Risk Sharing in a World with Processing Costs: Trading versus Banking

By Jos van Bommel

We analyze the bank versus exchange problem in a Diamond Dybvig (1983) economy with exogenous transaction processing costs. We find that processing costs in the market enables the bank to overcome the *side trade* threat (Jacklin (1987)) and offer some desirable liquidity insurance. Moreover, in the bank equilibrium processing costs are proportional to consumption, while in the market economy early and late consumers incur equal costs. These two effects explain that for a given level of *aggregate* processing costs, the bank economy is superior. On the other hand, the number of transactions in the bank economy is larger. It is for this reason that if processing costs are proportional to transaction value, and independent of the mechanism used, the exchange economy is superior.

I. INTRODUCTION

One of the primary explanations for the existence of banks is that they provide a mechanism to share liquidity risk. Diamond and Dybvig (1983) ('DD') showed that if agents can invest in long term projects that return, per unit invested, a certain cashflow $R > 1$ at t_2, but are subject to an unverifiable risk of becoming 'impatient' at interim date t_1, they may form coalitions that collect endowments, invest some of the proceeds, and hold cash to distribute to those who become impatient early. Such a DD bank returns, per unit deposited, at the option of the depositor, either r_1 at t_1 or r_2 at t_2. Diamond and Dybvig showed that if agents are sufficiently risk averse, there exists a schedule $\{r_1, r_2\}$ with $1 < r_1 < r_2 < R$, that makes agents *ex-ante* better off than the schedule $\{1, R\}$ which would be obtained in a pure exchange economy.

The DD bank has been studied extensively due to the fragility of the suggested arrangement: If patient agents somehow believe that other patient agents are withdrawing at t_1, they too will withdraw at t_1, thereby causing a 'bank run'. Diamond and Dybvig themselves suggest two remedies against this bad equilibrium: suspension of convertibility and deposit insurance.[1] There is now an extensive literature that uses the DD-model to investigate financial fragility.[2]

Jos van Bommel, Saïd Business School, Park End Street, Oxford, OX1 1HP, UK. Tel: +44 (0)1865 288905; jos.vanbommel@sbs.ox.ac.uk

[1] Under suspension of convertibility, the bank stops paying when at t_1 the projected amount (proportion of impatient agents $\times$ r_1) has been withdrawn. This takes away the patient agents' worries that there will not be enough in the bank at t_2, and removes their incentive to 'run'. Deposit insurance refers to the situation where governments make up the shortfall in case of a bank run and *ex-post* tax the agents.
[2] See Dowd (1992) or chapter seven of Freixas and Rochet (1997) for an overview of this literature.

Instead of studying bank runs, this paper investigates how reasonable the by Diamond and Dybvig suggested liquidity insurance arrangement is in the first place. Jacklin (1987) showed that the DD bank does not constitute a subgame perfect Nash equilibrium in the presence of a market. The reason is that if a bank offers $\{r_1, r_2\}$ with $1 < r_1 < r_2 < R$, agents can increase their payoffs by bypassing the bank and investing their entire endowment in the long term project. If they become impatient at t_1, they sell their claims to patient depositors. The suggested deviation is known as *side trading*.[3] Because patient depositors will be willing to pay up to r_1R/r_2 for a long term claim, side traders could potentially receive a schedule $\{r_1R/r_2,\ R\} \gg \{r_1,\ r_2\}$.

This arbitrage opportunity is also discussed in Bhattacharya and Gale (1987), who use a DD model to consider the interplay between banks and central banks with the former facing liquidity risk. As in DD, the existence of an interbank market causes intermediaries to overinvest in the illiquid asset. Central banks can improve on the allocation only if they monitor their member banks' liquidity to avoid side trading. Similarly, Haubrich and King (1990) consider a model in which agents, instead of having uncertain intertemporal consumption preferences, face income shocks. Also in this scenario, banks cannot provide liquidity insurance in the presence of a contemporaneous exchange mechanism. Recently, Allen and Gale (2004), who study a DD model with uncertain asset payoffs and aggregate preferences, corroborate that if financial intermediaries are redundant if agents can trade assets on their own.[4]

To overcome the Jacklin critique, Wallace (1988) assumes that agents are spatially separated, making trade impractical. Depositors are represented as campers who may wake up hungry at night, but do not want to wake up fellow campers to trade with them. A food dispenser (interpreted as a cash machine) then increases welfare. In a similar vein, Diamond (1997) assumes that agents, apart from running a liquidity risk, may also be unable to trade. A deposit taking bank then increases welfare both by centralizing the short term investment, and by offering some liquidity insurance. Freeman (1988) argues that trade between agents require buyers of second hand capital to conduct costly due diligence, because sellers may not be around when the assets pay off. Intermediaries are not subject to these costs because they are infinitely lived, and can be sued in case of nonpayment. We believe that neither solution is a fully satisfactory description of our

A recent contribution comes from Goldstein and Pauzner (2002), who study the trade off between liquidity insurance and bank vulnerability.

[3] von Thadden (1999) showed that, apart from side *trading*, also side *investing* can constrain the bank's ability to provide liquidity insurance. If there is a security that returns $R_1 > 1$ over one period (or the asset itself can be liquidated for R_1 after one period), the bank is naturally constraint by $r_2 < r_1R_1$, even if agents are isolated.

[4] Bhattacharya and Padilla (1996) extend Jacklin's concerns to an overlapping generations setting. They show that in the presence of an intergenerational market, banks not only face the threat of side trading agents, but also need to protect themselves against arbitraging colleague banks: if banks offer too generous payoffs to depositors, long lived banks are inclined to open deposits with each other. Qi (1994) and Fulghieri and Rovelli (1998) show however that if interbank lending can be prohibited, banks dominate exchanges in OLG settings.

mostly anonymous and liquid financial markets. In this paper we provide a novel argument to overcome the Jacklin critique and explain the existence of liquidity providing banks in the presence of markets: transaction processing costs.

Benston and Smith (1976) argued that "the *raison d'être* for the [financial intermediation] industry is the existence of transactions costs". Their main argument is that banks enjoy economies of scale, scope, reputation, and networks in coping with transaction costs. A much investigated source of transaction costs are adverse selection and moral hazard. In exchanges such externalities lead to positive bid/ask spreads and limited market depth. In banking, information asymmetries require costly screening and monitoring. In either case, the resulting inefficiencies can be interpreted as transaction costs. Boot and Thakor (1997) acknowledge frictions in both mechanisms, and go on to argue that pure information asymmetries are better dealt with in a market setting, where a competitive price efficiently aggregates information regarding the project's value. Pure moral hazard on the other hand is better dealt with in a bank economy, where monitoring can be delegated and coordinated.

In this paper we study an economy without adverse selection or moral hazard, and justify transaction costs using a more mundane explanation: handling and administration costs, the expenses that are incurred by intermediaries to pay for fixed assets (buildings, automatic teller machines, communication technology, etc.), and meet operating expenses (salaries, electricity, rent, postage, etc.). Although it is generally accepted that the processing of financial transactions is a major task of banks, there is no study that investigates how processing costs affect the liquidity insurance and delegated monitoring tasks of financial intermediaries. This paper provides a preliminary investigation in this until now unexplored area.

Notice that we do not *explain* and *compare* processing costs in market and bank economies, but instead treat processing costs as exogenous. The aim of this paper is to set up a horse race between the bank and exchange economies, to find which mechanism prevails for different processing cost hurdles.[5]

We find that, in our adjusted DD model, a main advantage of the market mechanism is the significantly lower transaction volume. The reason is that in the bank economy all goods flow through the bank which stores uninvested consumption goods centrally. In the exchange economy on the other hand, agents store uninvested funds at home, and exchange only part of their assets.[6] We find that if processing costs as a proportion of value transacted are the same for both economies, the market economy is the only feasible equilibrium.

