Should Banks Create Money?

Christian Wipf*

April 2021

Abstract

The paper compares the welfare properties of the current fractional reserve banking (FB) system versus a narrow banking (NB) system where inside money is fully backed by reserves issued by the central bank. The analysis shows that under sufficient competition FB is beneficial compared to NB because of higher interest payments on inside money. Since under FB inside money funds loans, banks have higher income on their asset side which they pass on to inside money holders in the form of higher interest if competition is sufficiently high. This improves welfare because it compensates the inside money holders against inflation. A calibrated version of the model suggests however that the welfare gains of FB are relatively small, below 0.15% of GDP p.a. in the US between 1984 and 2008. (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

The appropriate role or the "division of labor" between central banks and commercial banks in the money supply of an economy has been increasingly debated since the recent financial crisis and structural changes like the decrease in cash use. In Switzerland for example there was a public referendum on forcing banks to fully back their monetary liabilities with reserves, i.e. imposing a 100% reserve narrow banking (NB) system.[1] And many central banks think about offering electronic check money not only to banks but also to the general public with go under the name of "Central Bank Digital Currencies" (CBDC) or "reserve for all" (RFA) proposals. Under such a system there would be no direct regulation on banks liability issuance but they would face competition from central banks.[2]

Usually the topic is phrased as an efficiency vs. stability tradeoff. The current fractional reserve banking (FB) system, where commercial banks provide the majority of the money supply in the form of demand deposits partially backed by reserves, is supposed to have efficiency benefits but at the cost of increasing financial instability. Our understanding of this tradeoff however, is rather one-sided. While, following the seminar paper by Diamond and Dybvig [1983] we understand the fragility of FB relatively well, the potential benefits of FB are much less clear. For example Brunnermeier and Niepelt [2019] argue any FB allocation can be replicated with an NB or CBDC allocation given appropriate transfers and open-market operations by a fiscal authority and a central bank. This echoes the quantity theory which essentially argues that the quantity and the composition of money should be irrelevant for real allocations and welfare. On the other hand FB dominated monetary history and the classical analysis of Diamond and Dybvig [1983] suggest it should increase beneficial long-term investment. Finally some authors like Chari and Phelan [2014] argue FB is inefficient and explain the historical dominance of this system with a pecuniary externality which makes it privately but not socially optimal. Against this background the paper abstracts from instability and focuses on the potential benefits of FB. The specific questions are: Does FB have a socially useful role compared to NB proposals like 100% reserves? If yes, what is it? And how important is it quantitatively?

To address these questions I first build a banking model based on Berentsen et al. [2007] with the following main ingredients: A central bank issues potentially interest bearing reserves and controls inflation and the nominal interest rate. Commercial banks issue deposits and bonds, finance loans and hold reserves. Deposits are usable as money while bonds are illiquid. Under FB banks partially

[1] For the policy discussion in other countries including Iceland and the Netherlands, see `https://sovereignmoney.site/` and `https://internationalmoneyreform.org/`. See Pennacchi [2012] for an overview of the history of narrow banking and related proposals.

[2] For an overview on CBDC-proposals, see Niepelt [2020] and Bank for International Settlements [2018]. Besides a lot of conceptual work a few central banks (e.g. Uruguay, Sweden) have already implemented first pilot projects.

back deposits with reserves while under NB they fully do it. Banks have market power in the deposit market. Finally there are households who use deposits or reserves to buy consumption goods. Holding money is costly in the economy, i.e. the central bank sets a positive nominal interest rate and does not run the Friedman rule. This "inflation tax" is the basic inefficiency in the model. To quantify the implications of the model I calibrate it with US-data from 1984–2008.

The main theoretical result is that FB is socially useful compared to NB proposals if the banking sector is sufficiently competitive (otherwise the two systems are equivalent). The economics behind this result is that FB provides a means of payment with a higher return (a higher interest rate on deposits) than NB. Under FB banks fund loans with deposits and since these normally pay a higher return that reserves banks have a higher income on their asset side under FB. With sufficient competition they pass on this higher income to their liability side in the form of higher interest payments on deposits. This is beneficial because it compensates the deposit holders against for inflation. The calibrated version of the model suggests however, that the welfare gains of FB due to higher interest payments are relatively small. Around 0.1%/0.3% of GDP at a nominal interest rate of 5%/10% and below 0.15% of GDP p.a. in the US between 1983 and 2008. If the central bank pays interest on reserves these welfare gains are even smaller. The analysis suggests that the efficiency benefits of FB are relatively small. Conversely this means that – to the extend that NB proposals are able to provide financial stability – the costs of financial safety are relatively small.

The rest of the paper is organized as follows: Section 2 shows the basic environment. Then the model (section 3) is presented under perfect (section 3) and imperfect (section 4) competition and section 5 addresses interest on reserves. Finally I calibrate the model to US data in 6.

2 Environment

The environment is based on Lagos and Wright [2005] and follows Berentsen et al. [2007]. Time is discrete and continues forever. Every period is divided into two sequential competitive markets called *first market* (FM) and *second market* (SM). There is a perishable consumption good q in the FM and x in the SM.

There is a unit mass of infinitely lived households or agents. They discount future periods with β, are anonymous and cannot commit. At the beginning of every period agents face a *preference shock* which determines what they can do in the FM. With probability $s \in (0, 1)$ an agent is a *seller* and can only produce (weakly) convex disutility of production $c(q)$. With the inverse probability $1 - s$ an agent is a *buyer* and can only consume with strictly concave utility of consumption $u(q)$, satisfying the Inada-conditions. In the second market all agents can consume x and produce h with a quasi-linear utility function $U(x) - h$.

