INEFFICIENT PROVISION OF LIQUIDITY

OLIVER HART

HARVARD UNIVERSITY & NBER

AND

LUIGI ZINGALES*

UNIVERSITY OF CHICAGO, NBER & CEPR

JULY 2011

(First Version March 2011)

Abstract

We study an economy where the lack of a simultaneous double coincidence of wants creates the need for a relatively safe asset (money). We show that, even in the absence of asymmetric information or an agency problem, the private provision of liquidity is inefficient. The reason is that liquidity affects prices and the welfare of others, and creators do not internalize this. This distortion is present even if we introduce lending and government money. To eliminate the inefficiency the government must restrict the creation of liquidity by the private sector.

Key Words: liquidity, money, banking.
JEL Codes: E41 G21, E51.

* We would like to thank David Abrams, Eduardo Azevedo, Mathias Dewatripont, Douglas Diamond, Douglas Gale, Itay Goldstein, Christian Julliard, Jonathan Klick, Colin Mayer, Raghu Rajan, Andrei Shleifer, Jeremy Stein, Elu von Thadden, and participants at the Chicago, ILE/Wharton, LBS, and LSE finance seminars, the Oxford law and finance seminar, the NYU law and economic seminar, the University of Trento, and the NBER corporate finance meeting, for helpful comments. Oliver Hart gratefully acknowledges financial support from the U.S. National Science Foundation through the National Bureau of Economic Research. Luigi Zingales gratefully acknowledges financial support from the Center for Research in Security Prices (CRSP), the Stigler Center, and the Initiative on Global Markets at the University of Chicago.

1. INTRODUCTION

Historically, one of the primary functions of the banking sector has been to create means of payment ("money like" liabilities) either in the form of bank notes (private money) or credit. Yet most models of banking focus on other aspects, such as risk sharing or monitoring. Less attention has been paid to the role of banks in creating liquidity for transactions purposes and how this interacts with the government's role in this area. Does a competitive banking sector generate the socially optimal amount of the means of payment? How does the liquidity created by the banking sector affect the equilibrium level of goods and asset prices? Does the degree of competition in the banking sector affect the amount of liquidity created and hence the level of prices?

In this paper we try to answer these questions. We show that, even in the absence of asymmetric information or an agency problem, the amount of liquidity generated by the banking sector is inefficient. A competitive banking sector generates an excessive quantity of means of payment, while a monopolistic one a scarcity. The distortion arises from two externalities the creation of money has in general equilibrium: more money increases the equilibrium prices 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 will buy. These externalities are pecuniary in nature but unlike standard pecuniary externalities they have welfare consequences.

A competitive bank, which ignores the externality imposed on other buyers, will generate too much liquidity. A monopolistic bank that serves the buyers, and ignores the welfare of the sellers, generates too little liquidity. There is one intermediate level of competition that generates the efficient amount of liquidity but it would be fortuitous for the economy to end up there.

We start by assuming that future income (be it labor income or investment income) is not pledgeable (or contractible). Eventually, however, we relax this assumption, requiring only that human capital is inalienable and a worker can quit at any time.

In such a world we are able to show that transactional needs create a demand for relatively safe assets (liquidity). The intuition is that transactional needs generate a form of risk aversion even in risk neutral people. When an agent has the opportunity/desire to buy, having twice as much pledgeable wealth in some states does not compensate her for the risk of having half as much in other states, because in the latter states she is wealth constrained. As a result, agents are willing to hold relatively safe assets even if they have a lower yield.

We analyze these issues in a simple economy consisting of two groups of agents: doctors and builders. Doctors buy building services from builders and then builders buy doctors' services from doctors, or the other way round. Agents are also endowed with wheat. There is a lack of a simultaneous double coincidence of wants: each builder requires a doctor at a different date and typically one with different skills from the doctor he is building for, and vice versa for builders. Since future income is not pledgeable this generates a need for liquidity. Wheat is costly to carry and can easily rot, so banks arise in our model as depositary institutions, which store wheat and issue notes. An alternative to depositing wheat in a bank is to invest it in higher return activities that take place outside the banking sector. Initially, we suppose that the returns from these activities are not pledgeable.

For much of the paper we simply assume that low return activities inside the banking sector are pledgeable and that high return activities outside are not. Later we delve more deeply into what determines whether an asset's return can be collateralized for the purposes of liquidity.

We start by studying the effect of the supply of notes on equilibrium prices and on social welfare when banks are purely passive institutions that "notify" all deposits, i.e., an agent who deposits a unit of wheat receives a note of equal value. In notifying the deposits of a doctor who buys building services before he sells doctors services, a bank imposes a negative externality on doctors in the same position at other banks: raising the amount of money increases the price of building services, which is bad for them since they consume these services. Because of this externality too much wheat endowment is stored to create liquidity instead of being invested in socially productive activities.

The distortion in liquidity that we identify is present even when banks can control the notification process, as long as they act in a competitive way. A monopolistic bank, in contrast, ends up under-producing money; this is the standard result that a profit maximizing monopolist restricts production. We show that the distortion is robust to the introduction of lending and government money. We introduce government money by supposing that the government can impose sales taxes and that agents can pay these sales taxes with government notes.
While the introduction of government money potentially crowds out the need for private money, government money is costly to the extent that sales taxes impose deadweight losses.

As noted, in most of the paper we take the distinction between collateralizable and non-collateralizable assets as exogenous. In Section 6 we go into this more deeply. Suppose that activities that earn a higher expected return are also riskier. Assume that the returns from these activities are positively correlated and that uncertainty about them is resolved before trading takes place. Then a high return realization provides large amounts of liquidity for the economy, while a low return realization provides low amounts of liquidity. However, since there are diminishing returns to liquidity—the value of liquidity falls to zero when the gains from trade have been exhausted—this induces the equivalent of risk aversion in agents: the yield on the high return asset is discounted in the good state, and as a result the safe asset is favored for liquidity purposes. Given this interpretation of the model our main result is that an unregulated market economy allocates an excessive amount of resources to assets whose returns are stable.

Our paper is related to two literatures: that on banking and that on money. Much of the banking literature is concerned with the role played by banks and the need for bank regulation (see, for example, Dewatripont and Tirole (1994)). One branch of the banking 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 the 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), Allen and Gale (1998) and (2007), and Bhattacharya and Gale (1987)), or to reduce future adverse selection (Gorton and Pennacchi (1990)). There are also 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 reason 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 externalities in the creation of money. A general theory of banking would bring all these frictions together.

A recent 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 claims 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 Stein (2011). In his model, however, it is assumed that agents have a discontinuous demand for a riskless claim (money). 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 richer in terms of implications for monetary policy. In this respect, the two models can be seen as complementary. Our results are also related to those in chapter 7 of Holmstrom and Tirole (2011). They show that investors tend to hold an excessive amount of liquidity to take advantage of potential firesales. Our model shows that this excess of liquidity is not specific to the storage role of money; it applies also to its transactional role.

Our paper is also related to the huge literature on money. Much of this literature is concerned with the role money plays in general equilibrium (e.g., Hahn (1965)). To create such a role, one needs to dispense with the traditional Walrasian auctioneer and explicitly introduce an exchange process. 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 externality we identify in our paper. The role money plays in our model (i.e., to address the lack of double coincidence of wants) is similar to that in 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 medium of exchange. Another large part of the literature on money analyzes the role of inside money on monetary policy, as in Brunnermaier and Sannikov (2011), 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 externalities in the creation of liquidity.

Finally, our approach is close 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 they have sold their own 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

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 bank lending. In Section 5 we study the interaction between private money and government money. In Section 6 we consider some extensions, while in Section 7 we discuss the implications this paper has for the nature of money. Conclusions follow.

2. The Framework

We consider an economy that lasts three periods:

1 ---------------------------------------2---------------------------------------3

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. Doctors and builders can consume wheat in any period and there is no discounting. Both doctors and builders have an endowment of wheat in period 1 equal to e. As will become clear we will assume that e is "large".

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 sum of the quantities of wheat consumed by individual $i = d, b$ in each period; 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. We normalize the price of wheat to be 1 in all periods. Let p_b and p_d be the price respectively of building and doctor services. In words, doctors and builders have a constant marginal utility of wheat, a constant marginal utility of the service provided by the other group of agents, and a quadratic disutility of labor.

