Money and the Size of Transactions

Joseph Zeira

The Hebrew University of Jerusalem and CEPR

February 2006

Abstract

Consumers make transactions of different sizes over time. This paper shows that this assumption, coupled with the assumption of costly trade in assets, can be helpful in modeling liquidity. In such a model assets with different cost structures are used to pay for different sizes of transactions. This can explain the demand for money itself, the precautionary demand for money, and the holding of both cash and demand deposits. Consumers use cash for small transactions, demand deposits for larger transactions, and savings for the largest transactions. Finally, the paper shows that modeling banks as suppliers of liquidity helps in understanding their dominance in financial intermediation.

JEL Classification: E40, E41, E51.
Keywords: Transactions, Liquidity, Demand for Money, Demand Deposits, Banks.

Contact:
Joseph Zeira
Department of Economics
The Hebrew University of Jerusalem
Mt. Scopus
Jerusalem 91905
Israel
Tel: 972 2 588 3256
E-mail: mszeira@mscc.huji.ac.il

Money and the Size of Transactions

1. Introduction

This paper claims that the theory of liquidity can benefit significantly from incorporating the observation that transactions come in different sizes. People purchase vegetables in the market on one day, purchase a concert ticket on another day, and buy a car in the following month. Each purchase involves a different size of payment. This simple fact of life, together with the fact that assets differ in their transaction costs, mean that each asset is used to purchase a different size-type of transactions. Also, payments and their sizes are not perfectly anticipated and their arrival is stochastic to a large extent. Hence, there are positive demands for assets with different degrees of liquidity. This can help us in understanding many monetary issues, such as the basic need for money, the precautionary demand for money, the portfolio choice between bonds and money, the emergence of commercial banks, and why banks specialize in financial intermediation.

The paper uses a very simple framework of analysis to demonstrate its main idea. It assumes that due to taste shocks people consume different amounts in each period of time. These amounts also differ from what these people produce and sell, namely from their income. That gives rise to a demand for money, as money is used to transfer income over time when other stores of value have higher transaction costs. Money gives us the flexibility to purchase different amounts in different periods of time and this flexibility is what we usually call 'liquidity.' Note that the type of money issued is less important here, as long as it is a durable and divisible asset.

Taste shocks are not known in advance. Hence, people leave aside some money in order to be able to perform large purchases in the future if needed. This is the precautionary demand for money. It is a result of future taste shocks and of risk aversion. Hence, this framework can formalize the precautionary demand for money and examine how it is related in general equilibrium to variables like the rate of inflation or to parameters like the distribution of transaction size.

The role of money in transferring income over time becomes less crucial if there are other stores of value like land, capital and bonds, which have higher rates of return. The paper next studies the demand for money in the presence of such an asset, if it has transaction costs, as in the well-known Baumol-Tobin analysis. Note, that transaction costs alone are insufficient to derive a positive demand for money. If transactions are of the same size, one asset is superior and the other is not held. Only under the assumption of different sizes of transactions money pays for small purchases, while large purchases are paid from savings. Since consumers anticipate both small and large transactions, they hold positive amounts of both assets, including money. Hence, this Baumol-Tobin result depends crucially on having different sizes of transactions. This paper therefore sheds new light on the Baumol-Tobin approach.[1]

Next, the paper discusses the coexistence of cash and demand deposits. It introduces commercial banks to the analysis as suppliers of an additional transaction technology, demand deposits, when holding cash is costly, due to possibility of theft or loss.[2] Deposits eliminate this cost, but create a transaction cost, which is either the cost of going to the bank, or the cost of processing a check. As a result people face two competing transaction technologies, with different cost structures. Since future

[1] One could say that the different size assumption is already hidden in the original Baumol-Tobin framework, as large income arrives once in a while, and small amounts are consumed daily.
[2] The idea that demand deposits emerge when holding cash is risky appears also in a recent paper by He, Huang and Wright (2004).

transactions come in many potential sizes, people hold both cash and deposits. They use cash for small transactions, and demand deposits for larger ones.

The paper then turns to describe lending by banks out of their deposits. This raises several issues, like financial intermediation, optimal size of reserves, risk of run on the bank, etc. The paper briefly shows that its framework can shed light on these issues. It focuses on understanding why commercial banks are leading financial intermediaries. This is because they obtain funds at a lower cost than other financial intermediaries, as they do not pay interest on deposits. Hence, banks are successful financial intermediaries not because of superior information, but rather because they have cheaper funds, as they pay their depositors with liquidity.

The analysis in this paper uses a series of models of overlapping generations with taste shocks. The models progress gradually from one to more assets. They are quite similar models, which differ only to highlight different issues. Overlapping generations models are used for tractability only. A model with infinitely lived consumers, where time periods are shorter, would fit better our main idea of variable sizes of transactions, but is harder to solve analytically. Appendix A outlines such a model.

This paper belongs to a vast literature that tries to explain the demand for money within a framework of utility maximization. This literature began with Sidrausky (1967), where money appears in the utility function, as it facilitates transactions and enables consumers to increase leisure. This approach is very popular, due to its tractability, but is vulnerable to the Lucas critique, as utility is implicit and hence changes with policy. A second line of research, “cash-in-advance”, began with Lucas (1980) and assumes that money has to be held ahead of time to pay for transactions. Romer (1986), Starr (2003) and others try to embed the Baumol (1952) and Tobin (1956) framework within a Ramsey model. Another approach has been followed by Wallace (1980) and Bewley (1980, 1983), who focus on money as a store of value that transfers income over time to accommodate for income and taste shocks. A recent approach is the search theory of money that formalizes the tradition of lack of 'double coincidence of wants.' This theory models people who produce and consume different types of goods, and thus need to search for trading partners. Money reduces such search costs. This theory began with Diamond (1984) and Kiyotaki and Wright (1989, 1991 and 1993).[3]

The specific contribution of this paper to the literature can be described as combining the approach of Wallace and Bewley, of money as a store of value, with that of Baumol and Tobin, which relates liquidity to transaction costs. The paper is also related to the search theory of money as it belongs to the tradition of money as a way to overcome lack of 'double coincidence of wants.' Only here, the mismatch is not of types of goods transacted, but of sizes, or values.

The idea of different sizes of transactions has it roots already in Patinkin (1956), who presents a model of demand for money where timing of transactions is uncertain. Lindbeck (1963) further develops the idea and states that "lack of perfect synchronization between payments and receipts" is important for explaining the transaction motive for money. He further stresses the importance of uncertainty, not only with respect to timing of transactions, but also to their values.[4] This idea is discussed also in Niehans (1978). This paper embeds these ideas in a general equilibrium framework. Two recent papers that have also distinguished between transactions by size are Prescott

[3] Recent examples of this line of research are Kocherlakota (1998), Wallace (2001), Lagos and Wright (2004), and Rocheteau and Wright (2003).

(1978) and Whitesell (1992), who focus on cash vs. demand deposits. But they assume that there are many transactions with different sizes within the same period of time. This paper assumes instead that each transaction holds in a separate period of time. Hence, it creates a connection between liquidity and store of value of money.[5]

The structure of the paper is as follows. Section 2 presents the first model, which describes the demand for money as a store of value. Section 3 formalizes the precautionary demand for money. Section 4 adds bonds with transaction costs and shows that the demand for money is still positive. Section 5 introduces bank deposits as a new transaction technology. Section 6 discusses bank loans and financial intermediation, while Section 7 summarizes. Appendix A describes a Ramsey version of the model and Appendix B presents proofs to the propositions.

2. Model I: The Demand for Money

This is the basic model in the paper, which shows how the role of money as a medium of exchange and its role as a store of value are closely related when transactions differ in size over time. This model shows how money can facilitate transactions and increase welfare so that it has a positive demand. This is shown by a simple model, where people wish to consume different quantities than what they produce and thus need a medium of exchange, which also stores value over time.

Consider an economy with one physical good, which is used for consumption only. The consumption good is not durable. The economy consists of overlapping generations, and each generation has a continuum of people of size 1. Each person lives 2

---

[4] See Lindbeck (1963), pages 21-23.

periods, works in both and produces a fixed amount y in each period. The individual consumes in both periods, but consumption can differ over time, due to inter-temporal preferences. The utility function of an individual is:

(I.1) $$U = \theta \log c_1 + (1-\theta)\log c_2,$$

where c_1 is consumption in first period of life, c_2 is consumption in second period and θ is a taste shock, which is realized in first period of life. It is assumed that taste shocks differ across individuals and are uniformly distributed on [0, 1] in each generation. Finally assume that lending and borrowing are impossible, as the cost of contracting is infinite. This assumption is relaxed from Model III on, when bonds with finite transaction costs are added to the analysis. But the assumption of transaction costs is crucial to this paper, as discussed in the introduction.

As there are no debt contracts and as the consumption good is not durable, each individual consumes exactly y in each period of life. Next we introduce an asset called money, which is both durable and divisible. For simplicity assume that this is fiat money. Although the analysis holds for commodity money, like silver or gold coins, as well, we deal with fiat money only, to simplify the analysis of inflation.[6] Denote the outstanding amount of money in period t by M_t and the money price of the physical good at time t by P_t. We assume that money is issued by the government, that increases it at a fixed rate μ, $\mu \geq 0$, namely:

(I.2) $$M_t = M_{t-1}(1+\mu).$$

---

[5] A more remotely related paper is Berentsen and Rocheteau (2003), which is a search theoretic model of money, where traders differ with how much they favor each good.