However, the assumption of mechanism-independent proportional processing costs may not be realistic. Bank transactions and market transactions are of a completely different nature. Demandable debt securities can be administered by an automated teller machine, whereas in the market more sophisticated

[5] In a similar vein, Hoerova (2005) compares welfare in bank, exchange, and hybrid economies for different probabilities of panic-driven bank runs.

[6] In the standard DD model the transaction value that flows through the bank equilibrium is $1 + \lambda r_1 + (1 - \lambda) r_2$. The Jacklin-robust bank has $r_1 = 1$, $r_2 = R$, so that the total transaction value is $1 + R - \lambda(R - 1)$. The transaction value in exchange equilibrium is only $2\lambda(1 - \lambda)$.

mechanisms are needed to match supply and demand. We consider different proportional processing costs for banks and exchanges.

Processing costs may not be truly *dead weight*. Financial contract processing provides many jobs in the economy. Therefore we also look at the cases where the *aggregate* processing costs are the same in both economies. We find that in this comparison the bank economy offers a Pareto superior allocation. Moreover, there exists a significant region in the processing cost space where the bank offers the Pareto-superior allocation, while incurring higher aggregate processing costs.

The superiority of banks in a world with given aggregate processing costs, derives from two main advantages. First, processing costs make *side trading* less attractive, enabling banks to offer some liquidity insurance by narrowing the payoff gap between early and late withdrawals.[7] Second, in a bank economy transaction value is proportional to consumption, while in an exchange economy the transaction value is equal for early 'poor' consumers and late 'rich' consumers.

The remainder of this paper is organized as follows: In the next section we describe the model. In Section III we derive the equilibrium in the pure exchange economy and in Section IV we characterize the bank risk sharing arrangement. Section V compares the two mechanisms. Section VI concludes and makes conjectures on extensions of the model. Proofs are in the appendix.

II. THE MODEL

Our model closely follows Diamond and Dybvig (1983): there is a continuum of utility maximizing agents who interact on three dates. At t_0 agents are born with an endowment of one unit of a non-perishable consumption good. Agents may become impatient, with probability λ, immediately before t_1, in which case they die soon after t_1. If agents remain patient, they die soon after t_2. Types are not verifiable, so that no contingent claims can be traded. Agents enjoy decreasing marginal utility over total lifetime consumption and have no time preferences apart from their limited life. For tractability we further assume that a utility function with a constant relative risk aversion coefficient: $U(C_1, C_2) = \frac{1}{1-\lambda}(C_1 + C_2)^{1-\gamma}$.

Agents can invest in a long term project that pays, per unit of consumption good invested at t_0, R units at t_2. The payoff R is net of processing cost in the primary market. The project is illiquid in the sense that it is never economical to interrupt the project and recover a liquidation value. The only means of liquidating a project is by selling it as a going concern in the secondary market. Agents also have access to a storage technology.

To share liquidity risk, agents may trade one-period project-backed bonds with face value R for contemporaneous goods in the secondary market which opens at t_1. The price (in goods per bond), at which exchange takes place is denoted p. The

[7] If agents have a coefficient of relative risk aversion less than one, the bank would aim to *widen* the gap between early and late withdrawals.

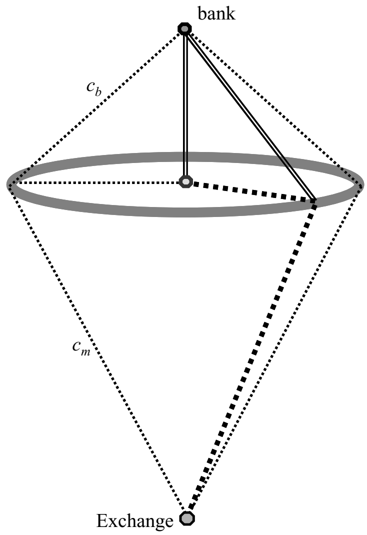

Figure 1: Geographic Representation of Processing Costs.

This figure visualizes the processing costs in a three dimensional map. The production technology is located in the centre, agents live on the grey circle. In a bank economy consumption goods follow the path indicated by the double line. In a market economy, goods move along the dotted path. Processing costs are proportional to the distances in the figure.

exchange is subject to a proportional processing fee c_m contemporaneous goods (cash) for both buyers and sellers: selling agents receive $(1 - c_m)p$ per bond sold, while buying agents spend $(1 + c_m)p$ per bond bought.

Alternatively, agents may form a coalition. Such an institution (the DD bank) takes deposits at t_0, invests part of it in the long term asset, and offers its clients, per unit of consumption good deposited, a demandable debt security that pays either r_1 at t_1 or r_2 at t_2, at the option of the depositor. The coalition is subject to a cash processing cost of c_b per dollar deposited or withdrawn: depositors who bring one unit to the bank and withdraw at t, thus consume $(1 - c_b)^2 r_t$.

One way to interpret the processing costs is the spillage caused by a "leaking bucket" which is carried over physical distances. Figure 1 shows the three-dimensional geography. Agents are located along a circle around the production technology. The radius of the circle denotes the processing cost in the primary market. The bank is located above the production technology at an equal distance of c_b from the agents, while the exchange is situated under the production technology at equal distance of c_m from the agents.[8] However, we want to emphasize

[8] If we denote the radius of the circle c_p, standing for the transaction cost in the primary market, the *gross* payoff to the production technology would be $R(1 - c_p)^{-2}$.

that Figure 1 is merely a visual aid. Distances proxy for costs of order taking, accounting, clearing, settling, etc. The visualization explicitly distinguishes between three different types of processing costs: primary market processing costs (c_p), bank processing costs (c_b) and secondary market processing costs (c_m). To keep our horse race between fair, we assume that banks incur the same primary market processing costs as agents. Letting banks derive scale economies in the primary market would affect our conclusion in a trivial way. Because primary market processing costs cannot be avoided we define the return on the long term project (R) net of processing costs. Storage takes place 'at home' (in the exchange economy), or at 'the bank'. That is, there are no processing costs for storage.

We let processing costs be variable only to simplify the exposition: making processing costs fixed leads to the same results. With fixed processing costs participants would employ a mixed strategy in which they invest all or nothing.

Finally, we allow for potential side trading by giving late dying depositors the opportunity to withdraw funds from the bank and then trade in the market at t_1.

III. THE MARKET ECONOMY

In an equilibrium where only an exchange operates, at t_0 agents invest some proportion y of their endowment in the long term project store the remainder. Agents who become impatient sell project-backed bonds in the market to patient agents. Patient agents spend z on buying project backed bonds. If the after processing costs return on such bonds is less than or equal to unity, they also consume early.[9] Ex ante, an agent's maximization problem is:

$$\max_{\substack{y\in[0,1]\\ z\in[0,1-y]}} \lambda U\left(1-y+(1-c_m)yp^*\right) + (1-\lambda)U\left(1-y-z+\left(y+\frac{z}{(1+c_m)p^*}\right)R\right) \tag{1}$$

We look for a symmetric equilibrium in which all agents invest y^* in projects and z^* project-backed bonds. In equilibrium we need (y^*, z^*) to solve (1) and the market clearing condition to hold:

$$p^* = \frac{(1-\lambda)z^*}{\lambda y^*(1+c_m)} \tag{2}$$

The first derivative of (1) with respect to z is:

$$\frac{\partial E[U]}{\partial z} = (1-\lambda)U'(C_2^*)\left(\frac{R}{(1+c_m)p^*}-1\right) = 0 \tag{3}$$

[9] Patient agents can, instead of consuming $1-y-z$ early, also store this amount to be consumed at t_2. Since agents only care about total consumption, it does not affect their utility.