In period 1 each agent learns whether he will first buy or sell. Ex ante both events are equally likely. For convenience, we assume that in the east side of town doctors buy builders' services in period 2, while builders buy doctors' services in period 3. In the west side of town, the order is reversed. After each agent learns whether he will first buy or sell (equivalently, whether he is in the east or west side of town), he can invest part of his wheat endowment in a project that pays off at the end of period 3 (in the form of wheat), and whose return is $\overline{R} > 1$; the rest of the endowment is (costlessly) stored. For the moment we take the return from the investment project to be perfectly certain. The investment opportunity is specific to the agent.

We have deliberately set up the model to be very symmetric; this helps with the welfare comparisons later.

In periods 2 and 3 the market meets and the doctors and builders trade in the order determined in period 1. Throughout the paper we will analyze the east part of town, where doctors buy in period 2 and builders in period 3; the reverse case is completely symmetrical.

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 simultaneous* double coincidence of wants: builders and doctors are in either the market for buying or the market for selling: they cannot do both at the same time. Hence, even if the builder a doctor buys from wants the doctor services from his customer, he cannot buy them at the same time as he is selling building services.

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 in period 3 *and* from the return from their investment project. This is the assumption made in classic Walrasian or Arrow-Debreu theory and it is easy to compute the Walrasian equilibrium.

Clearly both the doctors and the builders invest all their endowment in the investment project since this maximizes their wealth; and under complete markets maximizing wealth is a necessary condition for maximizing utility. Thus, a doctor solves the following maximization problem:

(2.1) $$\text{Max } x_d + b_d - \frac{1}{2} l_d^2$$

$$\text{S.T.} \quad x_d + p_b b_d \leq p_d l_d + e\bar{R}.$$

The solution is

(2.2) $$l_d = p_d \quad \text{if} \quad p_b \geq 1$$

$$= \frac{p_d}{p_b} \quad \text{if} \quad p_b < 1$$

$$b_d = 0 \quad \text{if} \quad p_b > 1$$

$$b_d = \frac{p_d^2}{p_b^2} + \frac{e\bar{R}}{p_b} \quad \text{if} \quad p_b < 1$$

$$0 \leq b_d \leq \frac{p_d^2}{p_b^2} + \frac{e\bar{R}}{p_b} \quad \text{if} \quad p_b = 1$$

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, a builder solves:

(2.3) $$\text{Max } x_b + d_b - \frac{1}{2} l_b^2$$

$$\text{S.T.} \quad x_b + p_d d_b \leq p_b l_b + e\bar{R}.$$

The solution is

(2.4) $$l_b = p_b \quad \text{if} \quad p_d \geq 1$$

$$= \frac{p_b}{p_d} \quad \text{if} \quad p_d < 1$$

$$d_b = 0 \qquad \text{if} \quad p_d > 1$$

$$d_b = \frac{p_b^2}{p_d^2} + \frac{e\bar{R}}{p_d} \qquad \text{if} \quad p_d < 1$$

$$0 \leq d_b \leq \frac{p_b^2}{p_d^2} + \frac{e\bar{R}}{p_d} \qquad \text{if} \quad p_d = 1$$

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) $$b_d = l_b$$

(2.6) $$d_b = l_d$$

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 $e\bar{R}$ ((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: In the unique Walrasian equilibrium all wheat is invested and $p_b = p_d = b_d = l_b = d_b = l_d = 1$. *The utilities of the doctors and builders are* $U_d = e\overline{R} + \frac{1}{2}$, $U_b = e\overline{R} + \frac{1}{2}$, *respectively, and total welfare (social surplus) equals* $W \equiv U_d + U_b = 2e\overline{R} + 1$.

Note that the Walrasian allocation and prices are independent of the initial endowment e and $\overline{R}$ (except for consumption of wheat, which varies one to one with $e\overline{R}$). Also the Walrasian equilibrium achieves maximal social surplus; this follows from the first theorem of welfare economics and the symmetry of the parties. In what follows we will refer to the Walrasian equilibrium allocation as the first best.

3. Introducing Banks

We now suppose that parties cannot pledge their future labor income or investment income—the returns from these activities can be diverted or hidden. We also assume that it is impossible for individuals to carry wheat around with them when they trade; it is too cumbersome or the wheat would rot or be stolen. In the absence of any further assumption the model now becomes trivial. Trade between doctors and builders is impossible—they have nothing to trade with— and so each agent invests and eats his wheat in period 3. We have $U_d = e\overline{R}$ and $U_b = e\overline{R}$.

However, we now introduce storage facilities. These storage facilities are perfectly secure in the sense that wheat deposited at a facility in period 1 will remain there and be intact at the end of period 3.

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 f notes, where each note is a claim on a unit of wheat at the end of period 3; he can then use these notes to pay builders. The builders in turn can use these notes to pay doctors. At the end of period 3 the holders of the notes can go to the storage facility and redeem them for an equal number of units of wheat. We call these storage facilities banks. At the moment, these are completely passive institutions, which just store and issue notes on a one-for-one basis.

3.1. Individual Optimization

Let's consider first the doctors. In period 1, a doctor invests $e - f_d$ units of wheat in the investment project and deposits f_d units of wheat in the bank, receiving f_d in notes. In period 2 he uses these notes to purchase $\frac{f_d}{p_b}$ units of building services. In period 3, he will choose l_d to maximize $p_d l_d - \frac{1}{2} l_d^2$, i.e., set $l_d = p_d$. Note that it is too late for the doctor to buy more building services and so his marginal return from work is p_d rather than $\frac{p_d}{p_b}$. A doctor's labor yields revenue p_d^2 in the form of notes, which he redeems for wheat at the end of period 3; in addition he incurs an effort cost of $\frac{1}{2} p_d^2$, and so his net utility is $\frac{1}{2} p_d^2$. Finally, he will also receive and consume the payoff from his investment in the form of wheat.

Therefore, a doctor's (expected) utility when he buys first is

$$\frac{f_d}{p_b} + \frac{1}{2} p_d^2 + (e - f_d)\bar{R} \quad . \tag{3.1}$$

Each doctor chooses f_d to maximize (3.1), taking prices as given. The first order condition for an interior solution is

$$\frac{1}{p_b} = \bar{R}. \tag{3.2}$$

Similarly, a builder's utility is

$$\frac{f_b}{p_d} + \frac{1}{2} \frac{p_b^2}{p_d^2} + (e - f_b)\bar{R}.$$

Note that a builder's marginal return from work is $\frac{p_b}{p_d}$, since she will use her income to buy doctor services. We shall see shortly that builders will be a corner, so their first order condition is

(3.3) $$\frac{1}{p_d} \leq \overline{R} .$$

3.2 Market Equilibrium

We solve for the equilibrium under the conjecture that $f_b = 0$. Since only $f_d > 0$, we will drop the subscript and set $f_d \equiv f$. We also conjecture that p_d and p_b are less than one. In due course we will show that these conjectures are correct.

We work backwards, starting with the market for doctors in period 3. After the doctors have bought building services, the builders find themselves with a quantity f of notes. Hence, their demand for doctor services will be given by $\frac{f}{p_d}$ .The supply of doctor services is

(3.4) $$l_d = p_d .$$

Hence, market clearing requires

(3.5) $$\frac{f}{p_d} = p_d .$$

Similarly, the market-clearing condition in the building services market is

(3.6) $$\frac{f}{p_b} = \frac{p_b}{p_d}$$

since the labor supply of builders is $\frac{p_b}{p_d}$. Combining this with (3.5) yields

(3.7) $$p_d = f^{\frac{1}{2}},\ p_b = f^{\frac{3}{4}} .$$

Substituting the equilibrium price into (3.2) we have:

(3.8) $$f^{-\frac{3}{4}} = \overline{R}$$

or

$$f = \hat{f} = \frac{1}{\overline{R}^{\frac{4}{3}}}$$

where we assume that e is larger than $\hat{f}$, so the solution is not at the corner. Since $\overline{R}>1$, $\hat{f} <1$, which implies $1> p_d > p_b$. Hence (3.3) holds as an inequality and the builders will be at a corner solution with $f_b = 0$, as initially conjectured.

