Rewarding Prudence: Risk Taking, Pecuniary Externalities and Optimal Bank Regulation

Michael Kogler

May 2015 Discussion Paper no. 2015-12
Rewarding Prudence:
Risk Taking, Pecuniary Externalities and Optimal Bank Regulation

Michael Kogler

Author's address: Michael Kogler
FGN-HSG
Varnbuelstrasse 19
CH-9000 St.Gallen
Phone +41 71 224 2156
Fax +41 71 224 2874
Email michael.kogler@student.unisg.ch
Website www.fgn.unisg.ch

1 An earlier version of this paper circulated under the title 'Moral Hazard, Bank Risk Taking, and Prudential Regulation: A Welfare Analysis'. I thank Christian Keuschnigg, Jochen Mankart, seminar participants at University of St.Gallen, and participants at the 7th RGS Doctoral Conference in Economics in Dortmund and the Warsaw International Economic Meeting (WIEM) for helpful discussions and comments. Financial support of the Swiss National Science Foundation (Project no. 100018/1_146685/1) is gratefully acknowledged.
Abstract

This paper provides a new rationale for macroprudential regulation and studies its optimal design, implementation, and distributional consequences. In a partial equilibrium model where bank risk taking is subject to moral hazard, we show that although private contracting can solve the bank's agency (i.e., risk shifting) problem, the market outcome is constrained-inefficient: The combination of moral hazard and competition for deposits that are not supplied elastically leads to a pecuniary externality as raising deposits ultimately increases the deposit rate and exacerbates risk shifting of all other banks. As a result, banks are too large, have too much leverage, and take excessive risk. This generic inefficiency provides a strong rationale for regulation even in the absence of classical frictions such as social cost of bank failure or incorrectly priced deposit insurance. The pecuniary externality can be internalized by standard regulatory tools such as capital requirements or by issuing a specific number of banking licenses. Related to the idea of financial restraint, optimal regulation creates rent opportunities to reward prudent banks. This also leads to redistribution from depositors to banks and firms with access to the capital market such that optimal regulation is not a Pareto improvement.

Keywords

Bank Regulation, Bank Competition, Risk Taking, Moral Hazard.

JEL Classification

D60, D62, G21, G28.
1 Introduction

The recent financial crisis provides many insights for economists and policymakers. This paper focuses on two of them: First, the importance of interactions and interconnections between banks gave rise to a new regulatory paradigm, macroprudential regulation. Previous frameworks like Basel II were strongly influenced by the predominant microprudential approach and proved to be inadequate given their primary focus on the stability of a single financial institution. Although still somewhat vague, macroprudential regulation has become an important benchmark when shaping a new regulatory framework (e.g., Basel III). In particular, this approach takes a more systemic perspective and addresses problems like procyclicality, contagion, and pecuniary externalities. Second, the crisis forcefully demonstrated the importance of asymmetric information in the banking industry where adverse selection and moral hazard are widespread phenomena. This problem has been well recognized in the literature, which shows how private contracting can, in principle, resolve it: Bank owners need to invest sufficient equity such that the incentives for prudent lending and diligent monitoring are preserved. However, this private solution may fail to internalize externalities, which, in turn, provides a rationale for regulation.

This paper combines the two insights and studies a typical form of interaction between banks - competition for deposits - in the context of a standard agency problem - risk shifting - that can be addressed by private contracting. The normative analysis identifies a novel pecuniary externality: Banks fail to internalize that raising deposits exacerbates risk shifting of all other banks through its effect on the deposit rate. Therefore, banks are too large and have too much leverage, which erodes profits and results in excessive risk taking. In particular, this inefficiency is generic in the sense that it prevails even in the absence of typical frictions such as social cost of bank failure and incorrectly priced deposit insurance or government guarantees and thus provides a strong rationale for intervention. This aspect becomes particularly relevant in the context of the European Banking Union where attempts to internalize such cost through self-insurance of banks\(^1\) are made. Moreover, the paper derives the design and describes the key mechanisms of optimal regulation as well as its distributional consequences: The main idea is to reward prudent banks which relates to the paradigm of financial restraint emphasized by Hellmann, Murdock, and Stiglitz (1997, 2000). The regulator creates rent opportunities for banks by limiting leverage and deposit demand, which keeps the funding cost artificially low and raises profit margins. The prospect of earning a rent, in turn, increases the private cost of risk taking and strengthens the incentive for prudent lending. We find that different instruments including capital requirements or imperfect competition can be applied to create such rent opportunities eventually implementing a constrained-efficient allocation. The distributional implications are largely determined by this *modus operandi* which yields new insights about how the cost of bank regulation are shared.

For that purpose, this paper develops a static equilibrium model of banking which entails features of moral hazard and risk taking models as well as an endogenous resource allocation determined by bank lending and capital structure. The latter allows for a full-fledged model of depositors instead of the elastic deposit supply common in many moral hazard models in banking.\(^2\) This modification is crucial as the deposit rate then also reflects the scarcity of capital, which depends on size and capital structure of banks, and thus creates the possibility for a welfare-reducing pecuniary externality. In addition, the explicit model of the deposit side also allows for a comprehensive study of distributional consequences.

The remainder of this paper is organized as follows: Section 2 reviews the related literature. Subsequently, section 3 outlines the model; and section 4 provides a first-best benchmark, introduces moral hazard, and analyzes both the market equilibrium and the allocation chosen by a social plan-

---

1. Banks contribute to the Single Resolution Funds and to *ex ante* financed deposit insurance.
2. For example, Suarez and Sussman (1997) or Repullo (2013).
 ner subject to the same informational constraints. Eventually, section 5 discusses three extensions of the model, and section 6 concludes.

## 2 Literature

This paper builds on three strands of the literature: (i) pecuniary externalities and macroprudential regulation, (ii) moral hazard and risk shifting, and (iii) bank risk taking.

First of all, pecuniary externalities (i.e., price externalities) are the source of second-best inefficiencies in sectors characterized by imperfect information and incomplete markets as shown by Greenwald and Stiglitz (1986). These frictions are particularly relevant in banking and pecuniary externalities thus provide a rationale for macroprudential regulation. The problem is essentially that atomistic agents fail to internalize price reactions which affect incentives, financial or collateral constraints. Gersbach and Rochet (2012) show that the bank’s failure to internalize the effect of reallocating capital on asset prices results in excessive downsizing during after an adverse shock. Using a double moral hazard model that allows for collusion between banker and borrower, Tressel and Verdier (2014) identify a pecuniary externality associated with the return on bank capital that provides a rationale for macroprudential regulation. Bianchi (2011) provides a quantitative model that implies economically significant welfare gains from correcting pecuniary externalities. Other common mechanisms that involve pecuniary externalities in finance include amplification effects due to financial constraints, for example, Suarez and Sussman (1997), Caballero and Krishnamurthy (2003) and Lorenzoni (2008), and fire sales (i.e., the sale of assets at a dislocated price because of liquidity needs or regulatory requirements), for example, Shleifer and Vishny (1992), Diamond and Rajan (2011) or Korinek (2011).

Second, moral hazard is a well-known phenomenon in banking and corporate finance. In contrast to Modigliani and Miller (1958), the capital structure becomes key as it influences choices like risk and effort. A classical implication of moral hazard is risk shifting (or asset substitution); seminal contributions include Jensen and Meckling (1976) and Stiglitz and Weiss (1981). Essentially, firms favor riskier activities if they are funded by a large share of debt. Intuitively, owners protected by limited liability can shift the downside risk onto debtholders and increase their value by engaging in excessively risky activities. This holds a fortiori for banks, which are highly levered; it might be exacerbated by guarantees and deposit insurance. The problem is usually addressed by ensuring that a bank has sufficient equity. Applications of the risk shifting problem include, for example, Suarez and Sussman (1997) who show that it is the source of endogenous cycles even in the absence of stochastic shocks, Boot and Ratnovski (2012) who study risk shifting associated with the trading activities of a bank that engages both in trading and traditional banking, Repullo (2013) who characterizes how the capital structure mitigates risk shifting focusing on its cyclical properties, or Acharya, Mehra, and Thakor (2013) who examine the optimal capital structure of banks when there is both risk shifting and managerial rent seeking. The empirical relevance of risk shifting by banks is documented in studies, for example, by the Office of the Comptroller of the Currency (1988), Esty (1996) or Laudier, Thesmar, and Sraer (2011).

Risk shifting is at the core of a closely related literature that focuses how on risk taking depends on bank competition and capital regulation: It emphasizes the role of the bank’s charter or franchise value (i.e., the present value of its future profits) as a disciplining device similar to equity since it discourages risk taking by making bank failure privately costly. Since intense competition lowers future profits, the so-called ‘charter value hypothesis’ emphasized, for example, by Keeley (1990), Matutes and Vives (2000), and Allen and Gale (2000, 2004), postulates a negative relationship

---

3Characterized by a high payoff if successful but a low success probability

4Merton (1977) shows that bank shareholders’ value of deposit insurance is equivalent to a put option and increases in risk (volatility).
between bank competition and financial stability. However, this view has recently been challenged: Boyd and De Nicoló (2005) show that this positive relation is reversed as soon as risk is determined by borrowers who benefit from intense competition among banks. Martinez-Miera and Repullo (2010), however, argue that lower bank profits reduce the buffer to absorb loan losses giving rise to a u-shaped relationship between competition and financial risk. The evidence is mixed: Keeley (1990) and Demirguc-Kunt and Detragiache (1997, 1998) find that increased bank competition following deregulation reduces charter values resulting in higher bank risk and a higher frequency of banking crises whereas Jayaratne and Strahan (1998) obtain a negative effect of deregulation on loan losses. Beck, Demirgüç-Kunt, and Levine (2006) estimate a positive effect of bank concentration on the probability of a systemic crisis, and Beck, De Jonghe, and Schepens (2013) find considerable cross-country heterogeneity in the relationship of competition and risk taking. A similar argument applies to capital regulation: In general, one needs to distinguish between inside and outside equity. While the disciplining effect of the former is widely acknowledged, the effects of the latter are more ambiguous: Rochet (1992) shows that the effectiveness of capital requirements depends on whether banks maximize value or utility, which is related to the completeness of financial markets. Hellmann et al. (2000) argue that capital requirements put equity at risk but its cost reduce charter values. Hence, they should be complemented by deposit rate controls. Repullo (2004), however, shows that the second effect vanishes in case of intense competition. Besanko and Kanatas (1996) disentangle risk taking and monitoring and stress the adverse dilution effect. Eventually, Hakenes and Schnabel (2011) find counteracting effects of tighter capital requirements in a model where risk is jointly determined by banks and borrowers.

This paper combines the three literatures: It is most closely related to Allen and Gale (2000, 2004) and Suarez and Sussman (1997), from which the models of risk taking and moral hazard are borrowed. The key innovation is that deposits and the resource allocation are fully endogenized by modelling a two-sector economy à la Gersbach and Rochet (2012): This enables us to connect the bank’s agency problem to the economy-wide resource allocation thus replacing the common assumption in moral hazard models that deposits are elastically supplied, which opens the ground for feedback effects such as pecuniary externalities. While applied risk taking models⁵ indeed include this mechanism using a reduced-form deposit supply, a full-fledged model is required for a comprehensive welfare analysis that is a prerequisite to clearly identify inefficiencies and to derive and design optimal policies.

3 The Model

This paper develops a static, partial equilibrium model of moral hazard in banking. The economy consists of two sectors (banking and traditional sector) characterized by different technologies, which both produce a homogeneous good, as well as three types of agents (bankers, traditional firms, and investors). The technologies are described by

**Assumption 1** The banking technology includes a continuum of risky investment projects (loans) characterized by constant returns to scale and a binary payoff; the project’s success probability \( p \) is decreasing in the return if successful \( R \). The traditional (frictionless) technology is risk-free and characterized by the aggregate production function \( F(X) \) with \( F(0) = 0 \) and \( F'(X) > 0 > F''(X) \) satisfying the Inada conditions \( \lim_{X \to 0} F'(X) = \infty \) and \( \lim_{X \to 1} F'(X) = 0 \).

The idea of two technologies is borrowed from Gersbach and Rochet (2012) and goes back to Lorenzoni (2008). Importantly, only bankers have access to the banking technology, which can be justified by specific skills for monitoring or loan collection. Effectively, both technologies are operated by firms or entrepreneurs, which only play a passive role in the baseline model. Those in the

⁵For example, Allen and Gale (2000), Boyd and De Nicoló (2005), Hakenes and Schnabel (2011).
banking sector are funded by bank loans while those in the traditional sector issue corporate bonds and are directly financed by investors. Following Holmström and Tirole (1997), one may interpret the two sectors as small, credit-constrained firms and large corporations respectively. This interpretation of the banking technology is consistent with the fact that small, credit-constrained firms are, in general, riskier and more productive (as captured by the linear technology). Alternatively, one could interpret the banking technology as trading in the spirit of Boot and Ratnovski (2012) since this activity is also risky and subject to moral hazard (risk shifting).\(^6\)

The economy is populated by three types of risk-neutral agents: First, a continuum of measure one of identical bankers (owner-managers) is endowed with private wealth\(^7\) \(K_B > 0\) and each of them operates a bank that invests an amount \(L\) in the banking technology.\(^8\) The bank is funded with bank capital (inside equity) \(K \leq K_B\) and deposits \(D\). The banker is protected by limited liability. Second, a continuum of identical investors, each saving an exogenous amount \(V > 0\), solve a portfolio choice problem as they can either deposit their savings with the bank or purchase corporate bonds. Third, traditional firms (e.g., corporations) directly funded by corporate bonds invest an amount \(X\) and operate the frictionless technology. The total resources that can be invested in both sectors are normalized to one such that \(V = 1 - K_B\). If the endowment is not invested, it perishes and cannot be consumed. Figure 1 graphically illustrates the model.

---

\(^6\) However, in our setup, banks would engage only in trading.

\(^7\) An equity endowment is standard in corporate finance models with moral hazard, e.g., Holmström and Tirole (1997), and also applied in the banking literature, e.g., Rochet (1992) and Gersbach and Rochet (2012).

\(^8\) Since firms are not modeled explicitly, this is a shortcut for providing credit to firms that invest in such projects.
3.1 Agents

This section introduces the three agents of the model, bankers, traditional firms, and investors, and outlines their key decisions.

3.1.1 Bankers

The model builds on Allen and Gale (2000, 2004) to which we add equity. The banker’s definition does not only refer to the owner-manager but may as well include large shareholders whose objectives are aligned with the manager’s. Essentially, the banker’s key task is risk taking, that is, determining the risk-return profile of his portfolio.

First of all, the banker opens a bank, injects an amount $K$ of his private wealth $K_B$ as inside equity, and raises deposits $D$ from investors. He invests the remainder $K_B - K$ in corporate bonds. Then, the amount $L = D + K$ is invested in the banking technology, which is a shortcut for providing loans to small businesses. Each loan yields a binary return $R \in [R, \overline{R}]$ if successful (henceforth: target return) and zero else. The success probability is characterized by:

Assumption 2 The success probability is a function of the return if successful and satisfies:

$$p'(R) < 0, \quad p''(R) \leq 0, \quad p(R) = 1, \quad p(\overline{R}) < 1, \quad \forall R \in [R, \overline{R}]$$

There exists $R \in [R, \overline{R}]$ such that $p(R)R > 1$.

This is essentially the return structure of Allen and Gale (2000), which goes back to Stiglitz and Weiss (1981): High-yield loans are riskier while safer loans promise a smaller return if successful. Hence, the expected return $p(R)R$ is an inverse u-shaped function of $R$, which exceeds one at its maximum. The negative relationship between risk and return gives rise to the well recognized risk shifting problem of debt as soon as $R$ is unobservable. As Allen and Gale (2000), we assume that the portfolio consists of a continuum of identical loans with perfectly correlated returns but relax this assumption in section 5.3. Hence, the bank fails with probability $1 - p(R)$. This means that there is no role for a capital buffer in case of failure such that the only purpose of bank capital is to alleviate an agency problem. The banker’s objective is to maximize expected profit of both investments

$$\pi_B = p(R)(RL - bD) + \gamma(K_B - K)$$

which equals expected revenue $p(R)RL$ net of deposit repayment ($b$ denotes the gross deposit rate) plus the (gross) return of corporate bonds. The banker determines lending $L$ and capital structure (i.e., equity $K$) as well as the target return $R$. He suggests a contract that specifies deposit repayment: A deposit rate $b$ if the bank succeeds and zero if it fails. Eventually, one can define the capital loan ratio $\kappa = \frac{K}{L}$ in order to better compare the model with the related literature.

3.1.2 Traditional Firms

The model of the traditional sector builds on Gersbach and Rochet (2012)$^9$: The frictionless technology is operated by firms which, due to the absence of risk and informational frictions, can access the capital market and directly borrow from investors. More precisely, there exists a continuum of measure one of these firms each with a unit-size project and an idiosyncratic (net) productivity $x$ distributed according to some continuous and twice differentiable distribution function $G(x)$ on $[0, \infty)$. They issue corporate bonds$^{10}$ that promise a risk-free (gross) return $\gamma$. As

$^9$In contrast to Gersbach and Rochet (2012), we disentangle investors and traditional firms; they rely on a shortcut where investors directly invest in the frictionless technology.

