

Scope and Limits of Bank Liquidity Creation*

Diemo Dietrich[†]

Thomas Gehrig[‡]

August 11, 2022

Abstract

In standard banking models a preference for liquidity arises because investors want to take precautions against sudden expenditure needs. We propose that investors may also want to preserve flexibility in case better investment opportunities arrive later. The co-existence of both investor types is crucial for the scope and limits of bank liquidity creation. When standard financial frictions apply, co-existence can result in welfare losses relative to a world with only a single investor type. In a friction-free world, however, such co-existence entails welfare gains. In either case, policy recommendations based only on a single motive for liquidity demand may be seriously misguided.

Keywords: investment opportunities · expenditure needs · liquidity demand · competitive insurance markets

JEL Classification: D11, D86, E21, E22, G21, L22

*Without implicating them we are grateful for the comments from Toni Ahnert, Ron Anderson, Sonny Biswas, Mike Burkhart, Matthieu Chavaz (discussant), Bartosz Gebka, Hans Gersbach, Charles Goodhart, Martin Hellwig, Esa Jokivuolle, Kostas Koufopoulos, Alex Mürmann, Martin Oehmke, Paul Pfleiderer, Jean-Charles Rochet, Eva Schliephake, Vania Stavrakeva, Günter Strobl, Uwe Vollmer, Josef Zechner, Jean-Pierre Zigrand, and seminar participants at the joint University of Zürich/ETH Seminar, the VGSF Research Seminar, the SRC Seminar at LSE, and participants at the Economic Society European Meeting 2021, the London FIT workshop, the 37th GdRE/Banque de France Symposium on Money, Banking and Finance, the Money Macro & Finance Conference 2021 in Cambridge, and the Verein für Socialpolitik Conference 2021.

[†]University of Greifswald, Germany, diemo.dietrich@uni-greifswald.de

[‡]University of Vienna, A, and CEPR and VGSF, e-mail: thomas.gehrig@univie.ac.at

1 Introduction

Already Adam Smith (1776) considered the ability of society to create, and utilize, liquidity as key for efficiently directing savings to investments. But what precisely is liquidity in this context? In this paper, we place uncertainty about the timing of future productive investment opportunities into the focus of attention. Such uncertainty generates a preference for liquidity as it induces a desire in investors to be flexible and able to withdraw from previous investments in order to take advantage of more lucrative opportunities as and when they arrive. While establishing a direct link between liquidity and the efficiency of capital allocation, this perspective has not gained much attention in a banking context.¹

Instead, the banking literature has widely built on the notion that liquidity preferences arise from a desire to *take precautions* against sudden consumption needs (due to shocks to agents' marginal rate of intertemporal substitution, see Bryant, 1980; Diamond and Dybvig, 1983). We study competitive banking sectors where the canonical precautionary motive *co-exists* with liquidity preferences based on the desire to *preserve flexibility* for future investment opportunities. With this approach, we primarily aim to explore the nature of banks' liquidity creation as an equilibrium outcome in a model with multiple sources of liquidity demand. We do so because, empirically, banks create liquidity through a variety of on- and off-balance sheet activities (Berger and Bouwman, 2009), suggesting that a preference for liquidity can arise for various reasons. This view gives rise to a range of new questions. Above all, how does the *co-existence* of diverse motives for liquidity preferences affect market outcomes? Do market outcomes deviate from an efficient creation of liquidity? What intermediation arrangements will arise? To what extent do standard credit frictions matter?

Our secondary aim is to relate such a general framework to the canonical models of bank liquidity creation that focus on only one source of liquidity preferences. The latter have been frequently used for policy analysis and policy advice, particularly for banking regulation. Notably, the maturity transformation that features prominently in those specific models has been widely interpreted as a source of systemic risk and, indeed, turned into the dominant manifestation of systemic risk in banking. In

¹Since Jones and Ostroy (1976, 1984) a similar real option value has been ascribed to safe and liquid assets such as outside money, while disregarding additional potential efficiency gains attainable through financial intermediation.

our analysis it turns out that maturity transformation is not a general outcome in the banking market as widely suggested by the canonical model. Rather pooling or even non-existence of equilibrium can occur in equilibrium, when multiple source of liquidity preferences are considered.

In our general model with multiple sources of liquidity demand, financial frictions are key. For standard credit frictions and information asymmetries about the nature of liquidity preferences, we document a variety of equilibrium outcomes. For example, there is a range of parameter values for which both types of liquidity preferences can be efficiently served, independently of each other. There is also another, wide range of parameter values for which the liquidity preferences interact with significant consequences. There, inefficiencies arise, with either too much or too little maturity transformation, and an equilibrium may not even exist at all. Interestingly, in the limiting case, when taking precautions becomes the dominant motive within the total population, liquidity creation may not converge to the one that obtains in a specific model that considers exclusively the precautionary motive. Likewise, when preserving flexibility for future investment opportunities gains dominance in a mixed population of investors, there is not necessarily convergence to the equilibrium that obtains in total absence of the precautionary motive. Equilibria in models that consider only a single motive for liquidity preferences are thus not necessarily robust. Last not least, when there are no significant financial frictions, the equilibrium outcome contrasts sharply with the conventional view, too. This is because the co-existence of motives actually generates scope for synergies. Banking sectors that serve both types of liquidity preferences simultaneously can take advantage of such synergies without engaging in any maturity transformation. Specifically, banks hold only liquid reserves and grant short-term credit lines while raising funds through, e.g., demand deposits (held by investors who want to take precautions) and equity shares (held by investors who want to preserve flexibility).

It is important to note that considering the co-existence of various motives for liquidity preferences in an economy opens a wider perspective than the mere co-existence of deposit taking and lending within a single bank. In this latter sense Kashyap et al. (2002) argue that banks can economize on costly reserve holdings through diversifying liquidity outflows from both sides of their balance sheet. By contrast, our interest lies in the competitive relations between the underlying liquidity motives themselves. For example, without financial frictions, banks take advantage of hitherto unnoticed

synergies. Specifically, credit lines are utilized only by lucky investors who can actually find more lucrative opportunities. Therefore, credit lines generate higher returns for banks than long-term investments, and these higher returns ease a bank's constraints on its liquidity creation for all investors, including for those who take precautions against sudden consumption needs. With financial frictions, while credit lines and deposit taking do not create synergies, the co-existence of various motives for liquidity preferences affects equilibrium bank liquidity creation in non-trivial ways, and can even cause a market breakdown.²

Our analysis has a number of implications. Firstly, and importantly, the nature of liquidity matters, not only but especially for banks' maturity transformation.³ Secondly, financial frictions are key determinants for a possible interdependence of different liquidity services. For example, depending on which frictions hold, the maturity transformation performed by banks may be less prevalent or excessive under co-existence than in cases with only a single motive for liquidity preferences. Thirdly, while in equilibrium each bank business model earns zero profits, serving the desire for preserving flexibility requires lower levels of reserves and earns higher returns on assets than serving the need for taking precautions. This can have implications for, e.g., internal incentive schemes as managers in different business lines of a bank should be assessed according to different performance criteria. Finally, the scope for creating liquidity for those with a liquidity preference for precautionary reasons is fading if returns on long-term projects decline. This can imply that pure-strategy equilibria do not exist and equilibrium outcomes are thus effectively indeterminate. If such a decline of returns on long-term projects can be related to an economy-wide (exogenous) fall in the level of interest rates, hitherto unnoticed – and unintended – consequences emerge from a zero-interest rate environment.

The role of bank liquidity creation for investment finance has been analysed extensively in the literature, most notably by Holmström and Tirole (1998), but with a different perspective. While those authors also stress the liquidity implications of a limited pledgeability of future returns, our framework emphasizes the possibility of future investment opportunities as an additional motive for

²In principle one could consider an even richer setup with more than two motives for liquidity preferences. However, two motives already generate a fairly rich menu of equilibrium outcomes that contrast sharply to models with single motives.

³This parallels the thrust of Gehrig and Jackson (1998) in a trading context.

a preference for liquidity. Also, in Holmström and Tirole (1998), credit frictions are not absolute. Therefore, banks still offer lines of credit to firms. Unlike credit lines, however, the deposit contract, which obtains in our model when credit frictions are absolute, also provides for higher long-term returns to those unlucky investors who do not find investment opportunities.

Credit frictions have been also identified to generate a demand for liquid, marketable financial assets, such as volatile bubbles (Martin and Ventura, 2012) and fiat money (Dietrich et al., 2020). In our model, credit frictions generate the maturity transformation banks typically engage in. A range of reasons has been identified for the credit friction utilized in the present paper. For example, only the investor has the specific skills needed to successfully manage and complete the project (Hart and Moore, 1994), their consumption is not observable (Wallace, 1988), or penalties like future exclusion from financial markets are ineffective for enforcing loans (Kehoe and Levine, 1993).

Finally, by concentrating our analysis on intermediaries, we sidestep the role of markets in providing liquidity. This is partly because Geanakoplos and Walsh (2018) establish that due to pecuniary externalities competitive markets tend to under-provide liquidity in the single-motive model of Diamond and Dybvig (1983) under fairly mild conditions. Moreover, in a dynamic context, market equilibria are inherently unstable in the single-motive case (Dietrich and Gehrig, 2021).

The paper is organized as follows: Section 2 presents the details of the model. Section 3 provides the analysis of equilibrium outcomes in the presence of frictions. Section 4 contains the key insights for the frictionless economy. Section 5 discusses some implications of our analysis and briefly reviews the main features of our setup. Section 6 concludes the paper. All technical proofs are relegated to the Appendix.

2 Setup

Consider an economy populated by investors and banks. There are three dates $t \in \{0, 1, 2\}$, with one good at each date. The good can be consumed or used for production in one of three technologies.

Technologies The technologies are: storage, long-term production and short-term production. Each technology features constant returns to scale. Storage is one-for-one and can be used at dates $t \in \{0, 1\}$. Long-term production has to be initiated at $t=0$, and takes two periods until $t=2$ to produce the good. Per-unit-returns are $R > 1$. Long-term production cannot be prematurely liquidated at date $t=1$. Henceforth, we refer to long-term production also as R -technology. Short-term production opportunities arise at date $t=1$, to produce $Q > R$ per unit of investment after one period at date $t=2$. Accordingly, short-term production is called Q -technology.

Investors There is a continuum of investors, each endowed with one unit of the good at $t=0$. All investors have access to the R -technology at date $t=0$, and to storage at dates $t=0$ and $t=1$. There are two types of investors. One type of investors values consumption c only at date $t=2$. As of date $t=0$, there is a probability $\mu \in]0, 1[$ that an investor of this type gets lucky at date $t=1$ as she will gain access to the Q -technology. With probability $1 - \mu$ she will remain without access to the Q -technology. Getting lucky is uncorrelated across investors of this type. Although only interested in consumption at the final date, a long-term commitment to the illiquid R -technology is not optimal for them as these investors want to preserve their *flexibility* in case they get lucky at date $t=1$. Henceforth, we call investors who seek to preserve their flexibility F -investors.

An investor of the other type does not know at date $t=0$ when she needs to consume. Specifically, with probability $\lambda \in]0, 1[$ she will value only consumption c at date $t=1$, whereas with probability $1 - \lambda$, she will value consumption only at date $t=2$. Getting impatient, i.e. having to consume early, is uncorrelated across investors of this type. For them, a long-term commitment to the illiquid R -technology is not optimal because these investors need to take *precautions* against sudden expenditure needs. They are henceforth called P -investors.

The investors' *Bernoulli* utility function u is independent from their type, twice continuously differentiable, and satisfies $u'(c) > 0$, $u''(c) < 0$, $\lim_{c \rightarrow 0} u'(c) = \infty$, and $\lim_{c \rightarrow \infty} u'(c) = 0$. To simplify the exposition, we divide investors into groups, each of mass one. In every group there are either F -investors or P -investors, with γ and $1 - \gamma$ as the shares of P -investor groups and F -investor groups, respectively, in the total population. As the probabilities μ and λ are deterministic and common

knowledge at date $t = 0$, the law of large numbers applies, i.e. a share μ in a group of F-investors is lucky, and a share λ in a group of P-investors is impatient.

Banks There is a continuum of penniless banks. They have access to storage and to the R -technology, but not to the Q -technology. Banks are perfectly competitive (Bertrand competition) and maximize expected profits. At date $t = 0$, investors can exchange their endowments for contracts offered by banks. A *contract* $\mathcal{D} = (r_1, r_2)$ is a sequence of payments $\{r_t\}_{t \in \{1,2\}}$ a bank makes to investors at $t = 1$ and $t = 2$, respectively. A *business model* $\mathcal{M} = (r_1, r_2, y)$ consists of a contract \mathcal{D} and a portfolio share held in storage $y \in [0, 1]$, and is *sustainable* if designed to earn non-negative profits.⁴

Frictions There are potentially three types of frictions. Firstly, at date $t = 0$, the ex-ante motive for the liquidity preference is private information. Accordingly, investors are free to choose between all contracts banks offer. Secondly, at date $t = 1$, the realized consumption need is private information, i.e. only the individual P-investors learns whether they get impatient and need to consume immediately, or patient and can wait until date $t = 2$. Similarly, access to the Q -technology is private information as only the individual F-investor learns at date $t = 1$ whether they are lucky and can invest in the profitable new opportunity or not. Therefore, contracts cannot be made contingent on the ex-post realization of liquidity needs of investors. Thirdly, while storage and the R -technology are available to investors and banks alike, and are thus fully contractible, the Q -technology is specific to F-investors who are not able to credibly pledge the returns they realize with this superior technology at date $t = 2$. In what follows we consider two different scenarios, one without any frictions and the other with all three frictions in place.

