WORKING PAPERS
RESEARCH DEPARTMENT

WORKING PAPER NO. 09-26
BANKING: A MECHANISM DESIGN APPROACH

Fabrizio Mattesini
University of Rome Tor-Vergata

Cyril Monnet
Federal Reserve Bank of Philadelphia

Randall Wright
University of Wisconsin
and Visiting Scholar,
Federal Reserve Bank of Philadelphia

October 15, 2009

RESEARCH DEPARTMENT, FEDERAL RESERVE BANK OF PHILADELPHIA
Ten Independence Mall, Philadelphia, PA 19106-1574 • www.philadelphiafed.org/research-and-data/

Banking: A Mechanism Design Approach*

Fabrizio Mattesini
University of Rome Tor-Vergata

Cyril Monnet
Federal Reserve Bank of Philadelphia

Randall Wright
University of Wisconsin
and Federal Reserve Bank of Philadelphia

October 15, 2009

Abstract

We study banking using the tools of mechanism design, without a priori assumptions about what banks are, who they are, or what they do. Given preferences, technologies, and certain frictions – including limited commitment and imperfect monitoring – we describe the set of incentive feasible allocations and interpret the outcomes in terms of institutions that resemble banks. Our bankers endogenously accept deposits, and their liabilities help others in making payments. This activity is essential: if it were ruled out the set of feasible allocations would be inferior. We discuss how many and which agents play the role of bankers. For example, we show agents who are more connected to the market are better suited for this role since they have more to lose by reneging on obligations. We discuss some banking history and compare it with the predictions of our theory.

*We thank many colleagues for their helpful comments and discussions. Wright thanks the NSF for research support. The views expressed in this paper are our own and do not necessarily reflect those of the Federal Reserve Bank of Philadelphia or the Federal Reserve System. This paper is available free of charge at www.philadelphiafed.org/research-and-data/publications/working-papers/.

1 Introduction

Our goal is to study banking without making a priori assumptions about what banks are, who they are, or what they do. To this end, we adopt the approach of mechanism design. This method, in general, begins by describing an economic environment, by which we mean preferences, technologies, and certain frictions – including spatial or temporal separation, information problems, commitment issues, etc. One then tries to describe the set of allocations that are attainable, respecting both resource and incentive feasibility constraints. Sometimes one also describes allocations that are optimal according to particular criteria. One then looks at these allocations and tries to interpret the outcomes in terms of institutions that can be observed in actual economies. We want to see if something that looks like banking emerges as an outcome of this exercise. To reiterate, we do not take a bank as a primitive concept. Our primitives are preferences, technologies, and frictions; and we want to see if something like banking arises endogenously.

Much has been written about the virtues of mechanism design in general.[1] Our particular approach is close to that advocated by Townsend (1987, 1990). He describes the method as asking if institutions that we see in the world, such as observed credit or insurance arrangements, can be derived from simple but internally consistent economic models, whereby internal consistency we mean that one cannot simply assume a priori that some markets are missing, contracts are incomplete, prices are sticky, etc. Of course, something that looks like missing markets or incomplete contracts may emerge, but the idea is to lay out the environment explicitly and derive this as an outcome.[2] Simple models, with minimal frictions, may not generate arrange-

[1] Seminal contributors to mechanism design have recently been honored with a Nobel prize; go to http://nobelprize.org/nobel_prizes/economics/laureates/2007/ecoadv07.pdf for a general description of the approach.

[2] As Townsend (1988) puts it: "The competitive markets hypothesis has been viewed primarily as a postulate to help make the mapping from environments to outcomes more precise ... In the end though it should be emphasised that market structure should be endogenous to the class of general equilibrium models at hand. That is, the theory should explain why markets sometimes exist and

ments that resemble those in actual economies; for example, they typically predict that credit and insurance work better than the institutions we observe. So, one asks, what additional complications can be introduced to bring the theory more in line with what we see? We want to apply this method to study banking.