$U(x)$ is strictly concave in x and also satisfies the Inada-conditions.

The efficient quantities for consumption in the FM, q^* and q_s^*, and in the SM, x^*, are given by equalizing marginal utility of consumption with marginal disutility of production in both markets:

$$\frac{u'(q^*)}{c'(\frac{1-s}{s}q^*)} = 1, \quad q_s^* = \frac{1-s}{s}q^*, \quad U'(x^*) = 1 \tag{1}$$

Since households are anonymous buyers cannot issue debt to buy goods from sellers in the FM. Thus buyer find it useful to hold money, which is provided by a *central bank* and by *commercial banks* described next.

The central bank issues outside fiat money M or reserves. Reserves are potentially interest bearing with interest rate i_m and the growth rate is $\gamma = \frac{M}{M_{-1}}$.[3] The central bank manages the outside money supply by lump-sum transfers τ to agents in the SM.

Outside money is the numeraire in the economy. Let $1/\phi$ be the price of x in terms of outside money in the SM (ϕ is the value of money) and let p be the price of q in terms of outside money in the following SM. Denote (gross) inflation π as the ratio of prices between two consecutive SM, i.e. $\pi = \frac{\phi_{-1}}{\phi}$. Under stationarity $\phi M = \phi_{-1}M_{-1}$ and therefore $\gamma = \pi$. So the (long-run) inflation rate equals the growth rate of the outside money supply and the central bank perfectly controls inflation and the nominal interest rate, $\frac{\gamma}{\beta}$, the product of inflation and the real interest rate $1/\beta$.

Throughout the paper we assume that holding money is costly. This implies the nominal interest rate γ/β which is the opportunity cost of holding money must be positive (inflation above the Friedman rule) and above the interest rate on reserves (if $i_m = 0$ the two conditions coincide).

$$\frac{\gamma}{\beta} - 1 > 0 \quad , \quad \frac{\gamma}{\beta} - 1 > i_m \tag{2}$$

This assumption introduces a basic inefficiency into the environment in the form of an "inflation tax" on real FM activity. Note that the preference shock aggravates this basic inefficiency. If holding money is costly, the risk to be a seller with (costly) idle money holdings in the FM makes acquiring money even less attractive ex-ante.

Commercial banks are profit-maximizing firms. In contrast to households, they can commit and monitor agents at no cost. The first property enables them to issue debt and the second property to make loans. Bank assets are loans l

[3] In the whole paper I denote variables in a representative period t without subscript, the period before t with subscript -1 and the period after with $+1$.

(interest rate i) and reserves m. They issue deposits d (interest rate i_d) and bonds (interest rate i_b). Deposits are usable as means of payment in the FM (i.e. they are liquid) but bonds are not.[4] On the other hand deposits are subject to a reserve requirement. A bank issuing d deposits must back them at least with a fraction α of reserves, i.e. $\alpha d \leq m$. Generally the reserve requirement will bind and therefore if $\alpha < 1$ we will have $d > m$ and banks are in a FB system if $\alpha = 1$ we will have $d = m$ and we are in a NB system. Under NB banks can finance loans only with bonds while under FB the finance them with bonds and deposits.

All contracts are specified in terms of outside money (i.e. they are nominal), formed in the FM after the preference shock and fully redeemed in the following SM.[5] In the main text I assume an infinite amount of perfectly competitive banks. In section 4 and in the calibration I also introduce and use Cournot competition with N banks in the deposit market while keeping the loan/bond market perfectly competitive. Banks have two roles in this environment. They create new liquidity (issue money) and they can reallocate liquidity after the preference shock, i.e. they provide insurance against liquidity risk.

Figure 1 summarizes the sequence of events in this economy in a representative period t: In the SM in $t-1$ households acquire outside money m (due to quasi-linear preferences they will acquire the same money holdings). Then, after the preference shock, they interact with banks. Buyers need money to buy goods in the following SM. By depositing and borrowing they acquire deposits d or more reserves, depending on the rate of return of the two means of payment. Sellers on the other hand acquire bonds b because these have a higher return than deposits/reserves and sellers don't need money in the FM. In the following SM sellers redeem deposits and bonds and buyers repay loans l. And the acquire new (identical) money holdings m_{+1}.

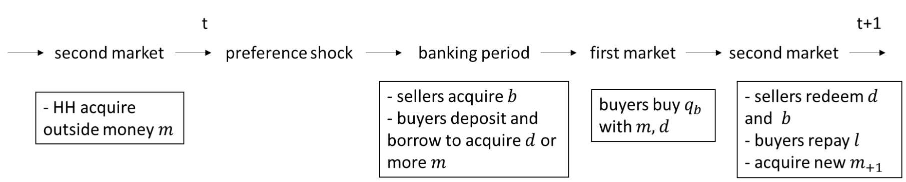


Figure 1: sequence of events

[4]Deposits might be interpreted as checkable deposits and bonds would be less liquid forms of liabilities like time deposits.

[5]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.