Where $C_2^* \equiv 1 - y - z + (y + \frac{z}{(1+c_m)p^*})R$ denotes the equilibrium consumption of the patient type. There are thus two kinds of equilibrium. If $p^* < \frac{R}{(1+c_m)}$, agents spend all their stored goods on buying project-backed bonds in the secondary market, or, $z = 1 - y$. In this case the equilibrium can be found by simultaneously solving:

$$p^{\#} = \frac{(1-\lambda)(1-y)}{\lambda y(1+c_m)} \tag{4}$$

which is the market clearing condition with z substituted by $1 - y$, and first order condition:

$$\lambda U'\left(C_1^{\#}\right)((1-c_m)p^{\#} - 1) + (1-\lambda)U'\left(C_2^{\#}\right)\left(R - \frac{R}{(1+c_m)p^{\#}}\right) = 0 \tag{5}$$

where the left-hand side is the first derivative with respect to y of (1) with $z = 1 - y$, and $C_1^{\#} = 1 - y + y(1-c_m)p^{\#}$ and $C_2^{\#} \equiv (y + \frac{1-y}{(1+c_m)p^{\#}})R$.

If the solution of this system of equations has $p^{\#} = \frac{(1-\lambda)(1-y)}{\lambda y(1+c_m)} < \frac{R}{(1+c_m)}$, we have an equilibrium, or $p^* = p^{\#}$. If we have $p^{\#} > \frac{R}{(1+c_m)}$ we must have an equilibrium of the second kind, which has $p^* = \frac{R}{(1+c_m)}$, and agents do not spend all stored goods on buying bonds. In this case the equilibrium (z^*, y^*) can be found by simultaneously solving (2) and first order condition

$$\lambda U'(C_1^*)\left(R\frac{(1-c_m)}{(1+c_m)} - 1\right) + (1-\lambda)U'(C_2^*)(R-1) = 0 \tag{6}$$

where $C_1^* = 1 - y + y(1-c_m)p^*$.

In the appendix we prove that for all parameter values a unique solution exists. The following proposition characterizes the market equilibrium:

Proposition 1.

If $R > 1$, $0 < \lambda < 1$, $\gamma > 0$, *and* $c_m \geq 0$,

(i) there exists a unique pure strategy equilibrium $\{y^*, p^*\}$
(ii) $\frac{\partial y^*}{\partial c_m} < 0$; $\frac{\partial y^*}{\partial \lambda} < 0$; $\frac{\partial y^*}{\partial \gamma} < 0$; $\frac{\partial y^*}{\partial R} \gtrless 0$
(iii) There exists a critical processing cost, $\underline{c}_m(\lambda, \gamma, R)$, *above which patient agents do not convert all their savings into bonds at* t_1 $(z^* > 0)$
(iv) There exists a critical processing cost, $\bar{c}_m \equiv \min(1, |\frac{R-1}{1+(2\lambda-1)R}|) > \underline{c}_m(\lambda, \gamma, R)$ *above which there is no trade at all.*

The existence of processing costs affects the equilibrium in important ways. First, in the $c_m = 0$ case, agents are indifferent with respect to their investment decision, so that mixed strategies with constant aggregate investment co-exist with the pure strategy equilibrium that was forwarded by Diamond and Dybvig (1983). However, if $c_m > 0$, not only aggregate but also individual investment is uniquely determined. Mixed strategies on secondary market participation may still obtain.

Second, we find that investment y^* depends not only on the probability of becoming impatient λ as in DD, but also on c_m, γ, and R. In accordance with intuition, equilibrium investment decreases in the processing cost and risk aversion. We find that equilibrium investment *decreases* in R if $\gamma > 1$. The reason is that investing more increases the payout to patient agents, but decreases the payout to impatient agents.

Finally, also the equilibrium price depends on λ, R, γ, and c_m. The price at the intermediate date may either in- or decrease in the parameters. The direction of dependence depends on the kind of equilibrium. In particular, we find that p^* initially increases in c_m (in the first type of equilibrium) and decreases for $c_m > \underline{c}_m$.

In the exchange economy there are two critical processing costs. First, we have $\underline{c}_m(\lambda, \gamma, R)$, above which patient agents consume early. The second critical processing cost is determined by newborns and impatient agents. For $\lambda > 1/2$, newborns will not invest even a small amount if $\lambda(1 - c_m)p + (1 - \lambda)R < 1$. And, of course, impatient investors will not sell their claims if $c_m > 1$.[10]

Figure 2 illustrates how the equilibrium depends on processing costs. The figure shows the $y^*(c_m)$ and $p^*(c_m)$ functions for a *base case*, with $\gamma = 2$, $\lambda = 1/2$, and $R = 2$, and for several variations. The sharp angles in the $y^*(c_m)$ and $p^*(c_m)$ curves denote the first critical processing cost. In the base case, trade is significantly reduced for $c_m > 0.55$, when patient investors no longer exchange all their savings for bonds. Both trade and investment disappears when $c_m > 1$. Figure 2 also shows that lower risk aversion leads to marginally higher equilibrium investment, lower prices, and a later kink (lower $\underline{c}_m$). A higher R leads to marginally *lower* investment, higher prices and a later kink (higher $\underline{c}_m$), whereas a higher probability of becoming impatient leads to lower investment and an earlier kink. Notice that in the $R = 3$ case there is still investment if $c_m > 1$, but no trade, and that in the $\lambda = 0.7$ case, there will be no investment if $c_m > 0.58$, even though the (reservation) price is $\frac{R}{1+c_m} = 1.27$.

IV. THE BANK ECONOMY

In this section we explore how a Diamond Dybvig bank acts as a mechanism to share liquidity risk. The bank offers demand deposits which gives the holder the right to withdraw r_1 at t_1 or r_2 at t_2 per unit of consumption good deposited. As mentioned earlier, the bank faces a variable processing cost c_b per unit transacted.

Following Diamond and Dybvig, we look for a vector $\boldsymbol{r}^* = \{r_1^*, r_2^*\}$ that maximizes agents' *ex ante* expected utility. Unlike Diamond and Dybvig we explicitly

[10] Notice that for small λ and $c_m > 1$, newborns may still *invest*. In such cases there is no trade, and impatient agents simply forego their claims. In such cases we have $C_1 = 1 - y$ and $C_2 = 1 - y + yR$.

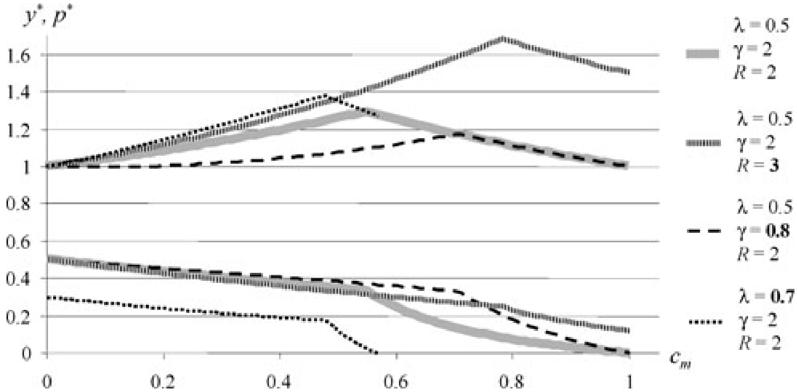

Figure 2: The Exchange Equilibrium.

We calculate the equilibrium price p^* and investment y^* (lower lines) that obtain in a Diamond and Dybvig exchange economy where risk averse agents invest in a long term asset with return R and need to consume with probability λ. Buying and selling comes with a variable processing cost. The figure shows p^* and y^* as a function of processing cost c_m for a *base case* which has $\lambda = 0.5$, $R = 2$, and CRRA parameter $\gamma = 2$ (the thick grey line), as well as for variations with different R, γ, and λ.

take into account that agents may deviate from the equilibrium by either investing their entire endowment in bonds at t_0, and, in case of becoming impatient, offering them for sale to patient depositors, or by storing their entire endowment and, if remaining patient, buying bonds in the market at t_1.

