division of surplus between suppliers of bank capital (bank owners) and its infra-marginal customers (borrowers). Our analysis yields a closed-form expression for the cost of debt for bank-funded borrowers that encodes this equilibrium price, which has a familiar empirical counterpart; it is the shadow value of bank capital, an object that has been estimated in a recent influential literature (see, e.g., Koijen and Yogo, 2015, Kisin and Manela, 2016).[6] By microfounding cross-sectional asset demand based on agency and regulatory frictions, our paper provides tangible predictions for the pricing of assets held by regulated levered institutions, contributing novel insights to the emergent literature on institutional asset pricing (see e.g., Koijen and Yogo (forthcoming), He and Krishnamurthy (2013)).

Our approach yields transparent and testable predictions regarding the effects of various policy interventions and shocks. A general takeaway of our analysis is that bank loans with reservation prices close to the shadow value of bank capital are those that have the highest propensity of being affected by any type of shock or intervention affecting banks and borrowers' alternatives to bank finance. Shocks to an economy's bank capital move only the supply curve, thereby changing the identity of the marginal borrower. The impact of capital injections on allocative efficiency therefore depends on the social surplus created by the marginal borrower type's investment opportunities. Yet, due to the above-mentioned wedge, this social surplus may be negative. The pricing implications of shocks to the supply of bank capital also follow immediately. An increase in the supply lowers the shadow value of bank capital, thereby reducing bank loans' equilibrium yields.

The existing literature intensely debates the merits of changes to regulatory bank capital requirements (see, e.g., Admati et al., 2011, Allen et al., 2011, 2015, Begenau, forthcoming). Our analysis contributes to this debate by identifying the compositional effects of these policies, which ultimately shape aggregate effects. A change in capital requirements is a text-book credit supply shock, as it does not change a borrower's investment opportunities and implied reservation interest rates. However, it constitutes a demand shock for bank capital. This is because an increase in capital ratio requirements raises the amount of bank capital effectively required to fund any given loan (quantity channel) and reduces the value added

[6]For example, in the context of the insurance industry, Koijen and Yogo (2015) find a shadow cost of capital between \$0.10 and \$5.53 per dollar of statutory capital for the cross section of insurance companies in their sample.

any given loan provides to bank equity holders (surplus channel). As a result, increases in the overall capital ratio requirement (the capital to assets ratio) lead the demand curve to shift downwards and to fan out to the right. Moreover, the ranking of borrowers within the demand curve may change due to a skin-in-the-game effect — the reservation prices of borrowers whose bank-dependent surplus depends more on the above-mentioned put wedge (e.g., risky borrowers) fall more than those of other borrowers do. Thus, an increase in ratio requirements generally causes the ranking of borrower types in the demand curve to become better aligned with the ranking based on social surplus. Despite the increased reliance on bank capital, overall lending to surplus-generating borrowers can therefore expand if surplus-destroying risky borrowers start to be unprofitable and thus rationed, an effect that frees up previously used capital. On the other hand, if increases in ratio requirements are insufficient to cause substantive changes in the ranking of borrowers within the demand curve, such policy changes primarily lead to the rationing of marginal borrowers, which may be good borrowers.

Our theory also allows analyzing the overall equilibrium effects of targeted changes in the capital charges associated with specific classes of securities. Even such targeted changes have externalities on other types of borrowers, in particular non-targeted marginal borrowers. For example, if the risk weights of a subset of infra-marginal borrowers are increased, but these increases are insufficient to cause those borrowers to become rationed, this policy merely induces the rationing of additional marginal borrowers. Moreover, our analysis highlights that setting capital charges for various asset classes should not be based only on evaluations of a loan's riskiness (which is the primary focus of existing regulations), but also on a borrower's bank dependence. Setting high risk weights for borrowers that are non-bank dependent is beneficial independent of a loan's riskiness.

Finally, we analyze the effects of improvements in the efficiency and accessibility of public markets or other bank alternatives available to borrowers. This analysis sheds light on time-series trends associated with financial innovations, such as the development of junk bond markets in the 1980s, securitization and shadow banking in the 2000s, and the ongoing development of FinTech funding platforms, such as those facilitating crowdfunding. Moreover, it may be applied to cross-country comparisons (say USA vs. Italy), or to evaluate policy initiatives aiming to give borrowers better access to non-bank finance, such as the European Union's "Markets in financial instruments directive" MiFID II. If these bank alternatives are less subject to distortions associated with government bailouts or deposit insurance, they will compete with banks for only those types of borrowers that are viable under such lower subsidies; that is, those borrowers that tend to have fundamentally better and safer investment opportunities. As a result, the *relative* ranking of high-risk borrowers in the demand curve for bank capital improves, implying that banks will tend to shift their portfolios towards these borrowers. Consistent with these predictions, Hoshi and Kashyap (1999, 2001) show empirically that deregulations leading up to the "Japanese Big Bang" allowed large corporations to switch from banks to public capital markets, which caused banks to take greater risks. If policy makers take a macroprudential approach to regulating the entire financial system, they can counteract this perverse behavior by increasing capital requirements in response to the increased availability of non-bank finance.

Relation to the literature. As in Holmstrom and Tirole (1997), banks in our model can create social value by lending to borrowers that would otherwise be credit-rationed by public markets.[7] This advantage emanates from banks' ability to monitor borrowers and thereby reduce moral hazard, consistent with Diamond (1984). Relative to Holmstrom and Tirole (1997), our framework features not only general distributions of borrowers that differ along multiple dimensions (state-contingent cash flows, bank dependence, and regulatory risk classifications), but also non-linearities in banks' objectives associated with the above-discussed elements emphasized by the empirical literature, that is, implicit government subsidies for debt and capital regulations. We propose a novel conceptual approach to transparently characterizing credit market outcomes in this class of general equilibrium economies, introducing loans' implicit demand for bank capital as a central metric. This characterization yields precise predictions for allocations and a novel cross-sectional asset pricing relation that reveals how securities' return premia are affected by both capital regulations and implicit debt subsidies that introduce "reaching for yield" incentives.

In our general equilibrium setting with a cross-section of borrowers, risk-taking is not only associated with heterogeneous portfolio strategies across banks,[8] but also causes distortions in the cross-section of asset prices.[9] In particular, our framework can rationalize apparent "bubbles" in risky asset classes that are subject to low regulatory risk weights. Such pricing effects do not emerge in partial equilibrium settings, such as for example the one considered in Rochet (1992), who shows that banks typically choose specialized, risky portfolios when their deposits are insured (see also Repullo and Suarez, 2004). Moreover, our cross-

[7]In Chemmanur and Fulghieri (1994), borrowers can also choose between bank loans and publicly traded debt, but their analysis focuses on incentives for information production in distress.

[8]Kahn and Winton (2004) show that such "segmentation" may even obtain within a bank by creating subsidiaries without mutual recourse.

[9]Becker and Ivashina (2015) provide empirical evidence of reaching for yield behavior by life insurers, consistent with predictions of Pennacchi (2006).

sectional asset pricing relation for securities held by regulated levered financial institutions is a novel contribution to the recent literature on the pricing of securities when intermediaries are marginal investors.[10]

Finally, our paper relates to the literature that explores the role of competition for financial stability and banks' risk-taking incentives. Marcus (1984) and Keeley (1990) highlight that competition *between* banks reduces a bank's value of staying solvent and thus, encourages risk-taking.[11] In our model, banks compete not only with each other but also with investors in public markets. Yet, as borrowers have heterogeneous access to these markets, this channel has additional compositional implications, consistent with the above-mentioned evidence on the Japanese Big Bang.

2. Model Setup

We consider a discrete-state economy with two dates, 0 and 1.[12] At date 1, the aggregate state of the world $s \in \Sigma$ is realized. The ex-ante probability of state s is denoted by $\pi_s > 0$. The economy consists of three types of agents, *entrepreneurs*, *investors*, and *bankers*. All agents in the economy are risk-neutral, have a rate of time preference of zero, and have access to a risk-free outside investment opportunity yielding a net-return of $r_F \geq 0$.

2.1. Entrepreneurs

Entrepreneurs are agents with real investment opportunities. We also refer to these agents as *firms*, *borrowers*, or *issuers*. There is a continuum of firms of total measure one, indexed by $f \in \Omega_f$.[13] Each firm f is owned by a cashless entrepreneur who has access to a project that requires a fixed-scale investment I at time 0, and produces state-contingent cash flows C_s at time 1.[14] Firm cash flows $C_s(q, a)$ are affected by the entrepreneur's discrete fundamental

[10]See, e.g., Garleanu and Pedersen (2011) and He and Krishnamurthy (2013).

[11]Related implications of competition for regulation have also been studied in Boot et al. (1993), Hellmann et al. (2000), and Repullo (2004).

[12]In Appendix B, we discuss the robustness of our main insights to multi-period settings.

[13]Formally, $f = (f_1, f_2)$ with $f_i \in [0, 1]$ for $i \in \{1, 2\}$ and $\Omega_f = [0, 1] \times [0, 1]$. The double continuum assumption for firms will ensure that firms are atomistic relative to banks.

[14]A simple way to capture an investment scale decision in our type of environment is to consider firms that own collections of fixed-scale investment opportunities.

type $q_f \in \Omega_q$ and her *unobservable* binary action $a_f \in \{0,1\}$. Going forward, we will at times omit firm subscripts when doing so does not create ambiguity.

Firms are subject to limited liability and have access to monitored financing from banks and unmonitored financing from public markets. In public markets, investors and banks compete for firms' securities. Both investors and bankers can observe the firm fundamental q, implying that there is no asymmetric information about fundamentals between issuers and providers of capital. There is, however, a moral hazard problem. Shirking, $a = 0$, allows the entrepreneur to enjoy a private benefit of $B(q)$ when unmonitored, and 0 when monitored by banks.[15]

Assumption 1 *Parameters satisfy the following relations:*

1) $\frac{\mathbb{E}[C_s(q,0)]}{1+r_F} + B(q) < I \quad \forall q,$

2) $\frac{C_s(q,0)}{1+r_F} < I \quad \forall s, \forall q.$

The first condition implies that no project generates positive social surplus (including the private benefit) under shirking. The second assumption is made for expositional reasons. It simplifies the entrepreneur's incentive problem when unmonitored finance is provided and implies that debt is the optimal contract (see Lemma 1 below).

2.2. Investors

There is a continuum of competitive investors with sufficient wealth to finance all projects in the economy. At date 0, investors have access to the following investment opportunities: (1) securities issued by firms in public markets, (2) bank deposits and bank capital (equity), and (3) the risk-free outside investment opportunity. Competition, capital abundance, risk-neutrality, a zero rate of time preference, and access to an outside investment opportunity yielding a return of $r_F \geq 0$ imply that investors' demand an expected rate of return of r_F on all investments in equilibrium.

Financing of firms via public markets requires that the borrower's stake in her company provides her with sufficient incentives to exert effort $(a = 1)$, as part 1 of Assumption 1

[15]More generally, similar qualitative results obtain as long as banks strictly reduce the private benefit of shirking.

renders financing under shirking $(a = 0)$ infeasible. Going forward, we denote by

$$NPV\left(q\right) \equiv \frac{\mathbb{E}[C_s\left(q,1\right)]}{1+r_F} - I \tag{1}$$

the project's value added under high effort. Securities purchased by investors must allow them to break even on their investment. Taken together, a firm with fundamental q can obtain financing from investors in public markets if there exists a security with promised state-s cash flows, $CF_s \geq 0$, that satisfies both the entrepreneur's IC constraint and investors' IR constraint:

$$\frac{\mathbb{E}\left[\max\left\{C_s\left(q,1\right) - CF_s, 0\right\}\right]}{1+r_F} \geq B\left(q\right) + \frac{\mathbb{E}\left[\max\left\{C_s\left(q,0\right) - CF_s, 0\right\}\right]}{1+r_F}, \tag{IC}$$

$$\frac{\mathbb{E}\left[\min\left\{C_s\left(q,1\right), CF_s\right\}\right]}{1+r_F} \geq I. \tag{IR}$$

Lemma 1 *A firm with fundamental q can obtain unmonitored finance from investors in public markets if and only if $NPV\left(q\right) \geq B\left(q\right)$. Under unmonitored finance, debt is an optimal contract and the value of an entrepreneur's equity is $NPV\left(q\right)$.*

A firm cannot receive unmonitored finance — and is thus bank-dependent — if its value added NPV is small relative to the moral hazard rent B. While our model relates bank-dependence to moral hazard rents, one may more generally view the parameter $B(q)$ as any firm fundamental that determines bank-dependence in reduced form.[16] Note that our setup leaves full flexibility on how a particular fundamental type q is associated with state-contingent cash flows $C_s(q, 1)$ and the bank dependence parameter $B(q)$.

2.3. Banks

There is a continuum of competitive bankers $b \in \Omega_b$ of mass 1.[17] Bankers have access to a costless monitoring technology that allows them to eliminate an entrepreneur's private benefit from shirking, $B\left(q\right)$.[18] As a result, banks can effectively raise entrepreneurs' pledgeable income.

[16] Empirically, large firms are more likely to have access to public markets than small- and medium sized firms do (see e.g., Gertler and Gilchrist (1994) or Iyer et al. (2014)).

[17] In Appendix B, we discuss the robustness of our analysis with respect to the possibility that banks have market power.

[18] As discussed in Appendix B, key insights of our analysis also apply when banks have to incur costs to monitor borrowers and when banks differ in their monitoring abilities.

At time 0, each banker has positive initial wealth in the form of cash, and bankers' aggregate wealth is E_I.[19] Since the distribution of wealth is not important for our key results, we presume that aggregate wealth is uniformly distributed among bankers, implying that E_I also corresponds to bankers' initial per-capita wealth. Banks may also raise external funds in the form of outside equity capital E_O and deposits D. We denote by A the total amount invested in firms and by M the total amount invested in the risk-free outside investment opportunity. Thus, we obtain the following balance sheet identity in terms of book values:

$$A + M = E + D, \tag{2}$$

where we define $E \equiv E_I + E_O$ as the total book equity capital. Banks can invest in firms via bank loans or via unmonitored bonds issued in public markets. Regarding these investments we make two assumptions. First, firm projects requiring bank monitoring are funded by a loan that is fully held on the balance sheet of the monitoring bank.[20] Second, banks can invest only in bonds that are at least pari passu with other debt issued by a firm (but not junior debt or equity).[21] These assumptions ensure that we can abstract from security design and the origination and trading of synthetic (derivative) securities.[22]

External financing frictions. Banks are subject to limited liability and face external financing frictions, consistent with the literature on the bank lending channel. As our paper's contribution is focused on micro-founding the aggregate demand for bank capital in the presence of general cross-sectional borrower distributions, we model the supply side in a parsimonious and flexible way.[23] For a bank to raise a net-amount E_O of new equity capital, investors need to put up $c(E_O)$ units of cash, where for $E_O > 0$, the function $c(\cdot)$ satisfies the properties $c(E_O) \geq E_O$, $c'(E_O) > 1$, and $c''(E_O) \geq 0$. For $E_O \leq 0$, the function is given by $c(E_O) = E_O$. That is, a bank raises $c(E_O)$ units from investors, but due to costly frictions obtains in net only E_O units of new equity bank capital, with the remainder being absorbed

[19]In Appendix B, we discuss the implications of legacy assets for our model's predictions.

[20]This assumption ensures that our model captures the "skin-in-the-game" requirement that is typical for models with moral hazard.

[21]In practice, investments in firms' equity do not play an important role on the asset side of banks' balance sheets. This may, in part, be explained by stringent capital requirements: under Basel III, U.S. banks are subject to a risk-weight of 300% for publicly traded stocks and 400% for non-publicly traded equity exposures.