3 The Model

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:

$$\max_{l,d,d'} = d + d' + l(1+i) - (l+d)(1+i_d) - d'(1+i'_d) \qquad (3)$$
$$s.t. \quad \alpha(l+d) \leq d + d'$$

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:

$$l = \frac{d'}{\alpha} + \frac{1-\alpha}{\alpha} d \qquad (4)$$

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:

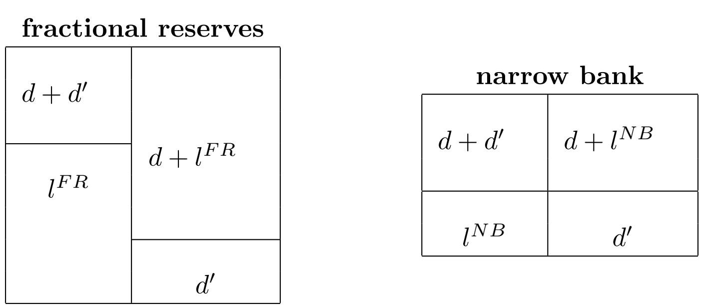
$$1 + i_d = \alpha + (1 - \alpha)(1 + i) \tag{5}$$
$$1 + i'_d = 1 + i \tag{6}$$

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'.



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 (4). 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 (5). 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, labor 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.

$$W(m, n, d', l) = \max_{x,h,m_{+1}} U(x) - h + \beta V(m_{+1}) \tag{7}$$
$$s.t. \quad x + \phi m_{+1} = h + \phi(\tau + m) + n\phi(1+i_d) + d'\phi(1+i'_d) - l\phi(1+i)$$

The first-order condition for consumption in the second market is $U'(x^*) = 1$ and is thus always efficient.[6] The first order condition for optimal (positive) outside money holdings then solves:

$$\phi = \beta V'(m_{+1}) = sV'_s(m_{+1}) + (1-s)V'_b(m_{+1}) \tag{8}$$

(8) 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:

$$\begin{aligned} W_m &= \phi \\ W_n &= \phi(1+i_d) \\ W'_d &= \phi(1+i'_d) \\ W_l &= -\phi(1+i) \end{aligned} \tag{9}$$

[6]For this to hold x^* needs to be big enough such that h is positive for both agents. This can be achieved by an appropriate scaling of $U(..)$.

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

$$(1-s)x_b + sx_s = x^* = (1-s)h_b + sh_s0 \tag{10}$$

$$m_{+1} = M \tag{11}$$

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.

$$V_b(m) = \max_{q_b, l_b, d_b, d_b'} \quad u(q_b) + W(m - d_b - d_b', l_b + d_b - pq_b, d_b', l_b) \tag{12}$$

$$s.t. \quad pq_b \leq d_b + l_b$$

$$d_b + d_b' \leq m$$

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 (9) 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 (9) and with λ denoting the multiplier for the constraint):

$$q_b: \qquad u'(q_b) = p(\phi(1+i_d) + \lambda)$$

$$l_b: \qquad \lambda + \phi(1+i_d) = \phi(1+i)$$

$$d_b: \qquad \lambda + \phi(1+i_d) \geq \phi(1+i_d')$$

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 (12) is given by:

$$u'(q_b) = p\phi(1+i) \tag{13}$$

$$l_b = pq_b \tag{14}$$

$$d_b = 0 \quad , \quad d'_b = m \tag{15}$$

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

$$V'_b(m) = \phi(1 + i'_d) \tag{16}$$

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.

$$V_s(m) = \max_{q_s, l_s, d_s, d'_s} \quad -c(q_s) + W(m - d_s - d'_s, d_s + l_s + pq_s, d'_s, l_s) \tag{17}$$

$$s.t. \quad d_s + d'_s \leq m$$

The envelope conditions (9) 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

$$c'(q_s) = p\phi(1 + i_d) \tag{18}$$

$$l_s = 0 \tag{19}$$

$$d_s = 0 \quad , \quad d'_s = m \tag{20}$$

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

$$V'_s(m) = \phi(1 + i'_d) \tag{21}$$

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:

$$\begin{aligned} d' &= (1-s)d'_b + sd'_s = m \\ d &= (1-s)d_b + sd_s = 0 \\ l &= (1-s)l_b \\ (1-s)q_b &= sq_s \end{aligned} \tag{22}$$

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

$$(1-s)l_b = \frac{m}{\alpha} \tag{23}$$

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 (16) and for a seller (21) with the condition for optimal outside money holdings (8) to get:

$$\phi = \beta\phi_{+1}(1 + i_{d'+1}) \tag{24}$$

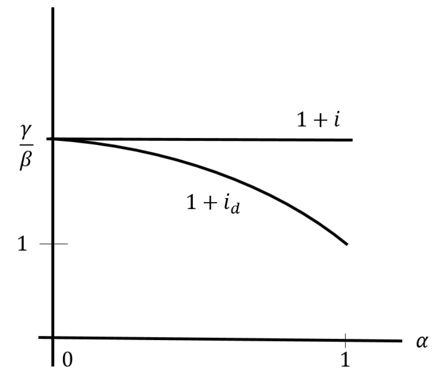
To get the equilibrium interest rates we apply stationarity ($\gamma = \phi/\phi_{+1}$) to (24) and use the relations on interest rates from the bank problem, (5) and (6).

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

$$1 + i'_d = \frac{\gamma}{\beta} \tag{26}$$

$$1 + i_d = (1-\alpha)\frac{\gamma}{\beta} + \alpha \tag{27}$$

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 (13) with optimal production (18) and use the equilibrium expressions for the interest rates and market clearing (22):

$$\frac{u'(q_b)}{c'(\frac{1-s}{s}q_b)} = \frac{1+i}{1+i_d} = \frac{\gamma/\beta}{(1-\alpha)\gamma/\beta+\alpha} \tag{28}$$

