b
UNIVERSITÄT
BERN

Faculty of Business, Economics
and Social Sciences

Department of Economics

Should Banks Create Money?

Christian Wipf

20-15

August 2020

DISCUSSION PAPERS

Schanzeneckstrasse 1
CH-3012 Bern, Switzerland
http://www.vwi.unibe.ch

Should Banks Create Money?

Christian Wipf*

August 2020

Abstract

The paper compares the welfare properties of two competing organizations of the monetary system: The current fractional reserve banking system versus a narrow banking system where inside money is fully backed by outside money issued by the central bank. Using a New Monetarist model, the analysis shows that fractional reserve banking is beneficial because of the interest payments on inside money. Since inside money funds loans, it pays interest, compensating the agents for the inflation tax and thus reducing the welfare costs of inflation. Since narrow banking provides no such compensation fractional reserve banking typically dominates narrow banking in terms of welfare. This also holds if outside money pays interest. Only if fractional reserve banking is sufficiently constrained, narrow banking can yield higher welfare. (JEL: E42, E51, G21)

*University of Bern, Schanzeneckstrasse 1, P.O. Box 8573, CH-3001 Bern, Switzerland, christian.wipf@vwi.unibe.ch. I thank Cyril Monnet, Aleksander Berentsen, Dirk Niepelt, Lukas Altermatt, Mohammed Ait Lahcen and Lukas Voellmy for discussions and comments.

1 Introduction

Contemporary monetary systems are characterized by a mixture of public and private means of payment. Central banks issue cash and reserves (outside money) and commercial banks issue demand deposits (inside money). Demand deposits are claims on outside money redeemable on demand. Typically banks operate under a fractional reserve banking system. They issue inside money in excess of the outside money they hold against redemptions. This system has been consistently criticized. Opponents of fractional reserve banking typically argue that it brings instability and question social benefits. They propose to separate the monetary function of banks from their other functions. If banks issue inside money they should back it fully with very safe and liquid assets while other assets should be funded by non-monetary liabilities like long-term debt or equity. The most prominent example of such a “narrow banking” proposal is the “Chicago plan” from 1933 which called for a full backing of inside money by outside money (reserves). A similar plan was advocated by Friedman [1960] and related proposals came up after the recent financial crisis.[1]

The paper aims to analyze the long run welfare implications of such proposals. Although interest in these questions has increased recently, the macroeconomic literature on the topic (where money is explicitly modelled as a nominal asset) is still surprisingly scant given the long history and the fundamental nature of the debate.[2] The specific focus of the paper are the potential benefits of fractional reserve banking in an environment where banks have a “monetary” role, i.e. their liabilities circulate as means of payment (inside money).[3] An emphasis lies on the concrete definition of narrow banking. Narrow banking means that banks must back their demandable (monetary) liabilities fully with very safe and liquid assets but they can acquire other assets if they fund them with non-monetary liabilities. In many models banks are restricted to issue one type of liabilities (an example is Chari and Phelan [2014]) which severely limits narrow banking systems. In the model presented here banks will thus be able to issue other liabilities besides inside money.

I use a basic “New Monetarist” model building on Berentsen et al. [2007]. A preference shock divides agents into consumers (buyers) and producers (sellers) and buyers acquire money to buy goods from sellers. There is outside money issued by the central bank and inside money issued by perfectly competitive commercial banks. Holding money is costly because the inflation rate lies above the Friedman rule. This “inflation tax” depresses real activity and is the basic

[1] In Switzerland there was a vote to introduce a narrow banking system in June 2018. See Pennacchi [2012] for an overview of the history of narrow banking and related proposals.

[2] This excludes two- or three-periods real models like Faure and Gersbach [2019], Benigno and Robatto [2019], Jackson and Pennacchi [2019] and Stein [2012]. Wile often providing useful intuitions, these models lack crucial aspects of monetary economies like inflation.

[3] The instability of fractional reserve banking systems has been extensively studied in the literature following Diamond and Dybvig [1983]. However, these models typically ignore the monetary role of bank liabilities and derive their demandable nature from liquidity shocks. A more recent example is Andolfatto et al. [2016]

inefficiency in the model. The uncertainty from the preference shock aggravates this basic inefficiency. The risk of ending up as a seller with costly idle money balances makes acquiring money even less attractive ex ante.

Banks have two roles in this economy. They issue means of payment (inside money) and they provide liquidity insurance against the preference shock by reallocating money balances after the shock hits. Banks acquire outside money and loans and finance them by issuing inside money and non-monetary debt. Fractional reserve banks partially back their inside money with outside money while narrow banks fully do it. I also consider a "constrained" fractional reserve banking system where banks can only issue inside money and no non-monetary debt.[4]

The analysis shows that fractional reserve banking is beneficial. If banks issue more inside money with respect to outside money welfare rises. This is interesting because the quantity theory of money would predict that the quantity of money and its composition (between inside and outside money) should be irrelevant for the equilibrium real allocation and welfare.

The reason why the quantity theory doesn't apply here are the interest payments on inside money. In equilibrium the interest rate on inside money is a weighted average between the loan rate (which equals the inflation tax by the Fisher equation) and the return on outside money (which is one in nominal terms). If banks issue more inside money, lending increases and the asset mix of banks shifts towards more loans. Thus the interest rate on inside money rises. Higher interest payments on inside money induce sellers to produce more for the same price (or to produce the same for a lower price) and this increases the value of real balances for buyers. This is how interest on inside money compensates the agents for the inflation tax and reduces the welfare costs of inflation. The compensation is partial however, because the interest on inside money (the weighted average) is always below the inflation tax.

The mechanism is equivalent to an economy where the central bank pays interest on outside money (see Rocheteau and Nosal [2017] and section 5). The result complements the usual argument of models in the spirit of Diamond and Dybvig [1983] where fractional reserve banking is beneficial because it increases high-return long term investment.

The paper also shows that fractional reserve banking dominates narrow banking in terms of welfare. This is not surprising given the first result. Under fractional reserve banking banks could also choose to back the inside money they issue fully with outside money, i.e. they could choose to become narrow banks voluntarily. And under perfect competition private and social interest are typically aligned. Thus the fact that banks do choose to become fractional and not narrow banks indicates that welfare is higher under fractional reserve

[4]This restriction could reflect a distortion in the choice of the liability structure of banks. It could be privately beneficial for banks to choose a higher level of inside money as liabilities than socially optimal (e.g. because of deposit insurance).

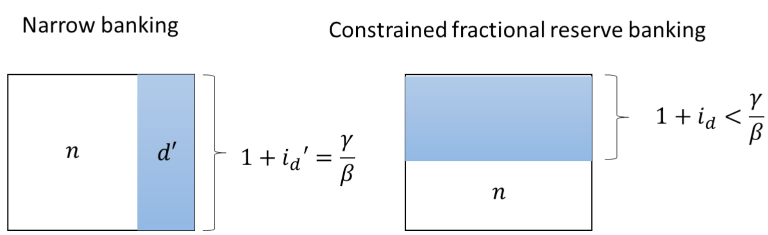
banking. But narrow banking also improves welfare compared to an economy without banks. Under narrow banking non-monetary debt pays an interest rate equal to the inflation tax. Since this debt is held by sellers after the preference shock they are perfectly compensated for the inflation tax. The same is true under fractional reserve banking, non-monetary debt is also offered at an interest rate equal to the inflation tax. Thus fractional reserve banking achieves the same insurance against the preference shock and on top of that also offers a partial compensation for the inflation tax on inside money by offering a positive interest rate. Narrow banking does not offer this kind of compensation because the interest rate on inside money is zero. The allocation in the narrow banking economy is the same as in the basic model of Berentsen et al. [2007]. Thus the paper shows how the Berentsen et al. [2007] model can be interpreted as a narrow banking economy.

Brunnermeier and Niepelt [2019] derive an interesting equivalence result. They argue that every fractional reserve banking allocation can be replicated under a narrow banking system when accompanied with appropriate transfers/open-market operations by a fiscal authority or a central bank. In section 5 I show that indeed, if the central bank pays interest on outside money under a narrow banking regime, any fractional reserve banking allocation can be replicated. However, this equivalence breaks down if we allow for interest payments on outside money in both systems. The model shows that for any central bank interest payment on outside money below the inflation tax the interest rate on inside money (and thus also welfare, for the arguments given above) under fractional reserve banking will be higher than under narrow banking and the equivalence result does not hold. This also relates to the proposal by Friedman [1960] that a narrow banking system should be accompanied by interest on reserves.

There is a debate on whether it matters if banks are modelled as institutions who influence the money supply by issuing partially backed inside money"ex nihilo" or as institutions who only intermediate already existing funds like outside money. This is exactly the difference between fractional reserve and narrow banks in the model. Fractional reserve banks create inside money "ex nihilo" and influence the money supply while narrow banks can be seen as pure intermediators of outside money as in Berentsen et al. [2007]. Thus the paper shows that it can make a difference how exactly banks are modelled and provides a counterexample to Andolfatto [2018] who finds no substantial effect in his model.

Finally the paper shows that if fractional reserve banking is constrained i.e. if fractional reserve banks can only issue inside money but no non-monetary debt, fractional reserve banking only dominates narrow banking if banks can issue a sufficiently high quantity of inside money. In this economy fractional reserve banks only issue inside money at an interest rate below the inflation tax. In this case sellers holding inside money after the preference shock are only imperfectly compensated for the inflation tax. Thus the insurance against the preference shock is imperfect in contrast to the narrow banking economy where sellers can still use non-monetary debt at an interest rate equal to the inflation tax. On the other hand fractional reserve banking still has the advantage of providing a partial compensation against the inflation tax which narrow banking has not since interest on inside money is zero.

The following figure highlights this difference. In the narrow banking economy we have a full compensation against the inflation tax on the cash deposited by agents against non-monetary debt d' but no compensation on inside money n. In the constrained fractional reserve banking economy we have a *partial* compensation of the inflation tax (because $1 + i_d$ is below the inflation tax) on the *full* stock of inside money n. If fractional reserve banks are sufficiently constrained in issuing inside money the interest rate on inside money is low and narrow banks can yield higher welfare.



The rest of the paper is organized as follows: Section 2 shows the basic environment. Then the the full model (section 3) and the model with constrained fractional reserve banking (section 4) are presented. Finally I address the question whether the systems can be equivalent in terms of welfare if the central bank pays interest on outside money (section 5).