Note that the level of trade of doctors' services is $\frac{f}{p_d} = f^{\frac{1}{2}} < 1$ and of builders' services is $\frac{f}{p_b} = f^{\frac{1}{4}} < 1$. In other words trade levels are lower than in the Walrasian equilibrium. (Recall that in the Walrasian equilibrium $p_d = p_b =1$ and one unit of each service is traded.)

Proposition 2: If $\overline{R}>1$ and $e > \hat{f}$, there is less trade and prices are lower in the market equilibrium when agents cannot pledge future income than in the Walrasian equilibrium.

3.3 Social Optimum

Obviously the market equilibrium is not first best optimal given that it operates below the Walrasian equilibrium level of trade. We now show that the market equilibrium is also not second best optimal: a planner operating under the same constraints as the market can do better. Recall that ex ante it is not known who will buy first: doctors or builders. Thus the expected utility of each group is $\frac{1}{2}U_d + \frac{1}{2}U_b$ .The social optimum is obtained by maximizing $U_d + U_b$ taking into account the effect of f on prices.

That is, the planner maximizes

(3.9) $$U_d + U_b = f^{\frac{1}{4}} + \frac{1}{2}f + \frac{1}{2}f^{\frac{1}{2}} + (e-f)\overline{R} + e\overline{R}\,.$$

The first order condition is

(3.10) $$\frac{1}{4}f^{-\frac{3}{4}} + \frac{1}{2} + \frac{1}{4}f^{-\frac{1}{2}} = \overline{R}\,.$$

Comparing the left-hand side of (3.8) and (3.10) we have

Proposition 3: If $\bar{R}$ >1 and e > $\hat{f}$, the market equilibrium leads to an excessive amount of private money f, and an excessive trade level, relative to the second best social optimum.

Proof: Since the l.h.s. of both (3.8) and (3.10) are decreasing in *f,* to prove that the solution of (3.8) is large than the solution of (3.10) it is enough to show that $f^{-\frac{3}{4}} > \frac{1}{4} f^{-\frac{3}{4}} + \frac{1}{2} + \frac{1}{4} f^{-\frac{1}{2}}$. This can be rewritten as $\frac{1}{2}(f^{-\frac{3}{4}} - 1) + \frac{1}{4}(f^{-\frac{3}{4}} - f^{-\frac{1}{2}}) > 0$, which is true as long as $f < 1$. When *f=1,* the inequality becomes an equality. Trade levels are higher in the market equilibrium since equilibrium trade levels, given by $f^{\frac{1}{2}}$ for doctor services and $f^{\frac{1}{4}}$ for building services, are monotonically increasing in f .

The social and private returns from varying f are illustrated in Figure 1.

[Figure 1 here]

There are two types of inefficiency. With respect to the Walrasian equilibrium, there is too little trade but at the same time too much wheat is invested in liquidity-creating unproductive storage instead of productive projects. This inefficiency is due to the lack of pledgeabilty of future income. In the Walrasian equilibrium there is no liquidity problem since future (labor) income can be pledged and so no wheat is invested in unproductive storage. In addition, there is an inefficiency with respect to the second best. The creation of more means of payments imposes a positive externality on the builders (who see the price of their building services go up) and a negative externality on the other doctors (who see the price of what they are buying go up). In standard models the effect of these "pecuniary" externalities is second order and thus does not create a divergence between social and private optimality. Here, however, while the positive externality on other builders is second order, the negative externality on other doctors is first order, since the doctors are liquidity constrained. (The builders who sell before they buy are not liquidity constrained.) Thus the market equilibrium yields too much liquidity and too much trade relative to the second best, even though too little trade relative to the first best.

Notice that if $\bar{R}$ =*1*, the social and private solutions do not differ: they are both $f = 1$(as long as $e \geq 1$). In this case the first best is achieved. There is still a divergence between private and social incentives, but this divergence is infra-marginal.

As we have noted, as long as $\bar{R} > 1$, the economy will operate below the Walrasian equilibrium level of trade, regardless of the quantity of endowment *e*. Because high return projects cannot be collateralized, and the only investments that can be collateralized (wheat storage) have a lower yield, there is an opportunity cost of creating liquidity. Thus the optimal amount of liquidity is too low from a first best efficiency point of view: in the first best there would be enough liquidity to generate a trade of one unit of each service. This conclusion is reminiscent of Friedman (1969)'s famous result that with a non-negative rate of inflation people hold too little money. In Friedman, however, this inefficiency is in the form of a shoe-leather cost; here it is in the form of missed trading opportunities. The ultimate source of inefficiency in our paper is lack of pledgeability: the inefficiency would disappear if all labor income could be pledged or if the return from the non-collateralizeable investments were equal to 1.

3.4 Overproduction of Collateral

In our model there are three types of activities: non-collateralizeable investments (high return projects), collateralizeable investments (low return storage), and future labor income. Interestingly, it would be natural to think that if future labor income is contractible, the inefficiency would disappear. In Section 4 we show this is not the case.

Note that if some high return projects are (or become) collateralizeable, their prices will jump from $\frac{1}{\bar{R}}$ to $\left(\frac{1}{\bar{R}}\right)^2$, since they can be used both to produce $\bar{R}$ in period 3 and to buy goods in period 1, which has a return of $\frac{1}{p_b} = \bar{R}$. This conclusion can be generalized: assets that can be collateralized to back credit will trade at a higher price than otherwise identical assets that cannot be collateralized to produce credit. As a result, if we add an early period where effort is exerted to produce various assets, there will be an overproduction of assets that can be collateralized to produce liquidity (Madrigan and Philippon, 2011). Therefore, there will be excess resources invested in collateralizeable assets and/or an overproduction of low-yielding assets that can be collateralized to produce liquidity, a result similar to the one Madrigan and Philippon (2011) obtain for houses.

4. BANK LENDING

Let's now consider the case where banks have some ability to seize payments that go through the banking sector. Specifically, suppose that all payments for building and doctor services take place through check transfers and that the bank is able to seize these before they are cashed for consumption. In other words, labor income is now contractible. However, a bank cannot force anyone to work: human capital is inalienable (as in Hart and Moore (1994)). That is, all a bank can do is to ensure that someone who defaults has zero consumption, apart from their investment income.

We continue to assume that investment income cannot be seized. In Section 6.1 we show that with a different timing of the model this assumption can be relaxed.

Given that doctors go first, only doctors will want to borrow. Builders, who buy second, will obtain no advantage from borrowing. A bank, knowing that it will have no power over builders at the end of period 3, will insist on being repaid before it approves the builders' payment to doctors. But then builders will have to repay their debt before they buy doctors' services, making their borrowing useless.

Let β be the amount borrowed by each doctor. By borrowing an amount β an individual doctor can consume $\frac{\beta}{p_b}$ of building services, which is attractive if he has to pay back only β (given p_b <1). Of course, a bank needs to make sure that it will be repaid. In period 3, a doctor exerts $\frac{1}{2}p_d^2$ of effort, receiving in exchange p_d^2 in terms of payment. His net utility is $\frac{1}{2}p_d^2$. Thus, he cannot borrow more than $\frac{1}{2}p_d^2$: if he did he would prefer not to work in period 3, default, and consume nothing (except for his investment income).

A doctor's utility is now given by

(4.1) $$\frac{f+\beta}{p_b}+\frac{1}{2}p_d^2-\beta+(e-f)\overline{R}\ .$$

Notice that given $p_b<1$, a doctor's utility is increasing in β; thus a doctor will borrow up to the constraint.