The rest of the equilibrium allocation is then given by:

$$q_s = \frac{1-s}{s}q_b \tag{29}$$

$$\phi N = Q\frac{c'(q_s)}{1+i_d} \tag{30}$$

$$\phi M_{-1} = \alpha\phi N \tag{31}$$

$$x_b = x_s = x^* \tag{32}$$

$$h_s = x^* - \phi M_{-1}i - c'(q_s)q_s \tag{33}$$

$$h_b = x^* + \frac{q_b c'(q_s)(1+i)}{1+i_d} - \phi M_{-1}i \tag{34}$$

Where $N = (1-s)l_b = \frac{M_{-1}}{\alpha}$ denotes total inside money (total loans) issued and $Q = (1-s)q_b = sq_s$ is the total amount consumed in the first market. Furthermore we have that total deposits against non-monetary debt d' equal the total stock of outside money $d' = M_{-1}$ and total deposits against inside money are zero. After the banking period banks hold all outside money in this economy. As already mentioned earlier we need a scaling condition on $U(..)$ to ensure that x^* is sufficiently high for $h_s > 0$. Since second market consumption is always efficient at x^* we can focus on the allocation in the first market for welfare analysis. 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 stationary equilibrium with banks, inside money and non-monetary debt in which:*

i) first market consumption solves (28)

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

[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 28.

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-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, (24), 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 (24) 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.

$$0 < i_d, i'_d < \gamma/\beta - 1 \tag{35}$$

Any interest rates satisfying (35) 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.

[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.

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

$l_b \in (pq^*, \infty)$. The economy achieves the first best allocation equivalent to an economy with direct credit.

4 Imperfect Competition

(incomplete)

The approach here follows Chiu et al. [2019]. Cournot competition with N banks in the deposit market but the bond/loan market is still perfectly competitive.

The equilibrium loan rate i and bond rate i_b still equal nominal interest rate:

$$i = i_b = \frac{\gamma}{\beta} - 1$$

But the deposit rate is now given by:

$$i_d = \begin{cases} i_m & \text{if} \quad k \geq \bar{k} \\ \frac{1+i_d^{PC}}{1+k} - 1 > i_m & \text{if} \quad k < \bar{k} \end{cases}$$

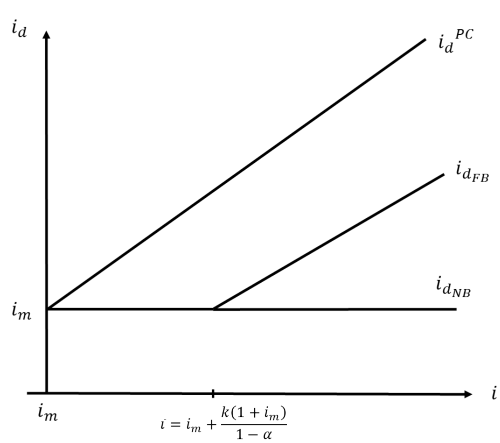
where k is the markup due to imperfect competition

$$k = \frac{1}{\epsilon_D N} \quad , \quad \epsilon_D = \frac{(1+i_d)}{\phi D}\frac{\partial \phi D}{\partial i_d} \quad , \quad \bar{k} = \frac{(1-\alpha)(i-i_m)}{1+i_m}$$

and i_d^{PC} is the interest on deposits under perfect competition from the last section:

$$i_d^{PC} = (1-\alpha)i + \alpha i_m \quad \in (i_m, i)$$

The following figure plots the interest on deposits under imperfect competition:



The welfare results are modified in the following way:

$$\begin{array}{lll} W_{FB} > W_{NB} & \text{if} \quad k < \bar{k}, i > \bar{\iota} & i_{d_{FR}} > i_{d_{NR}} = i_m \\ W_{FB} = W_{NB} & \text{if} \quad k \geq \bar{k}, i \leq \bar{\iota} & i_{d_{FR}} = i_{d_{NR}} = i_m \end{array}$$

FB dominates NB only if the deposit market is sufficiently competitive resp. the nominal interest rate is sufficiently high.

5 Interest on reserves

The results from Brunnermeier and Niepelt [2019] suggest that every FB allocation can be replicated with a NB 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 a FB equilibrium given by 1 and adapt a NB system, i.e. change α to 1. From proposition 1 we know that this reduces q_b and welfare compared to the original FB allocation. Can the central bank replicate the initial allocation? Suppose she pays interest i_m on outside money in the second market.[11] 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:

$$\tau + i_m M_{-1} = M - M_{-1} = (\gamma - 1) M_{-1} \tag{36}$$

Now reconsider the bank problem (3). If $i_m < i_d$ and $i_m < i'_d$ (and $i > i_d$), the reserve constraint (4) will bind. Then free entry (zero profits) yields a similar relationship to (5) between the i_d and the other interest rates:

$$i_d = \alpha i_m + (1 - \alpha) i \tag{37}$$

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. On the other hand, the relationship between the interest rate on non-monetary debt and the loan rate is unchanged. As in equation (6) we have $i'_d = i$. With interest on outside money commercial banks can pay interest on inside money even under NB (with $\alpha = 1$) now. From (37) we get

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

that $i_d = i_m$ under NB. The rest of the problem being unchanged, equilibrium consumption under NB with interest on outside money is still given by (28).

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

Thus indeed a NB economy where the central bank pays interest on outside money can replicate any FB allocation given by (28). The condition is that the central bank pays an interest rate on outside money equalling the interest on inside money commercial banks pay under FB, i.e. $i_m = i_d = (1-\alpha)i$. According to (36) the central bank finances these interest payments by lower transfers or higher taxes in the SM. In equilibrium this means buyers work more and sellers work less in the SM. But these changes do not matter for welfare since both have linear disutility.