Given this set up, the problem for the bank becomes:

$$\max_{r} \lambda U\big((1-c_b)^2 r_1\big) + (1-\lambda) U\big((1-c_b)^2 r_2\big) \tag{7}$$

Subject to the budget constraint:

$$(1-\lambda r_1)\, R \geq (1-\lambda) r_2 \tag{8}$$

And the incentive compatibility constraints:

$$\lambda U\big((1-c_b)^2 r_1\big) + (1-\lambda) U\big((1-c_b)^2 r_2\big) \geq U(1) \tag{9}$$

$$\begin{aligned} &\lambda U\big((1-c_b)^2 r_1\big) + (1-\lambda) U\big((1-c_b)^2 r_2\big) \\ &\quad > \lambda U\left(\frac{(1-c_m) r_1}{(1+c_m) r_2} R\right) + (1-\lambda) U(R) \end{aligned} \tag{10}$$

$$\begin{aligned} &\lambda U\big((1-c_b)^2 r_1\big) + (1-\lambda) U\big((1-c_b)^2 r_2\big) \\ &\quad > \lambda U(1) + (1-\lambda) U\left(\frac{(1-c_m) r_2}{(1+c_m) r_1}\right) \end{aligned} \tag{11}$$

Constraint (9) requires that banking is better than storing. Constraints (10) and (11) require that banking is better than side trading.[11] In all three equations, the left hand side is the expected utility from banking. The right-hand side of (10) is the expected utility of an agent who at t_0 invests her entire endowment in bonds and sells them to patient depositors in case she becomes impatient. Because patient depositors stand to receive $(1 - c_b)^2 r_2$ at t_2 if they hold on to their deposit, and $\frac{(1-c_b)^2 r_1}{(1+c_m)p} R$ if they withdraw and buy bonds, they are willing to pay up to $\frac{R r_1}{(1+c_m) r_2}$ per bond. A single deviator could thus pocket $\frac{(1-c_m) r_1}{(1+c_m) r_2} R$ by side-trading. The right hand side of (11) is the expected utility of an agent who stores his endowment and, if he stays patient, buys bonds from impatient depositors, who are willing to *sell* deposit-backed bonds for $\frac{R r_1}{(1-c_m) r_2}$.[12]

Because budget constraint (8) is binding, maximization problem (6) is straightforward. If $\gamma > 1$, the bank would like to set $r_1 > 1$ and $r_2 < R$, but may find itself constraint by (10) or (9). If, on the other hand, $\gamma < 1$, the bank wants to set $r_1 < 1$, $r_2 > R$, and constraint (11) or (9) may bite. If it is impossible to satisfy both side trade constraints, or if the bank cannot offer agents a utility higher than the storing option, a bank equilibrium does not exist.

The following proposition describes the equilibrium in the bank economy:

Proposition 2.

If $0 < \lambda < 1$, $0 < \gamma < \infty$, $1 < R < \infty$, $c_m > 0$,

(i) There exists a critical processing cost $\underline{c}_b(\lambda, \gamma, R, c_m)$ *so that for* $c_b < c_b$, *there exists an unconstrained bank equilibrium.*
(ii) There exists a critical processing cost $\bar{c}_b(\lambda, \gamma, R, c_m)$ *so that for* $c_b > \bar{c}_b$, *no bank equilibrium exists* $\bar{c}_b(\lambda, \gamma, R, c_m)$ *weakly increases in* c_m.
(iii) If $c_m = c_b$, *there is no bank equilibrium.*

Not surprisingly, a bank is only feasible as a risk sharing vehicle if its processing costs are sufficiently low. If c_m is high, and c_b sufficiently low, the bank may offer the first best risk sharing schedule, that is, the solution to maximization problem (6). If $c_b > \underline{c}_b$, a bank may be able to offer a second best, side-trade constrained, schedule $\{r_1^*, r_2^*\}$.

Competitive banks aim to offer a payoff schedule $\{r_1^*, r_2^*\}$ so as to best cater to their clients preferences but are constrained by the threat of side-trading. The side trade threat is mitigated if processing costs in the market are higher, so that with higher c_m, the bank can offer (weakly) 'better' payoff schedules, and hence sustain a higher processing costs c_b. We find that the critical processing costs do not only depend on c_m, but also on the probability of becoming impatient, the agent's risk aversion, and the return on the long term asset.

[11] The reason that (10) and (11) are strict inequalities is that the bank cannot afford a single deviating depositor.
[12] Or, they are willing to sell their $(1 - c_b)$ deposits, each of which pays $(1 - c_b) r_2$ at t_2, for $r_2(1 - c_b)^2/(1 - c_m)$, or $r_2(1 - c_b)/(1 - c_m)$ per unity-denominated deposit.

The most interesting result of our analysis is contained in proposition 2iii), which says that if banks and exchanges face equal positive proportional processing costs, banks cannot exist in equilibrium. Notice that if $c_m = c_b = 0$, we are back in Jacklin's Diamond and Dybvig world, in which a potential bank *can* exist but cannot improve on the market's allocation. The proof of proposition 2iii) (see appendix) shows that for side trade constraint (10) to hold we need $r_1 < 1$, in which case constraint (11) cannot hold. A more intuitive way to see that the exchange mechanism is superior in case of equal proportional processing costs is by looking at the transaction values in both economies. In the bank economy, the total transaction value is $1 + \lambda r_1 + (1 - \lambda) r_2$, while in the market economy it is only $2\lambda y^* p^*$. For small c's, the schedule of the side trade proof bank, $\{r_1^*, r_2^*\}$, approaches $\{1, R\}$ while the market equilibrium $\{y^*, p^*\}$ goes to $\{1 - \lambda, 1\}$. From this follows that the processing cost free bank processes $1 + \lambda + (1 - \lambda)R$, more than four times the transaction value in the exchange economy, $2\lambda(1 - \lambda)$.

Figure 3 graphically illustrates how the bank equilibrium depends on processing costs. The figure shows the payoffs r_1 and r_2 for two parameter constellations.

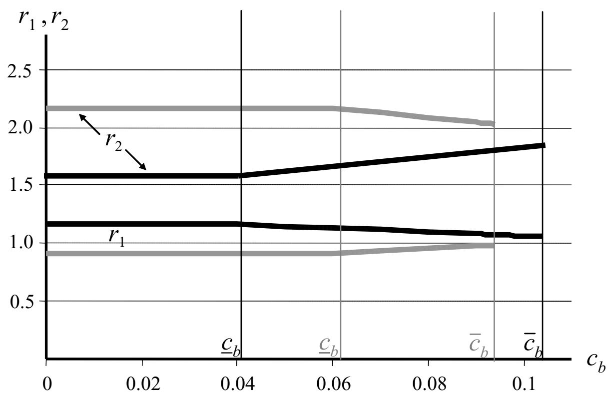

Figure 3: The Bank Equilibrium.

We compute agents' payoffs when agents use a Diamond Dybvig bank to share liquidity risk. Risk averse agents, who become 'impatient' with probability λ before the long term technology pays R, deposit their unit endowments in the bank. The bank invests part of the deposits in the long term technology and stores the remainder. Depositors may withdraw r_1 at time t_1, or r_2 at t_2. All transactions with the bank are subject to a processing cost c_b. If agents decide to trade outside the bank, they incur a processing cost c_m. The figure shows the equilibrium payoffs r_1 and r_2 for the case with $\lambda = 0.5$, $R = 2$, $c_m = 0.1$ and CRRA parameter $\gamma = 2$ (the black lines) as a function of bank processing cost c_b. The grey lines show the equilibrium payoffs for the situation $\lambda = 0.5$, $R = 2$, $c_m = 0.1$, $\gamma = 0.8$. Critical processing costs $\underline{c}_b$, for unconstraint risk sharing and $\bar{c}_b$ for constraint risk sharing are denoted on the x-axis.