[6] For the analysis of inflation of commodity money through debasements, see Sussman and Zeira (2003).

The new money issued is used by the government to purchase the physical good, at an amount $M_{t-1}\mu / P_t$. Hence, fiscal expenditures are equal to the receipts of inflation tax.

To solve the model we first describe utility maximization of an individual born in period t with a taste shock θ. The individual sells output y, buys consumption of first period and keeps an amount of money $m_t(\theta)$ for next period. This amount is equal to

$$\text{(I.3)} \qquad m_t(\theta) = \arg\max_{m \geq 0} \left\{ \theta \log\left(y - \frac{m}{P_t} \right) + (1-\theta)\log\left(y + \frac{m}{P_{t+1}} \right) \right\}.$$

Note that P_{t+1} is the expected price in next period and due to rational expectations it is known in period t, since aggregates are deterministic in this economy. The constraint $m \geq 0$ is the liquidity constraint of the consumer in first period of life.

We distinguish between two cases in utility maximization (I.3). In the first case the individual has a low taste shock, low consumption in first period, and does not reach the liquidity constraint. In the second case the consumer has a high θ, wants to purchase a large amount and the liquidity constraint is binding. The two cases divide the population into two sets, according to taste shocks. The low taste shocks are in set A and the high taste shocks are in B. A person in A is described by point A in Figure 1. In this set the amount of real balances of money, denoted $l_t(\theta)$, is derived from first order condition:

$$\text{(I.4)} \qquad \frac{m_t(\theta)}{P_t} = l_t(\theta) = y\left(1 - \theta - \theta \frac{P_{t+1}}{P_t} \right) = y(1 - \theta - \theta \Pi_{t+1}).$$

We use the notation Π_t for the gross rate of inflation in period t. Equation (I.4) shows that the individual demand for real balances depends negatively on the expected rate of inflation. A person belongs to set A as long as the amount of money kept for second period is positive, namely as long as:

(I.5) $$\theta \le \theta_t \equiv \left(1+\Pi_{t+1}\right)^{-1}.$$

The set A is defined by equation (I.5) and it decreases as the rate of inflation rises. The reason is that inflation increases the incentive to consume in the present.

[Insert Figure 1 here]

Set B, where $\theta > \theta_t$, is shown by point B in Figure 1. In this case the liquidity constraint is binding, no money is left to second period and the individual consumes exactly y in each period. Clearly, the introduction of money increases utility for those in set A. Hence, this model can justify the introduction of money to the economy. Note that in this model money raises welfare only if it is durable, namely if it stores value.

We next derive the monetary equilibrium. To do it we sum up the demands for money by the young in set A. We consider only the amounts kept by the end of the period, as money used for transactions within a period only changes hands and is not counted as net demand for money. Hence, the aggregate demand for real balances is

(I.6) $$\frac{M_t^d}{P_t} = \int_0^1 l_t(\theta)d\theta = y\int_0^{\theta_t}\left(1-\theta-\theta\Pi_{t+1}\right)d\theta = \frac{y}{2\left(1+\Pi_{t+1}\right)}.$$

Hence, the demand for real balances depends negatively on the expected rate of inflation. The demand for liquidity is shown in Figure 2.

[Insert Figure 2 here]

To complete the derivation of the monetary equilibrium we add the monetary dynamics to Figure 2, through the horizontal curve at 1 + μ. Clearly, the intersection of the two curves is the rational expectations stable equilibrium. The rate of inflation is therefore equal to the rate of monetary expansion, namely $\Pi_t = 1+\mu$ for all t, and the equilibrium price level in period t is:

(I.7) $$P_t = \frac{2(2+\mu)}{y} M_t .$$

Hence, money in this economy has a positive value, even though it does not affect utility directly. It increases utility indirectly by supplying liquidity.

It is interesting to analyze a variant of this benchmark model, where the taste shocks have different volatility. Consider the same model except that taste shocks are uniformly distributed on [½ - σ, ½ + σ] in each generation, where $0 \leq \sigma \leq$ ½. It can be shown that in this case the demand for real balances is equal to:

(I.8) $$\frac{M_t^d}{P_t} = \frac{y}{4\sigma}\left[\frac{1}{2}+\sigma-\left(\frac{1}{2}-\sigma\right)\Pi_{t+1}\right]\left[\left(1+\Pi_{t+1}\right)^{-1}-\frac{1}{2}+\sigma\right].$$

Clearly, the demand for real balances depends negatively on the expected rate of inflation as in the benchmark model. Note that in this case inflation is bounded, as demand for money is zero for inflation above (½+σ)/(½-σ). Another implication of (I.8) is that the demand for money depends negatively on shocks variability σ. This result stresses the role of the assumption of different sizes of transactions in this model of money.

This section therefore shows how money is useful when people have different sizes of transactions over time. They use money to transfer income from one period to another if consumption in first period is low enough. For money to function this way it must be a store of value. Hence, the roles of money as a means of transaction and as a store of value are closely related when transactions differ in size over time.

Finally, the model can be applied to the issue of optimal monetary policy. We look for the rate of inflation that maximizes average utility across individuals in each generation. To do it we need to reformulate monetary policy. In the original model money finances public consumption, but as we do not want to confuse the issues of optimal monetary policy and fiscal policy we assume instead that money printed by the government is distributed equally as a subsidy to all individuals at all ages. This is therefore a version of the famous 'helicopter drop' example. If the rate of monetary expansion is μ, the size of subsidy for each person in each period is:

$$s_t = \frac{1}{2}\frac{\mu M_{t-1}}{P_t} = \frac{1}{2}\frac{\mu}{\Pi_t}\frac{M_{t-1}}{P_{t-1}}. \qquad \text{(I.9)}$$

Note that inflation subsidizes all, but taxes only money holders, namely those with low θ. Hence, those who do not hold money receive a net subsidy, while those who hold much money are net-taxed. Proposition 1 assesses the overall welfare effect of inflation.

Proposition 1: The steady state rate of inflation which maximizes average utility across individuals is zero.

Proof: In Appendix B.

Hence the optimal rate of inflation in this simple model of money is zero.[7] Note though that this policy is optimal on average only, as agents differ significantly: some gain from this policy, and some lose, depending on how much money they hold.

3. Model II: The Precautionary Demand for Money

According to our basic approach people hold money for transactions in which the amount they wish to purchase is larger than the amount they earn as income. Hence, money enables them to meet an opportunity for a large purchase. The more money they have

[7] Note that the average subjective discount rate is equal to zero, so that optimal monetary policy here fits the famous Friedman rule, despite the difference in modeling.

available, the greater the chance that they will be able to purchase without being liquidity constrained. This leads to a precautionary demand for money. One of the reasons people hold money is to be able to make large unanticipated purchases. Our framework of analysis enables us to formalize this concept. Although it is usually impossible to separate between the precautionary demand and other motives for holding money, we present here a simple extension of our model that is capable of that.

Consider a closed economy with only one physical good, which is used for consumption and is perishable. Population consists of overlapping generations, each a continuum of size 1. Every person lives 3 periods, produces in each period an amount y and consumes in the three periods according to the following utility function:

(II.1) $$U = \frac{1}{2}\log c_1 + \theta \log c_2 + (1-\theta)\log c_3.$$

The taste parameter θ differs across people and is independent and uniformly distributed on [0, 1]. Furthermore, the taste parameter is stochastic, unknown to the individual in first period of life and revealed only in second period. We assume debt contracts are infinitely costly and the only asset in the economy is money. Its quantity is M_t and the nominal price of the physical good is P_t. Similar to model I, the rate of monetary expansion is fixed and equal to μ.

A person born in t earns y in first period of life, consumes $c_{1,t}$ and keeps savings in money, with real value $l_{1,t} = y - c_{1,t}$. In second period of life she consumes her income plus her money savings from last period, minus $l_{2,t+1}$, the real amount of money left in the end of the second period. In third period of life she consumes income plus money saved in second period. The individual realizes her taste shock θ in the second period of life and then she maximizes utility for the second and third periods:

(II.2) $$l_{2,t+1}(\theta) = \arg\max_{l \geq 0} \{\theta \log(y + l_1 / \Pi_{t+1} - l) + (1-\theta)\log(y + l / \Pi_{t+2})\}.$$

This maximization can lead to two possible outcomes, which divide the taste shocks into two sets, as in Model I. In set A the consumer consumes a small amount in second period of life and has sufficient liquidity. In set B the consumer wants to purchase more and is liquidity constrained. The two cases can be described in a diagram similar to Figure 1. In set A the demand for money is determined by the FOC and is equal to:

(II.3) $$l_{2,t+1}(\theta) = y(1 - \theta - \theta\Pi_{t+2}) + (1-\theta)\frac{l_{1,t}}{\Pi_{t+1}}.$$

Hence lifetime utility in set A is:

(II.4) $$U_{A,t}(\theta) = \frac{1}{2}\log(y - l_{1,t}) + N(\theta) - (1-\theta)\log\Pi_{t+2} + \log\left[y(1 + \Pi_{t+2}) + \frac{l_{1,t}}{\Pi_{t+1}}\right].$$

The function N denotes: $N(\theta) = \theta\log\theta + (1-\theta)\log(1-\theta)$. Set A includes the small taste shocks, for which the demand for money in second period is positive. Hence, the borderline between sets A and B is:

(II.5) $$\theta_{t+1} \equiv \frac{y\Pi_{t+1} + l_{1,t}}{y\Pi_{t+1} + l_{1,t} + y\Pi_{t+1}\Pi_{t+2}}.$$

In set B, where $\theta > \theta_{t+1}$, the consumer faces a liquidity constraint, leaves no money for third period and consumption in second and third periods are $y + l_{1,t} / \Pi_{t+1}$ and y respectively. In this set lifetime utility is

(II.6) $$U_B(\theta) = \frac{1}{2}\log(y - l_{1,t}) + \theta\log\left(y + \frac{l_{1,t}}{\Pi_{t+1}}\right) + (1-\theta)\log y.$$

Since in first period of life the individual does not know her future taste shock θ, the demand for money in this period is determined by maximizing expected utility:

(II.7) $$EU_t = \int_0^{\theta_{t+1}} U_{A,t}(\theta)d\theta + \int_{\theta_{t+1}}^1 U_{B,t}(\theta)d\theta.$$

The following proposition describes the results of this expected utility maximization.