2 Environment

Basic structure: The environment follows a standard model in the style of Lagos and Wright [2005] as presented in Berentsen et al. [2007]. Time is discrete and continues forever. Every period is divided into two sequential competitive markets called *first* and *second* market. There is a perishable consumption good produced and consumed in both markets denoted q in the first market and x in the second.

Agents: There is a unit mass of infinitely lived agents. They discount future periods with β and they cannot commit. At the beginning of every period agents face a *preference shock* which determines what they can do in the first market. With probability $s \in (0, 1)$ an agent is a *seller* and can only produce and with the inverse probability $1 - s$ an agent is a *buyer* and can only consume. Sellers have a (weakly) convex disutility of production $c(q)$ and buyers utility of consumption is strictly concave $u(q)$ and satisfies the Inada-conditions. In the second market all agents can consume and produce and their preferences are represented by a utility function $x - h$, i.e. they consume(produce) with linear (dis)utility. Denote the efficient quantity of buyer consumption in the first market with q^* given by:

$$\frac{u'(q^*)}{c'(\frac{1-s}{s}q^*)} = 1 \tag{1}$$

where $q^* = \frac{s}{1-s}q_s^*$.

Outside money, monetary policy and prices: There is a stock M of outside fiat money called "cash" issued by the central bank evolving at rate $\gamma > 0$, i.e. $M = \gamma M_{-1}$. The growth rate of the cash supply γ is the monetary policy tool of the central bank. She manages the cash supply by lump-sum cash transfers τ to agents in the second market. Since agents have unit mass the transfer/tax per agent is $\tau = M - M_{-1} = (\gamma - 1)M_{-1}$.

Let p be the price of consumption good q in terms of money in the first market and ϕ is the value of fiat money in the second market in terms of consumption goods x (i.e. the inverse of the price level in the second market). Denote (gross) inflation π as the ratio of the prices between two consecutive second markets, i.e. $\pi = \frac{\phi}{\phi_{+1}}$.

Banks, financial contracts and inside money: There is also an infinite amount of perfectly competitive, profit-maximizing firms (banks). In contrast to agents, they can commit and monitor other agents at no cost. The first property enables them to issue debt and the second property enables them to make loans.

Banks can issue two types of debt: *inside money* and *non-monetary debt.* Inside money is debt usable as means of payment in the first market. Banks need to back it at least with a fraction $\alpha \in (0, 1)$ in outside money. This constraint should capture the idea that the transactions with inside money in the first market generate redemptions from in- and outflows of inside money between banks and to settle and clear these flows banks need some outside money[5] Non-monetary debt is not usable as means of payment in the first market (suppose e.g. it has a longer maturity). But banks don't need to back it with outside money. Both types of debt are nominal, interest bearing claims on outside money. Also both types of debt are fully redeemed in the following second market after they are issued. Bank loans are inside money loans, which are also paid back in the next second market, denominated in outside money.[6] Bank contracts are formed in a *banking period* after the preference shock.

[5] In appendix A.4 I present a model where the banks outside money holdings are endogenous arising from redemptions from inside to outside money before the first market.

[6] With linear utility in the second market there is no gain from spreading the redemption of debt or the repayment of loans over multiple periods. Thus assuming this kind of contracts is not constraining in this environment.

Role for money and banks: The role for money in this environment is motivated by limited commitment and anonymity. Since agents cannot commit and are anonymous, buyers cannot issue debt in the first market and sellers require immediate compensation for the goods they produce. Buyers must give sellers "something" if they want to consume in the first market. This why agents (buyers) hold (inside or outside) money.

Banks have two roles in this environment. They can issue liabilities that circulate as means of payment in the first market, i.e. they can create money. And they can reallocate money after the preference shock which is valuable because acquiring money is costly, i.e. banks provide insurance against liquidity risk.

Equilibrium: I focus on stationary and symmetric equilibria. In a stationary environment the value of aggregate real balances is constant over time implying $\phi M = \phi_{+1} M_{+1}$. Since the stock of cash grows at γ also the price level in the second market must grow at γ or $\phi/\phi_{+1} = \gamma$ in a stationary equilibrium. By setting γ the central bank can thus also determine long-run inflation and γ can also be interpreted as the long run inflation target of the central bank.

Throughout the paper I assume holding money is costly, i.e. $\gamma > \beta$ and the central bank does not follow the Friedman rule. The assumption introduces a basic inefficiency into the environment in the form of an "inflation tax" which banks can potentially alleviate. Agents hold too little money for first best consumption in the first market, i.e. inflation acts like a tax on consumption/production in the first market. Since inside money is a claim on outside money the inflation tax also applies to inside money. The preference shock aggravates this basic inefficiency from the inflation tax. If acquiring money is costly, the risk to be a seller with (costly) idle money holdings in the first market makes acquiring money even less attractive ex-ante.

Sequence of events: Figure 1 summarizes the sequence of events in this economy: Agents acquire outside money in the second market. In the banking period after the preference shock, they can can deposit this outside money in banks and borrow to acquire inside money and non-monetary debt. In the first market they consume and produce and finally in the second market all inside money and all non-monetary debt is redeemed and the loans are repaid.

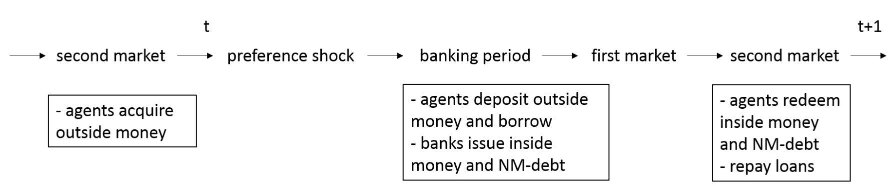


Figure 1: sequence of events

3 Unconstrained fractional reserve banking

3.1 Banks

Banks issue inside money and non-monetary debt in the banking period. They issue inside money against cash deposits d and as loans l. They must back this inside money with at least a fraction $\alpha \in (0, 1)$ of outside money. The interest on inside money is i_d and the interest on loans is i. Banks can also issue non-monetary debt against cash deposits d' at interest i'_d. A representative bank maximizes the nominal value of her assets (cash and loans) minus the value of liabilities (inside money, non-monetary debt) subject to having enough cash to satisfy the reserve constraint. The problem of a representative bank is:

$$\max_{l,d,d'} = d + d' + l(1+i) - (l+d)(1+i_d) - d'(1+i'_d) \tag{2}$$
$$s.t. \quad \alpha(l+d) \leq d + d'$$

If $i > i_d$ and $i_d, i'_d > 0$ the bank would like to make a loan as big as possible and to have as little cash deposits as possible. The reserve constraint will bind and the bank will not hold excess-reserves. Suppose this holds, then we get the following relationship between loans and cash deposits:

$$l = \frac{d'}{\alpha} + \frac{1-\alpha}{\alpha} d \tag{3}$$

Note that the loan size increases exponentially as α decreases for a given d or d' and the increase is stronger for a marginal increase in d' than in d. For example if $\alpha = 0.5$ and the bank takes a cash deposit against inside money, her lending capacity increases by 1. If the bank instead takes a cash deposit against non-monetary debt, her lending capacity increases by two. This is because issuing inside money triggers further cash acquisitions over the reserve constraint who converge to $\frac{\alpha}{1-\alpha}$ in the end while issuing non-monetary debt does not have this consequence.

Using the binding reserve constraint we can rewrite the objective function as:

$$\max_{d,d'} \quad \frac{d}{\alpha}(\alpha + (1-\alpha)(1+i) - (1+i_d)) + \frac{d'}{\alpha}(\alpha + (1+i) - \alpha(1+i'_d) - (1+i_d))$$

Thus if banks use both types of debt and we apply free entry (zero profits) we find that the interest rate on inside money and the interest rate on non-monetary debt satisfy:

$$1 + i_d = \alpha + (1 - \alpha)(1 + i) \tag{4}$$

$$1 + i'_d = 1 + i \tag{5}$$

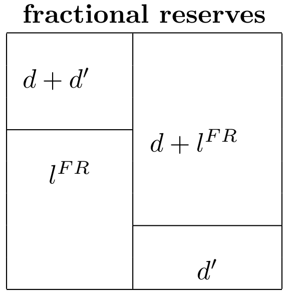
Thus the interest rate on inside money i_d is a weighted average of the return on cash (1 in nominal terms) and the return on loans $(1+i)$. As the gross loan rate will be bigger than one, the interest on inside money must be below the loan rate. This also implies it is below the interest on non-monetary debt, $i'_d > i_d$.

Also i_d increases if α goes down. This is because a lower α shifts the asset mix of the bank from assets with no return (cash) to assets with return (loans). Consequently, the bank pays an interest on its liabilities (inside money) closer to the loan rate under zero profits. The spread between loan rate and interest on inside money decreases. For example if α decreases from 0.5 to 0.25 and the loan rate stays constant at $i = 0.2$ the return on inside money increases from $i_d = 0.1$ to 0.15. If $\alpha \to 1$ the interest on inside money goes to zero. This is the case of narrow banking where issued inside money must be backed fully with outside money. If $\alpha \to 0$ the interest on inside money approaches the loan rate. In this case banks don't need to back inside money with outside money thus for the bank there is no difference between issuing inside money and non-monetary debt. i'_d does not depend directly on α.

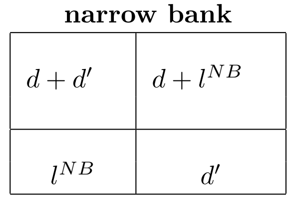
The two types of debt yield the following trade-off for the bank: Non-monetary debt has the advantage that it increases cash holdings without increasing inside money (which would trigger further cash holdings). Thus loans can increase by $1/\alpha$ at the margin with non-monetary debt while in the case of deposits against inside money they can only increase by $(1-\alpha)/\alpha$. The disadvantage of non-monetary debt are the higher funding costs since $i'_d > i_d$.

We define a *fractional reserve banking system* as an economy where banks don't fully back their issued inside money with outside money, i.e. $\alpha \in (0, 1)$. A *narrow banking system* is an economy where bank fully back their issued inside money with outside money, i.e. $\alpha = 1$. If banks acquire loans they must fund them with non-monetary debt. The following figure shows the balance sheets of a fractional reserve and a narrow bank for the same amount of deposits against inside money d and non-monetary debt d'.

fractional reserves



narrow bank



In a fractional reserve banking system the issued inside money $(d+l^{FR})$ exceeds the outside money deposited $(d+d')$ while under narrow banking they must be equal by definition. This implies that fractional reserve banks can lend more $(l^{FR} > l^{NB})$ because they don't have to back inside money $1:1$ with outside money. Narrow banks' lending capacity is constrained by the cash deposits against non-monetary debt $l^{FR} = d'$. But fractional reserve banks can lend according to (51). The last difference concerns the interest rate on inside money. Under narrow banking banks only accept cash deposits for inside money d if they pay zero interest, i.e. if $i_d = 0$, see (4). Otherwise they would set $d = 0$. Thus in a narrow banking system inside money and outside money are perfect substitutes.