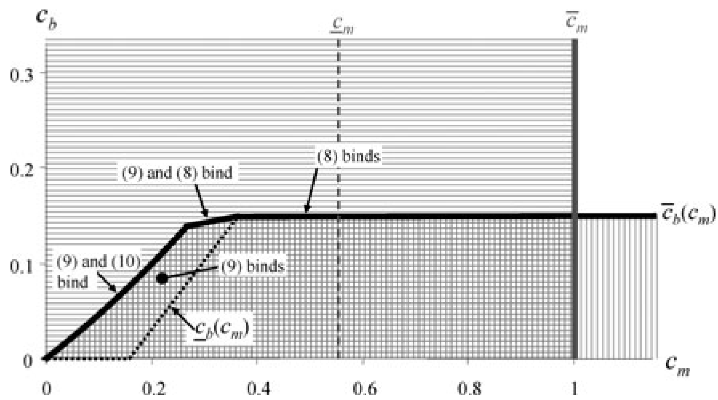

Figure 4: Critical Processing Costs.

In a Diamond Dybvig economy risk averse agents who become illiquid with probability λ before a long term technology returns R, share risk using either a market mechanism or a bank. For $\lambda = 1/2$, $R = 2$, and agents' CRRA parameter $\gamma = 2$, we compute bank and market equilibria in the presence of processing costs c_m for market trade and c_b for bank transactions. The pattern of horizontal lines, bordered by the line $\bar{c}_m$, marks the region in which a market economy is exists. The pattern of vertical lines, below line $\bar{c}_b(c_m)$, marks the region where a bank economy is exists.

Clearly, if $\gamma > 1$, the bank offers $r_1 > 1$, $r_2 < R$. For low enough c_b, the bank can chose the unconstrained solution given by (7). If c_b reaches a first critical level, schedule $\{r_1^*, r_2^*\}$ becomes constraint by equation (9), forcing the bank to decrease r_1 and increase r_2 to avoid side traders to play the "invest and sell to patient depositors" game. If $\gamma < 1$, the bank offers $r_1 < 1$, $r_2 > R$, until it sees itself constrained by side traders who may play "store and buy deposits from impatient depositors".

Figure 4 shows the critical bank processing costs $\underline{c}_b$ and $\bar{c}_b$ as a function of c_m, for given values of λ, R, and γ. As indicated, the critical processing costs increase in c_m because a higher market processing costs decreases the side-trade threat. For small c_m, the zero processing cost bank can offer some ex-post wealth transfer (set $r_1 > 1$, $r_2 < R$), but is constrained by side trade threat (10). The wealth transfer it can offer then decreases in c_b.[13] The c_b where also constraint (9) or (11) starts to bite determines $\bar{c}_b$. For small c_m, constraint (9) bites before (11) does. For large c_m, banks can offer first best risk sharing until c_b is so large that autarky becomes superior.

Figure 4 also shows the critical processing costs for the exchange mechanism, $\underline{c}_m$ and $\bar{c}_m$, as well as the areas where both mechanisms are feasible. We find that

[13] To see this, notice that an increase in c_b decreases the lhs of the binding constraint (10), but not the rhs. To keep its depositors on board, the bank needs to decrease the r_1/r_2 ratio. This action reduces the side trader's utility more than it does the depositors' utility.

there is a substantial area where both an exchange and a bank economy constitute a Nash equilibrium. Notice however that in this (thatched) area banks and markets cannot *co*-exist. The reason for this is twofold: If both types of equilibrium are feasible, one will always offer higher utility as the other one. Moreover, the market price in the bank equilibrium is different from the bid-price (or asking price, in case $\gamma < 1$) for bonds in the corresponding bank equilibrium.

V. BANKS VERSUS EXCHANGES

In this section we investigate which equilibrium is more likely to prevail. To answer this question, we calculate the agents' *ex ante* expected utilities in the bank and market economy for those (c_m, c_b) combinations where both types of equilibrium are *Nash*. Not surprisingly, we find a diagonal line that divides the thatched area of Figure 4 into two. For (c_m, c_b) coordinates above the iso utility line (high c_b, low c_m) the market economy is the Pareto superior equilibrium. In the area under the line, the bank arrangement offers superior allocations.

We find that the iso utility line is not only determined by the aggregate processing costs: when we draw the iso processing cost line, we find it to be different. Moreover, it lies in bank territory, implying that for a given level of *aggregate* processing costs, the bank equilibrium offers the Pareto superior allocation, and that there exists a substantial region where the bank offers a superior allocation *despite* the fact that it incurs higher aggregate processing costs than the exchange. The above observations are formalized in proposition 3:

Proposition 3.

For all $0 < \lambda < 1$, $0 < \gamma < \infty$, $1 < R < \infty$, $c_m, c_b \geq 0$,

(i) There exists a region in (c_m, c_b)-space where the bank equilibrium offers agents a higher expected utility than the exchange.

(ii) There exists a region in (c_m, c_b)-space where the bank equilibrium offers a higher expected utility than the exchange, and where the aggregate processing costs incurred are higher in the bank equilibrium than in the exchange equilibrium.

The proof of proposition 3(i) is straightforward. It is easy to see that the expected utility in the exchange equilibrium decreases in c_m, while it (weakly) increases in the bank equilibrium. Hence there must a line $c_m(c_b)$ above which banking offers a higher expected utility.

Proposition 3(ii) stems from two sources. First we have that, unlike the exchange, the bank is able to offer agents some degree of *ex-post* wealth transfer (Jacklin, 1987). Due to the desirability of such liquidity insurance, the bank can sustain higher aggregate processing costs and still survive the Nash equilibrium criterion. Because ex-post wealth transfer is desirable unless $\gamma = 1$, the wedge between the iso-utility line and the iso-processing costs increases in $|\gamma - 1|$.

A second, more subtle, explanation for the bank's superiority for given aggregate processing costs is that in the bank equilibrium the processing costs are optimally allocated between early diers and late diers: early diers end up paying $c_b + (1 - c_b)r_1 c_b$ in processing costs, while late diers pay $c_b + (1 - c_b)r_2 c_b$. In the exchange equilibrium, early and late diers pay exactly the same processing costs: $\lambda c_m y^* p^*$. In the presence of risk aversion ($\gamma > 0$), the former allocation of processing costs is superior by Jensen's inequality. It is this phenomenon that causes the wedge between the iso utility line and the iso processing cost line to be positive even for $\gamma = 1$.

Figure 5 illustrates the analysis of this section. For parameter values $\lambda = 1/2$ and $R = 2$, we plotted the iso utility and iso processing cost lines for three different values for γ. The first panel shows, for $\gamma = 2$, the regions where exchanges, respectively banks, are the more reasonable risk sharing mechanisms. Notice that in the light-grey speckled area, a bank economy may offer lower expected utility, but is still a Nash equilibrium: if everybody banks, a *single* deviator cannot do better by side trading. Similarly, in the dark-grey speckled area, the exchange is an *unintuitive* but valid Nash equilibrium, simply because if all agents choose the exchange mechanism a *single* agent will not start a bank.[14]

The wedge between the iso-processing cost line and the iso-utility line denotes the area where the bank is the Pareto superior risk sharing vehicle even though aggregate processing costs in the market equilibrium are lower. The wedge widens for $c_m > \underline{c}_{\mathrm{m}}$ because in this area the amount of trade (and hence aggregate processing costs) in the exchange quickly reduces.