This answer however, is incomplete, in the sense that we compared a NB economy with $i_m > 0$ with a FB economy without this feature. Suppose we allow for interest payments on outside money in both systems. Does the equivalence result still hold? The answer is no. From (37) we see that $i_d > i_m$ as long as $i_m < i$. In other words, as long as the central bank interest payments are below the nominal interest rate interest on inside money under FB (i_d) will be higher than interest payments on inside money under NB (i_m). Thus for any $i_m < i$ the FB allocation dominates the NB allocation and the results from proposition 1 go through. The NB allocation is never equivalent to the FB 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.[12]

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

6 Quantitative Assessment

We now want to get a sense of how quantitatively important these social benefits of fractional reserve banking due to higher interest on deposits are. Following

[12] 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.

the literature on the welfare costs of inflation pioneered by Lucas [2000],[13] we calibrate the model to the US economy using yearly data on interest rates, monetary aggregates and nominal GDP. Thus we interpret the model as representing yearly steady states given by particular combinations of the nominal interest rate i and the interest rate on inside money i_d. We construct a velocity-equation relating money balances held to nominal GDP to which we fit the parameters s, σ, B and N. In the model total money balances held are total deposits D which also equals nominal output in the FM. Nominal output in the SM is given by x^*/ϕ so total nominal GDP is given by $N + x^*/\phi$. This gives us the following velocity- or real money demand equation:

$$v = \frac{D}{D + x^*/\phi} = \frac{1}{1 + \frac{x^*}{\phi D}} \tag{39}$$

Since real deposit holdings ϕD decrease in the nominal interest rate i, v also decreases in i. Following the literature we parametrize $u(q) = \frac{q^{1-\eta}}{1-\eta}, c(q) = q$ and $U(x) = B\log(x)$. This implies.

$$x^* = B \tag{40}$$

$$q_b = \left(\frac{1+i_d}{1+i}\right)^{1/\sigma} \tag{41}$$

$$\phi D = (1-s)\frac{(1+i_d)^{\frac{1-\sigma}{\sigma}}}{(1+i)^{\frac{1}{\sigma}}} \tag{42}$$

Note that $\frac{\partial \phi D}{\partial i_d} > 0$, i.e. real deposits decrease in the deposit rate if $\sigma < 1$ with elasticity $\epsilon_D = \frac{1-\sigma}{\sigma}$.

In the following we will use yearly, averaged US-data from 1984 until 2008.[14] What is the rationale for choosing this time period? From 1933 the Glass-Steagall Act prohibited commercial banks to pay interest on bank deposits under Regulation Q. With this regulation the basic mechanism of the model was prohibited. Only in the 1980s regulation Q was gradually lifted. Particularly, in 1982 banks were allowed to issue interest-bearing money market deposit accounts (MMDA). Following Lucas and Nicolini [2015] we assume that after two years (1984) banks had adjusted to the new instrument and had moved funds

[13] In the New Monetarist literature see in particular Lagos and Wright [2005], Craig and Rocheteau [2008] and Berentsen et al. [2015]. See also Lucas [2013] for similar arguments with a cash-in-advance model.

[14] Lucas and Nicolini [2015] argue that money demand models often fail to match the behaviour of real money demand at high frequency "in the sense that temporary changes in the short term interest rates have relatively little impact on real money balances" (p. 58). As a robustness exercise I thus also used HP-filtered data with a smoothing parameter of 100. While – unsurprisingly – the fit of such a model is slightly better, the parameter estimates and the welfare results are practically identical to using yearly data.

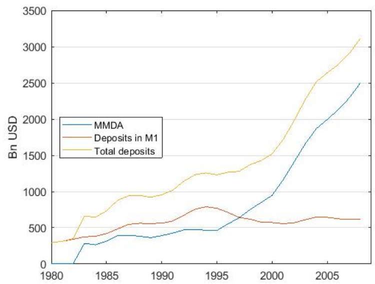
from normal deposit accounts to MMDA's. Thus from then on the basic mechanism of the model was (at least partly) operative. Finally, we restrict attention until 2008 because nominal interest rates were essentially zero after 2008 and we developed the model under the assumption of a positive nominal interest rate ($i > 0$). The following table summarizes the correspondence between model variables and the data:

Model	**Data**	**Data source**
deposits D	M1 – currency + MMDA	FRED, Lucas and Nicolini [2015]
reserves M	M0 – currency	FRED
nominal interest rate i	3-Month T-Bill rate	FRED
interest on deposits i_d	weighted average interest deposits in M1 + interest on MMDA	FRED Lucas and Nicolini [2015]

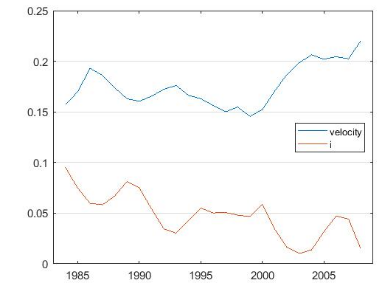
For inside money N we use the deposit component of M1 combined with the MMDA data provided by Lucas and Nicolini [2015]. We add MMDAs because they provide a similar economic function to traditional deposit accounts and as has been shown by Lucas and Nicolini [2015] money demand is not stable without this correction. We abstract from currency because the model is silent on this. For outside money M we use M0 minus currency. This is a measure of reserves hold by commercial banks and thus $\alpha = M/D$ measures bank reserve holdings. For the nominal interest rate i we use the 3-month T-bill rate. The interest rate on inside money i_d is calculated based on data from Lucas and Nicolini [2015] on the interest rate on the deposits in M1 and on MMDA accounts, weighted by the shares of the two aggregates.[15]

The following figure plots a) the evolution of MMDA, deposits in M1 and total deposits D since 1980, b) velocity v against the nominal interest rate, c) the evolution of bank reserve holdings α since 1980 and d) the nominal interest rate i, the interest rate on deposits from the data i_d and the interest rate implied by the theoretical model, $i_d = (1 - \alpha)i$ from (27).