3.2 Second market

A representative agent may bring outside money (m) inside money n, non-monetary debt d' and some own debt l into the second market. He chooses consumption x, work h and his new holdings of outside money m_{+1}. $V(m_{+1})$ denotes the expected value of entering the next period with m_{+1} units of outside money where $V(m_{+1}) = sV_s(m_{+1} + (1-s)V_b(m_{+1})$ i.e. the expected value of entering next period with m_{+1} units of money is the value as a buyer/seller times the respective probabilities.

$$W(m,n,d',l) = \max_{x,h,m_{+1}} x - h + \beta V(m_{+1}) \tag{6}$$
$$s.t. \quad x + \phi m_{+1} = h + \phi(\tau + m) + n\phi(1+i_d) + d'\phi(1+i'_d) - l\phi(1+i)$$

The first order condition for optimal (positive) outside money holdings solve:

$$\phi = \beta V'(m_{+1}) = sV'_s(m_{+1} + (1-s)V'_b(m_{+1}) \tag{7}$$

(7) implies that agents want to choose the same amount of cash to bring into the next period - independent of m, n, d' and l. This is a consequence of the linear utility function introduced by Lagos and Wright [2005]. The envelope conditions to the problem are:

$$\begin{aligned} W_m &= \phi \\ W_n &= \phi(1+i_d) \\ W'_d &= \phi(1+i'_d) \\ W_l &= -\phi(1+i) \end{aligned} \tag{8}$$

Finally market clearing for the output good and for money solves:

$$(1-s)h_b + sh_s = (1-s)x_b + sx_s \tag{9}$$
$$m_{+1} = M \tag{10}$$

3.3 Banking period and first market

I focus on the case where agents use only inside money in the first market. In appendix A.3 I show that buyers will strictly prefer to acquire inside money instead of outside money if interest on inside money is positive, i.e. if $i_d > 0$ and I assume they also acquire inside money if $i_d = 0$. Since after the preference shock all uncertainty is resolved, the problem of the banking period and the first market for buyers or sellers can be taken together.

buyer problem

A buyer arrives with m units of outside money in the banking period. There he decides how much of this he should deposit for inside money d_b and for non-monetary debt d'_b and how much he should borrow l_b. Then, in the first market he chooses how much to consume q_b given the amount of inside money $n = d_b + l_b$ he has.

$$V_b(m) = \max_{q_b, l_b, d_b, d'_b} \quad u(q_b) + W(m - d_b - d'_b, l_b + d_b - pq_b, d'_b, l_b) \tag{11}$$
$$s.t. \quad pq_b \leq d_b + l_b$$
$$d_b + d'_b \leq m$$

It is clear that the buyer should deposit all his cash be it for inside money or non-monetary debt, i.e. the second constraint must bind and $d'_b = m - d_b$. From the envelope conditions (8) the marginal value of inside money and non-monetary debt dominate the marginal value of cash. Second, I focus on an interior solution for borrowing. A buyer would not borrow if he already brings sufficient outside money balances m for his unconstrained level of consumption. But this could only happen if acquiring money is costless, i.e. if the inflation tax is zero or $\gamma = \beta$ which is not what we assume. The problem yields the following first-order conditions (also using (8) and with λ denoting the multiplier for the constraint):

$$\begin{aligned} q_b : &\quad u'(q_b) = p(\phi(1 + i_d) + \lambda) \\ l_b : &\quad \lambda + \phi(1 + i_d) = \phi(1 + i) \\ d_b : &\quad \lambda + \phi(1 + i_d) \geq \phi(1 + i'_d) \end{aligned}$$

The last constraint is formulated with a weak inequality meaning that if the inequality is strict the buyer wants to choose $d_b = m$ and $d'_b = 0$. From the banking problem we know that $i > i_d$ in equilibrium. Thus the constraint in the first market must bind and the buyer will be liquidity constrained, $\lambda = \phi(i - i_d)$. We also know from the banking problem that $i = i'_d$ in equilibrium. This implies the buyer is indifferent between depositing his outside money for inside money or for non-monetary debt and any combination of $d_b + d'_b = m$ is fine. The third condition holds at equality. Without loss of generality we will assume that the buyer deposits all his cash for non-monetary debt. Thus the solution to problem (11) is given by:

$$u'(q_b) = p\phi(1+i) \tag{12}$$
$$l_b = pq_b \tag{13}$$
$$d_b = 0 \quad , \quad d'_b = m \tag{14}$$

And the marginal value of outside money for a buyer is:

$$V'_b(m) = \phi(1+i'_d) \tag{15}$$

seller problem

A seller also arrives with m units of outside money in the banking period. He can deposit his outside money for inside money d_s or for non-monetary debt d'_s and he can borrow l_s. In the first market he chooses production q_s.

$$V_s(m) = \max_{q_s, l_s, d_s, d'_s} \quad -c(q_s) + W(m - d_s - d'_s, d_s + l_s + pq_s, d'_s, l_s) \tag{16}$$
$$s.t. \quad d_s + d'_s \leq m$$

The envelope conditions (8) and the relations on interest rates derived in the banking problem $i = i'_d > i_d$ significantly simplify the analysis. Also sellers will deposit all their cash for inside money or non-monetary debt. But since inside money has no liquidity value for they they strictly prefer to deposit for non-monetary debt, i.e. $d'_s = m$ and $d_b = 0$. Also sellers don't borrow if $i > i_d$, so $l_s = 0$. Thus the optimality conditions for the sellers are

$$c'(q_s) = p\phi(1+i_d) \tag{17}$$
$$l_s = 0 \tag{18}$$
$$d_s = 0 \quad , \quad d'_s = m \tag{19}$$

and the marginal value of outside money for a seller is:

$$V'_s(m) = \phi(1+i'_d) \tag{20}$$

Finally we have the market clearing conditions in the first market. Denote total bank demand for deposits against non monetary debt as d' and total bank demand for deposits against inside money as d and total bank supply of loans as l. Using the optimality conditions from above we have the following market clearing conditions in the first market:

$$
\begin{aligned}
d' &= (1-s)d'_b + sd'_s = m \\
d &= (1-s)d_b + sd_s = 0 \\
l &= (1-s)l_b \\
(1-s)q_b &= sq_s
\end{aligned}
\tag{21}
$$

Combine this with the binding reserve constraint of banks to get:

$$
(1-s)l_b = \frac{m}{\alpha} \tag{22}
$$

3.4 Equilibrium

We first solve for the equilibrium interest rates. We combine the expressions for the marginal value of outside money for a buyer (15) and for a seller (20) with the condition for optimal outside money holdings (7) to get:

$$
\phi = \beta\phi_{+1}(1 + i_{d'+1}) \tag{23}
$$

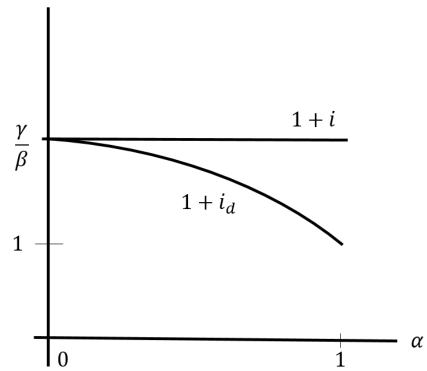
To get the equilibrium interest rates we apply stationarity ($\gamma = \phi/\phi_{+1}$) to (23) and use the relations on interest rates from the bank problem, (4) and (5).

$$
1 + i = \frac{\gamma}{\beta} \tag{24}
$$

$$
1 + i'_d = \frac{\gamma}{\beta} \tag{25}
$$

$$
1 + i_d = (1-\alpha)\frac{\gamma}{\beta} + \alpha \tag{26}
$$

Thus the loan rate and the interest on non-monetary debt must equal the inflation tax and are independent of α. This can be interpreted as a Fisher equation: the interest rates are the real interest rate ($1/\beta$) times inflation ($\phi/\phi_{+1} = \gamma$). The interest on inside money is then the weighted average of the inflation tax and the return on cash and is thus below the inflation tax and decreasing in α. We get the following picture for the evolution of the interest rates as a function of α:



To get equilibrium consumption in the first market combine optimal consumption (12) with optimal production (17) and use the equilibrium expressions for the interest rates and market clearing (21):

$$\frac{u'(q_b)}{c'(\frac{1-s}{s}q_b)} = \frac{1+i}{1+i_d} = \frac{\gamma/\beta}{(1-\alpha)\gamma/\beta + \alpha} \tag{27}$$

The following proposition summarizes the most important results:

Proposition 1. *Suppose holding money is costly (i.e. $\gamma > \beta$) and the reserve constraint is interior (i.e. $\alpha \in (0,1)$), then there is a unique stationary equilibrium of an economy with banks, inside money and non-monetary debt in which:*

i) first market consumption solves (27)

ii) first market consumption is below first best q^* and the inefficiency increases in the inflation tax γ/β and the reserve constraint α but is independent of the preference shock s.

iii) welfare is higher than in an economy without banks (see A.1), i.e. fractional reserve banking and inside money creation are essential.

iv) as $\alpha \to 1$ (the economy becomes a narrow banking economy) the allocation approaches an economy without banks and preference shock as in A.2.

v) as $\alpha \to 0$ the allocation approaches the first best allocation (1).

First the proposition shows that inside money is not neutral in this economy (*ii*)). The more inside money relative to cash (the lower α) the higher consumption in the first market and the higher welfare. Thus the quantity theory of money claiming that the quantity and the composition of inside and outside money are irrelevant does not hold here. This implies that fractional reserve banks are essential, i.e. they improve the allocation compared to an economy without banks and inside money creation (*iii*)).