[22]While security design would be an interesting extension, our assumption ensures that we can focus on *issuer* risk classifications (introduced below), avoiding the need to specify classifications for all possible security types that an individual firm might issue.

[23]See, e.g., Decamps et al. (2011) and Bolton et al. (2013) for similar reduced-form specifications.

by issuance costs.[24] Going forward, we will refer to this remainder, $(c(E_O) - E_O)$, as *net issuance costs*. In contrast, paying dividends (which implies $E_O < 0$) is not subject to any frictions. Similarly, the process of issuing deposits is frictionless. A wedge between banks' costs of raising debt on the one hand and equity on the other is a general property of models where moral hazard impedes outside financing, and debt provides better incentives (Innes, 1990, Tirole, 2006). Such a wedge may also arise because of adverse selection (Gorton and Pennacchi, 1990), or due to equity claims' lack of monetary services (Stein, 2012).

Bank regulation. We take two features of real-world regulations pertaining to banks as primitives of our economy. First, bank deposits are effectively insured by FDIC insurance and/or implicit bailout guarantees. Second, banks are subject to capital requirements. Recent empirical evidence suggests that these two features are of first-order importance for banks' leverage choices in practice (see Jiang, Matvos, Piskorski, and Seru (2020)). Although there is a substantial literature that sheds light on the potential reasons why these two institutional features might exist,[25] a variety of economic forces, including political economy frictions (incentives for holding office, lobbying, competition between countries, etc.), are likely responsible for their historical emergence and persistence. As it is not the purpose of this paper to rationalize these institutions based on one particular economic force, we take them as given and analyze their implications for credit supply decisions.

In the following, we describe how our model captures these institutional features. First, promised payments of bank deposit contracts are fully insured by the government,[26] and any shortfalls are financed by lump-sum taxes that are levied from investors. As common in the literature, we thus abstract from deposit insurance premia,[27] which are quite insensitive to banks' asset risk in practice (see, e.g., Kisin and Manela, 2016). This approach is also in line with our objective to capture the effects of implicit bailout guarantees, for which banks do not pay insurance premia. Yet, we also discuss in Appendix B that the key insights of our

[24]The costs may be interpreted as equity issuance costs or, for example, as capturing the effect of debt overhang in reduced form (see Admati et al. (2018)). Whether the costs are purely private or also have a social component is irrelevant for the positive analysis of banks' decisions. Exploiting proprietary access to Swedish banks' internal profitability metrics, Qi (2019) provides direct evidence that banks internally view equity issuances as costly.

[25]See Diamond and Dybvig (1983) for deposit insurance and Bianchi (2016) or Chari and Kehoe (2016) for bailouts.

[26]If guarantees were imperfect, the deposit rate would reflect a bank's default risk, but less than justified by a bank's asset risk. The qualitative results of our analysis would be unaffected in this case.

[27]See, e.g., Hellmann et al. (2000) and Repullo and Suarez (2013). See also Pennacchi (1987, 2006) and Iannotta et al. (2018) for analyses of deposit insurance pricing and implications for bank regulation and financial system risks.

conceptual approach are robust to deviations from this specification.

Second, banks are subject to capital regulations that may be contingent on risk classifications of the issuers in which a bank invests. Risk classifications are denoted by ρ, and take values in the discrete set Ω_ρ. The empirical counterpart of these risk classifications might be credit ratings and/or asset classifications, which are used in regulations in practice. Going forward, we refer to the pair (q, ρ) as an issuer's *type*. We impose the technical condition that if any issuer in the economy is of the type (q, ρ), there is a also strictly positive mass of firms of this type, $m(q, \rho) > 0$.[28] Whereas the risk classification ρ is *verifiable* for regulatory purposes, the firm fundamental q is not (see, e.g., Grossman and Hart, 1986, for the definition of verifiability). Yet, as we do not impose any restrictions on the relation between ρ and q, our model can in principle capture any degree of verifiability in the context of regulations.

Let $x(q, \rho)$ denote a bank's portfolio weight corresponding to issuers of type (q, ρ), and let $\mathbf{x}$ denote the vector of portfolio weights for all issuer types. Due to shortsale constraints for bank loans, the portfolio weights must satisfy $x(q, \rho) \geq 0$. As in the regulatory frameworks of Basel I-III, bank capital regulation prescribes that the equity capital-to-assets ratio of every bank, $e \equiv \frac{E}{A}$, be above some minimum threshold $e_{\min}(\mathbf{x})$ that is a weighted average of asset-specific capital requirements $\underline{e}(\rho)$:

$$e_{\min}(\mathbf{x}) \equiv \sum_{\forall q, \rho} x(q, \rho) \cdot \underline{e}(\rho) \,. \tag{3}$$

Note that whereas a bank's investment strategy $x(q, \rho)$ conditions on the full type (q, ρ), the regulatory capital requirement parameter $\underline{e}(\rho)$ conditions only on the verifiable component ρ. In line with regulations in practice, it is useful to recast $\underline{e}(\rho)$ as the product of a risk-weight, $rw(\rho)$, and an overall level of capital requirements, $\underline{e}$, that is,

$$\underline{e}(\rho) = rw(\rho) \cdot \underline{e}. \tag{4}$$

Bankers' Objective. Competitive banks take equilibrium yields $y(q, \rho)$ charged to firms of type (q, ρ) as given. The state-contingent rate of return for an investment in an issuer of type (q, ρ) is given by:

$$r^s(q, \rho) = \min\left\{ y(q, \rho), \frac{C_s(q, 1)}{I} - 1 \right\}. \tag{5}$$

[28]This assumption ensures that an infinitesimal bank's asset demand never exceeds the total supply of firms with a given existing type (q, ρ).

Equation (5) reflects that a bank, after lending an amount I, receives a borrowing firm's total cash flow $C_s(q,1)$ whenever the firm defaults. The overall rate of return on a bank's portfolio in state s, which we define as r_A^s, is given by:

$$r_A^s(\mathbf{x}) = \sum_{\forall q,\rho} x(q,\rho) \cdot r^s(q,\rho). \tag{6}$$

Due to deposit insurance, investors are willing to provide deposit finance to banks at a promised interest rate of $r_D = r_F$, regardless of the asset holdings of a bank. Thus, after raising a net-amount of outside equity E_O and deposits D, the total market value of a bank's equity is:

$$E_M = \frac{\mathbb{E}\left[\max\left\{(1 + r_A^s(\mathbf{x}))A + (M - D)(1 + r_F), 0\right\}\right]}{1 + r_F}, \tag{7}$$

which accounts for a bank's limited liability. Before raising outside finance, a banker's objective is to maximize the value of her equity stake, i.e., the market value of the inside equity, which we denote by $E_{M,I}$. Competition implies that the value outside equity holders obtain must be equal to the cash they put up, $c(E_O)$. Thus, we obtain:

$$E_{M,I} = \max_{E_O,M,D,\mathbf{x}} \left\{E_M - c(E_O)\right\}. \tag{8}$$

It is useful to express this objective function in terms of the equity ratio $e = \frac{E_I + E_O}{A}$. Using this definition and the balance sheet identity (2), we can eliminate the variables D and M, and write the *expected* rate of return on bank book equity (ROE) before the cost of outside equity as:

$$r_E(\mathbf{x},e) \equiv \mathbb{E}\left[\max\left\{r_F + \frac{r_A^s(\mathbf{x}) - r_F}{e}, -1\right\}\right], \tag{9}$$

which reflects the fact that equity returns are a *convex* function of asset returns. Non-linearities emerge since equity holders are not only subject to limited liability ex post, but also benefit from the anticipation of deposit insurance payments or bailouts ex ante; debt is priced to promise a yield equal to r_F irrespective of a bank's ability to pay depositors in all states of the world. Using (9), we obtain the equivalent maximization problem:

$$E_{M,I} = E_I + \max_{E_O,e,\mathbf{x}} \left[(E_I + E_O) \frac{r_E(\mathbf{x},e) - r_F}{1 + r_F} - (c(E_O) - E_O) \right], \tag{10}$$

s.t.

$$e \geq e_{\text{min}}(\mathbf{x}), \tag{11}$$

$$\mathbf{x} \geq \mathbf{0}. \tag{12}$$

This latter representation highlights that the market value is equal to the book value of initial equity holders' investment, E_I, plus the net present value of the loan portfolio from bank equity holders' perspective, minus the net issuance costs for outside equity, $(c(E_O) - E_O)$. The regulatory constraint (11) is another channel introducing non-linearities to a bank's maximization problem.

3. Analysis

We now analyze the competitive equilibrium of the economy.

Definition 1 *A Competitive Equilibrium is a yield function, an investment and effort strategy for each entrepreneur, an outside equity, equity ratio, and portfolio strategy for each banker, and an investment strategy for each investor such that:*

a) Given its type (q, ρ), the entrepreneur of each firm f decides whether to raise I units of capital at the equilibrium yield $y(q, \rho)$, and whether to shirk or not to maximize her expected utility.

b) Each banker b chooses net outside equity E_O, her equity ratio $e \geq \sum_{q,\rho} x(q, \rho) \cdot \underline{e}(\rho)$, and the vector of portfolio weights $\mathbf{x} \geq 0$ to maximize (10).

c) Investors decide on investments in the risk-free outside investment opportunity, firm debt, bank deposits, and bank outside equity to maximize their expected utility.

d) Markets for debt, deposits, and bank capital clear.

Our analysis of the equilibrium proceeds as follows. We first study the optimal behavior of an individual bank in partial equilibrium, that is, taking prices as given. In a second step, we determine the prices of all assets in the economy in general equilibrium.

3.1. Bank Optimization in Partial Equilibrium

It is convenient to separate the maximization problem of an individual bank (10) into two steps; a problem of optimal outside equity issuance on the one hand, and the jointly optimal portfolio and leverage choice on the other, that is,

$$E_{M,I} - E_I = \max_{E_O} \left[\frac{(E_I + E_O)(\max_{e,\mathbf{x}}[r_E(\mathbf{x},e)] - r_F)}{1 + r_F} - (c(E_O) - E_O), \right] \tag{13}$$

First, consider the inner (ROE) maximization problem, given the exogenous yields on loans $y(q, \rho)$:

$$\max_{\mathbf{x},e} [r_E(\mathbf{x}, e)] \text{ s.t. } e \geq e_{\min}(\mathbf{x}). \tag{14}$$

Given a solution $(\mathbf{x}^*, e^*)$ to this maximization problem, we define the set of a bank's failure states:

$$\Sigma_F(\mathbf{x}^*, e^*) \equiv \left\{ s \in S : \frac{r_A^s(\mathbf{x}^*) - r_F}{1 + r_F} < -e_{\min}(\mathbf{x}^*) \right\}. \tag{15}$$

In these states, a bank's assets are insufficient to cover the promised liabilities. We also define $\Sigma_S(\mathbf{x}^*, e^*)$ as the set of complementary survival states.

The following Lemma characterizes banks' optimal capitalization and portfolio choices.

Lemma 2 *Optimal bank capitalization e^* and portfolio choices $\mathbf{x}^*$ satisfy the following properties:*

i) ***Capitalization:*** *The capital constraint binds, that is, $e^* = e_{\min}(\mathbf{x}^*)$, if either*

1) there exists a portfolio $\mathbf{x}$ that yields $r_E(\mathbf{x}, e_{\min}(\mathbf{x})) > r_F$, or

2) for an optimal portfolio $\mathbf{x}^*$, failure states exist, $\Sigma_F(\mathbf{x}^*, e_{\min}(\mathbf{x}^*)) \neq \emptyset$.

ii) ***Portfolio choice:*** *In any bank failure state,* $s \in \Sigma_F(\mathbf{x}^*, e^*)$*, all issuer types in the loan portfolio experience sufficiently low excess returns relative to their respective capital requirement.*

$$\frac{r^s(q,\rho) - r_F}{1 + r_F} < -\underline{e}(\rho).$$

In contrast, in bank survival states, all loans in the portfolio feature sufficiently high excess returns, i.e., $\frac{r^s(q,\rho)-r_F}{1+r_F} \geq -\underline{e}(\rho)$*, and deliver the same levered return conditional on survival, that is,* $\mathbb{E}\left[\left.\frac{r^s(q,\rho)-r_F}{\underline{e}(\rho)}\right| \Sigma_S\right] = k$ *for all* $(q,\rho) : x(q,\rho) > 0$*.*

Capitalization. Part i.1 of Lemma 2 states that if the equilibrium loan yields allow banks to obtain a positive expected excess return on bank capital, banks have a strict incentive to choose the minimum capital-to-assets ratio allowed by the regulatory capital constraint. To understand part i.2, observe that upon bank default in some state s, government transfers to bank depositors are strictly decreasing in e. Total payments to all security holders are thus increasing in leverage, a key departure from the Modigliani-Miller benchmark. While these transfers accrue *ex post* to depositors, competition among investors on the deposit rate ensures that the present value of these transfers is passed on to bank equity holders *ex ante*. The present value of these transfers is the value of a put (see Merton, 1977).[29] Thus, shareholder value maximization requires the value of the put be maximized by minimizing the capital-to-asset ratio for any optimal portfolio $\mathbf{x}^*$.[30]

Portfolio choice. Lemma 2 highlights that optimally designed bank portfolios may consist of multiple, imperfectly correlated issuer types. Such portfolios exhibit correlated downside [tail risks in that for each state s, the losses on each investment either wipe out the associated regulatory capital cushions $\underline{e}(\rho)$, or none of them. Choosing exposures to correlated tail risks is an optimal response to convexity in a bank's objective function implied by deposit guarantees. To further illustrate the implications of these optimal portfolio choices, consider an example of a bank that can invest in safe US treasuries or risky Greek bonds. Suppose yields are such that investing exclusively in Greek bonds yields the same ROE as investing exclusively in US treasuries. Then, starting from a portfolio invested only in Greek bonds, the

[29]Once we endogenize loan yields in general equilibrium, banks pass on part of the put value to firms.

[30]Consistent with this prediction, Kisin and Manela (2016) show empirically that capital requirements are indeed effectively binding for the largest banks in the US economy. Moreover, these equilibrium features are consistent with the empirical results of Jiang, Matvos, Piskorski, and Seru (2020)), which suggest that deposit subsidies and capital regulation are indeed the two first-order determinants of banks' level of capitalization in practice.

bank will receive a *strictly* lower ROE if it marginally increases the portfolio weight of US treasuries. This is because the expected return on treasuries across the bank's *survival* states must be strictly lower than that for Greek bonds.[31] Conversely, starting from a portfolio with 100% US treasuries, a bank also strictly lowers its ROE when marginally increasing the portfolio weight of Greek bonds. After such a marginal deviation, the bank still does not default, and thus, lacks the benefit of a bailout put. Therefore, it cannot assign the same marginal value to a Greek bond as when being exclusively invested in Greek bonds. In short, bank specialization can naturally occur in our environment, shedding light on related recent evidence (see Rappoport et al., 2014).[32]

Outside equity issuances. Given a solution e^* and $\mathbf{x}^*$ yielding $r_E(\mathbf{x}^*, e^*)$, we can now characterize the incentives of an individual bank to issue outside equity (see the outer maximization problem in equation (13)).

Lemma 3 *A bank gains from marginally increasing date-0 capital as long as:*

$$\frac{r_E(\mathbf{x}^*, e^*) - r_F}{1 + r_F} > c'(E_O) - 1. \tag{16}$$

When deciding on equity issuances, a bank simply compares its expected date-1 expected excess return on bank capital, $r_E(\mathbf{x}^*, e^*) - r_F$, discounted at rate r_F, with the date-0 marginal net issuance costs for new bank capital, $(c'(E_O) - 1)$.

