Document de travail de l'OFCE

FINANCIAL INTERMEDIARIES AND TRANSACTION COSTS

N° 2010-02

JANUARY 2010

Augusto Hasman
SKEMA Business School and OFCE

Margarita Samartin
Business Department, Universidad Carlos III de Madrid

Jos van Bommel
Universidad Cardenal Herrera in Valencia

OFCE - Centre de recherche en économie de Sciences Po
69, quai d'Orsay - 75340 Paris Cedex 07 - Tél/ 01 44 18 54 00 - Fax/ 01 45 56 06 15
www.ofce.sciences-po.fr

Financial Intermediaries and Transaction Costs

Augusto Hasman, Margarita Samartín, and Jos van Bommel*

August 2009

We present an overlapping generations model with spatial separation and agents who face unsystematic liquidity risk. In a pure exchange economy, agents engage in life cycle portfolio rebalancing. In an intermediated economy, intergenerational banks or mutual funds cater to diversified clienteles so as to avoid rebalancing transactions. In equilibrium, these intermediaries pay redemptions with portfolio income and never sell secondary assets. We also find that the pure exchange economy has a downward sloping yield curve and is inherently cyclical.

JEL Classification: D91, G21, E43

Keywords: Financial Intermediation, Overlapping Generations, Liquidity

---

* Hasman is at SKEMA Business School and OFCE DRIC in Sophia Antipolis, Samartín at Universidad Carlos III in Madrid, and van Bommel at Universidad Cardenal Herrera in Valencia. Our e-mails are: augusto.hasman@sciences-po.fr, samartin@emp.uc3m.es, jozef.vanbommel@uch.ceu.es.
We thank participants of the finance workshop at the University of Oxford for valuable feedback. All errors are ours. This research is partly funded by the Spanish Ministry of Education and Culture, project SEJ2007-67448/ECON.

1. Introduction

The incumbent explanations for the existence of financial intermediaries are that they (i) provide liquidity risk sharing, (ii) solve inefficiencies due to asymmetric information, and (iii) align incentives through active monitoring. A rich literature on each of the above theories exist.[1] On a fourth generally accepted *raison d´être* for financial intermediation, (iv) scale economies in transactions and logistics, a formal model is conspicuously absent.

In this paper, we use a three period overlapping model in which agents may need to consume before a long term production technology pays a riskless dividend. To share liquidity risk, agents trade secondary claims for consumption goods in a pure exchange market or open accounts with financial institutions in an intermediated economy. Our basic framework is thus the same as the OLG Diamond Dybvig (1983) model employed by Qi (1994), Bhattacharya and Padilla (1996), and Fulghieri and Rovelli (1998), who all investigate the relative merits of financial intermediaries and markets in sharing liquidity risk.[2] The distinguishing feature of our model is the existence of transaction costs, which we model as dead weight shoe leather costs that are incurred whenever agents interact.

We find that in the pure exchange economy, a centrally located market opens where early consumers sell secondary claims to newborns and late consumers. In equilibrium, newborns invest their endowment in a mix of short term and long term claims, and sell these claims when they have unexpected consumption needs. Agents without consumption needs also trade, because they need to reinvest the dividends from their maturing assets. Hence, in equilibrium, early consumers travel to the market twice while late consumers travel three times.

[1] The seminal paper addressing the liquidity insurance argument is Diamond and Dybvig (1983). Banks´ circumvention of information asymmetries was conjectured by Leland and Pyle (1977), and modelled by Boyd and Prescott (1986) and Broecker (1990), among others. Diamond (1984) demonstrates how delegating loan-monitoring to financial institutions can be economical. See Freixas and Rochet (1997) for a comprehensive review of the banking literature.

[2] In single generation models, the papers by Diamond Dybvig (1983), Jacklin (1987), Jacklin and Bhattacharya (1988), Hellwig (1994), Alonso (1996), Samartín (2001), Diamond (1997) and von Thadden (1998) among others, analyze the relative degrees of risk sharing provided by banks and markets. For a review of this literature, see Dwyer and Samartín (2008).

As an alternative to the exchange mechanism, a centrally located financial intermediary may open. In following with the literature, we will denote this institution a *bank*, even though open-end mutual funds and insurance companies take on similar roles in centralizing liquidity and reducing transaction costs. Agents open deposits with the bank and the bank invests in the technology, so that agents undertake only two trips: one to deposit, and one to withdraw. Because the depositor clientele of the bank is diversified and stationary, no costly rebalancing transactions are required, making the intermediated economy superior to the pure exchange economy.

Apart from the reduced number of transactions, our analysis uncovers an additional advantage of intermediated economies relative to exchange economies. We show that while the latter are inherently cyclical, the former achieve non-cyclical allocations.[3] We find that for small transaction costs, the exchange economy with a starting date cannot escape from a severe cyclical pattern, and that intergenerational welfare can be improved upon if transaction costs are introduced, because it dampens cyclicality. We also find that the starting up phase of an intermediated economy is significantly shorter than the time to stationarity in an exchange economy.

Benston and Smith (1976) argue that that banks economize on transactions costs. They interpret transaction costs as costs of transportation, administration, search, evaluation, and monitoring, among others, and argue that banks enjoy economies of scale, scope, and networks in these tasks. In this paper we disregard transaction costs due to information asymmetries or moral hazard, and instead focus only on the most mundane, yet unavoidable processing costs of the *shoe leather* type. Since Baumol (1952), Tobin (1956) and Orr and Mellon (1961), who analyze the trade-off between holding cash and financial assets in the presence of shoe leather costs, these simplest of transaction costs have been largely ignored in micro economic theory of banking.[4]

Exceptions are models of Diamond (1997) and Qian *et al.* (2004), who assume that agents face an exogenous risk of facing (prohibitive) transaction costs. They show, in

---

[3] Azariadis (1981) first observed that OLG models can have cyclical equilibria. Bhattacharya, Fulghieri and Rovelli (1998) first mention the cyclicality in the OLG Diamond Dybvig economy.

[4] Macro economists *do* acknowledge the importance of shoe leather costs. Lucas (2000) for instance argues that the excessive shoe leather costs brought on by inflation may account for as much as 1% of GNP when moving to the Friedman optimum.

a stationary and in an OLG model respectively, that this transaction cost risk gives intermediaries an advantage over pure exchange mechanisms.

Papers that study the effect of transaction costs on mutual funds include Cherkes *et al.* (2008), who examine how closed end funds can save on transaction costs. Their analysis shows that retail investors find funds attractive, even at prices above net asset value, because they offer them lower round trip transaction costs than home made portfolios of illiquid stocks. Chordia (1996) investigates the case for open end funds, and shows that, apart from achieving economies of scale in transaction costs and improving on diversification, mutual funds provide transaction cost sharing.[5] We show that funds can avoid transaction costs altogether. In our model economy (in which there is no aggregate risk), an open end fund will never sell secondary assets. Because intermediaries have large diversified clienteles of overlapping generations, total redemptions will exactly be paid from dividend income.

In the pure exchange equilibrium on the other hand, agents will rebalance their portfolio continuously, moving to shorter term securities as they grow older. We argue that a prime role of intermediaries is the cancelling out of such life cycle rebalancing.

