Abstract

We study the role of banks in creating inside money. We show that, even in the absence of asymmetric information or an agency problem, the private provision of inside money is inefficient. The reason is that inside money affects prices and the welfare of non-bank customers, and banks do not internalize this. Under competition, banks tend to generate too much money relative to what is socially efficient, while a monopolistic bank chooses too little money. We argue that a regulator can eliminate these inefficiencies by promoting competition and imposing a ceiling on the loan-to-deposit ratio of each bank.
1. Introduction

Banking is probably the most regulated industry in the world (Barth et al. (2004)). One reason for this is to be found in political economy. Banks are where the money is. This feature attracts not only robbers like Willy Sutton, but also governments in need of revenues (Rajan and Zingales (2003)). More recently, the most popular explanation for bank regulation is that it is an antidote to the too-big-to-fail doctrine (Stern and Feldman (2004)). Time inconsistency problems force governments to rescue large banks in trouble. This creates a moral hazard problem that needs to be addressed by regulation (Farhi and Tirole (2009), Chari and Kehoe (2010)). Once again this is a political economy problem. Yet, is there an economic problem, a market failure, which justifies government intervention in banking?

Most economists would answer this question affirmatively by appealing to Diamond and Dybvig (1983)’s classic argument that demandable deposits expose banks to inefficient bank runs. However, if demandable deposits cause this inefficiency, why don’t banks choose different contracts (Allen and Gale (2007))? Furthermore, as Walter Bagehot already suggested in 1873, the risk of inefficient bank runs can be mitigated by the creation of a lender of last resort: in a crisis the lender of last resort lends at a penalty rate to solvent but illiquid banks that have adequate collateral. So why do we need capital requirements, liquidity requirements, restrictions on activities, limitations on bonuses, and all the other regulations that we observe in the banking sector, but not in the rest of the economy? Why do these regulations also apply to small banks that are clearly not too big to fail?

This paper tries to analyze these questions by focusing on a unique role played by banks, the creation of a means of payment. Historically, banks were the only institutions whose liabilities were treated like money in ordinary exchange. We study the efficiency property of the amount of money created by the banking sector (inside money). In particular, we analyze the general equilibrium effects that the availability of money has on prices and identify two pecuniary externalities in the creation of inside money by banks. More money increases the equilibrium price of the goods that those with the money buy; but it also increases the wealth of the agents supplying these goods and so the prices of the goods they buy. Short of a central planner, no economic agent is able to internalize all these externalities. A monopolistic bank, which does not internalize the surplus of the sellers, will generate too little money. A competitive bank, which ignores the externality imposed on other buyers, will generate too much. As a result,
we show that, even in the absence of asymmetric information or an agency problem, the private provision of inside money is generally inefficient.

We analyze these inefficiencies arising from the transactional role played by money in a basic economy in which there are two groups of agents: doctors and builders. The economy lasts two periods, and hence there is no (simple) role for fiat money. Doctors buy building services from builders and then builders buy doctors’ services from doctors. The lack of a double coincidence of wants (each builder requires a doctor different from the one he is building for) generates a need for money. Doctors have an endowment of wheat, but wheat is costly to carry and can easily rot. Banks arise in our model as depositary institutions, which store wheat and issue notes. Banks can issue an amount of notes inferior to the value of the wheat deposited (overcollateralized notes) or superior to that amount (banks extend loans). We study the determinants of this supply of notes and its effect on equilibrium prices and on social welfare.

To the extent that banks have market power, they partially internalize the effect that note creation has on equilibrium prices. Oligopolistic banks create too little inside money, to keep prices of services consumed by their customers (in our model doctors) low. In particular, oligopolistic banks do not internalize the positive externality that higher prices have on their non-customers (in our model builders). This effect is present regardless of the ownership structure of banks (whether they are mutually owned or owned by outsiders), although, as we shall see, it is stronger in profit-maximizing banks.

A bank also imposes a negative externality on doctors at other banks: raising the amount of inside money increases the price of building services, which is bad for them since they consume these services. For large banks, this second effect is small since there aren’t many other doctors, but for small banks it dominates the first effect. Thus, we find that, in a competitive market, we have the opposite problem: the amount of inside money created by the banking sector is excessive. Some wheat endowment is stored to create liquidity instead of being invested in profitable opportunities.

It follows that there is a particular level of competition that delivers the socially efficient quantity of inside money. This level, however, is contingent on the return on the alternative use of liquidity. When this return changes (as it is likely to do over the business cycle), the efficient level of competition changes as well.
Regulators can more easily restrict the amount of lending than force banks to lend against their will. Therefore, the above market inefficiency is best addressed by promoting competition and entry in the banking sector, while imposing some ceiling on the ratio between lending and deposits (what the New Zealand central bank calls the core funding ratio).

There exists a large literature on the role played by banks and the need for bank regulation (see, for example, Dewatripont and Tirole (1994)). One branch of this literature, starting with Diamond (1984) and continuing with Holmstrom and Tirole (1997), focuses on the asset side of banks: their role in monitoring loans. Another branch of this literature, starting with Diamond and Dybvig (1983), focuses on the liability side of banks: the ability of banks to provide risk sharing in the face of liquidity needs (Diamond and Dybvig (1983) and Allen and Gale (1998) and (2007)), or to reduce future adverse selection (Gorton and Pennacchi (1990)). Finally, there are some papers, such as Diamond and Rajan (2001), which try to integrate the two sides, showing how demand deposits are critical in making credible the ability of lenders to extract a repayment for their loans.

Holmstrom and Tirole (1998) and (2011) focus on moral hazard on the side of suppliers of liquidity, rather than on the side of users. They show that, in the presence of aggregate uncertainty, the state power to tax future income creates liquidity for the corporate sector, improving its ability to invest.

In all these papers the source of friction is either some informational asymmetry or some agency problem (or both). While these problems are important, they are not the only ones relevant for banks. One important reason why banks are unique is that they issue liabilities that are used as a means of payment. Our goal is to analyze the implications of this role. For this reason, we abstract from all the other frictions and focus on the pecuniary externalities in the creation of inside money. A general theory of banking would bring all these frictions together.

A new strand of the banking literature, which deals with the transactional role of deposits, introduces behavioral features. This strand derives the uniqueness of banks from the misperception by depositors that their claim are safe (Gennaioli et al. (2010), Rotemberg (2010)) or from the banks’ ability to arbitrage irrationally exuberant markets and rationally priced deposits (Shleifer and Vishny (2010)). We do not introduce behavioral aspects here.

Our result that the creation of inside money is excessive is similar to a conclusion of Stein (2011). In his model, however, it is assumed that agents have a discontinuous demand for a
riskless claim (money), while we do not make such an assumption. Similarly, his inefficiency arises from an assumed friction in the financial markets (that patient investors cannot raise additional money), while ours arises endogenously. Stein’s model, however, is much richer in terms of implications for monetary policy. The two models can be seen as complementary.

There is also a huge literature on money. Much of this literature is concerned with the role money plays in a general equilibrium model (e.g., Hahn (1965)). To create such a role, one needs to introduce explicitly an exchange process in the standard Arrow-Debreu model, dispensing with the traditional Walrasian auctioneer. Ostroy and Starr (1990) provide an excellent survey of attempts in this direction. As far as we can tell, none of these attempts analyze the pecuniary externality we identify in our paper. The role money plays in our model (i.e., addressing the lack of double coincidence of wants) is similar to Kyotaki and Wright (1989). Their focus, however, is on what goods can become money and how. Our focus is to what extent private banks can provide the efficient quantity of the medium of exchange.