3.2. Prices and Allocations in General Equilibrium

We now analyze how prices and allocations are determined in general equilibrium. As highlighted in the introduction, a key feature of our approach is to derive the effective demand curve for bank capital, rather that a demand curve for credit. This approach is instructive as bank capital is the key scarce resource through which equilibration occurs. We derive a novel issuer-specific metric that allows us to construct this aggregate demand curve: an issuer type's effective *reservation price* for bank capital. This reservation price encodes all dimensions of issuer heterogeneity, and yields a univariate score that determines which issuers in

[31] Recall that we started with the supposition that exclusively investing in Greek bonds (and defaulting in some states) yields the same ROE as exclusively investing in US treasuries (and not defaulting).

[32] Moreover, in Appendix B, we discuss how these results extend to environments where banks differ ex ante in terms of characteristics such as legacy asset holdings.

the economy obtain bank finance. Importantly, this price has to be determined conditional on banks' optimal capitalization and portfolio choices (see Lemma 2) *and* conditional on an issuer's security being held by banks with portfolios that are the best matches for that security. These two conditions are relevant since portfolios will generally differ across banks in equilibrium (that is, multiple optimal bank portfolios coexist) and since, due to non-linearities, a security's co-movement with the rest of a bank's portfolio matters for the value to equity holders).

Going forward, we will refer to p as the date-0 market value bank equity holders obtain per unit of bank capital, that is, $p \equiv \frac{E_M}{E}$. In equilibrium, the marginal value attained per unit of bank capital is equalized across all loans provided in equilibrium, otherwise it would be optimal for a bank to deviate.[33] In contrast, interest rates generally differ across loans, accounting for bank and borrower specific attributes. Nonetheless, the interest rate on a loan in an efficient portfolio is increasing in the equilibrium value of p, as the net present value of the loan to bank equity holders is an increasing function of the interest rate charged.

We will first construct the aggregate supply and demand correspondences for bank equity, which we denote by $E^S = S(p)$ and $E^D = D(p)$ respectively. Market clearing then determines the equilibrium market price of bank capital p^*, and the equilibrium quantity E^*. Second, given E^* and p^*, we determine the equilibrium composition and pricing of credit in closed-form.

3.2.1. Aggregate Equity Supply and Demand

Aggregate supply of bank equity. Given Lemma 3, we immediately obtain the aggregate inverse supply function for bank equity:

$$S^{-1}(E) \equiv c'(E - E_I).$$

Note that this function represents the marginal cost of increasing bank capital at date 0. How this marginal cost relates to the *required return* on equity capital in equilibrium will be a result of our analysis below. As paying dividends is not associated with an additional cost, the inverse supply function is equal to one for $E < E_I$.

Aggregate demand for bank equity. To derive the aggregate demand for bank equity we

[33] The profitability index (see Berk and DeMarzo, 2014) of a loan is then equal to $(p-1)$.

initially determine for each issuer type her *effective* reservation price per unit of bank equity. Next, we construct the aggregate demand curve by aggregating across all issuer types in the economy.

An issuer type's effective reservation price per unit of bank equity is measured as a present value accruing to bank equity holders. The payments encoded in this reservation price come from both the issuer and the government (via deposit insurance). Thus, this metric is affected by both the traditional credit demand side (issuers) and factors affecting the credit supply side (regulations, government subsidies, and banks' optimal response to them). These two components of the reservation price are determined by the two Lemmas we have established thus far: first, the issuer's outside option in public markets (Lemma 1) pins down the maximum interest rate that an issuer is willing to pay for a bank loan. Second, banks' optimal leverage and portfolio decisions (Lemma 2) affect the magnitude of expected government subsidies, which are internalized by bank equity holders as debt is priced competitively.

Lemma 4 *An issuer of type (q, ρ) has the following reservation price per unit of bank capital:*

$$p^r (q, \rho) = 1 + \frac{NPV (q) \, \mathbb{1}_{\{B(q)>NPV(q)\}} + PUT (q, \rho)}{I \underline{e} (\rho)}, \tag{17}$$

where we define the date-0 put value:

$$PUT (q, \rho) \equiv \frac{\mathbb{E} \left[\max \left\{ I(1 - e(\rho))(1 + r_F) - C_s(q, 1), 0 \right\}\right]}{1 + r_F} \geq 0, \tag{18}$$

and where the demanded quantity of bank capital at this reservation price is $I \underline{e} (\rho)$.

The numerator of the ratio on the right-hand side of equation (17) reflects the *incremental* private surplus that bank financing of an issuer type (q, ρ) generates in excess of the surplus attainable under public market financing. As highlighted above, this incremental surplus has to be determined conditional on banks' optimal capitalization and portfolio choices and conditional on an issuer's security being held by banks with portfolios that are the best matches for that security. The incremental surplus derives from two sources.[34] First, it is attained for all projects that are bank-dependent (where $B(q) > NPV(q)$), as these projects would be credit-rationed under unmonitored public market financing. Second, incremental private surplus is attained whenever there is a positive probability that the government will cover a

[34]In Appendix B, we highlight that the concept of this reservation price naturally extends to the presence of other sources of bank-dependent surplus.

shortfall in payments to depositors that effectively funded this issuer type (captured by the term PUT) — this shortfall depends on the regulatory capital cushion for a given security, $\underline{e}(\rho)$, and a security's risk properties. Finally, the total incremental surplus is scaled by the effective equity capital demanded by the issuer, $I\underline{e}$, yielding the *per-unit* premium of the reservation price in excess of 1.

In our analysis below, we will further develop an asset pricing relation that reveals how security prices provide useful information on the magnitudes of PUT values in the cross-section of securities. Moreover, regulators have access to confidential data from stress tests specifically gauging securities' tail risk behavior in various adverse scenarios, which is directly relevant for quantifying the magnitudes of these PUT values across banks and securities.

Lemma 4 allows us to construct an *aggregate* demand correspondence by sorting issuer types according to their reservation prices $p^r(q,\rho)$. At a price p, all borrower types with $p^r(q,\rho) \geq p$ demand a quantity of bank equity equal to $I\underline{e}(\rho)$. Since issuer types are discrete, the inverse demand function is a step function (see Figure 1), which implies that the associated demand is a correspondence. Let $[\cdot,\cdot]$ denote the range operator, and let $m(q,\rho)$ denote the mass of issuers of type (q,ρ). Then the aggregate demand correspondence for bank equity, $D(p)$, is given by:

$$D(p) \equiv \left[\sum_{(q,\rho):p^r(q,\rho)>p} I \cdot \underline{e}(\rho) \cdot m(q,\rho), \sum_{(q,\rho):p^r(q,\rho)\geq p} I \cdot \underline{e}(\rho) \cdot m(q,\rho) \right]. \tag{19}$$

As Lemma 4 derived the reservation prices $p^r(q,\rho)$ in terms of exogenous parameters, the aggregate demand for bank equity is also expressed analytically. Since the reservation prices are both a function of social surplus and deposit insurance subsidies, issuers with the highest reservation price for bank equity are not necessarily those that create the greatest societal value. Going forward, we denote by $D^{-1}(E)$ the inverse aggregate demand function associated with (19).

Figure 1 illustrates the potential misalignment of the equilibrium demand for bank equity with the social surplus created by bank finance. Throughout, our graphs follow the familiar convention of price theory — we plot inverse demand functions, where the quantity of bank equity is plotted on the horizontal axis, and the price of bank equity on the vertical axis. The figure introduces an example with three issuer types that we will revisit at various points

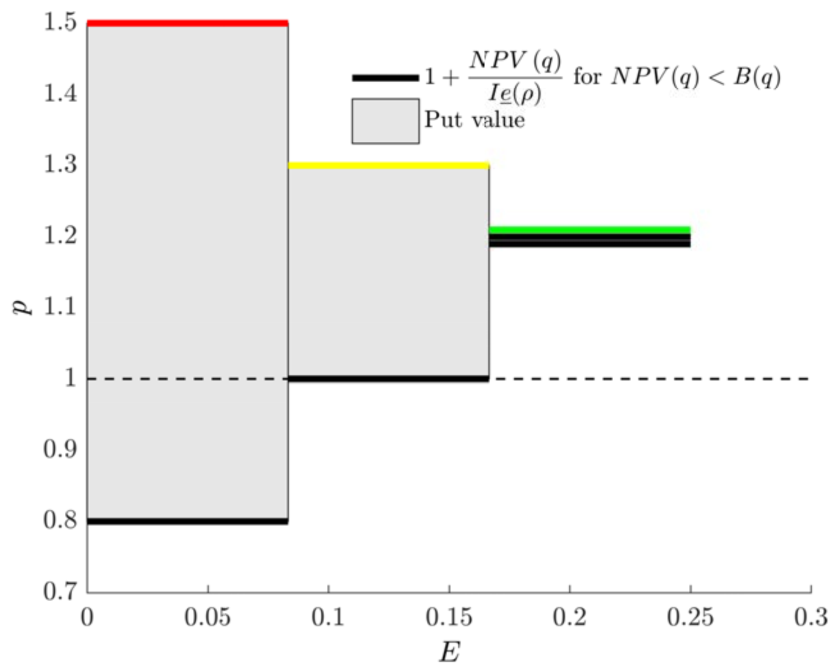

Figure 1. Demand for bank capital and bank-dependent private surplus. The graph illustrates the aggregate demand for bank capital in an economy with three issuer types, two equiprobable aggregate states, $r_F = 0$, $I = 1$, a general capital requirement of $\underline{e} = 25\%$, and $B(q) = 0.15$ for all issuer types. The three issuer types' reservation prices are indicated by the green, yellow, and red lines. Jointly, these reservation prices determine the aggregate demand correspondence. The green type is a good (positive NPV), safe borrower without access to unmonitored finance and project cash flows $C = (1.05, 1.05)$. The yellow type is a good, risky borrower with public market access and project cash flows $C = (1.8, 0.6)$. The red type is a bad (negative NPV), risky borrower and project cash flows $C = (1.5, 0.4)$. The black solid line indicates the social surplus (social NPV) that bank financing generates in excess of an issuer type's outside option from unmonitored finance, per unit of bank equity used. Since the yellow issuer type has access to unmonitored finance, the social value generated by bank financing is zero. For each type, the area between the reservation price and the black solid line measures the put value. Since the green issuer type is safe, the associated put value is zero.

of our analysis below. Throughout, these three issuer types will be indicated by the colors red, yellow, and green. Following a traffic light analogy, these color assignments indicate whether a bank should fund a borrower type, if the objective was to maximize social surplus. The red issuer type represents high-risk, negative-NPV borrowers, the yellow type high-risk, positive-NPV firms with access to public markets, and the green type bank-dependent, low-risk, positive-NPV issuers (see the figure caption for parameter values).

Figure 1 plots two curves, the aggregate inverse demand curve (in red, yellow, and green), and a curve representing the issuer types' *bank-dependent* social surplus per unit of equity capital used (in black). The vertical difference between these two curves, highlighted by the grey-shaded area, represents the wedge due to deposit insurance. The magnitude of this wedge is evidently issuer type-specific, revealing distortions in the ranking of issuers based on private surplus (green, yellow, red) relative to the one based on social surplus (black). In fact, in this example, the ranking is exactly inverted — the red type's reservation price is the highest even though the social surplus its projects create is the lowest (and negative); the green type's reservation price is the lowest but its bank-dependent social surplus is the highest. We will explore the implications of this misalignment and its dependence on various features of the economy in our comparative statics analyses below.

The following proposition derives the equilibrium price and quantity of aggregate bank equity.

Proposition 1 (Price and Quantity of Bank Capital) *The equilibrium amount of bank equity capital is given by*

$$E^* = \max\{E \geq 0 : D^{-1}(E) \geq S^{-1}(E)\}, \tag{20}$$

implying that aggregate outside equity issuances (or dividend payments) amount to

$$E_O^* = E^* - E_I. \tag{21}$$

The equilibrium value per unit of bank capital is given by:

$$p^* = \frac{E_M}{E^*} = S^{-1}(E^*). \tag{22}$$

To discuss the intuition underlying Proposition 1, we simply extend our example from Figure 1 by incorporating an inverse supply function. Figure 2 illustrates a standard case where the equilibrium is characterized by the intersection of demand and supply, that is, by the condition $D^{-1}(E^*) = S^{-1}(E^*)$.[35] In equilibrium, the market value of a unit of bank capital is p^*. This price is also the Lagrange multiplier on banks' equity capital constraint, a shadow value that a recent literature has estimated for banks and insurance companies (see, e.g., Koijen and Yogo, 2015, Kisin and Manela, 2016).

Given this equilibrium price, the distribution of surplus follows immediately. Bank surplus is positive if and only if p^* is strictly greater than 1, that is, if bank capital is scarce

[35] Due to discontinuities in the inverse demand function, $D^{-1}(E)$, it is also possible that demand and supply do not intersect. Such a case will be illustrated below in Figure 3.

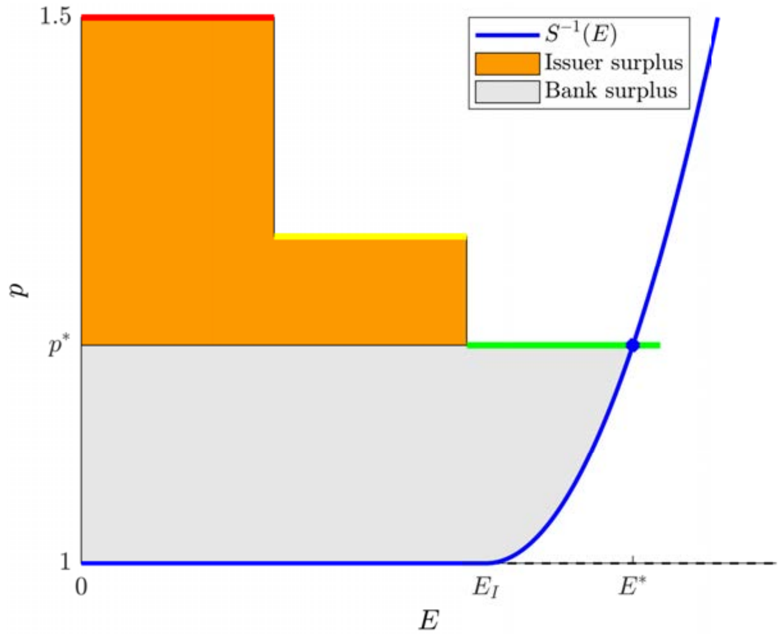

Figure 2. Equilibrium price and quantity of capital. The graph extends Figure 1 by adding an inverse supply function. The supply of bank capital is given by: $S^{-1}(E) = c'(E) = 1 + 50\,(\max\{E - E_I, 0\})^2$. The equilibrium quantity E^* and price p^* are indicated by the blue circle. The marginally funded borrower type is the green type. The incremental surplus that issuers obtain above and beyond the surplus attainable from public market finance is illustrated by the orange-shaded area. The grey-shaded area measures the surplus accruing to banks' initial equity holders.

$(E^* > E_I)$. On the other hand, the issuer surplus per unit of bank equity is given by the difference between an issuer's reservation price and the equilibrium price, that is, by $(p^r(q,\rho) - p^*)$. As standard in price theory, the marginal type receives zero surplus, and all inframarginal issuer types have reservation prices weakly greater than p^*. In Figure 2, we indicate issuers' incremental surplus from bank finance and bank surplus by the orange- and grey-shaded areas, respectively.