Our model also shows that, in the absence of aggregate risk, transaction costs lead to a downward sloping yield curve. The reason for this is that equilibrium investments in long term securities come with lower *per period* transaction costs. In a general equilibrium model Vayanos (1998) describes this effect for an economy with transaction costs and aggregate risk. He shows that illiquid stocks may trade at higher yields, because they internalize lower expected transaction costs.[6]

The rest of the paper is organized as follows: the basic model is presented in the next section. Section 3 analyzes the equilibrium in an unbounded pure exchange economy, while section 4 investigates how a competitive bank can improve on the above allocation. Section 5 juxtaposes the exchange and bank equilibrium in an economy

---

[5] Funds can effectively enforce an ex-ante desirable allocation of transaction costs: early consumers end up paying less and late consumers paying more. Van Bommel (2008) makes a similar argument for a Diamond Dybvig bank.

[6] Other papers that study the effect of transaction costs on asset prices are Constantinides (1986), Aiyagari and Gertler (1991), Heaton and Lucas (1996), Vayanos and Vila (1999), and Huang (2003).

with a starting date and a first generation. We discuss several possible extensions to our model in section 6, and summarize our findings in section 7. The appendix contains the proofs, an analysis of the model with variable transaction costs, and an analysis of the bank's start up phase.

2. The model

We analyze an infinite horizon, single consumption good economy, where time is represented by the unbounded set of integers, representing the dates t on which agents interact. Henceforth we refer to periods between two dates as years. On every date, a continuum of agents is born on a circle with radius one, where they live throughout their lives. Agents are born with an endowment of one unit of consumption good and live either one year, with probability λ, or two years, and consume only on the date they die. All agents have expected utility preferences, modelled by an instantaneous utility function with constant relative risk aversion γ: $U(C)=\frac{1}{1-\gamma}C^{1-\gamma}$. Agents born at date t learn their type after t but before t+1. Types are not verifiable. We assume that the population is large enough so that there is no uncertainty on the aggregate distribution of agents in the population.[7]

The economy is endowed with two technologies to produce goods over time. The first technology, *storage*, allows agents to costlessly transfer consumption from one period to the next, at the same location. The second, *long term*, technology allows agents to convert one unit of consumption at date t into $R > 1$ units of consumption at $t + 2$. This technology, which cannot be interrupted at $t + 1$, is located in the midpoint of the circle where it is operated by a firm which issues, inelastically, two-year bonds with face value R payable at the midpoint for a price of unity.

Naturally, agents can buy 2-year bonds and share liquidity risk by trading secondary claims. To do this, agents travel to a centrally located *market*, where they exchange one-year bonds for consumption goods. The price for such bonds, denoted p_t, is set by a Walrasian auctioneer to whom buyers and sellers submit their demand and supply

[7] A continuum of newborns of size one, λ impatient one-year olds, (1-λ) patient one-year olds, and (1-λ) two-year olds. The certain aggregate distribution of types, which is common in this literature, is usually justified in terms of the law of large numbers. Duffie and Sun (2007) provide a model of independent random matching that makes the law of large numbers hold exactly.

curves. We abstract away from moral hazard so that buying (selling) bonds equates to riskless lending (borrowing).

Agents may also form a financial intermediary. In this case, agents set up a *bank* where they deposit their endowments and receive a demandable debt security that entitles the holder to withdraw r_1 after one year or r_2 after two years. We shall assume throughout this article that exchanges and banks make zero profit.

The cost of travel, denoted c, is the distinguishing feature of our model. For tractability we assume this cost to be independent of the value of claims or goods transported. The fixed cost assumption produces an exchange equilibrium in which newborns employ mixed strategies and hold portfolios that consist either of only two-year bonds, or of only one-year bonds. In the appendix we show that a variable transaction cost leads to qualitatively the same insights as the fixed cost model. We further assume, without loss of generality, that the transaction cost is levied on the way back home.

In the following we will characterize the equilibrium allocations, denoted $\{C_1,C_2\}$, in the exchange and bank economies. Naturally, we will compare agents' *ex-ante* utility in both economies.

3. The exchange economy

In a pure exchange economy, newborns and patient one-year olds can store or buy one- or two-year bonds. To save space we shall limit our attention to the case where transaction costs are low enough for storage to be always dominated by investing. Because there is no uncertainty, and transaction costs are fixed, newborns choose between two strategies: they either spend their entire endowment on buying two-year bonds, which they will sell if they become impatient, or they buy only one-year bonds and *roll over* these bonds if they remain patient. We denote these strategies *LT* and *ST* respectively. Notice that in the *LT* strategy all agents travel twice, while in the *ST* strategy, patient agents travel three times.

Clearly, in equilibrium we need both types of agents. This implies that agents employ mixed strategies. In equilibrium, agents must be indifferent between both strategies.

Hence, we need that:

$$
\begin{aligned}
&E[U]_{LT} \equiv \lambda U\left((1-c)p_{t+1}-c\right)+(1-\lambda)U\left((1-c)R-c\right) \\
&= E[U]_{ST} \equiv \lambda U\left(\frac{(1-c)R}{p_t}-c\right)+(1-\lambda)U\left(\left(\frac{(1-c)R}{p_t}-c\right)\frac{R}{p_{t+1}}-c\right) \quad \forall t
\end{aligned} \tag{1}
$$

The first bracketed term denotes the consumption of an agent who employs the *LT* strategy and becomes impatient: when born she buys 1-*c* two-year bonds and sells them at her first birthday for p_{t+1}. The fourth bracketed term is the consumption of an agent who plays *ST* and turns out patient: she buys $\frac{1-c}{p_t}$ one-year bonds when born, so that on her first birthday she receives $\frac{(1-c)R}{p_t}$ in midpoint of the circle. Because she is patient, she reinvests the proceeds (net of transaction costs) in one-year bonds. One year later she travels again and consumes $\left(\frac{(1-c)R}{p_t}-c\right)\frac{R}{p_{t+1}}-c$.

Notice that we do not consider the case where newborns buy one-year bonds, and then roll them over into two-year bonds when they stay patient. We will show that this strategy is strictly dominated by playing *LT* or *ST*.

It can be easily verified that there is a unique stationary price process, $p_t^*=p^*$ that solves (1) for all *t*. With this p^*, we can derive the equilibrium probability θ^* with which newborns play *LT* from the following market clearing condition:

$$
p^* = \frac{(1-\theta^*)(1-c)+(1-\theta^*)(1-\lambda)\left(\frac{R(1-c)}{p^*}-c\right)}{\theta^*(1-c)\lambda} \tag{2}
$$

In expression (2) the numerator is the equilibrium supply of goods, and the denominator the equilibrium supply of bonds. The following proposition gives the solution for the unique full investment equilibrium in the unbounded market economy.