<u>Proposition 2:</u> There is a unique optimal amount of $l_{1,t}$ that maximizes the expected utility (II.7). This amount of money is proportional to income: $l_{1,t} = yh(\Pi_{t+1}, \Pi_{t+2})$, where h depends negatively on the expected rates of inflation in t+1 and t+2. Furthermore, if inflation is not too high, $l_{1,t}$ is positive. Only at high inflation it falls to zero.

Proof: In Appendix B.

Hence, individuals in this economy are willing to sacrifice consumption in first period of life in order to reduce the risk of being liquidity constrained in second period of life. We therefore view $l_{1,t}$ as the precautionary demand for money. If the rate of inflation is high, the losses from holding money outweigh the benefits from reducing the risk of liquidity constraint and the precautionary demand for money falls to zero.

We next describe the equilibrium. The demand for real balances is the sum of the demand by grownups and the demand by the young:

(II.8) $$\frac{M_t^d}{P_t} = l_{1,t} + \int_0^{\theta_t} l_{2,t}(\theta)d\theta = \\ = y\left\{h(\Pi_{t+1}, \Pi_{t+2}) + \int_0^{\theta_t}\left[1 - \theta - \theta\Pi_{t+1} + (1-\theta)\frac{h(\Pi_t, \Pi_{t+1})}{\Pi_t}\right]d\theta\right\}.$$

The following proposition analyzes the equilibrium and the relation of real balances to inflation.

Proposition 3: The equilibrium rate of inflation depends on the previous period real balances. At the steady state the rate of inflation is equal to the rate of monetary expansion μ. The demand for money falls with inflation.

Proof: See Appendix B.

In order to justify our interpretation of $l_{1,t}$ as the precautionary demand for money, consider the case of no inflation, namely $\mu = 0$, and assume that people have access to a better asset than money: it not only stores value, but also enables them to borrow. In other words, they have a bond with the same interest rate as money, 0, but that can be used both to save for next period and to borrow in order to purchase in the present. In this case liquidity of grownups is never constrained. It can be shown that in this case expected utility is maximized at $l_{1,t} = 0$. Hence, the amount of money held by the young is purely a precautionary demand for money.

4. Model III: Money and Bonds

So far money has been the only asset in the economy. This section shows that it is used to transfer income over time even if there are additional assets, as long as there are transaction costs. The asset added in this section can be described either as bonds (indexed) or as physical capital. For the sake of simplicity assume that this asset has a constant positive real rate of return r. This asset has a higher return than money, but it is less liquid, since it is costly to invest in this asset and to withdraw the investment for use. More specifically, assume that finding an investor and signing a lending contract has utility cost of size x, and withdrawing the loan after a period or two has a utility cost as well, of size z. While we do not impose any restriction on z, except that it is positive, we assume that the cost of contracting x is sufficiently high, to simplify the analysis:

(III.1) $$x > \log(1+r).$$

The critical assumption in this analysis, that $z > 0$, follows the Baumol-Tobin tradition, of a costly exchange between bonds and money. Below we discuss the difference between this model and previous models in this tradition.

The rest of the model is similar to the previous models in the paper. Each person lives three periods, produces an amount y in the first period of life only, and consumes in the second and third periods of life.[8] The utility from consumption is

(III.2) $$U = \theta \log c_2 + (1-\theta)\log c_3.$$

The taste shock θ is revealed only in the second period of life. We assume for simplicity that taste shocks are uniformly distributed on [0, 1]. As in the previous models we assume that the government increases the amount of money at a constant rate μ. To simplify notation this section analyzes the steady state of the economy only. In the steady state the rate of inflation is constant and equal to the rate of monetary expansion, so:

(III.3) $$\Pi = 1+\mu.$$

We first analyze the representative consumer, who decides not only on consumption levels but also on the composition of her portfolio. In first period the individual saves all income and divides it between money and lending:

$$y = \frac{m_1}{P} + b_1 = l_1 + b_1,$$

where l is real balances and b is real amount of bonds.

[8] This assumption is made for simplicity only. Since the main focus of this model is not saving, but portfolio decisions, the main results of the model are not affected by it.

In the second period of life there can be three cases, depending on the taste shock θ. In case A, the transaction size in second period is smaller than money, so that some money is left after the transaction. It is not invested in bonds due to assumption (III.1), so it is kept as money for third period of life.[9] In case B, the size of transaction is larger, so all money is used, but no bonds are used, to avoid the cost z. In case C, the transaction is even larger, so it is paid by all the money and by some savings. This involves going to the investor and retrieving some of the loan back.

In case A the consumer leaves l_2 real balances of money and keeps it to third period of life. This amount is determined by the following utility maximization:

$$\max_{l_2}\left\{\theta\log\left(\frac{l_1}{\Pi}-l_2\right)+(1-\theta)\log\left[b_1(1+r)^2+\frac{l_2}{\Pi}\right]-x-z\right\}.$$

The optimal amount of money left is therefore:

(III.4)
$$\begin{aligned} l_2(\theta) &= \Pi^{-1}\left[(1-\theta)l_1-\theta(1+r)^2\Pi^2 b_1\right]= \\ &= \Pi^{-1}\left[(1-\theta)l_1-\theta(1+r)^2\Pi^2+\theta(1+r)^2\Pi^2 l_1\right]\end{aligned}$$

Calculating the optimal levels of consumption we get that utility in A is equal to:

(III.5)
$$U_A(\theta)=N(\theta)-(2-\theta)\log\Pi+\log\left[l_1-l_1(1+r)^2\Pi^2+y(1+r)^2\Pi^2\right]-x-z.$$

The set A includes all individuals with non-negative l_2, namely if the taste shock is smaller than the threshold θ_1:

(III.6)
$$\theta_1=\frac{l_1}{y(1+r)^2\Pi^2-[(1+r)^2\Pi^2-1]l_1}.$$

In set B the consumer is mildly liquidity constrained. She uses all her money, but is not sufficiently constrained to call off a loan. In this case consumption levels are

---

[9] To show this consider the case where it pays most to invest the money in bonds, namely $\theta = 0$. If a person puts all money in bonds, utility is: $\log[b_1(1+r)^2+l_1(1+r)]-2x-z$. If not utility is: $\log[b_1(1+r)^2+l_1]-x-z$. Due to (III.1) the individual prefers to keep money and not invest in bonds.

$c_2 = l_1 / \Pi$ and $c_3 = b_1(1+r)^2$ in the second and third periods respectively. Optimal utility is

(III.7) $$U_B(\theta) = 2(1-\theta)\log(1+r) - \theta\log\Pi + \theta\log l_1 + (1-\theta)\log(y-l_1) - x - z.$$

In set C the taste shock θ is larger and so is consumption. The liquidity constraint is so binding that the consumer goes to the investors and retrieves part of the loan, of size *d*. The consumer maximizes

$$\theta\log\left(\frac{l_1}{\Pi} + d\right) + (1-\theta)\log\{[b_1(1+r) - d](1+r)\} - x - 2z.$$

The amount of loan called off is therefore equal to:

$$d(\theta) = \theta y(1+r) - l_1\left[\theta(1+r) + \frac{1-\theta}{\Pi}\right].$$

Hence, utility in this case is

(III.8) $$U_C(\theta) = N(\theta) + (1-\theta)\log(1+r) - \log\Pi + \log[y(1+r)\Pi - l_1(1+r)\Pi + l_1] - x - 2z.$$

The borderline between sets B and C is determined by indifference between the two states, namely by the condition: $U_B(\theta_2) = U_C(\theta_2)$. This condition boils down to

(III.9) $$\begin{aligned} &N(\theta_2) - (1-\theta_2)\log(1+r)\Pi + \log\left[y(1+r)\Pi - l_1((1+r)\Pi - 1)\right] - \\ &\quad - \theta_2\log l_1 - (1-\theta_2)\log(y-l_1) = z. \end{aligned}$$

This threshold level θ_2 is binding if it is smaller than 1. If U_B is everywhere higher than U_C then θ_2 is equal to 1. It is easy to see that this does not happen if l_1 is low.