The equilibrium in the market for doctors' services will be given by

$$\frac{f+\beta}{p_d}=p_d$$

and the equilibrium in the market for building services will be given by

$$\frac{f+\beta}{p_b}=\frac{p_b}{p_d}.$$

Since doctors will borrow as much as possible we will have

$$\beta=\frac{1}{2}p_d{}^2=\frac{1}{2}(f+\beta)$$

or

(4.2) $$\beta=f\ .$$

Hence,

$$p_d=(2f)^{\frac{1}{2}} \qquad \text{and} \qquad p_b=(2f)^{\frac{3}{4}}.$$

From (3.2) the FOC in a competitive market is

(4.3) $$(2f)^{-\frac{3}{4}}=\bar{R}$$

or

$$\hat{\hat{f}}=\frac{1}{2}\frac{1}{\bar{R}^{\frac{4}{3}}}=\frac{1}{2}\hat{f}$$

where $\hat{f}$ is the solution of (3.8) when there was no borrowing and $\hat{\hat{f}}$ is the solution of (4.3) when there is borrowing. Since $2\hat{\hat{f}}=\hat{f}$, it follows that the presence of lending keeps prices constant. However, lending cuts in half the amount of wheat that is notified, reducing welfare losses, since more can be invested productively. Without lending e - $\hat{f}$ is invested in the risky project, while with lending, the investment rises to $e-\hat{\hat{f}}$.

It is easy to see that $U_d(\hat{\hat{f}})>U_d(\hat{f})$, since

$$U_d(\hat{\hat{f}})=\frac{\hat{\hat{f}}+\beta}{p_b}+\frac{1}{2}p_d^2+(e-\hat{\hat{f}})\bar{R}-\beta$$

and

$$(2\hat{\hat{f}})^{\frac{1}{4}}+(e-\hat{\hat{f}})\bar{R}=(\hat{f})^{\frac{1}{4}}+(e-\frac{1}{2}\hat{f})\bar{R}>(\hat{f})^{\frac{1}{4}}+\frac{1}{2}\hat{f}+(e-\hat{f})\bar{R}=U_d(\hat{f}).$$

Similarly, we have that

$$U_b(\hat{\hat{f}}) = \frac{1}{2}(\frac{p_b}{p_d})^2 + e\overline{R} = \frac{1}{2}(2\hat{\hat{f}})^{\frac{1}{2}} + e\overline{R} = \frac{1}{2}\hat{f}^{\frac{1}{2}} + e\overline{R} = U_b(\hat{f}).$$

The planning solution is obtained by maximizing the sum of the utilities of a doctor and a builder (i.e., the expected utility of each):

$$W = \frac{f+\beta}{p_b} + \frac{1}{2}p_d^2 + (e-f)\overline{R} - \beta + \frac{1}{2}(\frac{p_b}{p_d})^2 + e\overline{R}.$$

Substituting the value of p_d and p_b, we obtain

$$W = (2f)^{\frac{1}{4}} + (e-f)\overline{R} + \frac{1}{2}(2f)^{\frac{1}{2}}.$$

Thus, the FOC is

(4.4) $$\frac{1}{2}(2f)^{-\frac{3}{4}} + \frac{1}{2}(2f)^{-\frac{1}{2}} = \overline{R}.$$

By comparing (4.3) and (4.4) we have

Proposition 4: If $\overline{R}$ >1, the market equilibrium leads to an excessive amount of private money even in the presence of lending. Lending, however, increases welfare.

Proof: Since the l.h.s. of both (4.3) and (4.4) are decreasing in *f,* to prove that the solution of (4.3) is larger than the solution of (4.4) it is enough to show that $(2f)^{-\frac{3}{4}} > \frac{1}{2}(2f)^{-\frac{3}{4}} + \frac{1}{2}(2f)^{-\frac{1}{2}}$,

or $(2f)^{-\frac{3}{4}} > (2f)^{-\frac{1}{2}}$, which is always true since $2f < 1$ by (4.3).

Lending, thus, does not resolve the tension between private and social objectives. Nevertheless, lending does improve welfare since it increases the volume of trade without sacrificing the higher return of the alternative investment. Interestingly, lending is not a perfect substitute for the notification of wheat. To see why, look at equation (4.2). The amount of feasible lending is directly related to the amount of notes present in the system. The reason is that lending faces a repayment constraint. With no notes in the system, the amount borrowed by doctors equals the purchasing power in the hands of builders, which in turn equals the revenue received by doctors for their services. But if the revenue equals the debt, it is not in the interest of the doctors to work, given that they have to exert costly effort. Hence, the doctors will default. To have a functioning lending market, we need a minimum amount of deposits.

Finally, it is important to emphasize that we have only scratched the surface of borrowing. If we introduce uncertainty in the proceeds from trade, then borrowing will be risky. (In principle it could be state-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.

5. Interaction between Private and Public Money

So far we have ignored any role of the government in providing liquidity. In this finite horizon economy, to introduce government money we need to specify why it is accepted. Following a long tradition (e.g., Cochrane (1998)), we assume that government money is valuable because one can pay taxes with it. We suppose that each agent receives an amount m of government money in period 2 after investments have been made but before any trading takes place, and will have to pay some taxes in period 3. We assume that the agent has an option to pay taxes either in dollars or an equal number of units of wheat. This anchors the value of money.

In this type of model, it is generally assumed that taxation takes the form of a non-distortionary lump sum tax. What is ignored is that even a lump sum tax has some potential distortions. What happens if an individual refuses to pay the lump sum tax? Presumably, he would be thrown in jail. To make this a credible threat, however, the government would have to build prisons in advance, which is in itself distortionary.

Rather than working with a lump sum tax and a jail threat (the results are similar but less elegant), we prefer to use a standard model of a distortionary tax, such as a consumption tax. Specifically, we assume that the government can impose a mill tax on those who turn wheat into flour.

Assume that each agent can obtain λ units of flour in period 3 at the cost of $\frac{1}{2}c\lambda^2$ units of wheat. This activity occurs at facilities that can easily be monitored by the government, and so the sales tax cannot be avoided. One unit of flour yields one unit of utility. An agent's utility is now:

Doctors: $U_d = x_d + b_d - \frac{1}{2} l_d^{\,2} + (1-t)\lambda_d - \frac{1}{2} c \lambda_d^{\,2}$

Builders: $U_b = x_b + d_b - \frac{1}{2} l_b^{\,2} + (1-t)\lambda_b - \frac{1}{2} c \lambda_b^{\,2}$

where t is the tax rate on flour.

We assume that this transformation from wheat into flour can be financed out of the return from the investment. We suppose that this return is always high enough so that agents are not at a corner solution, and so λ_d, λ_b satisfy the first order condition

(5.1) $$\lambda_d = \lambda_b = \frac{1-t}{c}.$$

Hence, budget balance for the government implies

(5.2) $$m = \frac{t(1-t)}{c}.$$

5.1 No Private Borrowing

For simplicity let's start with the case where there is no private borrowing. Consider the market for doctors. The number of notes in the hands of builders will be *f*+*m* (from trading with doctors) plus *m* of their own. Thus, market equilibrium is given by

(5.3) $$\frac{f+2m}{p_d} = p_d.$$

We assume that $f + 2m < 1$; otherwise (5.3) would be replaced by p_d =1.

On the other hand, in the market for builders the number of notes available to doctors is *f*+*m,* so the market equilibrium is given by

(5.4) $$\frac{f+m}{p_b} = \frac{p_b}{p_d}.$$

Solving (5.3) and (5.4) yields

(5.5) $$p_d = (f+2m)^{\frac{1}{2}} \quad \text{and} \quad p_b = (f+m)^{\frac{1}{2}}(f+2m)^{\frac{1}{4}}.$$

We can write the utilities of doctors and builders as

(5.6) $$U_d = \frac{f+m}{p_b} + \frac{1}{2}p_d^2 + (e-f)\bar{R} + \frac{1}{2c}(1-t)^2$$

(5.7) $$U_b = \frac{m}{p_d} + \frac{1}{2}(\frac{p_b}{p_d})^2 + e\bar{R} + \frac{1}{2c}(1-t)^2$$

where we use (5.1) to substitute for λ_d, λ_b.

The competitive equilibrium is characterized by the first order condition of (5.6) with respect to *f*, where prices are taken as given:

(5.7) $$\frac{1}{p_b} = \bar{R}\,.$$

As noted we require the solution to (5.5) and (5.7) to satisfy $f + 2m < 1$ for these formulae to be correct. This will be true as long as *m* is not too large. We assume this in what follows.