Assumptions For our analysis we make two technical assumptions.

A1 (Relative Risk Aversion)

The coefficient of relative risk aversion exceeds one, i.e., $-cu''(c)/u'(c) > 1$.

⁴In line with the literature, we do not allow for re-depositing after a withdrawal. Re-depositing is particularly relevant in a dynamic context with overlapping generations (see Bhattacharya and Padilla, 1996).

Without this assumption, many results will just be reversed for $-cu''(c)/u'(c) < 1$.

A2 (Single Crossing Condition)

In the (r_1, r_2) space, indifference curves of P-investors and of F-investors cross only once.

Single crossing is a standard assumption in mechanism design theory. In our context, it is satisfied, for example, if relative risk aversion is constant, and ensures that the set of probabilities μ , for which equilibrium types obtain, is convex.

In what follows, we let c_R and c_Q denote the consumption by an F-investor with and without investment opportunity, respectively. Resources per F-investor directed to the R -technology and to storage are x_R and y^F , respectively, and investment into the Q -technology per F-investor, who has actually access to it, is x_Q . Similarly, we let c_1 and c_2 denote a P-investor's consumption if impatient and patient, respectively, while x^P and y^P are resources per P-investor directed to the R -technology and to storage, respectively. Finally, an *economy* \mathcal{E} is a description of F-investors, P-investors, and technologies, i.e. $\mathcal{E} = (u, \gamma, \lambda, \mu, Q, R)$.

For the analysis in the following sections it is helpful to consider first the optimal solution to the respective investors' risk-sharing problems for the hypothetical case that F-investors can be distinguished from P-investors whilst all other frictions remain.

The F-investors' problem All F-investors are ex-ante identical. Therefore, sharing their risks optimally among themselves is a solution to the following problem

$$\begin{aligned} & \max_{(c_R, c_Q, x_R, x_Q, y^F) \in \mathbb{R}_+^4 \times [0,1]} \mu u(c_Q) + (1 - \mu) u(c_R) \\ \text{s.t.} & \left\{ \begin{array}{l} x_R + y^F = 1 \\ \mu x_Q = y^F \\ (1 - \mu) c_R = R x_R \\ c_Q = Q x_Q \\ c_R \geq x_Q \\ c_Q \geq c_R \end{array} \right. \end{aligned} \tag{1}$$

The first constraint is the resource constraint at date $t=0$; the second constraint states that investment in the Q -technology at date $t=1$ is limited to what has been stored at date $t=0$. The third constraint states that consumption by F-investors without access to the Q -technology is equal to what is generated with the R -technology. The fourth constraint states that consumption by F-investors with access to the Q -technology is equal to what is generated with this technology.⁵ The final two lines are the incentive constraints, ensuring that unlucky F-investors have no incentive to pretend they got access to the Q -technology, and that lucky F-investors have no incentive to pretend they got no access to it.

Disregarding for now the final two constraints, a solution to this problem satisfies the first four constraints and the first-order condition

$$u' \left(\frac{R(1-y^d)}{1-\mu} \right) - \frac{Q}{R} u' \left(\frac{Qy^d}{\mu} \right) = 0 \quad (2)$$

with y^d as optimal storage. Banks implement the solution to problem (1) because, due to Bertrand competition, they either operate a business model $\mathcal{M} = (r_1^d, r_2^d, y^d)$, with $r_1^d = x_Q = y^d/\mu$ and $r_2^d = Rx_R/(1-\mu) = R(1-y^d)/(1-\mu)$ or leave the market.

The contract (r_1^d, r_2^d) features certain characteristics worthy further elaboration. Let c_Q^d and c_R^d be the consumption by F-investors with and without access to the Q -technology, respectively, associated with the business model (r_1^d, r_2^d, y^d) . Then, condition (2) implies $c_Q^d > c_R^d$. For relative risk aversion equal to one, i.e. $-cu''(c)/u'(c) = 1$, we obtain $cu'(c) = u'(1)$. Hence, $Ru'(R) = Qu'(Q)$, and the first-order condition (2) requires $c_Q^d = Q$ and $c_R^d = R$. Accordingly, the bank's business model satisfies $r_1^d = 1$, $r_2^d = R$ and $y^d = \mu$. For relative risk aversion greater one, i.e. $-cu''(c)/u'(c) > 1$, we obtain $Ru'(R) > Qu'(Q)$. Therefore, condition (2) requires $R < c_R^d < c_Q^d < Q$. Accordingly, the bank's business model satisfies $r_1^d < 1$, $r_2^d = R(1-\mu r_1^d)/(1-\mu) > R$, and $y^d < \mu$. Note $r_1^d < 1$ implies that the contract entails a penalty for early withdrawals. Finally, by Assumption 1, the contract (r_1^d, r_2^d) is also incentive compatible.⁶

⁵Stating these constraints directly with equality is innocent. In short, equality follows from non-satiation together with $Q > R > 1$. The latter implies that it is neither efficient to keep any storage between date $t=1$ and $t=2$ nor to use the R -technology for the consumption by investors with access to the Q -technology.

⁶F-investors without access to the Q -technology are better off withdrawing at $t=2$ as $r_1^d < r_2^d$. F-investors with access to the Q -technology consume Qr_1^d if they withdraw at $t=1$ and r_2^d if they withdraw at $t=2$ and $Qr_1^d > r_2^d$ by condition (2).

The P-investors' problem All P-investors are ex-ante identical, and their problem is well known since Diamond and Dybvig (1983). Sharing their risks optimally among themselves is a solution to the following problem

$$\begin{aligned} & \max_{(c_1, c_2, x_R, y) \in \mathbb{R}_+^3 \times [0, 1]} \lambda u(c_1) + (1 - \lambda) u(c_2) \\ & \text{s.t.} \begin{cases} x_R + y = 1 \\ \lambda c_1 = y \\ (1 - \lambda) c_2 = R x_R \\ c_1 \leq c_2 \end{cases} \end{aligned} \quad (3)$$

The restrictions are the feasibility constraints. The first line is the resource constraint at date $t = 0$; the second line requires that consumption by impatient P-investors is limited by the stored goods available at date $t = 1$; the third line is derived from the constraint that total consumption is limited by the total availability of stored and produced goods.⁷ The last line ensures incentive compatibility if information about early consumption needs is private. Disregarding the final constraint for now, the solution to the problem satisfies $c_1^\delta = y^\delta / \lambda$ and $c_2^\delta = R(1 - y^\delta) / (1 - \lambda)$ where y^δ solves the first-order condition

$$u' \left(\frac{y^\delta}{\lambda} \right) - R u' \left(\frac{R(1 - y^\delta)}{1 - \lambda} \right) = 0. \quad (4)$$

Again, competitive banks implement optimal risk-sharing by operating a business model $\mathcal{M} = (r_1^\delta, r_2^\delta, y^\delta)$ with $r_1^\delta = y^\delta / \lambda$ and $r_2^\delta = R(1 - y^\delta) / (1 - \lambda)$. For relative risk aversion equal to one, a bank's payment to those who withdraw early is $r_1^\delta = 1$. Then, from the feasibility constraint for $t = 2$, we obtain $r_2^\delta = R$. For relative risk aversion greater one, the bank pays an insurance benefit at date $t = 1$, i.e. more than a P-investor has deposited with the bank in the first place. Specifically, the bank pays $r_1^\delta = y^\delta / \lambda > 1$ and $r_2^\delta = R(1 - \lambda r_1^\delta) / (1 - \lambda) \in]r_1^\delta, R[$. Accordingly, P-investors receive a subsidized rate for early withdrawal, rather than a penalty as it was the case with F-investors. Finally, this contract also ensures incentive compatibility if information about early consumption needs is private.⁸

⁷Equality in these three constraints follows from non-satiation together with $R > 1$.

⁸Impatient P-investors have no choice but to withdraw at date $t = 1$. Patient P-investors are strictly better off by waiting until date $t = 2$ since $r_2^\delta > r_1^\delta$

3 Co-existence of liquidity motives with frictions

In the present section we study the co-existence of F-investors and P-investors in a frictional market, i.e. when both their identity and the individual liquidity event are private information and banks cannot enforce loan repayments. The credit friction implies that banks offer only deposit contracts (r_1^F, r_2^F) and (r_1^P, r_2^P) to investors, but no loans.

3.1 Equilibrium concept

We consider competitive deposit markets and focus on equilibria in the spirit of Rothschild and Stiglitz (1976). Specifically, we consider pure-strategy equilibria, where each bank is limited to offering one deposit contract, and investors choose from all contracts offered by banks to maximize their expected utility but cannot randomize their choice. It is useful to begin with a definition of incentive compatible contracts.

Definition 1 (Incentive Compatible Contracts)

Let $\mathcal{D}^F = (r_1^F, r_2^F)$ be the contract for F-investors, and $\mathcal{D}^P = (r_1^P, r_2^P)$ the contract for P-investors. An incentive compatible menu of contracts $\{\mathcal{D}^F, \mathcal{D}^P\}$ satisfies

$$\mu u(Qr_1^F) + (1 - \mu)u(r_2^F) \geq \mu u(Qr_1^P) + (1 - \mu)u(r_2^P) \quad (5)$$

$$\lambda u(r_1^P) + (1 - \lambda)u(r_2^P) \geq \lambda u(r_1^F) + (1 - \lambda)u(r_2^F) \quad (6)$$

$$Qr_1^F \geq r_2^F \quad (7)$$

$$r_1^F \leq r_2^F \quad (8)$$

$$r_1^P \leq r_2^P \quad (9)$$

$$\mu u(Qr_1^F) + (1 - \mu)u(r_2^F) \geq \sup \{ \mu u(Qy^F + R(1 - y^F)) + (1 - \mu)u(R(1 - y^F) + y^F) : y^F \in [0, 1] \} \quad (10)$$

$$\lambda u(r_1^P) + (1 - \lambda)u(r_2^P) \geq \sup \{ \lambda u(y^P) + (1 - \lambda)u(R(1 - y^P) + y^P) : y^P \in [0, 1] \} \quad (11)$$

Condition (5) requires that F-investors prefer the contract intended for F-investors over the contract intended for P-investors, with strict inequality for $(r_1^F, r_2^F) \succ_F (r_1^P, r_2^P)$. Condition (6) requires that P-investors prefer the contract intended for P-investors, with strict inequality for $(r_1^P, r_2^P) \succ_P (r_1^F, r_2^F)$.

These two incentive constraints need to be satisfied at date $t = 0$. For contracts to be incentive compatible, there are also incentive constraints at date $t = 1$ when investors know their status. Specifically, condition (7) requires that F-investors with access to the Q -technology must not be better off by pretending to have no access; condition (8) that F-investors without access to the Q -technology must not be better off by pretending to have access; and condition (9) that patient P-investors must not be better off by pretending to be impatient. Finally, contracts must be such that depositing with banks makes investors better off than autarky. This holds provided the expected utility associated with their contract is at least as large as the expected utility an investor achieves in autarky, i.e. if contracts satisfy the participation constraints (10) and (11).

We can now define a banking equilibrium.

Definition 2 (Banking Equilibrium)

A perfect-competition, pure-strategy banking equilibrium is an incentive compatible menu of contracts $\{\mathcal{D}^F, \mathcal{D}^P\}$ such that the associated business models $\{\mathcal{M}^F, \mathcal{M}^P\}$ are sustainable, while no bank can profitably enter the market with another contract $\mathcal{D}' \notin \{\mathcal{D}^F, \mathcal{D}^P\}$.

A bank's business models is *sustainable* if the bank does not make a loss and would thus be strictly better off exiting the market. A business model $\mathcal{M}^F = (r_1^F, r_2^F, y^F)$ of offering contracts only to F-investors is sustainable if $\mu r_1^F \leq y^F$ and $(1 - \mu)r_2^F \leq R(1 - y^F)$; a business model $\mathcal{M}^P = (r_1^P, r_2^P, y^P)$ of offering contracts only to P-investors is sustainable if $\lambda r_1^P \leq y^P$ and $(1 - \lambda)r_2^P \leq R(1 - y^P)$; and a business model $\mathcal{M}^{\text{Pool}} = (r_1^{\text{Pool}}, r_2^{\text{Pool}}, y^{\text{Pool}})$ of offering the same contract, a pooling contract, to F-investors and to P-investors alike, i.e. $\mathcal{D}^F = \mathcal{D}^P = (r_1^{\text{Pool}}, r_2^{\text{Pool}})$, is sustainable if $(\gamma\lambda + (1 - \gamma)\mu)r_1^{\text{Pool}} \leq y^{\text{Pool}}$ and $(1 - (\gamma\lambda + (1 - \gamma)\mu))r_2^{\text{Pool}} \leq R(1 - y^{\text{Pool}})$. Provided either of these inequalities is strict, the respective business model is associated with strictly positive profits.

In equilibrium, there is *no profitable market entry* by banks with contracts other than \mathcal{D}^F and \mathcal{D}^P . Therefore, operating banks make zero profits. A business model associated with a contract only for F-investors thus satisfies $(1 - \mu)r_2^F = R(1 - \mu r_1^F)$; a business model associated with a contract only for P-investors satisfies $(1 - \lambda)r_2^P = R(1 - \lambda r_1^P)$; and a business model associated with one contract for both investor types, satisfies $(1 - (\gamma\lambda + (1 - \gamma)\mu))r_2^{\text{Pool}} = R(1 - (\gamma\lambda + (1 - \gamma)\mu)r_1^{\text{Pool}})$.