Another large slice of this literature on money analyzes the role of inside money on monetary policy, as in Brunnermaier and Sannikov (2010), Diamond and Rajan (2006), and Kashyap and Stein (2004). Our model is silent on this.

Our approach resembles that of Kiyotaki and Moore (1997) and (2002). Like us, they rely on some limited pledgeability of future income. Their main focus, however, is on the multiplier/contagion effect that the failure of one intermediary can have on the overall system. Our paper, instead, is concerned with the pecuniary externality in the creation of inside money.

Our paper is perhaps closest to that of Mattesini et al. (2009). They study banking using the tools of mechanism design. They consider an economy with two groups of consumers who want to trade with each other. As in our model, there is a timing problem: the first group has to buy from the second group before the first group has sold its output. Mattesini et al. (2009) analyze how claims on deposits with third parties are a better means of exchange than claims on individual wealth. They study the social optimum but not the market equilibrium. In contrast, we take the superiority of third party deposits as given, and study the divergence between the market equilibrium and the social optimum. Another difference between the papers is that Mattesini et al. use an infinite horizon model while ours is a two period one.

The rest of the paper proceeds as follows. In Section 2 we lay out the framework and describe the Walrasian equilibrium. In Section 3 we analyze the effect of the introduction of
banks as storage facilities. In Section 4 we introduce the possibility of risky investments. In Section 5 we analyze some extensions and in Section 6 we discuss some possible form of regulation. Conclusions follow.

2. The Framework

We consider an economy that lasts two dates:

1----------------------------------------2

Agents consume at dates 1 and 2. There are two types of agents in equal numbers: doctors and builders. The doctors want to consume building services and the builders want to consume doctor services. At date 1 doctors have an endowment of wheat equal to \( e > 0 \). Builders have a zero endowment of wheat.

We write agents’ utilities as:

Doctors: \[ U_d = x_d + b_d - \frac{1}{2} l_d^2 \]

Builders: \[ U_b = x_b + d_b - \frac{1}{2} l_b^2 \]

where \( x_i \) is the quantity of wheat consumed by individual \( i = d, b \) at date 1 and/or date 2; \( b_d \) is the quantity of building services consumed by the doctors; \( l_d \) is the labor supplied by the doctors; \( d_b \) is the quantity of doctor services consumed by the builders; and \( l_b \) is the labor supplied by the builders. We assume constant returns to scale: one unit of builder labor yields one unit of building services and one unit of doctor labor yields one unit of doctor services. Let \( p_b \) and \( p_d \) be the price respectively of building and doctor services in terms of wheat.

In words, doctors and builders have a constant marginal utility of wheat, a constant utility of the service provided by the other group of agents, and a quadratic disutility of labor.

At date/period 2 the market meets twice. First, doctors buy from builders, and then builders buy from doctors. We assume there are many doctors and many builders, and so the prices for both services are determined competitively. It is crucial for our analysis that there is no double coincidence of wants: the builder a doctor buys from requires the doctor services of
another doctor: receiving the doctor services of this doctor would be useless (one can imagine that builders and doctors have symmetric but different skills).

2.1. A benchmark: the Walrasian equilibrium

In an ideal world the doctors could pledge to pay the builders out of income from supplying doctor services that they will earn later in period 2. This is the assumption made in classic Walrasian theory and it is easy to compute the Walrasian equilibrium.

The doctors solve the following maximization problem:

\[
\begin{align*}
\text{(2.1)} & \quad \text{Max } x_d + b_d - \frac{1}{2} l_d^2 \\
\text{S.T.} & \quad x_d + p_b b_d \leq p_d l_d + e.
\end{align*}
\]

The solution is

\[
\begin{align*}
(2.2) & \quad l_d = p_d \quad \text{if } p_b \geq 1 \\
& \quad = \frac{p_d}{p_b} \quad \text{if } p_b < 1 \\
& \quad b_d = 0 \quad \text{if } p_b > 1 \\
& \quad b_d = \frac{p_b^2}{p_b^2} + \frac{e}{p_b} \quad \text{if } p_b < 1 \\
& \quad 0 \leq b_d \leq \frac{p_d^2}{p_b^2} + \frac{e}{p_b} \quad \text{if } p_b = 1
\end{align*}
\]

The intuition is that, if \( p_b > 1 \), doctors prefer wheat to building services, if \( p_b = 1 \) they are indifferent, and, if \( p_b < 1 \), they prefer building services. The marginal utility of wealth for doctors is 1 if \( p_b \geq 1 \) and \( \frac{1}{p_b} \) if \( p_b < 1 \), and this affects their labor supply decision. If \( p_b \geq 1 \), doctors choose labor supply to maximize \( p_d l_d - \frac{1}{2} l_d^2 \); and, if \( p_b < 1 \), they maximize \( \frac{p_d}{p_b} l_d - \frac{1}{2} l_d^2 \).

Similarly, the builders solve:

\[
\begin{align*}
\text{(2.3)} & \quad \text{Max } x_b + d_b - \frac{1}{2} l_b^2
\end{align*}
\]
S.T. \[ x_b + p_d d_b \leq p_b l_b. \]

The solution is

\[
(2.4) \quad l_b = \begin{cases} p_b & \text{if } p_d \geq 1 \\ \frac{p_b}{p_d} & \text{if } p_d < 1 \end{cases}
\]

\[
\begin{align*}
d_b &= 0 & \text{if } p_d > 1 \\
d_b &= \frac{p_b^2}{p_d^2} & \text{if } p_d < 1 \\
0 \leq d_b \leq \frac{p_b^2}{p_d^2} & \text{if } p_d = 1
\end{align*}
\]

Again the marginal utility of wealth for builders is 1 if \( p_d \geq 1 \) and \( \frac{1}{p_d} \) if \( p_d < 1 \).

For markets to clear we must have

\[
(2.5) \quad b_d = l_b
\]

\[
(2.6) \quad d_b = l_d
\]

On the one hand, (2.5) and (2.6) cannot be satisfied if either \( p_b > 1 \) or \( p_d > 1 \) (demand will be less than supply for building/doctor services, respectively). On the other hand, we cannot have both \( p_b < 1 \) and \( p_d < 1 \) because then the demand for wheat would be zero, while the supply is \( e \) ((2.5) and (2.6) imply that the wheat market clears). Hence, either \( p_b < 1 \) and \( p_d = 1 \), or \( p_b = 1 \), \( p_d < 1 \), or \( p_b = p_d = 1 \). It is easily seen that the first case is inconsistent with (2.5) and the second with (2.6).

We are left with \( p_b = p_d = 1 \). It is immediate that (2.2)-(2.6) hold if \( b_d = l_b = d_b = l_d = 1 \).

Hence, we have established
Proposition 1: The unique Walrasian equilibrium satisfies \( p_b = p_d = b_d = l_b = d_b = l_d = 1 \). The utilities of the doctors and builders are \( U_d = e + \frac{1}{2} \), \( U_b = \frac{1}{2} \), respectively, and total welfare (social surplus) equals \( W \equiv U_d + U_b = e + 1 \).

Note that the Walrasian allocation and prices are independent of the initial endowment \( e \) (except for the doctors’ consumption of wheat, which varies one to one with \( e \)). Also the Walrasian equilibrium achieves maximal social surplus, which is no surprise given the first theorem of welfare economics and the symmetry of the parties.