Proposition 1 *(market equilibrium)*

For low enough c, there exists a unique stationary market equilibrium with fully invested agents (no storage). For the equilibrium price we have:

$$
p^* \in \left(\frac{\sqrt{c^2+4R(1-c)^2}-c}{2(1-c)}, \sqrt{R}\right) \tag{3}
$$

Newborn agents spend their entire endowment on two year bonds with probability θ^**, and invest all in one-year bonds otherwise. The equilibrium periodic investment in two year bonds is given by:*

$$\theta^* = \frac{(1-c)p^* + (1-\lambda)\left(R - c(R+p^*)\right)}{(1-c)(\lambda p^* + 1)p^* + (1-\lambda)\left(R - c(R+p^*)\right)} \tag{4}$$

A closed form solution for p^* can only be found for certain values for γ. For brevity and relevance we focus on the interval in which p^* lies. This interval, given by (3), follows from equation (1). The upper bound for p^*, $\sqrt{R}$, makes the *LT* strategy dominant: with $p^* = \sqrt{R}$ impatient consumption is independent of the strategy played, while patient consumption is higher under the *LT* strategy (see (1)). Similarly, if p^* = $\frac{\sqrt{c^2 + 4R(1-c)^2} - c}{2(1-c)}$, the *ST* strategy would dominate. The expression for θ^* follows from rearranging equation (2).

The key point is that $p^* < \sqrt{R}$, which implies that the return on a one-year bond is higher than the one year return on a two-year bond. This also is the reason why one-year old patient agents do not buy two-year bonds. Since their holding period is one year with certainty, patient one-year olds will opt for one-year bonds.

Notice that proposition 1 limits itself to the full investment case. We shall later see that if c is zero or is higher than a threshold, multiple cyclical equilibria exist. Moreover, if the economy has a starting date (and hence a unique equilibrium), the economy may be by stuck in severe cyclicality. In section 5 we shall show that for a wide range of parameter values, the presence of transaction costs in a pure exchange economy with a starting date may enable convergence to the one-periodic equilibrium characterized in proposition 1 and hence *improve* overall welfare vis-à-vis the transaction cost free economy. But before considering the truncated economy, we shall first investigate how a financial intermediary can improve on the market allocation.

4. The intermediated economy

We now consider how financial intermediaries can improve on the market allocation. Following the literature, we consider the simplest of intermediaries: infinitely lived deposit institutions whose only source of capital are deposits. In section 5 we consider how such intermediaries emerge in an economy with a starting date. In a stationary economy the intermediaries issue demandable debt securities that give depositors the right to withdraw, per unit of consumption good deposited, either r_1 after one year, or r_2 after two years. Naturally, the former payout is earmarked for the impatient agents, and the latter for the patient types. If we denote y the bank's periodic investment in the production technology, we can write its objective function as follows:

$$\max_{r_1, r_2} \lambda U\left((1-c)r_1 - c\right) + (1-\lambda)U\left((1-c)r_2 - c\right) \tag{5}$$

subject to:

$$\lambda r_1 + (1-\lambda)r_2 + y \le 1 + yR \tag{6}$$

$$r_1 \le \sqrt{R} \tag{7}$$

$$r_2 \le R \tag{8}$$

Equation (6) gives the bank's budget constraint, while equations (7) and (8) are *no-arbitrage* conditions, which need to hold to avoid that competing banks open up immediately next to the incumbent and invest with their neighbors instead of investing in the production technology.[8] Naturally, (7) and (8) will bind so that we immediately have:

Proposition 2 *(bank economy)*

If c is sufficiently small, there exists a bank equilibrium in which banks offer, per unit of good deposited, $r_1 = \sqrt{R}$ *one year after deposit, or* $r_2 = R$ *after two years, at the option of the depositor. The periodic investment is given by:*

$$y^* = (1-c)\frac{\lambda\sqrt{R} + (1-\lambda)R - 1}{R-1} \tag{9}$$

[8] The no arbitrage condition is due to Bhattacharya and Padilla (1996). Qi (1994) showed that if this condition does not hold, banks can improve on the market allocation by building an asset buffer that benefits all generations.

and all agents consume $\{(1-c)\sqrt{R}-c,\ (1-c)R-c\}$.

We have essentially the same bank as suggested by Bhattacharya and Padilla (1996), who show that in an economy without government intervention, the threat of arbitrage destroys the bank's superior risk sharing ability and that the best schedule it can offer its clients is the market allocation. However, because in our model the payouts are offered at the central location, and agents incur shoe-leather costs, the bank offers the agents a better allocation than the exchange mechanism:

Corollary *(banks versus markets)*

For positive transaction cost c, the bank allocation is Pareto superior to the one-periodic exchange allocation.

Proof: Agents who choose the *LT* strategy in the market economy receive an consumption schedule $\{(1-c)p^{*}-c,\ (1-c)R-c\}$, whereas bank depositors consume $\{(1-c)\sqrt{R}-c,\ (1-c)R-c\}$. The bank allocation is superior because $p^{*}<\sqrt{R}$. *Q.E.D.*

The reason for the superiority of banks lies of course in the difference in total travel. In the market economy, the total number of trips to the midpoint of the circle is $2+\theta(1-\lambda)$ per generation, while in the bank equilibrium only two trips are made. Interestingly, it is the impatient agents who bear these transaction costs. Patient agents are equally well off in both economies.

Notice that the equilibrium allocation in the bank economy avoids roll over arbitrage (Qi, 1994): a patient agent who withdraws and redeposits reaps $\left((1-c)\sqrt{R}-c\right)\sqrt{R}-c$, clearly less than her consumption when she would hold on to her deposit. Similarly, side trading á la Jacklin (1987) is dominated. A newborn who buys two year bonds and offers them for sale to patient bank depositors in case she becomes impatient, fetches at most $\sqrt{R}-\frac{c}{1-c}$ per bond, and hence consumes at most $(1-c)\sqrt{R}-2c$ < $(1-c)p^{*}-c$ if she becomes impatient, and $(1-c)R-c$ if she remains patient.[9]

5. Starting up the economy

Now that we have shown that in the unbounded OLG economy banks offer agents a higher expected utility, a natural question to ask is how the economy develops when there is a starting date and a first generation.

5.1. The exchange economy

Because in a hypothetical first year, there are no one-year bonds for sale, agents of the first generation will either store their entire endowment or play *LT*. In equilibrium they need to be indifferent so that we have:

$$\begin{aligned} & E[U]_{LT,0} \equiv \lambda U\left((1-c)p_1 - c\right) + (1-\lambda)U\left((1-c)R - c\right) \\ = \ & E[U]_{Store,0} \equiv \lambda U\left(1\right) + (1-\lambda)U\left((1-c)\frac{R}{p_1} - c\right) \end{aligned} \tag{10}$$

Clearly there exists a unique p_1 which solves this equality. We shall denote it p_1^*.

Similarly, the second and consecutive generations need to be indifferent between playing *LT* and either storing or playing *ST*:

$$E[U]_{LT} = \max\left(E[U]_{ST}, E[U]_{Store}\right) \tag{11}$$

Where $E[U]_{LT}$ and $E[U]_{ST}$ are defined in (1) and $E[U]_{Store}$ is the same as in the second line of (10) but with p_t instead of p_1. Clearly, also to (11) a unique solution exists. Since we have p_1^*, we can find $\boldsymbol{p}^* \equiv \{p_1^*, p_2^*, p_3^*, ...\}$ by recursively solving (11).