A separating equilibrium is *credit-constrained* if the credit friction constitutes the only constraint that is actually binding. Hence, the contract \mathcal{D}^F maximizes the expected utility of F-investors subject only to the zero-profit condition $(1 - \mu)r_2^F = R(1 - \mu r_1^F)$ and the contract \mathcal{D}^P maximizes the expected utility of P-investors subject only to the zero-profit condition $(1 - \lambda)r_2^P = R(1 - \lambda r_1^P)$. That is, private information about an investor's type or about an investor's realized liquidity event does not imply that any of the incentive constraints (5) through (9) are binding. Such equilibrium is necessarily separating.

A separating equilibrium is called *incentive-constrained* if, in addition to the credit friction, at least one of the incentive constraints arising from the private information about the investor type, (5) or (6), is binding. For example, if (5) is binding, then banks cannot profitably stay in, or enter, the market with a business model $(r_1^P, r_2^P, y^P) = (r_1^\delta, r_2^\delta, y^\delta)$ because they would attract not only all P-investors but also all F-investors, which renders such business model unviable. In any case, banks' contract offers to the F-investors satisfy the zero-profit condition $(1 - \mu)r_2^F = R(1 - \mu r_1^F)$, while banks' contract offers to the P-investors satisfy the zero-profit condition $(1 - \lambda)r_2^P = R(1 - \lambda r_1^P)$.

In a *pooling equilibrium* F-investors and P-investors obtain one and the same contract, i.e. $\mathcal{D}^F = \mathcal{D}^P$, and this contract satisfies the joint zero-profit constraint that obtains if banks pool the resources of all investors, i.e. $(1 - (\gamma\lambda + (1 - \gamma)\mu))r_2^{\text{Pool}} = R(1 - (\gamma\lambda + (1 - \gamma)\mu)r_1^{\text{Pool}})$. Banks offering separating contracts cannot profitably enter the market in such pooling equilibria. This is because either F-investors and P-investors would both prefer the pooling contract over the separating contracts, or P-investors prefer the F-investors' contract, F-investors prefer the P-investors' contract, or both.

For the following analysis, it is helpful to think of the zero-profit conditions as graphs in a (r_1, r_2) space. For any business model, these graphs are linear and go through $(1, R)$, regardless of the value for λ , μ , and γ . The contract that serves best the F-investors' liquidity preference is characterized by a pair (r_1^d, r_2^d) for which the F-investors' indifference curve is tangent to the banks' zero-profit line $r_2 = R(1 - \mu r_1)/(1 - \mu)$, provided the banks' business model is targeted solely at F-investors. Recall that by Assumption 1, this contract implies $r_1^d < 1$ and $r_2^d > R$. The contract that serves best the P-investors' liquidity preference is characterized by a pair (r_1^δ, r_2^δ) for which their indifference curve is

tangent to the banks' zero-profit line $r_2 = R(1 - \lambda r_1)/(1 - \lambda)$ and by Assumption 1, satisfies $r_1^\delta > 1$ and $r_2^\delta < R$. Finally, a bank's zero-profit line is steeper for a larger μ and λ , respectively.

3.2 Rare investment opportunities

If the proportion of impatient P-investors is not smaller than the proportion of lucky F-investors, $\mu \leq \lambda$, the zero-profit line for P-investors is steeper than the respective zero-profit line for F-investors. The efficient P-investors' contract (r_1^δ, r_2^δ) lies inside the set of feasible contracts for F-investors. Therefore, F-investors prefer their own credit-constrained contract (r_1^d, r_2^d) over the efficient contract for P-investors. Intuitively, from an F-investor's perspective, the insurance benefit of a contract for P-investors to those withdrawing early is small relative to what one has to give up when remaining patient. This makes the P-investors' contract sufficiently unattractive to F-investors. A similar argument can be made for the incentives of P-investors. The contract intended for F-investors is unattractive to P-investors because, as P-investors are more likely to withdraw early, the penalty associated with an F-investors' contract is particularly costly for P-investors. However, credit-constrained separation equilibria not only exist for $\mu \leq \lambda$, but even for $\mu > \lambda$ up to a critical level $\bar{\mu} < 1$.

Proposition 1 (Credit-constrained Separation)

Consider economies $\mathcal{E} = (u, \gamma, \lambda, \mu, Q, R)$, where u satisfies Assumption 1 and Assumption 2. Then, for every $R > 1$, $Q > R$, and $\lambda \in]0, 1[$, there is $\bar{\mu} \in]\lambda, 1[$ such that a credit-constrained separation equilibrium exists if and only if $\mu \leq \bar{\mu}$. In credit-constrained separation equilibria, the marginal rate of substitution between r_1 and r_2 is lower for F-investors than for P-investors, i.e. $-\frac{\mu}{1-\mu} \frac{u'(Qr_1^F)}{u'(r_2^F)} Q > -\frac{\lambda}{1-\lambda} \frac{u'(r_1^P)}{u'(r_2^P)}$.

Proof: See Appendix A. □

Figure 1 illustrates equilibria that involve credit-constrained separation for some $\mu \in]\lambda, \bar{\mu}[$. F-investors strictly prefer the solution (r_1^d, r_2^d) to their problem (1) over the solution (r_1^δ, r_2^δ) to the P-investors' problem (3), as (r_1^δ, r_2^δ) lies below the F-investors' indifference curve going through their

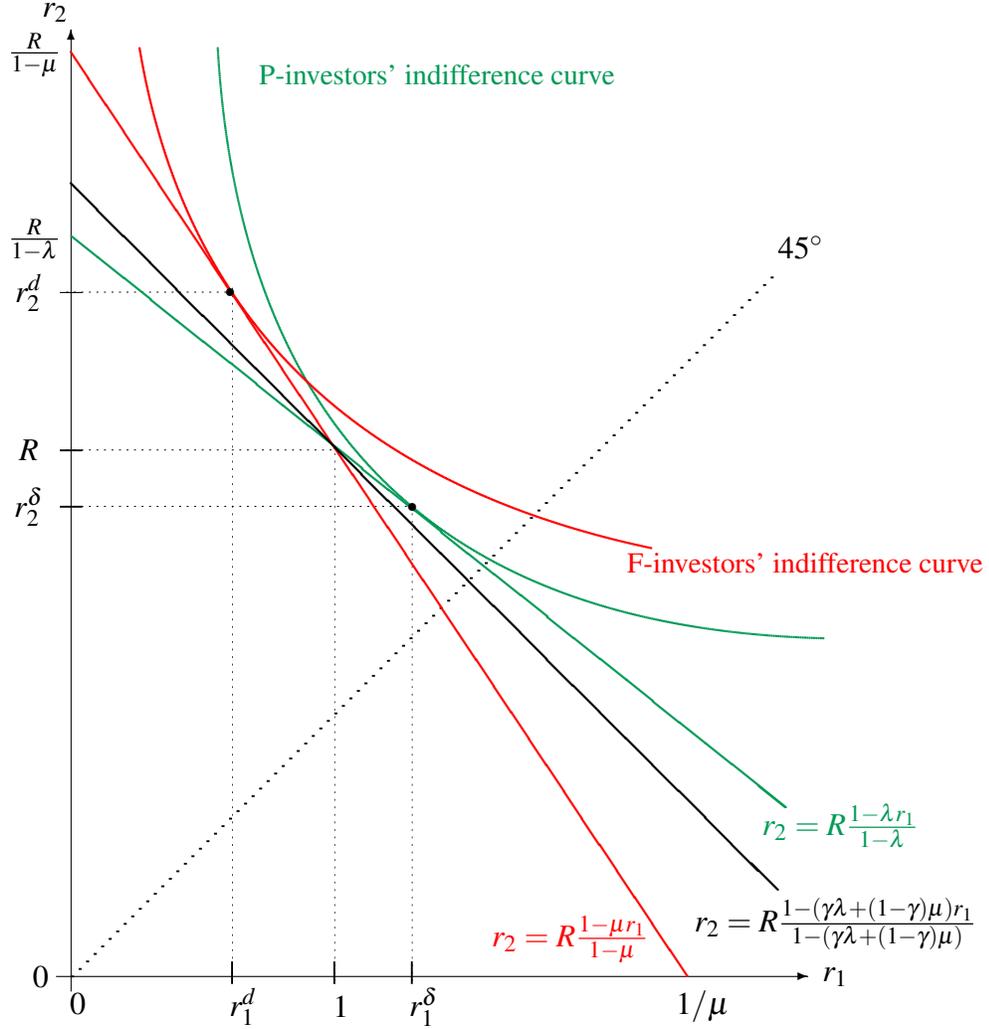


Figure 1: Credit-constrained Separation.

own contract (r_1^d, r_2^d) . Similarly, P-investors strictly prefer (r_1^δ, r_2^δ) over (r_1^d, r_2^d) . Under Assumptions 1 and 2, these preference relations imply that the indifference curve of F-investors is flatter than the indifference curve of P-investors, as can be seen at the intersection of both curves. It is not possible for any bank to profitably enter the market by offering a contract designated either exclusively to F-investors or exclusively to P-investors, because F-investors as well as P-investors already enjoy the best allocation possible given the credit friction. Also, a bank cannot profitably enter the market with a pooling contract. This is because the zero-profit constraint associated with pooling, $r_2 = R \frac{1 - (\gamma\lambda + (1-\gamma)\mu)r_1}{1 - (\gamma\lambda + (1-\gamma)\mu)}$, does not facilitate any contracts that are Pareto-improvements to the separating contracts (r_1^δ, r_2^δ) and (r_1^d, r_2^d) .

To sum up, provided the share μ of lucky F-investors is not too large, a banking equilibrium efficiently provides for the liquidity needs of P-investors and F-investors subject only to the credit constraint. P-investors are insured against the risk of the need to consume early, while F-investors are insured against the risk of missing a better investment opportunity. Both motives require some liquidity management, but optimal contracts stipulate different solutions. While the insurance payment is front-loaded in the contract with P-investors, and back-loaded in the contract with F-investors, nobody has an incentive to hide their own motive for their liquidity preference.

Corollary 1 (Bank Reserves)

For $\mu \leq \lambda$ the reserve holdings of the P-investor bank are larger than the F-investor bank, i.e. $y^P > y^F$. Accordingly, expected returns for a bank specializing on F-investors exceed those of a bank focusing on P-investors.

Business models (r_1^F, r_2^F, y^F) associated with optimal deposits for F-investors thus require lower reserve holdings than business models (r_1^P, r_2^P, y^P) associated with optimal deposits for P-investors. This is because F-investors require reserves below their probability of getting lucky μ , i.e. $y^F = y^d < \mu$, while P-investors require reserves in excess of their probability to consume early λ , i.e. $y^P = y^\delta > \lambda$. Therefore, $y^F < y^P$ for $\mu \leq \lambda$. The differences in bank portfolios have direct implications for the returns on bank assets. As those are determined by $y + R(1 - y)$ the returns on assets are higher for a bank with F-investors than for a bank with P-investors for $y^d < y^\delta$. To the extent that the different business models are offered in-house by a single (universal) bank, the Corollary implies the return on assets be applied differently across services.⁹

3.3 Frequent investment opportunities

Now consider economies with a relatively high probability of getting access to highly productive investment opportunities, i.e. $\mu > \bar{\mu}$. Under such conditions, how will equilibrium outcomes be affected by the co-existence of the two motives for liquidity demand? It turns out that the outcomes

⁹This implication is in contradiction to real-world conduct as presented in Pennacchi and Santos (2021), according to which management compensation is based on total return on equity, aggregated across all product lines.

can vary substantially, depending on the specific characteristic of the economy at hand: we consider separating equilibria with inflated insurance for P-investors and pooling equilibria, and also show that pure-strategy equilibria may not exist altogether.

Incentive-constrained separation with inflated insurance for P-investors Suppose the P-investors' marginal rate of substitution between r_1 and r_2 exceeds the rate for F-investors for all realizations of (r_1, r_2) , yet the probability μ of accessing the Q -technology is sufficiently large such that F-investors prefer the efficient contract (r_1^δ, r_2^δ) for P-investors over the credit-constrained contract (r_1^d, r_2^d) for F-investors. Therefore, separation constrained solely by the credit friction breaks down, as the incentive constraint for F-investors (5) is violated for $(r_1^F, r_2^F) = (r_1^d, r_2^d)$ and $(r_1^P, r_2^P) = (r_1^\delta, r_2^\delta)$.¹⁰

Figure 2 illustrates this case. If both contracts are on the same indifference curve for F-investors, they weakly prefer their own contract, (r_1^d, r_2^d) , over the contracts offered to P-investors, (r_1^P, r_2^P) . Banks make zero-profits with P-investors if (r_1^P, r_2^P) is on the respective zero-profit line. In Figure 2, there are thus two potential contracts, characterized by the intersection of the F-investors' indifference curve and the P-investors' zero-profit line. One contract is to the north-west of (r_1^δ, r_2^δ) , and the other to the south-east. F-investors are indifferent between these two. However, as long as the contract to the south-east of (r_1^δ, r_2^δ) satisfies $r_1^P < r_2^P$, P-investors strictly prefer this one because their marginal rate of substitution between r_1 and r_2 exceeds the rate for F-investors.¹¹

Such equilibrium thus implies an even larger insurance benefit to P-investors relative to the case where F-investors (who aim for flexibility) are absent. Given the incentive constraint of F-investors, (r_1^P, r_2^P) is the best separating contract P-investors can get. Also, a bank cannot profitably enter the market with a pooling contract as a pooling business model does not facilitate contracts that would generate zero profits and be a Pareto-improvement to the two separating contracts, (r_1^d, r_2^d) and (r_1^P, r_2^P) . The following Proposition generalizes these insights.