VI. SUMMARY AND DISCUSSION

In this paper we study a Diamond Dybvig model with exogenous processing cost to compare welfare in a bank economy with that in a pure exchange economy. We find that banks process significantly more goods than an exchange. This plays to the advantage of the exchange and leads us to conclude that, if processing costs are proportional to transaction value, bank economies cannot exist if agents are free to trade outside the coalition.

On the other hand, processing costs in the market give banks the opportunity to offer some degree of desirable liquidity insurance. This is the main reason that if aggregate processing costs are fixed, bank economies offer their constituents higher expected utility than exchange economies. Another advantage of the bank arrangement is that transaction values are proportional to consumption. If agents are risk averse, this allocation is better than in the market economy, where both early and late dying agents transact the same amount, and hence incur the same absolute processing costs. Finally, we find that for low levels of processing costs, there are two feasible equilibria, even though one results in higher welfare.

[14] Or, in the light (dark) grey area, only the exchange (bank) equilibrium meets the intuitive criterion of Cho and Kreps (1996).

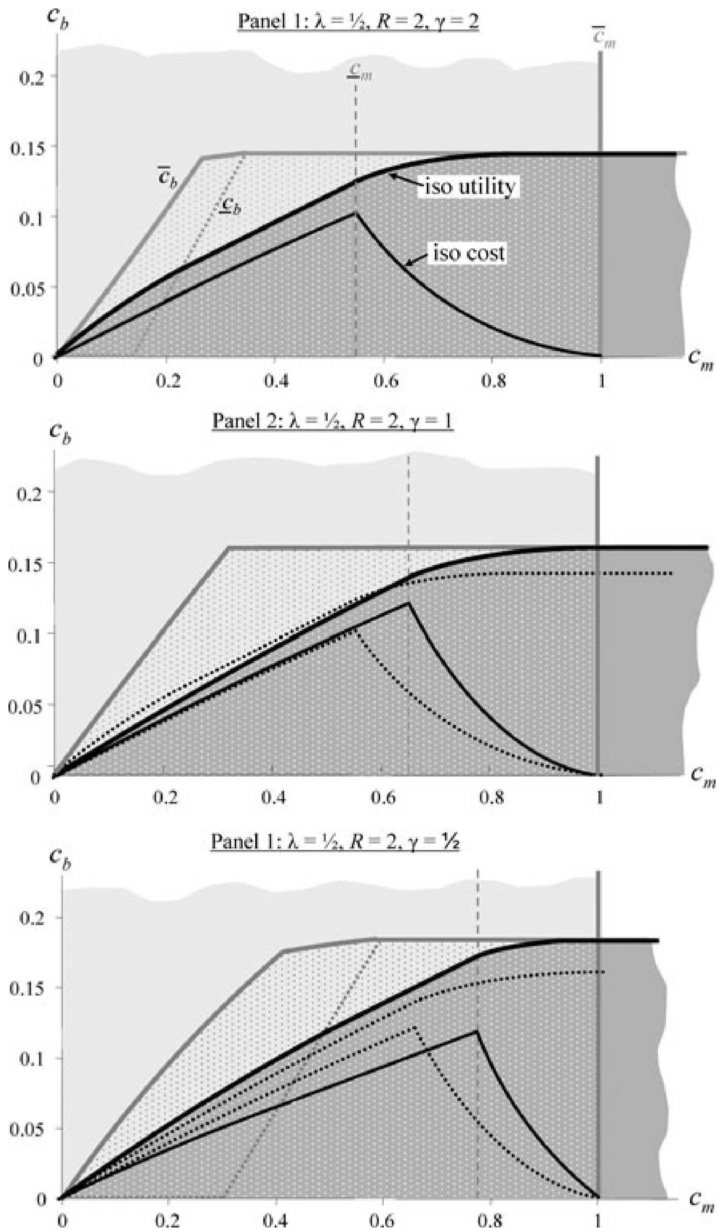

Figure 5: Bank versus Exchange Equilibrium.

This figure shows the areas in (c_m, c_b)-space in which exchange and bank equilibria exist. In the light grey region, the exchange results in the higher expected utility. In the dark grey region the bank equilibrium Pareto-dominates. In the speckled areas both equilibriums exist.

VII. APPENDIX

PROOF OF PROPOSITION 1 (EXCHANGE EQUILIBRIUM)

It is easy to establish that for $c_m = 0$, the equilibrium is $\{y^*, p^*\} = \{1 - \lambda, 1\}$. (Diamond and Dybvig, 1983). For small c_m the price constraint will not be binding so that in equilibrium, $p = \frac{(1-\lambda)(1-y)}{\lambda y(1+c_m)}$. The first order condition (3) can then be rewritten as:

$$\lambda(1-y)^{1-\gamma}(1+K_1)^{-\gamma}(K_1 - y(1+K_1)) + y^{1-\gamma}K_2\left(1 - \frac{y}{1-\lambda}\right) = 0 \quad \text{(A1)}$$

where $K_1 = \frac{(1-\lambda)(1-c_m)}{\lambda(1+c_m)}$ and $K_2 = R^{1-\gamma}\ (1-\lambda)^{1+\gamma}$

To prove equilibrium uniqueness for small c_m we note that the first term of (A1), $T_{1,1}(y) \equiv \lambda(1 - y)^{1-\gamma}(1 + K_1)^{-\gamma}\ (K_1 - y\,(1 + K_1))$ crosses the x-axis once, from above, on the interval (0, 1), at $y^{T_1=0} \equiv \frac{K_1}{1+K_1}$. The second term, $T_{1,2}(y) \equiv y^{1-\gamma}K_2(1 - \frac{y}{1-\lambda})$ also crosses the x-axis once, from above, at $y^{T_2=0} \equiv 1 - \lambda$.

Because $y^{T_1=0} = \cdots = 1 - \lambda - \frac{2\lambda c_m}{1-c_m+2\lambda c_m} \leq y^{T_2=0}$, and because the sign of the second derivatives of both terms do not change over the (0, 1) range, there is a unique solution to (A1).[15]

To prove that $\frac{\partial y_0^*}{\partial c_m} < 0$ in the region where the price constraint is not binding (c_m small), we use the fact that that $T_{1,1}$ decreases in c_m for all $y \in (y^{T_1=0}, 1)$:

$$\frac{\partial T_{1,1}}{\partial c_m} = \frac{\partial T_{1,1}}{\partial K_1}\frac{\partial K_1}{\partial c_m} = \lambda(1-y)^{1-\gamma}\,(1+K_1)^{-\gamma-1}\,(1+K_1(1-\gamma) - y(1+K_1)(1-\gamma))\frac{\partial K_1}{\partial c_m} \quad \text{(A2)}$$

If $\gamma < 1$, the last bracketed term decreases continuously in y. The minimum value on $(y^{T_1=0}, 1)$ is thus reached at $y = 1$. In this case the last bracketed term is still positive, making $\frac{\partial T_{1,1}}{\partial c_m}$ negative, because $\frac{\partial K_1}{\partial c_m} < 0$. If $\gamma > 1$ the last bracketed term of (A2) increases continuously in y. It becomes positive (and hence $\frac{\partial T_{1,1}}{\partial c_m}$ negative) at

$$y^{T_1'>0} \equiv \frac{1+K_1(1-\gamma)}{(1+K_1)(1-\gamma)} = \frac{K_1}{1+K_1} - \frac{1}{(1+K_1)\,(\gamma-1)} < y^{T_1=0}$$

Combining $\frac{\partial T_{1,1}}{\partial c_m} < 0$ with the observation that $\frac{\partial T_{1,2}}{\partial c_m} = 0$, we have $\frac{\partial y^*}{\partial c_m} < 0$.