It can be easily seen from (10) that the first market price is smaller than the stationary price: $p_1^* < p^*$. Because on the second date of a bounded economy one-year bonds are relatively cheap, the second generation prefers playing *ST* over storing. The equation which solves (11) for the second generation then gives $p_2^* > p^*$, which implies that the

[9] The reservation price for the bond makes patient depositors indifferent between early withdrawing and purchasing bonds and sticking to their deposit. The consumption of a depositor who withdraws and buys bonds is $\frac{(1-c)\sqrt{R}-c}{p}R-c$, while his reservation consumption is $(1-c)R-c$.

third generation may prefer storing over buying one-year bonds. If it does, the relevant equation to find p_3^* is equal to (10), with p_3 instead of p_1. This means that *if* the third generation stores, all consecutive odd generations will, and we have a two-periodic equilibrium without full investment. If, however, the third generation does not store, and divides itself *ST* and *LT* players, then all consecutive odd generations will have *ST* players and the price will two-periodically converge to the stationary level, p^*.

We therefore have that in a bounded economy, one of three types of equilibrium obtains: (i) autarky, in which case the first generation, and hence consecutive generations, do not interact and recur to storing only, (ii) a perpetually cyclical equilibrium in which the first and consecutive odd generations store, and (iii) a cyclical start-up that converges to the steady state equilibrium characterized in proposition 1.

To gain insight into the different types of equilibrium, we compute the equilibria for different parameter values. Figure 1 shows how, for $\lambda = ½$ and $\gamma = 2½$, the equilibrium type depends on parameters c and R. Not surprisingly, for low values of R and high transaction costs c, it is more difficult to start up the economy.

We also see that there exists a wedge-shaped region where a full investment exchange equilibrium exists in the unbounded economy, but the economy with a starting date is stuck in autarky. Naturally, this is due to the fact that the first generation that invests faces more payoff risk than a stationary generation. It is not surprising then that the size of the wedge increases in γ.

There is also a region where a bounded exchange economy gets stuck in cyclicality. We find that the size of this triangular region sharply decreases in γ. For $\lambda = ½$, convergence always obtains for $\gamma > 3$, while for $\gamma < 2$ convergence only obtains for very high values of R. Risk aversion thus helps an economy to converge to stationarity. The reason is that the risk averse agents of the first generation store more, which helps to increase p_1^*, (and decrease p_2^*, etc.)

Finally, there is a parameter region where the bounded economy is better off with transaction costs than without. Above the dashed line in figure 1, the presence of a transaction costs enables the exchange economy to converge to a one-periodic price

p^*, while in the absence of transaction costs, the price process would be two-periodic $\{..,1,R,1,R,\ldots\}$. Although total intergenerational consumption in the economy with transaction costs is lower, and the second and consecutive even start-up generations enjoy lower utility than in the transaction cost free economy, the reduced cyclicality causes long run welfare to be higher. Naturally, the region where the economy is better off with transaction costs increases in risk aversion parameter γ.

Figure 2 illustrates the convergence process for an economy with $\lambda = ½$, $R = 2$, $c = 0.10$, and $\gamma = 4$. Not surprisingly, the speed to convergence increases in γ. To illustrate convergence we chose $\gamma = 4$ because for lower values of γ, convergence is significantly slower, so that either the convergence shape, or the zigzag pattern would be difficult to depict in a graph.

5.2. The intermediated economy

In section 4 we saw that in the stationary bank equilibrium, impatient one-year olds are paid partly from the proceeds of investments made one year before they were born. In the bounded economy, the first generation's impatient agents can only be paid from the deposits made by the second generation, which limits the amount of risk sharing. To reach the stationary level, the bank needs to build up a buffer during a start up phase. Qi (1994) and Bhattacharya *et al.* (1998) show that this can be achieved by 'holding out' early generations. A bank in the start up phase will invest as much as possible in the technology while offering early agents the utility they would obtain if they would be ostracized. To minimize the payouts in early periods, the bank sets r_1 as high as possible, and is constrained only by the rollover constraint $r_1^2 \geq r_2$. To meet its r_1 obligation on early dates, the bank may need to store. Appendix *C* gives the precise maximization problem of the bank in an economy with a starting date.

We find that for most parameter values, the bank can be started up within a single period, which means that it can already offer the second generation the stationary schedule $\{\sqrt{R}, R\}$. Not surprisingly, the length of the start-up phase decreases in the *R/c* ratio.

We also find that whenever a bank equilibrium exists, it can be started up. This is because there is no cyclicality in an intermediated economy, and the bank's allocation is superior to the market allocation due to transaction costs.

Figure 3 gives the number of start up periods as a function of R and c, if $\lambda = ½$ and $\gamma = 2½$. For comparison, the equilibrium regions in the exchange economy (see figure 1) are displayed in gray. As can be seen from the figure, there is a narrow parameter region for which there would be no trade in the market economy, but a bank equilibrium exists (and can be started up). Appendix *C* gives an illustration of such a bank.

6. Discussion

In the previous sections we characterized equilibrium allocations in simple models of pure exchange and intermediated economies, and showed that intermediaries can save on transaction costs by servicing a large diversified clientele of agents who face uncorrelated liquidity shocks and fixed transaction costs. In the following we conjecture that the key insights of our model continue to hold in more realistic models.

Under our first simplifying assumption agents can only become impatient on a single intermediary date. It is easy to see that if we assume more impatience dates before the technology pays off, our results would be stronger. This is because with more impatience dates, more rebalancing would take place in exchange economy. If there exists a continuum of potential impatience dates, agents continuously rebalance, slanting their portfolio to ever more short term securities as they get older (Vayanos, 1998; Vayanos and Vila, 1999). By catering to a continuum of agents, banks, mutual funds, pension plans or insurance companies do not need to rebalance their asset portfolios, as long as they continuously rebalance their clientele of depositors and investors.

A second unrealistic assumption concerns the gestation period, which we assumed to equal the maximum life span of the agents. Also this assumption can be altered without affecting the key insights of our model. Similarly, the presence of multiple production technologies, of different gestation lags, would not change the key insights of our analysis. It can be shown that in the pure exchange economy long term projects attract investment, and at least one production technology has a gestation lag longer than the minimum time until potential impatience, agents will rebalance their portfolios to shorter duration securities as they get older, and hence incur avoidable transaction costs.

This will also be the case if the long term assets are infinitely lived, as equities are. In this case, surviving agents in the exchange economy will, over their life cycle, rebalance their portfolio from predominately low dividend yield to high dividend yield stocks. Our model formalizes this intuition, and shows that, in the presence of transaction costs, agents align consumption duration with portfolio duration, because long term securities fetch lower prices on intermediary dates.[10]

A third simplifying assumption in our model is that of no payoff risk. A natural extension to our model is to consider a long term technology that offers a stochastic dividend. It can be shown that also in such a model newborn agents will buy a mix of long and short term securities and rebalance their portfolios over their life.[11] In the intermediated economy on the other hand, mutual funds will offer their investors state dependent redemption schedules (based on *net asset values*), that *ex-ante* offer them higher expected utility than the market economy due to the avoidance of excessive dividend reinvestment and maturity rebalancing.

Finally, we assumed that agents' liquidity shocks are independent, and that there is no aggregate liquidity risk. However, it can be shown that our results continue to hold if agents' liquidity shocks are less than perfectly correlated, and aggregate risk exists. In this case, banks will hold cash buffers alongside a portfolio of long term and short term assets. The optimal cash buffer depends mostly on the covariances of depositors' liquidity shocks. Edgeworth (1888) first discussed the bank's cash-inventory problem in the presence of stochastic net withdrawals.