2.2. A benchmark: equilibrium when doctors cannot borrow but can pay builders directly with wheat

Suppose now that the doctors cannot pledge their future income from doctor services to pay for building services. One can imagine that the doctors can hide this income in such a way that the builders cannot get hold of it. However, the doctors can carry their wheat around with them and can pay builders directly with it. (Spot trades in contrast to borrowing contracts are enforceable.) In turn, the builders can use the wheat they receive from doctors to pay for doctor services.

Note first that if \( e \geq 1 \), then the above Walrasian equilibrium can be sustained. The doctors need only 1 to pay for the building services they want and their endowment is enough to cover this. The builders also get sufficient wheat from the doctors to purchase doctor services. From now on we assume \( e < 1 \), although much of our analysis of banking in Sections 3-5 is also relevant for the case \( e \geq 1 \).

We solve for the equilibrium backwards. After the doctors have bought building services, the builders find themselves with a quantity \( e' \leq e \) of wheat. Hence, their demand for doctor services will be given by \( \frac{e'}{p_d} \) if \( p_d \leq 1 \) and 0 if \( p_d > 1 \). The supply for doctor services is given by

\[
 l_d = p_d ,
\]

as in (2.2), since any income received by doctors is spent on wheat (it is too late to consume more building services). Clearly \( p_d > 1 \) cannot clear the doctor services market since the demand
for doctor services would be zero and the supply strictly positive. Hence, market clearing requires

\[(2.8) \quad \frac{e'}{p_d} = p_d, \]

that is,

\[(2.9) \quad p_d = (e')^{\frac{1}{2}}.\]

Since \( e' \leq e < 1 \), we see from (2.9) that \( p_d < 1 \).

Now consider the first round of trading. The demand for building services is given by

\[(2.10) \quad b_d = 0 \quad \text{if} \quad p_b > 1 \]
\[b_d = \frac{e}{p_b} \quad \text{if} \quad p_b < 1 \]
\[0 \leq b_d \leq \frac{e}{p_b} \quad \text{if} \quad p_b = 1 \]

The reason is that, if \( p_b > 1 \), doctors prefer wheat to building services; if \( p_b < 1 \), they prefer building services to wheat; and, if \( p_b = 1 \), they are indifferent. Since \( p_d < 1 \) the supply of building services is

\[(2.11) \quad l_b = \frac{p_b}{p_d}. \]

It is easy to see that the building services market cannot clear if \( p_b \geq 1 \) (if \( p_b > 1 \), demand is zero, and supply is positive, while if \( p_b = 1 \), demand is less than 1 and supply exceeds 1). Hence, market clearing \((b_d = l_b)\) requires

\[(2.12) \quad \frac{e}{p_b} = \frac{p_b}{p_d}. \]

Since \( p_b < 1 \), doctors spend all their endowment on building services. Hence, \( e' = e \).

Combining this with (2.9) and (2.12) yields
Also, from (2.7), (2.11),

\begin{equation}
  l_d = d_b = e^{\frac{1}{2}}, \quad l_b = b_d = e^{\frac{1}{4}},
\end{equation}

and from the budget constraints

\begin{equation}
  x_d = e \quad \text{and} \quad x_b = 0.
\end{equation}

It follows that

\begin{equation}
  U_d = \frac{1}{2}e + e^4, \quad U_b = \frac{1}{2}e^{\frac{1}{2}},
\end{equation}

and total welfare is

\begin{equation}
  W = \frac{1}{2}e + \frac{1}{2}e^2 + e^4,
\end{equation}

which is strictly less than what is obtained in the Walrasian equilibrium (e+1).

**Proposition 2.** Suppose e<1. If doctors cannot pledge future income but can pay builders directly with wheat, the equilibrium is

\begin{align*}
  p_d &= e^{\frac{1}{3}}, \quad p_b = e^{\frac{1}{4}}, \\
  l_d &= d_b = e^{\frac{1}{3}}, \quad l_b = b_d = e^{\frac{1}{4}}, \\
  U_d &= \frac{1}{2}e + e^4, \\
  U_b &= \frac{1}{2}e^{\frac{1}{2}}, \quad W = \frac{1}{2}e + \frac{1}{2}e^2 + e^4. \quad \text{Welfare is strictly below the first-best level.}
\end{align*}

It is interesting to observe that the small endowment of wheat (e<1), in combination with the inability of the doctors to pledge future income, has ambiguous effects on the doctors’ utility, but always hurts the builders, relative to the Walrasian outcome. The doctors’ utility is

\begin{equation}
  U_d = \frac{1}{2}e + e^4,
\end{equation}

which will be below e +1/2 for small e, but above it for large e ; while the utility of builders is

\begin{equation}
  U_b = \frac{1}{2}e^{\frac{1}{2}}, \quad W = \frac{1}{2}e + \frac{1}{2}e^2 + e^4. \quad \text{Welfare is strictly below the first-best level.}
\end{equation}

However, the doctors’ restricted trading imposes a negative externality on the builders, and they are always worse off. As we shall see, these effects will be important in what follows.
3. Introducing Banks

We now suppose that it is impossible for the doctors to carry wheat with them when they trade with builders; it is too cumbersome or the wheat would rot or be stolen. In the absence of any further assumptions the model now becomes trivial. The doctors would eat their wheat and no trade between doctors and builders would occur. We would have $U_d = e$ and $U_b = 0$.

However, we now introduce storage facilities. These storage facilities are perfectly secure in the sense that wheat deposited at a facility at date 1 will remain there and be intact at the end of period 2. To simplify matters, we will take the number of storage facilities to be given (see below). However, in a more general model one could suppose that there is a fixed cost of setting up a storage facility and that the number of storage facilities is endogenous.

Storage per se does not change anything since there is no advantage to doctors from storing wheat rather than consuming it right away. However, let us suppose that claims can be issued on the wheat deposited in a storage facility. In particular, if a doctor deposits $f$ units of wheat he will receive $g \leq f$ notes, where each note is a claim on a unit of wheat at the end of period 2; he can then use these notes to pay builders. (It will be important that $g$ can be less than $f$, i.e., not all the wheat deposited yields notes.) The builders in turn can use these notes to pay doctors. At the end of period 2 the holders of the $g$ notes can go to the storage facility and redeem them for $g$ units of wheat. Putting aside any fees, the doctor who initially deposited $f$ would then get back his deposit minus the notes he received: $f-g$.

We call these storage facilities banks.

The idea that banks issue notes in an amount below deposits may seem counter-intuitive. We offer some more appealing interpretations in Section 4 when we introduce risky investments.

Let us assume that the mass of doctors equals 1 (which is also the mass of builders) and that there is a fixed number of banks, $\frac{1}{\alpha}$, where $0 < \alpha < 1$. Each bank serves a fraction $\alpha$ of the doctors. For simplicity we assume that each bank is a monopolist with respect to its constituency of doctors; however, we doubt that much would change if we allowed several banks to compete for the same constituency of doctors.
Note that the case $\alpha = 0$ can be interpreted as the (limiting) situation where every doctor can set up his own bank.

We will distinguish between mutual (or cooperative) banks and banks with outside owners. We start with the former.