Figure 2 illustrates three relevant types of equilibrium outcomes that we will highlight throughout our analysis: (1) over-investment in surplus-destroying (red) issuer types (2) under-investment in bank-dependent (green) issuer types, and (3) crowding-out of public market financing in the sense that (yellow) issuer types with access to public markets obtain bank finance in equilibrium.

The following proposition shows how the equilibrium price of bank capital p^* in combination with the aggregate demand correspondence (19) directly characterizes the composition and pricing of credit in the economy.

Proposition 2 (Composition of Credit and Pricing) *All issuer types with* $p^r(q,\rho) > p^*$ *and a fraction* $\xi \in [0,1)$ *of borrower types with* $p^r(q,\rho) = p^*$ *are financed by banks.*[36] *These issuer types' equilibrium debt yields,* $y(q,\rho)$*, satisfy the following equilibrium relation for the expected return on debt:*

$$\mathbb{E}\left[r^s(q,\rho)\right] = r_F + \underline{e}(\rho)\left(r_E^* - r_F\right) - \frac{PUT(q,\rho)(1+r_F)}{I}. \tag{23}$$

Of the remaining issuers in the economy, only issuer types with $NPV(q) \geq B(q)$ *obtain unmonitored finance from public markets, and their expected return on debt satisfies:*

$$\mathbb{E}\left[r^s(q,\rho)\right] = r_F. \tag{24}$$

The expected excess return on bank capital follows from the price of equity p^**:*

$$r_E^* - r_F = (p^* - 1)(1 + r_F). \tag{25}$$

Proposition 2 provides a closed-form representation of the composition and pricing of credit.[37] The proposition highlights that the difference between a borrower's reservation price and the shadow price of bank capital, $(p^r(q,\rho) - p^*)$, is a *sufficient statistic* for bank funding. A borrower obtains bank funding if this statistic is weakly positive.

Equation (23) reveals that a CAPM type asset pricing relation holds for all securities held by banks. Yet, contrary to the classic CAPM, a security's expected return is not a linear function of its beta with respect to an aggregate risk factor. Instead, a security's expected return increases with its regulatory risk weight, which is interacted with the expected excess return on bank capital, $(r_E^* - r_F)$. This component of the expected return does not represent a risk premium, but rather compensation for a security's use of banks' scarce capital, which could be used profitably to extend loans to other (marginal) borrowers. In addition, the expected return is dampened by a security-specific term, $\frac{PUT\cdot(1+r_F)}{I}$, that reflects the implicit pass-through of deposit subsidies per unit of investment.

[36]Here, $\xi = \frac{E^* - \sum_{(q,\rho):p^r(q,\rho)>p^*} I\cdot\underline{e}(\rho)\cdot m(q,\rho)}{\sum_{(q,\rho):p^r(q,\rho)=p^*} I\cdot\underline{e}(\rho)\cdot m(q,\rho)}$.

[37]In knife-edge cases where multiple issuer types (q,ρ) have the same reservation price $p(q,\rho)$, a tie-breaker rule can ensure the uniqueness of the equilibrium allocation in terms of the masses of each issuer type that obtain bank finance. One such tie-breaker rule is to assume that among issuer types with identical reservation prices, banks rank issuer types according to the incremental social surplus they create under bank finance, $NPV(q)\,\mathbb{1}_{\{B(q)>NPV(q)\}}$.

Risky securities that contribute to a bank's tail risk and that have low regulatory risk weights tend to have larger PUT values. As a result, the pricing relation (23) predicts that these types of securities, if held by regulated levered institutions, may command *negative* expected excess returns relative to the risk-free rate. This prediction, which is unusual for a rational framework, sheds light on empirical evidence suggesting that in certain sub-periods of the leverage cycle, fixed income securities yield negative expected excess returns (see, e.g., Greenwood and Hanson, 2013).[38] Conversely, this pricing relation also predicts that the returns of securities held by regulated levered institutions are informative about both capital scarcity and the magnitude of implicit bailout subsidies.

As regulatory risk classifications (e.g., based on credit ratings) affect a security's risk weight, they crucially affect both the pricing and the allocation of bank credit. We will discuss this issue in more detail in our comparative statics analysis below. Finally, equation (25) provides a mapping between the price of equity and banks' expected excess returns on *book* equity. When bank capital is not scarce, $p^* = 1$, it does not yield an excess return relative to other outside investment opportunities.

Remarkably, the tractable pricing relation (23) holds for *all* securities of bank-funded issuers despite the fact that marginal investors across various securities differ — the cross-section of banks is generally exposed to heterogenous risks (due to heterogenous equilibrium investment strategies). The following corollary highlights the diversity of banks' investment portfolios.

Corollary 1 (Heterogeneous bank portfolios) *Suppose two issuer types that do not exhibit correlated tail risks (see Lemma 2) are funded in equilibrium, then the two issuer types must be financed by different banks.*

A bank typically invests in a continuum of borrowers (i.e., an infinite number), but as these borrowers exhibit correlated tails risks, doing so does not yield diversification with respect to relevant tail risks.

[38] Banks' securities holdings account for about 20% of their assets (see Laux and Leuz, 2010, Abbassi et al., 2016). In addition, insurance companies, which are also regulated levered institutions that may be subject to implicit too-big-to-fail guarantees, hold a large fraction of corporate debt.

4. Positive and Normative Implications

In this section, we derive positive and normative implications of our model. To do so, we will analyze and illustrate how equilibrium outcomes vary as a function of the bank capital supply, regulatory capital ratio requirements, public market development. In this context, we will repeatedly consider a useful summary measure of efficiency — the total surplus that firm investment creates in the economy, that is, the sum of the surpluses created by all projects financed in equilibrium. For brevity, we will refer to this object simply as *total surplus* going forward.

4.1. Equity Capital Supply

As highlighted in the introduction, the financial accelerator literature following Bernanke and Gertler (1989) has identified bank net worth as a key state variable affecting growth and allocative efficiency. A key object of interest for our study is how variation in this aggregate state variable has *heterogenous* effects across different borrower types in the economy. In practice, various economic shocks can lead to declines or increases in bank capital. For example, a macroeconomic downturn is typically associated with higher loan default rates, and correspondingly, declines in bank net worth. On the other hand, equity capital injections by governments during crises can increase aggregate bank capital (see, e.g., Giannetti and Simonov, 2013).

The following Corollary to Propositions 1 and 2 summarizes how changes to aggregate bank capital affects prices and allocations in the economy. To streamline the presentation, we focus on economies where the finite number of borrower types (q, ρ) have distinct reservation prices p^r, which eliminates knife-edge cases.

Corollary 2 *A decline in the aggregate amount of inside bank capital E_I*

1. *weakly increases the equilibrium price of bank capital p^*, the expected return on bank capital r_E^*, and loan yields $y(q, \rho)$,*
2. *weakly decreases aggregate investment, but weakly increases unmonitored funding by public markets.*
3. *The local effect on total surplus from firm investment is*

(a) *negative, if the marginal borrower type satisfies* $0 < NPV < B$,

(b) *neutral, if the marginal borrower type satisfies* $NPV > B$,

(c) *positive, if the marginal borrower type satisfies* $NPV < 0$.

Figure 3 illustrates the effects of shocks to banks' inside equity, building on our earlier example with three issuer types (Figures 1 and 2). These shocks affect only the equity supply curve, shifting it outwards (or inwards), from the solid blue line to the dashed blue line (or dotted black line). As a result of the considered increase (to the dashed blue line), the equilibrium price of equity p^* drops from the reservation price of the green issuer type to one, reflecting that bank equity capital is no longer scarce. Whereas some issuers of the green type (who have positive-NPV projects) were rationed at the initial level of equity (solid blue line), this is no longer the case after the increase. While more abundant equity capital resolves this rationing of green issuer types, it does not reduce allocative inefficiencies caused by the funding of red issuer types.

On the other hand, the considered decrease in equity capital (to the dotted black line) causes the equilibrium price of equity p^* to rise to the reservation price of the yellow issuer type. As a result, all issuers of the green type are rationed, reducing total surplus. The remaining issuer types that receive bank funding either destroy surplus (red types) or could also be funded by public markets (yellow types).

More generally, negative shocks to a financial system's capital affect borrowers in the order in which they are ranked in the demand curve for bank capital. Marginal borrowers are most affected. In contrast, borrowers that are ranked high in the demand curve are least likely to be rationed. The fundamental quality of borrowers with reservation prices close to the shadow price of bank capital thus determines whether such rationing has positive or negative implications for total surplus. Our explicit formula (17), provides clear predictions on how various borrower characteristics determine these reservation prices.

4.2. Capital Ratio Requirements

A quickly-developing macroeconomic literature evaluates capital requirements as a macroprudential tool used by policy makers to stabilize and support economic growth. Whereas this literature typically directly specifies banks' investment technologies, our objective is to shed light on relevant compositional effects. In particular, in this section, we examine the im-

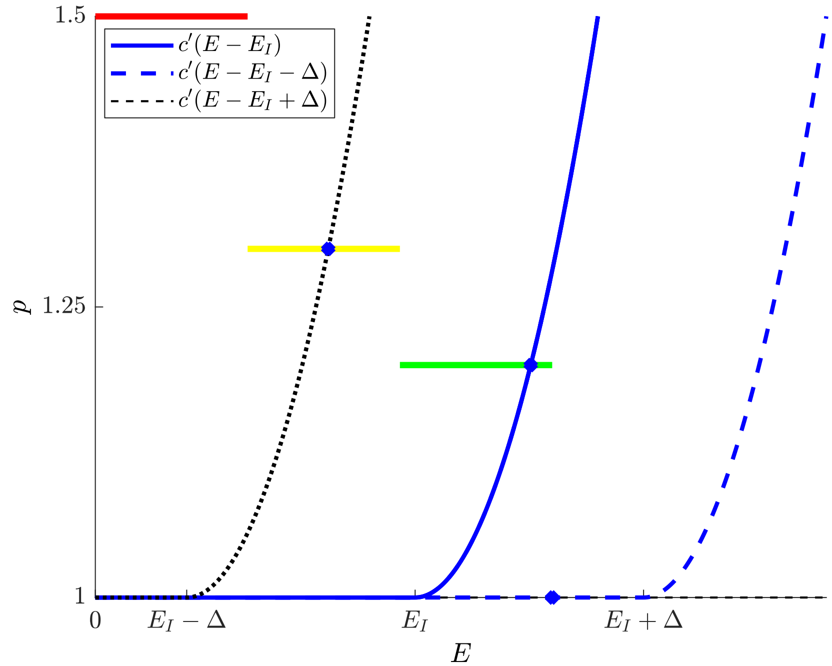

Figure 3. Bank equity capital supply. The graph illustrates how equilibrium outcomes are affected by an increase or decrease in inside capital E_I relative to the baseline level considered in Figure 2. We consider changes of magnitude $\Delta = 0.125$.

plications of changes to overall equity capital ratio requirements $\underline{e}$, and of risk-classification specific changes, that is, adjustments to the risk weights $rw(\rho)$.

Corollary 3 *The following comparative statics with respect to capital ratio requirements $\underline{e}$ apply:*

1. *An increase in capital ratio requirements from $\underline{e}$ to $\underline{e} + \varepsilon$ (with $\varepsilon > 0$)*
   - *(a) weakly increases loan yields for at least one borrower type in the economy. It may decrease loan yields for some borrower types if and only if capital is scarce before the increase.*
   - *(b) weakly decreases aggregate investment,*
   - *(c) weakly increases total surplus if bank capital is not scarce after the increase.*
2. *For $D^{-1}(E^*) > p^*$, a marginal increase in $\underline{e}$ is compensated by additional equity issuances, leaving aggregate bank funding unchanged.*

For $D^{-1}(E^*) = p^*$, a marginal increase in $\underline{e}$ strictly reduces the fraction of firms of the marginal type that receive bank funding.

(a) If the marginal issuer type is bank-dependent and has a negative (positive) NPV, this reduction in bank funding has a positive (negative) impact on total surplus.

(b) If the marginal issuer type is not bank-dependent, then total surplus is unaffected.

Conventional wisdom suggests that an increase in system-wide capital ratio requirements increases the cost of capital for all firms. Since Part 1a) confirms that this effect applies for at least one borrower type, it holds for all firms provided that they are *identical*. Yet our framework reveals that when borrowers are heterogeneous, this result generally no longer applies. To understand why this is the case, it is instructive to revisit the three channels through which capital requirements affect a firm's cost of capital according to our asset pricing relation (23): (1) the required use of costly equity capital funding $\underline{e}(\rho)$, (2) the PUT value, and (3) the expected return on bank equity r_E^*. The first two channels unambiguously support the conventional wisdom based on standard partial equilibrium intuition: Ceteris paribus, the increased use of scarce capital (channel 1) and the reduction in the put value (channel 2) both increase the firm's cost of capital. However, the effect on the expected return on bank equity, which is determined in equilibrium, is ambiguous. A sufficient decline in that rate may overturn the impact of the first two channels.[39] Our subsequent discussion of an example illustrated in Figure 4 highlights that, for instance, safe borrowers may get strictly better financing terms after capital ratio requirements have been *increased*.

An increase in overall capital ratio requirements $\underline{e}$ might a priori be considered a credit supply shock for all borrowers in the economy. Yet, consistent with the results thus far, it is actually only the group of marginal borrowers that is affected in terms of quantities by *small* adjustments to capital requirements, highlighting the importance of those borrowers that have reservation prices close to the shadow price of bank capital. Moreover, the implications of *larger* shocks crucially depend on the cross-sectional distribution of potential borrowers, and the responses in their reservation prices.

To illustrate the effects of both small and large shocks, we revisit our baseline example with three issuer types introduced in Figures 1 and 2. We start by considering increases in the overall capital ratio requirements $\underline{e}$, and then consider more targeted interventions that change the risk-weights of securities with specific risk classifications ρ.

[39]Of course, the equilibrium rate of return can only decrease if bank capital is scarce to begin with.

Overall capital ratio requirements. The four panels of Figure 4 illustrate demand and supply curves under distinct capital ratio requirements $\underline{e}$. As an initial reference point, Panel A simply replicates the baseline parameterization of Figure 2. Panels B to D, in turn, illustrate the effects of gradual increases in the equity ratio requirement $\underline{e}$ (small, medium, and large) relative to this benchmark.