Also for mutual funds, redemptions and contributions are unlikely to be perfectly correlated, so that they too hold cash. A cash buffer reduces the need to engage in uninformed but costly liquidity trades but also reduces the fund's expected return. Not surprisingly, the optimal cash balance is an interior optimum, so that mutual funds engage in significant liquidity trade on their shareholders behalf. Edelen (1999)

[10] This intuition was first described in (Thaler and Shefrin, 1981), who attributed the observed consumption/cashflow duration matching to self-control. Notice that in our model intermediaries do not have a comparative advantage in life cycle rebalancing driven by tax reasons (Miller, 1977) or risk preferences. Clearly, the suggested generational diversification is incompatible with age dependent tax- or risk preferences. In any case, whether life cycle risk rebalancing is optimal is still an open question. See Merton (1969) and Samuelson (1969) among others.

[11] Maturity rebalancing in the face of liquidity risk is studied by Eisfeldt (2007).

estimates the associated costs of liquidity driven trades for mutual funds to be 1.5 – 2% per annum. Chordia (1996) argues that to reduce forced liquidity trades triggered by shareholder redemptions, funds charge upfront fees, called *loads*. Yan (2006) develops a model to solve a mutual fund's cash inventory problem, and finds evidence that fund cash holdings are determined by a trade off between expected variance of the *flow* (contributions minus redemptions) and transaction costs in the asset markets.

In our model we assume away aggregate risk so as to best illustrate the main trust of this paper, which is that through intergenerational pooling of shareholders and depositors, intermediaries are able to cancel rebalancing transaction that would occur in the free market. Correlated liquidity shocks and aggregate flow risk obfuscate this advantage but do not eliminate it.

7. Summary and conclusions

In this paper we formalize the widely held intuition that the comparative advantage of financial intermediaries lies in their ability to internalize transaction costs by holding buffers of assets with different durations.

Our model demonstrates that centrally located intergenerational banks can economize on transaction costs that would otherwise be incurred due to life cycle duration rebalancing. In a simple model with spatial separation and shoe leather costs in which agents face consumption risk, excessive trade obtains in the pure exchange economy because equilibrium prescribes late dying agents to reinvest dividends and incur avoidable transaction costs. We show that an intermediary that offers demand deposits can internalize the rebalancing trades, and improve welfare.

Naturally, our shoe leather cost aims to capture the plethora of other types of transaction processing costs, such as communication, search, administration, and even paying attention. It is by netting out these tasks that financial intermediaries add value, while offering their clients immediacy. Our financial intermediary can be interpreted as a bank with automatic teller machines, insurance companies, or an open end mutual fund.

Appendix A: Proofs

Proof of proposition 1

We first show that there exists a unique one-periodic equilibrium. Substituting p for p_t and p_{t+1} in equation (1) and rewriting gives:

$$\lambda\left\{U\big((1-c)p-c\big)-U\left((1-c)\frac{R}{p}-c\right)\right\}+(1-\lambda)\left\{U\big((1-c)R-c\big)-U\left((1-c)\frac{R^2}{p^2}-c\frac{R}{p}-c\right)\right\}=0 \qquad \text{(A1)}$$

Differentiating shows that the lhs of (A1) is continuously increasing in p for all relevant parameter values. The derivative of the first term is positive. The derivative of the second term of (A1) is:

$$(1-\lambda)\left(2\frac{(1-c)R^2}{p^3}-\frac{cR}{p^2}\right)U'\left((1-c)\frac{R^2}{p^2}-c\frac{R}{p}-c\right) \qquad \text{(A2)}$$

which is positive unless $p>\frac{2(1-c)R}{c}$. The latter inequality cannot hold in equilibrium because if it would, nobody would buy one-year bonds.

Also, the lhs of (A1) is positive if $p=\sqrt{R}$ and negative for $p=\frac{\sqrt{c^2+4R(1-c)^2}-c}{2(1-c)}$.

This proves that there is a unique one periodic equilibrium, and that it lies between abovementioned values.

We still need to prove that no two-periodic equilibrium exists. A potential two-periodic equilibrium solves the following system of equations:

$$\lambda\left\{U\big((1-c)p_j-c\big)-U\left((1-c)\frac{R}{p_i}-c\right)\right\}+(1-\lambda)\left\{U\big((1-c)R-c\big)-U\left((1-c)\frac{R^2}{p_ip_j}-c\frac{R}{p_j}-c\right)\right\}=0 \quad \text{(A3)}$$

And

$$\lambda\left\{U\big((1-c)p_i-c\big)-U\left((1-c)\frac{R}{p_j}-c\right)\right\}+(1-\lambda)\left\{U\big((1-c)R-c\big)-U\left((1-c)\frac{R^2}{p_jp_i}-c\frac{R}{p_i}-c\right)\right\}=0 \quad \text{(A4)}$$

Assume $p_i > p^*$, then clearly, from both (A3) and (A4), we must have $p_j < p^*$.
Consider the lhs's of (A3) and (A4) functions of p_j, for given $p_i > p^*$. The functions cross at p_i, and increase continuously. The derivatives are:

$$\frac{\partial lhs(\text{A3})}{\partial p_j}=\lambda(1-c)U'\big((1-c)p_j-c\big)+(1-\lambda)\left(\frac{(1-c)R^2}{p_ip_j^{\,2}}-c\frac{R}{p_j^{\,2}}\right)U'\left((1-c)\frac{R^2}{p_ip_j}-c\frac{R}{p_j}-c\right) \qquad \text{(A5)}$$

$$\frac{\partial lhs(\text{A4})}{\partial p_j}=\lambda\frac{(1-c)R}{p_j^{\;2}}U'\left((1-c)\frac{R}{p_j}-c\right)+(1-\lambda)\frac{(1-c)R^2}{p_i p_j^{\;2}}U'\left((1-c)\frac{R^2}{p_j p_i}-c\frac{R}{p_i}-c\right) \qquad \text{(A6)}$$

Clearly $\frac{\partial lhs(\text{A4})}{\partial p_j}>\frac{\partial lhs(\text{A3})}{\partial p_j}\ \forall p_j<p^*$, which proves that for any $p_i > p^*$, no p_j exists that simultaneously satisfies equations (A3) and (A4).

The periodic investment $\theta^* = \frac{(1-c)p^*+(1-\lambda)\left(R-c(R+p^*)\right)}{(1-c)(\lambda p^*+1)p^*+(1-\lambda)\left(R-c(R+p^*)\right)}$, follows from rewriting the market clearing condition (2).

Proof of Proposition 2

The optimal bank allocation follows immediately from solving (5) subject to (7) and (8). The expression for the periodic investment, $y^*=(1-c)\frac{\lambda\sqrt{R}+(1-\lambda)R-1}{R-1}$, follows from rewriting (6).