**Mutual Banks**

We suppose that each bank offers the doctors in its constituency the following service: a doctor can deposit any amount of wheat $f$ and receive notes (or checks) equal to $\sigma f$, where $\sigma$ is the ratio between the amount of notes and the amount of deposits (all measured in terms of units of wheat). Initially, we consider only fully-backed deposits, thus $0 \leq \sigma \leq 1$. Hence $\sigma$ is a policy instrument of the bank, which is the same for all customers; moreover, we assume that the bank can announce and commit to it.\(^1\)

Consider a doctor who initially deposits $f$ at his bank and receives $\sigma f$ in notes. Suppose that this doctor uses all these notes for the purchase of building services in the first half of period 2, but then in the second half of period 2 acquires $f'$ in notes from builders who purchase his doctoring services. Typically these $f'$ notes will be liabilities of other banks. Thus at the end of period 2 the doctor will go to his own bank and ask for (and receive) $(1-\sigma)f$ units of wheat; and will go to other banks and ask for (and receive) $f'$ units of wheat. Similarly, customers of other banks who have acquired the $\sigma f'$ notes of the original doctor will go to his bank and have them redeemed for wheat. (Of course, if doctor visits to other banks are costly it would be more efficient for banks to settle some transactions among themselves. We will assume that doctor visits are costless and so will not deal with this issue in the current paper.)

Consider a single bank’s choice of $\sigma$ given that the bank serves a fraction $\alpha$ of the population of doctors and that the average choice of other banks is $\bar{\sigma}$. We know from Section 2.2 that even if every doctor deposits all his endowment $e$ and $\sigma = 1$ for all banks, then $p_b < 1$ and $p_d < 1$. A fortiori this must be true when $f \leq e$ and $\sigma \leq 1$. Thus we can focus on situations where $p_b < 1$ and $p_d < 1$.

\(^1\) One can imagine more complicated arrangements between banks and customers, e.g., banks could put limits on how much each customer can deposit. For simplicity, however, we restrict attention to the above contracts.
Suppose that an individual doctor deposits $0 \leq f \leq e$ in his bank. At date 1, the doctor consumes $e - f$ units of wheat. He also receives $\sigma f$ in notes from his bank, which he uses to purchase building services in the first half of period 2. In the second half of period 2, he will choose $l_d$ to maximize $p_d l_d - \frac{1}{2} l_d^2$, i.e., set $l_d = p_d$. This yields revenue $p_d^2$ in the form of notes, which he redeems for wheat at the end of period 2 (it is too late to buy more building services); in addition he incurs an effort cost of $\frac{1}{2} p_d^2$. Finally, he will also receive $(1 - \sigma)f$ units of wheat from his own bank at the end of period 2. The doctor’s utility is therefore

\[(3.1)\]

\[e - f + \frac{\sigma f}{p_b} + \frac{1}{2} p_d^2 + (1 - \sigma)f\]

which he maximizes with respect to $0 \leq f \leq e$. It is easy to see that the maximum is achieved at $f = e$ (the wheat deposited in the bank yields a higher return than wheat consumed given that $p_b < 1$), and so maximized utility is

\[(3.2)\]

\[\frac{\sigma e}{p_b} + \frac{1}{2} p_d^2 + (1 - \sigma)e.\]

The mutual bank chooses $\sigma$ to maximize the utility of a representative member, given by (3.2), taking into account the effect of the bank’s choice of $\sigma$ on prices, $p_b$ and $p_d$.

Let us consider this price effect. Given $\sigma$ the total value of notes in circulation will be $\alpha \sigma e + (1 - \alpha) \bar{\sigma} e$; the first term represents the contribution of this bank and the second term the contribution of the other banks. Since doctors use all their notes on building services, the demand for building services is

\[(3.3)\]

\[\frac{\alpha \sigma e + (1 - \alpha) \bar{\sigma} e}{p_b},\]
while the supply is, as in (2.4), \( \frac{P_b}{P_d} \). (Builders will spend all the proceeds from their building services on doctor services.) Equating these yields

\[ (3.4) \quad p_b^2 = (\alpha \sigma e + (1 - \alpha) \tilde{\sigma} e) p_d. \]

In the market for doctors, demand is

\[ (3.5) \quad \frac{\alpha \sigma e + (1 - \alpha) \tilde{\sigma} e}{p_d}, \]

since the builders use all the notes received from doctors to buy doctor services; and supply is \( p_d \). Combining this with (3.4) yields

\[ (3.6) \quad p_b = (\alpha \sigma e + (1 - \alpha) \tilde{\sigma} e)^{\frac{3}{4}}, \]

\[ p_d = (\alpha \sigma e + (1 - \alpha) \tilde{\sigma} e)^{\frac{1}{4}}. \]

Substituting into (3.2), we see that the utility of a representative doctor at the bank choosing \( \sigma \) is

\[ (3.7) \quad \frac{\sigma e}{(\alpha \sigma e + (1 - \alpha) \tilde{\sigma} e)^{\frac{3}{4}}} + \frac{1}{2} (\alpha \sigma e + (1 - \alpha) \tilde{\sigma} e) + (1 - \sigma) e \]

\[ = \frac{\sigma e^{\frac{1}{4}}}{(\alpha \sigma + (1 - \alpha) \tilde{\sigma})^{\frac{3}{4}}} - \sigma e(1 - \frac{1}{2} \alpha) + e + \frac{1}{2} (1 - \alpha) \tilde{\sigma} e. \]

We study a Nash equilibrium in which each bank chooses \( \sigma \) to maximize (3.7), taking \( \tilde{\sigma} \) as given. Let \( y = \alpha \sigma + (1 - \alpha) \tilde{\sigma} \) and \( z = (1 - \alpha) \tilde{\sigma} \). Then, maximizing (3.7) is equivalent to maximizing

\[ (3.8) \quad y^{\frac{1}{4}} - \frac{z}{y^{\frac{3}{4}}} - e^{\frac{3}{4}}(1 - \frac{1}{2} \alpha)(y - z) \]

with respect to \( y \). It is easy to see that (3.8) is strictly concave in \( y \). Thus, there is a unique maximizer \( y \) and hence a unique maximizer \( \sigma \) of (3.7), given \( \tilde{\sigma} \).
Moreover, the optimal $y$ is strictly increasing in $z$. It follows that, if two different banks choose different values of sigma in equilibrium, i.e., they face different values of $z$, then they will choose different values of $y$. But $y$ equals the average value of sigma over all banks, and must therefore be the same for each bank. It follows that the equilibrium sigma is the same for all banks, i.e., any Nash equilibrium is symmetric.

Differentiating (3.7) and setting $\sigma = \hat{\sigma}$, we may conclude that the equilibrium level of $\sigma$, if $0 < \sigma < 1$, satisfies

\begin{equation}
\frac{1}{4} \alpha + \frac{1}{4} (1 - \alpha) + \frac{3}{4} (1 - \alpha) = e^{\frac{3}{4}(1 - \frac{1}{2} \alpha) \sigma^{\frac{3}{4}}},
\end{equation}

which we can rewrite as

\begin{equation}
\sigma = \left(\frac{1}{4}\right)^{\frac{1}{3}} \frac{1}{1 - \frac{1}{2} \alpha} e^{\frac{3}{4}(1 - \frac{1}{2} \alpha) \sigma^{\frac{3}{4}}}.
\end{equation}

Finally, recognizing that for small $\alpha$ the solution of (3.10) may exceed 1, in which case we will be at a corner, we may write the overall equilibrium as

\begin{equation}
\sigma^* = \min \{\left(\frac{1}{4}\right)^{\frac{1}{3}} \frac{1}{1 - \frac{1}{2} \alpha} e^{\frac{3}{4}(1 - \frac{1}{2} \alpha) \sigma^{\frac{3}{4}}}, 1\}.
\end{equation}