$^{10}$This is only one possibility; any financial claim on the firm can be issued since Modigliani-Miller applies here.
a result, only sufficiently productive firms with \( x > \gamma \) earn a positive profit and invest. Aggregate investment in the traditional sector is:

\[
X = \int_{\gamma}^{\infty} dG(x)
\]  

(2)

Therefore, a fraction \( 1 - G(\gamma) \) of traditional firms invests and issues bonds; the supply of bonds equals \( X = 1 - G(\gamma) \). The remaining fraction of firms, \( G(\gamma) \), is not activated. Moreover, \( \gamma = G^{-1}(1 - X) \), and the aggregate production function characterizing the frictionless technology is:

\[
F(X) = \int_{G^{-1}(1-X)}^{\infty} x dG(x)
\]  

(3)

This function satisfies assumption 1, the Inada conditions guarantee \( X \in [0, 1] \). Importantly, the marginal product equals the bond return \( \gamma \):

\[
F'(X) = G^{-1}(1 - X) = \gamma
\]  

(4)

Intuitively, the marginal product equals the idiosyncratic productivity of the marginal firm and the bond return. One may also interpret (4) as firms’ participation constraint. Aggregate profits of traditional firms are \( \pi^F = F(X) - \gamma X \).

### 3.1.3 Investors

The model of the investor also follows Gersbach and Rochet (2012): Each investor, for example, a household, saves an exogenous amount \( V = 1 - K_B \) for end of period consumption. For that purpose, he can either deposit the savings with a bank or purchase corporate bonds. The bank promises a (gross) deposit rate \( b \) if successful (with probability \( p(R) \)) and zero else; bonds yield a safe return of \( \gamma \). Hence, investors solve a portfolio choice problem by allocating their savings between deposits \( D \) and bonds \( X = V - D \) as to maximize their expected end-of-period value:

\[
\pi^I = \max_D p(R)bD + \gamma(V - D)
\]  

(5)

Investors choose the asset that promises a higher expected return; they invest in both if expected returns are equalized:

\[
p(R)b = \gamma
\]  

(6)

Intuitively, investors need to be compensated by an ‘actuarially fair’ deposit rate for bearing the risk of bank failure. Accordingly, one can interpret (6) as a participation constraint.\(^{11}\)

### 3.2 Markets

Two markets exist in this economy: a deposit market where banks raise deposits from investors and a bond market where investors and bankers purchase corporate bonds issued by traditional firms. In equilibrium, both markets simultaneously clear:

\[
D = L - K
\]  

(7)

\[
X = V - D + K_B - K
\]  

(8)

Condition (7) can also be interpreted as a bank balance sheet identity. Walras’ law implies that the bond market clears as soon as the deposit market clears and vice versa because adding up the

\(^{11}\)Alternatively, one may think of deposit insurance with a fairly priced insurance premium \((1 - p)\gamma/p\) that ensures a balanced budget of the deposit insurance fund.
two market clearing conditions yields the aggregate resource constraint \( X + L = 1 \), which requires that total investment in both sectors equals total resources.

4 Equilibrium Analysis

This section examines three equilibria: First, a frictionless benchmark without asymmetric information ("first best") is outlined. Second, the moral hazard (risk shifting) problem of banks is introduced. Subsequently, we derive and characterize two equilibria with asymmetric information: a market or laissez-faire equilibrium and an allocation chosen by a social planner or regulator subject to the same informational constraints (henceforth: constrained social optimum). Thereby, the inefficiency of the market equilibrium is shown and implementation strategies and distributional implications of the constrained social optimum are discussed. A numerical example that illustrates the equilibria and their properties is included in appendix A.3.

4.1 First Best

This section characterizes a benchmark economy without any informational frictions. In particular, bank risk taking is observable and contractible. Then, a social planner or regulator maximizes social welfare defined as the sum of expected payoffs \( W = \pi^B + \pi^I + \pi^F \). Welfare coincides with aggregate expected output of both sectors. The choice is only subject to an aggregate resource constraint.

Program 1 The regulator maximizes social welfare \( W \) by choosing target return \( R \), bank lending \( L \), and investment in the traditional sector \( X \)

\[
W = \max_{R,L,X} F(X) + p(R)RL
\]

subject to the aggregate resource constraint \((RC)\)

\[ L + X = 1 \]

Solving program 1 yields first-order conditions \((RC)\) and

\[
p(R) + p'(R)R = 0 \quad (10)
\]

\[
F'(X) = p(R)R \quad (11)
\]

Target return and success probability are determined independently of the bank’s size and capital structure: It maximizes the expected loan return \( p(R)R \) such that the marginal gains of increasing the return equal the marginal cost due to the reduced success probability. Furthermore, traditional investment \( X \) and, by the resource constraint, bank lending \( L \) are determined as to equalize the (expected) marginal returns of both technologies: The left-hand side of (11) captures the marginal return of the traditional and the right-hand side the constant expected return of the banking technology. Therefore, the marginal gain of additional bank lending equals its opportunity cost, the forgone return from traditional investment.

The bank’s capital structure (i.e., deposits \( D \) and equity \( K \)) is indeterminate as (11) and \((RC)\) only pin down the resource allocation between the two sectors. Therefore, any capital loan ratio \( \kappa \) between zero and \( \frac{K}{L} \) is consistent with equilibrium. Furthermore, the prices \( \gamma \) and \( b \) are irrelevant for the first-best allocation as they only determine the income distribution. Nevertheless, they can be computed: The corporate bond return follows from the productivity of the marginal firm according to (4) and the deposit rate from investors’ participation constraint (6).
Importantly, condition (11) implies a zero profit margin of banks: Since the expected deposit rate, \( p(R)b \), equals the bond return, which follows from the marginal product of the traditional sector \( F'(X) \), and marginal returns are equalized across sectors, it also equals the expected return from bank lending, \( p(R)R \). In other terms, as \( L \) is not determined at the bank level, each bank expands until the rise of the equilibrium bond return and deposit rate drives the profit margin down to zero. Note that a bank only earns a positive profit if partly funded by inside equity (i.e., \( K > 0 \)). The zero profit margin is a necessary equilibrium condition given a constant returns to scale technology and perfect competition. Combining first-order conditions and resource constraint establishes:

**Proposition 1** The first-best equilibrium allocation \( \{\tilde{R}, \tilde{L}, \tilde{X}, \tilde{\gamma}, \tilde{b}\} \) is characterized by (4), (6), (10), (11), and (RC). The target return maximizes the expected loan return; expected marginal returns of both technologies are equalized such that the bank’s profit margin is zero. The capital structure is indeterminate and a larger supply of bank capital \( K_B \) has no welfare effect. The first best can be implemented as a competitive market equilibrium.

**Proof:** See appendix A.2.

The neutrality with respect to capital structure and wealth distribution implies that equity has no advantage over debt and that social welfare cannot be increased by \textit{ex ante} redistribution (e.g., from investors to bankers to create additional equity). This is a classical Modigliani-Miller result. Consequently, the bank’s capital structure (i.e., \( K \) and \( D \)) is indeterminate.\(^{12}\) Furthermore, the first best can be achieved by the market mechanism in the absence of any frictions as implied by the first fundamental theorem of welfare economics.

Figure 2 illustrates the first best: \( L \) on the horizontal axis is bank lending, the vertical axis measures the returns. The upward-sloping curve \( F'(X) \) is the marginal return of the frictionless technology\(^{13}\) while the horizontal curve indicates the expected return of the banking technology \( p(R)R \) determined according to (10). Optimality requires that the marginal returns of both technologies are equalized, and \( \tilde{L} \) is chosen accordingly. The shaded area below the two curves represents social welfare defined as aggregate output of both sectors. Accordingly, the rectangle below the horizontal line to the left of \( \tilde{L} \) is the (expected) revenue of the banking sector and the area below the \( F'(X) \)-curve to the right of \( \tilde{L} \) is the output of the traditional sector \( F(\tilde{X}) \). The surplus is shared among the three agents: Investors earn an expected income \( \pi^I \) equal to the entire rectangle below the horizontal line to the right of \( K_B \), corporate profits \( \pi^F \) are captured by the area between their

---

\(^{12}\) Note that the allocation only implies \( \tilde{K} \leq \min\{\tilde{L}, K_B\} \) and \( \tilde{D} \leq \min\{\tilde{L}, V\} \).

\(^{13}\) Note that this curve intersects the horizontal axis at \( L = 0 \) due to the Inada conditions.
marginal return and the horizontal line, and bankers achieve expected profits \( \pi^B \) represented by the rectangle below the horizontal line to the left of \( K_B \). Since expected return and deposit rate coincide, the bank earns zero expected profits on its deposit-funded investments, and the profit consists of the return on equity.

### 4.2 Moral Hazard and Risk Shifting

The agency problem is essentially the same as in Suarez and Sussman (1997), who build on the moral hazard part of Stiglitz and Weiss (1981). They focus on entrepreneurial moral hazard but their approach has also been applied to banks, for instance, by Repullo (2013). In particular, the return \( R \) if successful is unobservable or not verifiable. As a result, moral hazard and limited liability lead to risk shifting: Debt reduces the banker’s marginal cost of risk taking thus distorting the choice of the target return. Compared to Suarez and Sussman (1997), we introduce a key innovation: The required return on bank deposits is variable and endogenous as it is given by investors’ opportunity cost (i.e., the bond return).\(^{14}\) Consequently, the supply of deposits is not perfectly elastic but depends on the allocation of resources in the economy. This modification is justified because deposits are ultimately scarce given that savers can invest in alternative assets. In particular, it is a common feature in applied risk taking models such as Allen and Gale (2000) or Boyd and De Nicoló (2005) which include a reduced-form inverse deposit supply. These models, however, do not discuss the agency problem so prominently and in such detail. Endogenous opportunity cost are crucial because they are the source of a pecuniary externality that is the very reason of the inefficiency.

In general, a deposit contract that specifies deposit repayment needs to satisfy the participation constraint of investors (6). It requires the deposit rate \( b \) to be fair such that the expected repayment \( p(R)b \) equals the opportunity cost, namely, the (safe) return on corporate bonds \( \gamma \). This condition always holds regardless of an agency problem; it is also fulfilled in the first best. As soon as target return \( R \) and success probability \( p \) are private information of the banker and not observed by investors, the banker cannot commit to a particular risk-return profile that he chooses after deposits are attracted. As a result, not all contracts that ensure participation are feasible because the banker chooses the target return \( R \) regardless of the underlying risk-return profile in the contract. Since he can deviate \textit{ex post}, deposits may not be fairly priced anymore and investors are better off purchasing corporate bonds only such that the bank cannot raise deposits in the first place. Only a contract based on a risk-return profile that is privately optimal \textit{ex post} for the banker is feasible. This implies that a contract also satisfies an incentive compatibility constraint (IC):

\[
R = \arg \max_{R} p(R)[R - b(1 - \kappa)]L \tag{12}
\]

The target return needs to maximize the banker’s expected profit given a deposit rate \( b \).\(^{15}\) We subsequently rely on the first-order approach, which is valid because the objective function is concave. The first-order condition of (IC) is:

\[
p(R) + p'(R)[R - b(1 - \kappa)] = 0 \tag{13}
\]

Loans \( L \) and equity \( K \) jointly affect risk taking through the capital loan ratio \( \kappa \). As in the first best, target return balances marginal gains and cost but the latter is now distorted if the bank is funded by debt. Moral hazard and limited liability give rise to the risk shifting problem well recognized in the literature and analyzed in related models such as Suarez and Sussman (1997) or

\(^{14}\) Another minor change is that the loan return is the choice variable, which determines the success probability, while in Suarez and Sussman (1997), the entrepreneur chooses the success probability associated with convex effort cost.

\(^{15}\) The banker’s income from corporate bonds, \( \gamma(K - K_B) \), is independent of \( R \) and thus omitted.
Hakenes and Schnabel (2011): Since the debt is a flat repayment, it reduces the residual income the banker can appropriate if successful. Therefore, the banker, who is protected by limited liability and only cares about the ‘upside’ of his investment, has little to lose if the bank fails such that he prefers more profitable but riskier loans. This can be seen in (13): Debt reduces the bank’s profit margin, which can be interpreted as a ‘static’ charter value\(^{16}\), and thus lowers the marginal cost of risk taking. As a result, riskier loans that promise a higher return if successful are less costly from the banker’s perspective if the bank is partly funded by debt. Importantly, risk shifting occurs even though deposits are fairly priced\(^{17}\) in equilibrium: Moral hazard and limited liability prevent the banker from committing to a prudent lending strategy (such as the first best) as long as the bank is partly funded by debt. Essentially, this feature is already in Jensen and Meckling (1976), where a firm always has an incentive to choose a (potentially) inefficient, risky portfolio although its creditors correctly price the debt, and in Repullo (2013), where risk shifting occurs even though banks pay a fair interest rate on uninsured deposits in equilibrium.

\[\begin{align*}
\text{Figure 3: Risk Shifting and Feasible Contract} \\
\end{align*}\]

Figure 3 illustrates the feasible deposit contract and risk shifting: The upper panel shows the two constraints, the lower panel plots expected loan against target return. The IC-curve includes all combinations of \(b\) and \(R\) that are incentive-compatible, the PC-curve all combinations that satisfy the participation constraint. Both curves are upward-sloping, the former since a higher deposit rate exacerbates risk shifting, the latter since investors require a compensation for a higher risk of failure. The set of feasible contracts is thus given by the arc of the IC-curve between the two intersections of IC and PC. Obviously, moving leftwards raises expected bank profits\(^{18}\) such that point B defines the second-best contract. The dashed, vertical line represents the first-best target return \(\hat{R}\), which is independent of the deposit rate and maximizes expected loan return. Hence, the first-best contract is given by point A. Risk shifting leads to an inefficient portfolio with too high target return and loan risk. Position and shape of both constraints depend on capital loan ratio \(\kappa\) and corporate bond return \(\gamma\): A higher capital loan ratio makes the IC-curve steeper and for

\(^{16}\)As the model is static, the disciplining effect is provided by the prospect of realizing the current profit at the end of the period.

\(^{17}\)Government guarantees or incorrectly priced deposit insurance are not necessary but are likely features of banks that would strengthen this result.

\(^{18}\)Higher expected loan return \(pR\) and lower deposit rate \(b\).
\( \kappa = 1 \), it coincides with the first-best target return. Consequently, the first best could, in principle, be replicated if the bank is fully funded by equity, which eliminates risk shifting. However, this can only be an equilibrium outcome under rather restrictive conditions (see section 4.3.3). A higher bond return \( \gamma \) shifts the PC-curve upwards as for any given risk-return profile, investors require a higher deposit rate. Obviously, the curves might not always intersect such that no deposit rate that is both fair and incentive-compatible exists. Hence, banks cannot attract deposits, and investors allocate their entire savings to the traditional sector. This case may occur if banks are poorly capitalized such that the risk shifting problem is severe (the IC-curve is rather flat) and if investors require a high return since corporate bonds are attractive (the PC-curve is shifted upwards). Hence, a feasible contract exists only if \( \kappa \geq \kappa_0 \); \( \kappa_0 \) is given by the tangency of the IC- and PC-curve and is an increasing function of \( \gamma \). Since both capital loan ratio and bond return are endogenous, the equilibrium resource allocation always ensures a feasible deposit contract.

Risk taking and deposit rate are thus jointly determined by the constraints (6) and (13) for given capital structure and opportunity cost. It is characterized by:

**Lemma 1** Target return satisfies \( \frac{\partial R}{\partial \kappa} < 0 \), and \( \frac{\partial R}{\partial \gamma} \geq 0 \), the deposit rate satisfies \( \frac{\partial b}{\partial \kappa} < 0 \), and \( \frac{\partial b}{\partial \gamma} > 0 \).

**Proof:** See appendix A.2.

A higher capital loan ratio reduces the bank’s debt burden thereby increasing residual income and the marginal cost of risk taking, which alleviates risk shifting. The reduced bank risk, in turn, allows lowering the deposit rate. The capital loan ratio can be raised in two ways: injecting more equity and deleveraging. Although the model, in principle, allows for both, ultimately only deleveraging occurs in equilibrium because the banker will inject his entire private wealth as inside equity in the first place (i.e., \( K = K_B \)) exactly as to mitigate risk shifting.\(^{19}\) A higher bond return \( \gamma \) tightens the participation constraint and requires a higher deposit rate. This, in turn, reduces marginal cost of risk taking inducing the banker to choose a higher target return. Essentially, a high capital loan ratio and low funding cost are substitutes on the ‘risk shifting-front’ as they both strengthen the banker’s incentive for prudent lending. Since the resource allocation is endogenous, they are interconnected and bank capital provides discipline through two different mechanisms: At the bank level, it directly increases the banker’s ‘skin in the game’. At the industry level, a higher equity share implies a smaller bank leverage, which lowers the demand for deposits such that more resources are allocated to the traditional sector. Hence, the bond return falls and the participation constraint is relaxed creating a positive feedback effect on each bank as a lower deposit rate discourages risk taking even further. Critically, the second effect is not internalized in a competitive market equilibrium and is the very reason for its inefficiency.