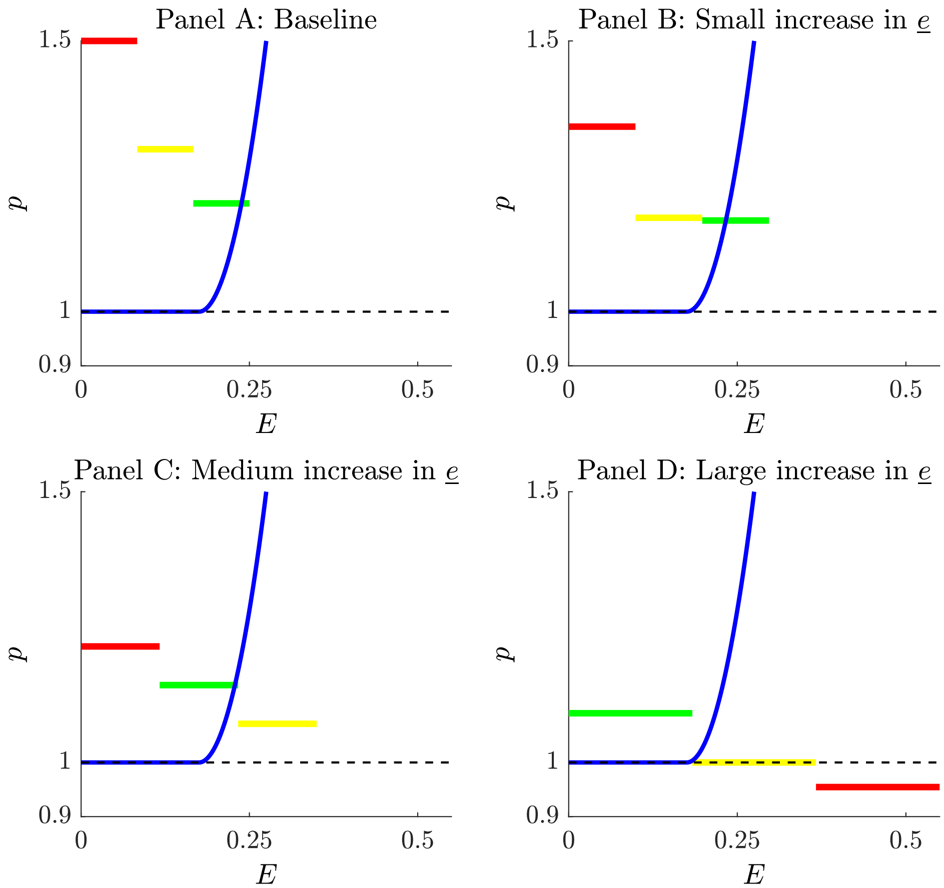

Figure 4. Capital ratio requirements. Panels A through D illustrate the effects of increases in capital ratio requirements. Panel A replicates the economy illustrated in Figure 2, where all borrower types are subject to a ratio requirement of $\underline{e} = 0.25$. Panels B through D consider gradual increases in capital ratio requirements, up to a level of $\underline{e} = 0.55$ in Panel D.

Changes to the overall capital ratio requirement affect only the demand curve, causing three types of adjustments. First, all issuer types' reservation prices are reduced, which is graphically reflected by a downward adjustment in the demand curve. This effect follows immediately from the fact that reservation prices reflect incremental private surplus attainable *per unit* of equity used (see equation (17)). As ratio requirements are increased, more units of equity are required to fund any borrower type, lowering the per-unit surplus. Second, the downward adjustments in reservation prices are issuer-type specific. Those issuer types whose reservation price is more reliant on the PUT-component of private surplus exhibit stronger downward adjustments. As a result, the ranking of issuer types within the demand curve can change as $\underline{e}$ is increased. Third, the demand curve pans out to the right, that is, the width of each borrower type on the demand curve increases, as more equity capital is required to fund the borrowers of any type.

The graphs reveal that changes to overall capital ratio requirements are a fairly blunt tool. On the one hand, increases can have the desirable effect of aligning the private ranking of borrower types with the ranking based on social surplus — the "large increase" in $\underline{e}$ considered in Panel D achieves this result. A better alignment obtains as greater skin in the game reduces distortions introduced by the PUT component affecting the demand for bank capital. On the other hand, increases in ratio requirements can also cause the rationing of surplus-generating bank-dependent borrowers — the "small increase" considered in Panel B for example shows a case where that type of rationing is more severe than in the baseline economy with the lowest ratio requirements.

More generally, the graphs highlight that changes to ratio requirements potentially have a non-monotonic effect on the rationing of good, bank-dependent borrowers (green types). Whereas small increases in ratio requirements worsen rationing of good borrower types, medium and large increases can completely alleviate their rationing. This result obtains as small increases in ratio requirements broadly increase the demand for equity without changing the ranking of borrower types within the demand curve. Yet, for large enough increases in ratio requirements, good borrowers obtain a higher ranking, thus giving them priority in access to bank finance. Moreover, since the financing of other borrowers eventually becomes unprofitable (see Panel D), the existing equity capital is only used to fund green types. Somewhat paradoxically, increases in capital ratio requirements can therefore free up bank capital and cause it to be no longer scarce (cf. Panel A versus Panel D), as reflected by a price of equity capital $p^* = 1$ (and, equivalently, $r_E^* = r_F$). This general equilibrium effect also makes it possible that the green firm type experiences a strict decrease in its cost of capital, see (23), which allows the entrepreneur to extract the entire NPV from the project.

Yet, high ratio requirements are not per se a guarantee for improved allocative efficiency. If ratio requirements were increased beyond the level considered in Panel D, the total equity capital required to fund all borrowers of the green type would increase further (graphically, the width of the green types demand segment would increase), and at some point, surplus-generating bank-dependent borrowers would again be rationed.

Overall, these illustrations highlight that increases in ratio requirements can have the desirable effect of better aligning the private demand for bank capital with the ranking based on social surplus. Yet, they also reveal potential adverse effects due to the increased reliance on bank capital for the funding of *any* borrower type, a channel that can cause the rationing of surplus-generating bank-dependent borrowers.

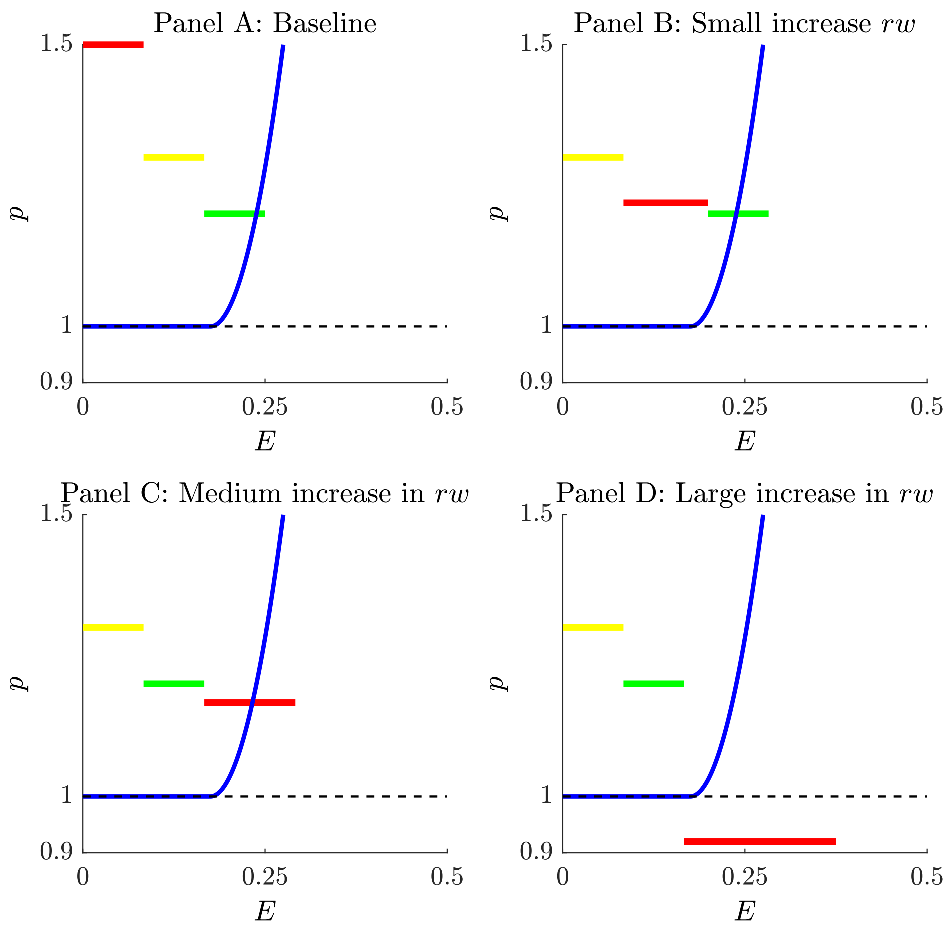

Figure 5. Risk weights. Panels A through D illustrate the effects of increases in the risk-weight applying to red borrower types, assuming that regulatory risk classifications perfectly identify this type. Panel A replicates the economy illustrated in Figure 2, where all borrower types are subject to a ratio requirement of $\underline{e} = 0.25$ (that is, all borrower types have a risk weight of 1). Panels B through D consider gradual increases in the red type's risk weight, up to a level of 2.5 in Panel D.

Risk weights. Next, we consider policy makers' opportunity to undertake more targeted adjustments to capital requirements, specifically by changing risk weights that are contingent on the risk classifications ρ. One of the major changes in the regulatory frameworks from Basel I to Basel II was the introduction of such risk weights that are contingent on external ratings. A similar system of risk-based capital requirements was introduced for U.S. insurance companies in 1994. Yet, the ratings used in regulations in practice are generically noisy and incomplete, that is, they pool multiple types of borrowers. In fact, regulations used in practice even pool borrowers of multiple ratings classes. For example, capital regulations applying to U.S. insurance companies impose the same "risk based capital charges" no matter if a corporate bond is rated AAA, AA, or A (see Becker and Opp, 2013, Becker et al., 2020). Due to the associated pooling of borrowers, changing risk weights for specific risk classifications then generically involves the same types of trade-offs as the ones discussed above for overall capital ratio requirements (in the provided examples, effectively three borrower types were pooled under one risk classification). In particular, whereas increasing risk weights tends to reduce the funding of surplus-destroying risky borrowers of a given risk-classification, they can also cause bank-dependent surplus-generating borrowers with the same risk classification to be rationed.

Yet, even when risk classifications are perfectly precise, changes to risk weights generally have non-trivial implications. In particular, counter to conventional wisdom, focused risk weight changes do not per se imply a credit supply shock for the targeted borrowers. Instead, it is again the group of marginal borrowers that is affected by local changes. We illustrate the effects of small and large changes to risk weights in Figure 5. The figure follows a format similar to that of Figure 4. Panel A again replicates the baseline parameterization from Figure 2, and Panels B through D consider changes to capital requirements. Yet, now, only the risk weights applying to borrowers of the red type are increased. This type of policy intervention thus presumes that the regulator has access to regulatory risk classifications that perfectly identify only the borrowers of the red type. Conditional on having access to these precise classifications, there is no downside to imposing *large* increases in the risk weight for red types, as investment in these risky types projects' always reduces expected total surplus. In fact, sufficiently large changes that cause the rationing of red types (see Panels C and D) can help free up capital that is then directed to borrowers of the green type, which were previously partially rationed (see Panels A and B).

Yet, even when policy interventions can be targeted with that much precision, small increases in risk weights can harm allocative efficiency. In particular, the change from the baseline level to the one considered in Panel B *reduces* total surplus. This result obtains as the considered risk weight increase is insufficient to cause the rationing of red borrowers. Instead, red types remain inframarginal borrowers and simply use more of banks' equity capital — graphically, the red segment of the demand curve widens. As a result, additional marginal green borrowers are crowded out, causing increased rationing of beneficial bank-dependent investment. This result reveals potentially important interactions and spill-over effects occurring even when a policy maker can adjusts risk weights based on perfectly precise risk classifications.

4.3. Development of Public Markets

The development and accessibility of credit from sources other than banks varies considerably across countries (see, e.g., Rajan and Zingales, 1995, 1998). Moreover, countries have been affected, to varying degrees, by long-term trends associated with financial innovations. These trends have had the implication that borrowers have obtained better access to alternatives to the funding provided by regular banks. For example, important innovations have included the development of junk bond markets in the 1980s, securitization and shadow banking in the 2000s, and most recently, the development of FinTech funding platforms, such as those facilitating crowdfunding. Despite this variation in the cross-section and over time, the rules governing bank capital requirements have changed very infrequently, and following the Basel accords, a large set of countries has instituted very similar rules. In this section, we analyze how a given set of rules for capital requirements can have starkly different allocative implications across economies that differ in borrowers' access to non-bank funding, which we broadly term "public market development." In the context of our model, access to public markets is affected by both the surplus a borrower's projects generate (NPV) and the moral hazard rent B that is attainable absent bank monitoring. The more developed public markets are, the lower is this moral hazard rent, and the fewer firms have to rely on banks as the sole source of finance.

Figure 6 illustrates the implications of improvements in public markets that lower the moral hazard frictions in these markets for all borrower types. As detailed in the figure's caption, the graphs again build on our baseline Figure 2, subject to a few adjustments. In Panel A, the moral hazard friction in public markets is large ("High B"), implying that both green and yellow borrower types do not have access to this source of finance. Lacking this outside option, these borrower types are highly profitable for banks, as measured by their high reservation prices for bank capital. Given these high reservation prices, banks use their

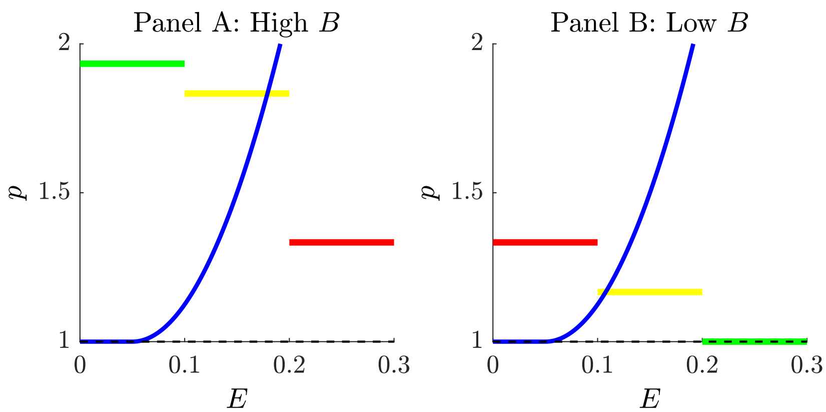

Figure 6. Changes in public market development. The figure illustrates the effect of a decrease in the parameter B for all borrower types from 0.3 (Panel A) to 0.15 (Panel B). The Panels of the figure build on our previous benchmark parameterization shown in Figure 2, subject to the following adjustments: the green type now has cash flows $C = (1.28, 1.28)$, the general capital requirement is $\underline{e} = 30\%$, and $E_I = 0.05$.

scarce capital to extend credit to green and yellow borrower types only. Surplus-destroying red borrower types are rationed.

In contrast, in Panel B, the moral hazard frictions in public markets are lower ("Low B"), causing green and yellow types to have access to these markets. Moreover, since borrowers of the green type also have safe cash flows, these borrowers do not create any incremental private surplus with bank finance (as the PUT component is also zero). As a result, the green type's reservation price for bank capital drops to one, causing this type to move off the banking sector's balance sheet. In contrast, the yellow type, while not bank dependent, does generate some incremental private surplus with bank funding, as the PUT value is positive. Yet, the yellow borrower type's reservation price does drop relative to the regime with less developed public markets depicted in Panel A, as the PUT value becomes the sole source of incremental private surplus from bank finance. Finally, the reservation prices of surplus-destroying risky red borrower types are unaffected by the change in public market development, as public markets are in any case not a feasible source of funding for these borrowers. As a result, red borrowers end up becoming those with the relatively highest reservation prices, and therefore start to obtain bank finance. In sum, the model reveals banks' increased incentives to focus on reaching for yield (instead of using monitoring abilities) after public markets become more efficient and a greater competitive threat.

5. Case Studies

In this section, we illustrate how our conceptual approach of an aggregate demand function for bank capital can be used to shed light on important crisis episodes that have been in the focus of extensive empirical research. As already mentioned in the introduction, a key contributing factor to the *Japanese crisis* were deregulations that improved public market access for large firms. Our framework predicts that such increased competition faced by banks for a subset of borrowers naturally causes safe large firms to rank lower in the aggregate demand curve for bank capital, and conversely, riskier firms to rank relatively higher (see also the comparative statics analysis in Section 4.3). This mechanism can help explain the crowding out of safe bank lending documented by Hoshi and Kashyap (1999, 2001) and Caballero et al. (2008). We now discuss the more recent financial crisis and the subsequent European debt crisis through the lens of our framework.