Why is more inside money (a lower α) beneficial? In section 3.1 on banks we saw how a lower α leads to increases in lending and a shift in the asset-mix of banks from assets with no return (cash) to return bearing assets (loans) which allows banks to pay higher interest on inside money. This is beneficial because this interest (partially) compensates the agents for the inflation tax, which is the basic inefficiency in the economy. Since the equilibrium interest rate on inside money is below the inflation tax in equilibrium ($1+i_d < \gamma/\beta$) this compensation is always partial. The mechanism is identical to an economy with only outside money where the central bank pays interest on cash in the second market.[7]

Second: since welfare decreases in α fractional reserve banking with $\alpha \in (0,1)$ dominates narrow banking with $\alpha = 1$ in terms of welfare. This makes intuitive sense. Under fractional reserve banking banks could always decide to become narrow banks voluntarily, i.e. choose to hold more outside money than they are obliged to. Since perfect competition aligns private and social interests the fact that banks don't do this under fractional reserve banking indicates that fractional reserve banking is welfare improving.

However, also narrow banking is essential, i.e. welfare under narrow banking is higher than without banks as in appendix A.1. We saw in the section on banks that narrow banking cannot offer interest on inside money, $i_d = 0$. So narrow banks cannot provide a compensation against the inflation tax and the welfare costs of inflation. But they are useful because they perfectly insure agents against the preference shock. The easiest way to see this is by verifying that the allocation of the narrow banking model is exactly the same as in an economy without preference shock and banks (*iii*)), see A.1. The intuition is that since the deposit rate on non-monetary debt exactly equals the inflation tax ($1+i'_d = 1+i = \frac{\gamma}{\beta}$) agents that turn out to be sellers are perfectly compensated for the inflation tax on the cash they acquired. So the risk of being a seller who cannot use the cash in the first market disappears and the allocation-worsening role of the preference shock is eliminated. Narrow banks can be seen as a substitute for a market to borrow/lend cash after the preference shock.[8]

The narrow banking is equivalent to an allocation with only outside money and banks reallocating this outside money after the preference shock. This is the basic version of Berentsen et al. [2007]. Thus the proposition shows how the model of Berentsen et al. [2007] can be interpreted as a narrow banking economy where banks issue fully backed inside money and non-monetary debt.[9]

Finally a few comments on the equilibrium if $\alpha = 0$ (v). In this case banks don't need outside money to back the inside money they issue. This means they are only willing to take cash deposits if the interest rate on inside money (and non-

[7] The equilibrium allocation of an economy with interest on outside money is given by $\frac{u'(q)}{c'(q)} = \frac{\gamma}{\beta}\frac{1}{1+i_m}$ where $1+i_m$ is the gross interest on cash by the central bank, see Rocheteau and Nosal [2017] p.140. Note that this expression is exactly identical to equation 27.

[8] See [Rocheteau and Nosal, 2017], chapter 8.5 for this equivalence.

[9] In footnote 9 of Berentsen et al. [2007] the authors also make the interpretation of their model as a narrow banking economy.

monetary debt) is zero, i.e. $i_d = i'_d = 0$. Zero profits then implies that also the loan rate is zero. However, the condition on agents optimal outside money holdings, (23), tells us that if interest rates are zero agents would not be willing to hold outside money if the inflation tax is positive ($\gamma > \beta$). Thus this cannot be an equilibrium. Now suppose equilibrium interest rates are $i = i_d > 0$ and $i'_d > 0$ At these interest rates bank demand for cash deposits (either for inside money or non-monetary debt) is zero. They only make loans in inside money and since $i = i_d$ they also make zero profits. However, for this equilibrium to exist also the supply of cash deposits must be zero i.e. agents don't want to hold cash anymore. This is satisfied if the marginal costs of holding outside money are higher than the marginal benefits, i.e. if (23) is an inequality. In equilibrium we must thus have that both interest rates on inside money and on non-monetary debt are positive but below the inflation tax, i.e.

$$0 < i_d, i'_d < \gamma/\beta - 1 \tag{28}$$

Any interest rates satisfying (28) would be an equilibrium if $\alpha = 0$. In such an economy outside money has no function anymore, it is a pure inside money economy.[10] Because the loan rate and the interest on inside money are identical buyers are never liquidity constrained and their holdings of inside money can be anything from the quantity to consume the first best to infinity, i.e. $l_b \in (pq^*, \infty)$. The economy achieves the first best allocation equivalent to an economy with direct credit.

3.5 The non-neutrality of inside money

To see the beneficial effects of higher interest on inside money more clearly look at supply and demand for the consumption good in the first market. Equilibrium supply from the seller side is implicitly defined by optimal production (17), $c'(q) = p\phi(1 + i_d)$, and equilibrium demand is given by the buyer optimality condition $u'(q) = \gamma/\beta p\phi = (1 + i)p\phi$. Figure 2 shows demand and supply according to these equations. Specifically it depicts a shift from initial supply $S(\alpha_1)$ to the new supply curve $S(\alpha_2)$ for a decrease in α from an arbitrary value α_1 to a lower value α_2. Note that supply weakly increases in the interest on inside money from the properties of the utility functions. Higher interest payments on inside money induce the sellers to produce more for the same relative price (or to produce the same for a lower price). The figure shows that as a result of this change equilibrium consumption (and production) will clearly increase and the relative price $p\phi$ decreases.

[10]Since all prices and contracts were defined in outside money such an economy would have to use another numeraire.

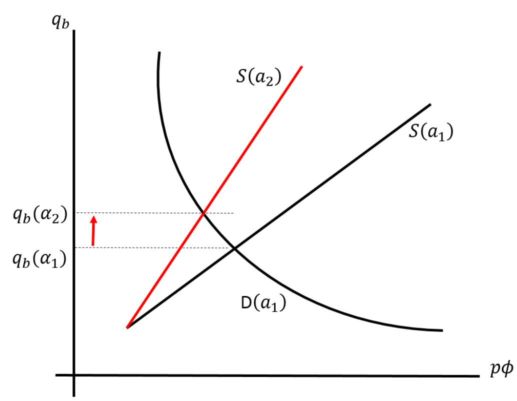


Figure 2: The equilibrium effects of an increase in interest of inside money

How can we see the non-neutrality from a quantity theory perspective? Look at the inside money a buyer brings to the first market. In equilibrium a buyer deposits his outside money for non-monetary debt and acquires inside money only by borrowing. Thus his inside money holdings are just l_b. Using market clearing in the first and the second market, (22) and (9), we get that the inside money holdings of a buyer are $l_b = \frac{M_{-1}}{\alpha(1-s)}$. Equilibrium consumption in the first market can then be written as:

$$q_b = \frac{n}{p} = \frac{M_{-1}/(\alpha(1-s))}{p} \tag{29}$$

If the quantity theory would hold with respect to inside money this would imply that the price level in the first market rises 1 : 1 with the amount of inside money available for the buyer. For example if α decreases from 0.5 to 0.25 the amount of inside money available for a buyer l_b more than doubles. However, since also interest on inside money goes up (per unit) sellers accept this amount of inside money at a lower price than a 1 : 1 increase. Thus the denominator of (29) rises less than the denominator of (29) and real consumption q_b is not constant but rises.

4 Constrained Fractional reserve banking

In this section, I assume banks are restricted to issue inside money under fractional reserve banking. But the narrow banking allocation is the same as before. This shows how the fact that banks can issue different types of liabilities matters for the results. The allocation could be interpreted as a situation where the private interests of banks and the social interests are not aligned and banks - for some reason - have a private benefit of issuing more (or only) inside money (e.g. because it's cheaper) than socially optimal i.e. the choice of the liability structure of banks is distorted. The balance sheet of banks with only inside money as liabilities is:

constrained
fractional reserves

d	$d + l^{FR}$
l^{FR}	

For banks this means d' is zero in (2) and we only have i_d and i as interest rates. If $i > i_d > 0$ also here the bank wants to make a loan as big as possible and to hold as little outside money as possible. Thus the reserve constraint will bind and instead of (51) we get

$$l = \frac{1-\alpha}{\alpha}. \tag{30}$$

Then, under zero profits (4) still holds.

$$1 + i_d = \alpha + (1-\alpha)(1+i) \tag{4}$$

For buyers and sellers the basic problem is still (11) and (16). For them the difference is that they cannot deposits their outside money for non-monetary debt anymore. In this case both will deposit their outside money for inside money and we have:

$$d_b = d_s = m \tag{31}$$

The change also affects the marginal value of outside money. For buyers (15) becomes now:

$$V_b'(m) = \frac{u'(\frac{m+l}{p})}{p} = \phi(1+i) \tag{32}$$

and for sellers (20) becomes

$$V_s'(m) = \phi(1+i_d) \tag{33}$$

since they now deposit for inside money and not for non-monetary debt. This has implications for optimal outside money holdings of buyers. Instead of (23) (7) now yields:

$$\phi = \beta\phi_{+1}[(1-s)(1+i_{+1}) + s(1+i_{d+1})]. \tag{34}$$

or under stationarity:

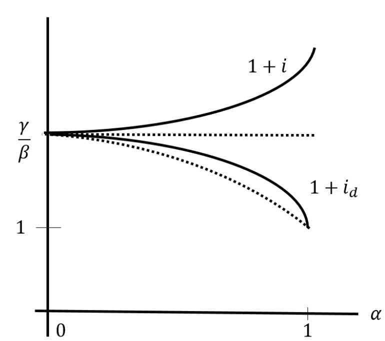
$$\frac{\gamma}{\beta} = (1-s)(1+i) + s(1+i_d) \tag{35}$$

Combining (35) with the relation for interest rates from the bank problem (4) yields the following equilibrium interest rates:

$$1 + i_d = \frac{(1-\alpha)\gamma/\beta + \alpha(1-s)}{1-\alpha s} \tag{36}$$

$$1 + i = \frac{\gamma/\beta - \alpha s}{1-\alpha s} \tag{37}$$

The big difference to before is that now the loan rate increases with α (before it was independent of α). This is because before both agents could deposit their outside money at rate i'_d which was equal to i. Thus it was irrelevant whether an agent turned out to be a buyer or a seller and $1 + i'_d = 1 + i = \gamma/\beta$ in equilibrium. Now since there is a spread between the loan rate and the deposit rate $i > i_d$ the preference shock matters because as a seller an agent gets less. Thus as a buyer the agent must be compensated with an interest $1+i > \gamma/\beta$ in equilibrium and $1+i_d < \gamma/\beta$. As i_d is decreasing in α this also implies $1+i$ must rise with α otherwise the agent would not hold outside money in equilibrium. The following figure shows this evolution in equilibrium (the dashed lines are the interest rates of the economy without preference shock).