We can next calculate the expected utility of the individual in first period of life, which is when portfolio decisions are made:

(III.10) $$EU = \int_0^{\theta_1} U_A(\theta)d\theta + \int_{\theta_1}^{\theta_2} U_B(\theta)d\theta + \int_{\theta_2}^{1} U_C(\theta)d\theta.$$

The individual determines the amount of money by equating the marginal utility of money to zero or by equating the cost and the benefits of holding money. Increasing l_1 is costly as it reduces future income from bonds. But it is also beneficial, as it reduces the probability of selling bonds and thus reduces the probability of suffering the utility cost z. In other words, increasing the amount of money reduces the probability of hitting the liquidity constraint. Proposition 3 shows that the individual always prefers to hold a positive amount of money, despite its lower rate of return.

Proposition 4: The amount of money held by the young is positive at any rate of inflation. It is proportional to income: $l_1 = yh$, it is negatively related to the rate of inflation, and it is positively related to z. If the transaction cost x is sufficiently high, the consumer holds money only and no bonds, so that: $l_1 = y$.

Proof: In Appendix B.

In this specific model the overall demand for money is by the young and by the grownups in set A, who do not use all their money in second period of life. Hence the equilibrium in the money market is determined by the following equilibrium condition:

(III.11)
$$\frac{M}{P} = l_1 + \int_0^{\theta_1} l_2(\theta)d\theta = yh\left\{1 + \frac{1}{2\Pi}\frac{h}{(1+r)^2\Pi^2 - h[(1+r)^2\Pi^2 - 1]}\right\}.$$

This condition determines the price level in this economy. It can be shown that the right hand side, which is the demand for real balances, depends negatively on the nominal interest rate $(1+r)\Pi$.

In this paper money is used to transfer income over time. In model III there is another asset – bonds – that can do the same job. This asset has a higher rate of return, but it also has transaction costs. In other words, the two assets have different cost structures. If all transactions were of the same size, then they would all be performed either by money or by bonds, and the two assets would not co-exist. Hence the assumption that transactions come in different sizes is crucial, as some transactions are better financed by money and some by savings. Thus, there are positive demands for the two assets, money and bonds. Hence, this model follows the Baumol-Tobin tradition, but clarifies that the Baumol-Tobin assumption works only if transactions come in different sizes. Interestingly in the original Baumol-Tobin model all purchases are the same, but they are not equal to income. Hence this model makes this assumption explicit and stresses its importance. This model also clarifies how money and savings are used for different types of purchases, where money pays for smaller transactions, while larger transactions are paid from saving accounts.

5. Model IV: Commercial Banks – Cash and Demand Deposits

This section introduces commercial banks to our framework. Since our approach to money focuses on transactions, we introduce banks as an innovation in transactions technology. Banks offer an alternative means of payment to cash. In addition to holding cash people can deposit money in the bank. Such deposits have a clear benefit of reducing the risk involved with carrying large amounts of cash, due to loss or theft. The new asset, demand deposits, is safer, but has costly transactions. The way demand deposits are used by consumers, whether by going to the bank to draw cash for each purchase or by using checks, is less important for our analysis. What is important is that in any way there is a cost involved in each transaction. In comparison, cash transactions are not costly, but holding cash is costly due to risk of loss or theft. These are therefore two transaction technologies, or two assets, which have different cost structures, as do money and bonds. In a similar way it is shown that both forms of money – cash and demand deposits – are held. One is used for small size transactions and the other for large ones. Hence, banks offer an alternative transaction technology, but since this technology is not completely costless and has a different cost structure, banks do not completely crowd out the use of cash, and the two forms of money coexist.

To demonstrate this idea, consider a similar economy to those in previous models. The economy has overlapping generations, each of size 1, and each person lives three periods, produces y in the first period of life, and consumes in the second and third periods of life where utility is

(IV.1) $$U = \theta \log c_2 + (1-\theta)\log c_3.$$

The taste shock θ is random, is revealed only in the second period of life, and is uniformly distributed on $[0, \tau]$, where $\tau < 1$. This upper bound is assumed for tractability of the solution only. For further simplicity assume that there are no bonds but money only. People can hold cash, which faces risk of theft or loss, or they can deposit money in banks, which are safe. Going to the bank to deposit or withdraw money has utility cost of size v.[10] In order to model the risk of holding cash as simple as possible assume that too much cash gets stolen. Namely, if the amount of cash held by a person exceeds a real value E, it is stolen. Since it is reasonable that this threshold of theft is relative to average

[10] Cost of check payment is usually in income and not utility, but the results of the model are the same.

income, we assume that it is equal to $E = ey$, where $e > 0$. Hence, e measures danger of theft in the economy.[11] We further assume that $e < 1 - \tau$. As for monetary policy, we assume as in former models that the government increases the amount of money at a fixed rate μ. For simplicity we analyze the steady state only, where the rate of inflation is constant over time. Hence, it is equal to the rate of monetary expansion: $\Pi = 1 + \mu$.

A person earns income y in first period of life and divides it to two forms of money. A real amount f is held in cash and a real amount d is deposited in the bank, so that $f + d = y$. Demand deposits pay no interest. In the second period of life the person realizes her θ and four cases can occur. In case A, a small amount is purchased with some of the cash, and some cash, $f / \Pi - c_2$, is left for third period. In case B the transaction is larger and all the cash is used in period 2, but no bank deposit is used to avoid the cost of going to the bank. In case C the consumer wishes to consume more in period 2 and both cash and some demand deposit are used, but some of the demand deposit is left in the bank. This is done because the anticipated purchase in period 3 is too large to be paid by cash only. In case D all the money in the bank is withdrawn, some is used for purchase and some remains as cash. Since $1 - \theta$ is larger than e the consumer is constrained by the threat of theft and keeps only ey as cash. Hence, small transactions are paid by cash and large transactions by demand deposits. The following proposition describes some characteristics of the equilibrium.

Proposition 5: The amount of cash held by each young consumer is $e\,y$. Both demands for cash and for demand deposits are proportional to income y. The demand for cash depends

[11] An alternative modeling strategy is to assume that cash depreciates at some rate. In this case the model is

positively on the average size of taste shocks τ, but demand deposits and the overall amount of money depend negatively on it.

Proof: In Appendix B.

Analyzing the effect of inflation on the demands for cash and for deposits is more cumbersome. To simplify the calculation we assume that the cost of going to the bank v is sufficiently high so that case C disappears, as the consumer prefers to keep money in cash for third period rather than go once more to the bank. Hence, θ_2 is determined by the intersection of U_B and U_D:

(IV.2) $$\theta_2 \log e + (1-\theta_2)\log(1-e) = \theta_2 \log(1-e\Pi) + (1-\theta_2)\log(e\Pi).$$

Note that if there is no inflation and $\Pi = 1$, this θ_2 is equal to ½. If the rate of inflation rises, θ_2 decreases. Based on the proof to proposition 4 the aggregate demand for cash is:

(IV.3) $$F = ye\left(2 + \frac{e}{2\tau\Pi} - \frac{\theta_2}{\tau}\right).$$

Hence, the effect of inflation on the demand for cash is ambiguous. The demand for bank deposits is:

(IV.4) $$D = y(1-e)\left(1 + \frac{\theta_2}{\tau\Pi}\right).$$

Hence, the demand for bank deposits depends negatively on the rate of inflation. Clearly inflation affects demand deposits by more than it affects cash holdings.

The overall demand for money in this case is:

(IV.5) $$L = y\left[1 + d + \frac{e^2}{2\tau\Pi} + \frac{\theta_2}{\tau}\frac{1-e-e\Pi}{\Pi}\right].$$

---

fully equivalent to model III of money and bonds.

Hence the demand for money is falling with the rate of inflation.

Note that in this section the outstanding amount of money must be equal to the total amount of cash and demand deposits since banks do not lend yet and as a result they hold a reserve ratio of 100%. In the next section we add lending to the model and examine its effects on money and on financial intermediation.

6. Model V: Banks and Financial Intermediation

The situation described in model IV, where banks have large stocks of cash in their safes, tempts banks to lend some of the money in order to earn interest on it. This leads us to the issues of bank lending, optimal reserves, and how these are related to the monetary equilibrium. This section presents a preliminary analysis of these issues, using a model similar to the previous one. The main additions to model V are that we bring back bonds or physical capital to the model, to enable bank lending, and we also introduce cost to banking activity. Let us assume that there is a class of investors, who have projects that require investment of 1 in one period. There are good and bad investors. The return on a project run by a good investor in next period is $1 + R$. A bad project yields 0. It is possible to perfectly monitor and screen investors before lending to them, but this monitoring requires work of q workers per project selected. This gives rise to financial intermediaries as delegated monitors as in Diamond (1984), and the resulting net rate of return for consumers who lend to these financial intermediaries is: $r = R - y\,q$.