However, given the advantage of equity on the ‘risk shifting front’, bankers would prefer full equity funding, which obviously does not reflect reality. In such models, an interior equity share is ensured by making bank capital scarce\(^{20}\) or more expensive\(^{21}\) than deposits.\(^{22}\) In this framework, equity is not expensive because banker and investor face similar opportunity cost (the bond return \( \gamma \)), which is the reason why he prefers injecting as much equity as possible. The fixed supply (endowment) of bank capital eventually limits the lending capacity thus creating marginal cost of raising the equity share even further by deleveraging. Another advantage of focusing on inside equity is that it unambiguously alleviates the agency problem; in case of outside equity, the target

\(^{19}\)This is due to the incentive effect of equity which implies an excess return over bonds. See, section 4.3.

\(^{20}\)For example, Suarez and Sussman (1997) where it is effectively given by a deterministic first-stage production or Gersbach and Rochet (2012).

\(^{21}\)For example, Repullo (2013), Hakenes and Schnabel (2011)

\(^{22}\)Moreover, there are models, e.g., Acharya et al. (2013), with multiple agency problems, some of which (e.g., rent seeking) can be resolved by using debt such that the capital structure has an interior solution.
return would have to be ex post observable for profit sharing and the positive incentive effect would prevail only the absence of dilution effects or shareholder-manager conflicts such that the banker indeed maximizes the joint surplus.

In principle, this contract is achieved without any intervention as investors only deposit their savings with the bank if the offered deposit rate is both fair and incentive-compatible. Hence, the agency problem itself provides no rationale for intervention because it is resolved by private contracting. In the spirit of the representation hypothesis of Dewatripont and Tirole (1994), one may, however, argue that implementing this second-best contract is difficult (e.g., due to lack of sophistication and coordination problems) and expensive (e.g., due to duplication cost) for investors such that it is preferable if a regulator acts on their behalf.

4.3 Market Equilibrium

This section derives the competitive market or laissez-faire equilibrium under information asymmetry; all agents are unconstrained by any regulatory requirements.

4.3.1 The Banker’s Problem

The banker determines bank lending and capital structure as well as target return and deposit rate (i.e., the deposit contract) as to maximize expected profits. His choice is subject to several constraints: The deposit contract has to be feasible and to satisfy incentive compatibility and participation constraint. In addition, equity is restricted by the banker’s wealth endowment, which is captured by a capital availability constraint. Hence, the banker solve

Program 2 The banker maximizes expected bank profit \( \pi^B \) by choosing target return \( R \), deposit rate \( b \), loans \( L \), and equity \( K \)

\[
\pi^B = \max_{R, b, L, K} p(R)[R - b(1 - \kappa)]L + \gamma(K_B - K) \tag{14}
\]

subject to the incentive compatibility constraint \( (IC) \)

\[
R = \arg \max_R p(R)[R - b(1 - \kappa)]L
\]

the investors’ participation constraint \( (PC) \)

\[
p(R)b = \gamma
\]

and the capital availability constraint \( (CA) \)

\[
K_B \geq K
\]

This optimization problem is similar to Suarez and Sussman (1997) and Repullo (2013); in particular, objective function and incentive compatibility constraint only differ in their choice variables as \( b \) and \( \kappa \) are taken as given in \( (IC) \). Target return and deposit rate follow from incentive compatibility and from participation constraint as discussed above. The allocation of the banker’s wealth \( K_B \) between equity and bonds is given by

\[
\gamma = p(R)b + \frac{\lambda p'(R)b}{L} - \zeta \tag{15}
\]
where $\lambda < 0$ and $\zeta$ denote the Lagrange multipliers of (IC) and (CA) respectively. If the banker purchases bonds, he earns the bond return $\gamma$; if he invests his wealth in his own bank, he reduces the expected deposit repayment and, on top of that, alleviates risk shifting by increasing the capital loan ratio such that the expected loan return is higher. This is captured by the term $\frac{\lambda p(\tilde{R})b\kappa}{L}$.

Due to this positive incentive effect, equity earns an implicit excess return over bonds and the banker, therefore, injects the entire endowment in his own bank as equity. This can be seen since (15) and (PC) jointly imply $\zeta > 0$ as long as the incentive compatibility constraint binds (i.e., $\lambda < 0$). By complementary slackness, the capital availability constraint binds and $K = K_B$. Moral hazard is the very reason why the banker is not indifferent between the two assets like in the first best but strictly prefers equity to bonds. If his wealth rises, he is willing to provide even more equity to further alleviate risk shifting, and if it is large enough, he will even prefer full equity funding ($\kappa = 1$), which completely eliminates risk shifting. As soon as the supply of bank capital is scarce, however, the banker trades off portfolio quality against lending scale. Note that due to $K = K_B$, the capital loan ratio is determined solely by the amount of loans $L$ and any increase of $\kappa$ necessarily involves deleveraging. Moreover, the choice of $L$ is characterized by the first-order condition

$$p(R)(R - b) = \frac{\lambda p'(R)b\kappa}{L}$$

(16)

This condition requires that the banker’s marginal gains and cost of bank lending are equalized. In the first best (i.e., if $\lambda = 0$) are no marginal cost at the bank level (since the technology is linear and the marginal return constant) and (16) eventually coincides with the zero profit condition (11). In the presence of moral hazard, however, there exist marginal cost of a higher $L$ as *ceteris paribus* the capital loan ratio declines, which exacerbates risk shifting and reduces the expected loan return. Contrary to the first best, (11), this condition implies that the bank earns a positive profit margin on its externally funded investments because the expected cost of deposit finance $p(R)b$ are smaller than the expected revenue from lending $p(R)R$, that is, the banker earns an informational rent. This rent can be interpreted as a static counterpart of charter value as the opportunity of earning this rent alleviates risk shifting.

### 4.3.2 Traditional Firms, Investors, and Market Clearing

Investment by traditional firms is determined at the extensive margin, and only sufficiently productive firms invest and issue bonds (i.e., $x \geq \gamma$). Hence, (4) implies that the bond return equals the marginal product of the frictionless technology, $F'(X) = \gamma$. Investors maximize the end-of-period value of their savings $V = 1 - K_B$ by allocating them between deposits and corporate bonds as described in section 3.1.3. Hence, deposits balance (expected) returns from bonds and deposits, which is already captured by the participation constraint (PC) in program 2.

Eventually, both bond and deposit market clear in equilibrium, and the market clearing conditions (7) and (8) hold.

### 4.3.3 Equilibrium Allocation

Combining the privately optimal choices of the agents yields the following proposition:

**Proposition 2** The equilibrium allocation $\{\tilde{R}, \tilde{L}, \tilde{X}, \tilde{K}, \tilde{D}, \tilde{b}, \tilde{\gamma}\}$ is characterized by (4), (7), (8), (13), (PC), (CA), and

$$F'(X) = \frac{p(R)R}{1 + \frac{\lambda p'(R)b\kappa}{p(R)L}}$$

(17)

The first best can be implemented as a competitive market equilibrium with $\kappa = 1$ if

$$K_B \geq G[p(\tilde{R})\tilde{R}]$$

(18)
Otherwise, moral hazard distorts resource allocation and risk taking: Compared to the first best, there is underinvestment in the banking sector (i.e., $\hat{L} < \bar{L}$) and bank loans are more profitable but riskier (i.e., $\hat{R} > \bar{R}$) resulting in a higher probability of bank failure. The banker invests his entire private wealth in the bank (i.e., $\hat{K} = \mathbb{K}_B$).

**Proof**: See appendix A.2.

The central condition is (17), which characterizes the resource allocation. It follows from combining the first-order conditions of traditional firms and bank, (4) and (16), with (PC). It requires that $L$ and $X$ balance marginal returns of the two technologies taking into account a potential risk shifting effect. In principle, the first best can be replicated as a market equilibrium even if the target return is unobservable: This requires full equity funding (i.e., $\kappa = 1$ and $K = \bar{L}$), which eliminates risk shifting associated with debt such that the first-best deposit contract is feasible (point A in figure 4). However, this is only possible if the supply of bank capital $K_B$ is large enough: Condition (18) requires it to be at least as large as first-best bank lending $\bar{L} = G[p(\hat{R})\hat{R}]$, which depends on the relative productivities of both technologies. Then, the incentive compatibility constraint is slack (i.e., $\lambda = 0$) and the conditions (17) and (13) coincide with their first-best counterparts (11) and (10). Nevertheless, full equity funding is not an appropriate description of reality, banks usually funded by very a small share of equity. Therefore, we subsequently focus on the case where bank capital is scarce (i.e., $K_B$ is small) and condition (18) is not satisfied. Since banks are partly funded by deposits, moral hazard and risk shifting are present such that the incentive compatibility constraint binds ($\lambda < 0$). This reduces the bank’s marginal return compared to the first best: First, risk shifting directly lowers the expected return, $p(R)R$, as discussed above. Second, higher bank lending reduces the capital-loan ratio thereby further weakening the incentive to invest prudently such that portfolio quality deteriorates (see, lemma 1). This effect is captured by the denominator of (17), which is larger than one and drives a wedge between the effective marginal returns of both sectors allowing the banker to earn an informational rent (i.e., a positive profit margin). Given the lower returns of bank lending, less productive firms in the traditional sector attract funding such that its marginal product decreases as well, $F'(\hat{X}) < F'(\bar{X})$. Compared to the first best, more resources are allocated to the traditional sector resulting in underinvestment in banks compared to the first best. This arises as scarce bank capital is needed to preserve incentives thereby limiting banks’ lending capacity. This is a typical rationing result due to asymmetric information.

\[\text{Figure 4: Market Equilibrium}\]

---

23Recall that under moral hazard, the banker injects his entire private wealth as inside equity. The relation between lending scale and leverage thus becomes mechanical.
In general, moral hazard distorts the banker’s risk, lending, and leverage choices resulting in excessive risk taking as well as underinvestment in the banking sector. Consequently, welfare is lower than in the first best as illustrated in figure 4: The blue-shaded area below the curves corresponds to social welfare in the market equilibrium \( \hat{W} \). The distortions of risk (lower expected revenue) and size (smaller banks) result in a deadweight loss shown by the gray-shaded area. This consist of ‘quality’ loss due to an inefficiently risky loan portfolio as well as a ‘quantity’ loss due to the reallocation of resources towards the less efficient traditional sector. The former essentially represents the agency cost of debt emphasized by Jensen and Meckling (1976). Moreover, one can observe that the banker now earns an informational rent as the lending spread is strictly positive \( [p(R)R > \gamma = p(R)b] \) but his return on equity is lower. Whether overall bank profits increase compared to the first best is thus ambiguous. Corporate profits increase as due to a lower bond return the profit of each firm increases and more firms find it profitable to invest. In contrast, investors’ expected income is reduced as they earn lower returns on both assets.

4.4 Constrained Social Optimum

Unless the supply of bank capital is large enough to ensure full equity funding, the market equilibrium is constrained-inefficient because of a pecuniary externality: When raising deposits by choosing size and leverage, banks fail to internalize that attracting these resources changes the relative price of deposits. As traditional investment declines, the productivity of the marginal firm increases, which implies a higher bond return. This, in turn, tightens investors’ participation constraint and requires a raise of the deposit rate that reduces profit margins and exacerbates the risk shifting problem of all other banks resulting in inefficiently risky loan portfolios. Since moral hazard establishes a mechanical relation between equity and the banker’s wealth endowment, the pecuniary externality is directly associated with the lending choice. Note that a single bank does not influence the deposit rate (because it is infinitesimally small) but instead the fact that all banks choose their size neglecting the price reaction causes an aggregate externality. Consequently, banks in the market equilibrium are too large and have too much leverage. This inefficiency is similar in kind to Lorenzoni (2008) and Gersbach and Rochet (2012) but the externality is associated with the deposit rate instead of asset prices. Critically, its presence hinges on the combination of moral hazard with perfect competition\(^\text{24}\) and a deposit supply that is not perfectly elastic: Moral hazard makes risk taking sensitive to the deposit rate, perfect competition implies non-internalization due to price taking behavior, and the inelastic supply of deposits links the resource allocation to the deposit rate through endogenous opportunity cost in the participation constraint. In the first best, in contrast, the target return is independent of the deposit rate and the pecuniary externality vanishes. If deposits were supplied elastically as in Suarez and Sussman (1997) and Repullo (2013), raising deposits would not increase investors’ opportunity cost and would only directly affect risk taking through the capital loan ratio. Although the reason why bank capital is necessary in this framework is essentially an agency problem of each bank, the pecuniary externality is an aggregate phenomenon therefore giving rise to macroprudential regulation.

This section examines the allocation chosen by a risk-neutral social planner or regulator subject to the same informational constraints as the market participants.\(^\text{25}\) In particular, the regulator cannot restore the first best unless condition (18) holds, and the allocation is only constrained-efficient.

4.4.1 The Regulator’s Problem

Acting as a social planner, the regulator maximizes social welfare \( W \). In contrast to the first best, the regulator cannot observe the target return and needs to design a deposit contract that is feasible

\(^{24}\)Imperfect competition leads to partial internalization but due to a different mechanism; see, section 4.5.2.

\(^{25}\)Essentially, the regulator cannot observe the banker’s choice of the loan return and not directly control \( R \).
and preserves the banker’s incentives. Therefore, prices are not irrelevant as the contract specifies the deposit rate $b$. Obviously, setting $b = 0$ would eliminate risk shifting and restore the first best. However, the regulator can usually not force investors to deposit their savings with the bank at a zero (gross) rate. Hence, the choice is subject to investors’ participation constraint. For notational simplicity, we combine the participation constraints of traditional firms and investors, (4) and (6), thereby eliminating $\gamma$. Therefore, the regulator solves the following optimization problem:

**Program 3** The regulator maximizes welfare $W$ by choosing target return $R$, equity $K$, lending $L$, traditional investment $X$, and the deposit rate $b$

$$W = \max_{R,b,L,X,K} F(X) + p(R)RL$$

subject to the banker’s incentive compatibility constraint (IC)

$$R = \arg \max_R p(R)[R - b(1 - \kappa)]L$$

the combined participation constraint (CP)

$$p(R)b = F'(X)$$

the aggregate resource constraint (RC)

$$L + X = 1$$

and the capital availability constraint (CA)

$$K_B \geq K$$

### 4.4.2 Equilibrium Allocation

Solving program 3 yields the following proposition:

**Proposition 3** The equilibrium allocation $\{R^*, L^*, X^*, K^*, b^*, \gamma^*\}$ is characterized by (4), (7), (13), (CP), (RC), (CA), and

$$F'(X) = \frac{p(R)R}{1 + \frac{\lambda p'(R)\kappa}{p(R)R} \left[ 1 - \frac{(1-\kappa)F''(X)L}{\lambda p'(R)\kappa} \right]}$$

If condition (18) holds, banks are fully equity-funded and the constrained social optimum coincides with the first best. Otherwise, the regulator allocates fewer resources to the banking sector than in the market equilibrium (i.e., $L^* < \hat{L}$). Banks are better capitalized (i.e., $\kappa^* > \hat{\kappa}$) and the deposit rate is lower (i.e., $b^* < \hat{b}$) such that the profit margin increases resulting in the choice of a safer and more efficient loan portfolio (i.e., $R^* < \hat{R}$). A larger supply of bank capital is welfare-improving as long as $K_B < G[p(\hat{R})\hat{R}]$.

**Proof:** See appendix A.2.

The central condition is (20), which pins down the constrained-efficient resource allocation. It requires the social marginal return of both technologies to be equalized taking into account all incentive effects. Importantly, the social return of bank lending is lower than the private in (17) because it includes the pecuniary externality. This is captured by the third term in the denominator that represents the response of investors’ opportunity cost to a change of lending. Hence,
the pecuniary externality depends on the shape of the production function, which is intuitive as its curvature indicates the responsiveness of marginal returns (i.e., opportunity cost) to traditional investment. Note that the adverse welfare effect of the pecuniary externality is conditional on moral hazard: As soon as \( \lambda = 0 \), it vanishes because risk taking is then independent of the deposit rate as shown in section 4.1 and the mechanism through which a bank’s deposit demand affects risk taking of other banks breaks down. In the presence of moral hazard, however, the pecuniary externality reduces the social marginal return of bank lending as lending increases deposit demand and eventually raises the deposit rate \( b \) and exacerbates risk shifting of all other banks. If \( X = \hat{X} \) and \( L = \hat{L} \) in (20), traditional investment yields an excess social return; the regulator thus reallocates additional resources to the traditional sector. As a result, the marginal product of this technology falls due to the concavity of \( F(X) \) whereas a higher capital loan ratio and lower funding cost alleviate risk shifting and raises the social marginal return of bank lending. In other terms, a regulator solves the quality-quantity trade-off more in favor of quality. In general, this line of reasoning is consistent with the prominent finding of the literature that bank competition may weaken financial stability: In more competitive markets, banks are larger such that deposit demand raises the deposit rate thereby eroding profits and exacerbating moral hazard.

![Figure 5: Constrained Social Optimum](image)

Unless the supply of bank capital is large enough, this allocation is not first-best because moral hazard still distorts the risk and lending choices but it is constrained-efficient. Hence, social welfare is lower than in the first best but there are welfare gains compared to the market equilibrium as illustrated in figure 5. There are two counteracting effects: On the one hand, the bank’s portfolio is less risky and more efficient resulting in a higher expected net revenue (‘quality effect’). On the other hand, banks are smaller and more resources are invested in the comparatively less productive frictionless technology (‘quantity effect’). However, the positive quality effect dominates since this allocation is chosen as to maximize social welfare.