It is easy to see from (3.11) that $\sigma^*$ is decreasing in $\alpha$ and either $\sigma^* = 1$ for all $0 < \alpha < 1$ (if $e < \left(\frac{1}{2}\right)^{\frac{4}{3}}$) or there is a cut-off $\bar{\alpha}$, where

\begin{equation}
\frac{1 - \frac{3}{4} \bar{\alpha}^{\frac{4}{3}}}{1 - \frac{1}{2} \bar{\alpha}} = e,
\end{equation}

such that

\begin{equation}
\sigma^* = 1 \text{ for } \alpha \leq \bar{\alpha} \text{ and } \sigma^* < 1 \text{ for } \alpha > \bar{\alpha}.
\end{equation}

We have established
Proposition 3. In the market equilibrium, if \( e < \left(\frac{1}{2}\right)^{\frac{4}{3}} \), each bank chooses \( \sigma^* = 1 \), regardless of \( \alpha \). If \( e \geq \left(\frac{1}{2}\right)^{\frac{4}{3}} \), each bank chooses \( \sigma^* = 1 \) if \( \alpha \leq \bar{\alpha} \), where \( \bar{\alpha} \) is given by (3.12), and \( \sigma^* < 1 \) if \( \alpha > \bar{\alpha} \).

In other words, small banks choose \( \sigma^* = 1 \), whereas large banks may choose \( \sigma^* < 1 \).

Let us turn now to the socially optimal choice of \( \sigma \). The builders choose \( l_b = \frac{p_b}{p_d} \), receive revenue \( \frac{1}{2} \left( \frac{p_b}{p_d} \right)^2 \) (which they use to consume doctor services) and incur an effort cost of \( \frac{1}{2} \left( \frac{p_b}{p_d} \right)^2 \). Thus, their utility is

\[
U_b = \frac{1}{2} \left( \frac{p_b}{p_d} \right)^2 = \frac{1}{2} \sigma^2 e^2
\]

from (3.6) with \( \bar{\sigma} = \sigma \).

(3.2) and (3.6) tell us that the utility of a doctor is

\[
U_d = \sigma^2 e^2 + \frac{1}{2} \sigma e + (1 - \sigma) e
\]

Thus, total welfare is

\[
W = U_d + U_b = \sigma^2 e^2 + \frac{1}{2} \sigma e + \frac{1}{2} \sigma^2 e^2 + e
\]

It is easy to show that \( \frac{dW}{d\sigma} > 0 \) for \( 0 \leq \sigma \leq 1 \), and so \( W \) is maximized at \( \sigma = 1 \).

We have established:

Proposition 4. In the market equilibrium banks choose too low a level of \( \sigma \) in terms of social efficiency. The problem is worse for large banks (\( \alpha \) close to 1) and disappears for small banks (\( \alpha \) close to 0).
The intuition is as follows. Large mutual banks restrict $\sigma$, i.e., issue too few notes, to lower the price of building services; this helps their members since their members consume these services. In doing this, however, large banks ignore the positive externality they impose on builders, who gain from high prices since this allows them to buy more doctor services. Small banks choose a high $\sigma$ because their impact on prices is limited. In this specification of the model, the level chosen by small banks is socially optimal, but in Section 4 we shall show this is not always the case.

Note that the social planner achieves the same outcome as in Section 2.2, where doctors could pay builders directly with wheat.

At this point a question arises: What would happen if each bank served a fraction $\alpha$ of doctors and builders, rather than exclusively doctors? One problem is that the two constituencies have different preferences and so it is not clear how a mutual bank would behave. Suppose for simplicity that the bank maximizes the sum of the utilities of its customers. Then it is easy to show that $\sigma = 1$ in the market equilibrium for all $\alpha$. In other words the inefficiency that we have identified in this section disappears. However, we will see that this is not the case anymore in Section 4 when we introduce a risky investment.

**Outside Owned Banks**

Suppose now that banks are owned by outsiders rather than the doctors themselves. Each bank approaches individual doctors in its constituency with the following proposal: “You pay a fee $F$ and in return you can decide how much to store with us, $0 \leq f \leq e$, and we will give you notes equal to $\sigma f$ and at the end of date 2 you can withdraw $(1 - \sigma)f$ units of wheat.” Here we again assume that each bank announces and commits to a particular level of $\sigma$, $0 \leq \sigma \leq 1$.

Each doctor always has the option to turn down the bank’s offer. If he does so, he will consume his wheat $e$, and consume no building services. However, he can supply his labor in the market for doctors in the second half of period 2, receiving $\frac{1}{2} p_d^2$ in net terms. Thus, his utility is

\begin{equation}
\text{(3.17)} \quad e + \frac{1}{2} p_d^2.
\end{equation}

In contrast, if the doctor accepts the bank’s offer, his utility will be
(3.18) \[ e = f + \sigma \frac{f}{p_b} + \frac{1}{2} p_d^2 + (1-\sigma)f - F. \]

The doctor will choose \(0 \leq f \leq e\) to maximize (3.18), yielding \(f = e\), and so maximized utility is

(3.19) \[ \sigma \frac{e}{p_b} + \frac{1}{2} p_d^2 + (1-\sigma)e - F. \]

A bank’s profits are given by \(F\) and hence a profit maximizing bank will choose \(F\) so that (3.17) and (3.19) are equal. If an individual bank chooses \(\sigma\), while the average choice of other banks is \(\bar{\sigma}\), \(p_b\) and \(p_d\) are given by (3.6), and so

(3.20) \[ F = \sigma \frac{e}{p_b} + \frac{1}{2} p_d^2 + (1-\sigma)e - e - \frac{1}{2} p_d^2. \]

As before, it is easy to show that the Nash equilibrium is unique and symmetric. Setting the derivative of (3.20) with respect to \(\sigma\) equal to zero, substituting \(\sigma = \bar{\sigma}\), and taking account of corner solutions, yields

(3.21) \[ \sigma^* = \text{Min}\{(1-\frac{3}{4} \alpha)^{-1} \frac{1}{e}, 1\}. \]

Comparing (3.21) and (3.11), we see that a profit-maximizing bank chooses a lower level of \(\sigma\) than does a similarly competitive mutual bank. The intuition is that the profit-maximizing bank ignores the fact that a higher \(\sigma\) increases \(p_d\), which increases the income doctors obtain from supplying doctor services. The point is that this increase is also enjoyed by doctors who do not contract with the bank and so the bank cannot charge for it.

Note, however, that in this model competitive banks still get things right: if \(\alpha\) is small the solution to (3.21) is \(\sigma^* = 1\).

4. Risky Investments

4.1 Only banks have access to the risky technology
So far we have ignored any possible use of the deposited wheat. Suppose now that each bank can take some of the funds deposited by the doctors and allocate them to an investment with the following payoff per dollar (constant return to scale):

\[
R = \frac{(1+r)}{\varepsilon} \quad \text{with probability } \varepsilon
\]
\[= 0 \quad \text{with probability } 1-\varepsilon
\]

The expected return of the investment is \((1+r)\), which exceeds 1 (we assume \(r>0\)).

We will focus on the limiting case where \(\varepsilon \to 0\) and the investment becomes infinitely risky. In this section we also assume that all uncertainty about the investment is resolved at the end of period 2, after all trade has occurred.