Financial crisis in the U.S. (2007/08). The fact that sophisticated financial institutions were holding large amounts of "toxic" structured securities on their balance sheets was a key reason for the severity of the 2007/08 financial crisis (Diamond and Rajan, 2009). In the terminology of our model, this observation raises the question why so many risky assets ranked highly in the aggregate demand curve for bank capital, even when the underlying investments in real estate were inefficient from an ex-ante perspective. An explanation consistent with our model is that the popular practice of securitization in the pre-crisis period generated an unusually large supply of securities with a high PUT value.

A key force behind this increased supply was the possibility to economize on capital requirements by securitizing a loan pool even if the risk of the loan pool was ultimately still borne by the bank (see, e.g., Acharya et al., 2013). Since the "savings" in regulatory capital requirements for securitization tranches were linked to their ratings, profit-maximizing credit agencies in turn responded to the demand for highly-rated securities by increasing their supply (see, e.g., Opp et al., 2013). As a result, by 2007, 60% of collateralized debt obligations were rated AAA (Fitch, 2007). At the same time, the very design of the structuring process implied that the highly rated tranches were exposed to high tail risk , akin to "economic catastrophe bonds" (see Coval et al. (2009a), Coval et al. (2009b)). In sum, the combination of high tail risk, rating-contingent capital requirements, and rating inflation generated a large supply of securities with high PUT value, causing severe distortions in the aggregate demand curve.

These distortions have several immediate implications. First, if we view subprime homeowners as a borrower type in our model, our framework predicts "real" overinvestment in the housing sector. Second, since overall capital requirements in the pre-crisis period were so low that bank capital was not scarce, the reaching-for-yield-behavior by competitive financial institutions implied that the put value was passed on to borrowers in the form of too low loan yields, consistent with empirical evidence for low risk-premia in the pre-crisis period (Muir, 2017). Within our framework, when banks (and similarly, insurance companies) become marginal investors in publicly traded debt, they may bid up prices to the point where these securities earn negative expected excess returns (see equation 23), consistent with empirical evidence by Greenwood and Hanson (2013).[40] Via this risk-taking mechanism, our theory thus also predicts a rational overvaluation of the underlying real estate, relative to a frictionless benchmark.

Further, recent empirical research has produced more detailed micro-level evidence identifying the risk-taking channel underlying our narrative. Relying on institutional imperfections of capital regulation, Becker and Ivashina (2015) and Iannotta et al. (2018) have identified "reaching-for-yield" behavior by both insurance companies and banks, respectively, by exploiting variation of "risk" *within* capital requirement buckets.[41] Based on this reaching-for-yield behavior, our framework predicts that risk signals used for regulation, such as credit ratings, will be reflected in prices (controlling for cash flow characteristics q). A recent study by Kisgen and Strahan (2010) finds direct evidence in support of this implication.

European debt crisis (2010/12). In the aftermath of the Financial crisis, European banks substantially increased their portfolio share of government bond holdings precisely at a time when the credit risk of these sovereign debt positions went up due to rising budget deficits. For example, the portfolio share that Italian banks allocated to Italian government bonds increased from 5% in 2008 to over 10% in 2012 (see SEB, 2018). A higher ranking of sovereign debt in the aggregate demand is consistent with the view that the private sector lacked profitable investment opportunities, whereas the PUT value associated with sovereign debt increased substantially. A key factor for this increase in PUT value (and, hence, reservation prices) was that the increase in the sovereigns' credit risk was not counterbalanced by corresponding increases in regulatory risk-weights. Instead, the *Capital Requirement Directive*

[40] See also footnote 38 for related evidence.

[41] For example, capital regulations applying to U.S. insurance companies impose the same "risk based capital charges" no matter if a corporate bond is rated AAA, AA, or A (see Becker and Opp, 2013). Within our model, this may be interpreted as variation of q holding $\underline{e}(\rho)$ fixed.

assigns a zero-risk weight for "exposures to Member States' central government [...] denominated and funded in the domestic currency of that central government" (see Hannoun, 2013), *regardless of credit risk*. Consistent with the view that risk-taking incentives were instrumental for the increase in the portfolio share allocated to sovereigns, banks' overall portfolios exhibited correlated tail risks (see prediction in Lemma 2), which was further facilitated by a removal of concentration limits for sovereign debt exposures by Eurozone regulators: A "home-bias" in sovereign debt holdings in the sense of Greek banks holding Greek sovereign debt (see empirical evidence by Acharya and Steffen (2015)) ensures that losses on sovereign debt positions occur precisely in states of the world where the bank defaults on obligations to its own creditors.[42]

In turn, the aggregate consequences of risk-taking behavior by European banks were far more severe than a redistribution of wealth from tax payers to bank equity holders.[43] First, the lack of "market discipline" induced by banks' risk-taking behavior aggravated the magnitude of the European sovereign debt crisis by facilitating excessive borrowing ex ante. Second, empirical evidence by Acharya et al. (2014) shows that bank risk-taking caused negative real effects by crowding out lending to small and medium-sized firms: Since public markets are not as developed in Europe (see, e.g., Rajan and Zingales, 1995), many of these firms did not possess a viable outside option to bank finance so that credit rationing resulted from the above described change in the ranking of borrowers in the aggregate demand curve.

6. Conclusion

An influential literature in macroeconomics and banking highlights bank capital as a key state variable affecting aggregate economic outcomes. In this study, we propose a transparent and flexible framework to analyze which types of borrowers in an economy are most affected by shocks relating to bank capital and the regulations governing it. To do so, we develop a novel approach to characterizing the credit market equilibrium based on a micro-founded aggregate demand function for bank capital. Despite the presence of multi-dimensional borrower heterogeneity, this approach yields transparent predictions for the composition and pricing of credit.

[42]See further discussion in Appendix B, where we address how our results extend to legacy assets.

[43]If there are positive marginal social cost of public funds (as in Farhi and Tirole (2017)), then even pure transfers to the banking sector are distortionary.

The demand curve central to our analysis is based on borrowers' reservation prices for bank capital. These reservation prices are shown to have an economically intuitive representation and provide sharp predictions on the behavior of bank funding. In particular, the difference between a borrower's reservation price and the shadow value of bank capital is a sufficient statistic for the provision of bank credit.

Existing empirical studies analyzing micro-level bank data typically recognize that credit demand and supply are materially affected by borrower heterogeneity and factors linking credit demand and supply curves across borrower-bank pairs (Khwaja and Mian, 2008). To limit confounding factors, this literature often focuses on outcome variation at the borrower level, which, however, provides limited insights on compositional effects at the aggregate level. The approach proposed in this paper — to determine loans' reservation prices for bank capital — might provide useful conceptual guidance for future studies analyzing the complex behavior of the composition of credit and its importance for macroeconomic stability and efficiency. Such studies will be particularly valuable in light of the emergent COVID-19 pandemic, which severely affects both the financial sector and the cross-sectional distribution of borrowers.

References

Abbassi, P., R. Iyer, J.-L. Peydró, and F. R. Tous, 2016, Securities trading by banks and credit supply: Micro-evidence from the crisis, *Journal of Financial Economics* 121, 569–594.

Acharya, V. V., T. Eisert, C. Eufinger, and C. W. Hirsch, 2014, Real effects of the sovereign debt crisis in europe: Evidence from syndicated loans, Technical report, New York University.

Acharya, V. V., P. Schnabl, and G. Suarez, 2013, Securitization without risk transfer, *Journal of Financial economics* 107, 515–536.

Acharya, V. V., and S. Steffen, 2015, The "greatest" carry trade ever? understanding eurozone bank risks, *Journal of Financial Economics* 115, 215 – 236.

Admati, A. R., P. M. DeMarzo, M. F. Hellwig, and P. C. Pfleiderer, 2011, Fallacies, irrelevant facts, and myths in the discussion of capital regulation: Why bank equity is not expensive, Working paper series, Stanford University.

Admati, A. R., P. M. DeMarzo, M. F. Hellwig, and P. C. Pfleiderer, 2018, The leverage ratchet effect, *Journal of Finance* 73, 145–198.

Allen, F., E. Carletti, and R. Marquez, 2011, Credit market competition and capital regulation, *Review of Financial Studies* 24, 983–1018.

Allen, F., E. Carletti, and R. Marquez, 2015, Deposits and bank capital structure, *Journal of Financial Economics* 118, 601–619.

Atkeson, A. G., A. d'Avernas, A. L. Eisfeldt, and P.-O. Weill, 2018, Government guarantees and the valuation of american banks, Working Paper 24706, National Bureau of Economic Research.

Bahaj, S., and F. Malherbe, 2018, The forced safety effect: How higher capital requirements can increase bank lending, Technical report, University College London.

Becker, B., and V. Ivashina, 2015, Reaching for yield in the bond market, *Journal of Finance* 70, 1863–1902.

Becker, B., and M. Opp, 2013, Regulatory reform and risk-taking: replacing ratings, Working Paper 19257, National Bureau of Economic Research.

Becker, B., M. Opp, and F. Saidi, 2020, Regulatory forbearance in the u.s. insurance industry: The effects of eliminating capital requirements, Working Paper DP14373, CEPR.

Begenau, J., forthcoming, Capital requirements, risk choice, and liquidity provision in a business cycle model, *Journal of Financial Economics* .

Berk, J. B., and P. M. DeMarzo, 2014, *Corporate finance*, fourth edition (Pearson Education).

Bernanke, B., and M. Gertler, 1989, Agency costs, net worth, and business fluctuations, *The American Economic Review* 79, 14–31.

Bianchi, J., 2016, Efficient bailouts?, *American Economic Review* 106, 3607–59.

Bolton, P., H. Chen, and N. Wang, 2013, Market timing, investment, and risk management, *Journal of Financial Economics* 109, 40 – 62.

Boot, A., S. I. Greenbaum, and A. Thakor, 1993, Reputation and discretion in financial contracting, *American Economic Review* 83, 1165–83.

Caballero, R. J., T. Hoshi, and A. K. Kashyap, 2008, Zombie lending and depressed restructuring in japan, *American Economic Review* 98, 1943–77.

Chari, V. V., and P. J. Kehoe, 2016, Bailouts, time inconsistency, and optimal regulation: A macroeconomic view, *American Economic Review* 106, 2458–93.

Chemmanur, T. J., and P. Fulghieri, 1994, Reputation, Renegotiation, and the Choice between Bank Loans and Publicly Traded Debt, *Review of Financial Studies* 7, 475–506.

Chodorow-Reich, G., 2014, The employment effects of credit market disruptions: Firm-level evidence from the 2008–9 financial crisis, *The Quarterly Journal of Economics* 129, 1–59.

Coval, J. D., J. W. Jurek, and E. Stafford, 2009a, Economic catastrophe bonds, *American Economic Review* 99, 628–66.

Coval, J. D., J. W. Jurek, and E. Stafford, 2009b, The economics of structured finance, *Journal of Economic Perspectives* 23, 3–25.

Decamps, J.-P., T. Mariotti, J.-C. Rochet, and S. Villeneuve, 2011, Free cash flow, issuance costs, and stock prices, *Journal of Finance* 66, 1501–1544.

Diamond, D. W., 1984, Financial intermediation and delegated monitoring, *Review of Economic Studies* 51, 393–414.

Diamond, D. W., and P. H. Dybvig, 1983, Bank runs, deposit insurance, and liquidity, *Journal of Political Economy* 91, 401–19.

Diamond, D. W., and R. G. Rajan, 2009, The credit crisis: Conjectures about causes and remedies, *American Economic Review* 99, 606–610.

Drechsler, I., A. Savov, and P. Schnabl, 2017, The deposits channel of monetary policy, *The Quarterly Journal of Economics* 132, 1819–1876.

Duffie, D., 2018, Prone to fail: The pre-crisis financial system, Working paper, Stanford GSB.

Farhi, E., and J. Tirole, 2017, Shadow banking and the four pillars of traditional financial intermediation, Technical report, National Bureau of Economic Research.

Fitch, 2007, Inside the ratings, Information brochure, Fitch.

Garleanu, N., and L. H. Pedersen, 2011, Margin-based asset pricing and deviations from the law of one price, *Review of Financial Studies* 24, 1980–2022.

Gertler, M., and S. Gilchrist, 1994, Monetary policy, business cycles, and the behavior of small manufacturing firms, *The Quarterly Journal of Economics* 109, 309–340.

Giannetti, M., and A. Simonov, 2013, On the real effects of bank bailouts: Micro evidence from japan, *American Economic Journal: Macroeconomics* 5, 135–67.

Gorton, G., and G. Pennacchi, 1990, Financial Intermediaries and Liquidity Creation, *Journal of Finance* 45, 49–71.

Greenwood, R., and S. G. Hanson, 2013, Issuer Quality and Corporate Bond Returns, *Review of Financial Studies* 26, 1483–1525.

Grossman, S. J., and O. D. Hart, 1986, The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration, *Journal of Political Economy* 94, 691–719.

Hannoun, H., 2013, Sovereign risk in bank regulation and supervision: Where do we stand?, Technical report, Conference presentation, Financial Stability Institute High-Level Meeting, Bank for International Settlements, Basel, Switzerland.

He, Z., and A. Krishnamurthy, 2013, Intermediary asset pricing, *American Economic Review* 103, 732–770.

Hellmann, T. F., K. C. Murdock, and J. E. Stiglitz, 2000, Liberalization, moral hazard in banking, and prudential regulation: Are capital requirements enough?, *American Economic Review* 90, 147–165.

Holmstrom, B., and J. Tirole, 1997, Financial intermediation, loanable funds, and the real sector, *The Quarterly Journal of Economics* 112, pp. 663–691.

Hoshi, T., and A. Kashyap, 1999, The japanese banking crisis: Where did it come from and how will it end?, Working Paper 7250, National Bureau of Economic Research.

Hoshi, T., and A. Kashyap, 2001, *Corporate Financing and Governance in Japan: The Road to the Future* (Mass.: MIT Press).

Iannotta, G., G. Pennacchi, and J. Santos, 2018, Ratings-based regulation and systematic risk incentives, *Review of Financial Studies* .

Innes, R. D., 1990, Limited liability and incentive contracting with ex-ante action choices, *Journal of Economic Theory* 52, 45–67.

Iyer, R., J.-L. Peydró, S. da Rocha-Lopes, and A. Schoar, 2014, Interbank liquidity crunch and the firm credit crunch: Evidence from the 2007–2009 crisis, *Review of Financial Studies* 27, 347–372.

Jiang, E., G. Matvos, T. Piskorski, and A. Seru, 2020, Banking without deposits: Evidence from shadow bank call reports, Working paper series, Stanford University.

Jiménez, G., S. Ongena, J.-L. Peydró, and J. Saurina Salas, 2012, Credit supply and monetary policy: Identifying the bank balance-sheet channel with loan applications, *American Economic Review* 102, 2301–26.

Jiménez, G., S. Ongena, J.-L. Peydró, and J. Saurina Salas, 2014, Hazardous times for monetary policy: What do twenty-three million bank loans say about the effects of monetary policy on credit risk-taking?, *Econometrica* 82, 463–505.

Kahn, C., and A. Winton, 2004, Moral hazard and optimal subsidiary structure for financial institutions, *Journal of Finance* 59, 2531–2575.

Kashyap, A. K., O. A. Lamont, and J. C. Stein, 1994, Credit conditions and the cyclical behavior of inventories, *The Quarterly Journal of Economics* 109, 565–592.

Kashyap, A. K., J. C. Stein, and D. W. Wilcox, 1993, Monetary policy and credit conditions: Evidence from the composition of external finance, *American Economic Review* 83, pp. 78–98.