Whereas the first best is neutral with regard to the initial wealth distribution, a higher endowment of bankers now translates into a larger supply of equity, which improves social welfare as long as equity is scarce and banks are partly funded by debt. Two effects or a combination of them are possible\(^{26}\): For a given bank size, more equity raises the capital loan ratio, discourages risk shifting and leads to a better portfolio quality. For a given risk-return profile, it allows for larger banks thereby allocating more resources to the more productive banking sector. Importantly, this means that the regulator could, in principle, improve welfare through some \textit{ex ante} redistribution from investors to bankers. He could even restore the first best if an amount \( \hat{L} - K_B \) was redistributed.

\(^{26}\)For an illustration, see scenario (2) in the numerical example (appendix A.3).
4.5 Implementation

The constrained social optimum characterized in proposition 3 is rather flexible in terms of its implementation. Regulation is necessary only if the supply of bank capital is scarce enough. In this case, there are two general approaches to implement the constrained social optimum in a market economy: standard prudential regulation and imperfect competition. The latter also allows interpreting the constrained social optimum in the context of financial liberalization, a policy pursued in many countries during the last three to four decades.

4.5.1 Prudential Regulation

A natural approach is prudential regulation. We focus on two standard instruments: minimum capital requirements and deposit rate ceilings. While the former is standard and widely used, the latter though rarely in place today was a key component in previous regulatory frameworks especially in the U.S. (regulation Q). Since the regulator’s choice satisfies incentive compatibility and participation constraint, the constrained social optimum results as soon as the capital loan ratio is set at \( \kappa = \kappa^* \) or the (risk-adjusted) deposit rate at \( b = b^* \):

**Corollary 1** The constrained social optimum can be implemented in a market economy by minimum capital requirements \( \kappa \geq \kappa^* \) or deposit rate ceilings \( b \leq b^* \).

**Proof**: See appendix A.2.

Essentially, \( \kappa \) and \( b \) are mechanically related by the resource and investors’ participation constraint such that \( b = b^* \) follows as soon as \( \kappa = \kappa^* \) and vice versa. More precisely, the relation between \( L \) and \( b \) is fixed but since the banker prefers injecting his entire wealth given moral hazard (i.e., \( K = K_B \)), it extends to \( \kappa \) as well. The intuition is that regulation aims at internalizing the pecuniary externality associated with bank deposits on others’ risk shifting problem. For that purpose, the regulator needs to prevent the deposit rate from rising too strongly, which can be achieved either by limiting leverage and deposit demand or direct price control. This equivalence result, at first sight, contrasts with Hellmann et al. (2000), who argue that capital requirements are Pareto-inferior compared to deposit rate controls or even counterproductive as (outside) equity is costly reducing charter value. However, our model is static such that future profits and charter value, which may decrease when using costly outside equity, do not affect incentives. In addition, the focus on inside equity eliminates any shareholder-manager conflict, which may weaken the positive incentive effect of bank capital due to a dilution of the banker’s stake as shown by Besanko and Kanatas (1996).

4.5.2 Imperfect Competition

Since prudential regulation essentially aims at increasing the informational rent to induce prudent lending by keeping the deposit rate artificially low compared to the market, one might think of another institution that achieves a similar outcome: imperfect competition. Suppose that the banking sector is not perfectly competitive and banks compete for deposits in a Cournot fashion, which is a common assumption in the literature, for example, in Allen and Gale (2004) or Boyd and De Nicoló (2005). For that purpose, assume that instead of a continuum a finite number \( N \in \mathbb{N} \) of bankers each endowed with private wealth \( \kappa B_i \) exists. The key difference to perfect competition is that each banker takes into account how investors’ opportunity cost and deposit

---

27Note that size restrictions (i.e., an upper bound on lending or deposits) also work but they are less common.

28As in Allen and Gale (2004), aggregate bank capital endowment remains bounded if \( N \) increases; otherwise, as \( N \to \infty \), \( K \to \infty \) and, since banks are partly deposit-funded (i.e., \( \kappa < 1 \)), \( D \to \infty \), which is inconsistent with equilibrium if resources are limited.
rate react to his choice of loans and deposits given the optimal choices of all other bankers. As a result, banks are smaller and raise fewer deposits such that the lower deposit rate allows them to earn a rent (i.e., the lending spread increases). Compared to perfect competition (program 3), the optimization problem of bank $i$ is slightly adjusted since the bond return in the participation constraint is not taken as given anymore and replaced by an inverse deposit supply function:

**Program 4** The bank maximizes its expected profit $\pi^B_i$ by choosing target return $R_i$, deposit rate $b_i$, loans $L_i$, and equity $K_i$

$$\pi^B_i = \max_{R_i, b_i, L_i, K_i} p(R_i)[R_i - b_i(1 - \kappa_i)]L_i + \gamma(K_{Bi} - K_i)$$ (21)

subject to the incentive compatibility constraint (IC), the capital availability constraint (CA), and the adjusted investors' participation constraint (PC)

$$p(R_i)b = F'\left[1 - \sum_{i=1}^{N} L_i\right]$$

The term in square brackets equals traditional investment $X$. In a symmetric Cournot-Nash equilibrium, the allocation is described by:

**Lemma 2** The equilibrium allocation $\{R', L', X', D', K', b', \gamma'\}$ is characterized by (4), (7), (8), (13), (PC), (CA) and

$$F'(X) = \frac{p(R)R}{1 + \frac{\lambda p'(R)\kappa}{p(R)L} \left[1 - \frac{(1-\kappa_i)F''(X)L}{p(R)L} - \frac{1-\kappa}{p(R)b} \frac{F''(X)L}{N}\right]}$$ (22)

The banking sector is smaller (i.e., $L' < \hat{L}$), the deposit rate and the failure probability are lower (i.e., $b' < \hat{b}$ and $p' < \hat{p}$) compared to the perfectly competitive market equilibrium characterized by proposition 2. For $N \to \infty$, these two equilibria coincide.

**Proof:** See appendix A.2.

Condition (22) characterizes the resource allocation. It differs from perfect competition, (17), because the banker considers the deposit interest rate as endogenous and internalizes the feedback effects on his bank's profit: There is the typical direct effect due to lower funding cost (captured by the fourth term in the denominator) as well as an indirect effect due to changes in risk taking (captured by the third term). As a result, the private marginal return of bank lending is smaller than under perfect competition, and bankers reduce lending such that the overall deposit demand falls, which is a standard rationing result of the Cournot model. This can be seen in (22) as the lower expected return implies an increase in traditional investment and, by the resource constraint, a decrease in bank lending. As implied by lemma 1, the higher capital loan ratio and the lower opportunity cost of investors mitigate risk shifting. Furthermore, the relationship between the amount of loans $L$ and the number of banks $N$ is positive because the denominator in (22) decreases in $N$ thus raising the private marginal return of bank lending. Intuitively, the impact of a single bank on the deposit rate diminishes as more banks compete, and the profit reduction from additional lending is lower: In concentrated markets, the banking sector is small, profits are large and portfolios safe while in competitive markets, the banking sector is large, profits are small and portfolios risky. Moreover, if $N \to \infty$, an infinitely large number of very small banks operate, and the impact of a single bank on the deposit rate is negligible; price taking behavior is restored and the perfectly competitive outcome results.
Objective function and first-order condition of banker and regulator differ: The former internalizes an adverse feedback effect on his own profit, the latter the pecuniary externality that distorts risk taking of all other banks. Although both approaches create rent opportunities leading to more prudent lending than under perfect competition, the two outcomes are generally not the same, and the imperfectly competitive market equilibrium is, in principle, inefficient. To illustrate this, we focus on two extreme cases: In a monopsony, \( N = 1 \), the bank fully internalizes the deposit rate’s response. Therefore, (22) implies \( L'(1) < L^* \), and the monopsonistic bank is too small resulting in underinvestment in the banking technology. The intuition is that the banker too strongly reduces lending in order to lower the deposit rate and to increase the rent. Thus, the loan portfolio is safer and more efficient than in the constrained social optimum. The first effect, however, dominates such that the monopoly is inefficient overall. If \( N \to \infty \), in contrast, there is no internalization, and the allocation converges to the competitive outcome, which is constrained-inefficient as banks fail to internalize the pecuniary externalities leading to overinvestment by banks and excessive risk taking. This line of reasoning implies

**Corollary 2** There exists\(^{29}\)

\[
N^* \approx \frac{p(R^*)[2p'(R^*) + p''(R^*)(R^* - b^*(1 - \kappa^*))]}{p'(R^*)^2b^*(1 - \kappa^*)} > 1
\]

such that the constrained social optimum can be approximately implemented as an imperfectly competitive market equilibrium with \( N^* \) operating banks.

**Proof:** This follows from equalizing marginal returns of bank lending in both equilibria (22) and (20) and solving for \( N^* \). As soon as \( L = L^* \), \( X = X^* \) follows from the resource constraint such that (13) and (PC) jointly imply \( R' = R^* \) and \( b' = b^* \). \( N^* > 1 \) is implied by a property of the feasible contract (see (38) in appendix A.2). Q.E.D.

Note that this relation might not hold with equality due to the discrete nature of \( N \). Corollary 2 states that for a specific number of banks \( N^* \), the market mechanism incidentally achieves the constrained social optimum without any regulatory intervention apart from issuing licenses. This is possible though the objective functions of regulator and oligopsonistic bank differ. Since the degree of internalization of an oligopsonistic bank decreases in the number of banks whereas the pecuniary externality and, thus, the magnitude of the regulator’s corrective intervention are constant, there exists a particular value of \( N \) for which both effects are almost the same such that the bank incidentally chooses the efficient lending scale (i.e., \( L' = L^* \)). Consequently, the constrained social optimum can also be implemented by simply issuing \( N^* \) banking licenses. Alternatively, the regulator can issue \( N > N^* \) licenses and, for instance, impose capital requirements as discussed above. Both strategies yield an equivalent outcome although the latter appears more realistic and preferable.\(^{30}\)

For \( N < N^* \), however, prudential regulation cannot achieve constrained efficiency because banks are inefficiently small in order to extract rents.\(^{31}\) Therefore, financial liberalization, in the sense of increasing the intensity of competition by issuing more licenses, has a positive welfare effect as it reduces market power and allocates more resources to the productive banking technology. At the same time, however, pecuniary externalities become more damaging resulting in higher risk and weaker financial stability. If competition is intensified beyond a certain degree (i.e., if \( N > N^* \)), optimal policy requires accommodation by macroprudential regulation.

\(^{29}\)If \( N = 1 \), the denominator of (22) is larger than in (20). The marginal return of a monopsonistic bank is smaller than the social marginal return to bank lending implying that banks are smaller.

\(^{30}\)The reason is that regulation exactly achieves the constrained social optimum and appears more suitable as it adjusts at the intensive margin while otherwise some banks must enter or exit the market.

\(^{31}\)Then, the regulator would need to induce banks to expand.
This result has implications beyond the equivalence of imperfect competition and optimal regulation: The policy of liberalizing the banking sector (e.g., deregulation of deposit, removal of branching restrictions) and relying on capital requirements as the primary regulatory instrument instead that has been pursued in many countries since the 1970s and 1980s should, in principle, not have increased risk taking and weakened financial stability since the outcome is independent of the instrument. The fact that the frequency of banking crises, nevertheless, increased sharply indicates that (i) the banking sector might have been too concentrated prior to liberalization (i.e., $N < N^*$) and the increase in risk is an optimal response as the positive welfare effect of larger banks that employ a superior technology compared to traditional firms dominates or (ii) capital requirements imposed after liberalization (e.g., by the Basel accords) might have been insufficiently low. Importantly, the latter can be explained by the fact that the predominant microprudential approach likely neglected interactions between banks and aggregate phenomena such as pecuniary externalities.

4.6 Distributional Implications

In figure 5, one can observe that traditional firms gain from regulation whereas investors lose since the lower demand for deposits drives down deposit rate and bond return; the combined effect for firms and investors is negative. The change of the banker’s expected income is characterized by the very same quality-quantity trade-off than overall welfare: On the one hand, lower opportunity cost of investors reduce the deposit rate and strengthen incentives such that the profit margin is higher; on the other hand, banks are smaller. However, the first effect dominates. This is summarized by

**Corollary 3** *Prudential regulation is not a Pareto improvement since it only increases the surplus of bankers and firms with access to direct finance but reduces the surplus of investors.*

**Proof:** The welfare implications for investors and traditional firms follow from $\gamma^* < \hat{\gamma}$ and the fact that savings are constant; the banker’s gains are necessary for overall welfare to increase compared to the market equilibrium since the combined effect for firms and investors is negative. Q.E.D.

Optimal bank regulation is consistent with the idea of financial restraint as the regulator keeps interest rates (artificially) low thereby creating rents for banks and traditional firms. Low deposit rates are, however, necessary to alleviate risk shifting and to improve financial stability as well as overall efficiency. This is similar to Hellmann et al. (1997) who essentially argue that the government should create rent opportunities in order to induce efficient actions of banks. Moreover, given the interpretation that the banking technology is operated by (small) entrepreneurs exclusively relying on bank loans (due to agency problems and financial constraints), regulation has an asymmetric effect on the real sector: While traditional firms (e.g., large corporations) benefit from more attractive funding conditions and invest more, small businesses find it more difficult to obtain credit since the lending capacity of banks is constrained.

However, three caveats remain: First, investors’ savings are exogenous and constant, which is a reasonable assumption in a static model and keeps the analysis tractable. However, one can expect that investors in fact consume more and save less if the return falls. Consequently, the bond return decreases by a smaller amount than predicted and the welfare effects of bank regulation are less pronounced. Second, the welfare function implies that the regulator places the same weight on the surplus of all agents. If the relative welfare weight of investors, for example, were larger, allocation and distribution would differ from this benchmark. Third, in reality, banks are reluctant to accept

---

32Hellmann et al. (1997) define financial restraint as a set of policies that aim at creating rents (i.e., excess returns compared to a competitive benchmark) within the private sector (e.g., by setting deposit rate ceilings below their competitive level). This differs from financial repression where the government extracts rents from the private sector.
regulatory interventions and are spending large amounts on lobbying against regulation. Given their revealed preference, how can one reconcile it with the view that banks may even benefit from regulation? First of all, regulation addresses only one inefficiency in our analysis, namely, a pecuniary externality. In reality, many other inefficiencies exist (e.g., contagion, bankruptcy cost) that justify tighter regulation. Given the trade-off between quality and quantity discussed above, this may eventually reduce overall bank profits and only raise the profit margin, which is key for incentives. As an extension, we discuss in section 5.1 reduced-form social cost of bank failure that motivate further interventions. In addition, the focus on inside equity removes any scope for a shareholder-manager conflict on the bank’s side: The manager may also earn private benefits proportional to the bank’s size (e.g., prestige); then, the negative quantity effect is much stronger and regulation is more likely to reduce the bank’s (private) profit. Can the welfare gains be redistributed such that regulation is eventually a Pareto-improvement? The welfare gain of traditional firms can be redistributed by a lump-sum tax. Taxing bankers, however, requires a tax that does not distort risk taking and that is levied on an observable quantity. Even a lump-sum tax would distort risk taking by reducing residual income and the marginal cost of risk taking. A proportional tax on profit would be non-distorting but since loan return is unobservable, it violates the second property. Consequently, only the welfare gains of traditional firms can be redistributed, and the maximum tax revenue equals the welfare gain of traditional firms:

\[ T = \int_{\gamma^*}^{\hat{\gamma}} \min\{x, \hat{\gamma}\} - \gamma^* dG(x) = (\hat{\gamma} - \gamma^*)\hat{X} + \int_{\gamma^*}^{\hat{\gamma}} x - \gamma^* dG(x) \]  

Graphically, it corresponds to the area of \( F(X) \) between \( \gamma^* \) and \( \hat{\gamma} \) in figure 5. The first part of (23) captures the gain of firms that already operate in the market equilibrium due to a lower bond return, the second the surplus of rather unproductive firms that are not funded in the market equilibrium. The revenue can be used to compensate investors for the decline of their returns; the surplus of traditional firms is then the same as in the market equilibrium. Imposing even higher taxes on traditional firms would reduce their surplus compared to the market equilibrium such that they are worse off. Obviously, the maximum revenue \( T \) falls short of investors’ welfare loss \((\hat{\gamma} - \gamma^*)V\). Consequently, a redistributive tax only mitigates the adverse consequences for investors but not fully compensate them. Hence, regulation is not an efficiency gain according to the Pareto criterion. The very reason why investors cannot be compensated \textit{ex post} is that asymmetric information makes it impossible to tax the welfare gain of bankers in a non-distorting fashion.