Appendix B: Variable Transaction Costs

We now investigate the exchange economy equilibrium in the case of variable transaction costs. In this case the transaction costs are cV, where V is the value of the goods transported to the circle's midpoint, at the midpoint. This means that newborns can spend (1-c) on either one-year or two year bonds, and a patient agent who sells x one year bonds for p, can spend $\frac{x}{p}(1-c)$ on buying new bonds. In such an economy, newborns will buy both types of bonds. If they become impatient they sell the two-year bonds before they mature (as one-year bonds), and if they stay patient they roll over the one-year bonds. Denote y_t the amount that newborns invest in two year bonds. Their objective function is then:

$$\max_{y_t}\ \lambda U\left(\left((1-c-y_t)\frac{R}{p_t}+y_t p_{t+1}\right)(1-c)\right)+(1-\lambda)U\left(\left((1-c-y_t)(1-c)\frac{R^2}{p_t p_{t+1}}+y_t R\right)(1-c)\right) \qquad \text{(B1)}$$

The first order condition of this maximization problem is:

$$\left(\frac{R}{p_t}-p_{t+1}\right)\lambda U'\left(C_1^*\right)=\left(R-\frac{R^2(1-c)}{p_t p_{t+1}}\right)(1-\lambda)U'\left(C_2^*\right) \qquad \text{(B2)}$$

In a one-periodic stationary equilibrium both (B2) with $p_t = p_{t+1} = p^*$ and $y_t = y^*$ must hold together with the market clearing condition:

$$p^* = \frac{(1-y^*-c)+(1-\lambda)(1-y^*-c)\frac{R}{p^*}(1-c)}{y^*\lambda} \tag{B3}$$

A unique solution exists to these two equations. This can be seen by rewriting (B3):

$$y^* = \frac{(1-c)p^* + (1-\lambda)(1-c)^2 R}{p^{*2}\lambda + p^* + (1-\lambda)R(1-c)} \tag{B4}$$

Substituting this in (B2), and rewriting gives:

$$\left(R-p^{*2}\right)\lambda U'\left(C_1^*\right) = \frac{R}{p^{*2}}\left(p^{*2}-R(1-c)\right)(1-\lambda)U'\left(C_2^*\right) \tag{B5}$$

With

$$C_1^* = p^*(1-c)^2\frac{\lambda R + p^* + (1-\lambda)(1-c)R}{p^{*2}\lambda + p^* + (1-\lambda)(1-c)R} \tag{B6}$$

And

$$C_2^* = (1-c)^2 R\frac{p^* + (1-c)R}{p^{*2}\lambda + p^* + (1-\lambda)R(1-c)} \tag{B7}$$

It can be shown that for $p^* = \sqrt{R}$, the rhs of (B5) is positive (and clearly the lhs is zero). As we let p^* decrease, the rhs decreases, while the lhs increases so that a unique p^* exists that solves (B2) and (B3). This proves that also with variable transactions costs, and pure strategies, a unique *noncyclical* market equilibrium exists. Because the derivation is slightly more complicated and less intuitive, we opted for modelling the transaction costs be fixed.

Appendix *C*: Starting up the bank

First we observe that to minimize the time to stationarity, banks will maximize investment in the long term asset in early years, until the stationary investment y^* is reached. In the following denote y_t the bank's investment at t, s_t the investment in the storage technology and $\{r_1^t, r_2^t\}$ the payoffs distributed on date t. Let T be the time when $y_t = y^* = (1-c)\frac{\lambda\sqrt{R}+(1-\lambda)R-1}{R-1}$. Clearly at T, the bank can offer a schedule

$\{\sqrt{R}, R\}$ per deposited unit of consumption good. The maximization problem during the start-up phase is then:

$$\min_{\{r_1^t, r_2^t\}_{t=1..T}} T \tag{C1}$$

Subject to: $$y_0 = s_0 = r_2^1 = 0 \tag{C2}$$

$$(1-c)\left(\lambda r_1^t + (1-\lambda) r_2^t\right) + y_t + s_t \leq 1 - c + R y_{t-2} + s_{t-1} \quad \forall \ t > 0 \quad , \tag{C3}$$

$$\begin{aligned} &\lambda U\left((1-c) r_1^t - c\right) + (1-\lambda) U\left((1-c) r_2^{t+1} - c\right) \geq \\ &\qquad \max\left(U(1), U\left((1-c)(\lambda + (1-\lambda)R) - c\right)\right) \quad \forall \, t > 0 \end{aligned} \quad , \tag{C4}$$

and $$r_1^t r_1^{t+1} \geq r_2^{t+1} \qquad \forall \ t > 0 \tag{C5}$$

This maximization problem can be solved by letting the budget constraint (C3), participation constraint (C4) and roll-over constraint (C5) bind until *T*.

Table 1 illustrates the start up phase of such a bank in an economy where an exchange will not open: If $R = 1.33$, $\gamma = 2.5$ and $c = 0.10$. To see that an exchange equilibrium does not exist, solve (1) with the above parameter values. The solution gives $p^* = 1.1348$. If we now compute the agents' expected utility in with this p^*, we find it to be less than $U(1)$, the expected utility in autarky.[12] Table 1 shows that in a world with the above parameter values a bank equilibrium *does* exist, and can be started up in five years.

[12] With $U(C) = -\frac{2}{3} C^{-\frac{3}{2}}$, we find $E[U]_{LT} = E[U]_{ST} = -0.6709$, while $U(1) = -0.6667$.

References

Azariadis, C. (1981) Self fulfilling prophecies, *Journal of Economic Theory*, 25, 380-396.

Alonso, I. (1996) On avoiding bank runs, *Journal of Monetary Economics*, 37, 73-87.

Aiyagari, S.R., and Gertler, M. (1991) Asset returns with transaction costs and uninsured individual risk: a stage III exercise, *Journal of Monetary Economics*, 27, 309-331.

Baumol W.J. (1952) The transactions demand for cash: an inventory theoretic approach, *Quarterly Journal of Economics*, 66, 545-556.

Benston, G.J. and Smith, Jr., C.W. (1976) A transaction cost approach to the theory of financial intermediation, *Journal of Finance*, 31, 215-231.

Bhattacharya, S., Padilla, A.J. (1996) Dynamic Banking: A reconsideration, *Review of Financial Studies* 9, 1003-1031.

Bhattacharya, S., Fulghieri, P., and Rovelli, R. (1998) Financial intermediation versus stock markets in a dynamic intertemporal model *Journal of Institutional and Theoretical Economics*, 154, 291-319.

Boyd, J., and Prescott, E. (1986) Financial intermediary coalitions, *Journal of Economic Theory*, 38, 211-232.

Broecker, T. (1990) Credit worthyness tests and interbank competition, *Econometrica*, 58, 429-452.

Cherkes, M., Sagi, J., and Stanton, R. (2008) A liquidity-based theory of closed end funds, *Review of Financial Studies*, 22, 257-297.

Chordia, T. (1996) The structure of mutual fund charges, *Journal of Financial Economics*, 41, 3-39.

Constantinides, G. M. (1986) Capital market equilibrium with transaction costs, *Journal of Political Economy*, 94, 842-862.

Diamond, D.W., Dybvig, P.H. (1983) Bank runs, Deposit insurance and liquidity, *Journal of Political Economy* 91, 401-419.

Diamond, D.W. (1997) Liquidity, banks, and markets, *Journal of Political Economy* 105, 928-956.

Duffie, D., and Sun, Y. (2007) Existence of independent random matching, *The Annals of Applied* Probability, 17, 386-419.

Dwyer, G.P., Samartín, M. (2008) Theoretical explanations of why banks promise to pay par on demand, *Journal of Financial Stability*, forthcoming.