Consumers come in overlapping generations, each of size 1. Each person lives three periods, earns labor income y in the first period of life, and consumes in the second and third periods of life where utility is

(V.1) $$U = \theta \log c_2 + (1-\theta)\log c_3.$$

The taste shock θ is random, it is revealed only in the second period of life, and θ is uniformly distributed on [0, τ], where $\tau < 1$. There is a utility cost of going to the financial intermediary and create a loan contract or withdraw resources from the contract. This utility cost is x.[12] In addition to loans or bonds people can hold cash, which faces risk of theft or loss, or they can deposit money in banks, which are safe. Going to the bank to deposit or withdraw money has utility cost of size v. We assume realistically that $v < x$. Similar to model IV assume that if the amount of cash held is larger than $e\,y$, it is stolen.

Banks require a fixed investment and labor to run transactions and deposits. In order to form a bank an amount s is invested and n workers service the following amount of demand deposits:

(V.2) $$d = an^{\alpha}.$$

As for monetary policy, assume for the sake of simplicity that the supply of cash is fixed at M, so that only steady state equilibrium with fixed prices is analyzed.

We can now sketch the equilibrium of this economy. In first period of life consumers build a portfolio of three assets: cash, deposits, and savings, in the amounts f_1, d_1, and b_1 respectively. In second period of life they realize θ, consume, and keep cash, deposits and bonds in the amounts $f_2(\theta), d_2(\theta)$ and $b_2(\theta)$ respectively. Summing up over all individuals we get the overall deposits in banks in each period: $D = D_1 + D_2 = D(r)$, and the overall withdrawals from banks, which are equal. Hence, if banks are fully diversified in the population, so that the distribution of θ among their clients is the same as in the population, they can lend all their deposits and keep zero reserves. What is clear from this description is that banks can lend funds to investors and monitor them, just like other financial intermediaries, but the funds they use are less expensive, as they don't have to pay the interest rate of lenders. They pay zero interest to depositors and only supply them with liquidity. Hence, banks have an advantage in competition with other financial intermediaries.

To illustrate it we examine more closely banks' decisions. A bank lends its deposits d and its net profit is:

(V.3) $$(1+r)d - yn - (1+r)s = (1+r)an^{\alpha} - yn - (1+r)s.$$

Profit maximization in competition leads to the following amount of deposits in a bank:

(V.4) $$d = a^{\frac{1}{1-\alpha}}\left[\alpha(1+r)\right]^{\frac{\alpha}{1-\alpha}} y^{-\frac{\alpha}{1-\alpha}}.$$

The optimal profit of the bank is:

(V.5) $$(1+r)[(1-\alpha)d - s] = (1+r)\left[(1-\alpha)a^{\frac{1}{1-\alpha}}[\alpha(1+r)]^{\frac{\alpha}{1-\alpha}} y^{-\frac{\alpha}{1-\alpha}} - s\right].$$

Assume that these profits are positive. Hence banks have an advantage over financial intermediaries, who operate on zero profits. As a result banks lend the whole all deposits D, while other financial intermediaries lend only the remaining demand: $B - D$.

But banks can have an incentive and ability to increase their share in financial intermediation even by more, if financial intermediation has some increasing returns to scale. To demonstrate it assume that the required amount of labor for monitoring a project is q if the intermediary monitors only few projects in the sector the project belongs to, but it is equal to $q - \varepsilon$ if it monitors more than k projects in the sector. Hence, a bank has an interest to lend more in order to reduce costs. For that the bank enables

[12] Assuming that the costs of creating a contract and liquidizing it are different, as in Model III, leads to the

customers to open saving accounts. It therefore enters a competition with other financial intermediaries on individuals who wish to lend money at an interest rate r. The bank can win this competition since it operates on higher profits on average, so it can offer borrowers a slightly lower interest rate. Hence, banks end up lending much more than their demand deposits D and they take control over a large share of the market for financial intermediation.

To fully close the model we need to analyze how the market reacts to banks' excess profits. This depends to a large extent on the type of competition in the market for commercial banking. One possibility is that these profits create incentive for more banks to enter. Once the number of banks exceeds D/d, namely they need more deposits than what the public supplies, they turn to compete over the limited amount of deposits. The equilibrium in this case is determined by Cournot competition and by the zero profit entry condition. Another possibility is that banks run a Bertrand competition on prices. Namely, banks offer an interest rate on demand deposits. Again the equilibrium involves some zero profit condition. Note that this policy requires some degree of technical progress in banking, since payment of interest on demand deposits became possible only with computerization of banks.

Next consider a more realistic extension of the model, when banks hold positive reserves. For that add some uncertainty to money holding. This can be modeled within our framework by assuming that the distribution of taste shocks has a random aggregate component. Formally, assume that τ_t is random and changes from one period to the other. Hence banks need to keep some reserves in order to avoid the risk of illiquidity.

same results.

We do not fully solve this case here, but the intuition at least seems to be quite straightforward.

Focusing on liquidity shocks and reserves can highlight another issue in monetary economics, which is how central banks affect real activity. Note that central banks lend to banks who are hit by liquidity shocks τ_t. Its interest rate is lower than the market interest rate. Hence, it can be viewed as a subsidy to financial intermediation by banks. By changing the interest rate of lending to banks the central bank affects the cost of funds to banks and thus affects the interest rate they charge for their loans. Thus such a policy has real effects and not only nominal effects. The intuitive reason is that such a change in the interest rate by the central bank affects not only the stock of money, but the flow of income from the central bank to the economy, to commercial banks. This approach therefore views monetary policy as part of fiscal policy, where the central bank directly subsidizes the banking sector and through it investors. Note, that this view on the effect of monetary policy does not require any assumption of price rigidity. Monetary policy affects the economy as it directs real resources to some sectors in the economy.

7. Summary

This paper suggests that a useful way to understand the issue of liquidity is to combine the idea of transaction costs with the idea that transactions come in different and random sizes. Assets that differ in their transaction costs can be held in the portfolio only if transactions have different sizes. Otherwise only one asset is dominant. Thus the combination of the two assumptions can serve for a foundation to the theory of liquidity and therefore to the theory of money. In this framework people pay for small transactions with cash, for larger transactions with credit cards and demand deposits and for even larger transactions they withdraw money from their saving accounts.

This paper shows how this idea can be used to understand a number of issues in monetary economics, like precautionary demand for money, cash and deposits, financial intermediation by banks, etc. But this is just a preliminary examination of this idea, and it can be further explored in a number of ways. One way is to deepen the analysis of banks and their liquidity. Another direction of research is to fully introduce a central bank and discuss monetary policy beyond the discussion in Section 6. An interesting direction of research is to examine the optimal monetary policy in more complicated models than the one in Section 2. Another way to proceed is to apply the more realistic Ramsey model from Appendix A and solve it, at least numerically. Calibration methods can help to estimate transaction costs of various assets. Although the paper analyzes the emergence and use of money in a rather historical order: cash, demand deposits, etc., it reflects a strong belief that it has much to say on money in our times as well.

Appendix A: The Infinite Horizon Model

Consider an economy with a single physical good. Time is discrete. There is a continuous mass of size 1 of infinite horizon individuals in the economy. Output is produced by labor only and there is no capital. Each person produces a constant amount y in each period of time. Individuals derive utility from consumption, but they have taste shocks in each period, that affect their utility. The utility from consumption in period t of individual j is:

(A.1) $$E_t \sum_{s=0}^{\infty} \beta^s u[c_{t+s}(j), \theta_{t+s}(j)].$$

The utility function is concave in consumption and satisfies the Inada conditions. The taste shocks θ are independent and identically distributed over time and across individuals. Each shock is revealed in its time, but future taste shocks are unknown in advance.

If there is no money, consumers cannot shift income over time to meet their different taste shocks and must consume y in each period. The utility of consumer j in period t is therefore equal to:

(A.2) $$E_t \sum_{s=0}^{\infty} \beta^s u[y, \theta_{t+s}(j)].$$

Next assume that there is money in the economy. Assume that money is durable, coins or fiat money. Hence, it can be used to transfer income over time. Assume that the overall amount of money in the economy in period t is M_t. Let us denote the amount of money held by individual j by the end of period t by $m_t(j)$. Then the budget constraint in period t is described by:

(A.3) $$c_t(j) = \frac{m_{t-1}(j)}{P_t} + y - \frac{m_t(j)}{P_t}.$$

The consumer maximizes utility (A.1) given the budget constraints (A.3) and the money non-negativity constraints:

(A.4) $$m_t(j) \geq 0.$$

Clearly, optimal utility with money is higher than utility without money (A.2). Hence, money enables consumers to adjust better to taste shocks, namely to make transactions of different sizes over time.

We next turn to describe equilibrium in this economy. The equilibrium prices are the sequence $(P_t, P_{t+1},...)$ of nominal prices of the physical good in terms of money. As shown below these prices are fully known in advance, namely they are deterministic. The consumers react to these prices as they determine the rates of return of money. Hence, the individual maximizes (A.1), given the budget constraints (A.3) and (A.4) and taking the price levels as given.

It can be shown that the optimal real amount of money held by end of period *t* is a function *l* of initial money, current taste shock, and the expected rates of inflation:

(A.5) $$\frac{m_t(j)}{P_t} = l\left[\frac{m_{t-1}(j)}{P_t}, \theta_t(j), \pi_t\right],$$

where π_t is the vector of all future rates of inflation:

(A.6) $$\pi_t = \left(\frac{P_t}{P_{t+1}}, \frac{P_{t+1}}{P_{t+2}}, ...\right).$$

Equation (A.5) describes the individual demand for money. Note that this demand for money is positive, if the constraint (A.4) is not binding, namely if the consumer does not consume much in period *t* and can leave money to future consumption. But if the taste shock in t raises marginal utility of consumption in t by much, the constraint (A.4) becomes binding and the consumer does not leave money for future periods.