¹⁰If the P-investors' incentive constraint (6) is violated but not the F-investors' incentive constraint (5), then the P-investors' marginal rate of substitution between r_1 and r_2 cannot exceed the respective rate for F-investors.

¹¹If $r_1^P > r_2^P$, patient P-investors are better off pretending to be impatient and withdraw at date $t = 1$, which renders this contract incentive incompatible.

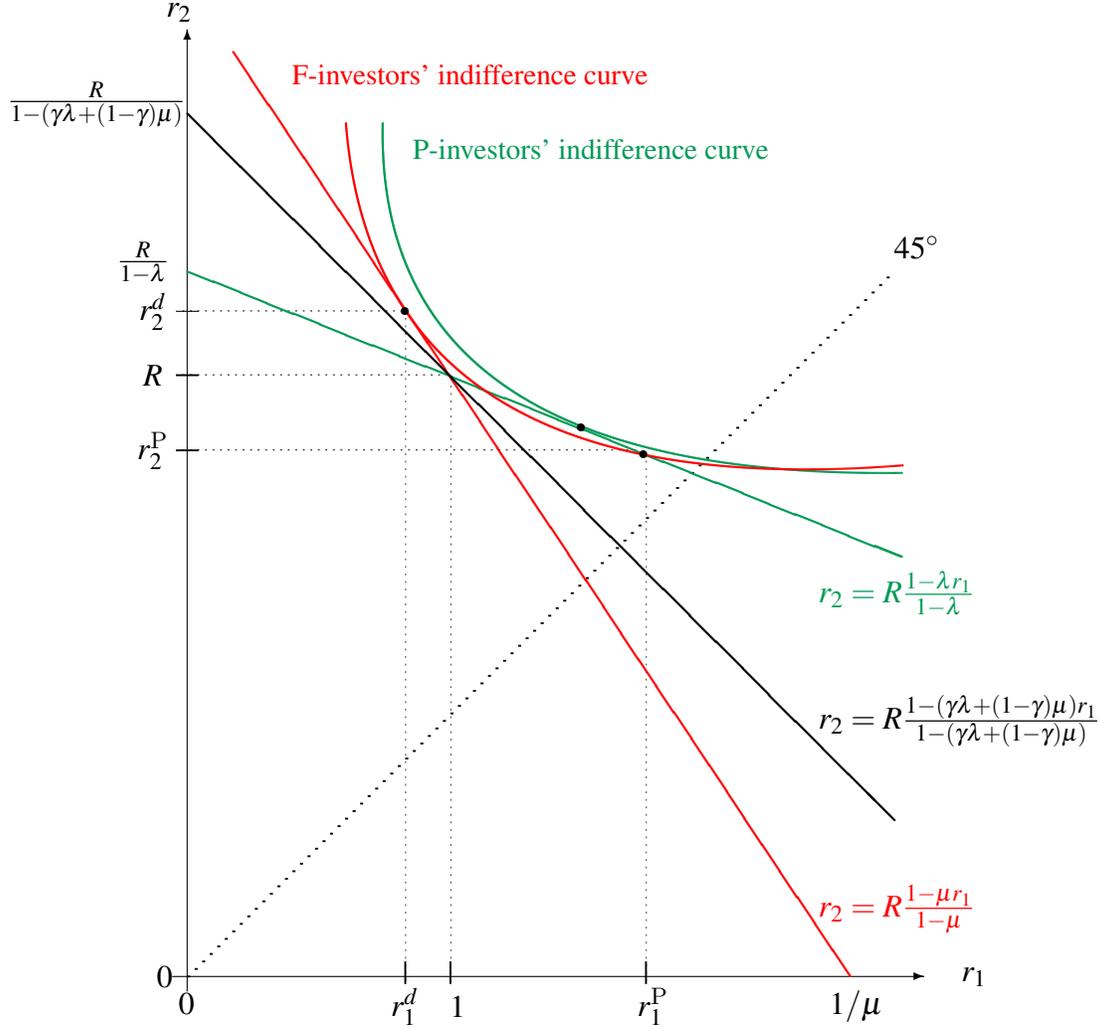


Figure 2: Incentive-constrained Separation with Inflated Insurance for P-investors.

Proposition 2 (Incentive-constrained Separation with Inflated Insurance for P-investors)

Consider economies $\mathcal{E} = (u, \gamma, \lambda, \bar{\mu}, Q, R)$ for which $\bar{\mu}$ is such that $(r_1^\delta, r_2^\delta) \sim_F (r_1^d, r_2^d)$ and $(r_1^\delta, r_2^\delta) \succ_P (r_1^d, r_2^d)$. Under Assumption 2, for each such economy \mathcal{E} there exist $\eta(\mathcal{E}) > 0$ such that there are economies $\mathcal{E}' = (u, \gamma, \lambda, \hat{\mu}, Q, R)$ with $\hat{\mu} \in]\bar{\mu}, \bar{\mu} + \eta(\mathcal{E})[$ where a separating equilibrium obtains in which the F-investors' contract (r_1^F, r_2^F) satisfies $r_1^F = r_1^d$ and $r_2^F = r_2^d$, and the P-investors' contract (r_1^P, r_2^P) satisfies $r_1^P > r_1^\delta$ and $r_2^P < r_2^\delta$.

Proof: See Appendix B. □

The next corollary states an interesting feature of the limits to inflated insurance for P-investors.

Corollary 2 (Populations dominated by P-investors)

The set of probabilities μ of accessing the Q -technology, for which equilibria with inflated liquidity insurance for P-investors obtain, converges to the empty set if the share of P-investors in the population γ approaches one.

Proof: See Appendix C. □

Intuitively, neither the indifference curves nor the zero-profit lines associated with separating contracts depend on the composition of the population, but the slope of the pooling zero-profit line, $r_2 = R \frac{1 - (\gamma\lambda + (1-\gamma)\mu)r_1}{1 - (\gamma\lambda + (1-\gamma)\mu)}$, does (see Figure 2). As γ goes to one, it converges to the zero-profit line for P-investors, $r_2 = R \frac{1 - \lambda r_1}{1 - \lambda}$. Therefore, the pooling zero-profit line eventually intersects a set of contracts enclosed by the P-investors' indifference curve going through (r_1^P, r_2^P) and the zero-profit line associated with P-investors. Their indifference curves being steeper than the F-investors' indifference curves, a set of pooling contract becomes thus available that are Pareto-improvements to the separating contracts (r_1^P, r_2^P) and (r_1^d, r_2^d) . Therefore, for any given μ for which inflated insurance for P-investors is an equilibrium provided $\gamma=0$, there is a $\bar{\gamma} < 1$ such that separating contracts with inflated insurance cannot be an equilibrium for all $\gamma \in]\bar{\gamma}, 1[$. Importantly, for a very large share of P-investors in the population, it is not the optimal contract for them that breaks separating contracts with inflated insurance.

Pooling Suppose the F-investors' marginal rate of substitution between r_1 and r_2 exceeds the respective rate for P-investors. Then, separation cannot exist in equilibrium. To see how, consider first two contracts between which F-investors are just indifferent. Of these two contracts, let one contract satisfy the zero-profit condition associated with P-investors, $r_2 = R(1 - \lambda r_1)/(1 - \lambda)$, and the other the zero-profit condition associated with F-investors, $r_2 = R(1 - \mu r_1)/(1 - \mu)$. Among these two contracts, P-investors then strictly prefer the contract intended for F-investors if and only if the marginal rate of substitution between r_1 and r_2 is higher for F-investors than for P-investors. Conversely, if we consider two contracts between which P-investors are just indifferent, again one contract satisfy-

ing the zero profits with P-investors, $r_2 = R(1 - \lambda r_1)/(1 - \lambda)$, the other zero profits with F-investors, $r_2 = R(1 - \mu r_1)/(1 - \mu)$, then F-investors will prefer the contract intended for P-investors.

While equilibria with separating contracts are, therefore, not possible, equilibria in which banks offer *pooling contracts* may still exist. Such pooling contracts specify identical payment schedules, $\mathcal{D}^F = \mathcal{D}^P = (r_1^{\text{Pool}}, r_2^{\text{Pool}})$, to all investors. Figure 3 illustrates this. Competitive banks with business models associated with pooling contracts offer payments satisfying $r_2 = R \frac{1 - (\gamma\lambda + (1 - \gamma)\mu)r_1}{1 - (\gamma\lambda + (1 - \gamma)\mu)}$, i.e. they are located on the pooling zero-profit line. Consider any contract on that line other than $(1, R)$, for example as in Point A. Given that the F-investors' marginal rate of substitution between r_1 and r_2 exceeds the respective rate for P-investors, there is a contract B such that F-investors are just indifferent between A and B, while P-investors strictly prefer B. Hence, a bank could profitably enter the market by offering contract B, pulling away P-investors from banks offering the pooling contract A. Left with only F-investors as clientele, contract A is no longer sustainable. Therefore, contract A cannot be an equilibrium. In turn, contract B as part of a separating equilibrium is not sustainable either, given the condition for the marginal rates of substitution between r_1 and r_2 . A similar argument can be made for pooling contracts to the north-west of $(1, R)$, ruling out pooling contracts on that upper branch of the pooling zero-profit line.

Next, consider the only remaining contract, $(1, R)$. Any contract on the P-investors' zero-profit line to the south-east of $(1, R)$ would not only make P-investors better off but also F-investors, and any contract on the F-investors' zero-profit line to the north-west of $(1, R)$ would not only make F-investors better off but also P-investors. Therefore, there are no separating contract offers which can break a pooling contract $(1, R)$. Indeed, as long as the slope of the pooling zero-profit line is between the slope of the indifference curve of the P-investors and the slope of the indifference curve of the F-investors, there are no other contracts on the pooling zero-profit line that would be Pareto-improvements to $(1, R)$ and thus attract both types of investors.

The following proposition formalizes these insights.

Proposition 3 (Pooling)

Consider economies $\mathcal{E} = (u, \gamma, \lambda, \mu, Q, R)$ with $\frac{\mu}{1 - \mu} \frac{u'(Qr_1)}{u'(r_2)} Q > \frac{\lambda}{1 - \lambda} \frac{u'(r_1)}{u'(r_2)}$ for all $(r_1, r_2) \in \mathbb{R}_+^2$. If

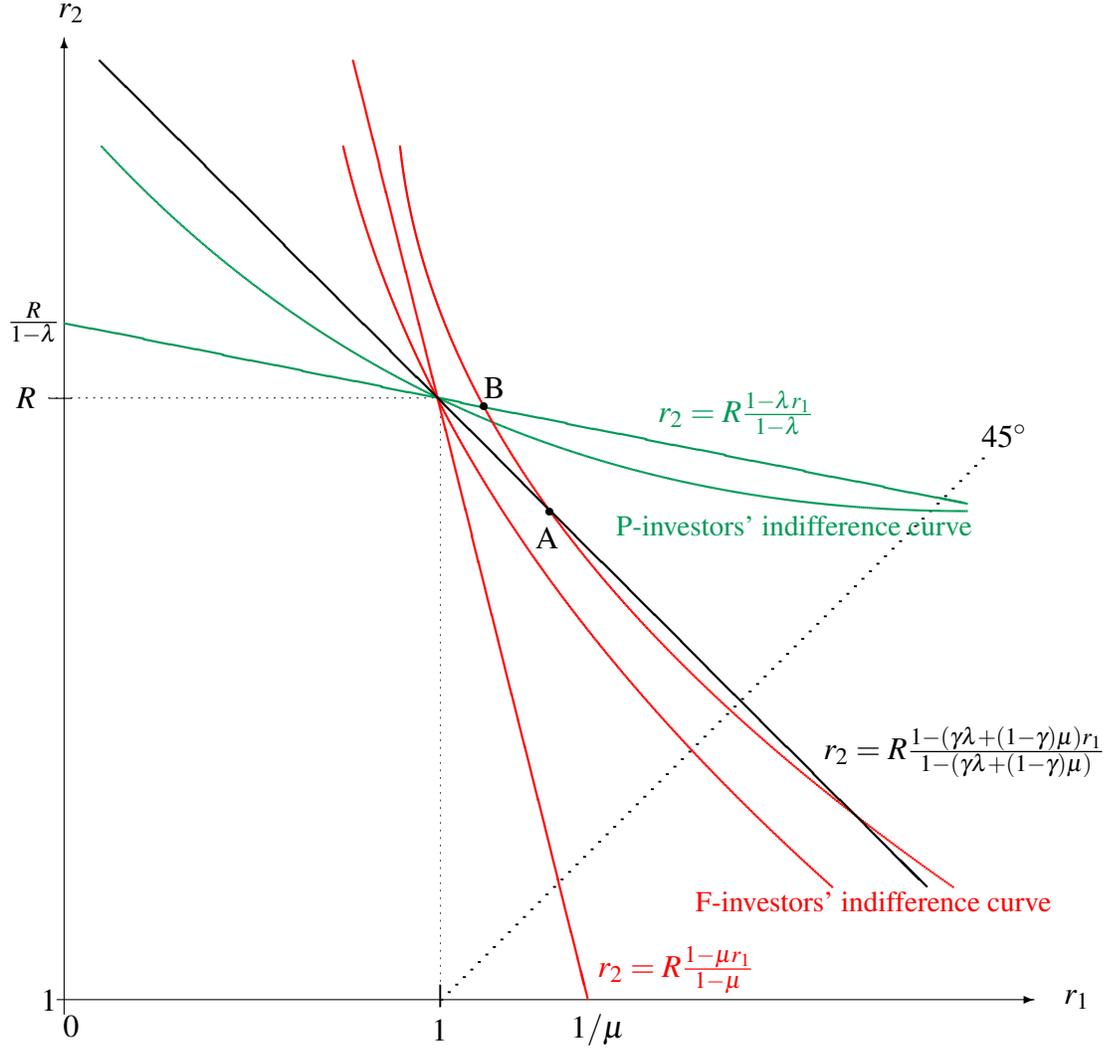


Figure 3: Pooling.

$$\left(\frac{\mu}{1-\mu} u'(Q)Q > \frac{\gamma\lambda+(1-\gamma)\mu}{1-(\gamma\lambda+(1-\gamma)\mu)} u'(R)R > \frac{\lambda}{1-\lambda} u'(1) \right).$$

$\frac{\mu}{1-\mu} u'(Q)Q > \frac{\gamma\lambda+(1-\gamma)\mu}{1-(\gamma\lambda+(1-\gamma)\mu)} u'(R)R > \frac{\lambda}{1-\lambda} u'(1)$ the only equilibrium is a pooling equilibrium. The contract is determined as $(r_1^{\text{Pool}}, r_2^{\text{Pool}}) = (1, R)$.