In contrast, the planner maximizes $W = U_d + U_b$ taking into account the effects of *f* and *m* on prices. In other words, the planner maximizes

(5.9) $$W = \frac{(f+m)^{\frac{1}{2}}}{(f+2m)^{\frac{1}{4}}} + \frac{1}{2}f + m + \frac{1}{c}(1-t)^2 - f\bar{R} + \frac{1}{2}\frac{f+m}{(f+2m)^{\frac{1}{2}}} + \frac{m}{(f+2m)^{\frac{1}{2}}} + 2e\bar{R}\,.$$

Several questions are worth asking. First, starting at the competitive equilibrium, does the planner want to introduce outside money, i.e., $m > 0$, given that the market will adjust to a new competitive equilibrium? The answer seems to be ambiguous, perhaps because the market provides too much of its own money (sets *f* too high).

A second question is: given that the planner can regulate *f,* does she want to set $m > 0$? Here the answer is affirmative. To see this, set *f* at the regulatory optimum, i.e., maximize *W* with respect to *f* when $m = 0$ (this is equivalent to maximizing (3.9)). Now consider a small change in m (or equivalently in *t*). By the envelope theorem,

(5.10) $$\frac{dW}{dm}\Big|_{m=0} = \frac{\partial W}{\partial m}\Big|_{m=0} = \frac{f^{\frac{1}{4}}\frac{1}{2}f^{-\frac{1}{2}} - f^{\frac{1}{2}}\frac{1}{4}2f^{-\frac{3}{4}}}{f^{\frac{1}{2}}} + 1 - \frac{2(1-t)}{c}\frac{dt}{dm} + \frac{1}{2}\frac{f^{\frac{1}{2}} - f\frac{1}{2}f^{-\frac{1}{2}}2}{f} + \frac{f^{\frac{1}{2}}}{f}$$

From (5.2), $\frac{dt}{dm} = \frac{c}{1-2t} = c$ at $m = 0$, and so

(5.11) $\frac{dW}{dm}\Big|_{m=0} = -1 + \frac{1}{f^{\frac{1}{2}}} > 0$

since (5.5) and (5.8) imply $f < 1$ when $m = 0$ (given $\bar{R} > 1$). It follows that it is always better for the planner to set $m > 0$ than $m = 0$.

A third question is: given that the regulator can choose *m* (and *t)* optimally, does she want to set $f = 0$, i.e., does outside money crowd out private money? Here the answer is: it depends. It is clear from (5.10) that f is small when $\bar{R}$ is large. In fact one can go further: $\frac{dW}{df}\Big|_{f=0}$ <0 when $\bar{R}$ is large at the optimal *m*, and so the socially optimal *f* is zero for large enough $\bar{R}$. On the other hand, we have seen that when $\bar{R}$ is close to 1 the private market equilibrium approximates the first best, and so private money is more efficient than outside money. We can carry out a similar exercise for variations in *c*. If c is small the deadweight costs of taxation are small and so outside money is more efficient than private money; the socially optimal f will be close to zero. On the other hand, if *c* is large, then the maximum possible value of *m* (given by $\frac{1}{4c}$) is small.[1] In other words, a pure outside money economy will achieve very little trade, and private money by itself can do better. Thus, the optimal *f* is bigger than zero.

In summary, private money and government money have different costs and benefits and hence are not perfect substitutes. Private money is costly because it crowds out productive investment. Government money is costly because it is backed by distortionary taxes that create deadweight losses.

5.2 Case with Private Borrowing

We now show that our result that some outside money is optimal ($m > 0$) generalizes to the case of borrowing. In the presence of private borrowing (5.3) and (5.4) are replaced by

(5.12) $$\frac{f + \beta + 2m}{p_d} = p_d$$

and the equilibrium in the market for building services will be given by

[1] This maximum value of *m* can be obtained from maximizing (5.2) with respect to *t*.

(5.13) $$\frac{f+\beta+m}{p_b}=\frac{p_b}{p_d}.$$

For these formula to apply we need $f+\beta+2m<1$. We then have

(5.14) $$p_d=(f+\beta+2m)^{\frac{1}{2}} \quad \text{and} \quad p_b=(f+\beta+m)^{\frac{1}{2}}(f+\beta+2m)^{\frac{1}{4}}.$$

We know that $\beta=\frac{1}{2}p_d^2$, from which it follows that

(5.15) $$\beta=f+2m.$$

Hence, total welfare is given by

(5.16) $$W=\frac{f+\beta+m}{p_b}+\frac{1}{2}p_d^2-\beta+(e-f)\overline{R}+\frac{1}{2c}(1-t)^2+\frac{m}{p_d}+\frac{1}{2}(\frac{p_b}{p_d})^2+e\overline{R}+\frac{1}{2c}(1-t)^2=$$

$$=\frac{(2f+3m)^{\frac{1}{2}}}{(2f+4m)^{\frac{1}{4}}}+\frac{1}{c}(1-t)^2-f\overline{R}+\frac{1}{2}\frac{2f+3m}{(2f+4m)^{\frac{1}{2}}}+\frac{m}{(4f+4m)^{\frac{1}{2}}}+2e\overline{R}.$$

The planner will maximize (5.16) with respect to *f* and *m* (given that β adjusts to satisfy (5.15)). Using the envelope theorem we have

(5.17) $$\frac{dW}{dm}\Big|_{m=0}=\frac{\partial W}{\partial m}\Big|_{m=0}=\frac{1}{2}(2f)^{-\frac{1}{2}}-2+\frac{3}{2}(2f)^{-\frac{1}{2}}>0$$

if $f<\frac{1}{2}$, which we know to be the case from (4.3). Hence, it is optimal to set $m>0$.

Interestingly, the introduction of government money does not eliminate the role for bank lending. In fact, it is complementary to bank lending, as shown by equation (5.15). The reason is similar to the one highlighted in Section 4: With no notes in the system, the amount borrowed by doctors equals the purchasing power in the hands of builders, which in turn equals the revenue received by doctors for their services. But if the revenue equals the debt, it is not in the interest of the doctors to work, given that they have to exert costly effort. To have a functioning lending market, thus, we need either private or outside money. Lending multiplies the effect from the injection of money.

5.3 Pigovian Taxation

So far we have only considered government money backed by a distortionary tax. In Section 3, however, we showed that the competitive equilibrium produced too much liquidity and that that government intervention could improve welfare. In this context it is natural to wonder whether a Pigovian tax on deposits can kill two birds with one stone: raise revenue to support more government money while *decreasing* the inefficiency associated with notification.

Surprisingly, the answer is that such a scheme will not work: the reason is that the tax will be raised at the same time deposits are made, worsening the constraint that agents face. Formally, let's assume that the government can require that anybody who deposits at a bank chooses a particular f and pays a total tax t.[2] We will show that the welfare maximizing t equals zero, i.e. the government does not want to impose any Pigovian tax.

A deposit tax will be assessed only on the early buyers (according to our convention, the doctors) since these are the only ones who deposit wheat. The total amount of money that can be supported by the tax revenue is $M=t$. Since the government cannot tell who needs to buy first, the money will be distributed equally between doctors and builders, each agent receiving an amount $m=\frac{t}{2}$.

Thus, the doctors will enter the buying phase with an amount of liquidity equal to $f-t+m$ and the builders, who will be paid by the doctors that amount, will enter their buying phase with an amount $f-t+2m$.

The equilibrium in the doctors' market requires that

(5.18) $$\frac{f-t+2m}{p_d}=p_d\ ,$$

which yields $p_d=f^{\frac{1}{2}}$.

The equilibrium in the builders' market requires that

(5.19) $$\frac{f-t+m}{p_b}=\frac{p_b}{p_d}\ ,$$

[2] One way for the government to do this is for them to impose a 100% marginal tax on deposits above f .

which yields $p_b = (f - \frac{1}{2}t)^{\frac{1}{2}} f^{\frac{1}{4}}$.

The expected utility of the doctors is given by $\frac{f-t+m}{p_b} + \frac{1}{2} p_d{}^2 + (e-f)\bar{R}$, while the expected utility of the builders is given by $\frac{m}{p_d} + \frac{1}{2}\left(\frac{p_b}{p_d}\right)^2$. Thus, the planner's problem is to maximize

(5.20) $$\frac{f-t+m}{p_b} + \frac{1}{2} p_d{}^2 + (e-f)\bar{R} + \frac{m}{p_d} + \frac{1}{2}\left(\frac{p_b}{p_d}\right)^2$$

subject to (5.18), (5.19) , and the doctors' individual rationality constraint (a doctor must be willing to deposit in a bank and pay the tax rather than deposit nothing).