Keeley, M. C., 1990, Deposit insurance, risk, and market power in banking, *American Economic Review* 80, pp. 1183–1200.

Khwaja, A. I., and A. Mian, 2008, Tracing the impact of bank liquidity shocks: Evidence from an emerging market, *American Economic Review* 98, 1413–42.

Kisgen, D. J., and P. E. Strahan, 2010, Do regulations based on credit ratings affect a firm's cost of capital?, *Review of Financial Studies* 23, 4324–4347.

Kisin, R., and A. Manela, 2016, The shadow cost of bank capital requirements, *Review of Financial Studies* 29, 1780–1820.

Koijen, R. S., and M. Yogo, forthcoming, A demand system approach to asset pricing, *Journal of Political Economy* .

Koijen, R. S. J., and M. Yogo, 2015, The Cost of Financial Frictions for Life Insurers, *American Economic Review* 105, 445–475.

Laux, C., and C. Leuz, 2010, Did Fair-Value Accounting Contribute to the Financial Crisis?, *Journal of Economic Perspectives* 24, 93–118.

Marcus, A. J., 1984, Deregulation and bank financial policy, *Journal of Banking and Finance* 8, 557 – 565.

Mas-Colell, A., M. D. Whinston, J. R. Green, et al., 1995, *Microeconomic theory*, volume 1 (Oxford university press New York).

Merton, R., 1977, An analytic derivation of the cost of deposit insurance and loan guarantees an application of modern option pricing theory, *Journal of Banking and Finance* 1, 3–11.

Muir, T., 2017, Financial crises and risk premia*, *The Quarterly Journal of Economics* 132, 765–809.

Opp, C. C., M. M. Opp, and M. Harris, 2013, Rating agencies in the face of regulation, *Journal of Financial Economics* 108, 46–61.

Peek, J., and E. S. Rosengren, 2000, Collateral damage: Effects of the japanese bank crisis on real activity in the united states, *American Economic Review* 90, 30–45.

Pennacchi, G., 2006, Deposit insurance, bank regulation, and financial system risks, *Journal of Monetary Economics* 53, 1–30.

Pennacchi, G. G., 1987, A reexamination of the over- (or under-) pricing of deposit insurance, *Journal of Money, Credit and Banking* 19, 340–360.

Qi, Y., 2019, Big broad banks: How does cross selling affect lending?, Unpublished dissertation, Stockholm School of Economics.

Rajan, R. G., and L. Zingales, 1995, What do we know about capital structure? some evidence from international data, *The Journal of Finance* 50, 1421–1460.

Rajan, R. G., and L. Zingales, 1998, Financial Dependence and Growth, *American Economic Review* 88, 559–586.

Rappoport, V., D. Paravisini, and P. Schnabl, 2014, Comparative advantage and specialization in bank lending, Technical report, London School of Economics and New York University.

Repullo, R., 2004, Capital requirements, market power, and risk-taking in banking, *Journal of financial Intermediation* 13, 156–182.

Repullo, R., and J. Suarez, 2004, Loan pricing under basel capital requirements, *Journal of Financial Intermediation* 13, 496–521.

Repullo, R., and J. Suarez, 2013, The procyclical effects of bank capital regulation, *Review of Financial Studies* 26, 452–490.

Rochet, J., 1992, Capital requirements and the behaviour of commercial banks, *European Economic Review* 36, 1137–1170.

Scharfstein, D., and A. Sunderam, 2016, Market power in mortgage lending and the transmission of monetary policy, *Unpublished working paper. Harvard University* .

SEB, 2018, Euro Area: Increased bank holdings of government debt, Technical report, SEB.

Stein, J., 2012, Monetary policy as financial-stability regulation, *Quarterly Journal of Economics* 127, 57–95.

Tirole, J., 2006, *Theory of Corporate Finance* (Princeton University Press, Princeton).

A. Proofs

A.1. Proof of Lemma 1

First, we show that if $NPV(q) < B(q)$, the borrower cannot raise financing under any contract. Assumption 1.1 implies that public financing requires high effort, i.e., $a = 1$. If the borrower exerts effort, the maximum value of the borrower's stake is given by $NPV(q)$, since the IR constraint and investor competition imply that investors' expected discounted payoff is equal to I, and $NPV(q)$ is equal to the difference between the present value of the firm's cashflows $\frac{\mathbb{E}[C_s(q,1)]}{1+r_F}$ and I. Second, as reflected by the IC constraint, the borrower's payoff under shirking is bounded from below by $B(q)$, due to limited liability. Hence, if $NPV(q) < B(q)$, it is impossible to jointly satisfy IC and IR.

We next show that whenever $NPV(q) \geq B(q)$, the borrower can raise financing with a debt contract that gives all surplus to the borrower, which also proves the optimality of debt. Set $CF_s = FV$ for all s. Then IR implies that $\frac{FV}{1+r_F} \geq I$. Moreover, using Assumption 1.2, we obtain that $\mathbb{E}[\max\{C_s(q,0) - FV, 0\}] = 0$ and the right hand side of IC achieves the lower bound $B(q)$ under any debt contract that satisfies IR. Since investors are competitive, the face value of debt is set such that IR binds, so that the borrower's payoff is $NPV(q)$. We have thus proven that whenever $NPV(q) \geq B(q)$, there exists a debt contract that satisfies IR and allows the borrower to extract the entire NPV.

Unlike in Innes (1990) the optimality of debt is implied by Assumption 1.2 rather than the joint assumption of the monotone likelihood ratio property (MLRP) and the monotonicity constraint of investors' payoff in firm cash flows. There are cash flow distributions that satisfy Assumption 1.2, but not MLRP, and vice versa.

A.2. Proof of Lemma 2

We analyze the individually optimal portfolio choice of a bank that faces a perfectly elastic supply of securities and takes as given the associated state-dependent returns $r^s(q, \rho)$. The bank's inner (ROE) maximization problem (14) is

$$\max_{e,\mathbf{x}} r_E(\mathbf{x},e) - r_F \text{ s.t. } e \geq e_{\min}(\mathbf{x}), \mathbf{x} \geq \mathbf{0}. \tag{26}$$

where

$$r_E(\mathbf{x},e) - r_F = \frac{1}{e}\mathbb{E}\left[\max\left\{r_A^s(\mathbf{x}) - r_F, -(1+r_F)\,e\right\}\right].$$

We note that $r_E(\mathbf{x},e) - r_F \geq 0$ if the bank chooses a strictly positive investment in a loan portfolio, $A > 0$. Otherwise, it would prefer to invest in cash or pay out dividends $E_O = -E_I$. We thus only consider the relevant case where a weakly positive excess return is attainable.

Leverage. Taking the partial derivative of $r_E(\mathbf{x},e)$ w.r.t. e yields

$$\frac{\partial r_E(\mathbf{x},e)}{\partial e} = -\frac{1}{e^2}\mathbb{E}\left[\max\left\{r_A^s(\mathbf{x}) - r_F, -(1+r_F)\,e\right\}\right] - \frac{1}{e}\Pr\left[\frac{r_A^s(\mathbf{x}) - r_F}{1+r_F} < -e\right].$$

Note that if $r_E(e,\mathbf{x}) > r_F$ for some $(e,\mathbf{x})$ then it must be the case that

$$\mathbb{E}\left[\max\left\{r_A^s(\mathbf{x}) - r_F, -(1+r_F)\,e\right\}\right] > 0.$$

It follows that $\frac{\partial r_E(\mathbf{x},e)}{\partial e} < 0$ if $r_E(\mathbf{x},e) > r_F$. Further, if $r_E(\mathbf{x},e) = r_F$ then $\frac{\partial r_E(\mathbf{x},e)}{\partial e} < 0$ as long as there is one state s with positive probability, where the bank defaults, that is, $\Pr[\frac{r_A^s(\mathbf{x})-r_F}{1+r_F} < -e] > 0$.

Thus, for any choice $(\mathbf{x},e)$ that yields $r_E(\mathbf{x},e) > r_F$ it is optimal to decrease e at the margin, unless the constraint $e \geq e_{\min}$ is already binding. Since decreasing e increases $r_E(\mathbf{x},e)$, the condition $r_E(\mathbf{x},e) > r_F$ remains satisfied after any decrease in e. Thus, for any $(\bar{\mathbf{x}},\bar{e})$ such that $r_E(\bar{\mathbf{x}},\bar{e}) > r_F$ it is the case that $\arg\max_e r_E(\bar{\mathbf{x}},\bar{e}) = e_{\min}$.

Further, for any choice $(\mathbf{x},e)$ that yields $r_E(\mathbf{x},e) = r_F$ and $\Pr[\frac{r_A^s(\mathbf{x})-r_F}{1+r_F} < -e] > 0$, marginally decreasing e also increases r_E (provided such a decrease is feasible, that is, the constraint $e \geq e_{\min}$ is not already binding). Since marginally decreasing e increases $r_E(\mathbf{x},e)$ (maintaining the condition that $r_E(\mathbf{x},e) \geq r_F$) and weakly enlarges the set of default states (maintaining $\Pr[\frac{r_A^s(\mathbf{x})-r_F}{1+r_F} < -e] > 0$), it is optimal to decrease e until the constraint $e \geq e_{\min}$ is binding. Formally, for any $(\bar{\mathbf{x}},\bar{e})$ such that $r_E(\bar{\mathbf{x}},\bar{e}) = r_F$ it is the case that $\arg\max_e r_E(\bar{\mathbf{x}},\bar{e}) = e_{\min}$ if $\Pr[\frac{r_A^s(\mathbf{x})-r_F}{1+r_F} < -e] > 0$.

This concludes the proof of the two statements about optimum leverage.

Portfolio choice. The analysis in the previous paragraph implies that it is optimal for banks to choose $e = e_{\min}$ as long as there exists a portfolio $\mathbf{x}$ such that $r_E(\mathbf{x}{,}e) > r_F$, or $r_E(\mathbf{x}{,}e) = r_F$ and $\Pr[\frac{r_A^s(\mathbf{x})-r_F}{1+r_F} < -e] > 0$. The following Lemma will be useful for characterizing the banks' portfolio choice.

Lemma 5 *For all (q,ρ) with $x^*(q,\rho) > 0$, we obtain*

$$\frac{\mathbb{E}\left[r^s(q,\rho) - r_F \middle| s : \frac{r_A^s(\mathbf{x}^*)-r_F}{1+r_F} > -e_{\min}(\mathbf{x}^*)\right]}{\underline{e}(\rho)} = \frac{\nu}{\Pr\left[s : \frac{r_A^s-r_F}{1+r_F} > -e_{\min}\right]} = k > 0. \quad (27)$$

where υ is the Lagrange multiplier on the constraint $\sum_{q,\rho} w(q,\rho) = 1$ and k is some positive constant.

Proof: Presume that such a portfolio $\mathbf{x}$ exists and that banks (optimally) choose $e = e_{\min}$. Then we can re-write the expected excess return on a bank's book equity as follows:

$$\begin{aligned} r_E(\mathbf{x}{,}e_{\min}) - r_F &= \mathbb{E}\left[\max\left\{\frac{r_A^s(\mathbf{x}) - r_F}{e_{\min}(\mathbf{x})}, -(1+r_F)\right\}\right] \\ &= \mathbb{E}\left[\max\left\{\frac{\sum_{q,\rho} x(q,\rho)\left[r^s(q,\rho) - r_F\right]}{\sum_{q,\rho} x(q,\rho)\underline{e}(\rho)}, -(1+r_F)\right\}\right] \\ &= \mathbb{E}\left[\max\left\{\sum_{q,\rho} \frac{r^s(q,\rho) - r_F}{\underline{e}(\rho)} \frac{x(q,\rho)\underline{e}(\rho)}{\sum_{\tilde{q},\tilde{\rho}} x(\tilde{q},\tilde{\rho})\underline{e}(\tilde{\rho})}, -(1+r_F)\right\}\right]. \quad (28) \end{aligned}$$

Defining $w(q,\rho) = \frac{x(q,\rho)\underline{e}(\rho)}{\sum_{\tilde{q},\tilde{\rho}} x(\tilde{q},\tilde{\rho})\underline{e}(\tilde{\rho})} \in [0,1]$ for all (q,ρ) as the new choice variables we obtain:

$$r_E(\mathbf{w}) - r_F = \mathbb{E}\left[\max\left\{\sum_{q,\rho} w(q,\rho)\frac{r^s(q,\rho) - r_F}{\underline{e}(\rho)}, -(1+r_F)\right\}\right].$$

Maximizing subject to the constraint that $\sum_{q,\rho} w(q,\rho) = 1$ and $w(q,\rho) \geq 0$ (short-sales constraint), we obtain for all (q,ρ) with $w^*(q,\rho) > 0$ the following condition at the optimum:

$$\frac{\partial r_E(\mathbf{w}) - r_F}{\partial w(q,\rho)} = \nu. \quad (29)$$

Further, we can write:

$$\frac{\partial r_E(\mathbf{w}) - r_F}{\partial w(q,\rho)} = \mathbb{E}\left[\frac{r^s(q,\rho) - r_F}{\underline{e}(\rho)} \middle| s : \frac{r_A^s - r_F}{1 + r_F} > -e_{\min}\right] \cdot \Pr\left[s : \frac{r_A^s - r_F}{1 + r_F} > -e_{\min}\right]. \tag{30}$$

Combining (29) and (30), we obtain (27) if $w^*(q,\rho) > 0$ (and, hence, $x^*(q,\rho) > 0$). ■

Correlated tail risks. First, note that we established in Lemma 5 that for any optimal choice $(\mathbf{x}^*, e^*)$ the expected excess asset return conditional on bank survival scaled by $\underline{e}(\rho)$ is identical across issuer types (q,ρ) with $x^*(q,\rho) > 0$. Suppose there is a type $(\tilde{q},\tilde{\rho})$ with $x^*(\tilde{q},\tilde{\rho}) > 0$ in the optimal portfolio that yields $\frac{r^s(\tilde{q},\tilde{\rho}) - r_F}{1+r_F} > -\underline{e}(\tilde{\rho})$ in some state s where the bank defaults, that is, where $\sum_{q,\rho} \frac{x^*(q,\rho) r^s(q,\rho) - r_F}{1+r_F} < -e_{\min}$. Then the bank could obtain a higher expected return on equity $r_E > r_E(\mathbf{x}^*, e^*)$ by investing only in this asset $(\tilde{q},\tilde{\rho})$, as it not only yields the same expected levered return across previous survival states (under the previous policy $(\mathbf{x}^*, e^*)$) but also allows the bank to survive in at least one additional state s.