Edelen, R. M. (1999) Investor flows and the assessed performance of open-end mutual funds, *Journal of Financial Economics*, 53, 439-466.

Edgeworth, F.Y. (1888) The mathematical theory of banking, *Journal of the Royal Statistical Society*, LI, 113-127.

Eisfeldt, A.L. (2007) Smoothing with liquid and illiquid assets, *Journal of Monetary Economics*, 54, 1572-1586.

Freixas, X., and Rochet, Rochet, J.-C. (1997) Micro-economics of banking, MIT Press, Cambridge, Massachusetts

Fulghieri, P., Rovelli, R. (1998) Capital markets, financial intermediaries, and liquidity supply, *Journal of Banking and Finance* 22, 1157-1179.

Heaton, J. and Lucas, D.J. (1996) Evaluating the effects of incomplete markets on risk sharing and asset pricing, *Journal of Political Economy*, 104, 443-487.

Hellwig, M.F. (1994) Liquidity provision, banking, and the allocation of interest rate risk, *European Economic Review*, 38, 1363-1389.

Huang, M. (2003) Liquidity shocks and equilibrium liquidity premia, *Journal of Economic Theory*, 109, 104-129.

Jacklin, C.J. (1987) Demand deposits, trading restrictions, and risk sharing, in: E. Prescott and N. Wallace (eds) ``Contractual Arrangements for Intertemporal Trade", University of Minnesota Press, Minneapolis, 26-47.

Lucas, Jr., R.E. (2000) Inflation and welfare, *Econometrica*, 68, 247-274.

Merton, R.C. (1969) Lifetime portfolio selection under uncertainty: The continuous-time case. *Review of Economics and Statistics*, 51, 247-257.

Miller, M. (1977) Debt and taxes, *Journal of Finance*, 32, 261-75.

Orr, D. and W.G. Mellon (1961) Stochastic Reserve Losses and Expansion of Bank Credit, *American Economic Review*, 51, 614-623.

Qi, J. (1994) Bank liquidity and stability in an overlapping generations model, *Review of Financial Studies* 7, 389-417.

Qian, Y., John, K., and John, T.A. (2004) Financial system design and liquidity provision by banks and markets in a dynamic economy, *Journal of International Money and Finance*, 23, 385-403.

Samartín, M. (2001) Banks increase welfare, *Financial markets, Institutions & Instruments* 10, 203-233.

Samuelson, P. (1969) Lifetime portfolio selection by dynamic stochastic programming, *Review of Economics and Statistics*, 51, 239-246.

Shefrin, H.M. and Thaler, R.H. (1981) An economic theory of self-control, *Journal of Political Economy*, 89, 392-406.

Tobin, J. (1956) The interest-elasticity of the transactions demand for cash, *Review of Economics and Statistics*, 38, 241-247.

Vayanos, D. (1998) Transaction costs and asset prices: A dynamic equilibrium model, *Review of Financial Studies*, 11, 1-58.

Vayanos, D., Vila, J.-L. (1999) Equilibrium interest rate and liquidity premium with transaction costs, *Economic Theory*, 13, 509-539.

von Thadden, E.L. (1998) Intermediated versus direct investment: optimal liquidity provision and dynamic incentive compatibility, *Journal of Financial Intermediation* 7, 177-197.

Yan, X.S. (2006) The determinants and implications of mutual fund cash holdings: theory and Evidence, *Financial Management*, 35, 67-91.

Figure 1: Equilibria in a Bounded Exchange Economy

We compute the equilibrium price prices in a bounded OLG Diamond Dybvig economy with $\lambda = ½$ and where agents have utility functions given by $U(C) = -\frac{2}{3}C^{-\frac{3}{2}}$ (CRRA parameter $\gamma = 2.5$), for varying fixed transaction costs c, and asset payoffs R. The areas between the black lines give the equilibrium type. Above the gray line a one-periodic stationary equilibrium exists in the unbounded economy. The dashed gray line denotes the area where overall welfare is higher in the bounded exchange economy with transaction costs than without transaction costs.

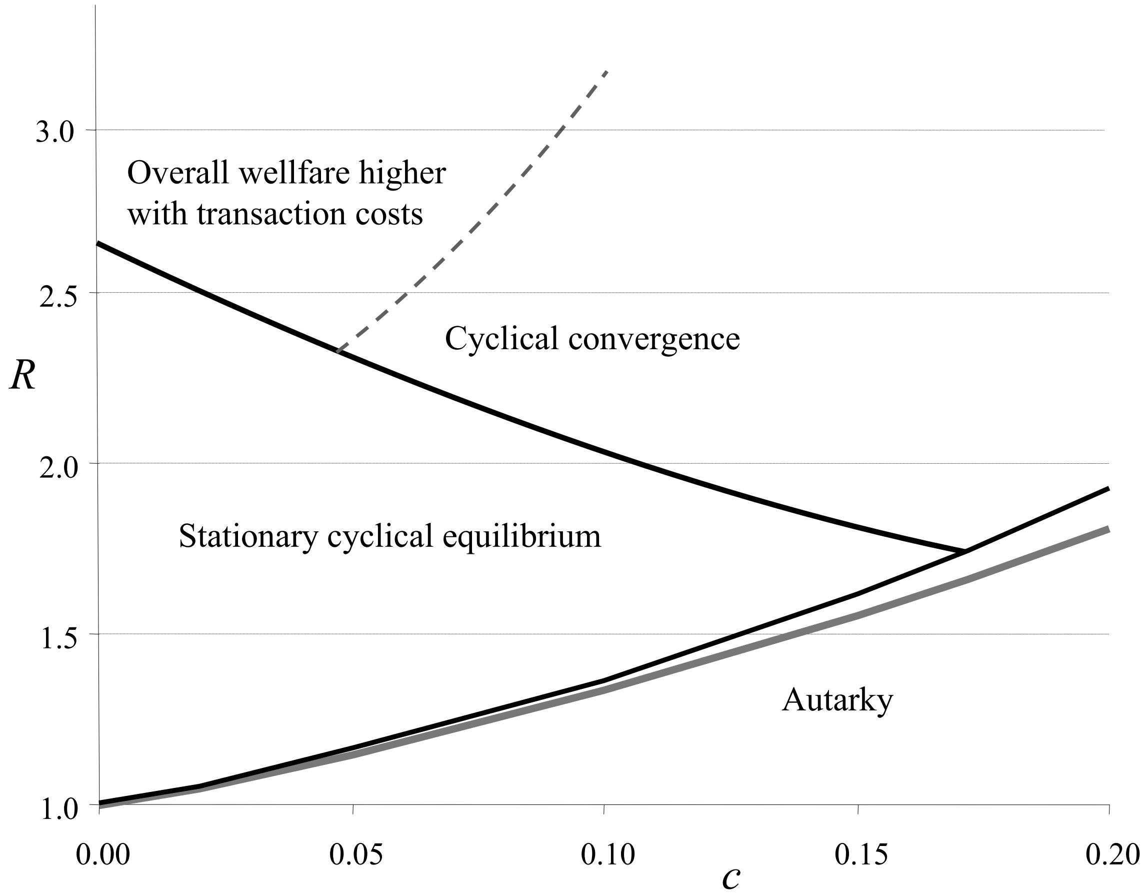

Figure 2: Market Equilibrium Start-up in a Bounded Economy

We compute the equilibrium prices in a bounded OLG Diamond Dybvig economy with transaction costs with $\lambda = ½$, $R = 2$, $c = 0.10$, and agents' utility function is given by $U(C) = -\frac{1}{3}C^{-3}$ (CRRA parameter $\gamma = 4$). The graph gives the equilibrium price of one year bonds as a function of time.