Proposition 5: It is optimal for the government to set t equal to zero.

Proof: Fix *f*. If we compare the expected utility at $t = m = 0$ and the expected utility when $\frac{t}{2} = m > 0$ we

have $f^{\frac{1}{4}} + \frac{1}{2} f + (e-f)\bar{R} + \frac{1}{2} f^{\frac{1}{2}} \geq \frac{f - \frac{t}{2}}{(f-\frac{t}{2})^{\frac{1}{2}} f^{\frac{1}{4}}} + \frac{1}{2} f + (e-f)\bar{R} + \frac{\frac{t}{2}}{f^{\frac{1}{2}}} + \frac{1}{2}\frac{(f-\frac{t}{2}) f^{\frac{1}{2}}}{f}$, which can

we rewritten as $f^{\frac{1}{4}} \geq \frac{(f-\frac{t}{2})^{\frac{1}{2}}}{f^{\frac{1}{4}}} + \frac{1}{2}\frac{\frac{t}{2}}{f^{\frac{1}{2}}}$. The right hand side equals the left hand side at $t = 0$ and

is decreasing in t, since $\frac{\partial rhs}{\partial t} = -\frac{1}{4}\frac{(f-\frac{t}{2})^{-\frac{1}{2}}}{f^{\frac{1}{4}}} + \frac{1}{4}\frac{1}{f^{\frac{1}{2}}} < 0$ when $f < 1$. Thus, the l.h.s. is always as big as the r.h.s. and the welfare is maximized at $t = m = 0$.

In most situations in which the social optimum differs from the competitive equilibrium, the social optimum can be implemented with a simple linear tax. Proposition 5 implies that this is not the case. Even if an appropriate linear tax can restore the right incentive on the margin, it will be suboptimal, since it will raise revenues, worsening the constraints of doctors.

The optimal level of deposits can be implemented with a nonlinear tax $t = 0$ if $f \leq f^*$ and $t = f$ if $f > f^*$, since in equilibrium this tax will raise no revenues. Yet, this tax is very informational intensive. It requires the planner to know the optimal level of deposits for each individual. This is trivial in the model, since everybody is the same, but it is not in reality.

In reality, however, there are better ways to deal with this problem, which do not require so much information. Since what matters is the aggregate level of deposits, the planner can limit herself to controlling that instead. Even with multiple banks, she can do this by issuing a fixed number of permits and allowing banks to accept deposits only if they own the right amount of permits.

This is exactly what central banks do with reserve requirements. The central bank requires each bank to deposit a fraction of its own deposits at the central bank. In this way a bank can accept a deposit only if it has enough deposits at the central bank. Thus, by controlling the level of its own deposits, a central bank can control the level of deposits in all the banks and (in this simple economy) the liquidity of the overall system.[3]

6. Extensions

6.1 Banks Investing in High Return Projects

We have assumed that high return projects can be undertaken only by the agents and not by the banks. What would happen if the banks could undertake the high return projects too? In this case each doctor would deposit 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 would then be used for trading. The doctors would endorse these shares to the builders in exchange for building services and the builders would further endorse them to doctors when they buy their services. There is no longer a trade-off between unproductive investments that provide liquidity and productive investments that don't. The first best can be achieved if agents invest all their endowment in the productive project as long as $e\bar{R} \geq 1$.

[3] This example applies to a situation of fiat money not of specie money, as in our model. However, even in a specie money economy central banks control the total amount of deposits with some fractional reserve.

A similar conclusion is reached if only individuals can invest in productive projects as long as these projects can be collateralized, i.e., their returns fully pledged.

We now show that these conclusions change if the high return projects are risky, highly correlated, and some uncertainty about their returns is resolved before the two rounds of trading. In this case the value of the notes backed by risky investments is volatile. Although individual agents are risk neutral, the transactional role played by money makes them risk-averse vis-à-vis fluctuations in the value of notes.

Agents dislike the uncertainty about the value of their claims because if the value of their claims drops below 1, the economy operates below full potential, while if the value of the claims is above 1, no extra benefits are generated.

Let's consider the case where banks can invest in the same projects as individuals. Suppose that each bank can choose how to divide its deposits between a risky project with an expected return of $\bar{R}$ and a safe project, and this choice is perfectly contractible, so there is no moral hazard involved. Let's also assume that uncertainty about returns gets fully resolved before the first round of trading. Finally, we assume that the risky projects are perfectly correlated and have the following payoff per dollar (constant return to scale):

(6.1)
$$\tilde{R} = \frac{\bar{R}}{\varepsilon} \text{ with probability } \varepsilon$$
$$= 0 \qquad \text{with probability } 1\text{-}\varepsilon$$

The expected return of the investment is $\bar{R}$,which exceeds 1. In this section we will focus on the limiting case where $\varepsilon \to 0$ and the investment becomes infinitely risky.

In this set up, since banks can commit to the mix of safe and risky investments, it is easy to show that doctors will deposit all their wheat e in the banks. (Builders will continue to invest in their own projects.) Thus, the crucial variable is the proportion μ of deposits that a typical bank invests in the riskless asset (storage),where $1-\mu$ is invested in the risky project.

Since trading will take place with banks' shares, the equilibrium prices will differ according to the state that is realized before the beginning of trading. Denote prices in the bad state by p_b, p_d and in the good state by p_b^g, p_d^g .

With probability $1-\varepsilon$ the risky investment will yield zero, the riskless asset one, and the value of a doctor's claim will be μe . By contrast, with probability ε the risky investment will yield $\frac{\bar{R}}{\varepsilon}$ and the value of the doctor's claim will be $\mu e + (1-\mu)\frac{e\bar{R}}{\varepsilon}$. In summary, if we let g be the value of the doctors' claim, then

$$g = g_1 = \mu e + (1-\mu)\frac{e\bar{R}}{\varepsilon} \text{ in the good state} \tag{6.2}$$

$$= g_2 = \mu e \quad \text{in the bad state}$$

A doctor's expected utility is

$$(1-\varepsilon)[\frac{\mu e}{p_b} + \frac{1}{2}p_d^2] + \varepsilon[\frac{g_2}{p_b{}^2} + \frac{1}{2}p_d^{g2}]. \tag{6.3}$$

From the analysis of Sections 3 and 4,

$$p_b = (\mu e)^{\frac{3}{4}} \quad p_d = (\mu e)^{\frac{1}{2}}$$

$$p_b^g = \min\left\{g_2^{\frac{3}{4}}, 1\right\} \quad p_d^g = \min\left\{g_2^{\frac{1}{2}}, 1\right\}$$

If we keep $\mu < 1$ constant and let $\varepsilon \to 0$, p_b^g converges to one and (6.3) becomes

$$\frac{\mu e}{p_b} + \frac{1}{2}p_d^2 + (1-\mu)e\bar{R}$$

which is identical to (3.1) with $\mu e = f_d$. Hence, all the solutions will be the same as in Section 3. Each (competitive) bank will choose $\mu e = f_d$, where f_d solves (3.2) .Also as $\varepsilon \to 0$ the welfare function will become

$$W \equiv U_d + U_b = (\mu e)^{\frac{1}{4}} + \frac{1}{2}(\mu e)^{\frac{1}{2}} + \frac{1}{2}\mu e + (1-\mu)e\bar{R} + e\bar{R} \tag{6.4}$$

which is identical to (3.9) with $\mu e = f$. Thus we have

<u>*Proposition 6.*</u> 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 3 with trading in fixed claims on deposits.

The intuition is that in the good state of the world the economy has plenty of liquidity and so the Walrasian equilibrium is realized. Since this state occurs with negligible probability, however, the value of liquidity in this state is very small. Thus liquidity is provided by the safe investment, which pays off in the bad state. The risky investment still has value in expected terms but only because it provides a huge amount of extra wheat consumption in the good state. For these reasons banks will hold a mix of safe and risky investments, just as individual doctors do in the model of Section 3. In the limit the two models are equivalent.

6.2 Non-competitive Banks

So far we have treated banks as purely passive institutions. In this subsection we explore how the results change once we allow banks to be of non-negligible size and to behave strategically. We start with the conjecture that, as in the competitive case, the only agents who want to deposit are doctors. We will then verify this conjecture.