### 5 Extensions

This section examines three extensions of the model - social cost of bank failure, a storage technology, and a loan market and entrepreneurial moral hazard - in order to obtain additional insights and policy implications as well as to evaluate the robustness of its results. On the one hand, we find that social cost of bank failure provide an additional rationale for intervention and justify even tighter regulation. On the other hand, we find that the welfare effects of pecuniary externalities are more complex but can still be negative under weaker assumptions.

\[ ^{33} \text{However, one might argue that although the regulator does not observe } R, \text{ he can anticipate bank behavior given capital structure and deposit rate. This information can be used to compute the necessary tax rate.} \]
5.1 Social Cost of Bank Failure

In reality, the banking sector is likely to be characterized by more frictions that require regulation than just pecuniary externalities, for example, bailout and bankruptcy cost, wealth losses of depositors, the loss of borrower-lender relationships or disruptions of the payment system. This section introduces some of these frictions as reduced-form social cost of bank failure.\footnote{A similar approach can be found, for example, in Estrella (2004) or Repullo (2013).} These cost require regulation tighter than in the standard model resulting in even safer but smaller banks.\footnote{Recall the discussion in section 4.4.2: Tighter regulation may reduce the banker’s expected payoff compared to the market equilibrium as soon as the quantity effect dominates, which may explain the bankers’ reluctance to accept regulatory interventions.}

Nevertheless, the interpretation of social cost is somewhat limited as deposits already earn a risk-adjusted, ‘fair’ deposit rate and potential wealth losses are fully internalized. Hence, the social cost rather mirror some additional externalities. More precisely, suppose there are social cost \( s > 0 \) per unit of loan, which reduce social welfare from (19) to:

\[
W = \max_{R,b,L,X,K} F(X) + [p(R)R - (1 - p(R))s] L
\]  

(24)

Hence, the social value of bank lending is lower than the bank’s private value as it includes all potential cost associated with bank failure. The constraints \((IC)\), \((PC)\), and \((CA)\) remain unchanged, and the regulator solves program 3 with a modified objective function. This yields the following proposition:

**Proposition 4** The allocation \( \{R^*, L^*, X^*, K^*, D^*, b^*, \gamma^*\} \) is characterized by (4), (7), (8), (13), \((PC)\), \((CA)\) and

\[
F'(X) = \frac{p(R)R - (1 - p(R))s}{1 + \frac{\lambda p'(R) \kappa}{p(R)L} \left[ 1 - \frac{1 - \kappa}{\kappa} \frac{F''(X)L}{L} \right]}
\]  

(25)

The banking sector is smaller (i.e., \( L^* < L^* \)) and safer (i.e., \( p^* > p^* \)) than in the baseline model.

**Proof:** See appendix A.2

The interpretation of this finding is straightforward: First, social cost of bank failure directly reduce the social marginal return of bank lending as shown in the numerator of (25). Second, the incentive effects associated with bank capital are stronger because it also reduces the probability that the social cost are incurred through its effect on risk shifting (i.e., Lagrange multipliers of \((IC)\) are \textit{ceteris paribus} larger in absolute terms increasing the wedge between marginal returns). Consequently, the regulator allocates fewer resources to the banking sector such that the deposit rate is lower and the profit margin larger. This reinforces the distributional implications for investors and firms summarized in corollary 3 but the effect on the banker’s surplus is now ambiguous because, provided \( s \) is sufficiently large, the negative quantity effect on total bank profits may prevail. This may be one explanation for the observed reluctance of bankers to accept regulation. Then, welfare gains are entirely driven by the lower expected social cost.

5.2 Storage Technology

A drawback of the baseline model is that no endogenous mechanism so far prevents negative returns for investors because they need to allocate their wealth between the two assets as their initial endowment otherwise perishes. Although this model is a real model and small negative real returns may occur even if nominal interest rates are nonnegative, the importance of the zero lower bound motivates the following extension: Investors can store an amount \( Q \); at the end of the
period, they receive a safe payoff of one per unit stored. One may think of holding cash provided that inflation is virtually zero. The main modification is that the bond return, and by the participation constraint, the expected deposit rate never fall below one because investors would then store their entire savings:

$$p(R)b = \gamma \geq 1$$

Traditional firms only invest if their idiosyncratic productivity exceeds the bond return. Combining both conditions, one obtains $$\gamma = \max \{F'(X), 1\}$$ such that investors’ opportunity cost of supplying deposits are either variable and equal to the marginal product of the frictionless technology or constant and equal to one, the value of storage. Traditional investment is limited by an upper bound because firms with an idiosyncratic productivity $$x < 1$$ cannot invest. Hence, we define $$\bar{X}$$, which follows from $$F'(\bar{X}) = 1$$, as the maximum traditional investment. From the resource constraint, it follows that if $$L > 1 - \bar{X} \equiv \bar{L}$$, traditional investment is below the maximum and the bond return larger than one. Thus, investors only hold bonds and deposits and do not use storage. If, in contrast, $$L < \bar{L}$$, traditional investment equals $$\bar{X}$$, the bond return is 1, and investors store $$Q = V - D - \bar{X} = \bar{L} - L > 0$$.

In the first best, the storage technology is never applied because $$p(\hat{R})\hat{R} = F'(\hat{X}) > 1$$ due to assumption 2, and the results derived in the baseline model prevail. In the market equilibrium, the banker’s optimization problem and first-order conditions are unchanged. However, investors’ opportunity cost are restricted by the lower bound. Combining the banker’s first-order condition (16) and the modified participation constraint yields:

$$\max \{F'(X), 1\} = \frac{p(R)R}{1 + \frac{b p'(R) \kappa}{p(R)L}}$$

Again, private marginal returns of both technologies are equalized in equilibrium: If the solution is $$L = \hat{L} > \bar{L}$$, investors’ opportunity cost are determined by the frictionless technology, and storage is not applied. The allocation thus coincides with the market outcome in the baseline model such that $$\hat{L} = \hat{L}$$ (see proposition 2). If the solution is $$\hat{L} \leq \bar{L}$$, the storage technology is the relevant alternative for investors and pins down their opportunity cost. Investors store part of their wealth $$Q > 0$$, which keeps the bond return constant such that the allocation differs from the baseline model. Since the private marginal return of bank lending is generally non-increasing in $$L$$, it follows that banks are then smaller than they would be in the absence of storage.

When determining the constrained social optimum, the regulator now also decides about the amount stored $$Q$$ and faces a different (combined) participation constraint because the opportunity cost are in fact given by $$\max \{F'(X), 1\}$$ as discussed above. Social welfare (i.e., expected payoffs to all agents) equals the combined output of banking and traditional sector plus the value of storage. The regulator solves

**Program 5** The regulator maximizes welfare $$W$$ by choosing target return $$R$$, equity $$K$$, lending $$L$$, traditional investment $$X$$, storage $$Q$$, and the deposit rate $$b$$

$$W = \max_{R, b, L, X, Q, K} \left( Q + F(X) + p(R)RL \right)$$

subject to the banker’s incentive compatibility constraint (IC), the capital availability constraint

---

36 An alternative interpretation consistent with the model is that investors can consume early (i.e., at the beginning of the period). In the absence of discounting, they only save if (net) returns are positive and otherwise immediately consume their entire wealth.

37 Without storage technology, it would be smaller than one.

38 This explains why the banker never stores his wealth even if he could access this technology: According to (16), he chooses bank size such that $$p(R)R \geq p(R)b \geq 1$$. 

26
(CA), the combined participation constraint (CP)

\[ p(R)b = \max\{F'(X), 1\} \]

the aggregate resource constraint (RC)

\[ L + X + Q = 1 \]

Solving this program yields the equilibrium allocation described by

**Proposition 5** The allocation \( \{R^\circ, L^\circ, X^\circ, Q^\circ, K^\circ, D^\circ, b^\circ, \gamma^\circ\} \) is characterized by \( (4), (7), (13), (CP), (RC), (CA) \)

\[
\max\{F'(X), 1\} = p(R)\frac{R}{1 + \frac{\lambda p'(R)\kappa}{p(R)\kappa} \left[ 1 - \mathbf{1}_\{X < \bar{X}\} \frac{(1 - \kappa) F''(X)L}{\kappa} \right]} \quad (28)
\]

and \( Q = \{L - L^\circ, 0\} \). Two cases exist:

- **If** \( L^\circ > \bar{L} \), the constrained social optimum is similar to the baseline model, \( L^\circ = L^* \), it is characterized by \( F'(X) > 1 \) and no storage.
- **If** \( L^\circ \leq \bar{L} \), the constrained social optimum differs from the baseline model, it is characterized by \( F'(X) = 1 \) and \( Q \geq 0 \). It coincides with the market outcome if (28) holds with equality.

**Proof:** See appendix A.2

If banks are large enough and attract sufficient deposits such that only productive traditional firms invest and investors’ net returns are strictly positive, the results of the baseline model prevail because the storage technology is not applied at all. Condition (28) coincides with its baseline counterpart (20). However, if banks are small, many resources are left to traditional firms which are, at the margin, rather unproductive such that investors would suffer from a decline of their returns. They can prevent this by storing part of their savings such that they earn a gross return of one. The allocation then differs from the baseline model and is characterized by \( L^\circ \leq \bar{L}, X^\circ = \bar{X}, Q^\circ = \bar{L} - L^\circ \geq 0 \). Since the social return of bank lending is generally non-increasing in \( L \), banks are smaller than in the absence of storage. Importantly, the allocation chosen by the regulator and the market outcome are similar because (26) and (28) coincide. However, this feature is not necessary in case \( L^\circ = \bar{L} \): As long as the marginal return of lending equals one for \( L = \bar{L} \) such that (28) holds with equality, both allocations are indeed the same. The social marginal return may, however, exceed one at the threshold but fall strictly below one just above. This is due to the discontinuity at \( \bar{L} \) such that (28) may not hold with equality. Bank lending in the market equilibrium, however, is then larger than \( \bar{L} \) such that the two allocations do not coincide.

**Figure 6** illustrates the impact of the storage technology. The first case of proposition 5 is captured in panels (1) and (2): They show market equilibrium and constrained social optimum with banks large enough (i.e., \( L > \bar{L} \)) such that investors earn positive net returns. Hence, storage is not applied, and the baseline model is replicated such that the allocations coincide with their counterparts (e.g., \( \hat{L} \) and \( L^\circ \)). One observes that internalizing the pecuniary externality leads to smaller banks, which mitigates risk shifting. Panel (3) illustrates the second case of proposition 5: Investors store part of their wealth, \( Q = \bar{L} - L^* > 0 \), ensuring a bond return of one. Importantly, market outcome and constrained social optimum coincide. Eventually, panel (4) illustrates the boundary case with bank lending \( L^\circ = \bar{L} \) in the constrained social optimum. In the market
equilibrium (shown in gray), however, banks are larger and of similar size than in the baseline model.

The key insight is that if investors store part of their wealth, they supply deposits elastically such that the opportunity cost are constant and independent of the resource allocation. Therefore, the mechanism of the pecuniary externality breaks down: Raising deposits has - at least marginally - no impact on investors’ opportunity cost such that it does not affect risk shifting of all other banks. This implies:

**Corollary 4** The pecuniary externality is welfare-reducing only if in a market equilibrium \( L > \bar{L} \) and storage is not applied.

**Proof:** Suppose \( \hat{L} \leq \bar{L} \) is the solution of (26) in the market equilibrium. Then, this condition coincides with (26) that determines the constrained social optimum such that both allocations are similar. If the solution of (32) is \( \hat{L} > \bar{L} \), it cannot follow from (28) as well because social and private marginal returns of bank lending differ if \( L \leq \bar{L} \). Thus, the market outcome differs from the welfare-maximizing allocation chosen by the regulator and is constrained-inefficient. Q.E.D.

As soon as \( L \leq \bar{L} \), investors partly rely on storage such that their opportunity cost are constant. The pecuniary externality vanishes and the market outcome is constrained-efficient as it coincides with the allocation chosen by the regulator. Importantly, this implies that regulation - at least in the absence of other frictions - is only necessary if investors earn a strictly positive net return such that they have no need to use storage.

These findings suggest that as soon as investors have the possibility to store their wealth and to protect themselves against negative (net) returns, the pecuniary externality is a relevant, welfare-reducing phenomenon only under some conditions that emerge endogenously. Essentially, the economy needs to be above the zero lower bound in the sense that investors earn positive returns and supply deposits not perfectly elastically. It requires that either the banking sector is large or the frictionless technology is relatively productive. The former ensures that banks attract a large deposit volume such that the remaining resources are small enough to fund only productive traditional firms. The latter makes sure that enough productive traditional firms exist and this sector generates positive net returns even if it needs to absorb a large amount of resources.\(^{39}\)

\(^{39}\)Note that there is a certain tension between these two characteristics as bank size is endogenous and generally
5.3 Entrepreneurial Moral Hazard and Diversification of Risk

This section introduces entrepreneurs and an endogenous loan market, which replaces the shortcut that banks directly invest. Importantly, entrepreneurs determine target return and loan risk reflecting the fact that they ultimately decide about investment and that banks cannot perfectly monitor their activities. This extension is motivated by a recent discussion in the literature: Boyd and De Nicoló (2005) show that the negative relationship between bank competition and financial stability that prominently featured in the literature is reversed as soon as borrowers determine risk. Intuitively, bank competition drives down lending rates and mitigates risk shifting by borrowers. However, such a setting limits the role of banks in risk taking. Therefore, we closely follow the stylized model of Hakenes and Schnabel (2011) that includes both bank and entrepreneurial risk shifting: Whereas a borrower chooses the project’s risk-return profile, the banker decides about the diversification of loans in its portfolio. Bank risk is thus jointly determined by entrepreneurs and banker. The aim of this extension is twofold: First, it explores the role of pecuniary externalities when there is bank and entrepreneurial moral hazard. In particular, it is a priori unclear whether banks are too large or too small in a market equilibrium. Second, the extension provides a robustness check.

5.3.1 Risk Taking and Moral Hazard Reconsidered

The main modification is joint risk taking by entrepreneurs and bankers. Loan risk and portfolio diversification are unobservable, which may lead to moral hazard and risk shifting of both agents as they trade off marginal gains and cost of a higher project return and a better diversified portfolio respectively. Entrepreneurs are modeled as follows: There is a continuum of measure one of risk-neutral, potential entrepreneurs with access to a unit-size project characterized by the banking technology (assumption 1): It yields a return $R$ with probability $p(R)$ and zero else; the success probability decreases in the return. Essentially, this corresponds to bank risk taking in the baseline model. Entrepreneurs have no private wealth but can borrow from banks at a (gross) lending rate $r$. Each entrepreneur faces opportunity cost $\theta$ (e.g., forgone labor income, home production) and maximizes the expected surplus:

$$\pi^E = \max_R p(R)(R - r) - \theta$$

(29)

The choice is characterized by the first-order condition

$$p(R) + p'(R)(R - r) = 0$$

(30)

Obviously, there is entrepreneurial risk shifting: A high lending rate reduces the marginal cost of investing in a more profitable but riskier project and encourages risk taking. Since a normative analysis requires a full-fledged loan demand, this is added in a simple way$^{40}$: Entrepreneurs are heterogeneous as they face opportunity cost of entrepreneurship uniformly distributed on the unit interval, $\theta \sim [0, 1]$. Only entrepreneurs with a positive expected surplus $\pi^E$ invest:

$$\theta \leq p(R)(R - r) \equiv \bar{\theta}(r)$$

Loan demand equals the fraction of investing entrepreneurs, $\bar{\theta}(r)$. It decreases in the lending rate $\bar{\theta}'(r) < 0$. The fraction $1 - \bar{\theta}$ does not activate the project as the outside option is more attractive. The banker’s model is complemented with portfolio diversification such that banks play an active even though loan risk is determined by borrowers. This captures another key feature of banks,

$^{40}$This differs from Hakenes and Schnabel (2011), who assume a reduced-form loan demand.
namely, the diversification of risks. Essentially, it is related to a bank’s business model: Some banks specialize and provide loans to borrowers with rather correlated risks such as regional banks or banks that primarily lend to a specific sector or type of borrower (e.g., mortgage banks). Others diversify and have a portfolio with weakly correlated returns such as universal or international banks. Following Hakenes and Schnabel (2011), we focus on the two extreme cases of perfect correlation and diversification\footnote{Portfolio correlation is either modeled numerically (e.g., Martinez-Miera and Repullo (2010)) or ruled out by assuming perfect correlation (e.g., Holmström and Tirole (1997), Repullo (2013), and Gersbach and Rochet (2012)).}: Loans are either perfectly correlated with probability \( z \) or perfectly uncorrelated with probability \( 1 - z \). Hence, either the loan portfolio as a whole and, thus, the bank succeeds with probability \( p(R) \) and fails otherwise or a fraction \( p(R) \) of loans is repaid with certainty. The probability of bank failure is \( z(1 - p(R)) \). Importantly, the bank is risk-free if its portfolio is perfectly diversified. The banker chooses the probability \( z \), which is associated with a u-shaped utility cost \( C(z) \) proportional to loans and satisfying \( C(z_0) = 0, \quad C''(z) > 0 \) where \( z_0 \in [0, 1] \) reflects the ‘natural correlation’ and any deviation from this level increases the cost. Accordingly, one can interpret \( C(z) \) as diversification cost for \( z < z_0 \) and as specialization cost for \( z > z_0 \).