To get equilibrium consumption in the first market we again combine the expressions for equilibrium interest rates with optimal buyer consumption (12) and seller production (17).

$$\frac{u'(q_b)}{c'(\frac{1-s}{s}q_b)} = \frac{1+i}{1+i_d} = \frac{\gamma/\beta - \alpha s}{(1-\alpha)\gamma/\beta + \alpha(1-s)} \tag{38}$$

The following proposition summarizes the most important results:

Proposition 2. *Suppose holding money is costly (i.e. $\gamma > \beta$) and the reserve constraint is interior (i.e. $\alpha \in (0,1)$), then there is a unique stationary equilibrium with only inside money in which:*

i) first market consumption solves (38)

ii) first market consumption is below first best consumption q^* and the inefficiency increases in the inflation tax γ/β, the reserve constraint α and the fraction of sellers in the economy s.

iii) welfare is lower than in the economy with both types of debt from proposition 1.

iv) welfare is higher than in an economy without banks as in A.1 and approaches this allocation as $\alpha \to 1$.

v) welfare is higher than in the narrow banking economy from proposition 1 if $\alpha < \tilde{\alpha}$ and lower if $\alpha > \tilde{\alpha}$ where $\tilde{\alpha} = \frac{\gamma/\beta}{\gamma/\beta + s} > 0.5$.

vi) as $\alpha \to 0$ the allocation approaches the first best allocation.

The most important result from this proposition is that constrained fractional reserve banking doesn't strictly dominate narrow banking in terms of welfare as before. From *iv*) the allocation with constrained fractional reserve banking approaches an economy without banks and we still know from proposition 1 that welfare with narrow banking is higher than in an economy without banks. As welfare increases when α decreases and approaches the first best allocation we have a threshold result. If banks can issue a sufficiently high quantity of inside money ($\alpha < \tilde{\alpha}$) welfare under constrained fractional reserve banking is higher while if banks are very constrained in the issuance of inside money ($\alpha > \tilde{\alpha}$) welfare under (unconstrained) narrow banking is higher.[11]

Remember that under narrow banking inside money pays zero interest, $i_d = 0$ but non-monetary debt pays an interest equal to the inflation tax $1 + i'_d = \gamma/\beta$. Thus agents who turn out to be sellers can deposit their outside money after the preference shock at an interest rate which perfectly compensates them for the inflation tax. Thus narrow banking achieves perfect insurance against the

[11] $\tilde{\alpha}$ must be above 0.5 since this is the number reached as $\gamma/\beta \to 1$ and $s \to 1$. So for any $\alpha < 0.5$ fractional reserves is always better.

preference shock (the basic inefficiency of the inflation tax on inside money however, is not addressed since inside money pays no interest). Note that now with constrained fractional reserve banking this is not the case. Agents who turn out to be sellers now deposit their outside money at $1 + i_d < \gamma/\beta$ since banks can only issue inside money. Thus the insurance against the preference shock is imperfect under constrained fractional reserve banking and narrow banking has a relative advantage in this respect. This explains why the threshold $\tilde{\alpha}$ decreases in s. The higher the risk of becoming a seller the more valuable is the perfect insurance of narrow banking and therefore the range where narrow banking dominates increases.

On the other hand fractional reserve banking has the advantage of partially compensating agents against the inflation tax on holding money because it pays interest on inside money $i_d > 0$ while under narrow banking interest on inside money is zero $i_d = 0$. This is the relative advantage of fractional reserve banking. It explains why the threshold $\tilde{\alpha}$ increases in the inflation tax γ/β. If the inflation tax is high holding money is very costly and thus the advantage of fractional reserve banking which provides a partial compensation for incurring this tax is very valuable. Thus the range where fractional reserve banking dominates increases in the inflation tax.

The following figure highlights this difference. In the narrow banking economy we have a full compensation against the inflation tax on the cash deposited by agents against non-monetary debt d' without compensation on inside money n. In the constrained fractional reserve banking economy we have a *partial* compensation of the inflation tax (because $1 + i_d$ is below the inflation tax) on the *full* stock of inside money n.

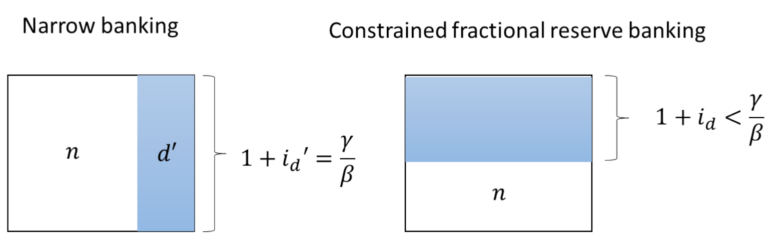


Figure 3: Interest under narrow and fractional reserve banking

In the fractional reserve banking economy from section 3 the relative advantage of narrow banking disappears. Banks now also offer non-monetary debt that perfectly compensates the sellers against the inflation tax as it also pays $1+i'_d = \gamma/\beta$. With two types of debt also fractional reserve banks offer perfect insurance against liquidity risk and the threshold result disappears. Fractional reserve banking then strictly dominates narrow banking as proposition 1 shows.

5 Interest on outside money and equivalence

Brunnermeier and Niepelt [2019] derive an interesting equivalence result. They argue that every fractional reserve banking allocation can be replicated with a narrow banking system when accompanied with appropriate transfers/open-market operations by a fiscal authority and a central bank. In this section I revisit this claim in the current framework.

Suppose we are in the baseline economy under a narrow banking system. From proposition 1 we know that forcing α to be 1 as a narrow banking system demands reduces welfare compared to fractional reserves. Now suppose the central bank pays interest i_m on outside money in the second market.[12] If an agent (or the bank) brings m units of outside money into the second market the central bank additionally gives him mi_m units of cash as interest. These interest payments enter the budget constraint of the central bank as an additional component. The seignorage revenue in period t, $M - M_{-1}$ (which can also be negative if the central bank reduces the money supply) is now used for transfers τ and interest payments $i_m M_{-1}$. In nominal terms:

$$\tau + i_m M_{-1} = M - M_{-1} = (\gamma - 1)M_{-1} \tag{39}$$

Now consider the bank problem of the baseline model from section 3. If the interest payments of the central bank are below the interest rates on inside money and non-monetary debt, i.e. if $i_d', i_d > i_m$ and $i > i_d$, the reserve constraint (51) will bind. Proceeding with the same steps as in the baseline model yields a similar relationship to (4) between the interest on inside money and the other interest rates:

$$1 + i_d = \alpha(1 + i_m) + (1 - \alpha)(1 + i) \tag{40}$$

The return on inside money is still a weighted average of the return on outside money and the return on loans taking into account that outside money now pays interest too. The relationship between the interest rate on non-monetary debt and the loan rate under zero profits on the other hand is unchanged. As in equation (5) we have $1 + i_d' = 1 + i$. Since the central bank pays interest on outside money commercial banks can pay interest on inside money even under narrow banking (with $\alpha = 1$) now. From (40) we get that $1 + i_d = 1 + i_m$ under narrow banking. The rest of the problem being unchanged equilibrium consumption under narrow banking but with interest on outside money is still given by (27). Only that interest on inside money now equals the interest paid by the central bank.

[12] On this see also Rocheteau and Nosal [2017] p.139-141.

$$\frac{u'(q_b)}{c'(\frac{1-s}{s}q_b)} = \frac{1+i}{1+i_d} = \frac{\gamma/\beta}{1+i_m} \tag{41}$$

Thus indeed a narrow banking economy where the central bank pays interest on outside money can replicate any fractional reserve banking allocation given by (27). The condition is that the central bank pays an interest rate on outside money equalling the interest on inside money commercial banks pay under fractional reserves. (39) tells us that the central bank interest payments are financed by lower transfers or higher taxes in the second market. In equilibrium this means buyers will work more and sellers consume more in the second market. But since both have linear (dis)utility these distributional changes do not matter for welfare.

However, this equivalence-argument is incomplete in the sense that we compared a narrow banking economy with interest on outside money with a fractional reserve economy without this policy. In fact, if we allow for interest payments on outside money in both systems the equivalence result disappears and fractional reserve banking again dominates narrow banking in terms of welfare as in proposition 1. From (40) we see that for interest rates below the inflation tax the interest rate on inside money, i_d, is always higher than the interest rate on outside money, i_m, since the loan rate equals the inflation tax in equilibrium. Thus for any central bank interest payment on outside money below the inflation tax the fractional reserve banking allocation is better than the narrow banking allocation and the results from proposition 1 go through. No matter what interest rate the central bank chooses, the narrow banking allocation is never equivalent to the fractional reserve banking allocation. Only in the limit when the central bank pays an interest rate equal to the inflation tax the two systems would be equivalent. In this case even the first best is achieved. This situation is similar to when the central bank chooses the Friedman rule ($\gamma = \beta$). In both cases outside money is costless to hold and banks who help with the creation and allocation of liquidity offer no social benefit.

Thus this model is an example where the equivalence proposition of Brunnermeier and Niepelt [2019] generally does not hold. The reason might be that they don't consider the possibility that policies chosen to counteract the changes from adopting a narrow banking system might also change the original allocation. The section also speaks to the proposal by Friedman [1960] who proposed that a narrow banking system should be accompanied by interest on reserves. The model shows that the proposal of Friedman would be welfare improving but fractional reserve banking would still dominate narrow banking as the interest rate on inside money would also incorporate the interest on outside money and still lie higher than the interest on outside money.

6 Conclusion

The paper analyzed the welfare implications of narrow banking compared to the current fractional reserve banking system. Abstracting from fragility issues and focusing on the "monetary" role of banks where bank liabilities circulate as means of payment (inside money) the analysis showed that fractional reserve banking is beneficial because of the interest payments on inside money. Since inside money funds loans, it pays interest, compensating the agents for the inflation tax and thus reducing the welfare costs of inflation. The paper thus provides a more "monetary" argument for efficiency gains from fractional reserve banking which complements the classical analysis from Diamond and Dybvig [1983] where fractional reserve banking mobilizes investment in long-term, high-return assets. The paper also connects to the literature on the welfare costs of inflation. As observed by Lucas [2000] the possibility of demand deposits to pay interest should be taken into account when estimating the welfare costs of inflation using a measure like M1 which is a sum of non-interest bearing outside money and possibly interest-bearing inside money. The paper formalized this observation. It also showed that fractional reserve banking generally dominates narrow banking in terms of welfare because of these interest payments although narrow banking is modelled more carefully than in other papers. In this respect the paper also demonstrated how Berentsen et al. [2007] can be interpreted as a narrow banking economy where banks issue inside money and a non-monetary liability like long-term debt. Finally the paper analysed a situation where fractional reserve banking is constrained in the issuance of liabilities which could be interpreted as a distortion in the bank liability choice e.g. because inside money is subsidized by deposit insurance. The paper shows that only then and if banks are very constrained in their issuance of inside money narrow banking can yield higher welfare.