Let us assume that there is a fixed number of banks, $\frac{1}{\alpha}$, where $0<\alpha<1$. Each bank serves a fraction α 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. In what follows we will report the results only for mutual (or cooperative) banks. Considering outside-owned banks will not ameliorate matters; in fact, it would make them worse.

We assume that each bank offers the doctors in its constituency the following service: a doctor can deposit an amount of wheat up to σ and receive notes (or checks) equal to σ. Hence, σ 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.

If $\sigma \geq \hat{f}$ obtained in (3.8), then each doctor will set $f=\hat{f}$ and the bank policy would be irrelevant. Hence, we focus on the possibility that $\sigma<\hat{f}$, where each doctor will deposit the full amount allowed, σ.

Consider a single bank's choice of σ given that the bank serves a fraction α of the population of doctors and that the average choice of other banks is $\hat{\sigma}$. We know from Section

2.2. that even if every doctor deposits as much as he wants, then $p_b < 1$ and $p_d < 1$. A fortiori this must be true when $\sigma < \hat{f}$. Thus we can focus on situations where $p_b < 1$ and $p_d < 1$.

The doctor's utility will be given by

(6.5) $$\frac{\sigma}{p_b} + \frac{1}{2} p_d^2 + (e - \sigma)\bar{R}.$$

The mutual bank chooses σ to maximize the utility of a representative member, given by (6.5), taking into account the effect of the bank's choice of σ on prices p_b and p_d.

Let us consider this price effect. Given σ the total value of notes in circulation will be $\alpha\sigma + (1-\alpha)\hat{\sigma}$; 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

(6.6) $$\frac{\alpha\sigma + (1-\alpha)\hat{\sigma}}{p_b},$$

while the supply is, as in (2.4), $\frac{p_b}{p_d}$. Equating these yields

(6.7) $$p_b^2 = (\alpha\sigma + (1-\alpha)\hat{\sigma})\, p_d.$$

In the market for doctors, demand is

(6.8) $$\frac{\alpha\sigma + (1-\alpha)\hat{\sigma}}{p_d},$$

since the builders use all the notes received from doctors to buy doctor services; and supply is p_d. Combining this with (6.3) yields

(6.9) $$p_b = (\alpha\sigma + (1-\alpha)\hat{\sigma})^{\frac{3}{4}},$$

$$p_d = (\alpha\sigma + (1-\alpha)\hat{\sigma})^{\frac{1}{2}}.$$

Substituting into (6.2), we see that the utility of a representative doctor at the bank choosing σ is

(6.10) $$\frac{\sigma}{(\alpha\sigma+(1-\alpha)\hat{\sigma})^{\frac{3}{4}}}+\frac{1}{2}(\alpha\sigma+(1-\alpha)\hat{\sigma})+(e-\sigma)\bar{R}\,.$$

We study a Nash equilibrium in which each bank chooses σ to maximize (6.10), taking $\hat{\sigma}$ as given. Let $y=\alpha\sigma+(1-\alpha)\hat{\sigma}$ and $z=(1-\alpha)\hat{\sigma}$. Then, maximizing (6.10) is equivalent to maximizing

(6.11) $$y^{\frac{1}{4}}-\frac{z}{y^{\frac{3}{4}}}-e^{\frac{3}{4}}(1-\frac{1}{2}\alpha)(y-z)$$

with respect to y. It is easy to see that (6.11) is strictly concave in y. Thus, there is a unique aximize y and hence a unique aximize σ of (6.10), given $\hat{\sigma}$.

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

(6.12) $$\sigma^{-\frac{3}{4}}+\frac{1}{2}\alpha[1-\frac{3}{2}\sigma^{-\frac{3}{4}}]=\bar{R}\,.$$

Let's now verify the conjecture that builders do not want to deposit any of their wheat. Define σ_d^* to be the solution of (6.12).This is the equilibrium level if and only if the builders do not want to deposit any more wheat. The utility of a builder who deposits σ_b is given by

$$\frac{\sigma_b}{p_d}+\frac{1}{2}\frac{p_b^2}{p_d^2}+(e-\sigma_b)\bar{R}\,.$$

Thus, no builder wants to deposit if

$$\frac{1}{p_d}\le\bar{R}\,.$$

Substituting the equilibrium level of $p_d=\sigma_d^{*\frac{1}{2}}$ we have

(6.13) $$\left[\frac{1-\frac{3}{4}\alpha}{\bar{R}-\frac{1}{2}\alpha}\right]^{\frac{2}{3}} \geq \frac{1-\frac{1}{4}\alpha}{\bar{R}}.$$

As long as (6.13) is satisfied (which is certainly true for a big enough $\bar{R}$), the solution determined by (6.12) is an equilibrium. As is easy to see, the level of $\bar{R}$ for which (6.13) is satisfied depends on α. For $\alpha = 0$, (6.13) is satisfied for any $\bar{R} > 1$. For $\alpha = 0.25$, (6.13) is satisfied for $\bar{R} > 1.23$.

Having established that (6.12) characterizes an equilibrium if (6.13) is satisfied, we can compare it with (3.10) to obtain:

Proposition 7. In a competitive market (α close to 0) banks choose too high a level of deposits with respect to what is socially efficient. In a monopolistic market (α =1) banks choose too low a level of deposits with respect to what is socially efficient.

Proof: For $\alpha = 0$, (6.12) becomes (3.8) and the previous result applies. For $\alpha = 1$ the l.h.s. of (6.12) becomes $\sigma^{-\frac{3}{4}} + \frac{1}{2} - \frac{3}{4}\sigma^{-\frac{3}{4}} = \bar{R}$. Since both (6.12) and (3.10) are decreasing in their arguments, to prove that the solution of (6.12) is smaller than the solution of (3.10) it is enough to show that $\frac{1}{4}f^{-\frac{3}{4}} + \frac{1}{2} + \frac{1}{4}f^{-\frac{1}{2}} > f^{-\frac{3}{4}} + \frac{1}{2} - \frac{3}{4}f^{-\frac{3}{4}}$, or $\frac{1}{4}f^{-\frac{1}{2}} > 0$, which is always true.

The intuition for the competition result is as before. The intuition for the monopolist one is simple. Large mutual banks restrict σ, 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 σ because their impact on prices is limited.

When it comes to builders, they cannot create liquidity to increase the price of their own services; hence, they face a trade-off similar to before: increasing their own purchasing power at the expense of the return on their investments.

Since a competitive banking sector generates too much liquidity and a monopolistic one too little, by continuity there exists a level of oligopoly α that delivers the efficient level of private money. Note, however, that this level is contingent upon $\bar{R}$; thus if the level of $\bar{R}$ changes with the business cycle, so does the level of competition that delivers the first best. In other words, unless the government can somehow fine tune the level of competition over time, this does not seem a very reliable method to achieve the first best level of money.

7. Implications for the Nature of Money

What is money? In our model money is a relatively safe asset that overcomes a pledgeability problem in a world where labor income is contractible but not fully pledgeable, because workers can quit. The need for this relatively safe asset arises for transactional purposes. Transactional needs generate a form of risk aversion even in risk neutral people. When an agent has the opportunity/desire to buy, having twice as much pledgeable wealth in some states does not compensate her for the risk of having half as much in other states, because in the latter states she is wealth constrained. As a result, agents are willing to hold relatively safe assets even if they have a lower yield.

In our model the cost of money is represented by its opportunity cost. Ideally, the best money is an asset that is valuable (in a fine horizon economy one needs a desirable asset to make the pledge credible) and whose opportunity cost is low. Before the introduction of fiat money, the main commodity that was used as money, gold, was not particularly useful per se. Hence, as noted by Sargent and Wallace (1983), the inefficiency of commodity money was not the distortion in the use of gold, but its overproduction.

Sargent and Wallace (1983) seem to suggest that the introduction of fiat money eliminates this problem. In contrast, our model indicates that the partial non-pledgeability of future labor income creates an additional demand for relatively safe assets. Hence, too many resources are invested in manufacturing these relatively safe assets. An example of this overproduction is the huge expansion of the finance sector in the first decade of the new millennium, an expansion that cannot be explained by any of the traditional roles performed by finance (Philippon (2008)). While the production of AAA mortgage-backed securities might have been privately optimal, our model suggests that it was not necessarily socially optimal.