Proof: See Appendix D. □

Note that a payment schedule $(1, R)$ also obtains in economies without banks but with asset markets.¹² There, all investors choose their own portfolio allocation between storage and the R -technology at date $t=0$, and then trade storage for R -projects in an asset market at date $t=1$ depending on their liquidity needs. For the asset market equilibrium to be arbitrage-free, equilibrium requires that the asset price equals one as only then storage and R -technology generate the same return between dates $t=0$ and $t=1$. With asset prices equal to one, impatient P-investors will sell their R -projects and consume one unit, and patient P-investors will use all their storage to buy R -projects and consume R units of the good. As for F-investors, those with access to the Q -technology will sell their holdings of R -projects and invest one unit in the new opportunity. F-investors without access use their storage to buy additional R -projects. In our setup, the co-existence of different motives for liquidity demand may also lead to a payment schedule $(1, R)$. However, it does so not for any but for some range of parameters and, most importantly, without any trading opportunities between investors.

Non-existence of pure-strategy equilibria To conclude the analysis of possible equilibria, economies can also be such that there is no contract that cannot be dominated by another contract.¹³

Proposition 4 (Non-existence of Equilibrium)

Consider economies $\mathcal{E} = (u, \gamma, \lambda, \mu, Q, R)$ with $\mu > \lambda$ and $\frac{\mu}{1-\mu} \frac{u'(Qr_1)}{u'(r_2)} Q > \frac{\lambda}{1-\lambda} \frac{u'(r_1)}{u'(r_2)}$ for all $(r_1, r_2) \in \mathbb{R}_+^2$. There is no equilibrium in pure strategies, provided the following condition $\frac{\mu}{1-\mu} u'(Q) Q > \frac{\gamma\lambda + (1-\gamma)\mu}{1-(\gamma\lambda + (1-\gamma)\mu)} u'(R) R > \frac{\lambda}{1-\lambda} u'(1)$ is violated.

Proof: See Appendix E. □

Under the conditions of this Proposition there is no viable contract that is not dominated by another contract. Figure 4 illustrates such case. Since the F-investors' marginal rate of substitution between r_1 and r_2 exceeds the respective rate for P-investors, neither separating contracts nor pooling contracts

¹²More generally, a payment schedule $(1, R)$ obtains when investors are allowed to trade directly with each other, be it trading certificates of deposits or lending to and borrowing from each other (Jacklin, 1987; Farhi et al., 2009)

¹³While mixed strategy equilibria will exist when randomization across contracts is allowed for (Dasgupta and Maskin, 1986), we do not pursue this possibility in this paper. By their very nature mixed strategy equilibria will induce added strategic uncertainty, and, hence, instability in market outcomes (see Gehrig and Ritzberger, 2022).

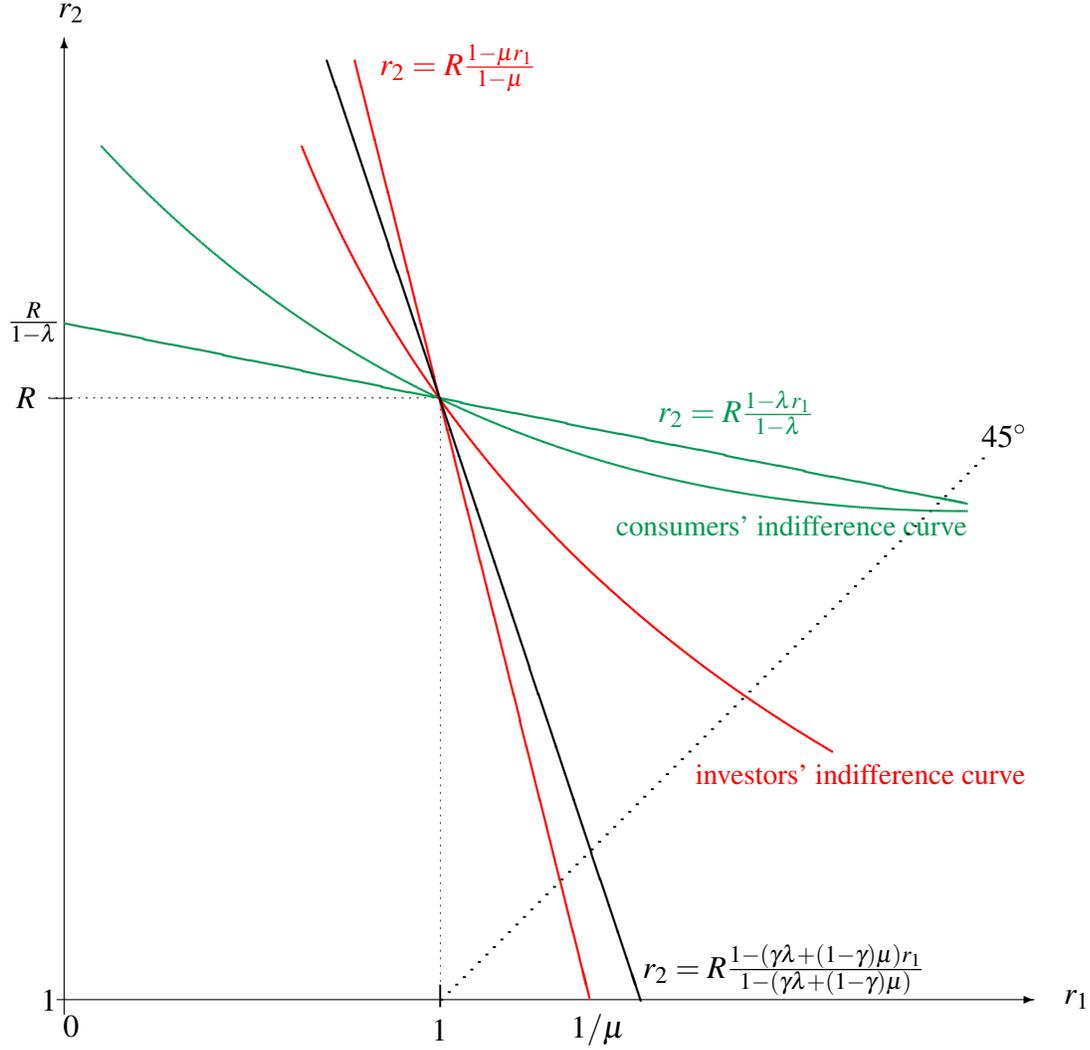


Figure 4: Non-existence of Equilibrium in Pure Strategies
 $\left(\frac{\gamma\lambda + (1-\gamma)\mu}{1 - (\gamma\lambda + (1-\gamma)\mu)} u'(R)R > \frac{\mu}{1-\mu} u'(Q)Q > \frac{\lambda}{1-\lambda} u'(1) \right)$.

other than $(1, R)$ are feasible in equilibrium by the arguments already made above. However, a pooling contract $(1, R)$ cannot be an equilibrium either. To see why, suppose banks were offering a pooling contract $(1, R)$. Then, another bank could profitably enter the market by offering another pooling contract, for there are Pareto-improvements to $(1, R)$ along the pooling zero-profit line — to the north-west of $(1, R)$ in Figure 4. As argued before, those contracts cannot be an equilibrium either given the marginal rates of substitution between r_1 and r_2 .

A pure-strategy choice of equilibrium contracts, i.e. one which does not apply lotteries over contracts, fails to exist here. Therefore, there is no stable market outcome. Interestingly, pure-strategy equilibria do not exist, if the population is highly unbalanced in either direction, i.e. if the proportion of P-investors, γ , is either very close to zero or to unity. The value of γ determines only the slope of the pooling zero-profit line. It converges to the F-investors' zero-profit line for $\gamma \rightarrow 0$ and to the P-investors' zero-profit line for $\gamma \rightarrow 1$. Corollary 3 summarizes the implications for the limiting cases.

Corollary 3 (Unbalanced Populations)

Consider economies $\mathcal{E} = (u, \gamma, \lambda, \mu, Q, R)$ with $\mu > \lambda$ and $\frac{\mu}{1-\mu} \frac{u'(Qr_1)}{u'(r_2)} Q > \frac{\lambda}{1-\lambda} \frac{u'(r_1)}{u'(r_2)}$ for all $(r_1, r_2) \in \mathbb{R}_+^2$. There is no equilibrium in pure strategies if the proportion of P-investors in the population, γ , is either very large or very low.

Proof: See Appendix F. □

Accordingly, for some economies with frequent investment opportunities, the market equilibrium that obtains if there was only a single motive for liquidity demand is not robust. That is, if the relative prevalence of that particular motive in the economy approaches one, there may be no trajectory on which the economy converges to an equilibrium in pure strategies (at least if restricted to equilibria in the spirit of Rothschild and Stiglitz, 1976).

4 Co-existence of liquidity motives without frictions

In this section we study the implications of the co-existence of F-investors and P-investors, and thus of two different motives for liquidity preferences, if for whatever reason, investor types can be differentiated at no cost and loans to F-investors can be fully enforced. It is well-known that intermediaries can deploy technologies to efficiently enforce loan contracts (Diamond, 1984). However, how the ability to enforce loan contracts effects the maturity transformation by banks in the framework of Diamond and Dybvig (1983) has not been explored. Equilibrium allocations for this case are straight forward and we keep it thus concise. Yet, they are again in stark difference to what is known from the

canonical framework. We begin with describing optimal allocations, followed by possible strategies how banks can implement those allocations.

4.1 Allocation

Consider the allocation that is optimal under the economy-wide feasibility constraints. Storage, the long-term R -technology, and the short-term Q -technology are all constant returns to scale. Hence, the Q technology dominates the R -technology in terms of producing consumption goods available at date $t=2$. On the other hand, storage dominates in terms of providing both, early consumption goods and funds for investment in the Q -technology. Accordingly, it is optimal to store all endowments from all investors between dates $t=0$ and $t=1$, and then use this storage to fund the Q -technology and the consumption by impatient P-investors. The returns on the Q -technology will then fund the consumption by F-investors and by patient P-investors.

While the optimal allocation of funds between storage, long-term production, and short-term production is determinate, Pareto-optimal allocations of consumption are indeterminate. For the sake of brevity, we present only the allocation that provides P-investors with the same consumption profile as if F-investors would not exist. At date $t=0$, all endowments are put into storage until date $t=1$. Once the future investment opportunities arrive and uncertainty about consumption needs is resolved, this storage is partly used to provide for impatient P-investors, $c_1^\delta = y^\delta / \lambda$, with y^δ satisfying the first-order condition (4). The remainder of the stored endowments, $1 - \gamma y^\delta > 1 - \gamma$, is invested in the Q -technology. At date $t=2$, the Q -technology will produce $Q(1 - \gamma y^\delta)$, which will be distributed to patient P-investors and all F-investors. Each patient P-investor gets $c_2^\delta = R(1 - y^\delta) / (1 - \lambda)$, leaving for each F-investor an amount of $Q + \frac{\gamma}{1 - \gamma}(Q - R)(1 - y^\delta) > Q$. Therefore, P-investors receive the consumption plan that corresponds to the first-best in case of isolation, and F-investors will be able to consume more than they could by providing for themselves. The reason is that by pooling the endowments of all investors, the comparatively unproductive investment in the R -technology can be avoided, in which P-investors would have to invest if they were left to their own devices. Instead, all goods for consumption at date $t=2$ are produced with the comparatively more productive Q -technology.