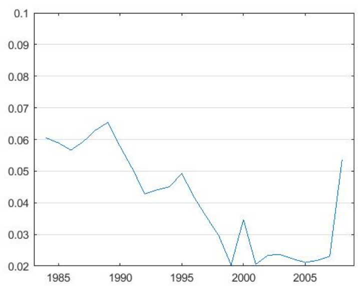
[15] We use a linear extrapolation for the years where the return is not specified.



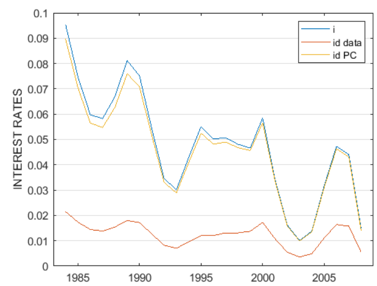
(a) deposit components



(b) velocity and i



(c) reserve ratio α



(d) i and i_d

Subfigure a) shows how important it is to include MMDA in total deposits. While still 0 in 1980 they strongly increase to over 280 bn USD in 1983, accounting for over 40% of total deposits. After 1995 MMDA's dominate the evolution of total deposits. The strong increase of MMDA after the mid-90s is related to the introduction of so called "sweep"-deposit accounts in 1994. This was another response to a relaxation of regulation Q. Sweep-accounts essentially allowed banks to automatically move funds from traditional deposit accounts to MMDA's.[16] Subfigure b) shows that bank reserve holdings α are close to zero. The mean for 1984 to 2008 is just 4%. Subfigure c) shows that the negative relationship between velocity and the nominal interest rate implied by the theory seems plausible for the data. As already pointed out above this hinges crucially on the inclusion of MMDA's in total deposits. Finally subfigure d) shows that the model clearly fails in explaining the evolution of the interest rate on deposits, i_d. The model predicts $i_d = (1-\alpha)i$, which implies essentially $i_d \approx i$ since α is decreasing towards zero in the considered period (see c)). In the data however, the interest rate on deposits is much lower. To account for this spread we estimate the imperfect competition model developed in section 4.

The calibration proceeds in two steps: First we estimate s, σ and B minimizing the sum of squared residuals between velocity in the data and in the model, (39):

[16] See Lucas and Nicolini [2015] and Berentsen et al. [2015].

$$\min_{s,\sigma,B} \quad \sum_{1984}^{2008} \left(\frac{D_t}{GDP_t} - \frac{1}{1 + \frac{B(1+i_{d_t})}{(1-s)q_{b_t}}} \right)^2$$

Second, we estimate the competition parameter N to fit the interest on deposits, taking the estimate of σ from above as given:

$$\min_{N\in\{1,2,\ldots,\infty\}} \quad \sum_{1984}^{2008} \left(i_{d_t} - \max\left\{ 0, \frac{1+i_{d_t}^{PC}}{1+k} - 1 \right\} \right)^2$$

The results of the first step are $s = 0.5, \sigma = 0.2$ and $B = 1.96$, which is very similar to similar calibrations.[17] The numbers imply $\epsilon_D = 0.25$ and $k = 0.4$. Note that the choice of s does not affect the results and following Lagos and Wright [2005] we will set it to $s = 0.5$.[18] The following figure plots velocity in the data and the estimated velocity from the model against the nominal interest rate. The overall fit is quite good. Only between the mid-1990s and 2002 the model has difficulties to match the data. The reason is that in this period velocity decreases and the nominal interest rate is constant or also slightly decreases. But since the theoretical model predicts a negative relationship between the two variables it cannot account for this development.

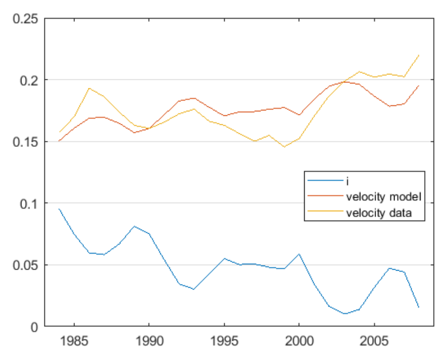
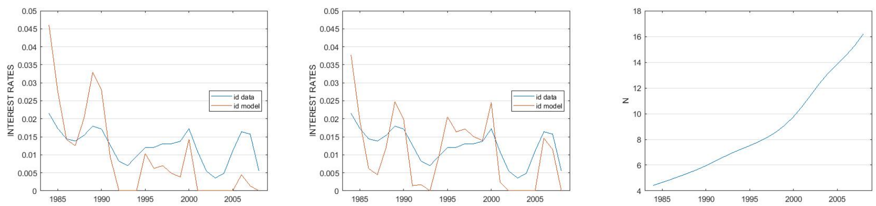


Figure 3: Model/data velocity against the nominal interest rate

[17]Lagos and Wright [2005] find $\sigma = 0.15, B = 1.97$ for a similar specification, Craig and Rocheteau [2008] find $\sigma = 0.14, B = 1.82$. The estimated parameters of Berentsen et al. [2015] are slightly higher, $\sigma = 0.31$ and $B = 2.2$, but they estimate the model to match average targets not yearly data.