The broader message of the paper is that narrow banking systems in the spirit of the Chicago Plan where banks must back inside money fully with non-interest bearing outside money have efficiency costs in terms of foregone interest payments. Paying interest on outside money as proposed by Friedman [1960] would improve welfare but fractional reserve banking would still dominate narrow banking in such an environment. The interest rate on inside money would also incorporate the interest on outside money and still lie higher than the interest on outside money.

Various extensions could be addressed in further work. The model presented here is essentially a model of efficient liquidity provision and allocation. It could be augmented by banks having a role in capital accumulation and investment. It would also be interesting to quantify the welfare gains from inside money creation. Some quantitative estimates from New Monetarist models on the welfare costs of inflation are available. They could be complemented with a quantitative estimate of this model to provide a measure for the quantitative importance of these welfare gains which could then be set into relation to estimated costs of financial fragility. Another direction would be deviations from perfect competition. For example the discussion of the desirability of private seignorage by banks is impossible in a model where banks make zero profits. Finally a more complete analysis of the two banking systems should include financial fragility.

References

David Andolfatto. Reconciling Orthodox and Hererodox Views on Money and Banking. *working paper*, 2018.

David Andolfatto, Aleksander Berentsen, and Fernando M. Martin. Financial Fragility in Monetary Economics. *working paper*, 2016.

Pierpaolo Benigno and Roberto Robatto. Private money creation, liquidity crises, and government interventions. *Journal of Monetary Economics*, 106: 42 – 58, 2019.

Aleksander Berentsen, Gabriele Camera, and Christopher Waller. Money, credit and banking. *Journal of Economic Theory*, 135(1):171 – 195, 2007.

Markus K. Brunnermeier and Dirk Niepelt. On the equivalence of private and public money. *Journal of Monetary Economics*, 106:27 – 41, 2019.

V.V. Chari and Christopher Phelan. On the social usefulness of fractional reserve banking. *Journal of Monetary Economics*, 65:1 – 13, 2014.

Douglas W. Diamond and Philip H. Dybvig. Bank Runs, Deposit Insurance, and Liquidity. *Journal of Political Economy*, 91(3):401 – 419, 1983.

Salomon Faure and Hans Gersbach. Money Creation in Different Architectures. *CEPR discussion paper*, 2019.

Milton Friedman. *A Program for Monetary Stability.* Fordham University Press, 1960.

Timothy Jackson and George Pennacchi. How Should Governments Create Liquidity? *working paper*, 2019.

Ricardo Lagos and Randall Wright. A unified framework for monetary theory and policy analysis. *Journal of Political Economy*, 113(3):463–484, 2005.

Robert E. Lucas, Jr. Inflation and welfare. *Econometrica*, 68(2):247–274, 2000.

George Pennacchi. Narrow banking. *Annual Review of Financial Economics*, 4 (1):141–159, 2012.

Guillaume Rocheteau and Ed Nosal. *Money, Payments, and Liquidity.* MIT Press, 2017.

Jeremy C. Stein. Monetary policy as financial stability regulation. *The Quarterly Journal of Economics*, 127(1):57–95, 2012.

Stephen D. Williamson. Liquidity, monetary policy, and the financial crisis: A new monetarist approach. *American Economic Review*, 102(6):2570–2605, May 2012.

Appendix A

A.1 Economy with outside money and preference shock

In an economy without banks the value of an additional unit of cash for a buyer in $t+1$ is $\frac{u'(q_{b+1})}{p_{+1}}$ and for a seller just ϕ_{+1} since he cannot deposit and earn interest. Thus optimal cash holdings (7) for an agent in this economy solve:

$$\phi = \beta[(1-s)\frac{u'(q_{b+1})}{p_{+1}} + s\phi_{+1}].$$

Optimal production against cash is given by $c'(q_s) = p\phi$. Thus the stationarity equilibrium consumption in the first market without banks q_b solves:

$$\frac{u'(q_b)}{c'(\frac{1-s}{s}q_b)} = \frac{\gamma/\beta - s}{1-s}. \tag{42}$$

We see the RHS of equation (42) is equal to the RHS of 43 if $s = 0$ (i.e. there is no preference shock) and that it is increasing in s. By the same logic as in A.1 we can thus conclude that if $s > 0$ the allocation is worse than in 43 and the inefficiency increases in s.

A.2 Economy with outside money and no preference shock

Rocheteau and Nosal [2017] (p.138) show there exists a unique stationary equilibrium where all buyers and sellers have access to the first market (i.e. $\sigma = 1$ in their model) if $\gamma > \beta$ which solves:

$$\frac{u'(q_b)}{c'(q_s)} = \frac{\gamma}{\beta} \tag{43}$$

A.3 Use of inside and outside money in the first market

In this section we want to show that if both inside and outside money are used in the first market agents (weakly) prefer using inside money if the interest rate is non-negative. To make this point we will look at the choice of means of payments for buyers and neglect non-monetary debt and borrowing. Suppose a buyer arrives with m units of outside money. Still we denote the amount of outside money the buyer deposits in the bank for inside money as d_b and the amount of outside money she keeps as m'. Thus $m = d_b + m'$. To simplify we will also slightly change the interpretation of the price p in the first market. Before we assumed that it is expressed in terms of inside money and we formulated the means-of-payment-constraint as $pq_b \leq d_b + l_b$. Now we define p in terms of outside money in the next second market. Thus we can write: $pq_b \leq d_b(1+$

$i_d) + m'$. We will also just use outside money in the value function for the next second market. We can rewrite this modified buyer problem as:

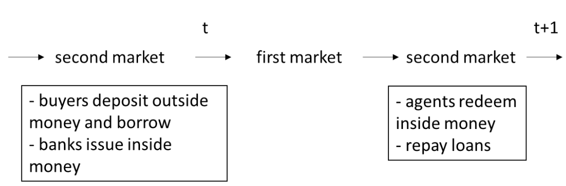
$$V_b(m) = \max_{q_b, d_b, m'} \quad u(q_b) + W(d_b(1+i_d) + m' - pq_b) \tag{44}$$
$$s.t. \quad pq_b \leq d_b(1+i_d) + m'$$
$$d_b + m' = m$$

Using $m' = m - d_b$ in the problem we see that in both the right-hand side of the constraint and the amount of outside money holdings in the next second market you get the positive term $d_b i_d$. Thus the marginal benefits of depositing are positive if $i_d > 0$ and a buyer would like to set d_b as high as possible, i.e. $d_b = m$ and $m' = 0$. This illustrates that a buyer strictly prefers to use inside money if the interest on inside money is positive.

A.4 A model with early redemptions

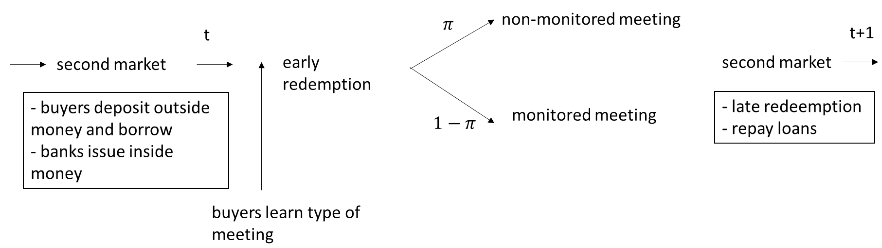
In the baseline model the need for banks to hold outside money was motivated by the assumption that clearing and settling the transactions with inside money in the first market takes a minimal amount of outside money proportional to the inside money used (α). However, the concrete clearing/settlement process was not modelled. Building on Williamson [2012] I now motivate the outside money holdings of banks differently. As in the baseline model buyers hold inside money but some of them will be able to only use outside money in the first market. These buyers will thus redeem their inside money for outside money before the first market and banks hold outside money to satisfy these redemptions. The model explains how the reserve ratio α can be interpreted as a function of the share of outside money transaction in the economy and inflation. It also confirms the basic logic of the baseline model from proposition 1 that there is a negative relationship between the reserve ratio and welfare.

Compared to the baseline model I will make some simplifying assumptions. First I assume there is a unit mass of buyers and sellers each and there is no preference shock. This limits the role of banks to the provision of liquidity and means only buyers will acquire money. I will also ignore the possibility that banks issue non-monetary debt and buyers will already acquire inside money by depositing and borrowing in the second market. Without preference shock this does not change the problem. Compared to figure 1 the sequence of events then simplifies to:



Following Williamson [2012] I will also depart from perfect competition in the first market and assume bilateral meetings between buyers and sellers where buyers make take-it-or-leave-it (TIOLI) offers to sellers. And I will assume linear disutility of working in the first market, $c(q) = q$, for sellers and ln-utility for buyers, $u(q) = \ln(q)$.[13]

The crucial novelty is that now there are two types of meetings in the first market for buyers. With probability π buyers go to a *non-monitored* meeting where they can only use outside money. And with probability $1-\pi$ they go to a *monitored* meeting where they can use inside money like in the baseline model. Buyers acquire inside money by depositing cash and borrowing in the second market. But since they may need cash if they go to a non-monitored meeting the bank now allows for *early* redemptions. After knowing the type of meetings buyers can redeem inside money into outside money at interest rate i_{d1} before the first market. Inside money which is not redeemed early will be redeemed in the following second market at the interest rate i_{d2} as in the baseline model. The sequence of events is then as follows:



Given the simplifications we can focus on the buyers. The buyers acquire inside money by depositing outside money, d, and borrowing l. The total amount of inside money a buyer acquires in a second market is $n = l + d$. Let the amount consumed in a non-monitored meeting be q^c (for "cash") and q in a monitored meeting. Also let n^c be the amount of inside money a buyer redeems before a non-monitored meeting and n' be the amount of inside money used in a monitored meeting where $n^c, n' \le n$ for feasibility. In a non-monitored meeting the buyer redeems n^c units of inside money for cash at rate $1 + i_{d1}$. The value of this outside money is $\phi n^c(1 + i_{d1})$ for a seller in the next second market and since we assume TIOLI offers by buyers this is exactly the amount produced, i.e. $q^c = \phi n^c(1 + i_{d1})$. In a monitored meeting the buyer uses n' units of inside money to pay the seller where by the same reasoning we must have that $q = \phi n'(1 + i_{d2})$. The problem of a representative buyer can then be written as:

[13] These assumptions do not fundamentally change the results. Assuming bilateral meetings with TIOLI offers and linear seller utility in the baseline model yields the same allocation $u'(q) = \frac{1+i}{1+i_d}$ as in the competitive model, (27).

$$\max_{d,l,n^c\leq n,n'\leq n} \quad -\phi_{-1}d+\beta\pi[u(\phi n^c(1+i_{d1}))+\phi(n-n^c)(1+i_{d2})] \tag{45}$$
$$+\beta(1-\pi)[u(\phi n'(1+i_{d2}))+\phi(n-n')(1+i_{d2})]$$
$$-\beta\phi l(1+i)$$

We can distinguish four possible cases solving this problem:

a) costless inside money ($i_{d1} < i_{d2} = i$): In this case inside money is costless to hold and the buyer holds so much that he is unconstrained in both types of meetings. The unconstrained consumption levels are:

$$u'(\tilde{q}^c)=\frac{1+i_{d2}}{1+i_{d1}} \quad , \quad u'(q^*)=1 \tag{46}$$

i.e. in the non-monitored meeting the buyer wants to consume $\tilde{q}^c$ and in the monitored meeting he consumes the first best quantity q^*. The return on inside money equals the inflation tax $1+i_{d2}=1+i=\frac{\phi}{\beta\phi_{+1}}$ and the buyer doesn't use all his inside money in both types of meetings ($n^c, n' < n$). The amount of real inside money holdings ϕn is undetermined when inside money is costless to hold. To be unconstrained in the monitored meeting a buyer needs at least $q^*/(1+i_{d2})=n^*$ units of real inside money and to be unconstrained in the non-monitored meeting the buyer needs at least $\tilde{q}^c/(1+i_{d1})=\tilde{n}$ units. With ln-utility we have $\tilde{n}=n^*$ i.e. buyers need the same amount of real inside money to be unconstrained in both types of meetings and real inside money holdings ϕn must lie in (n^*,∞). Without loss of generality we can assume that buyers choose $\phi n = n^*$ in case *a)*. This means buyers will always redeem (or use) all inside money in both types of meetings and we don't need to worry about partial redemptions.

b) costly inside money with spread ($i_{d1} < i_{d2} < i$): In this case the return on inside money is below the inflation tax and buyers are constrained in both types of meetings ($n^c = n' = n$). The consumption levels and real inside money holdings solve:

$$q^c=\phi n(1+i_{d1}) \quad , \quad q=\phi n(1+i_{d2}) \tag{47}$$
$$\text{with} \qquad \phi n=\frac{1}{1+i}$$

c) costly inside money without spread ($i_{d1}=i_{d2}=i_d < i$): In this case inside money is costly to hold and its return is the same when redeemed early or late.[14] Also here buyers are constrained in both types of meetings and consumption in both meetings and real inside money holdings solve:

$$u'(q)=\frac{1+i}{1+i_d} \quad , \quad \phi n=\frac{1}{1+i} \tag{48}$$

[14] We abstract from the case when $i_{d1}=i_{d2}=i$ which can never be feasible for the bank if she holds outside money.

In any of these cases the indifference condition between the two ways of acquiring inside money (borrowing and depositing cash) must hold as in the baseline model (see (23) and thus also here the loan rate equals the inflation tax:

$$\frac{\phi_{-1}}{\phi}\frac{1}{\beta} = 1 + i \tag{49}$$

Also, the following relationship between the interest rates must hold in any equilibrium:

$$0 \leq i_{d1} \leq i_{d2} \leq i \tag{50}$$

The inequalities ensure that buyers prefer using inside money ($i_{d1} \geq 0$), that only buyers going to a non-monitored meeting redeem early ($i_{d2} \geq i_{d1}$)[15] and that the solution for real inside money holdings is bounded ($i_{d2} \leq i$).

bank problem: The bank maximizes total cash profits under the assumption that buyers always redeem all inside money in non-monitored meetings. We consider a large bank with lots of buyers where the fraction of buyers going to a non-monitored meeting is approximately π. A representative bank who issued n units of inside money and holds d units of outside money therefore faces total early redemptions of πn units of inside money which she redeems with $1+i_{d1}$ per unit. Thus to withstand these redemptions the bank needs at least $\pi n(1+i_{d1})$ units of outside money. This constraint is similar to the reserve constraint in the baseline model. The problem of a bank who issued $n = l + d$ units of inside money is:

$$\begin{aligned} \max_{l,d} = d - \pi n(1+i_{d1}) + l(1+i) - (n - \pi n)(1+i_{d2}) \\ s.t. \quad \pi n(1+i_{d1}) \leq d \end{aligned}$$

We see early redemptions modify the bank problem in two ways (compare it to (2)). It decreases outside money holdings and it decreases the outstanding inside money of banks. Thus it decreases both assets and liabilities. We can reformulate the objective function as follows:

$$l(i - \pi i_{d1} + (1-\pi)i_{d2}) - d(\pi i_{d1} + (1-\pi)i_{d2})$$

Thus if $0 < \pi i_{d1} + (1-\pi)i_{d2} < i$ the bank wants to set loans as high as possible and cash deposits as low as possible. This implies the constraint should bind and we get:

$$l = \frac{1 - \pi(1+i_{d1})}{\pi(1+i_{d1})}d \tag{51}$$

Defining $\alpha = d/n$ as the ratio of bank outside to inside money holdings ("reserve ratio") like in the baseline model we can rewrite the interest rate on early

[15] We abstract from the possibility of belief-driven redemptions in the spirit of Diamond and Dybvig [1983]

redemptions as a function of the (now endogenous) reserve ratio:

$$1 + i_{d1} = \frac{\alpha}{\pi} \tag{52}$$

Using (51) on the objective function and applying the zero profit condition implies a similar condition for the interest rate for late redemption using (52):

$$1 + i_{d2} = \frac{(1-\alpha)(1+i)}{1-\pi} \tag{53}$$

Thus the rate on early (late) redemptions increases (decreases) in the reserve ratio and the rate on early redemptions is independent of the inflation tax. Interest on early redemptions can only be financed by holding more outside money while interest on late redemptions can be financed by loans with higher return also. Note also that paying a positive interest rate on early redemptions implies having a reserve ratio bigger than the fraction of buyers going to a non-monitored meeting.

The boundary conditions on the interest rates (50) translate into boundary conditions for the reserve ratio. $i_{d2} \leq i$ implies $\alpha \geq \pi$ and $i_{d1} \leq i_{d2}$ implies $\alpha \leq \tilde{\alpha}$ where

$$\tilde{\alpha} = \frac{1+i}{1-\pi+\pi(1+i)}\pi \tag{54}$$

Note that the interval $(\pi, \tilde{\alpha})$ is always non-empty since the fraction in (54) is always bigger than one. In equilibrium the bank must choose a reserve ratio lying between π and $\tilde{\alpha}$ to satisfy the conditions on the interest rates.

stationary equilibrium: As usual stationarity implies $\phi/\phi_{+1} = \gamma$ and thus the loan rate must equal the inflation tax or $1+i = \gamma/\beta$ from (49) as in the baseline model. Also market clearing for outside money implies $d = M_{-1}$. However, since we have two interest rates on inside money now but no additional equation we are missing one equation to pin down the equilibrium reserve ratio α^*. Below we will derive it assuming that equilibrium interest rates and leverage are chosen to maximize the expected steady state utility of a representative buyer. But let us first characterize the possible equilibria in the three cases:

a) costless inside money: We know in this case the interest on outside money equals the inflation tax, i.e. $i_{d2} = i = \gamma/\beta - 1$. From (53) and (52) this implies $\alpha = \pi$ and $i_{d1} = 0$ and consumption levels and real inside money holdings are $q = q^* = 1$, $q_c = \frac{1}{\gamma/\beta} = \phi n$. In this case the reserve ratio α is at the lowest possible value (π). The return for late redemptions is so high that holding inside money is costless and the return for early redemptions equals the return on outside money. Consumption in the monitored market is thus efficient and consumption in the non-monitored market is as low as in a pure outside money economy. There is no consumption risk-sharing between the two types of meetings for buyers.

c) costly inside money without spread: In this case the reserve ratio is at the highest possible value $\tilde{\alpha}$ such that the returns on early and late redemptions and the consumption levels in both types of meetings are equalized: $1 + i_{d1} = 1+i_{d2} = 1+i_d = \frac{\tilde{\alpha}}{\pi}$ and $q = q_c = \tilde{q}$. The bank chooses the maximal reserve ratio such that if she pays out all outside money early the return on early redemptions is not higher than for late redemptions. This achieves perfect consumption risk-sharing.

b) costly inside money with spread: In this case the bank chooses a reserve ratio between the two extreme cases i.e. $\alpha \in (\pi, \tilde{\alpha})$. Using the conditions from above, (47), this implies:

$$q^c = \frac{\alpha}{\pi(1+i)} \quad , \quad q = \frac{1-\alpha}{1-\pi}$$

and $q^* > q > \tilde{q} > q^c$ i.e. consumption in the monitored meeting is higher than in case *c)* but lower than in case *a)* and consumption in the non-monitored meeting is higher than in case *a)* but lower than in case *c)*. In this case the bank chooses some intermediate level of consumption risk-sharing.

Next we derive the optimal choice of the reserve ratio α^*. As mentioned above we will assume banks/buyers choose equilibrium interest rates and reserve ratio to maximize expected buyer steady state utility. Expected utility is given by:

$$W = \pi(u(q^c) - h^c) + (1-\pi)(u(q) - h) \tag{55}$$

where h^c (h) is the work effort of the buyer in the second market who has been in a non-monitored (monitored) meeting. From market clearing in the second market we must have $h^c = x^c$, i.e. work effort of buyers who have been in a non-monitored meeting must equal second market consumption of a seller in such a meeting. From the TIOLI-assumption we then have $x^c = \phi n(1 + i_{d1}) = q^c$. Analogue to this $h = x = \phi n(1 + i_{d2}) = q$ for agents in the monitored meetings and we can rewrite (55) as the formulation commonly used in models building on Lagos and Wright [2005] where welfare only depends on the surplus in the fist market:

$$W = \pi(u(q^c) - q^c) + (1-\pi)(u(q) - q) \tag{56}$$

The optimal reserve ratio $\alpha \in [\pi, \tilde{\alpha}])$ maximizes (56) using $u(q) = \ln(q)$, $q = \frac{1+i_{d2}}{1+i} = \frac{1-\alpha}{1-\pi}$ and $q^c = \frac{1+i_{d1}}{1+i} = \frac{\alpha}{\pi(1+i)}$. Note that the formulation of q and q^c includes the corner cases *a)* and *c)*.