Equation (A.5) also describes how the distribution of money across individuals evolves over time. Note that since the taste shock is independent across individuals it is also independent of the amount of money from period t-1. Hence, the distribution of money in period t is determined uniquely by the distribution in t-1, namely it evolves deterministically over time.

The aggregate demand and the aggregate supply of money determine the equilibrium in the market for money, which also determines the price of goods P_t:

$$\text{(A.7)} \qquad \frac{M_t}{P_t} = \int_0^1 l\left(\frac{m_{t-1}(j)}{P_t}, \theta(j), \pi_t\right) dj.$$

Since the taste shock is independent of the initial distribution of money from period t-1, it follows that the equilibrium price level is fully determined by the distribution of money in t-1 and is therefore not stochastic. Hence, present and future prices are fully determined by the initial distribution of money, and by future amounts of money.

In order to complete the proof of existence and uniqueness of equilibrium (which is only sketched here) note that the price levels in all periods depend on the path of future prices as shown by equations (A.5) and (A.7). A fixed-point argument shows the existence and uniqueness of an equilibrium price path. Note, that the equilibrium price is finite in each period. Namely money has a positive value despite the fact that it does not produce any direct utility. If there are no taste shocks the price of money is equal to zero and it has no value. To see this note that in this case all those who hold money in period t want to consume it since they have preference to the present: $\beta<1$. As a result the demand for goods exceeds y, but since the equilibrium amount of consumption must be y in this exchange economy, the equilibrium real balances become 0, namely money has no value. Hence, money has value in this economy only because transactions have different sizes.

The model can be extended by adding an asset that has higher return than money but is less liquid, as in the Baumol-Tobin approach and as done in Section 4 in the paper. If bonds pay a fixed real rate of interest *r*, but each sale of bonds has a fixed cost of real size *x*, people hold money none the less, for making small purchases. If the purchase is big enough, people might be willing to bear the cost *x* and not lose utility as a result of not making the full purchase. In a similar way all the main ideas of the paper can be developed within this framework as well, but the solution is less tractable, and possibly only numerical and not analytical. Hence, the paper uses a much simpler version, where life cycles are short and the model is solved fully.

Appendix B: Proofs

Proof of Proposition 1:

First, it can be shown that there exists a stable steady state equilibrium under this monetary policy as well. We focus from here on this steady state. In equilibrium, we get from (I.9), when we omit time subscripts:

(B.1) $$s = \frac{1}{2}\frac{\mu}{\Pi}\frac{M}{P} = \frac{1}{2}\frac{\Pi - 1}{\Pi}\frac{M}{P}.$$

Optimal behavior of each person is the same as in the benchmark model, except income is labor income plus the subsidy: $y + s$. Hence an individual demand for money is:

$$l(\theta) = (y+s)(1-\theta-\theta\Pi), \text{ } as\ long\ as\ \theta \le \frac{1}{1+\Pi}.$$

Hence the aggregate demand for money is:

(B.2) $$\frac{M}{P}=\frac{y+s}{2}\frac{1}{1+\Pi}.$$

Combining equations (B.1) and (B.2) together leads to the following steady state value of the subsidy:

(B.3) $$s=y\frac{\Pi-1}{4\Pi^2+3\Pi+1}.$$

We next examine the effect of this monetary policy on utility. Since inflation affects people differently we calculate the expected utility and get:

(B.4) $$\ln y+\ln 4+\ln(\Pi^2+\Pi)-\ln(4\Pi^2+3\Pi+1)+\frac{1}{1+\Pi}\ln(1+\Pi)- \\ -\ln\Pi\int_0^{\frac{1}{1+\Pi}}(1-\theta)d\theta+\int_0^{\frac{1}{1+\Pi}}[\theta\ln\theta+(1-\theta)]d\theta.$$

The derivative with respect to Π is equal to:

(B.5) $$\frac{1}{\Pi}\left(1-\frac{2\Pi+1}{2\Pi^2+4\Pi+2}\right)+\frac{2+\Pi}{(1+\Pi)^2}-\frac{8\Pi+3}{4\Pi^2+3\Pi+1}.$$

This derivative is declining and is equal to zero when $\Pi = 1$. QED.

<u>Proof of Proposition 2:</u>

Derivation of (II.7) with respect to $l_{1,t}$ yields

$$\frac{\partial EU_t}{\partial l_{1,t}}=\frac{\partial\theta_1}{\partial l_{1,t}}\left[U_A(\theta_{t+1})-U_B(\theta_{t+1})\right]+\int_0^{\theta_{t+1}}\frac{\partial U_A(\theta)}{\partial l_{1,t}}d\theta+\int_{\theta_{t+1}}^{1}\frac{\partial U_B(\theta)}{\partial l_{1,t}}d\theta.$$

Since the first item in the RHS is zero, as utility at the threshold between A and B is equal, the marginal utility of first period money is

(B.6)
$$\frac{\partial EU}{\partial l_1} = -\frac{1}{2}\frac{1}{y-l_1} + \theta_{t+1}\frac{1}{y\Pi_{t+1} + y\Pi_{t+1}\Pi_{t+2} + l_{1,t}} + \frac{1}{2}\frac{1-\theta_{t+1}^2}{y\Pi_{t+1} + l_{1,t}} =$$
$$= \frac{1}{2y}\left[-\frac{1}{1-h} + \frac{1}{\Pi_{t+1} + h} + \frac{\Pi_{t+1} + h}{\left(\Pi_{t+1} + \Pi_{t+1}\Pi_{t+2} + h\right)^2}\right],$$

where: $h = \frac{l_{1,t}}{y}$.

(B.6) is decreasing with liquidity h and as h approaches 1 it goes to -∞. Hence a maximum exists and is unique. It is equal to the h that equates (B.6) to 0, or by $h = 0$ if (B.6) is negative already at zero. Furthermore, liquidity in first period of life is proportional to income as $l_{1,t} = yh$. It can be shown that (B.6) is negatively related to both Π_{t+1} and Π_{t+2}. Hence, an increase in the rate of inflation reduces h and reduces the precautionary demand for money. Next examine when h is positive. Liquidity h is equal to zero when the marginal expected utility is:

$$\frac{1}{2y}\left[-1 + \frac{1 + (1+\Pi_{t+2})^2}{\Pi_{t+1}(1+\Pi_{t+2})^2}\right].$$

Hence, the precautionary demand for money is positive, $h > 0$, if:

(B.7)
$$\left(\Pi_{t+1} - 1\right)\left(1 + \Pi_{t+2}\right)^2 < 1.$$

Namely, if the rate of inflation is not too high this condition holds and the young hold a positive precautionary amount of money. If inflation is high (B.7) does not hold and the precautionary demand is zero. Q.E.D.

<u>Proof of Proposition 3:</u>

From calculation of the demand for money (II.8) and from (II.5) we get that the equilibrium in the money market is described by:

(B.8) $$\frac{M_t}{P_t} = y\left[h(\Pi_{t+1},\Pi_{t+2}) + \frac{\left(\Pi_t + h(\Pi_t,\Pi_{t+1})\right)^2}{2\Pi_t\left(\Pi_t + \Pi_t\Pi_{t+1} + h(\Pi_t,\Pi_{t+1})\right)}\right].$$

The dynamics of real balances are described by:

(B.9) $$\frac{M_t}{P_t} = \frac{M_{t-1}}{P_{t-1}}\frac{1+\mu}{\Pi_t}.$$

From equating (B.8) and (B.9) we get that the equilibrium rate of inflation satisfies:

$$\Pi_t = F\left(\frac{M_{t-1}}{P_{t-1}},\Pi_{t+1},\Pi_{t+2}\right).$$

Hence, the rational expectations equilibrium rate of inflation is a function of previous real balances:

$$\Pi_t = \Pi\left(\frac{M_{t-1}}{P_{t-1}}\right).$$

It can be shown that with this equilibrium function the steady state is stable. From (B.9) it follows that in the steady state the rate of inflation is equal to the rate of monetary expansion: $\Pi_t = 1+\mu$. Hence, real balances at the steady state are equal to:

$$\frac{M}{P} = y\left[h + \frac{(1+\mu+h)^2}{2(1+\mu)(2+3\mu+\mu^2+h)}\right].$$

Hence, real balances depend negatively on μ, directly and through h. Q.E.D.