8. Conclusions

We have built a simple framework to analyze the general equilibrium implications of the creation of liquidity by banks. To isolate these effects in this paper we consider either fully-backed money or lines of credit secured by a certain income stream. As a result there is no risk of banks' default. This risk is clearly a very important problem in reality, and we plan to consider it in future work.

In a world where some assets can be collateralized and other assets cannot, our paper shows that the assets that can be collateralized will trade at a premium with respect to what their yield would imply. More surprisingly, our paper shows that the competitive equilibrium will lead to an excess of collateralizeable assets. This distortion is present even if we introduce lending and government money. The source of this distortion is a (pecuniary) externality that arises from the creation of money. More money increases the equilibrium price of goods purchased by the agents who are liquidity constrained (think, for example, of the effect of the relaxed credit on the price of U.S. houses). This externality has welfare effects, because the buyers are liquidity constrained. The government can remove this distortion by imposing a tax on deposits or by restricting the creation of deposits through a reserve requirement. .

Our model also shows that this effect can be attenuated or even eliminated in a less than competitive banking sector. Yet, we run the risk of the opposite problem: with too little competition we will have too little liquidity. Therefore, it is unlikely that restricting competition is the best way for the government to remove the externality that we have identified. It would be better for the government to tackle the problem by directly targeting the creation of liquidity.

In our two period model, outside money is very costly because we compare its one period benefit with the deadweight cost of taxation necessary to support it. If the number of periods increases, the cost remains the same, but the benefits increase. In the limit, in an infinite horizon model, it may be possible to support outside money without any cost. Hence, outside money may be better able to reduce and possibly eliminate the inefficiency we have studied. We will consider this possibility in future work.

References

Allen, Franklin and Douglas Gale, "Optimal Financial Crises," *Journal of Finance*, 53 (1998), 1245-84.

Allen, Franklin and Douglas Gale, *Understanding Financial Crises*. (Oxford University Press, 2007).

Bhattacharya, Sudipto., and Douglas Gale, "Preference Shocks, Liquidity and Central Bank Policy," in W. Barnett and K. Singleton, eds., *New Approaches to Monetary Economics*. (New York: Cambridge University Press, 1987, 69-88).

Brunnermaier, Markus, and Yuliy Sannikov, "*A Macroeconomic Model with a Financial Sector*," Princeton University Working Paper, 2011.

Cochrane, John, "*A Frictionless Model of U.S. Inflation*," NBER Macroeconomics Annual, (April, 1998).

Dewatripont, Mathias, and Jean Tirole, *The Prudential Regulation of Banks*, (MIT Press, 1994).

Diamond, Douglas W., "Financial Intermediation and Delegated Monitoring," *Review of Economic Studies*, LI (1984), 393-414.

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

Diamond, Douglas W. and Raghuram G. Rajan, "Liquidity Risk, Liquidity Creation, and Financial Fragility: A Theory of Banking," *Journal of Political Economy*, 109 (2001), 287-327.

Diamond, Douglas W. and Raghuram G. Rajan, "Money in a Theory of Banking", *American Economic Review*, 96 (2006), 30-56

Farhi, Emmanuel, and Jean Tirole, "Collective Moral Hazard, Maturity Mismatch, and Systemic Bailouts," *American Economic Review*, (2011), forthcoming.

Friedman, Milton, *The Optimum Quantity of Money*, (Macmillan, 1969).

Gennaioli, Nicola, Andrei Shleifer, and Robert Vishny, "Neglected Risks, Financial Innovation, and Financial Fragility," *Journal of Financial Economics*, (2010), forthcoming.

Gorton, Gary, and George Pennacchi, "Financial Intermediaries and Liquidity Creation," *Journal of Finance*, 45 (1990), 49-71.

Hahn, Frank, "On Some Problems of Proving the Existence of an Equilibrium in a

Monetary Economy," in F.H Hahn and F.P.R Breckling, *The Theory of Interest Rates,* (London: McMillan , 1965, 126-135).

Hart, Oliver and John Moore, "A Theory of Debt Based on the Inalienability of Human Capital," *Quarterly Journal of Economics*, 109 (1994), 841-879.

Holmstrom, Bengt, and Jean Tirole, "Financial Intermediation, Loanable Funds and the Real Sector," *Quarterly Journal of Economics*, CXII (1997), 663-691.

Holmstrom, Bengt, and Jean Tirole, "Private and Public Supply of Liquidity," *Journal of Political Economy*, 106 (1998), 1—40.

Holmstrom, Bengt, and Jean Tirole, *Inside and Outside Liquidity,* (MIT Press, 2011).

Kashyap, Anil K and Jeremy C. Stein, "Cyclical implications of the Basel II capital Standards," Federal Reserve Bank of Chicago Economic Perspectives 1Q (2004).

Kiyotaki, Nobuhiro, and John Moore, "Credit Chains," unpublished manuscript (1997).

Kiyotaki, Nobuhiro, and John Moore, "Evil is the Root of All Money," *American Economic Review*, 92 (2002) 62-66.

Kiyotaki, Nobuhiro and Randall Wright, "On Money as a Medium of Exchange," *The Journal of Political Economy*, 97 (Aug. 1989), 927-954.

Madrigan, Virgiliu, and Thomas Philippon, "Household Leverage and the Recession," New York University, Working paper, 2011.

Mattesini, Fabrizio, Cyril Monnet, and Randall Wright, "Banking: A Mechanism Design Approach," Federal Reserve Bank of Philadelphia, Working Paper 09-26, 2009.

Ostroy, Joseph and Ross Starr, "The Transactions Role of Money," in Benjamin Friedman and Frank Hahn, eds., *Handbook of Monetary Economics,* Vol. I. (New York: North-Holland, 1990, 3-62).

Philippon, Thomas, "The Evolution of the US Financial Industry from 1860 to 2007: Theory and Evidence,'' New York University, Working Paper, 2008.

Rotemberg, Julio, "A Behavioral Model of Demand Deposit," Harvard Business School, Working Paper, 2010.

Sargent, Thomas and Neil Wallace, "A Model of Commodity Money," *Journal of Monetary Economics,* 12 (1983), *189-196.*

Shleifer, Andrei, and Robert W. Vishny, "Unstable Banking," *Journal of Financial Economics*, 97 (2010), 306 – 318, Stein, Jeremy, “Monetary Policy as Financial Stability Regulation,” Harvard University, Working Paper, 2011,

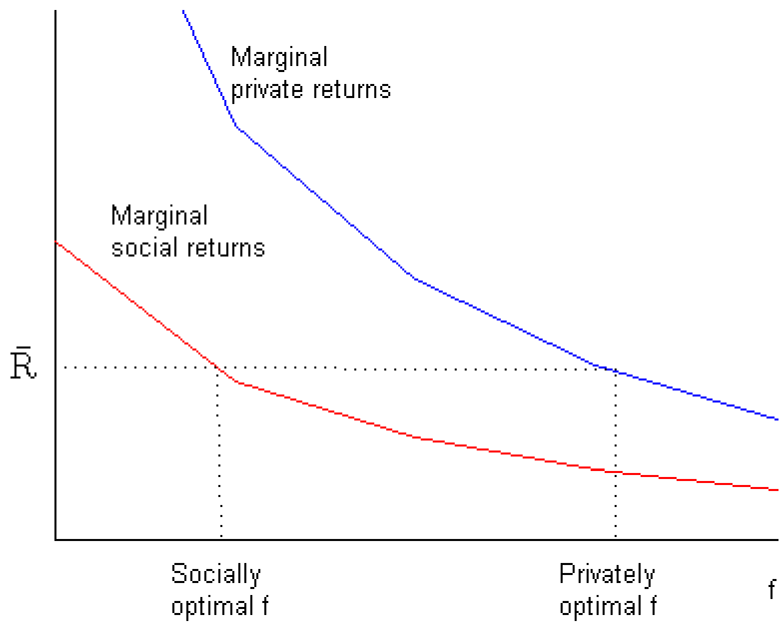

Figure I - Difference between private and social optimality