Hence, for small c_m there is a unique solution y^*, which decreases in c_m.

At some c_m, call it $\bar{c}_m(\lambda, \gamma, R)$, the price constraint becomes binding. At this point not only (A1) holds but also:

[15] The first derivative of $T_{1,1}$ may change sign. But because neither term has a "wiggle" (e.g., first concave then convex in y) on the interval $(y^{T_1=0}, y^{T_2=0})$, there can only be one solution to (A6).

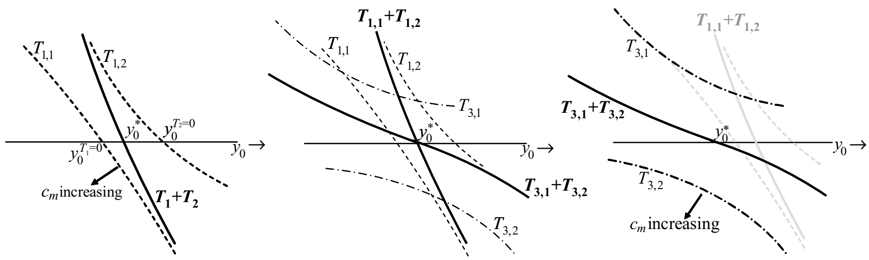

Figure A1: Derivation of the Equilibrium in the Exchange Economy.

In the first panel c_m is small. The equilibrium is found by letting equation (A1) hold. In the middle panel the price constraint just binds. Both equation (A1) and (A3) hold. In the right panel, c_m is high and the price constraint is binding. The equilibrium is found by letting equation (A3) hold.

$$\lambda\left(1-y\left(1-\frac{1-c_m}{1+c_m}R\right)\right)^{-\gamma}\left(\frac{1-c_m}{1+c_m}R-1\right) + (1-\lambda)\left(1+y\left(R-1\right)\right)^{-\gamma}\left(R-1\right)=0 \tag{A3}$$

which is the first order condition with p substituted by $\frac{R}{1+c_m}$. It can be easily verified that both terms, call them $T_{3,1}$ and $T_{3,2}$, continuously decrease in y and that the first term decreases in c_m, proving that also if c_m is increasing beyond $\bar{c}_m$, there is a unique solution, and that also then y_0^* decreases in c_m. See Figure A1.

$\frac{\partial y^*}{\partial \gamma} < 0$ follows from:

$$\frac{\partial \text{lhs(A1)}}{\partial \gamma} = -\lambda(1-y)^{1-\gamma}(1+K_1)^{-\gamma}\left(K_1 - y(1+K_1)\right)\ln\left((1-y)(1+K_1)\right) - y^{1-\gamma}K_2\left(1-\frac{y}{1-\lambda}\right)\ln\left(\frac{yR}{1-\lambda}\right) < 0 \quad \forall y \in \left(y^{T_{1,1}=0}, y^{T_{1,2}=0}\right) \tag{A4}$$

and

$$\frac{\partial \text{lhs(A3)}}{\partial \gamma} = -\lambda\left(1-y\left(1-\frac{1-c_m}{1+c_m}R\right)\right)^{-\gamma}\left(\frac{1-c_m}{1+c_m}R-1\right) \times \ln\left(1-y\left(1-\frac{1-c_m}{1+c_m}R\right)\right) - (1-\lambda)\left(1+y\left(R-1\right)\right)^{-\gamma}\left(R-1\right) \times \ln\left(1+y\left(R-1\right)\right) < 0 \quad \forall y \in (0, 1-\lambda) \tag{A5}$$

Similarly, $\frac{\partial y^*}{\partial \lambda} < 0$ because $\frac{\partial \text{lhs(A1)}}{\partial \lambda}$, $\frac{\partial \text{lhs(A3)}}{\partial \lambda} < 0$ on the relevant intervals.

Finally, $\frac{\partial y^*}{\partial R} \gtrless 0$ follows from $\frac{\partial \text{lhs(A1)}}{\partial R} = \frac{\partial T_{1,2}}{\partial R} = (1-\gamma)y^{1-\gamma}\frac{K_2}{R}(1-\frac{y}{1-\lambda}) \gtrless 0$

From the above expression we see that if $\gamma < (>)\, 0$, $\frac{\partial y^*}{\partial R} > (<)\, 0$. *Q.E.D.*

Proof of Proposition 2 (Bank Equilibrium)

Because utility strictly increases in consumption, budget constraint (8) holds, so that:

$$r_2 = \frac{1-\lambda r_1}{1-\lambda} R \tag{A6}$$

Hence, maximization problem (7) becomes:

$$\max_{r_1} \lambda U\big((1-c_b)^2 r_1\big) + (1-\lambda) U\left(R\frac{(1-c_b)^2}{1-\lambda}(1-\lambda r_1)\right) \tag{A7}$$

Which has first order condition:

$$r_1^{-\gamma} - R^{1-\gamma}(1-\lambda)^{\gamma}(1-\lambda r_1)^{-\gamma} = 0 \tag{A8}$$

Consider both terms as functions of r_1. The first term, $T_1(r_1) \equiv r_1^{-\gamma}$, is positive and decreasing on the same interval and has $\lim_{r_1 \downarrow 0} T_1(r_1) = \infty$. The second term, $T_2(r_1) \equiv -R^{1-\gamma}(1-\lambda)^{\gamma}(1-\lambda r_1)^{-\gamma}$, is negative and decreasing for all $r_1 \in (0, 1/\lambda)$ and has $T_2(0) = -R^{1-\gamma}(1-\lambda)^{\gamma}$ and $\lim_{r_1 \uparrow \frac{1}{\lambda}} T_2(r_1) = -\infty$.

This proves the existence of a unique solution to the unconstrained maximization problem.

It can be easily verified that the r.h.s. of constraint (10), with $r_2 = \frac{1-\lambda r_1}{1-\lambda}R$, is continuously increasing on $(0, 1/\lambda)$, while the r.h.s. of constraint (11) is continuously decreasing on $(0, 1/\lambda)$. This proves that there is at most one solution for the constrained maximization problem.

To prove proposition 2iii), let $c = c_b = c_m$. Side trade constraint (10) and (11) then become:

$$\lambda U\left(C_1^b\right) + (1-\lambda)U\left(C_1^b\right) > \lambda U\left(C_1^{di}\right) + (1-\lambda)U\left(C_2^{di}\right) \tag{A9}$$

$$\lambda U\left(C_1^b\right) + (1-\lambda)U\left(C_1^b\right) > \lambda U\big(C_1^{ds}\big) + (1-\lambda)U\left(C_2^{ds}\right) \tag{A10}$$

Where $\{C_1^b, C_2^b\} \equiv \{(1-c)^2 r_1, (1-c)^2 r_2\}$, denote consumption of impatient and patient agents who stick to the banking game, $\{C_1^{di}, C_2^{di}\} = \{\frac{(1-c)r_1}{(1+c)r_2}R, R\}$ the consumption schedule of a single agent who *d*eviates by *i*nvesting all in the long term technology, and sells it to patient depositors at t_1 in case he becomes impatient, and $\{C_1^{ds}, C_2^{ds}\} = \{1, \frac{(1-c)r_2}{(1+c)r_1}\}$ the consumption schedule of a *d*eviator who *s*tores and, if she remains patient, buys deposits of impatient depositors.

Assume $r_1 \geq 1$, then clearly $C_2^{di} > C_2^b$ by the budget constraint. Moreover,

$$C_1^{di} \equiv \frac{(1-c)r_1}{(1+c)r_2}R = \frac{(1-\lambda)(1-c)}{(1+c)(1-\lambda r_1)}r_1 \geq \frac{(1-c)}{(1+c)}r_1 > (1-c)^2 r_1 \equiv C_1^b$$

The first equality follows from the budget constraint, the second (weak) inequality follows from r_1 being ≥ 1. The final inequality follows from algebra. The entire expression shows that r_1 cannot be ≥ 1 without violating (10).