<u>Proof of Proposition 4:</u>

The marginal expected utility with respect to first period money satisfies:

$$\frac{\partial EU}{\partial l_1}=\frac{\partial\theta_1}{\partial l_1}\left[U_A(\theta_1)-U_B(\theta_1)\right]+\frac{\partial\theta_2}{\partial l_1}\left[U_B(\theta_2)-U_C(\theta_2)\right]+$$
$$+\int_0^{\theta_1}\frac{\partial U_A(\theta)}{\partial l_1}d\theta+\int_{\theta_1}^{\theta_2}\frac{\partial U_B(\theta)}{\partial l_1}d\theta+\int_{\theta_2}^{1}\frac{\partial U_C(\theta)}{\partial l_1}d\theta=$$
$$=\int_0^{\theta_1}\frac{\partial U_A(\theta)}{\partial l_1}d\theta+\int_{\theta_1}^{\theta_2}\frac{\partial U_B(\theta)}{\partial l_1}d\theta+\int_{\theta_2}^{1}\frac{\partial U_C(\theta)}{\partial l_1}d\theta.$$

Hence, the marginal expected utility is equal to:

(B.10)
$$\frac{\partial EU}{\partial l_1}=\frac{\theta_2^2-\theta_1^2}{2}\frac{1}{l_1}-\theta_1\frac{(1+r)^2\Pi^2-1}{y(1+r)^2\Pi^2-l_1[(1+r)^2\Pi^2-1]}-$$
$$-\frac{(1-\theta_1)^2-(1-\theta_2)^2}{2}\frac{1}{y-l_1}-(1-\theta_2)\frac{(1+r)\Pi-1}{y(1+r)\Pi-l_1[(1+r)\Pi-1]}.$$

Marginal expected utility is decreasing in l_1. We next show that the marginal expected utility at $l_1=0$ is infinite, and hence the optimal amount of money, where the marginal expected utility is zero, must be positive. To show this note that as $l_1\to 0$ both θ_1 and θ_2 approach zero as well. We next show that the first item in the RHS of (B.10) goes to ∞ when $l_1\to 0$, while it is easy to see that the other terms are negative and finite. From the definition of θ_2 in (III.9) we get that as l_1 and θ_2 converge to zero we get:

$$\theta_2\log l_1\xrightarrow[l_1\to 0]{}-z.$$

Hence, for l_1 close enough to 0:

$$\theta_2\geq-\frac{z}{2\log l_1}.$$

Hence:

$$\frac{\theta_2^2}{l_1}\geq\frac{z^2}{4(\log l_1)^2 l_1}\xrightarrow[l_1\to 0]{}\infty.$$

Since from (III.6) we get:

$$\frac{\theta_1^2}{l_1} \le \frac{l_1}{(y-l_1)^2(1+r)^4\Pi^4} \xrightarrow[l_1 \to 0]{} 0.$$

Namely the first item in the RHS of (B.10) goes to infinity and so does marginal expected utility, as money diminishes to zero. Hence the optimal amount of money must be positive.

From (B.10) and from the definitions of θ_1 and θ_2 we can see that l_1 is proportional to income y and hence: $l_1 = yh$. We next show that h depends negatively on $(1+r)\Pi$, the nominal interest rate. Using the following notation for the gross nominal interest rate, $I = (1+r)\Pi$, we can write the expected utility as:

(B.11) $$EU = \int_0^1 \max\left\{ \begin{array}{l} N(\theta) - (2-\theta)\log\Pi + \log[I^2 - h(I^2-1], \\ 2(1-\theta)\log(1+r) - \theta\log\Pi + \theta\log h + (1-\theta)\log(1-h), \\ N(\theta) + (1-\theta)\log(1+r) - \log\Pi + \log[I - h(I-1)] + -z \end{array} \right\} d\theta + \\ + \log y - z - x.$$

Clearly a rise in the nominal interest rate I reduces the marginal utility of holding l_1 and hence it reduces the optimal l_1.

Since optimal l_1 is proportional to y, it follows from (B.11) that the optimal expected utility can be written as:

$$\max EU = \log y + \phi(r, \Pi, z) - x.$$

Note that if the cost of finding an investor and signing a contract is high, the consumer might prefer not to hold bonds at all. In that case utility is equal to:

$$\int_0^1 [N(\theta) - (2-\theta)\log\Pi]d\theta + \log y.$$

Hence, if $x > \phi(r, \Pi, z) - \int_0^1 [N(\theta) - (2-\theta)\log\Pi]d\theta$, the consumer prefers to hold money only and does not go to the bonds market. QED.

Proof of Proposition 5:

Consider the four cases in second period of life. In case A the consumer purchases consumption with only part of her cash. Utility is:

$$\theta \log c_2 + (1-\theta)\log \frac{f + d - \Pi c_2}{\Pi^2} - 2v = \theta \log c_2 + (1-\theta)\log \frac{y - \Pi c_2}{\Pi^2} - 2v.$$

Optimization, subject to $c_2 \leq f/\Pi$, leads to: $c_2 = \theta y/\Pi$, $c_3 = (1-\theta)y/\Pi^2$, and cash held for next period is: $(f - \theta y)/\Pi$. Utility is:

$$U_A(\theta) = N(\theta) + \log y - (2-\theta)\log \Pi - 2v.$$

In case B consumption is larger, so all the cash is consumed and all demand deposits are kept for next period. Hence:

$$U_B(\theta) = \theta \log f + (1-\theta)\log(y - f) - (2-\theta)\log \Pi - 2v.$$

In case C consumption is financed by all the cash and by some demand deposits, while some demand deposit are kept in the bank. The consumer maximizes:

$$\theta \log c_2 + (1-\theta)\log \frac{y - \Pi c_2}{\Pi^2} - 3v.$$

Hence consumption in second period is $c_2 = \theta y/\Pi$, and in third period is $c_3 = (1-\theta)y/\Pi^2$, and utility is:

$$U_C(\theta) = N(\theta) + \log y - (2-\theta)\log \Pi - 3v.$$

In case D the consumer leaves only cash for third period. Assume that inflation is not too high so that: $e\Pi < 1-\tau$. In this case the consumer leaves the maximum amount of cash, namely $e\,y$. Hence: $c_2 = (1/\Pi - e)y$ and $c_3 = ey/\Pi$. Utility in this case is:

$$U_D(\theta) = \theta \log\left(\frac{1}{\Pi} - e\right) + (1-\theta)\log e + \log y - (1-\theta)\log \Pi - 2v.$$

Note that the borderline between cases A and B is $\theta_1 = f / y$. Denote the borderline between B and C (or D if C is empty) by θ_2. Note that the utilities for A, C and D do not depend on f. Hence the marginal expected utility is equal to:

$$\frac{1}{\tau}\int_{\theta_1}^{\theta_2} \frac{\partial U_B(\theta)}{\partial f} d\theta = \frac{1}{\tau}\int_{\theta_1}^{\theta_2} \left(\frac{\theta}{f} - \frac{1-\theta}{y-f}\right) d\theta.$$

Note that for $\theta \geq \theta_1 = f / y$, the integral is positive. Hence, the young consumer increases the amount of cash as much as possible up to the safety bound, so that $f = ey$.

Given the initial demand for cash we can calculate the amounts of cash in the second period of any individual and then the overall amount of cash. First we calculate the borderlines between the various sets. The borderline between B and C is described by:

(B.12) $$N(\theta_2) - \theta_2 \log e - (1-\theta_2)\log(1-e) = v.$$

The borderline between C and D is described by:

(B.13) $$N(\theta_3) - \theta_3 \log(1-\Pi e) - (1-\theta_3)\log(\Pi e) = v.$$

The overall amount of cash in the economy F is:

(B.14) $$F = ye + \frac{y}{\tau\Pi}\int_0^e (e-\theta) d\theta + \frac{y}{\tau}\int_{\theta_3}^{\tau} e d\theta = ye\left(\frac{e}{2\tau\Pi} + 2 - \theta_3\right).$$

Hence, the amount of cash is proportional to income y. Furthermore it depends positively on e, namely on the degree of safety in holding cash. As for τ, the distribution of taste shocks, it clearly has a negative effect on the demand for cash (note that θ_3 does not depend on τ).

The overall amount of demand deposits is:

(B.15)

$$D = y(1-e) + \frac{y}{\tau\Pi}\int_0^{\theta_2}(1-e)d\theta + \frac{1}{\tau}\int_{\theta_2}^{\theta_3}\frac{y(1-\theta)}{\Pi}d\theta =$$

$$= y\left(1-e+\frac{y\theta_3}{\tau\Pi}-\frac{ey\theta_2}{\tau\Pi}-\frac{y}{\tau\Pi}\frac{\theta_3^2-\theta_2^2}{2}\right).$$

Note that demand deposits depend negatively on τ. Q.E.D.

References

Baumol, William J., "The Transactions Demand for Cash: An Inventory Theoretic Approach," *Quarterly Journal of Economics*, 66 (1952), 545-556.

Berentsen, Aleksander, and Rocheteau, Guillaume, "Money and the Gains From Trade," *International Economic Review*, 44 (2003), 263-297.

Bewley, Truman, "The Optimum Quantity of Money," in *Models of Monetary Economics*, eds. John H. Kareken and Neil Wallace, Federal Reserve Bank of Minneapolis, Minneapolis, 1980, 169-210.

Bewley, Truman, "A Difficulty with the Optimum Quantity of Money," *Econometrica*, 51 (1983), 1485-1504.

Diamond, Douglas W., "Financial Intermediation and Delegated Monitoring," *Review of Economic Studies*, 51 (1984), 393-414.

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

Diamond, Peter A., "Money in Search Equilibrium," *Econometrica*, 52 (1984), 1-20.

He, Ping, Huang, Lixin, and Wright, Randall, "Money and Banking in Search Equilibrium," mimeo (2004).

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

Kiyotaki, Nobuhiro, and Wright, Randall, "A Contribution to the Pure Theory of Money," *Journal of Economic Theory*, 53 (1991), 215-235.