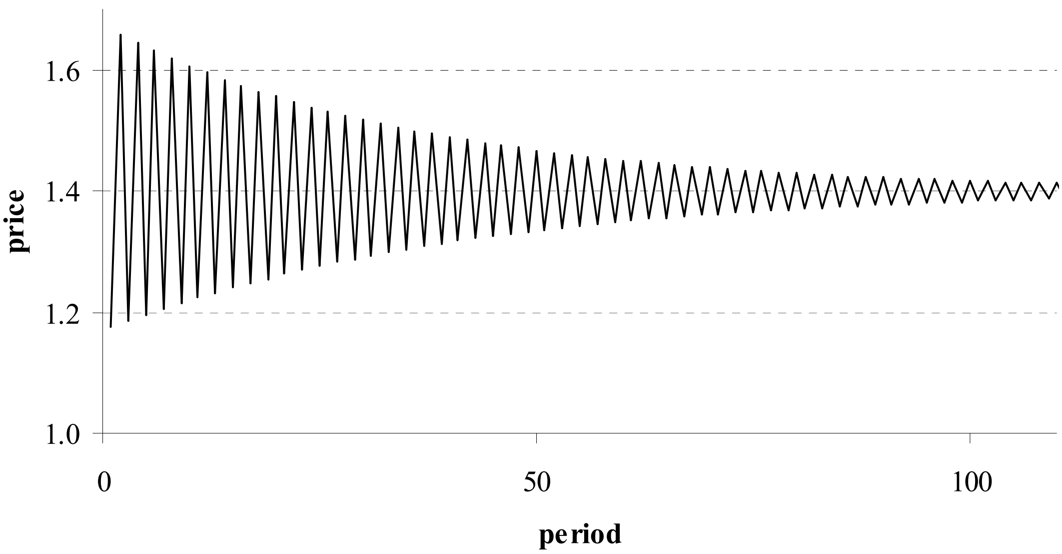

Figure 3: Start-up Bank Equilibria in a Bounded Economy

We compute the bank allocations in a bounded OLG Diamond Dybvig economy with $\lambda = ½$, $R = 2$, $c = 0.10$, and agents' utility function is given by $U(C) = -\frac{2}{3}C^{-\frac{3}{2}}$ (CRRA parameter $\gamma = 2.5$), for varying c and R. The areas between the black lines give the minimum number of periods it takes to start up the stationary bank equilibrium. Below the bottom black line, no bank equilibrium exists. In the region immediately above this line the bank's start up phase is at least three periods. In the background are the demarcation lines - in gray - for the exchange economy equilibria.

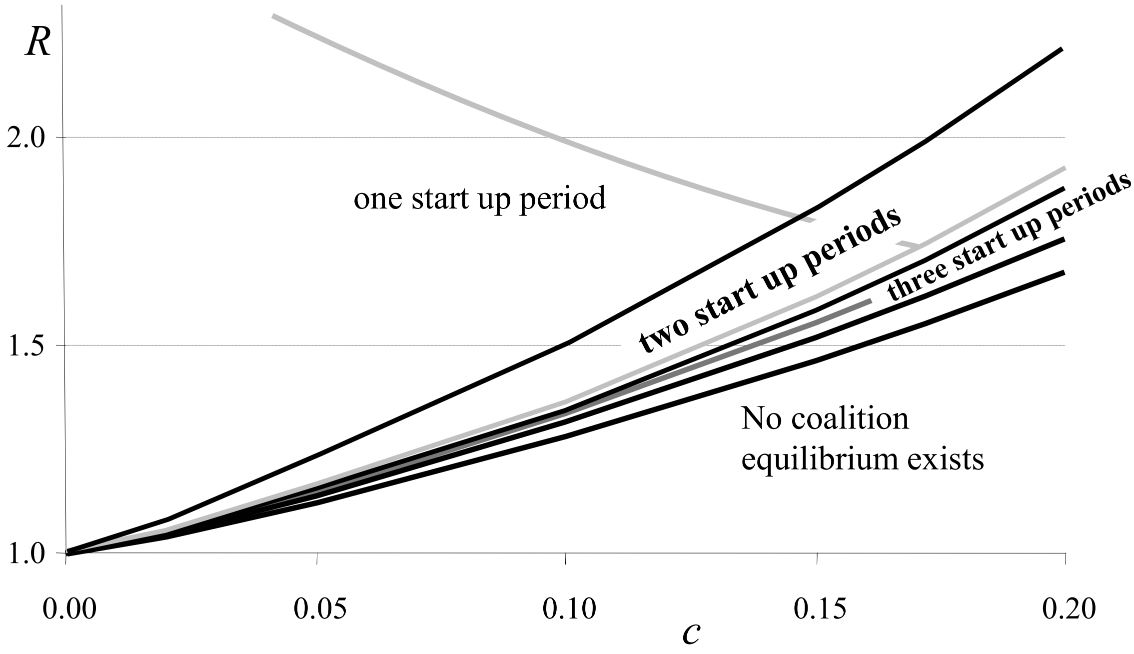

Table 1: Example of a Bank Start-up

For an economy with λ = ½, *R* = 1.33, γ = 2.5 and *c* = 0.10, we compute the payoff schedule for a competitive bank. To build an asset buffer it offers early generations their reservation utility, obtained in Autarky. The first two columns give the promised payoffs, per unit of consumption good deposited. The third-last column equals the second column of the previous row multiplied by λ(1-*c*). The inflow from investment equals the amount stored in the previous period plus the amount invested two periods earlier multiplied by *R*.

				Goods Inflow			Goods Outflow				
t	r_1	r_2	E[U]	From newborns	From investment	Total in	Invest	Store	Out to impatients	Out to patients	Total out
0	1.1462	1.3139	-0.6667	0.9		0.9000	0.6256	0.2744			0.9000
1	1.1462	1.3139	-0.6667	0.9	0.2744	1.1744	0.6313	0.0273	0.5158		1.1744
2	1.1462	1.3139	-0.6667	0.9	0.8656	1.7656	0.6389	0.0197	0.5158	0.5912	1.7656
3	1.1462	1.3139	-0.6667	0.9	0.8656	1.7656	0.6490	0.0095	0.5158	0.5912	1.7656
4	1.1553	1.3400	-0.6525	0.9	0.8656	1.7656	0.6586		0.5158	0.5912	1.7656
5	**1.1576**	**1.3400**	-0.6513	0.9	0.8697	1.7697	0.6586		0.5199	0.5912	1.7697
6	**1.1576**	**1.3400**	-0.6513	0.9	0.8825	1.7825	0.6586		0.5209	0.6030	1.7825
7	**1.1576**	**1.3400**	-0.6513	0.9	0.8825	1.7825	0.6586		0.5209	0.6030	1.7825
8				0.9	0.8825	1.7825	0.6586		0.5209	0.6030	1.7825