Conversely, suppose $\mathbf{x}^*$ is an optimal portfolio and there is an asset of type $(\tilde{q},\tilde{\rho})$ in the optimal portfolio with a strictly positive weight ($x^*(\tilde{q},\tilde{\rho}) > 0$) that yields $r^{\tilde{s}}(\tilde{q},\tilde{\rho}) \leq -\underline{e}(\tilde{\rho})$ in some state $\tilde{s}$ where the bank survives and has strictly positive equity value, that is, where $\sum_{q,\rho} w^*(q,\rho) \frac{r^s(q,\rho) - r_F}{\underline{e}(\rho)} > -(1 + r_F)$. Then it must be the case that in this survival state $\tilde{s}$ other assets in the portfolio yield $\frac{r^s(q,\rho) - r_F}{\underline{e}(\rho)} > -(1 + r_F)$, otherwise the bank would default in that state. For notational simplicity define the set of states where the bank survives under policy $(\mathbf{x}^*, e_{\min}(\mathbf{x}^*))$ as $\Sigma_S(\mathbf{x}^*, e_{\min}(\mathbf{x}^*))$. We showed in Lemma 5 that

$$\mathbb{E}\left[\frac{r^s(q,\rho) - r_F}{\underline{e}(\rho)} \middle| \Sigma_S\right] = k,$$

for all (q,ρ) with $x^*(q,\rho) > 0$. However, since asset $(\tilde{q},\tilde{\rho})$ performs worse than other assets in the portfolio in state $\tilde{s}$, that is, $\frac{r^{\tilde{s}}(\tilde{q},\tilde{\rho}) - r_F}{\underline{e}(\tilde{\rho})} < -(1 + r_F) \leq \frac{r^{\tilde{s}}(q,\rho) - r_F}{\underline{e}(\rho)}$ it must outperform, relative to the other assets in the portfolio in expectation in the other survival states, to ensure that equation (27) can hold, that is:

$$\mathbb{E}\left[\frac{r^s(\tilde{q},\tilde{\rho}) - r_F}{\underline{e}(\tilde{\rho})} \middle| \Sigma_S \setminus \tilde{s}\right] > \mathbb{E}\left[\frac{r^s(q,\rho) - r_F}{\underline{e}(\rho)} \middle| \Sigma_S \setminus \tilde{s}\right] \text{ for all } (q,\rho) \neq (\tilde{q},\tilde{\rho}) \text{ with } x^*(q,\rho) > 0.$$

If we set $w(\tilde{q},\tilde{\rho}) = 1$ and $w(q,\rho) = 0$ for all $(q,\rho) \neq (\tilde{q},\tilde{\rho})$ we obtain the following expected excess return on equity conditional on the states Σ_S:

$$\begin{aligned}
&(1-\Pr[\tilde{s}|\Sigma_S])\cdot\mathbb{E}\left[\left.\frac{r^s(\tilde{q},\tilde{\rho})-r_F}{\underline{e}(\tilde{\rho})}\right|\Sigma_S\setminus\tilde{s}\right]+\Pr[\tilde{s}|\Sigma_S]\cdot(-1-r_F)\\
>&(1-\Pr[\tilde{s}|\Sigma_S])\mathbb{E}\left[\left.\frac{r^s(\tilde{q},\tilde{\rho})-r_F}{\underline{e}(\tilde{\rho})}\right|\Sigma_S\setminus\tilde{s}\right]+\Pr[\tilde{s}|\Sigma_S]\frac{r^{\tilde{s}}(\tilde{q},\tilde{\rho})-r_F}{\underline{e}(\tilde{\rho})}\\
=&k,
\end{aligned} \tag{31}$$

that is, we obtain a conditional expected return that is greater than the one obtained from portfolio $\mathbf{x}^*$. Further, in failure states Σ_F, this new portfolio cannot yield equity holders lower returns than the previous portfolio $\mathbf{x}^*$, since equity holders are protected by limited liability. This implies that setting $x(\tilde{q},\tilde{\rho})=1$ and $x(q,\rho)=0$ for all $(q,\rho)\neq(\tilde{q},\tilde{\rho})$ increases r_E, contradicting the supposition that $\mathbf{x}^*$ was an optimal portfolio.

Thus, if $\mathbf{x}^*$ is an optimal portfolio then any asset (q,ρ) in this optimal portfolio with a strictly positive weight ($x^*(q,\rho)>0$) must yield $\frac{r^s(q,\rho)-r_F}{1+r_F}>-\underline{e}(\rho)$ in all states s where the bank survives and has strictly positive equity value.

A.3. Proof of Lemma 3

Recall that the optimal bank inside equity value can be written as follows:

$$E_{M,I}=E_I+\max_{E_O}\left[\frac{(E_I+E_O)(\max_{e,\mathbf{x}}[r_E(\mathbf{x},e)]-r_F)}{1+r_F}-(c(E_O)-E_O)\right].$$

Let $(\mathbf{x}^*,e^*)$ denote the optimal solution to the inner (ROI) maximization problem. It follows that if $(c'(0)-1)\geq\frac{r_E(\mathbf{x}^*,e^*)-r_F}{1+r_F}$ the bank optimally sets $E_O=0$ (note that c is weakly convex). Further, at any E_O where $(c'(E_O)-1)<\frac{r_E(\mathbf{x}^*,e^*)-r_F}{1+r_F}$ the bank can strictly increase its objective function at the margin by increasing E_O.

A.4. Proof of Lemma 4

The reservation price of an issuer is defined as the date-0 value added to bank equity holders per unit of allocated bank equity if the issuer is financed at her outside option. The derivation of the reservation price builds on results in Lemmas 1 and 2. First, if an issuer demands a loan to finance an investment of size I, optimal financing decisions by the banker (by Lemma 2) imply that the issuer "effectively" demands $I\underline{e}(\rho)$ units of bank equity. Bankers obtain the remaining funds of $I(1-\underline{e}(\rho))$ via (subsidized) deposits. Since the government transfers the difference between the promised repayment to depositors $I(1-e(\rho))(1+r_F)$ and the cash flows produced by banks assets (the cash flows generated by the borrower, $C_s(q,1)$) in bank default states, the present value of government transfers ultimately accruing to bank equity holders is

$$PUT(q,\rho) \equiv \frac{\mathbb{E}\left[\max\left\{I(1-e(\rho))(1+r_F)-C_s(q,1),0\right\}\right]}{1+r_F} \geq 0. \tag{32}$$

The value of $PUT(q,\rho)$ uses the optimality of portfolios with correlated tail risk (by Lemma 2) and that bankers hold senior loans with promised yields of $y(q,\rho) \geq r_F$.

Conditional on financing an issuer, the total *private* surplus shared between the bank equity holders and the issuer is, thus, given by $NPV(q)+PUT(q,\rho)$. Due to the borrower's outside option of unmonitored finance (see Lemma 1) the *maximum* value added that bankers can reap is given by

$$\Pi(q,\rho) = NPV(q) + PUT(q,\rho) - NPV(q)\,\mathbb{1}_{\{NPV(q)\geq B(q)\}}. \tag{33}$$

Scaling (33) by $I\underline{e}(\rho)$ and adding 1 yields the effective price that a banker receives per unit of bank equity if the borrower is financed at his outside option, i.e., the issuer's reservation price in (17).

A.5. Proof of Proposition 1

The result follows from standard general equilibrium analysis, see e.g., Mas-Colell et al. (1995).

A.6. Proof of Proposition 2

As is standard in general equilibrium theory, all issuer types (q,ρ) with a reservation price $p^r(q,\rho)$ above the equilibrium price p^* get financed. To obtain ξ note that after financing all issuers with $p^r(q,\rho) > p^*$, an amount of $E^* - \sum_{(q,\rho):p^r(q,\rho)>p^*} I \cdot \underline{e}(\rho)\cdot m(q,\rho)$ is left to fund issuers with $p^r(q,\rho) = p^*$. The total demanded capital by these issuers is $\sum_{(q,\rho):p^r(q,\rho)=p^*} I \cdot$

$\underline{e}(\rho) \cdot m(q, \rho)$. Hence, we obtain that

$$\xi = \frac{E^* - \sum_{(q,\rho):p^r(q,\rho)>p^*} I \cdot \underline{e}(\rho) \cdot m(q, \rho)}{\sum_{(q,\rho):p^r(q,\rho)=p^*} I \cdot \underline{e}(\rho) \cdot m(q, \rho)}. \tag{34}$$

To obtain the expected return on debt of bank finance borrowers we use the fact that all loans must yield the same ROE to bankers (or equivalently, the same price) if financed in optimal portfolios. That is,

$$\mathbb{E}\left[\max\left\{\frac{r^s(q, \rho) - r_F}{\underline{e}\,(\rho)}, -\,(1 + r_F)\right\}\right] = r_E^* - r_F, \tag{35}$$

where $r_E^* - r_F = p^*\,(1 + r_F)$. Multiplying (35) by $\underline{e}\,(\rho)$ and using basic algebra gives us:

$$\mathbb{E}\left[r^s(q, \rho)\right] = r_F + \underline{e}\,(\rho)\left[r_E^* - r_F\right] - \mathbb{E}\left[\max\left\{(1 - \underline{e}\,(\rho))\,(1 + r_F) - \left[1 + r^s(q, \rho)\right], 0\right\}\right] \tag{36}$$

Since $y(q, \rho) \geq r_F$, we obtain that $1 + r^s(q, \rho) = \frac{C_s(q,1)}{I}$ whenever $\frac{r^s(q,\rho) - r_F}{\underline{e}(\rho)} < -\,(1 + r_F)$. Thus, we get:

$$\mathbb{E}\left[r^s(q, \rho)\right] = r_F + \underline{e}\,(\rho)\left[r_E^* - r_F\right] - \frac{1 + r_F}{I}\frac{\mathbb{E}\left[\max\left\{I\,(1 - \underline{e}\,(\rho))\,(1 + r_F) - C_s\,(q, 1)\,, 0\right\}\right]}{1 + r_F}. \tag{37}$$

Using the definition of (18) we thus obtain (23).

B. Discussion of Modeling Assumptions

In this section, we discuss the robustness of the presented results with respect to various modeling assumptions. We highlight that key principles uncovered from our approach of considering a micro-founded aggregate demand function for bank capital continue to apply when various assumptions of our baseline model are relaxed. In this context, we will refer to the following broad definition of a borrower's reservation price for bank capital:

$$p^r = 1 + \frac{\text{Incremental private surplus from funding borrower with bank loan}}{\text{Bank capital needed to fund borrower}}. \tag{38}$$

In discussing implications of alternative modeling assumptions, we will repeatedly revisit this general representation of borrowers' reservation prices. In particular, we will evaluate which elements of equation (38) would be affected by additional economic channels not explicitly featured in our baseline model.

Market power. The proposed environment features the standard assumption that banks act competitively (as, e.g., in Holmstrom and Tirole, 1997). Yet, in principle, banks may have market power in the loan market (see evidence in Scharfstein and Sunderam (2016)) and/or in the deposit market (see evidence in Drechsler et al. (2017)). If banks had market power in the loan market, they would be able to extract a greater fraction of the surplus created when funding a borrower, that is, banks would receive higher prices per unit of bank capital. However, borrowers' *reservation prices* and the associated demand for bank capital are unaffected by this type of market power. As a result, key insights of our analysis regarding the demand curve would still apply if banks had market power in their interactions with borrowers.

On the other hand, if banks had market power in the deposit market, any investments yielding expected returns above the deposit rate (including storage investments) would generate additional private surplus. This source of surplus would imply an additional channel causing a wedge between the private ranking of borrowers within the demand curve based on reservation prices and the social ranking based on total surplus. In particular, investments in securities that are associated with higher risk weights could be financed less with "cheap" deposits, making these investments less attractive, ceteris paribus. While in the presented model, higher risk weights already cause borrowers to rank lower in the demand curve, this additional channel would add to the existing effect emerging from the PUT component affecting reservation prices. In particular, if safe storage investments (e.g., government bonds) were associated with very low risk weights, then banks would have a larger incentive to invest in these types of securities, shedding light on banks incentives to hold "safe" assets.

Ex-ante differences across banks. Our model reveals that even ex ante identical banks optimally choose heterogeneous portfolio strategies (see Corollary 1). If subgroups of banks additionally differed ex ante in terms of characteristics such as the probability of receiving government bailouts, legacy asset holdings, or monitoring technologies, these sources of heterogeneity would naturally lead to clientele effects. These clientele effects would lead to multiple bank capital demand curves, one for each subgroup of banks. For example, ceteris paribus, banks that are more likely to receive government bailouts would generate higher reservation prices with risky borrowers, as the PUT component of the reservation price would be higher. Moreover, banks could have heterogeneous monitoring technologies as represented by differing abilities to reduce moral hazard rents or differing monitoring costs. In this case, banks whose monitoring technologies are less efficient would also have greater risk taking incentives. As the monitoring-dependent surplus of these banks would be lower, the PUT component would be a relatively more important source of the private surplus shaping reservation prices. Similarly, if banks had different types of legacy assets, they would create more private surplus with those types of new borrowers that exhibit correlated tail risks with the existing assets. For example, as Greek banks are generically more exposed to Greek risk factors, this logic predicts that these banks have a comparative advantage specifically in holding Greek sovereign debt, rather than just any risky debt.[44]

Endogenous capital requirements and deposit insurance premia. The proposed modeling environment allows capturing many details of regulatory frameworks used in practice by putting effectively no restrictions on specifications for overall capital requirements, risk classifications, and risk weights. This framework can facilitate analyses of how regulators should optimally choose parameters of the regulatory environment when facing the plausible limitation that regulations can condition only on a given set of noisy but verifiable security risk classifications (akin to the coarse set of verifiable signals in the incomplete contracts literature following Grossman and Hart, 1986). These contractible risk classifications (e.g., credit ratings) generally pool multiple types of borrowers, and thus provide noisy and/or biased risk evaluations (for example, two borrower types (q, ρ) and (q', ρ) are pooled under the common regulatory risk classification ρ). Due to this type of pooling, setting risk weights for specific risk classifications then generically involves trade-offs. In particular, regulators typically face the dilemma that high risk weights on the one hand reduce the funding of surplus-destroying risky borrowers of a given risk-classification, but on the other hand they can also cause rationing of credit to bank-dependent surplus-generating borrowers with the same risk classification. These trade-offs emerging from imprecise risk classifications could also not be alleviated by additional regulatory tools used in practice, such as *deposit insurance premia*. As deposit insurance premia also have to rely on the same regulatory risk classifications of securities, they would operate similarly to risk weights in affecting the reservation prices of *all* borrowers pooled under a given risk classification ρ. In particular, deposit insurance premia would lower the incremental private surplus from bank lending for all borrowers of

[44] It is useful to relate this prediction regarding the effects of legacy assets to an interesting partial equilibrium analysis of Bahaj and Malherbe (2018). The authors show that a bank that has a risky legacy asset may not be willing to add a safe (otherwise good) lending opportunity to its portfolio, since doing so would reduce the overall put value for the bank. In our general equilibrium setting, this new safe asset would typically be purchased by a different bank with safer legacy assets. As a result, adding safe assets does not necessarily reduce the overall put value of the banking sector once a cross-section of banks is considered.

a given classification. Finally, analyses of this type could flexibly specify welfare functions incorporating additional allocative effects going beyond the surplus generated by borrowers (such as the costs of raising tax payer funds for bailouts).

Multi-period settings. To maintain its focus on compositional effects, our framework considers a two-period setup. The main economic principles developed in this paper would, however, extend to dynamic environments. In particular, as in our current setting, the equilibrium price of bank capital can generally be defined as the derivative of a bank's value function with respect to the level of its current capital. In multi-period settings, banks still effectively rank potential loans according to the value that these loans provide to equity holders, per unit of scarce capital they consume, a metric we have defined as a loan's reservation price. In a dynamic environment, these reservation prices would account for the continuation value ("franchise value") that banks forgo when defaulting. For example, in times with high franchise values, banks would effectively have more skin in the game, thus lowering the PUT components of loans' reservation prices and reducing reaching-for-yield incentives for banks. More generally, any time variation in economic prospects (future cash flows, bank dependence, regulations, etc.) would then make the magnitude of the PUT component and associated distortions time-dependent. Another interesting feature of dynamic environments is the notion that banks can retain profits to gradually build up equity capital, a channel that can help reduce the scarcity of bank capital. While these types of dynamics are undoubtedly relevant in practice, the main economic principles highlighted in this paper still apply in their presence, providing clear conceptual guidance on the determinants of the composition of credit.