[18]With the given functional forms the second term of the denominator in (39) reads $\frac{B(1+i_d)}{(1-s)q_b}$. Calibrating B and s separately will just estimate the same constant $K = B/(1-s)$ which can be satisfied for different combinations of B and s. B cannot be too small however (which implies s cannot be too small for a given K). Otherwise the condition for positive hours for sellers in the second market, $h_s > 0$ given by (33) cannot hold. Lagos and Wright [2005] arrive at the same conclusion.

The results for the competition parameter is $N = 6$ which (together with $\sigma = 0.2$) implies $\epsilon_D = 0.25$ and $k = 0.4$.[19] The following figure plots the interest rate on deposits i_d in the data against a) the model implied interest rate with $N = 6$ and b) with $N = 5$ between 1984–1991 and $N = 8$ between 1992–2008 and c) the implied yearly N to perfectly match the trend i_d using HP-filtered data with a smoothing parameter of 100.



(a) (b) (c)

Figure 4: (a) $N = 6$, (b) $N = 5$ (1984–1991), $N = 8$ (1992–2008), (c) implied yearly N to match trend i_d (HP-filtered data)

The figures show that the imperfect competition model has difficulties to match the interest rate in the data. The interest rate in the data is much less volatile than the interest rate from the model which closely follows the movements in the nominal interest rate. Also, assuming a constant N over the whole time period is probably not such a good approximation of reality. The fit with two different N for 1984–1991 and 1992–2008 is much better and also the implied N to match the trend in i_d is increasing. This indicates increasing competition in the sample period which seems plausible given the deregulation of the financial sector in this period. For the welfare calculations I use different values for N.

Now we quantify the welfare gains of FB. Following the welfare costs of inflation literature we ask: What fraction of steady state consumption/GDP Δ_{FB} would agents give up under FB to be equally well of as under NB? Thus Δ_{FB} solves:

$$\begin{aligned} W_{FB}(\Delta_{FB}) = 0.5[u(q_b \cdot \Delta_{FB}) - q_b] + U(x^* \cdot \Delta_{FB}) - x^* = W_{NB} \\ \text{with} \quad W_{NB} = 0.5[u(q_b|_{\alpha=1}) - q_b|_{\alpha=1}] + U(x^*) - x^* \end{aligned} \tag{43}$$

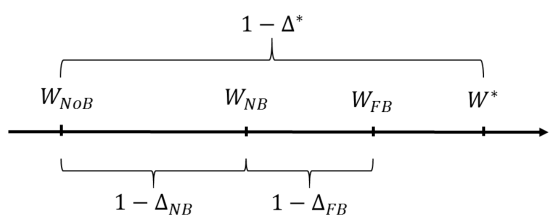
$1 - \Delta_{FB}$ then measures the benefits of FB due to a higher return on deposits. The following table shows these benefits for different levels of N setting α to the mean value of 4% and the interest rate on reserves to zero (this implies under NB the interest on deposits was 0):

[19] Chiu et al. [2019] get $N = 9$ in their calibration.

i	$1-\Delta_{FB}$ $(\alpha = 0.04, i_m = 0)$			
	$N=6$	$N=10$	$N=15$	$N=\infty$
2%	0%	0%	0%	0.02%
5%	0.02%	0.07%	0.09%	0.11%
10%	0.27%	0.32%	0.33%	0.35%
14.6%	0.56%	0.60%	0.62%	0.63%

Table 1: Welfare gains from FB due to higher i_d

The table shows that the welfare gains from FB due to higher interest payments are relatively small and increasing in the nominal interest rate. Around 0.1% of GDP at a $i = 5\%$ and around 0.3% of GDP at $i = 10\%$. At $i = 14.6\%$ – which corresponds to 10% inflation and a real interest rate using $\beta = 0.96$ – they are around 0.6% of GDP. The table also shows that the welfare effects of higher competition are relatively small. Going from a competition parameter of $N = 6$ to perfect competition ($N = \infty$) corresponds to differences in welfare gains below 0.1% of GDP.



To put these numbers into a broader perspective we also calculate the welfare gains of NB compared to an economy without banks as given in appendix A.1, denoted as $1-\Delta_{NB}$ where Δ_{NB} solves $W_{NB}(\Delta_{NB}) = W_{NoB}$ and W_{NoB} is the steady state welfare without banks. Since the main benefit of NB compared to no banks is liquidity insurance (without interest on reserves this holds exactly) $1-\Delta_{NB}$ mainly measures the benefits of liquidity insurance due to the banking sector. Furthermore we calculate the total welfare costs of inflation in the model, denoted by $1-\Delta^*$ which solve $W^*(\Delta^*) = W_{NoB}$. W^* is the steady state allocation with first best consumption q^*, which would correspond to a nominal interest rate of zero (Friedman rule). The figure above summarizes these relations and the following table contrasts the welfare gains of FB due to higher i_d with these other measures.

i	$1-\Delta^*$	$1-\Delta_{NB}$	$1-\Delta_{FB}$ ($\alpha = 0.04, N = 6$)
2%	0.07%	0.05%	0%
5%	0.36%	0.25%	0.02%
10%	1.04%	0.69%	0.27%
14.6%	1.69%	1.06%	0.56%

The table shows that the welfare gains from liquidity insurance ($1-\Delta_{NB}$) are significantly higher than the welfare gains due to higher interest payments on i_d under FB. Liquidity insurance compensates households for up to 75% of the total welfare costs in the economy while comparable numbers for the marginal benefits of FB, a higher interest on the means of payment, are significantly lower.