Now assume $r_1 < 1$. Then clearly $C_1^{ds} > C_1^b$. And from this we find that

$$C_2^{ds} \equiv \frac{(1-c)r_2}{(1+c)r_1} > \frac{(1-c)}{(1+c)}r_2 > (1-c)^2 r_2 \equiv C_2^b$$

Hence, r_1 cannot be < 1 without violating (11). *Q.E.D.*

PROOF OF PROPOSITION 3 (BANKS VERSUS EXCHANGE)

To prove proposition 3ii) we will show that for $\gamma = 1$ and small c_m, c_b, expected utility decreases in total processing cost faster in the market economy than it does in the bank economy, proving the existence of a wedge between the iso utility and iso processing costs line near the origin in (c_m, c_b) space. The wedge must also exist if $\gamma <> 1$, because in these cases the bank has an *additional* advantage over the exchange: it will offer (some) *ex-post* wealth transfer from late to early diers.

For $\gamma = 1$, we find the closed form exchange equilibrium from first order condition (A1). After some algebra we find:

$$y^* = \frac{\lambda(1-\lambda)(1-c_m)}{1-c_m+2\lambda c_m} + (1-\lambda)^2 \tag{A11}$$

$$p^* = \frac{1-c_m+2\lambda c_m(2-\lambda)}{(1+c_m)(1-c_m+2\lambda c_m(1-\lambda))} \tag{A12}$$

Hence, the expected utility in the exchange economy is:

$$\begin{aligned} E[U]_{market} = \lambda \ln\left(\frac{1-c_m+4\lambda c_m-2\lambda^2 c_m}{1+c_m}\right) \\ + (1-\lambda)\ln\left(\frac{1-c_m+2\lambda c_m-2\lambda^2 c_m}{1-c_m+2\lambda c_m}R\right) \end{aligned} \tag{A13}$$

After some more algebra we find that:

$$\lim_{c_m \to 0} \frac{\partial E[U]_{market}}{\partial c_m} = -2\lambda(1-\lambda) \tag{A14}$$

The total processing costs in the market equilibrium are:

$$TTC_{market} = 2\lambda y^* p^* c_m = \frac{2\lambda(1-\lambda)\big(1 - c_m + 4\lambda c_m - 2\lambda^2 c_m\big)c_m}{(1 - c_m + 2\lambda c_m)(1 + c_m)} \quad \text{(A15)}$$

So that:

$$\lim_{c_m \to 0} \frac{\partial TTC_{market}}{\partial c_m} = 2\lambda(1-\lambda) \quad \text{(A16)}$$

Hence, we have:

$$\lim_{c_m \to 0} \frac{\partial E[U]_{market}}{\partial TTC_{market}} = \frac{-2\lambda(1-\lambda)}{2\lambda(1-\lambda)} - 1 \quad \text{(A17)}$$

The bank equilibrium with $\gamma = 1$ has $r_1 = 1$, $r_2 = R$, so that the expected utility in the bank economy is:

$$E[U]_{bank} = \lambda \ln\left((1-c_b)^2\right) + (1-\lambda)\ln\left((1-c_b)^2 R\right) \quad \text{(A18)}$$

and:

$$\lim_{c_b \to 0} \frac{\partial E[U]_{bank}}{\partial c_b} = -2 \quad \text{(A19)}$$

The total processing costs in the bank economy are:

$$TTC_{bank} = c_b + (1-c_b)(\lambda + (1-\lambda)R)c_b \quad \text{(A20)}$$

so that:

$$\lim_{c_b \to 0} \frac{\partial TTC_{bank}}{\partial c_b} = 1 + \lambda + (1-\lambda)R \quad \text{(A21)}$$

Hence we find:

$$\lim_{c_m \to 0} \frac{\partial E[U]_{bank}}{\partial TTC_{bank}} = \frac{-2}{1 + \lambda + R(1-\lambda)} = -\frac{2}{2 + (R-1)(1-\lambda)} \quad \text{(A22)}$$

This shows that the decrease of expected utility as a function of increasing total processing costs is less for the bank than for the exchange. This proves that there is a wedge between the iso utility and iso processing costs lines in (c_m, c_b) space. *Q.E.D.*

VIII. REFERENCES

Allen, F. and D. Gale. 2004. "Financial Intermediaries and Markets." *Econometrica* 72:1023–1061.

Benston, G. J. and C. W. Smith. 1976. "A Transaction Cost Approach to the Theory of Financial Intermediation." *Journal of Finance* 31:215–231.

Bhattacharya, S. and D. Gale. 1987. "Preference Shocks, Liquidity, and Central Bank Policy." Pp. 69–88 in *New Approaches to Monetary Economics*, eds. W. A. Barnett and K. J. Singleton. Cambridge: Cambridge University Press.

Bhattacharya, S. and A. J. Padilla. 1996. "Dynamic Banking: A Reconsideration." *Review of Financial Studies* 9:1003–1031.

Boot, A. W. A. and A. V. Thakor. 1997. "Financial System Architecture." *Review of Financial Studies* 10:733–793.

Cho, I. and D. Kreps. 1987. "Signaling Games and Stable Equilibria." *Quarterly Journal of Economics* 102:179–221.

Diamond, D. W. and P. H. Dybvig. 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.

Dowd, K. 1992. "Models of Banking Instability: A Partial Review of the Literature." *Journal of Economic Surveys* 6:107–132.

Freeman, S. 1988. "Banking as the Provision of Liquidity." *Journal of Business* 61:5–64.

Freixas, X. and J-C. Rochet. 1997. *Micro-Economics of Banking.* Massachusetts: MIT Press.

Fulghieri, P. and R. Rovelli. 1998. "Capital Markets, Financial Intermediaries, and Liquidity Supply." *Journal of Banking and Finance* 22:1157–1179.

Goldstein, I. and A. Pauzner. 2005. "Demand Deposit Contracts and the Probability of Rank Runs." *Journal of Finance* 60:1293–1327.

Haubrich, J. G. and R. G. King. 1990. "Banking and Insurance." *Journal of Monetary Economics* 26:361–386.

Hoerova, M. 2005. "Run-prone Banking and Markets." Unpublished paper, Cornell University.

Jacklin, C. J. 1987. "Demand Deposits, Trading Restrictions, and Risk Sharing." Pp. in *Contractual Arrangements for Intertemporal Trade*, eds. E. Prescott and N. Wallace. Minneapolis: University of Minnesota Press.

Qi, J. 1994. "Bank Liquidity and Stability in an Overlapping Generations Model." *Review of Financial Studies* 7:389–417.

von Thadden, E. L. 1999. "Liquidity Creation through Banks and Markets: Multiple Insurance and Limited Market Access." *European Economic Review* 43:991–1006.

Wallace, N. 1988. "Another Attempt to explain an Illiquid Banking System: The Diamond and Dybvig Model with Sequential Service Taken Seriously." *Quarterly Review of the Federal Reserve Bank of Minneapolis* 1988:3–16.

IX. NOTES ON CONTRIBUTOR/ACKNOWLEDGMENTS

Jos van Bommel is University Lecturer of Finance at the Saïd Business School at the University of Oxford, and fellow at Linacre Collage. His research interests includes financial intermediation, corporate finance, and market microstructure. Dr. van Bommel has published in several journals, including the *Journal of Finance*, the *Journal of Banking and Finance* and the *Journal of Corporate Finance*.

I thank Tim Jenkinson and Dimitrios Tsomocos, and seminar participants at the finance workshops of the University of Vienna and the University of Oxford for valuable feedback. All errors are mine.