Kiyotaki, Nobuhiro, and Wright, Randall, “A Search-Theoretic Approach to Monetary Economics,” *American Economic Review*, 83 (1993), 63-77.

Kocherlakota, Narayana R., “Money is Memory,” *Journal of Economic Theory*, 81 (1998), 232-251.

Lagos, Ricardo, and Wright, Randall, “A Unified Framework for Monetary Theory and Policy Analysis,” *Journal of Political Economy*, 113 (2004), 463-.

Lindbeck, Assar, *A Study in Monetary Analysis*, Almquist&Wiksell, Stockholm, 1963.

Lucas, Robert, E., Jr., “Equilibrium in a Pure Currency Economy,” in *Models of Monetary Economics*, eds. John H. Kareken and Neil Wallace, Federal Reserve Bank of Minneapolis, Minneapolis, 1980, 131-145.

Niehans, Jürg, *The Theory of Money*, The Johns Hopkins University Press, Baltimore, 1978.

Patinkin, Don, *Money, Interest, and Prices: An Integration of Monetary and Value Theory*, Evanston, Illinois, Row Peterson and Co., 1956.

Prescott, Edward C., “A Multiple Means-of-Payment Model,” in *New Approaches to Monetary Economics*, eds. William A. Barnett and Kenneth J. Singleton, Cambridge University Press, Cambridge, 1987, 42-51.

Rocheteau, Guillaume, and Wright, Randall, “Money in Search Equilibrium, in Competitive Equilibrium, and in Competitive Search Equilibrium,” mimeo, 2003.

Romer, David, “A Simple General Equilibrium Version of the Baumol-Tobin Model,” *Quarterly Journal of Economics*, 101 (1986), 663-686.

Sidrauski, Miguel, “Rational Choice and the Patterns of Growth in a Monetary Economy,” *American Economic Review*, 57 (May 1967), 534-544.

Sussman, Nathan, and Zeira, Joseph, “Commodity Money Inflation: Theory and Evidence from France in 1330-1436,” *Journal of Monetary Economics*, 50 (2003), 1769-1793.

Starr, Ross M., “Why is there Money? Endogenous Derivation of 'Money' as the Most Liquid Asset: A Class of Examples,” *Economic Theory*, 21 (2003), 455-474.

Tobin, James, “The Interest Elasticity of the Transactions Demand for Cash,” *Review of Economics and Statistics*, 38 (1956), 241-247.

Wallace, Neil, “The Overlapping Generations Model of Fiat Money,” in *Models of Monetary Economics*, eds. John H. Kareken and Neil Wallace, Federal Reserve Bank of Minneapolis, Minneapolis, 1980, 49-82.

Wallace, Neil, “Wither Monetary Economics?” *International Economic Review*, 42 (2001), 847-870.

Whitesell, William C., “Deposit Banks and the Market for Payment Media,” *Journal of Money, Credit and Banking*, 21 (1992), 483-498.

Figures

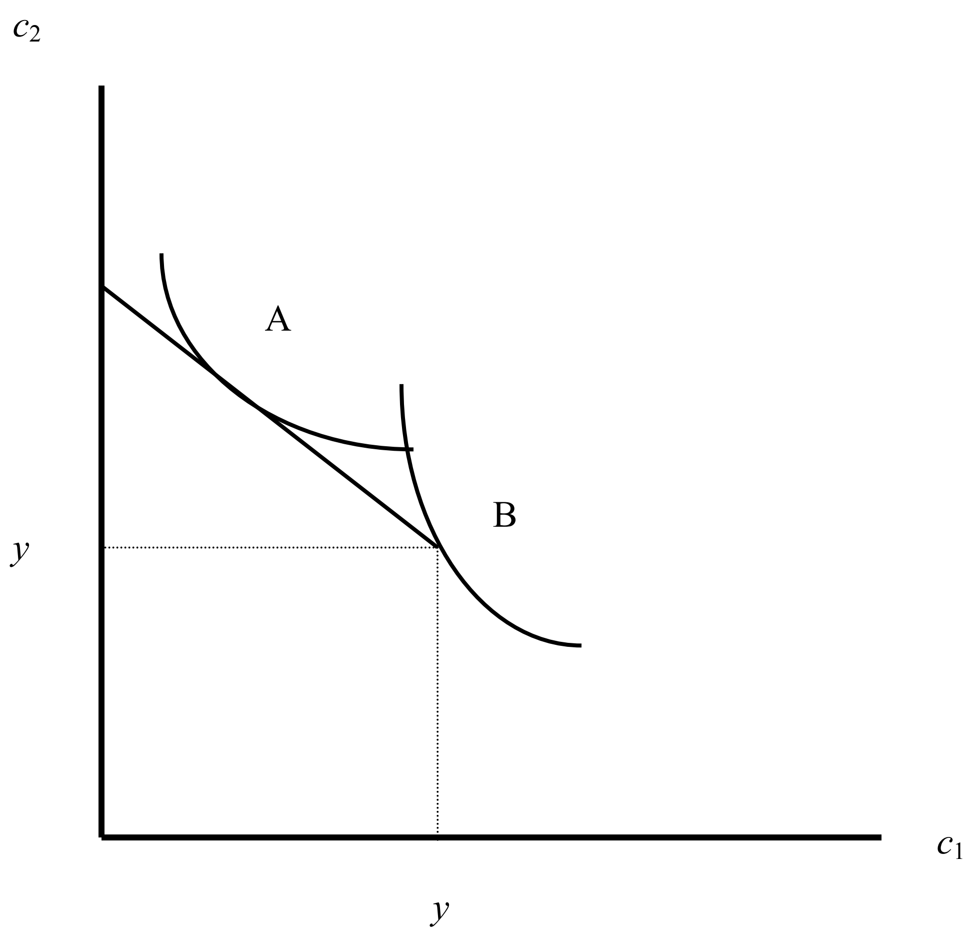

Figure 1

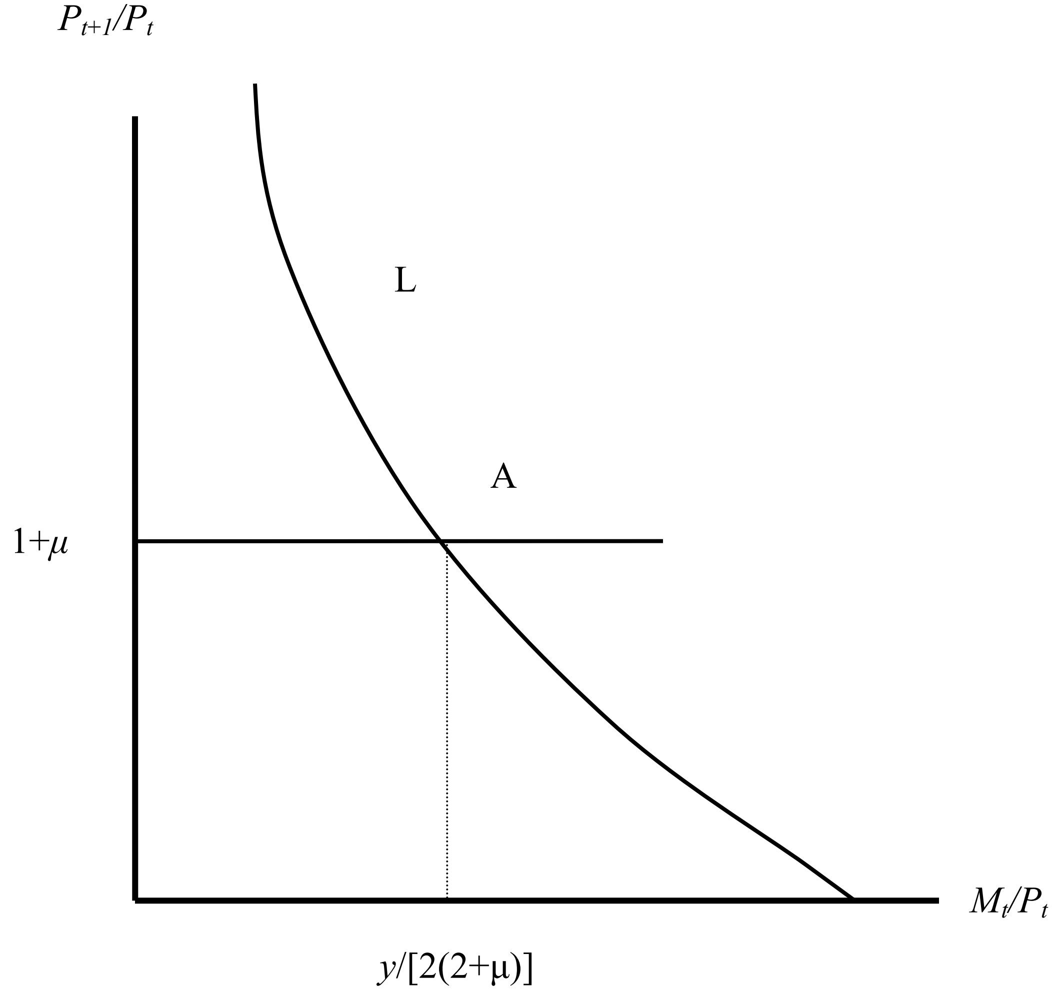

Figure 2