The contract specifies deposit repayment to investors. As in the baseline model, it satisfies investors’ participation constraint, which requires deposits to be fairly priced:

\[
\left[1 - z(1 - p(R))\right] b = \gamma \quad (31)
\]

Since both loan risk and portfolio diversification are unobservable, the contract is feasible if both actions are privately optimal ex post. Loan risk (i.e., project return \( R \)) is determined by entrepreneurs according to (30). The banker, in turn, decides about portfolio correlation (i.e., their business model) after funding is obtained and aims at maximizing expected profit. This is captured by the incentive compatibility constraint:

\[
z = \arg \max_z [p(R)r - \left[1 - z(1 - p(R))\right] b(1 - \kappa) - C(z)] \quad (32)
\]

The corresponding first-order condition is:

\[
(1 - p(R)) b (1 - \kappa) = C'(z) \quad (33)
\]

Moral hazard leads to overspecialization (\( z > z_0 \)) if the bank is funded by debt. Intuitively, specializing raises the risk of failure and lowers the expected value of the debt burden whereas expected loan return is unaffected. Hence, the banker is even willing to incur higher cost. Overspecialization is stronger in case leverage, deposit rate, and loan risk are high. Consequently, entrepreneurial and bank risk taking are determined by three conditions (30), (33), and (31) for a given lending rate, bond return, and capital structure. They are characterized by:

Lemma 3 \textbf{Target return satisfies} \( \frac{\partial R}{\partial r} > 0 \); \textbf{portfolio diversification and deposit rate satisfy} \( \frac{\partial z}{\partial \gamma} < 0, \quad \frac{\partial z}{\partial \kappa} \geq 0, \) and \( \frac{\partial z}{\partial r} > 0 \) as well as \( \frac{\partial \kappa}{\partial \gamma} < 0, \quad \frac{\partial \kappa}{\partial r} \geq 0, \) and \( \frac{\partial \kappa}{\partial \gamma} > 0 \).

\textbf{Proof:} See appendix A.2.

In contrast to the baseline model, loan risk only depends on the lending rate, which balances loan demand and supply. The finding that portfolio diversification increases in the capital loan ratio but decreases in the bond return is standard. The effect of the lending rate stems from the entrepreneur’s risk shifting response, which lowers the marginal gains from diversification. Consequently, higher lending rates increase both loan risk and portfolio correlation.
5.3.2 Equilibrium Allocation

We first briefly characterize the competitive market equilibrium that largely resembles the baseline model. However, there is one key difference: The project return is now determined by entrepreneurs according to (30). Therefore, bankers only consider their own risk shifting (i.e., overspecialization) problem when choosing bank size and capital structure. Accordingly, loan risk (project return $R$) and portfolio correlation follow from the incentive compatibility constraints of entrepreneur (30) and banker (33). The choices of investors, who allocate their savings between deposits and corporate bonds, and traditional firms, which invest if sufficiently productive, are similar to the baseline model. In equilibrium, bond, deposit, and loan market simultaneously clear; the latter requires $L = \bar{\theta}(r)$. Essentially, the equilibrium is characterized by:

**Lemma 4** In the market equilibrium, the resource allocation is determined by

$$F'(X) = \frac{p(R)R - C(z) - L}{1 - \lambda_1(1 - p(R))K_B}$$

(34)

where $\lambda_1$ denotes the Lagrange multiplier of the banker’s incentive compatibility constraint (33).

**Proof:** See appendix A.2.

This condition requires the private marginal returns of both technologies to be equalized taking into account bank risk shifting. Due to opportunity cost, however, the marginal return of bank lending is lower than in the baseline model.

The regulator’s objective is to maximize social welfare, the expected surplus of all four agents: $W = \pi^B + \pi^E + \pi^F + \pi^I$. In contrast to the standard model, two counteracting pecuniary externalities exist: Banks fail to internalize that (i) raising deposits tightens investors’ participation constraint increasing the deposit rate that leads to overspecialization of all other banks and (ii) their loan supply lowers the lending rate thus alleviating entrepreneurial risk shifting. The externalities are intertwined because each increase in lending requires a similar increase of deposits (unless the bank is fully equity funded). Since the regulator observes neither entrepreneurial nor bank risk taking, he designs a loan and a deposit contract that are feasible and satisfy the incentive compatibility constraints of entrepreneur and banker. However, the regulator cannot freely choose lending and deposit rate because participation of investors, traditional firms, and entrepreneurs needs to be ensured. This requires (expected) returns on deposits and bonds to be equalized (i.e., $(1 - z)(1 - p(R))b = F'(X)$ when using the combined formulation) and the loan volume to be consistent with the mass of entrepreneurs willing to invest (i.e., $\bar{\theta}(r) = L$). The regulator solves

**Program 6** The regulator maximizes social welfare $W$ by choosing project return $R$, diversification $z$, bank lending $L$, traditional investment $X$, equity $K$, deposit and lending rate $b$ and $r$:

$$W = \max_{R,z,b,r,L,X,K} F(X) + [p(R)R - C(z)]L - \int_0^L \theta d\theta$$

(35)

subject to incentive compatibility constraints of banker ($IC^B$)

$$z = \arg\max_z [p(R)r - [1 - z(1 - p(R))]b(1 - \kappa) - C(z)]L$$

and entrepreneur ($IC^E$)

$$R = \arg\max_R p(R)(R - r)$$

---

42 A more extensive derivation of the market equilibrium can be found in appendix A.2.

43 This effect is captured by the numerator which uses $r = R - \theta/p$. 

31
the combined participation constraint (CP)

\[ [1 - z(1 - p(R))] b = F'(X) \]

and entrepreneurs’ loan demand (LD)

\[ p(R - r) = L \]

the aggregate resource constraint (RC) and the capital availability constraint (CA).

Solving program 6 yields the constrained social optimum summarized in Proposition 6

The equilibrium allocation \{R', z', L', X', K', r', b', \gamma'\} is characterized by (4), (8), (30), (33), (CP), (LD), (RC), (CA), and

\[
F'(X) = \frac{p(R)R - C(z) - L}{1 - \frac{\lambda_1(1 - p(R)) \kappa}{(1 - PD)L} \left[ 1 - \frac{(1 - \kappa) \int F''(X)L}{\kappa} \right] - \frac{\lambda_2 p'(R) r}{(1 - PD)p(R)b}} \tag{36}
\]

where \( \lambda_1 < 0 \) and \( \lambda_2 < 0 \) are the Lagrange multipliers of (IC\text{E}) and (IC\text{B}) and \( PD = 1 - z(1 - p(R)) \) is a shortcut for the probability of bank failure. The pecuniary externalities are negative overall if

\[
-\frac{\lambda_1(1 - p(R)) b}{\varepsilon_D} + \frac{\lambda_2 p'(R) r}{\varepsilon_L} > 0 \tag{37}
\]

where \( \varepsilon_D > 0 \) and \( \varepsilon_L < 0 \) are the interest rate elasticities of deposit supply and loan demand.

Proof: See appendix A.2.

The key condition is (36); it requires that the resource allocation equals social marginal returns of both technologies taking into account all effects on incentives. In contrast to the private return, which is a critical determinant of the market outcome in (34), the social return includes the two counteracting pecuniary externalities as explained above. They are captured by the third and fourth term of the denominator of (37): The externality associated with the deposit rate is negative and lowers the social return compared to the private; the externality associated with the lending rate is positive and has an opposite effect. Consequently, if there is only bank moral hazard (i.e., \( \lambda_1 = 0 \) and \( R \) is first-best), raising deposits exacerbates risk shifting and leads to inefficient overspecialization.\(^{44}\) If there is only entrepreneurial moral hazard (i.e., \( \lambda_2 = 0 \) and \( z \) is first-best) such as in Boyd and De Nicoló (2005), a large loan supply discourages risk shifting and is welfare-improving. In a framework characterized by both entrepreneurial and bank moral hazard, it is a priori ambiguous whether bank lending is associated with an overall positive or negative externality on risk taking and welfare. This ambiguity is consistent with the positive findings of Hakenes and Schnabel (2011), who study the effect of capital requirements on risk taking.

Whether the social or the private marginal return of bank lending is larger or, equivalently, the pecuniary externalities are positive or negative overall has far-reaching consequences: In one case, banks are generically too large and have too much leverage, in the other case, they are too small. Thus, the regulator’s response fundamentally differs. Which of the two pecuniary externalities eventually prevails is shown in (37): It depends on (i) the interest rate elasticities of loan demand and deposit supply\(^{45}\) (i.e., inelastic supply and demand are associated with large price effects) and (ii) the severity of the agency problems of banker and entrepreneur (captured by the Lagrange

\(^{44}\)This essentially reproduces the result of the baseline model with a different risk variable.

\(^{45}\)The elasticity of deposit supply is expressed in terms of opportunity cost (safe bond return \( \gamma \)) in order to disentangle scarcity and risk, which are reflected by the deposit rate. The elasticity is positive as higher bond returns discourage traditional investment and investors supply more deposits.
multipliers). Hence, the overall welfare effect of the pecuniary externalities is negative if loan demand is rather elastic and entrepreneurial moral hazard of minor importance compared to deposit supply and bank moral hazard respectively. If (37) holds, the analysis leads to similar conclusions as the baseline model, in particular, the need for smaller banks funded by a larger share of equity. The key insight of this extension is a more complex relationship between bank lending and capital structure on the one hand and risk taking on the other hand if there is also entrepreneurial moral hazard: Pecuniary externalities associated with loans and deposits influence various interest rates and have counteracting effects. Therefore, it is not a priori clear whether banks are generically too large and a have too much leverage or whether the reverse is true. Unless loan demand is very inelastic and the entrepreneur’s agency problem severe, however, the results of proposition 3 prevail, and internalizing the pecuniary externalities reduces bank size and leverage.

6 Conclusion

This paper studies moral hazard and bank risk taking and examines to what extent it provides a rationale for macroprudential regulation. It develops a static model of banking with an endogenous allocation of resources between two sectors of the economy - banking and traditional sector - as well as a full-fledged supply of deposits. The main task of banks is risk taking that is subject to moral hazard. The innovation is that this approach enables us to identify a welfare-reducing pecuniary externality, which results from the combination of moral hazard, competition for deposits, and a deposit supply that is not perfectly elastic. Moreover, we provide an extensive discussion of various regulatory approaches to internalize the externality as well as its distributional consequences. The analysis yields six key insights: (i) In the presence of moral hazard, debt distorts risk taking and leads to the well-known risk shifting problem. Discipline can be provided by equity and low opportunity (i.e., funding) cost as they directly or through the effect on the profit margin raise the banker’s ‘skin in the game’ and increase his private cost of risk taking. (ii) The market outcome is constrained-inefficient because of a pecuniary externality: Banks fail to internalize that raising deposits eventually increases the deposit rate thereby exacerbating risk shifting of all other banks. As a result, banks are inefficiently large, have too much leverage, earn low profits, and take excessive risk. This inefficiency results from the combination of a standard agency problem and competition for deposits that are not supplied perfectly elastically. It provides a strong rationale for macroprudential regulation as it does not require the presence of (typical) frictions such as contagion, bailout cost, and incorrectly priced deposit insurance. (iii) The key idea of regulation is to reward prudent banks by allowing them to earn a rent in case of success; this makes risk taking privately costly and less attractive. (iv) Regulation is not a Pareto-improvement as the necessary reallocation of resources to the traditional sector implies a decline of investors’ returns that favors banks and firms in the traditional sector. Redistributive ex post taxation cannot fully offset this effect since raising revenue from banks would further distort their incentives. (v) The pecuniary externality can be internalized using different instruments: prudential regulation (capital requirements or deposit rate ceilings) or imperfect competition where a discrete number of banks competes for deposits in a Cournot fashion. Under imperfect competition, banks do not internalize the pecuniary externality but strategic interaction incidentally achieves the constrained social optimum if the number of banking licenses is appropriately chosen. Accordingly, if the banking sector is more concentrated, liberalization may improve social welfare; if it is more competitive, restraining competition has a positive welfare effect. (vi) If investors can protect themselves against negative net returns using a storage technology, the pecuniary externality is only welfare reducing if the economy is above the zero lower bound and storage is not applied. This requires a large banking sector or a relatively productive frictionless technology. Otherwise, the deposit supply is perfectly
elastic as investors’ opportunity cost are constant and independent of the resource allocation. If an entrepreneur borrows and determines loan risk while the bank chooses portfolio diversification as in Hakenes and Schnabel (2011), two counteracting pecuniary externalities associated with lending and deposit rate exist. Whether they are welfare-reducing overall crucially depends on the severity of banker’s and entrepreneur’s moral hazard problem and on the elasticities of loan demand and deposit supply.

References


A Appendix

A.1 List of Notations

<table>
<thead>
<tr>
<th>VARIABLES</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$b$</td>
<td>Risk-Adjusted Deposit Rate</td>
</tr>
<tr>
<td>$D$</td>
<td>Deposits</td>
</tr>
<tr>
<td>$K$</td>
<td>Bank Capital, Equity</td>
</tr>
<tr>
<td>$L$</td>
<td>Loans, Bank Lending</td>
</tr>
<tr>
<td>$N$</td>
<td>Number of Banks (Oligopolistic Case)</td>
</tr>
<tr>
<td>$Q$</td>
<td>Savings stored</td>
</tr>
<tr>
<td>$R$</td>
<td>Loan Return, Target Return</td>
</tr>
<tr>
<td>$r$</td>
<td>Lending Rate</td>
</tr>
<tr>
<td>$T$</td>
<td>Tax Revenue</td>
</tr>
<tr>
<td>$X$</td>
<td>Traditional Investment (Bonds)</td>
</tr>
<tr>
<td>$x$</td>
<td>Idiosyncratic Productivity of a Traditional Firm, $x \sim G(x)$</td>
</tr>
<tr>
<td>$z$</td>
<td>Diversification, Correlation of Loans</td>
</tr>
<tr>
<td>$\gamma$</td>
<td>Bond Return</td>
</tr>
<tr>
<td>$\kappa$</td>
<td>Capital Loan Ratio</td>
</tr>
<tr>
<td>$\theta$</td>
<td>Entrepreneur’s Opportunity Cost, $\theta \sim [0,1]$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>PARAMETERS</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$K_B$</td>
<td>Banker’s Private Wealth</td>
</tr>
<tr>
<td>$\Pi$</td>
<td>Maximum Loan Return</td>
</tr>
<tr>
<td>$\Pi^*$</td>
<td>Minimum Loan Return</td>
</tr>
<tr>
<td>$s$</td>
<td>Social Cost of Bank Failure</td>
</tr>
<tr>
<td>$V$</td>
<td>Investor’s Savings, $V = 1 - K_B$</td>
</tr>
<tr>
<td>$z_0$</td>
<td>'Natural’ loan correlation</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>DEFINITIONS</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$W$</td>
<td>Social Welfare</td>
</tr>
<tr>
<td>$\pi^b$</td>
<td>Bank’s Expected Profit</td>
</tr>
<tr>
<td>$\pi^e$</td>
<td>Entrepreneur’s Expected Profit</td>
</tr>
<tr>
<td>$\pi^r$</td>
<td>Traditional Firms’ Profit</td>
</tr>
<tr>
<td>$\pi^f$</td>
<td>Investors’ Expected Income</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>FUNCTIONS</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$C(z)$</td>
<td>Diversification and Specialization Cost</td>
</tr>
<tr>
<td>$F(X)$</td>
<td>Aggregate Production Function of Traditional Sector</td>
</tr>
<tr>
<td>$G(x)$</td>
<td>Distribution of Traditional Firms</td>
</tr>
<tr>
<td>$p(R)$</td>
<td>Loan’s Success Probability</td>
</tr>
</tbody>
</table>

A.2 Proofs and Derivations

Proof of Proposition 1 The independence of $R$ from capital structure and size immediately follows from (10). The zero profit margin is implied by (11) and $F'(X) = p(R)b$ from (4) and (6). The neutrality of the allocation w.r.t. the wealth distribution follows from (11), which shows that the bank’s size $L$ only depends on technological characteristics. The independence of social welfare from the supply of bank capital can be shown by an Envelope argument: $\frac{\partial W}{\partial K_B} = 0$. Eventually, we show that this allocation can be implemented as a market equilibrium: The banker chooses $R$, 

36
subject to the investors’ participation constraint (6) and the capital availability constraint $K \leq K_B$. The corresponding first-order conditions are (6), (10), and

$$p(R)b = p(R)R$$
$$p(R)b = \gamma + \zeta$$

where $\zeta$ denotes the Lagrange multiplier of the capital availability constraint. The bank expands until it earns zero expected profit (i.e., $b = R$); the participation constraint pins down the risk-adjusted deposit rate $b = \frac{\gamma}{p(R)}$. Therefore, $\zeta = 0$ and $K_B \leq K$. Traditional firms only invest if sufficiently productive such that $F'(X) = \gamma$ as implied by (4). Furthermore, bond and deposit markets clear, and (7) and (8) hold. Combining these results yields (11) such that the competitive market equilibrium coincides with the first-best allocation. Q.E.D.