4.2 Implementation

Provided banks know the individual motive for their customers' liquidity preference and can fully enforce loan repayments, a competitive banking sector can implement the optimal allocation. To see how, suppose all investors deposit their endowments in banks at date $t=0$. P-investors do so in exchange for a deposit contract which allows them to withdraw $r_1^P = c_1^\delta$ if they get impatient and $r_2^P = c_2^\delta$ if they remain patient. F-investors are granted credit lines to be drawn at a gross interest rate equal to Q at date $t=1$ and receive shares in the bank's equity which allows them to share the value of the bank's assets net of payments to P-investors at date $t=2$, i.e. $r_1^F = 0$ and $r_2^F = Q + \frac{\gamma}{1-\gamma}(Q-R)(1-\lambda r_1^P) > Q$. At the middle date $t=1$, lucky F-investors draw on their credit lines, borrowing all of the banks' remaining storage $1 - \gamma\lambda r_1^P$. At the final date $t=2$, F-investors settle their debt and pay $Q(1 - \gamma\lambda r_1^P)$ to banks. With these earnings, banks pay patient P-investors r_2^P and F-investors $r_2^F = Q + \frac{\gamma}{1-\gamma}(Q-R)(1 - \lambda c_1^P)$. Accordingly, we conclude:

Proposition 5 (Economies of Scope)

The co-existence of a precaution-driven and a flexibility-driven demand for liquidity entails efficiency gains from combining liquidity creation through credit lines with liquidity creation through deposit-taking. Provided banks can distinguish investors by their type and fully enforce loans to F-investors, banks can realize such economies of scope without engaging in maturity transformation.

Proof: Omitted. □

Under such ideal conditions, the business of accepting deposits and simultaneously granting lines of credit is a result of economies of scope.¹⁴ Interestingly, banks would not have to engage in any maturity transformation at all to reap these economies of scope. At date $t=0$, banks issue demand deposits to P-investors and equity shares to F-investors, both backed entirely by stored goods. From date $t=1$ onward, the banks' assets comprise the loans to F-investors and their liabilities are the

¹⁴Note, these economies arise here in absence of any incentive problems at the bank level. In Calomiris and Kahn (1991) and Diamond and Rajan (2001), for example, demand deposits are considered to provide incentives for banks to create value on behalf of their customers.

demand deposits still held by patient P-investors, with F-investors holding the residual claims on the banks' asset returns.

5 Discussion

Diversity of motives versus heterogeneity in risk preferences One may argue that some of our results can also be derived with heterogeneity within a single liquidity motive only. For example, upon inspection of Propositions 1 through 3, the same type of equilibrium phenomena would arise in a world in which all investors want to take precautions against sudden consumption needs but differ substantially with respect to their risk preferences: some have relative risk aversion of $-cu_1''(c)/u_1' < 1$ and others have relative risk aversion of $-cu_2''(c)/u_2' > 1$. Liquidity shocks for the more risk-tolerant investors arise with probability $\lambda_1 = \mu$ and for the more risk-averse investors with probability $\lambda_2 = \lambda$, where μ and λ are as in the main analysis above. In this case, the optimal contract for the more risk-tolerant investor is given by (r_1^1, r_2^1) with a penalty payment for early withdrawal $r_1^1 < 1 < R < r_2^1$, similar to the constrained-efficient F-Investor contract. The contract for the more risk-averse investors (r_1^2, r_2^2) is front-loaded $1 < r_1^2 < r_2^2 < R$ and resembles the efficient P-Investor contract. With such modification, our previous equilibrium analysis of the co-existence of different motives yields results that are identical for one with heterogeneity only in risk preferences.

However, and crucially, such a claim requires that the set of efficient contracts for one type of investors requires a penalty rate ($r_1^1 < 1$), while the other set of efficient contracts requires an insurance benefit ($r_1^2 > 1$). Therefore, the similarity in results obtains only if there is one group of sufficiently risk-tolerant investors (with $-cu_1''(c)/u_1' < 1$) and another group of quite risk-averse investors (with $-cu_1''(c)/u_1' > 1$). In other words, heterogeneity in risk-preferences alone will not be enough to generate all our phenomena within a Diamond and Dybvig (1983) model. Specifically, they do not obtain if either $-cu''(c)/u' > 1$ for all investors or $-cu_1''(c)/u_1' < 1$ for all investors.

Another distinguishing difference between the co-existence of different risk-preferences and the co-existence of different motives refers to the case without financial frictions. In the absence of investors who desire to preserve their flexibility for future, better investment opportunities, the potential

for economies of scope no longer exists. Therefore, with only precautionary investors of different risk attitudes in place, welfare will be much reduced (and banks' maturity transformation much stronger) relative to the case with an investment motive as an independent source of liquidity demand.

Low interest rate environment Since the long-term production generates safe returns R , they can be expected to be linked to the return on long-term government debt. How would equilibrium be affected in a low interest rate environment, i.e. if the long-term rate R converges to one? It is readily verified that in such an environment the deposit contract converges to a contract merely repaying P-investors their initial endowment, i.e. $\lim_{R \rightarrow 1} (r_1^\delta, r_2^\delta) = (1, 1)$. In other words, taking precautions loses relevance as a motive for liquidity demand, while preserving flexibility remains active since $Q > 1$.¹⁵

Interestingly, a low interest rate environment can contribute to instability as equilibria in pure strategies may cease to exist when the returns on the long-term production fall. The following example illustrates this. Suppose that the initial return with long-term production is $R=R_0$, and that for this value a pooling equilibrium just obtains, i.e. there is a small $\varepsilon > 0$ such that

$$\frac{\mu}{1-\mu} u'(Q)Q - \varepsilon = \frac{\gamma\lambda + (1-\gamma)\mu}{1-(\gamma\lambda + (1-\gamma)\mu)} u'(R_0)R_0 > \frac{\lambda}{1-\lambda} u'(1).$$

Suppose next that the return on the long-term technology, R , falls to one. For relative risk aversion larger one we obtain $\frac{d}{dR} (u'(R)R) < 0$, with $\lim_{R \rightarrow 1} u'(R)R = u'(1) > u'(Q)Q$. Therefore, a fall of R to one will lead to a violation of $\frac{\mu}{1-\mu} u'(Q)Q > \frac{\gamma\lambda + (1-\gamma)\mu}{1-(\gamma\lambda + (1-\gamma)\mu)} u'(R)R$ for sufficiently small ε , i.e. a pooling equilibrium, which exists and is the only equilibrium for $R=R_0$, ultimately fails to exist as $R \rightarrow 1$. In other words, a decrease in the long-term interest rate, as measured by R , increases the range of unstable outcomes. Clearly, this type of instability will not arise in a world with only a single motive for liquidity demand.

Bank regulation Our analysis so far can provide no reason why banks catering to both liquidity preferences are better (or worse) than banks specializing on one type of liquidity preference. But this

¹⁵Again, the same cannot be generally said if liquidity preferences arise only from a desire to take precautions with heterogeneity in risk aversion.

is due to our equilibrium concept as it restricts each bank to offer only contracts that do not generate losses. A case for, or against, in-house pooling of business lines can be made if we instead consider competitive equilibria where each individual bank can offer a *menu of cross-subsidizing contracts* (in the spirit of Miyazaki, 1977; Wilson, 1977; Spence, 1978). Then, should cross-subsidizing contracts prevail in equilibrium over banks offering only loss-free contracts, the two different motives could be expected to be served together. However, as we show in Dietrich and Gehrig (2022), banks offering menus of cross-subsidizing contracts do not prevail in equilibrium. Importantly, the possibility to offer menus of cross-subsidizing contracts can render equilibria with single-contract offers impossible such that no equilibrium exists at all. Accordingly, this suggests a reason for regulators to stop the practice of cross-subsidizing lines of business within a bank. This does not necessarily require the separation of ownership into several banking units but it does require to treat, and manage, separate business models separately from an organizational point of view. Accordingly, allowing bank shareholders to insist on the same, highest rate of return across all divisions within a multi-product bank, may be counter-productive.

Institutional indeterminacy Although we have made reference to deposits and equity shares as the contractual means that serve the interests of a specific type of investors, it should be clarified that the use of these terms is primarily to keep the presentation simple. We associate differences in contracts only with differences in the sequence of payments. Other, undoubtedly important features of contracts, such as being negotiable or tradable, and constraints often associated with certain contracts, such as sequential service, are not the focus of attention in this paper.

For example, that banks can refinance themselves entirely with equity in absence of frictions (see Section 4) is only one of many contractual solutions. What matters here is that banks do not engage in maturity transformation with either institutional arrangement. It is commonly understood that without frictions, little can be said about the specific institutional arrangement which maintains an allocation. This also includes whether banks set themselves up as universal banks, integrating several business lines via an internal capital market, or as separate entities which use an interbank loan market. Either arrangement will achieve the Pareto-optimal allocation in the absence of frictions. With an interbank

market, both banks store the endowments of their customers in the first period. After the first period, the bank specializing on P-investors grants the remainder of their stored goods $1 - \lambda c_1^\delta$ as a loan to the other bank for one period. If the interest rate on such interbank loans is R , the bank specializing on P-investors will pay its patient P-investors c_2^δ . The banks specializing in F-investors can grant its own stored funds, along with the funds borrowed from the other bank, as a loan to its lucky F-investors and thus implements the efficient allocation.

6 Concluding remarks

Our analysis reveals that the nature of liquidity preferences crucially matters for competitive market outcomes, particularly when different motives co-exist. In a simple framework we have shown that the co-existence of liquidity preferences based on a desire to take precautions with liquidity preferences based on a desire to preserve flexibility has the potential to benefit from economies of scope in a world without standard financial frictions. In the probably more likely scenario with frictions, the precise nature of these frictions as well as their interplay will affect the nature of market outcomes, and, therefore, potential policy implications. For example, focusing on one motive and one friction in isolation is likely to direct the policy debate towards bank runs, even for constellations, when their occurrence may not be likely because maturity transformation does not take place in equilibrium.

References

- Berger, A. N. and C. H. S. Bouwman (2009). Bank liquidity creation. *The Review of Financial Studies* 22(9), 3779–3837.
- Bhattacharya, S. and A. J. Padilla (1996). Dynamic Banking: A Reconsideration. *The Review of Financial Studies* 9(3), 1003–1032.
- Bryant, J. (1980). A model of reserves, bank runs, and deposit insurance. *Journal of Banking & Finance* 4(4), 335 – 344.

- Calomiris, C. W. and C. M. Kahn (1991). The role of demandable debt in structuring optimal banking arrangements. *The American Economic Review* 81(3), 497–513.
- Dasgupta, P. and E. Maskin (1986). Existence of equilibrium in discontinuous economies I: Theory. *Review of Economic Studies* 53(1), 1–25.
- Diamond, D. W. (1984). Financial intermediation and delegated monitoring. *The Review of Economic Studies* 51(3), 393–414.
- Diamond, D. W. and P. H. Dybvig (1983). Bank runs, deposit insurance, and liquidity. *Journal of Political Economy* 91(3), 401–419.
- Diamond, D. W. and R. G. Rajan (2001). Liquidity risk, liquidity creation, and financial fragility: A theory of banking. *Journal of Political Economy* 109(2), 287–327.
- Dietrich, D. and T. Gehrig (2021). On the instability of private liquidity provision. *Economics Letters* 209, 110117.
- Dietrich, D. and T. Gehrig (2022). Speculative and Precautionary Demand for Liquidity in Competitive Banking Markets. Centre for Economic Policy Research discussion paper, CEPR 15827.
- Dietrich, D., J. K. Shin, and M. Tvede (2020). Debt constraints and monetary policy. *Journal of Mathematical Economics* 87, 31 – 42.
- Farhi, E., M. Golosov, and A. Tsyvinski (2009). A theory of liquidity and regulation of financial intermediation. *The Review of Economic Studies* 76(3), 973–992.
- Geanakoplos, J. and K. J. Walsh (2018). Inefficient liquidity provision. *Economic Theory* 66, 213–233.
- Gehrig, T. and M. Jackson (1998). Bid-ask spreads with indirect competition among specialists. *Journal of Financial Markets* 1(1), 89 – 119.
- Gehrig, T. and K. Ritzberger (2022). Intermediation and price volatility. *Journal of Economic Theory* 201, 105442.

- Hart, O. and J. Moore (1994). A theory of debt based on the inalienability of human capital. *The Quarterly Journal of Economics* 109(4), 841–879.
- Holmström, B. and J. Tirole (1998). Private and public supply of liquidity. *Journal of Political Economy* 106(1), 1–40.
- Jacklin, C. J. (1987). Demand deposits, trading restrictions, and risk sharing. In E. D. Prescott and N. Wallace (Eds.), *Contractual Arrangements for Intertemporal Trade*, pp. 26–47. Minneapolis: University of Minnesota Press.
- Jones, R. and J. Ostroy (1976). Liquidity as flexibility. Discussion Papers no, UCLA.
- Jones, R. A. and J. M. Ostroy (1984). Flexibility and uncertainty. *The Review of Economic Studies* 51(4), 13–32.
- Kashyap, A. K., R. Rajan, and J. C. Stein (2002). Banks as liquidity providers: An explanation for the coexistence of lending and deposit-taking. *The Journal of Finance* 57(1), 33–73.
- Kehoe, T. J. and D. K. Levine (1993). Debt-constrained asset markets. *The Review of Economic Studies* 60(4), 865–888.
- Martin, A. and J. Ventura (2012). Economic growth with bubbles. *American Economic Review* 102(6), 3033–58.
- Miyazaki, H. (1977, Autumn). The Rat Race and Internal Labor Markets. *Bell Journal of Economics* 8(2), 394–418.
- Pennacchi, G. and J. Santos (2021). Why do banks target ROE? *Journal of Financial Stability* 54, 110856.
- Rothschild, M. and J. Stiglitz (1976). Equilibrium in insurance markets: An essay on the economics of imperfect information. *Quarterly Journal of Economics* 90, 629–649.
- Smith, A. (1776). *An Inquiry into the Nature and Causes of the Wealth of Nations*. London.