Obviously some frictions are needed, since models like Arrow-Debreu have no role for banks. As has been discussed often, frictionless models offers no role for any institution whose purpose is to facilitate the process of exchange. The simplest example is the institution of money. A classic challenge in monetary economics is to ask what frictions make money essential, in the following sense: Money is said to be essential when the set of allocations that satisfy incentive and other feasibility conditions is bigger or better with money than without it (Wallace 2001, 2008). We study the essentiality of banks in the same sense. Just like monetary economists ought not take the role of money as given, for this issue, we cannot take banks as primitive. In our environment, it is a good idea for the planner – or mechanism – to have some agents perform certain functions that resemble salient elements of banking: They take deposits, and their liabilities (claims on deposits) are used by others to facilitate the exchange process. This activity is essential: If it were ruled out, the set of feasible allocations would be inferior.[3]

The vast literature on banks is surveyed by Gorton and Winton (2002) and Freixas and Rochet (2008). Much of this research is based on informational frictions, including adverse selection, moral hazard, and costly state verification, that hinder the channeling of funds from investors to entrepreneurs. One can distinguish broadly three main strands. One approach originating with Diamond and Dybvig (1983) interprets banks as coalitions of agents providing insurance against liquidity shocks. Another approach pioneered by Leland and Pyle (1977) and developed by Boyd and

sometimes do not, so that economic organisation falls out in the solution to the mechanism design problem."

[3]Although we do not dwell on this issue in this paper, our approach is also useful to study whether the essentiality of banks and money is compatible.

Prescott (1986) interprets bank as coalitions sharing information in ways that induce agents to truthfully reveal the quality of investments. A third approach based on Diamond (1984) interprets banks as delegated monitors taking advantage of returns to scale. All of these papers provide useful insights into the functions of banking, although they often take as given institutional details that we want to endogenize, especially when we study how many agents should become bankers, which ones, and why, as well as their role in the payment process.[4]

The above information-based theory of banks has been criticized by Rajan (1998) on the grounds that it assumes agents have a limitless ability to contract. Rajan argues instead for a theory that relies on incomplete contracting. When the bank's abilities cannot be replicated in the market e.g., investment in reputation, then the bank is essential. Rajan focuses on limited enforcement to explain why it is difficult for two parties to contract. We agree with Rajan that limited enforcement is key to a theory of banking; however we go a step further and rather than taking the level of enforcement (or the degree of market incompleteness) as given, our mechanism design approach allows us to explain why enforcement is limited or inexistent for some agents.[5]

Compared with some previous works, we focus more on commitment issues and less on informational frictions, although imperfect monitoring is also an important element of the model. We highlight limited commitment because banking concerns

[4]Work on the Diamond-Dybvig model is a large branch of the banking literature; see Jacklin (1987), Wallace (1988, 1990), Peck and Shell (2003), Green and Lin (2003), Andolfatto et al. (2007), and Ennis and Keister (2008). Usually, if not always, Diamond-Dybvig models do not interpret the bank as a self-interested agent, but as a contract or a mechanism, nor do they derive which agents should be bankers. In the papers that emphasize information sharing or delegated monitoring, banks are agents, but their role is restricted to solving information problems, and again they typically do not derive endogenously which agents will be bankers. The fact that bank liabilities play a role in the payment system is usually not considered at all, but see Andolfatto and Nosal (2008), Huangfu and Sun (2008), Kiyotaki and Moore (2005), Williamson and Sanches (2008), He et al. (2005, 2008), Cavalcanti and Wallace (1999a, 1999b), Wallace (2005), and Mills (2008).

[5]Calomiris and Kahn (1991), Myers and Rajan (1998), and Diamond and Rajan (2001) belong to the incomplete markets approach to banking. These papers deal with the question of how to implement a given level of commitment and find that the structure of the bank portfolio matters.

intertemporal reallocation, and we want to take seriously incentives to make good on credit obligations. Our agents can use stored output as collateral to ameliorate commitment problems, but this does not work perfectly if collateral can be easily liquidated.[6] An implication is that delegated storage may be useful: If you deposit your output with a third party who has less incentive or ability to liquidate it for strategic reasons, others are more willing to give you credit. Thus, claims on deposits can be used to facilitate transactions, and this resembles banking. This activity can be part of an efficient arrangement even if the third party has an inferior storage technology. Thus, bank liabilities can be useful for payments even if dominated in return. Although other things being equal, it is obviously better if the bank has access to good storage or other investment opportunities.

The general idea is obviously correct that sellers often accept in payment the obligation of a third party – which can take the form of a note, check, credit/debit card, or other instrument issued by a commercial bank – when they would not accept your personal IOU. We want to ask, however, why the third party is less inclined than you to renege on obligations. In our approach, future rewards and punishments mitigate strategic behavior, but monitoring is imperfect (opportunistic deviations are detected only probabilistically). Agents with a higher likelihood of being monitored have a greater incentive to make good on obligations and, hence, are better suited to take on the responsibility of holding deposits. This is not a new insight – better monitoring is the key characteristic underlying banks in, for example, the Cavalcanti-Wallace (1999a, 1999b) model, but it is nonetheless a valid insight. However, it seems to pin a lot on the assumption that some agents are exogenously easier to monitor.