Proof of Lemma 1 Target return and deposit rate are jointly determined by the constraints (6) and (13). One may derive the Jacobian matrix:

$$J = \begin{bmatrix} 2p'(R) + p''(R)(R - b(1 - \kappa)) & -p'(R)(1 - \kappa) \\ p'(R)b & p(R) \end{bmatrix}$$

The Jacobian determinant is

$$|J| = (2p'(R) + p''(R)(R - b(1 - \kappa)))p + p'(R)^2b(1 - \kappa) < 0$$

(38)

The negative sign can be shown graphically: In figure 3, the IC- is steeper than the PC-curve in the intersection point, which is necessary as the former is a concave and the latter a convex function. The slopes can be related to the Jacobian. Thus, whenever a feasible contract exists (i.e., $\kappa < \kappa_0$), property (38) holds. Using Cramer’s rule, the sensitivities are:

$$\frac{\partial R}{\partial \kappa} = -\frac{p'(R)p(R)b}{|J|} < 0$$
$$\frac{\partial R}{\partial \gamma} = \frac{p'(R)(1 - \kappa)}{|J|} \geq 0$$
$$\frac{\partial b}{\partial \kappa} = \frac{p'(R)^2b}{|J|} < 0$$
$$\frac{\partial b}{\partial \gamma} = \frac{2p'(R) + p''(R)(R - b(1 - \kappa))}{|J|} > 0$$

Q.E.D.

Proof of Proposition 2 Replacing (IC) by the first-order condition (13), one may rewrite program 2 as a Lagrangian with three constraints and the multipliers $\lambda, \mu,$ and $\zeta$:

$$\mathcal{L} = p(R)[R - b(1 - \kappa)]L + \gamma(K_B - K) + \lambda[p(R) + p'(R)(R - b(1 - \kappa))] + \mu[p(R)b - \gamma] + \zeta[K_B - K]$$

The first-order conditions are

$$\frac{\partial \mathcal{L}}{\partial \lambda} = p(R)(R - b) - \frac{\lambda p'(R)bK}{L^2} = 0$$
$$\frac{\partial \mathcal{L}}{\partial \mu} = -p(R)(1 - \kappa)L - \lambda p'(R)(1 - \kappa) + \mu p(R) = 0$$
$$\frac{\partial \mathcal{L}}{\partial R} = \lambda[2p'(R) + p''(R)(R - b(1 - \kappa)] + \mu p'(R)b = 0$$

37
\[
\frac{\partial L}{\partial K} = p(R)b - \gamma + \frac{\lambda p'(R)b}{L} - \zeta = 0
\]

as well as the constraints (13), (PC), and (CA) and complementary slackness \(\zeta(K_B - K) = 0\). The first condition equals (16), which, together with (PC) and (4), yields (17). Solving the first-order conditions using (PC) yields the Lagrange multipliers:

\[
\begin{align*}
\lambda &= -\frac{p'(R)p(R)b(1 - \kappa)L}{[2p' + p''(R - b(1 - \kappa))]p + p'^2b(1 - \kappa)} \\
\mu &= \frac{[2p' + p''(R - b(1 - \kappa))]p(1 - \kappa)L}{[2p' + p''(R - b(1 - \kappa))]p + p'^2b(1 - \kappa)} \\
\zeta &= \frac{\lambda p'(R)b}{L}
\end{align*}
\]

Note that \(\lambda \leq 0, \mu \geq 0, \text{ and } \zeta \geq 0\) due to property (38); if \(L\) reaches its upper bound consistent with a feasible deposit contract (i.e., \(\kappa = \kappa_0\)), \(\lambda\) diverges to infinity.\(^{46}\) From the first-order condition (16), one can conclude that the banker never expands beyond \(L = \frac{K_B}{\kappa_0}\) because the marginal cost are prohibitively high. Thus, a feasible deposit contract exists in equilibrium. Complementary slackness requires

\[\zeta(K_B - K) = 0\]

When substituting for \(\lambda\), it follows that \(\zeta\) is an implicit function of the capital loan ratio \(\kappa = \frac{K_B}{L}\), with \(\zeta = 0\) if \(\kappa = 1\) and \(\zeta > 0\) else. Therefore, only two solutions are consistent with complementary slackness:

\[K = \{K_B, L\}\]

If \(K = L, \lambda = 0\) and the first-order conditions (17) and (11) coincide. Thus, the first best summarized in proposition 1 is restored. This can only be true if \(L \leq K_B\). Using the resource constraint, this implies \(\exists X \geq 1 - K_B : F'(X) = p(\hat{R})\hat{R}\) where \(R\) is given by condition (10). Substituting the \(F'(X) = G^{-1}(1 - X)\) from (4), yields condition (18). If \(K = K_B, \lambda < 0\) and the first-order condition (17) differs from (11) in the first best. It follows that \(F'(X) = p(R)b < p(R)\hat{R}\) such that there is a positive lending spread. Note that the banker always chooses lending \(L\) as to ensure a feasible deposit contract.\(^{47}\) In addition, \(\hat{R} \geq \hat{R}\) immediately follows from (13). Given this, comparing (17) to (11) implies \(F'(\hat{X}) < p(\hat{R})\hat{R} = F'(\hat{\hat{X}})\). Since \(F(X)\) is concave, it follows that \(\hat{X} > \hat{\hat{X}}\) and from the market clearing conditions (7) and (8) \(\hat{L} < \hat{\hat{L}}\). Eventually, the deadweight loss, \(\Delta W = \bar{W} - \hat{W}\), is:

\[
\Delta W = p(\hat{R})\hat{R}\hat{L} - p(\hat{\hat{R}})\hat{\hat{R}}\hat{\hat{L}} + F(\hat{X}) - F(\hat{\hat{X}})
\]

Expanding by \(p(\hat{R})\hat{R}\hat{L}\) and using \(p(\hat{R})\hat{R} = F'(\hat{\hat{X}})\) as well as \(L = 1 - X\), one obtains:

\[
\Delta W = [p(\hat{R})\hat{R} - p(\hat{\hat{R}})\hat{\hat{R}}\hat{\hat{L}} + F'(\hat{\hat{X}})(\hat{\hat{X}} - \hat{X})] + F(\hat{X}) - F(\hat{\hat{X}})
\]

The first term in square brackets, measuring the quality deterioration due to higher portfolio risk, is unambiguously positive. The second and third term reflect the distortion of the bank’s size resulting in smaller investment in the banking sector (captured by the second term, positive) but in larger investment in the traditional sector (captured by the third term, negative). The concavity of \(F(X)\) ensures that the combined size effect is positive overall such that there is a positive deadweight loss because of moral hazard. \(Q.E.D.\)

\(^{46}\)This is due to (38) which then holds with equality.

\(^{47}\)If \(L\) exceeds the upper bound, the Lagrange multiplier is infinitely large such that the marginal return of bank lending drops to zero and is inconsistent with (17).
Proof of Proposition 3 Replacing (IC) by the first-order condition (13), one can rewrite program 3 as a Lagrangian with four constraints and the multipliers $\lambda$, $\mu$, $\eta$, and $\zeta$:

$$
\mathcal{L} = F(X) + p(R)RL + \lambda [p(R) + p'(R)(R - b(1 - \kappa))]
+ \mu [p(R)b - F'(X)] + \eta [1 - L - X] + \zeta [K_B - K]
$$

The first-order conditions are

$$
\begin{align*}
\frac{\partial \mathcal{L}}{\partial L} &= p(R)R - \frac{\lambda p'(R)bK}{L^2} - \eta = 0 \\
\frac{\partial \mathcal{L}}{\partial X} &= F'(X) - \mu F''(X) - \eta = 0 \\
\frac{\partial \mathcal{L}}{\partial b} &= -\lambda p'(R)(1 - \kappa) + \mu p(R) = 0 \\
\frac{\partial \mathcal{L}}{\partial R} &= [p + p'(R)R]L + \lambda [2p' + p''(R)(R - b(1 - \kappa))] + \mu p'(R)b = 0 \\
\frac{\partial \mathcal{L}}{\partial K} &= \frac{\lambda p'(R)b}{L} - \zeta = 0
\end{align*}
$$

as well as the constraints (13), (CP), (RC), and (CA). From these conditions one obtains the Lagrange multipliers:

$$
\eta = F'(X) - \mu F''(X) > 0
$$

$$
\lambda = \frac{p'(R)p(R)b(1 - \kappa)L}{2p' + p''(R - b(1 - \kappa))[p + p^2b(1 - \kappa)]} < 0
$$

$$
\mu = \frac{p^2b(1 - \kappa)^2L}{2p' + p''(R - b(1 - \kappa))[p + p^2b(1 - \kappa)]} > 0
$$

$$
\zeta = \frac{\lambda p'(R)b}{L} \geq 0
$$

The signs again follow from property (38), which implies that the denominator is negative. Note that the Lagrange multipliers of (IC) and (CA) are the same functions as in the market equilibrium. Combining the first-order conditions w.r.t. $L$ and $X$ and substituting for $\eta$ and $\mu$ yields condition (20). In addition, deposits $D$ and bond return $\gamma$ follow from (4) and (8).

Complementary slackness again implies $K = \{K_B, L\}$. If $K_B \geq G[p(\tilde{R})\tilde{R}]$, the regulator replicates the first best. If this condition is not met, bankers' wealth is completely invested as inside equity ($K^* = K_B$). The right-hand side denominator in (20) is then strictly larger than in (17) due to the presence of the third, positive term, which implies that social returns are smaller than private marginal returns of banking. As a result, the same resource allocation in both equilibria (i.e., $L^* = \tilde{L}$) is impossible. We prove $L^* < \tilde{L}$ by contradiction: From $L^* > \tilde{L}$, it follows that $F'(X^*) > F'(\tilde{X})$.

Using the first-order conditions (17) and (20) yields

$$
\frac{1 + \varphi}{1 + \varphi^* \left[1 - \frac{\kappa^*}{\varphi^*} \frac{F''(X^*)L^*}{\kappa^*} \right]} > \frac{p(\tilde{R})\tilde{R}}{p(R^*)R^*}
$$

(39)

where $\varphi = \frac{\lambda p'(R)K_B}{p(R)L^2}$. The right-hand side of the inequality is larger than one as $L^* > \tilde{L}$ implies $\kappa^* < \kappa$ and $b^* > b$ such that $R^* > \tilde{R}$. Condition (39) requires $\varphi > \varphi^* \left[1 - \frac{\kappa^*}{\varphi^*} \frac{F''(X^*)L^*}{\kappa^*} \right]$. Since we assume $L^* > \tilde{L}$, this is only possible if $\varphi$ is an decreasing function of $L$ such that $\varphi > \varphi^*$. However, it can be shown that $\varphi$ is generally non-decreasing in $L$; it is zero for $L \leq K_B$ and diverges to infinity if $L$ is close to the upper bound consistent with a feasible deposit contract. Consequently, condition (39) does not hold for $L^* > \tilde{L}$, and only allocations with $L^* < \tilde{L}$ are possible. Q.E.D.
Proof of Corollary 1 To show that the constrained social optimum can be implemented by capital requirements, we add the regulatory constraint $\kappa \geq \kappa^*$ to program 2. The Lagrangian is:

$$L = p(R)[R - b(1 - \kappa)] + \gamma (K_B - K) + \lambda [p(R) + p'(R)(R - b(1 - \kappa))] + \mu[p(R)b - \gamma] + \psi[K_B - K] + \psi[\kappa - \kappa^*]$$

The Kuhn-Tucker conditions are

$$\frac{\partial L}{\partial L} = p(R)(R - b) - \frac{\lambda p'(R)bK}{L^2} + \frac{\psi K}{L^2} = 0$$
$$\frac{\partial L}{\partial K} = p(R)b - \gamma + \frac{\lambda p'(R)b}{L} - \zeta + \frac{\psi}{L} = 0$$
$$\frac{\partial L}{\partial \psi} = \kappa - \kappa^* \geq 0$$
$$\frac{\partial L}{\partial \psi} \psi = \psi(\kappa - \kappa^*) = 0$$

The other conditions are similar to the market equilibrium (see, proof of proposition 2). If the regulatory constraint does not bind, $\kappa > \kappa^*$, complementary slackness requires $\psi = 0$ and the first-order conditions coincide with those of the unregulated market equilibrium. Critically, the solution is then $\kappa = \hat{\kappa}$. Since $\kappa^* > \hat{\kappa}$, however, this outcome is impossible. Hence, the banker always chooses the boundary solution $\kappa = \kappa^*$. Since bank capital is scarce, (CA) binds such that $L = L^*$ and $X = X^*$ immediately follow. The bond return $\gamma^*$ follows from this allocation and leads, together with $\kappa^*$, to the constrained-efficient risk-return profile.

If the regulator imposes a deposit rate ceiling $b \leq b^*$, we replace the regulatory constraint in the banker’s optimization problem. Using a similar argument as above, one can show that the banker always chooses $b = b^*$. From figure 3, one may conclude that several pairs $\{\kappa, \gamma\}$ satisfy (IC) and (PC), once $b$ equals $b^*$. However, only $\kappa^*$ and $\gamma^*$ are consistent with resource constraint and the choice of traditional firms. If the banker chooses a slightly larger $\kappa$, incentive compatibility implies a smaller target return (i.e., the IC-curve in figure 3 shifts to the left) but the bond return eventually falls due to larger traditional investment and the participation constraint thus allows for a larger target return (i.e., the PC-curve shifts to the right). As a result, only a deposit rate smaller than $b^*$ reconciles incentives and fair pricing. However, the banker would never choose such a deposit rate because if the regulatory constraint is slack, his optimal choice would be $b > b^*$. 

Q.E.D.

Proof of Lemma 2 Bank $i$ solves program 6; the Lagrangian is:

$$L = p(R_i)[R_i - b_i(1 - \kappa_i)] + \gamma(K_{Bi} - K_i) + \lambda_i[p(R_i) + p'(R_i)(R_i - b_i(1 - \kappa_i))]$$
$$+ \mu_i[p(R_i)b_i - F'(1 - \sum_{i=1}^N L_i)] + \zeta_i[K_{Bi} - K_i]$$

The corresponding first-order conditions are

$$\frac{\partial L}{\partial L_i} = p(R_i)(R_i - b_i) - \frac{\lambda_i p'(R_i)b_iK_i}{L_i^2} + \mu_i F''(X) = 0$$
$$\frac{\partial L}{\partial b_i} = -p(R_i)(1 - \kappa_i)L_i - \lambda_i p'(R_i)(1 - \kappa_i) + \mu_i p(R_i) = 0$$
$$\frac{\partial L}{\partial R_i} = \lambda_i[2p'(R_i) + p''(R_i)(R_i - b_i(1 - \kappa_i))] + \mu_i p'(R_i)b_i = 0$$
$$\frac{\partial L}{\partial K_i} = p(R_i)b_i - \gamma + \frac{\lambda p'(R_i)b_i}{L_i} - \zeta_i = 0$$

40
as well as the constraints (13), (PC), and (CA). Since \( \zeta_i > 0 \), the capital availability constraint again binds such that \( K_i = K_B \). Rearranging the first-order condition w.r.t. \( L_i \) and substituting for \( \mu_i \) as well as using (PC) yields

\[
F'(X) = \frac{1 + \lambda p'(R_i) \kappa_i}{p(R_i) L_i} \left[ \frac{1 - (1 - \kappa_i) L_i F''(X)}{p(R_i)b_i} - \frac{(1 - \kappa_i)L_i F''(X)}{p(R_i)b_i} \right]
\]

which describes the optimal choice of bank \( i \). The Lagrange multipliers are determined as under perfect competition. In the symmetric Nash equilibrium, obviously \( R_i = R \) and, by (PC), \( b_i = b \).

In addition, \( L_i = \frac{L}{N} \) where \( L \) denotes the aggregate amount of loans (i.e., investment in banking sector). \( K_B = \frac{K}{N} \) and \( \kappa_i = \frac{K}{L} = \kappa \) then imply \( \lambda_i = \frac{\lambda}{N} \) and \( \mu_i = \frac{\mu}{N} \). Substituting this into (40) yields (21). For \( N \to \infty \) the third and fourth term in the denominator vanish, such that the allocation converges to the perfectly competitive equilibrium. \( \text{Q.E.D.} \)

Proof of Proposition 4

The optimization problem is similar to program 3 but with the objective function (24) instead of (19). Hence, the first-order conditions are the same as in the baseline model (see proof of proposition 3) but two of them differ:

\[
\frac{\partial L}{\partial L} = [p(R)R - (1 - p(R))s] - \frac{\lambda p'(R)bK}{L^2} - \eta = 0
\]

\[
\frac{\partial L}{\partial R} = [p + p'(R)(R + s)L + \lambda [2p'(R) + p''(R)(R - b(1 - \kappa))] + \mu p'(R)b = 0
\]

Using the above first-order conditions, it can be shown that

\[
\lambda' = -\frac{p'(R)[b(1 - \kappa) + s)p(R)L}{2[p'(R) + p''(R)(R - b(1 - \kappa))]}p + p^2b(1 - \kappa) < \lambda^*
\]

for all feasible values of \( R, b, \) and \( L \). Rearranging the first-order conditions yields (25). Compared to (20), the numerator is smaller and two terms in the denominator are larger such that the social marginal return of lending is smaller. Since the traditional sector earns an excess return if \( L = L^* \), the regulator allocates more resources to this sector. Hence, \( L' < L^* \). \( \text{Q.E.D.} \)

Proof of Proposition 5

Solving program 5 generally yields similar first-order conditions as in the baseline model (see proof of proposition 3). However, the first-order condition regarding \( X \) differs and there is an additional choice variable \( Q \):

\[
\frac{\partial L}{\partial X} = F'(X) - \mathbb{1}_{(X < \overline{X})}\mu F''(X) - \eta = 0
\]

\[
\frac{\partial L}{\partial Q} = 1 - \eta \leq 0
\]

If \( X < \overline{X} \), by definition \( F'(X) > 1 \) and \( \eta = F'(X) - \mu F''(X) > 1 \) such that \( \frac{\partial L}{\partial Q} < 0 \) and \( Q = 0 \).