Finally we also calculate how much consumption households in the historical US-FB system would have given up between 1984 and 2008 to be equally well of as in a NB system. In contrast to above we take the historical spread $i-i_d$ as given here, and don't use N in the calculation. And we consider a NB-system without interest on reserves as the alternative. The following figure shows the $1-\Delta_{FB}$ and $1-\Delta_{NB}$ with the actual spreads:

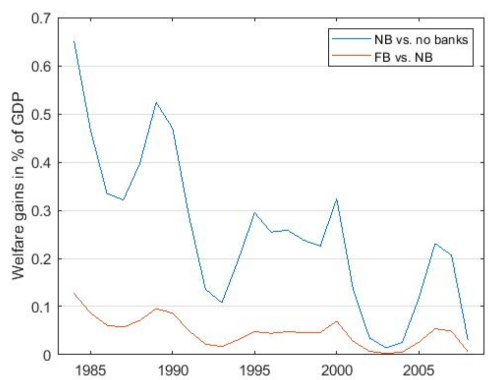


Figure 5: Historical welfare gains using i_d directly from US-data

Similar to before the historical welfare gains from having a FB system with higher interest payments is relatively small, below 0.15% of GDP around 1984 and even lower thereafter. Again, the gains from liquidity insurance are significantly higher.

So far we assumed the interest on reserves, i_m, was zero in both systems. Now we want to see how a positive interest on reserves alters the conclusions. The following figure plots the welfare gains from FB with $N = 6$ as a function of i_m for three different interest rates from above, $i = 5\%, 10\%$ and $i = 14.6\%$.

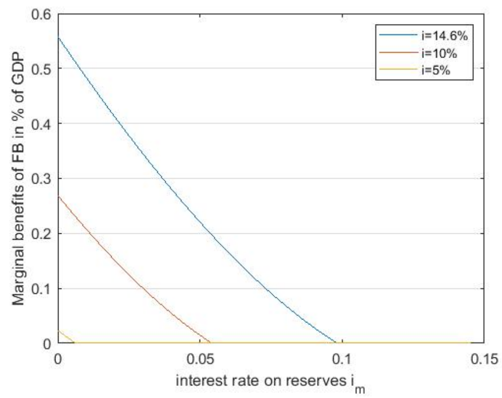


Figure 6: Welfare gains as a function of i_m

The figure shows how paying interest on reserves reduces the welfare gains from FB even further. For example with 10% inflation ($i = 14.6\%$) paying 5% interest on reserves reduces the welfare gains to 0.22% of GDP. Paying interest on reserves thus significantly reduces the costs of NB systems.

7 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.

References

Bank for International Settlements. Central bank digital currencies. Technical report, Committee on Payments and Market Infrastructures, 2018.

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

Aleksander Berentsen, Samuel Huber, and Alessandro Marchesiani. Financial Innovations, Money Demand, and the Welfare Cost of Inflation. *Journal of Money, Credit and Banking*, 47(S2):223–261, June 2015. doi: 10.1111/jmcb.12219.

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.

Jonathan Chiu, Mohammad Davoodalhosseini, Janet Hua Jiang, and Yu Zhu. Bank Market Power and Central Bank Digital Currency: Theory and Quantitative Assessment. Staff Working Papers 19-20, Bank of Canada, May 2019.

Ben Craig and Guillaume Rocheteau. Inflation and Welfare: A Search Approach. *Journal of Money, Credit and Banking*, 40(1):89–119, February 2008.

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

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

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

Jr. Lucas, Robert E. Glass-steagall: A requiem. *American Economic Review*, 103(3):43–47, May 2013.

Robert E. Lucas and Juan Pablo Nicolini. On the stability of money demand. *Journal of Monetary Economics*, 73(C):48–65, 2015.

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

Dirk Niepelt. Reserves for All? Central Bank Digital Currency, Deposits, and Their (Non)-Equivalence. *International Journal of Central Banking*, pages 211–238, 2020.

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.

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 (8) 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:

$$\frac{u'(q_b)}{c'(\frac{1-s}{s}q_b)} = \frac{\gamma/\beta - s}{1-s}. \tag{44}$$

We see the RHS of equation (44) is equal to the RHS of 45 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 45 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:

$$\frac{u'(q_b)}{c'(q_s)} = \frac{\gamma}{\beta} \tag{45}$$

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:

$$
\begin{aligned}
V_b(m) = \max_{q_b, d_b, m'} \quad & u(q_b) + W(d_b(1+i_d) + m' - pq_b) \\
s.t. \quad & pq_b \leq d_b(1+i_d) + m' \\
& d_b + m' = m
\end{aligned} \tag{46}
$$

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.

Appendix B

B.1 proof of proposition 1

Proof. To derive (28) 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 (25), (26) and (27) 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 (12) and (7) 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 (22) pins down a unique q_s even if $c(q_s)$ is not strictly convex (and (17) strictly concave). Thus the solution to (28) must be unique.

Comparative statics: Differentiate the left-hand side of (28), $\frac{u'(q_b)}{c'(q_s)}$, with respect to q_b

$$\frac{\partial(\frac{u'(q_b)}{c'(q_s)})}{\partial q_b} = \frac{u''(q_b)c'(q_s) - u'(q_b)c''(q_s)\frac{1-s}{s}}{c'(q_s)^2} < 0 \tag{47}$$

From the strict concavity of $u(q)$, $u''(q) < 0$ and therefore (47) 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 (28) $\frac{u'(q_b)}{c'(q_s)} = \frac{1+i}{1+i_d} > 1$ 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 (47). 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 (28) 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 (28) $\gamma/\beta\frac{1}{1+i_d}$ is strictly below the right-hand side of an economy without banks (44) $\frac{\gamma/\beta - s}{1-s}$. 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 28. At $\alpha = 1$ the right-hand sides of (45) in appendix A.2 and in (28) coincide.

■