Interpret \(\sigma\) now to be the fraction of deposited funds invested by the bank in the safe asset (perhaps literally put in the safe) and \(1-\sigma\) to be the fraction invested in the risky asset. The bank now offers the following deal to each depositor: deposit \(f\) and we will issue \(f\) notes (NOT \(\sigma f\)), where each note is a promise to pay a unit of wheat at the end of period 2. These notes, which are unsecured debt, are fully backed by the bank’s assets.

What is the value of a note issued by a bank? Each note-holder realizes that with probability \(\varepsilon\) -- if the risky investment pays off—the bank can redeem the note at par, while with probability \(1-\varepsilon\) its cash reserves will be \(\sigma\) (per dollar) and so it will redeem the note at \(\sigma\). Thus, as \(\varepsilon \to 0\), the expected value of the note is \(\sigma\). As in Section 3, \(\sigma\) can be interpreted as an indicator of the amount of inside money produced.

Suppose that each bank is a for-profit mutual; that is, each doctor, as well as obtaining banking services, receives an amount of any profit earned by his bank in proportion to his deposit.\(^2\) Assume also that each doctor is risk neutral. Then, a doctor’s utility, previously given by (3.1), becomes

\[
e - f + \frac{\sigma f}{p_b} + \frac{1}{2} p_a^2 + (1-\sigma) \frac{f}{\Phi} (1+r) \Phi
\]

where \(\Phi\) is the total deposit of doctors at this bank and the last term represents the bank’s expected profit.

Each doctor maximizes (4.2) subject to \(0 \leq f \leq e\). The solution is \(f = e\) and so a doctor’s utility is

\(^2\) A similar analysis can be carried out for banks with outside owners; we omit this.
The bank chooses $\sigma$ to maximize (4.3) subject to (3.6). As in Section 3, there is a unique symmetric Nash equilibrium, which can be computed as follows: set the derivative of (4.3) with respect to $\sigma$ equal to zero and substitute $\sigma = \hat{\sigma}$. Recognizing that we may have a corner solution, we obtain

$$
\sigma^* = \text{Min}\{(\frac{1 - \frac{3}{4} \alpha}{1 - \frac{1}{2} \alpha + r})^4 \frac{1}{e}, 1\}
$$

Clearly, the presence of $r$ in (4.4) lowers $\sigma$ relative to (3.10).

Now consider the socially optimal choice of $\sigma$. The utility of a doctor is

$$
U_d = \sigma^2 e^4 + \frac{1}{2} \sigma e + (1 - \sigma) e(1 + r),
$$

while the utility of a builder is

$$
U_b = \frac{1}{2} \frac{(p_b)^2}{p_d} = \frac{1}{2} \frac{1}{\sigma^2} e^2.
$$

Total welfare is

$$
W = U_d + U_b = \sigma^2 e^4 + \frac{1}{2} \sigma^2 e^2 + e(1 + r) - \sigma e(\frac{1}{2} + r)
$$

$W$ is strictly concave in $\sigma$ and so there is a unique maximizer. For $r$ close to zero, $\frac{dW}{d\sigma} > 0$ for $0 \leq \sigma \leq 1$ and so $\sigma = 1$ is optimal (as in Section 3.1). However, for a large $r$, we have an interior solution, characterized by the first order condition:

$$
\frac{1}{\sigma^\frac{3}{2}} + \frac{e^\frac{1}{2}}{\sigma^\frac{1}{2}} = 4e^{\frac{3}{2}}(\frac{1}{2} + r).
$$

In Section 3 we saw that large (mutual) banks choose too low a value of $\sigma$ in market equilibrium. To see that this is still true, take the solution of (4.4), set $\alpha = 1$, and substitute in (4.8). It is easy to check that the left-hand side of (4.8) exceeds the right-hand side, and so the socially optimal $\sigma$ is higher.
However, we now get a new effect. Suppose $\alpha$ is small, e.g. $\alpha = 0$ and $e > (\frac{1}{1+r})^3$. Then substituting the solution of (4.4) into (4.8), we can easily check that the left-hand side of (4.8) is less than the right-hand side. The conclusion is that small (e.g., competitive) banks choose too high a value of $\sigma$.

The intuition is the following. As we noticed in Section 3, a bank imposes a positive externality on builders: raising $\sigma$ increases $p_b$, which is good for them since they supply building services. However, a bank also imposes a negative externality on doctors at other banks: raising $\sigma$ increases $p_b$, which is bad for them since they consume building services. For large banks, this second effect is small since there aren’t many other doctors, but for small banks it dominates the first effect.

One may ask, why isn’t this effect second order in a competitive economy? After all, normally competitive equilibria are Pareto optimal. The reason seems to be that doctors are wealth-constrained given that they cannot pledge their future income. Thus the increase in $p_b$ caused by a small increase in $\sigma$ has a first order effect on them.

When $r=0$, we did not see this second effect because the social optimum was at a corner ($\sigma=1$), and so small banks couldn’t push $\sigma$ any higher.

Two further important points need to be considered. First, how much does our new result on the over-production of money by competitive banks depend on the assumption that each bank acts only on behalf of doctors. The answer is that it does not. The reason is that the effect of a bank on the utility of a builder, given by $\frac{1}{2}(p_b)^2$ in (4.6), is second order when $\alpha$ is close to zero. Hence a bank with a (small) constituency of builders and doctors would make the same choice of $\sigma$ as a bank representing only doctors. The intuition for this is that builders are not wealth-constrained. They supply their labor services before purchasing doctor services and so do not suffer from the pledgeability problems that doctors face.

Second, the reader may wonder whether it is enough to show that total surplus is not maximized in the competitive equilibrium. Does it follow from this that a Pareto improvement can be achieved? The answer is yes, under an additional assumption. Suppose that, starting at the competitive equilibrium, the planner slightly lowers $\sigma$. The ratio of $p_b$ to $p_d$ falls, and so builders are worse off. One thing the planner can do to compensate builders is to impose a small
lump sum tax on doctors with a corresponding lump sum subsidy for builders. Since each
group’s utility is linear in wheat and the sum of utilities rises, the tax and subsidy can be chosen
to obtain a Pareto improvement. Of course, this raises the question: how does the planner ensure
that the doctors pay the tax? After all we have supposed that builders cannot collect payments
promised by doctors. To answer this one probably has to suppose that the planner has some
powers that ordinary market participants do not: she can put tax scofflaws in jail.

4.2 Doctors have access to the risky technology
So far we have assumed that only banks have access to the risky positive-NPV investment. Let’s
now consider the case where this opportunity is available to doctors as well. In the absence of
access to the risky investment, we have seen that doctors deposit all their endowment in the
banks. The only reason they might not do so now is that they find it more profitable to invest in
the positive-NPV investment. Assume that doctors invest in the risky technology any amount not
deposited in banks. Then, a doctor’s utility becomes

\[(e - f)(1 + r) + \frac{\sigma f}{p_b} + \frac{1}{2} p_a^2 + (1 - \sigma) \frac{f}{\Phi} (1 + r)\Phi\]

where (as in 4.3) \(\Phi\) is the total deposit of doctors at this bank, the last term represents the
bank’s expected profit, \(f\) is the amount deposited in the bank, and \(e - f\) is the amount invested
in the risky project. Differentiating (4.9) with respect to \(f\) we obtain

\[(4.10)\]

\[\sigma \left[ \frac{1}{p_b} - (1 + r) \right] \]

The sign of 4.10 depends upon the relative size of \(p_b\) and \(r\). However, \(p_b\) is endogenous and in
equilibrium is given by \(p_b = (\sigma e)^{\frac{3}{4}}\). From (4.4) we have that

\[(4.11)\]

\[\left( \sigma e \right)^{\frac{3}{4}} \leq \frac{1 - \frac{3}{4} \alpha}{1 - \frac{1}{2} \alpha + r}\]

or
and hence (4.10) is always positive. Therefore, the doctors will deposit all their endowment in the bank and the equilibrium is the same as in Section 4.1.