Spence, M. (1978). Product differentiation and performance in insurance markets. *Journal of Public Economics* 10(3), 427–447.

Wallace, N. (1988). Another attempt to explain an illiquid banking system: the Diamond and Dybvig model with sequential service taken seriously. *Federal Reserve Bank of Minneapolis Quarterly Review* 12(4), 3–16.

Wilson, C. (1977). A model of insurance markets with incomplete information. *Journal of Economic Theory* 16(2), 167–207.

Appendix

A Proof of Proposition 1

The proof is by establishing six claims consecutively.

Claim 1: (r_1^d, r_2^d) and (r_1^δ, r_2^δ) satisfy the participation constraints for F-investors and P-investors, respectively.

The participation constraints are satisfied with strict inequality:

- For P-investors:

$$\begin{aligned} \lambda u(r_1^\delta) + (1 - \lambda)u\left(\frac{R(1 - \lambda r_1^\delta)}{1 - \lambda}\right) \\ > \lambda u(1) + (1 - \lambda)u(R) \\ > \sup\{\lambda u(y) + (1 - \lambda)u(R(1 - y) + y) \mid y \in [0, 1]\} \end{aligned} \quad (12)$$

The first inequality obtains since $r_1^\delta \in \arg \max\{\lambda u(r_1) + (1 - \lambda)u\left(\frac{R(1 - \lambda r_1)}{1 - \lambda}\right) \mid r_1 \in [0, \lambda^{-1}]\}$.

The second inequality obtains since $R > 1$ implies for all $y \in [0, 1]$ that $u(1) \geq u(y)$ and $u(R) \geq u(R(1 - y) + y)$, with $u(1) > u(y)$ for $y < 1$ and $u(R) > u(R(1 - y) + y)$ for $y > 0$.

- For F-investors:

$$\begin{aligned} \lambda u(Qr_1^d) + (1 - \lambda)u\left(\frac{R(1 - \lambda r_1^d)}{1 - \lambda}\right) \\ > \lambda u(Q) + (1 - \lambda)u(R) \\ > \sup\{\mu u(Qy + R(1 - y)) + (1 - \mu)u(R(1 - y) + y) \mid y \in [0, 1]\} \end{aligned} \quad (13)$$

The first inequality obtains since $r_1^d \in \arg \max\{\lambda u(Qr_1) + (1 - \lambda)u\left(\frac{R(1 - \lambda r_1)}{1 - \lambda}\right) \mid r_1 \in [0, \lambda^{-1}]\}$.

The second inequality obtains since $Q > R > 1$ implies for all $y \in [0, 1]$ that $u(Q) \geq u(Qy + R(1 - y))$ and $u(R) \geq u(R(1 - y) + y)$, with $u(Q) > u(Qy + R(1 - y))$ for $y < 1$ and $u(R) > u(R(1 - y) + y)$ for $y > 0$.

Claim 2: $(r_1^d, r_2^d) \succ_I (r_1^\delta, r_2^\delta)$ and $(r_1^\delta, r_2^\delta) \succ_C (r_1^d, r_2^d)$ for all $\lambda \geq \mu$.

1. For $\mu = \lambda$, the incentive constraints are satisfied for all investors:

- F-investors: $\lambda u(Qr_1^d) + (1 - \lambda)u\left(\frac{R(1 - \lambda r_1^d)}{1 - \lambda}\right) \geq \lambda u(Qr_1^\delta) + (1 - \lambda)u\left(\frac{R(1 - \lambda r_1^\delta)}{1 - \lambda}\right)$ for all $r_1^\delta \in [0, \lambda^{-1}]$, with strict inequality if $-\frac{u''(c)}{u'(c)}c \neq 1$, since $r_1^d \in \arg \max\{\lambda u(Qr_1) + (1 - \lambda)u\left(\frac{R(1 - \lambda r_1)}{1 - \lambda}\right) \mid r_1 \in [0, \lambda^{-1}]\}$.
- P-investors: $\lambda u(r_1^\delta) + (1 - \lambda)u\left(\frac{R(1 - \lambda r_1^\delta)}{1 - \lambda}\right) \geq \lambda u(r_1^d) + (1 - \lambda)u\left(\frac{R(1 - \lambda r_1^d)}{1 - \lambda}\right)$ for all $r_1^d \in [0, \lambda^{-1}]$, with strict inequality if $-\frac{u''(c)}{u'(c)}c \neq 1$, since $r_1^\delta \in \arg \max\{\lambda u(r_1) + (1 - \lambda)u\left(\frac{R(1 - \lambda r_1)}{1 - \lambda}\right) \mid r_1 \in [0, \lambda^{-1}]\}$.

Therefore, $(r_1^d, r_2^d) \succ_F (r_1^\delta, r_2^\delta)$ and $(r_1^\delta, r_2^\delta) \succ_P (r_1^d, r_2^d)$ for $\lambda = \mu$ and $-\frac{u''(c)}{u'(c)}c > 1$.

2. For $\lambda > \mu$, it suffices to consider the incentive constraints for F-investors and P-investors, respectively, letting λ increase for a given μ , starting from $\lambda = \mu$.

- F-investors: The LHS of condition (5) is not affected by changes in λ . Hence, the total effect on the differential of expected utilities is positive as the RHS of condition (5) changes according to

$$\frac{d}{d\lambda} \left(\mu u(Qr_1^\delta) + (1 - \mu)u(r_2^\delta) \right) = \mu Q u'(Qr_1^\delta) \frac{dr_1^\delta}{d\lambda} + (1 - \mu)u'(r_2^\delta) \frac{dr_2^\delta}{d\lambda} < 0 \quad (14)$$

as $\frac{dr_1^\delta}{d\lambda}, \frac{dr_2^\delta}{d\lambda} < 0$. The latter follows from applying the implicit function theorem to the P-investors' first-order condition (4). If written as

$$u'(r_1^\delta) - R u' \left(\frac{R(1 - \lambda r_1^\delta)}{1 - \lambda} \right) = 0$$

we have

$$\frac{dr_1^\delta}{d\lambda} = - \frac{R^2 u''(r_2^\delta) \frac{r_1^\delta - 1}{(1 - \lambda)^2}}{u''(r_1^\delta) + R^2 u''(r_2^\delta) \frac{\lambda}{1 - \lambda}} < 0,$$

and if written as

$$u' \left(\frac{R - (1 - \lambda)r_2^\delta}{\lambda R} \right) - R u'(r_2^\delta) = 0$$

we have

$$\frac{dr_2^\delta}{d\lambda} = -\frac{-u''(r_1^\delta)\frac{R-r_2^\delta}{\lambda^2 R}}{-u''(r_1^\delta)\frac{1-\lambda}{\lambda R} - Ru''(r_2^\delta)} < 0.$$

- P-investors: By the Envelope theorem, the LHS in condition (6) changes in response to increases in λ by $u(r_1^\delta) - u(r_2^\delta)$. The RHS in condition (6) changes in response to increases in λ by $u(r_1^d) - u(r_2^d)$. Hence, the total effect on the differential of expected utilities is $\left(u(r_1^\delta) - u(r_1^d)\right) - \left(u(r_2^\delta) - u(r_2^d)\right)$ which is positive since $r_1^\delta > r_1^d$ and $r_2^\delta < r_2^d$.

Claim 3: There is $\tilde{\mu} > \lambda$ such that $(r_1^\delta, r_2^\delta) \succ_F (r_1^d, r_2^d)$ for all $\mu \in]\tilde{\mu}, 1[$ and $(r_1^d, r_2^d) \succ_F (r_1^\delta, r_2^\delta)$ for all $\mu \in]0, \tilde{\mu}[$.

From the F-investors' first-order condition (2), we obtain $\lim_{\mu \rightarrow 1} y^d = \lim_{\mu \rightarrow 1} r_1^d = 1$. The LHS of condition (5) converges to $u(Q)$ and the RHS to $u(Qr_1^\delta) > u(Q)$ since $r_1^\delta > 1$. By the intermediate value theorem, there is thus $\tilde{\mu} > \lambda$ such that $(r_1^\delta, r_2^\delta) \succ_F (r_1^d, r_2^d)$ for all $\mu \in]\tilde{\mu}, 1[$. Since the utility differential $Z_I = (\mu u(Qr_1^d) + (1-\mu)u(r_2^d)) - (\mu u(Qr_1^\delta) + (1-\mu)u(r_2^\delta))$ is monotone in μ with $dZ_I/d\mu = \left(u(Qr_1^d) - u(Qr_1^\delta)\right) - \left(u(r_2^d) - u(r_2^\delta)\right) < 0$, the claim is established.

Claim 4: If Q is large, and λ small, there is $\hat{\mu} \in]\lambda, 1[$ such that $(r_1^d, r_2^d) \succ_P (r_1^\delta, r_2^\delta)$ for all $\mu \in]\hat{\mu}, 1[$ and $(r_1^\delta, r_2^\delta) \succ_P (r_1^d, r_2^d)$ for all $\mu \in]0, \hat{\mu}[$.

From the F-investors' first-order condition (2), we obtain $\lim_{\mu \rightarrow 1} y^d = \lim_{\mu \rightarrow 1} r_1^d = 1$. Therefore, $(r_1^d, r_2^d) \succ_P (r_1^\delta, r_2^\delta)$ holds for $\mu \rightarrow 1$ provided

$$\lambda u(1) + (1-\lambda)u(r_2^d) > \lambda u(r_1^\delta) + (1-\lambda)u(r_2^\delta). \quad (15)$$

The P-investors' contract (r_1^δ, r_2^δ) does not depend on μ or Q . The first-order condition (2), determining the F-investors' contract (r_1^d, r_2^d) , implies $dr_2^d/dQ > 0$ for all μ if $-cu''(c)/u'(c) > 1$. Hence, condition (15) is more likely to hold if Q is large or λ is small.

The utility differential $Z_P = (\lambda u(r_1^\delta) + (1-\lambda)u(r_2^\delta)) - (\lambda u(r_1^d) + (1-\lambda)u(r_2^d))$ is monotone in μ with

$$\frac{dZ_P}{d\mu} = -\left(\lambda u'(r_1^d)\frac{dr_1^d}{d\mu} + (1-\lambda)u'(r_2^d)\frac{dr_2^d}{d\mu}\right) < 0 \quad (16)$$

as $\frac{dr_1^d}{d\mu}, \frac{dr_2^d}{d\mu} > 0$. The latter follows from applying the implicit function theorem to the F-investors' first-order condition (2). If written as

$$Qu'(r_1^d) - Ru' \left(\frac{R(1 - \mu r_1^\delta)}{1 - \mu} \right) = 0$$

we have

$$\frac{dr_1^d}{d\mu} = - \frac{-R^2 u''(r_2^d) \frac{1-r_1^d}{(1-\mu)^2}}{Q^2 u''(r_1^d) + R^2 u''(r_2^d) \frac{\mu}{1-\mu}} > 0,$$

and if written as

$$Qu' \left(Q \frac{R - (1 - \mu)r_2^d}{\mu R} \right) - Ru'(r_2^d) = 0$$

we have

$$\frac{dr_2^d}{d\mu} = - \frac{u''(Qr_1^d) Q^2 \frac{r_2^d - R}{\mu^2 R}}{-u''(Qr_1^d) Q^2 \frac{1-\mu}{\mu R} - Ru''(r_2^d)} > 0.$$

By the intermediate value theorem, there is thus $\hat{\mu} \in]\lambda, 1[$ such that $(r_1^d, r_2^d) \succ_P (r_1^\delta, r_2^\delta)$ for all $\mu \in]\hat{\mu}, 1[$ and $(r_1^\delta, r_2^\delta) \succ_P (r_1^d, r_2^d)$ for all $\mu \in]0, \hat{\mu}[$ if (15) holds and the claim is established. If (15) does not hold, then $(r_1^\delta, r_2^\delta) \succ_P (r_1^d, r_2^d)$ for all $\mu \in]0, 1[$.

Claim 5 For $\mu \in]0, \bar{\mu}]$, there is no pooling contract which is a Pareto-improvement to $((r_1^d, r_2^d), (r_1^\delta, r_2^\delta))$.

The slope of the zero-profit constraint for the pooling contract is between the slopes of the two zero-profit constraints associated with a separating equilibrium. A necessary condition for a pooling contract, which lies on the pooling zero-profit line, to make P-investors better off than (r_1^δ, r_2^δ) is thus that $r_1 < 1$, while a necessary condition for a pooling contract to make F-investors better off than (r_1^d, r_2^d) is that $r_1 > 1$. As these two condition rule each other out, there is no Pareto-improvement through pooling.

By claims 1 through 5, there is $\bar{\mu} = \min \left\{ \mu \in]\lambda, 1[\mid (r_1^\delta, r_2^\delta) \succ_F (r_1^d, r_2^d) \wedge (r_1^d, r_2^d) \succ_P (r_1^\delta, r_2^\delta) \right\}$ such that $(r_1^d, r_2^d) \succ_F (r_1^\delta, r_2^\delta)$ and $(r_1^\delta, r_2^\delta) \succ_P (r_1^d, r_2^d)$ if and only if $\mu \in]0, \bar{\mu}[$.