Hence, the first-order conditions are equivalent to those of program 3 and the allocation coincides with the baseline. If \( X \geq \overline{X} \), by definition \( F'(X) \leq 1 \). Since \( F'(X) = \eta \) and \( \eta \geq 1 \), only \( X = \overline{X} \) is a solution of both conditions and \( Q = 1 - L^* - \overline{X} \geq 0 \). Condition (28) is obtained by eliminating \( \eta \) from the first-order condition w.r.t \( L \) using the two conditions above and by substituting for \( p(R)b \) using (CP). From (28), the resource allocation is determined by:

\[
1 = \frac{p(R)R}{1 + \frac{2p'(R)b}{p(R)L}} \quad \text{if } L \leq \overline{L}, \quad F'(X) = \frac{p(R)R}{1 + \frac{2p'(R)b}{p(R)L}} \left[ \frac{1 - (1 - \kappa)L F''(X) + \lambda p'(R)b}{\kappa} \right] \quad \text{if } L > \overline{L}
\]

The corresponding participation constraints are \( p(R)b = 1 \) and \( p(R)b = F'(X) > 1 \) respectively.
For high values of $L$, the social marginal return is similar to the baseline model, for low values, it coincides with the private return. Based on our considerations for the baseline model, the social marginal return is non-increasing in $L$. In addition, there is a discontinuous jump at $L = \mathcal{L}$. This implies that (i) whenever there exists an equilibrium with $L < \mathcal{L}$, there is none with $L > \mathcal{L}$ because, in this domain, the social marginal return is strictly below one; and (ii) whenever there exists an equilibrium with $L > \mathcal{L}$, there is none with $L < \mathcal{L}$ because the social marginal return is then larger than one. One needs to distinguish between three solutions implied by proposition 5: First, if the solution includes $\mathcal{L}^\circ > \mathcal{L}$, it automatically satisfies the equilibrium conditions of the constrained social optimum in the baseline model because (28) and (20) coincide. Second, if it implies $\mathcal{L}^\circ < \mathcal{L}$, the equilibrium involves storage and differs from the baseline model because conditions (28) and (20) are not the same. However, condition (28) equals (26) that characterizes the market equilibria such that these two allocations are the same. Third, if the solution is $\mathcal{L}^\circ = \mathcal{L}$, one further distinguishes between two cases: First, condition (28) has an exact solution and holds with equality. Then, it coincides with the market outcome. Second, (28) has no exact solution at all and the social marginal return is larger than one at $\mathcal{L}$ but smaller just above. The constrained social optimum is thus a boundary solution which differs from the market equilibrium because the private return does not exhibit such a jump such that it intersects $F'(X)$ at $L > \hat{L}$. One concludes that the constrained social optimum equals its baseline counterpart if $\mathcal{L}^\circ > \mathcal{L}$ but differs if $\mathcal{L}^\circ \leq \mathcal{L}$ and (ii) that constrained social optimum and market outcome coincide if $\mathcal{L}^\circ < \mathcal{L}$ or if $\mathcal{L}^\circ = \mathcal{L}$ is an exact solution of (28). Q.E.D.

**Proof of Lemma 3** The target return is determined by the entrepreneur’s first-order condition (30), which can be differentiated:

$$\frac{\partial R}{\partial r} = \frac{p'(R)}{2p' + p''(R - r)} > 0$$

As the banker’s diversification choice and the deposit rate are characterized by conditions (33) and (31), we derive the Jacobian matrix:

$$J = \begin{bmatrix} -C''(z) & (1 - p(R))(1 - \kappa) \\ -(1 - p(R))b & 1 - z(1 - p(R)) \end{bmatrix}$$

The Jacobian determinant is

$$|J| = -C''(z)[1 - z(1 - p(R))]p + (1 - p(R))^2b(1 - \kappa) < 0 \quad (41)$$

By similar considerations as in the baseline, it can be shown that the Jacobian determinant is negative.\textsuperscript{48} Using Cramer’s rule, the sensitivities are:

$$\frac{\partial z}{\partial \kappa} = \frac{(1 - p(R))[1 - z(1 - p(R))]b}{|J|} < 0 \quad \frac{\partial b}{\partial \kappa} = \frac{(1 - p(R))^2b^2}{|J|} < 0$$

$$\frac{\partial z}{\partial \gamma} = \frac{(1 - p)(1 - \kappa)}{|J|} \geq 0 \quad \frac{\partial b}{\partial \gamma} = -\frac{C''(z)}{|J|} > 0$$

$$\frac{\partial z}{\partial r} = \frac{p'(R)b(1 - \kappa)}{|J|} \frac{\partial R}{\partial r} \geq 0 \quad \frac{\partial b}{\partial r} = \frac{p'(R)b[C'(z) + C''(z)z]}{|J|} \frac{\partial R}{\partial r} > 0$$

Q.E.D.

\textsuperscript{48}The deposit rate implied by (PC) is finite while the deposit rate implied by (IC) can be infinitely large if $z \to 1$. Given that the former exceeds the latter for all $z \leq z_0$, the IC-curve needs to be steeper than the PC-curve at the intersection point. The slopes are related to $|J|$. 

42
Proof of Lemma 4 The entrepreneur chooses project return $R$ according to (30). Following Hakenes and Schnabel (2011), the banker maximizes expected profit subject to (IC), (PC), and (CA) with the corresponding multipliers $\lambda_1$, $\lambda_2$, $\mu_1$, and $\eta$.

$$\mathcal{L} = [p(R)r - [1 - z(1 - p(R))]b(1 - \kappa) - C(z)] L + \gamma(K_B - K) + \lambda_1 [(1 - p(R))b(1 - \kappa) - C'(z)] + \lambda_2 [(p(R) + p'(R)(R - r)] + \mu_1[(1 - z(1 - p(R))b - \gamma] + \zeta[K_B - K]$$

He chooses loans $L$, bank equity $K$, portfolio diversification $z$, and the deposit rate $b$. The first-order conditions are

$$\frac{\partial \mathcal{L}}{\partial L} = p(R)r - [1 - z(1 - p(R))]b - C(z) + \frac{\lambda_1(1 - p)bK}{L^2} = 0$$
$$\frac{\partial \mathcal{L}}{\partial K} = [1 - z(1 - p(R))]b - \gamma - \frac{\lambda_1(1 - p(R)b}{L} - \zeta = 0$$
$$\frac{\partial \mathcal{L}}{\partial b} = -[1 - z(1 - p(R))][(1 - \kappa)L + \lambda_1(1 - p(R))(1 - \kappa) + \mu_1(1 - z(1 - p(R))] = 0$$
$$\frac{\partial \mathcal{L}}{\partial z} = -\lambda_1 C''(z) - \mu_1 (1 - p)b = 0$$

as well as (30), (33), (PC), and (CA). Reformulating the first condition using (PC) and the first-order condition of traditional firms (4) to eliminate $b$ as well as $\hat{\theta}(r) = L$ to eliminate $r$ yields (34). Q.E.D.

Proof of Proposition 6 Replacing (IC) by the first-order conditions (30) and (33) and using $\hat{\theta}(r) = L$, one can rewrite program 6 as a Lagrangian with five constraints and the multipliers $\lambda_1$, $\lambda_2$, $\mu_1$, $\mu_2$, and $\eta$.

$$\mathcal{L} = F(X) + [p(R)R - C(z)] L - \frac{L^2}{2} + \lambda_1 [(1 - p(R))b(1 - \kappa) - C'(z)] + \lambda_2 [p(R) + p'(R)(R - r)]$$
$$+ \mu_1[(1 - z(1 - p(R))b - F'(X)] + \mu_2[p(R)(R - r) - L] + \eta[1 - L - X] + \zeta(K_B - K)$$

The first-order conditions are

$$\frac{\partial \mathcal{L}}{\partial L} = p(R)R - C(z) - L - \frac{\lambda_1(1 - p)bK}{L^2} - \mu_2 - \eta = 0$$
$$\frac{\partial \mathcal{L}}{\partial K} = -\lambda_1 (1 - p)b \frac{L}{L} - \zeta = 0$$
$$\frac{\partial \mathcal{L}}{\partial X} = F'(X) - \mu_1 F''(X) - \eta = 0$$
$$\frac{\partial \mathcal{L}}{\partial r} = -\lambda_2 p'(R) - \mu_2 p(R) = 0$$
$$\frac{\partial \mathcal{L}}{\partial b} = \lambda_1 (1 - p)(1 - \kappa) + \mu_1 (1 - z(1 - p(R))] = 0$$
$$\frac{\partial \mathcal{L}}{\partial R} = [p + p'(R)R][L - \lambda_1 p'(R)b(1 - \kappa) - \lambda_2 [2p'(R) + p''(R)(R - r)] + \mu_1 p'(R)zb = 0$$
$$\frac{\partial \mathcal{L}}{\partial z} = -C'(z)L - \lambda_1 C''(z) - \mu_1 (1 - p)b = 0$$

as well as (30), (33), (CP), (LD) and (RC). Combining the first, third and fourth condition yields (36) that characterizes the resource allocation. The Lagrange multipliers of (IC) are:

$$\lambda_1 = \frac{[1 - z(1 - p(R))]C'(z)L}{[1 - z(1 - p(R))]C'(z) - (1 - p)C'(z)} \leq 0$$
$$\lambda_2 = \frac{L}{2p' + p''(R)(R - r)} \left[ p + p'(R) R + \frac{p'(R)b(1 - \kappa)C'(r)}{(1 - PD)C''(z) - (1 - p)C'(z)} \right] \leq 0$$
\( \lambda_1 \) is zero as soon as \( \kappa = 1 \), which implies \( C'(z) = 0 \); \( \lambda_2 \) equals zero if \( r = 0 \) and \( \kappa = 1 \) (i.e., no entrepreneurial and bank risk shifting). Therefore, \( \zeta > 0 \) and the banker invests his entire private wealth as equity, \( K = K_B \), as long as \( \kappa < 1 \) (i.e., \( K_B \geq \tilde{L} \)) as in the baseline model.

The pecuniary externalities are welfare-reducing if the social marginal return of bank lending is smaller than the private. Since the Lagrange multipliers are similar functions both in the market equilibrium and the constrained social optimum, one can directly compare social and private returns using the first-order conditions (34) and (36) which yields:

\[
\frac{\lambda_1 (1 - p(R))(1 - \kappa) F''(X)}{1 - z(1 - p(R))} - \frac{\lambda_2 p'(R)}{p(R)} > 0
\]

This condition can be expressed as (37) using the interest rate elasticities of loan demand and deposit supply:

\[
\varepsilon_L = \frac{\partial L}{\partial r} = \frac{-p(R)r}{L} < 0 \quad \varepsilon_D = \frac{\partial D}{\partial \gamma} = \frac{-\gamma}{DF''(X)} > 0
\]

The former is derived from (30), the latter from the investors participation constraint of traditional firms (4). Q.E.D.

### A.3 Numerical Example

This numerical example serves a purely illustrative purpose. We use the following functional forms:

\[
p(R) = 1 - \frac{R - \tilde{R}}{\tilde{R}} \quad F(X) = A [X^\alpha - \alpha X]
\]

The production function satisfies the Inada conditions. The parameters are set as follows: \( \tilde{R} = 1 \), \( \tilde{R} = 2.4 \), \( A = 3 \), \( \alpha = 0.5 \), \( K_B = 0.15 \), \( V = 0.85 \). We compute four scenarios: The baseline example is scenario (1). In scenario (2), we assume that banks are better capitalized, which is captured by a larger supply of bank capital: \( K_B = 0.2 \) and \( V = 0.8 \). In scenario (3), we add social cost of bank failure with a cost parameter \( s = 0.3 \). Note that in this case the number of banking licenses is omitted because one would need to account for the social cost as well and cannot rely on corollary 3. Finally, scenario (4) includes a more productive traditional sector, which is implemented by setting \( A = 4 \). For each scenario, a first best (FB), market equilibrium (ME), and constrained social optimum (CS) are computed. The outcomes are shown in table 1.

In the baseline scenario, one can observe the effects of risk shifting when comparing first best and market equilibrium: Target return increases from 1.7 to 2.32, which reduces the success probability\(^49\) from 71% to 45%. At the same time, the banking sector shrinks (from 69% to 56% of investment) due to the need to use scarce bank capital. Risk shifting leads to a welfare loss (agency cost) of roughly 6.4% of aggregate output, which is entirely borne by investors whereas the welfare of bankers and traditional firms increases. When looking at the constrained social optimum, one concludes that the pecuniary externality that can be internalized by regulation accounts for almost one quarter of the welfare loss. Compared to the market equilibrium, the regulator chooses a slightly smaller and better capitalized banking sector, which alleviates risk shifting and results in a less risky loan portfolio. The expected payoffs of bankers and traditional firms increase by 26% and 7% respectively but investors receive a payoff that is 17% lower. In scenario (2), the supply of bank capital is 5 pp larger, which has no real effect in the first best. In the presence of moral hazard, however, this leads to banks that are both larger and better capitalized such that their loan port-

\(^{49}\) The high failure probabilities arise due to perfect correlation of loan returns and the liner specification of \( p(R) \).
folio is more efficient. Hence, the larger wealth endowment of bankers translates into improvement in quantity and quality. Note that the banker’s informational rent decreases (expected value falls from 0.24 to 0.23 in the CS). One can also observe a modest increase of social welfare. In scenario (3), bank failure is associated with social cost that directly lower welfare in all allocations. Target return and bank risk are adjusted downwards already in the first best. The choices in the market equilibrium are, however, unaffected as agents fail to internalize this externality. This explains the comparatively large welfare loss of 8.1% due to risk shifting and social cost. The additional internalization of the latter justifies tighter regulation (capital loan ratio 2 pp higher than in the baseline), which induces more prudent lending. This also explains why the welfare improvement of regulation is larger (almost one third). In scenario (4), the traditional sector is more productive: As a result, the first-best resource allocation has a larger traditional and a smaller banking sector. The latter also persists in the presence of moral hazard such that banks are better capitalized for a given wealth endowment. In the market equilibrium, for instance, the capital loan ratio is 4 pp higher compared to the baseline. This dominates the potentially adverse effect of higher funding cost and mitigates risk shifting overall. Interpreting the productivity increase as a positive technology shock, capital requirements are countercyclical (34% instead of 29%) to offset the adverse incentive effect of higher funding cost.
<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>FB ME CS</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Target return $R$</td>
<td>1.7 2.32 2.13</td>
<td>1.7 2.26 2.1</td>
<td>1.55 2.32 2.06</td>
<td>1.7 2.29 2.12</td>
</tr>
<tr>
<td>Success probability $p$</td>
<td>0.71 0.45 0.53</td>
<td>0.71 0.48 0.54</td>
<td>0.77 0.45 0.56</td>
<td>0.71 0.46 0.54</td>
</tr>
<tr>
<td>Bank lending $L$</td>
<td>0.69 0.56 0.51</td>
<td>0.69 0.58 0.53</td>
<td>0.67 0.56 0.48</td>
<td>0.61 0.49 0.44</td>
</tr>
<tr>
<td>Traditional investment $X$</td>
<td>0.31 0.44 0.49</td>
<td>0.31 0.42 0.47</td>
<td>0.33 0.44 0.52</td>
<td>0.39 0.51 0.56</td>
</tr>
<tr>
<td>Max. capital loan ratio $\kappa$</td>
<td>0.22 0.27 0.29</td>
<td>0.29 0.35 0.38</td>
<td>0.22 0.27 0.31</td>
<td>0.25 0.31 0.34</td>
</tr>
<tr>
<td>Deposit rate $b$</td>
<td>1.7 1.7 1.21</td>
<td>1.7 1.7 1.27</td>
<td>1.46 1.7 1.05</td>
<td>1.7 1.7 1.27</td>
</tr>
<tr>
<td>Welfare $W$</td>
<td>2.04 1.91 1.94</td>
<td>2.04 1.94 1.96</td>
<td>1.98 1.82 1.87</td>
<td>2.45 2.36 2.37</td>
</tr>
<tr>
<td>Exp. payoff bankers $\pi^B$</td>
<td>0.18 0.27 0.34</td>
<td>0.24 0.31 0.38</td>
<td>0.18 0.27 0.36</td>
<td>0.18 0.25 0.3</td>
</tr>
<tr>
<td>Exp. payoff trad. firms $\pi^F$</td>
<td>0.83 0.99 1.06</td>
<td>0.83 0.98 1.03</td>
<td>1.25 1.43 1.49</td>
<td>1.25 1.43 1.49</td>
</tr>
<tr>
<td>Exp. payoff investors $\pi^I$</td>
<td>1.02 0.65 0.54</td>
<td>0.96 0.65 0.55</td>
<td>0.05 0.09 0.06</td>
<td>1.02 0.67 0.58</td>
</tr>
<tr>
<td>Exp. social cost</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Capital requirement $\kappa^*$</td>
<td>0.29 0.38 0.31</td>
<td>0.31 0.34</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Licenses $N^*$</td>
<td>3</td>
<td>3</td>
<td>0.31</td>
<td>0.34</td>
</tr>
</tbody>
</table>

Table 1: Numerical Example