The intuition is simple. Investing in a bank is better than investing on one’s own account because it provides the additional benefit of using bank deposits to trade. It is true that banks may choose \( \sigma \) to restrict \( p_b \), but never to the point where (4.10) is negative and self-investment dominates.

5. Extensions

5.1 Trading in shares

The results obtained in Section 4 raise an important question: why does trading take place with fixed claims, rather than with shares in assets of the banks? One could imagine that each doctor deposits all his endowment in the bank in exchange for shares in the bank’s assets (in proportion to his fraction of the total funds deposited). These shares could then be used for trading. The doctors will endorse these shares to the builders in exchange for building services and the builders will further endorse them to doctors when they buy their services. In this world, where both doctors and builders are risk neutral, a choice \( \sigma = 0 \) would achieve the first best: it takes advantage of the higher return of the risky asset, and does not interfere at all with trading. Thus we are faced with the question: do mutual funds solve all problems?

The answer is affirmative only if we stick to the current timing of the resolution of uncertainty. If we assume that uncertainty is resolved earlier, then this result no longer holds.

Consider the other opposite extreme, where uncertainty is resolved after deposits are made, but before trading takes place. Suppose that the risky investments are perfectly correlated, i.e., there are two states of the world; in the good state all risky investments yield \( R = \frac{(1+r)}{\varepsilon} \) and in the bad state they all yield zero. Let \( \sigma \) be the investment in the riskless asset, and \( 1-\sigma \) in the risky one, of a typical bank.
With this new timing, equilibrium prices will differ according to the state. Denote prices in the bad state by $p_b, p_d$ and in the good state by $p_b, p_d$. Let each doctor deposit an amount $f$ of his endowment in a bank. Then, with probability $1 - \varepsilon$ the risky investment will yield zero and the value of the doctor’s claim will be $\sigma f$. By contrast, with probability $\varepsilon$ the risky investment will yield $R = \frac{(1+r)}{\varepsilon}$ and the value of the doctor’s claim will be

\begin{equation}
(5.1) \\
ge = \sigma f + (1 - \sigma) f \frac{1+r}{\varepsilon}.
\end{equation}

A doctor’s expected utility is

\begin{equation}
(5.2) \\
(1 - \varepsilon)[e - f + \frac{\sigma f}{p_b} + \frac{1}{2} p_d^2] + \varepsilon[e - f + \frac{g}{p_b} + \frac{1}{2} p_d^2].
\end{equation}

It is easy to show that this is increasing in $f$ (since $p_b \leq 1$) and so each doctor will set $f = e$. Hence, (5.1) can be rewritten as

\begin{equation}
(5.3) \\
ge = \sigma e + (1 - \sigma) e \frac{1+r}{\varepsilon},
\end{equation}

and (5.2) becomes

\begin{equation}
(5.4) \\
(1 - \varepsilon) \left[ \frac{\sigma e}{p_b} + \frac{1}{2} p_d^2 \right] + \varepsilon \left[ \frac{g}{p_b} + \frac{1}{2} p_d^2 \right]
\end{equation}

From the analysis of Sections 3 and 4,

\begin{align*}
p_b &= (\sigma e)^{\frac{3}{2}}, p_d = (\sigma e)^{\frac{1}{2}}, \\
p_b &= \text{Min} \left[ g^{\frac{3}{2}}, 1 \right], p_d = \text{Min} [g^{\frac{1}{2}}, 1].
\end{align*}

If we keep $\sigma < 1$ constant and let $\varepsilon \to 0$, $p_b$ converges to one and (5.4) becomes

\begin{equation}
\frac{\sigma e}{p_b} + \frac{1}{2} p_d^2 + (1 - \sigma)e(1 + r),
\end{equation}
which is identical to (4.3). Hence, all the solutions will be the same as in Section 4. In particular, for \( \varepsilon \to 0 \) the welfare function will become

\[
W = U_d + U_b = \frac{1}{2} \sigma^2 e^3 + \frac{1}{2} \sigma^2 e^2 + \frac{1}{2} \sigma e + (1 - \sigma) e (1 + r)
\]

which is identical to (4.7). Thus we have

**Proposition 4.** When uncertainty is resolved before trading takes place, the model of this section with trading in bank shares is equivalent to the model of Section 4 with trading in fixed claims on deposits.

### 5.2 Borrowing

Let's now consider the case where banks have some ability to prevent people from consuming future income and thus are able to lend against future income. One possibility is that all payments take place through check transfers and that the bank is able to seize salaries before these are cashed for consumption. In other words, the bank can ensure that someone who defaults has zero consumption.

We consider borrowing in the context of the model in Section 3, i.e., without uncertainty. A bank that lends is creating more inside money, which is equivalent to increasing \( \sigma \). Thus we would expect a bank to want to do this only if it is already setting \( \sigma = 1 \). Thus, suppose \( \sigma = 1 \) for every bank and let \( b \) be the amount that the bank lends to each customer. We assume that \( b \) is verifiable and that the bank can announce and commit to \( b \). By borrowing an amount \( b \) an individual doctor can consume \( \frac{b}{p_b} \) of building services, which is attractive if he has to pay back only \( b \) (given \( p_b < 1 \)). Of course, a bank needs to make sure that it will be repaid. In period 2, a doctor exerts \( \frac{1}{2} p_d^2 \) of effort, receiving in exchange \( p_d^2 \) in terms of payment. Thus, he cannot borrow more than \( \frac{1}{2} p_d^2 \); if he did he would prefer not to work in the second half of period two, default, and consume nothing.

A doctor's utility is now given by
The mutual bank chooses $b$ to maximize the utility of a representative member, given by (5.6), taking into account the effect of the bank’s choice of $b$ on prices $p_b$ and $p_d$, and subject to the constraint that a doctor does not default:

(5.7) \[ 0 \leq b \leq \frac{1}{2} p_d^2. \]

Consider the price effect. The total value of notes in circulation will be $e + ab + (1 - \alpha)\tilde{b}$, where $\tilde{b}$ is the average amount of the other banks’ lending. Since doctors use all their notes to buy building services, and builders use all their revenue from building services to buy doctor services, equilibrium in the builder and doctor markets yields

(5.8) \[ \frac{e + ab + (1 - \alpha)\tilde{b}}{p_b} = \frac{p_b}{p_d}, \]

(5.9) \[ \frac{e + ab + (1 - \alpha)\tilde{b}}{p_d} = p_d. \]

Hence

(5.10) \[ p_b = (e + ab + (1 - \alpha)\tilde{b})^{\frac{3}{4}}, \]

\[ p_d = (e + ab + (1 - \alpha)\tilde{b})^{\frac{1}{2}}. \]

(We will check below that $p_b \leq 1, p_d \leq 1$.)

In Nash equilibrium each bank maximizes (5.6) subject to (5.7) and (5.10). The first order condition for an interior solution, with $\tilde{b} = b$, yields
Note that $b + e \leq 1$, and so $p_b \leq 1$, $p_d \leq 1$ in (5.10).