Claim 6 $-\frac{\mu}{1-\mu} \frac{u'(Qr_1^d)}{u'(r_2^d)} Q > -\frac{\lambda}{1-\lambda} \frac{u'(r_1^d)}{u'(r_2^d)}$ and $-\frac{\mu}{1-\mu} \frac{u'(Qr_1^\delta)}{u'(r_2^\delta)} Q > -\frac{\lambda}{1-\lambda} \frac{u'(r_1^\delta)}{u'(r_2^\delta)}$ obtains for all $\mu \in]0, \bar{\mu}]$.

The proof is by contradiction. Suppose μ is such that an equilibrium with credit-constrained separation exists, i.e. $(r_1^d, r_2^d) \succsim_F (r_1^\delta, r_2^\delta)$ and $(r_1^\delta, r_2^\delta) \succsim_P (r_1^d, r_2^d)$. If either $-\frac{\mu}{1-\mu} \frac{u'(Qr_1^d)}{u'(r_2^d)} Q > -\frac{\lambda}{1-\lambda} \frac{u'(r_1^d)}{u'(r_2^d)}$ or $-\frac{\mu}{1-\mu} \frac{u'(Qr_1^\delta)}{u'(r_2^\delta)} Q > -\frac{\lambda}{1-\lambda} \frac{u'(r_1^\delta)}{u'(r_2^\delta)}$ would not hold, Assumption 2 implies that either $(r_1^d, r_2^d) \succ_P (r_1^\delta, r_2^\delta)$, or $(r_1^\delta, r_2^\delta) \succ_F (r_1^d, r_2^d)$, or both, would necessarily hold.

B Proof of Proposition 2

For any given (r_1, r_2) , the slope of the F-investors' indifference curve is

$$\frac{dr_2}{dr_1} = -\frac{\mu}{1-\mu} \frac{u'(Qr_1)}{u'(r_2)} Q$$

and the slope of the P-investors' indifference curve is

$$\frac{dr_2}{dr_1} = -\frac{\lambda}{1-\lambda} \frac{u'(r_1)}{u'(r_2)}.$$

By Assumption 2, if $(r_1^\delta, r_2^\delta) \sim_F (r_1^d, r_2^d)$ and $(r_1^\delta, r_2^\delta) \succ_P (r_1^d, r_2^d)$, the P-investors' indifference curve is steeper than the F-investors' indifference curve at $r_1 = r_1^\delta$ and $r_2 = r_2^\delta$, i. e.

$$-\frac{\bar{\mu}}{1-\bar{\mu}} \frac{u'(Qr_1^\delta)}{u'(r_2^\delta)} Q > -\frac{\lambda}{1-\lambda} \frac{u'(r_1^\delta)}{u'(r_2^\delta)}. \quad (17)$$

which together with (4) implies

$$-\frac{\bar{\mu}}{1-\bar{\mu}} \frac{u'(Qr_1^\delta)}{u'(r_2^\delta)} Q > -\frac{\lambda}{1-\lambda} R. \quad (18)$$

Let Z be defined by

$$Z := (\mu u(Qr_1^d) + (1-\mu)u(r_2^d)) - (\mu u(Qr_1) + (1-\mu)u(r_2))$$

with $r_1^d = y^d/\mu$, $r_2^d = R(1 - y^d)/(1 - \mu)$, and y^d solves (2). By definition, $\mu = \bar{\mu}$ implies $Z = 0$ for $r_1 = r_1^\delta = y^\delta/\lambda$ and $r_2 = r_2^\delta = R(1 - y^\delta)/(1 - \lambda)$, with y^δ solving (4). Concavity of u thus implies that there is (r'_1, r'_2) with $r'_1 < r_1^\delta = y^\delta/\lambda$ and $r'_2 > r_2^\delta = R(1 - y^\delta)/(1 - \lambda)$, which are also feasible as they satisfy $r'_2 = \frac{R(1 - \lambda r'_1)}{1 - \lambda}$, and for which $Z = 0$ also holds. However, since (r_1^δ, r_2^δ) maximizes the P-investors' expected utility subject only to their zero-profit constraint, $(r_1^\delta, r_2^\delta) \succ_P (r'_1, r'_2)$. Hence, in response to a marginal increase in μ , starting from $\bar{\mu}$, P-investors strictly prefer a marginal adjustment to a contract (r_1^δ, r_2^δ) over a marginal adjustment to a contract (r'_1, r'_2) . Therefore, banks offering marginal adjustment to a contract (r_1^δ, r_2^δ) will prevent other banks offering marginal adjustments to (r'_1, r'_2) from entering the market, even though both satisfy the F-investors' incentive constraint $Z = 0$.

Applying the implicit function theorem to $Z = 0$ we obtain $dr_2/d\mu = -(dZ/d\mu)/(dZ/dr_2)$ with

$$\frac{dZ}{d\mu} = u(Qr_1^d) - u(Qr_1) + u(r_2) - u(r_2^d), \quad (19)$$

$$\frac{dZ}{dr_2} = \mu \frac{Q}{R} u'(Qr_1) \frac{1 - \lambda}{\lambda} - (1 - \mu) u'(r_2). \quad (20)$$

Equation (19) follows by taking into account the Envelope theorem, according to which the effects of changes in y^d , induced by changes in μ , have no effect as the first-order condition (2) applies. Equation (20) follows by taking into account the zero-profit constraints, according to which $r_1 = (R - (1 - \lambda)r_2)(\lambda R)^{-1}$. Evaluating (19) at $r_1 = r_1^\delta$ and $r_2 = r_2^\delta$ yields $dZ/d\mu < 0$ because $r_2^d > r_2^\delta$ and $r_1^d < r_1^\delta$. Evaluating (20) at $r_1 = r_1^\delta$ and $r_2 = r_2^\delta$ yields $dZ/dr_2 < 0$ because of (18). Hence, $dr_2/d\mu < 0$ and $dr_1/d\mu = -((1 - \lambda)/\lambda R)(dr_2/d\mu) > 0$. By continuity, the result also applies to all $\mu > \bar{\mu}$ in some neighborhood of $\bar{\mu}$. Therefore, the Proposition obtains.

C Proof of Corollary 2

A necessary condition for incentive-constrained separation equilibria to exist with inflated insurance for P-investors is $-\frac{\mu}{1 - \mu} \frac{u'(Qr_1)}{u'(r_2)} Q > -\frac{\lambda}{1 - \lambda} \frac{u'(r_1)}{u'(r_2)}$ for all $(r_1, r_2) \in \mathbb{R}_+^2$. Consider two contracts (r_1^A, r_2^A) and (r_1^B, r_2^B) such that $(r_1^A, r_2^A) \sim_F (r_1^B, r_2^B) \sim_F (r_1^d, r_2^d)$, $r_2^A = R(1 - \lambda r_1^A)/(1 - \lambda)$, and $r_2^B =$

$R(1-\lambda r_1^B)/(1-\lambda)$. As

$$\lim_{\gamma \rightarrow 1} R \frac{1 - (\gamma\lambda + (1-\gamma)\mu) r_1}{1 - (\gamma\lambda + (1-\gamma)\mu)} = R \frac{1 - \lambda r_1}{1 - \lambda}.$$

pooling contracts $(r_1^{\text{Pool}}, r_2^{\text{Pool}})$ exist with $r_1^{\text{Pool}} \in]r_1^A, r_1^B[$ and $r_2^{\text{Pool}} = R \frac{1 - (\gamma\lambda + (1-\gamma)\mu) r_1^{\text{Pool}}}{1 - (\gamma\lambda + (1-\gamma)\mu)}$ such that $(r_1^{\text{Pool}}, r_2^{\text{Pool}}) \succ_F (r_1^d, r_2^d)$ and $(r_1^{\text{Pool}}, r_2^{\text{Pool}}) \succ_P (r_1^A, r_2^A)$ as well as $(r_1^{\text{Pool}}, r_2^{\text{Pool}}) \succ_P (r_1^B, r_2^B)$.

D Proof of Proposition 3

The proof proceeds in five steps.

Step 1: Under the condition of Proposition 3, when $-\frac{\mu}{1-\mu} \frac{u'(Qr_1)}{u'(r_2)} Q < -\frac{\lambda}{1-\lambda} \frac{u'(r_1)}{u'(r_2)}$ for all (r_1, r_2) , we have $(r_1^\delta, r_2^\delta) \succ_F (r_1^d, r_2^d)$ and $(r_1^\delta, r_2^\delta) \succ_P (r_1^d, r_2^d)$. Accordingly, credit-constrained separation cannot be an equilibrium.

Step 2: Pooling equilibria satisfy the pooling zero-profit constraint, $r_2 = R \frac{1 - (\gamma\lambda + (1-\gamma)\mu) r_1}{1 - (\gamma\lambda + (1-\gamma)\mu)}$. The associated zero-profit line intersect with the zero-profit lines associated with contracts intended for each of the investor types only in $(1, R)$.

Step 3: If $\frac{\mu}{1-\mu} u'(Q)Q > \frac{\gamma\lambda + (1-\gamma)\mu}{1 - (\gamma\lambda + (1-\gamma)\mu)} u'(R)R > \frac{\lambda}{1-\lambda} u'(1)$ the slope of the pooling zero-profit line lies between the P-investors' and the F-investors' marginal rates of substitution between r_1 and r_2 at $(1, R)$. In this case, for any contract $(\tilde{r}_1, \tilde{r}_2)$ on the pooling zero-profit line with $\tilde{r}_1 < 1$ and $\tilde{r}_2 > R$, there exists a contract (\hat{r}_1, \hat{r}_2) on the zero-profit line for F-investors (i.e. with slope $\mu/(1-\mu)$) that is equivalent for P-investors to $(\tilde{r}_1, \tilde{r}_2)$. Given the conditions on preferences $(\hat{r}_1, \hat{r}_2) \succ_F (\tilde{r}_1, \tilde{r}_2)$. Hence $(\tilde{r}_1, \tilde{r}_2)$ cannot constitute an equilibrium contract.

Step 4: Analogously, if $\frac{\mu}{1-\mu} u'(Q)Q > \frac{\gamma\lambda + (1-\gamma)\mu}{1 - (\gamma\lambda + (1-\gamma)\mu)} u'(R)R > \frac{\lambda}{1-\lambda} u'(1)$, then for any contract $(\tilde{r}_1, \tilde{r}_2)$ on the pooling zero-profit line with $\tilde{r}_1 > 1$ and $\tilde{r}_2 < R$, there exists a contract (\hat{r}_1, \hat{r}_2) on the zero-profit line for P-investors (i.e. with slope $\lambda/(1-\lambda)$) that is equivalent for F-investors to $(\tilde{r}_1, \tilde{r}_2)$. Given the conditions on preferences $(\hat{r}_1, \hat{r}_2) \succ_P (\tilde{r}_1, \tilde{r}_2)$. Hence $(\tilde{r}_1, \tilde{r}_2)$ cannot constitute an equilibrium contract either.

Step 5: Accordingly, contract $(1, R)$ is the only contract that is feasible and not dominated by any other contract. This proves the claim of the Proposition.

E Proof of Proposition 4

The proof is similar to the Proof of Proposition 3. However, since in this case the condition $\frac{\mu}{1-\mu} u'(Q)Q > \frac{\gamma\lambda + (1-\gamma)\mu}{1-(\gamma\lambda + (1-\gamma)\mu)} u'(R)R > \frac{\lambda}{1-\lambda} u'(1)$ is violated, the only potential pooling contract $(1, R)$ is dominated by either (r_1^d, r_2^d) for F-investors or by (r_1^δ, r_2^δ) for P-investors. Hence, no equilibrium contract obtains in this case. This proves the Proposition.

F Proof of Corollary 3

A necessary condition for pooling is

$$\frac{\mu}{1-\mu} u'(Q)Q > \frac{\gamma\lambda + (1-\gamma)\mu}{1-(\gamma\lambda + (1-\gamma)\mu)} u'(R)R > \frac{\lambda}{1-\lambda} u'(1).$$

However,

$$\lim_{\gamma \rightarrow 0} \frac{\gamma\lambda + (1-\gamma)\mu}{1-(\gamma\lambda + (1-\gamma)\mu)} u'(R)R = \frac{\mu}{1-\mu} u'(R)R > \frac{\mu}{1-\mu} u'(Q)Q$$

for $-cu''(c)/u'(c) > 1$. Therefore, for $\gamma \rightarrow 0$, $\gamma > 0$, Pareto-improving contracts to $(1, R)$ exist with $r_1 < 1$ and $r_2 = R \frac{1-(\gamma\lambda + (1-\gamma)\mu)r_1}{1-(\gamma\lambda + (1-\gamma)\mu)} > R$.

Similarly,

$$\lim_{\gamma \rightarrow 1} \frac{\gamma\lambda + (1-\gamma)\mu}{1-(\gamma\lambda + (1-\gamma)\mu)} u'(R)R = \frac{\lambda}{1-\lambda} u'(R)R < \frac{\lambda}{1-\lambda} u'(1)$$

for $-cu''(c)/u'(c) > 1$. Therefore, for $\gamma \rightarrow 1$, $\gamma < 1$, Pareto-improving contracts to $(1, R)$ exist with $r_1 > 1$ and $r_2 = R \frac{1-(\gamma\lambda + (1-\gamma)\mu)r_1}{1-(\gamma\lambda + (1-\gamma)\mu)} < R$.