First we will show that the optimal reserve ratio must be interior, i.e. $\alpha^* \in (\pi, \tilde{\alpha})$. Partially differentiating (56) with respect to α yields:

$$\frac{\pi - \alpha}{\alpha(1-\alpha)} + \frac{i}{1+i} \tag{57}$$

As can easily be checked the derivative is positive (negative) at $\alpha = \pi$ ($\alpha = \tilde{\alpha}$). Thus the marginal welfare effect of increasing α is positive (negative) at π ($\tilde{\alpha}$) and α^* must be interior. Setting (57) to zero then yields the optimal reserve ratio α^*:

$$\alpha^* = \frac{\sqrt{1 + 4\pi(1+i)i} - 1}{2i} \tag{58}$$

Clearly α^* depends positively on the share of cash transactions in the economy, π, and $\alpha^* \to 0$ as $\pi \to 0$ and $\alpha^* \to 1$ as $\pi \to 1$. So banks hold a higher share of outside money to the inside money issued if the fraction of buyers needing outside money is higher. On the other hand the partial derivative of the optimal reserve ratio with respect to the loan rate (the inflation tax) is also positive. The partial derivative is given by:

$$\frac{\partial \alpha^*}{\partial i} = 0.5 \frac{\sqrt{1 + 4\pi(1+i)i} - 1 - 2\pi i}{i^2 \sqrt{1 + 4\pi(1+i)i}} = \frac{\alpha^* - \pi}{i(2i\alpha^* + 1)} \tag{59}$$

The numerator is positive since $\alpha^* > \pi$. So if inflation rises (i goes up) the banks hold more outside money with respect to the inside money issued. This might seem paradoxical: if inflation goes up, holding outside money becomes more costly so why should banks hold relatively more? The answer again lies in the fact that banks find it optimal to do some consumption risk-sharing between the two types of meetings. Suppose inflation goes up but banks would not respond and keep the reserve ratio at the initial level. From the expressions for the interest rates, (52) and (53) we know the rate for early redemptions i_{d1} is unaffected and the rate for late redemptions i_{d2} rises with inflation. In fact since real inside money holdings are just $\phi n = \frac{1}{1+i}$ with ln-utility, consumption in monitored meetings, $q = \phi n(1 + i_{d2})$, would be unaffected and consumption in non-monitored meetings, $q^c = \phi n(1 + i_{d1})$ would decrease. Thus buyers in non-monitored meetings would be fully hit by the increase in inflation while buyers in monitored meetings would not be affected. Why is this not optimal? Since agents are risk averse and $q^c < q$, i.e. the marginal utility of consumption in the non-monitored market is higher, the bank finds it optimal to mitigate the consumption decline for non-monitored buyers although this is "inefficient" from the point of view of the sum of expected consumption in both types of meetings. Consumption in both types of meetings decreases but in the non-monitored meeting it decreases a little bit less than without increasing α^*.

Finally we want to see whether this model generates the same negative correlation between the reserve ratio and welfare as in the baseline model. In proposition 1 we found if α went up, expected welfare decreased. In this model an increase in α^* can have two sources as we just saw: an increase in the share of outside money transactions or an increase in inflation. How does welfare change if we increase α^* through increasing π or i? We can rewrite expected equilibrium welfare as a function of the parameters only where α^* is given by (58):

$$W(\alpha^*, i, \pi) = \pi(\ln(\frac{\alpha^*}{\pi(1+i)}) - \frac{\alpha^*}{\pi(1+i)}) + (1-\pi)(\ln(\frac{1-\alpha^*}{1-\pi}) - \frac{1-\alpha^*}{1-\pi}) \quad (60)$$

By the envelope theorem we can ignore the indirect effects of the parameters over α^* and partially differentiate (60) with respect to i and π to get the total effects. We get that $\frac{dW}{di} < 0$ and $\frac{dW}{d\pi} < 0$ i.e. expected welfare decreases both in inflation and in the share of buyers in the non-monitored market. Since higher inflation decreases consumption in both types of meetings the former is not surprising. The share of non-monitored meetings has two effects on welfare: it shifting weight from monitored to non-monitored meetings (this effect is negative since the surplus in monitored meetings is higher) and it decreases (increases) consumption in non-monitored (monitored) meetings over the interest rates. As it turns out the effects on the consumption levels just offset each other and the aggregate effect is negative. The model with early redemptions thus confirms the basic logic of the baseline model.

Appendix B

B.1 proof of proposition 1

Proof. To derive (27) we conjectured that $i = i'_d > i_d > 0$ in equilibrium. Note that if $\alpha \in (0,1)$ and $\gamma > \beta$ equilibrium interest rates given by (24), (25) and (26) satisfy this. Given that $i > i_d$ the buyer will always use all his money in the first market, $\lambda = \phi(i - i_d) > 0$. Then (11) and (6) are strictly concave in q_b, l_b, m and the first order conditions are sufficient for a unique maximum. Market clearing in the first market (21) pins down a unique q_s even if $c(q_s)$ is not strictly convex (and (16) strictly concave). Thus the solution to (27) must be unique.

Comparative statics: Differentiate the left-hand side of (27), $\frac{u'(q_b)}{c'(q_s)}$, with respect to q_b

$$\frac{\partial(\frac{u'(q_b)}{c'(q_s)})}{\partial q_b} = \frac{u''(q_b)c'(q_s) - u'(q_b)c''(q_s)\frac{1-s}{s}}{c'(q_s)^2} < 0 \tag{61}$$

From the strict concavity of $u(q)$, $u''(q) < 0$ and therefore (61) must decrease in q_b. At the first best allocation q^* (1) $\frac{u'(q^*)}{c'(q_s)} = 1$ and at any q_b solving (27) $\frac{u'(q_b)}{c'(q_s)} = \frac{1+i}{1+i_d} > 1$ since $i > i_d$ for $\alpha \in (0,1)$ and $\gamma > \beta$. Thus $\frac{u'(q^*)}{c'(q_s)} < \frac{u'(q_b)}{c'(q_s)}$ and therefore $q_b < q^*$ from (61). Thus we showed that for any $i > i_d$ the allocation will be inefficient. By the same argument any change *increasing* the right-hand side of (27) higher above 1 will decrease q_b further from q^*. Since $\frac{1+i}{1+i_d}$ increases in α we must have $\frac{\partial q_b}{\partial \alpha} < 0$ i.e. equilibrium consumption and welfare decreases in α. Also since $\frac{1+i}{1+i_d}$ increases in γ/β we must have $\frac{\partial q_b}{\partial \gamma/\beta} < 0$ i.e. equilibrium consumption and welfare decreases in the inflation tax.

iii)tov): The right-hand side of (27) $\gamma/\beta\frac{1}{1+i_d}$ is strictly below the right-hand side of an economy without banks (42) $\frac{\gamma/\beta - s}{1-s}$. Therefore by the same argument as above q_b (and welfare) are higher under fractional reserve banking than without banks. The results for *iv*) and *v*) are simply obtained by sticking in the limit values of α $(0,1)$ into the right-hand side of 27. At $\alpha = 1$ the right-hand sides of (43) in appendix A.2 and in (27) coincide.

■

B.2 proof of proposition 2

Proof. If $\alpha \in (0,1)$ and $\gamma > \beta$ equilibrium interest rates under constrained fractional reserve banking (37) and (36) satisfy $i > i_d > 0$. Then the same steps as in the first proof apply.

To verify *ii*) we use the same logic as in the first proof. To show that q_b and welfare in (38) is below first best q^* it is sufficient to show that the right-hand side of (38) is above 1. This holds if $\alpha \in (0,1)$ and $\gamma > \beta$. The comparative statics follow the same logic as above. If the partial derivative of the right-hand side of (38) with respect to the parameters s, α or γ/β increases the interest rate spread $\frac{1+i}{1+i_d}$ resp. the right-hand side of (38), q_b and welfare decreases. We have:

$$\frac{\partial \frac{1+i}{1+i_d}}{\partial \gamma/\beta} = \frac{\alpha(1-\alpha s)}{((1-\alpha)\gamma/\beta + \alpha(1-s))^2} > 0$$
$$\frac{\partial \frac{1+i}{1+i_d}}{\partial \alpha} = \frac{(1-s)^2\alpha + \gamma/\beta(\gamma/\beta - 1))}{((1-\alpha)\gamma/\beta + \alpha(1-s))^2} > 0$$
$$\frac{\partial \frac{1+i}{1+i_d}}{\partial s} = \frac{\gamma/\beta(1-\alpha) + \alpha^2(\gamma/\beta - 1))}{((1-\alpha)\gamma/\beta + \alpha(1-s))^2} > 0$$

Thus indeed $\frac{\partial q_b}{\partial \gamma/\beta} < 0$, $\frac{\partial q_b}{\partial \alpha} < 0$ and $\frac{\partial q_b}{\partial s} < 0$.

iii): As s goes to zero q_b given by (38) converges to the allocation without preference shock (27). Since increasing s decreases q_b from *ii*) q_b and welfare under unconstrained fractional reserve banking (27) must be higher.

iv): Compare the right-hand side of the no-bank equilibrium (42) with the right-hand side of (38) to see that the former is strictly higher for $\alpha \in (0,1)$ and identical for $\alpha = 1$.

v): Finding $\tilde{\alpha}$ requires equalizing the right-hand sides of the narrow banking equilibrium allocation (43), see proposition 1 *iv*) and the constrained fractional reserve banking equilibrium (38):

$$\frac{\gamma}{\beta} = \frac{\gamma/\beta - \tilde{\alpha}s}{(1-\tilde{\alpha})\gamma/\beta + \tilde{\alpha}(1-s)}$$

which yields

$$\tilde{\alpha} = \frac{\gamma/\beta}{\gamma/\beta + s}$$

For *vi*) use $\alpha = 0$ in the right-hand side of (38) which then yields q^* from (1).

■