We know that $b \geq 0$ and also that (5.7) and (5.10) imply

$$b \leq \frac{1}{2} (e + b),$$

that is,

$$b \leq e.$$  

We may conclude that the symmetric Nash equilibrium is characterized by

$$b = \text{Min}\{\text{Max} \left[ 0, \left( \frac{1 - \frac{3}{4}a}{1 - \frac{a}{2}} \right)^{\frac{4}{3}} - e \right], e \}.$$  

Comparing (3.11) and (5.14), we see that, as expected, $b > 0$ only if $\sigma = 1$ in the Nash equilibrium without borrowing.

Total welfare is given by

$$W = U_d + U_b = \frac{e + b}{p_b} + \frac{1}{2} p_d^2 - b + \frac{1}{2} \left( \frac{p_b}{p_d} \right)^2 = (e + b)^{\frac{3}{2}} + \frac{1}{2} (e + b) - b + \frac{1}{2} (e + b)^{\frac{3}{2}}.$$  

A social planner would maximize this subject to (5.13). It is easy to check that (5.15) is increasing in $b$ for $0 \leq b \leq 1 - e$ and so the social optimum is

$$b = \text{Min}\{e, 1 - e\}.$$
Clearly the planner chooses a higher value of $b$ than the market. But note that this is because there is no risky investment in this section. If we included a risky investment, as in Section 4, we would find that competitive banks ($\alpha$ close to zero) choose too high a level of $b$.

The conclusion is that the introduction of borrowing in the model of Sections 3 or 4 does not change our analysis significantly. The quantity of notes outstanding will still generally be socially inefficient.

It is important to emphasize that we have only scratched the surface of borrowing. If we introduce uncertainty, as in Section 5.1, then borrowing will be risky. (In principle it could be contingent.) Some loans may not be repaid, which might cause some banks not to be able to honor their claims. This may lead to contagion effects, as consumers cannot redeem claims and in turn default, leading other banks to default. (Contagion effects are analyzed in Kiyotaki and Moore (1997), (2002).) The analysis becomes much more complex, and richer, and we hope to explore the consequences in future work.

6. Regulation

We have shown that, even in the absence of asymmetric information or an agency problem, the private provision of inside money is generally inefficient. In a competitive equilibrium, banks generate too much money ($\sigma$ too high), while a monopolistic bank generates too little money ($\sigma$ too low), relative to what is socially efficient.

Since the optimal amount of inside money created is continuous in the level of competition, it follows that there is a particular level of competition that delivers the socially efficient quantity of inside money. This level, however, is contingent on the return on the alternative use of liquidity. When this return changes (as it is likely to do over the business cycle), the efficient level of competition changes as well.

An alternative is for the regulator to target $\sigma$ directly. Since $\sigma$ is both observable and verifiable, the regulator could impose the optimal $\sigma$ on all banks. This optimal level, however, also depends on the return on the risky investment. When the profitability of the risky investment is high, the socially optimal $\sigma$ corresponds to the solution of (4.8). By contrast if the
profitability of the risky investment is low, the optimal \( \sigma \) is \( \sigma = 1 \). Thus, from a practical point of view it is far from clear how to implement this solution.

Another possibility is for the regulator to have a total credit target, as some central banks do. Even assuming that the central bank is able to compute this target precisely, there is the problem of how to achieve it in practice, given that a central bank can restrict banks from lending, but cannot force them to lend if they do not want to. Given this asymmetry, the simplest solution is to ensure a very competitive environment (where we know that there is a tendency to produce too much credit) and then restrict credit through some aggregate and possibly bank-specific credit targets. Alternatively, in a competitive environment regulators can target the ratio of loans to deposits. In our model with lending this ratio will generally exceed the socially optimal level under competition. Thus, a regulator can easily impose a ceiling on this ratio and change this ceiling as a function of the business cycle and the alternative investment opportunities.

Of course, another way to tackle the shortage of inside money would be for the government to print its own money. If the government issues money, will we still need banks? In this finite horizon economy where bubbles are not possible, to be credible government money must be backed by future taxes, which will be distortionary. Banks can do better by creating inside money without any deadweight cost because banks’ liabilities (at least in the basic version without lending) are a claim on existing wealth rather than on future wealth.

In addition, the role for government-backed money exists only if banks are large since, under competition, banks provide too much inside money. This might be a reason why historically governments prefer to restrict competition in the banking sector, to ensure that there is room for government-backed money!

Before any attempt is made to implement our regulatory measures, it is important to remember that our analysis is incomplete, since it does not deal with the possibly systemic implications of bank failures. We plan to analyze this problem in future work. With this important caveat in mind, we now discuss how existing or proposed regulations can address the inefficiency identified in this paper.

One popular proposal (see Kotlikoff (2010)) is to restrict banks to hold only safe assets (in the context of our model, wheat). This proposal has two efficiency costs: First, it prevents
banks from lending to doctors, increasing the amount of liquidity and hence the level of economic activity and welfare. Second, it prevents banks from undertaking any positive NPV risky investments. In our model, if \( r \) is sufficiently high, and doctors do not have access to the same investment opportunities as banks, this restriction is socially inefficient. Narrow banking, however, could be an improvement if we start from a situation of a very concentrated banking market, since it would encourage banks to issue more notes. Yet, it is unclear why the right solution in this case would not be to reduce the concentration level.

Other proposals involve (increased) deposit insurance. It is not obvious how to introduce deposit insurance in the context of our model. First, we would need to establish a way for the government to tax to pay for the insurance. One possibility is a mill tax on consumption of wheat. Once this option is introduced, however, the government could find it valuable to issue riskless claims backed by future taxes regardless of deposit insurance. In fact, deposit insurance appears a suboptimal way to achieve the regulatory goals discussed in this paper, given the moral hazard it engenders. If the government can credibly commit to back all deposits in full, the amount of inside money will never be below the efficient level. Unfortunately, in the presence of this deposit guarantee, it is optimal for the banks, whether they are mutual or outside-owned, to choose \( \sigma = 0 \), so that taxpayers will have to back an amount equal to \( e \) with probability approaching one. However, it is impossible to finance an amount equal to \( e \) with a distortionary tax, while the total amount of wealth is equal to \( e \). Moreover, even if moral hazard was less extreme and a smaller amount of tax had to be raised, the deadweight cost would be very large.

7. Conclusions

We have built a simple framework to analyze the general equilibrium implications of the creation of inside money by banks. This model identifies the pecuniary externalities that arise in the creation of money. More money increases the equilibrium price of some goods (in our case building services), but it also increases the wealth of the agents supplying those goods (the

---

3 In the current version of the model lending to doctors is safe, while other investments are not. This does not need to be the case, though. If we introduced some idiosyncratic risk (doctors might die and be unable to pay back), the results would be the same as long as banks are sufficiently large to exploit the benefits of diversification.
builders) and so the price of the goods they buy (doctors’ services). Short of a central planner, no economic agent is able to internalize all these externalities. A monopolistic bank, which does not internalize the surplus of the builders, will generate too little money. A competitive bank, which ignores the externality imposed on other doctors, tends to produce too much.

Regulators can more easily restrict the amount of lending than force banks to lend against their will. Therefore, this inefficiency is best addressed by promoting competition and entry in the banking sector, while imposing some ceiling on the ratio between lending and deposits (what the New Zealand central bank calls the core funding ratio).

These policy recommendations, however, are tentative, since our current analysis does not look at possible contagion effects of bank failures. The framework we have developed, however, does allow for this analysis, which we plan to carry out in future work.
References