Therefore, we also assume that agents have different probabilities of gaining from market activity or have different stakes in the economic system. Even with equal

[6] If, for example, a debtor can consume the goods, his promise to deliver them out of storage may not be any more credible than a pledge to deliver goods out of future production. This is related to the discussion in Kiyotaki and Moore (2008) about why it is hard to borrow against capital. See also Mills and Reed (2008) and the references therein.

monitoring probabilities, those with a higher stake in the system are less inclined to deviate from proscribed behavior, because they have more to potentially lose from getting caught and punished (here punishment entails future autarky). Consistent with the history of banking, as we discuss in detail below, individuals with a greater connection to the market are better suited to be bankers because they have more to lose by reneging on obligations, even if they can be monitored as well as others. We then endogenize the monitoring intensity and find that it is cheaper to monitor those agents with a greater connection to the market. When we collect these results, we find that bankers should be agents with a combination of the following characteristics: They have access to good storage technologies or other investment opportunities; they can be more easily monitored because they have more to gain from the market and, therefore, more to potentially lose from strategic behavior.[7]

Even if these results are not too surprising, the analysis precisely identifies the relevant effects and the nature of the trade-offs when, for instance, we show how it may be desirable to sacrifice rate of return by depositing wealth with parties that are more easily monitored or have a bigger stake in the economic system. Similarly, when we choose which agents should be monitored and, hence, could be good bankers, we can show precisely how it is efficient to choose those with the right combination of a bigger stake in the system and better investment opportunities. Similarly, when we discuss the efficient number of banks, we can lay out the trade-off between having fewer bankers, which reduces monitoring costs, and having more deposits per bank, which increases incentives to misbehave. All of this comes directly out of a mechanism design approach, without primitive assumptions about what is a bank, who is a bank, or what banks do. And, as we discuss, these implications are useful for analyzing the history of banking. Therefore, we think the approach provides some new and interesting insights.

[7] The theory also predicts that agents who are more patient are better suited for the role of banker, although we emphasize this less.

The rest of the paper is organized as follows. Section 2 describes the basic economic environment, emphasizing the roles of temporal separation, limited commitment, and imperfect monitoring. Section 3 characterizes incentive feasible (IF) and optimal allocations in a baseline version of the model with one sector, by which we mean one group of ex ante homogeneous agents. Section 4 considers an economy with more than one sector, by which we mean different groups of ex ante heterogeneous agents (in general, they may have different attachments to the market, monitoring efficiencies, and storage technologies). This section contains the main result on the essentiality of activities that resemble banking: We show that it can be desirable to have agents make deposits with others (delegated storage), and claims on these deposits facilitate exchange. Section 5 provides results on which individuals are suited to be bankers, how to monitor them when it is costly, and why deposits can be useful for payments even if they are dominated in rate of return. Section 6 is the conclusion. Although the framework is quite tractable, there are some technical details involved in several of the proofs that we relegate to the Appendix.

2 The Environment

Time is discrete and continues forever. There are $N \geq 1$ different sectors, whereby a sector we mean a group of ex ante homogeneous agents, while in general, agents are heterogeneous across sectors. The role of sectors will be clear later. A representative sector has a set $\mathcal{A}$ of agents that each period is partitioned into three groups, $\mathcal{A}_0$, $\mathcal{A}_1$ and $\mathcal{A}_2$, with measures γ_0, γ_1, and γ_2. For now $\mathcal{A}$ is arbitrary (e.g., it could be finite or infinite). Agents take it as given that each period they belong to $\mathcal{A}_i$ with probability γ_i. Agents in $\mathcal{A}_1$ and $\mathcal{A}_2$ are called traders of type 1 and 2: They potentially produce, consume, and derive payoffs as described below. Agents in $\mathcal{A}_0$ are called nontraders; during that period they neither consume nor produce and derive a payoff normalized to 0. Among other things, this setup captures the idea that agents can have different stakes in the economy: A bigger γ_0 means that you have less to gain from participating in the market, which will play a critical role in the dynamic incentive conditions below.

In each period there are two goods, 1 and 2. Agents in $\mathcal{A}_1$ consume good 1 and produce good 2, whereas agents in $\mathcal{A}_2$ consume good 2 and produce good 1. Letting x_i and y_i denote consumption and production by type i, assume utility $U^i(x_i, y_i)$ is increasing in x_i, decreasing in y_i, and satisfies the usual convexity conditions. Also, $U^i(0,0) = 0$. A key friction is temporal separation: We divide each period in two and assume good i must be consumed in subperiod i. This implies a role for credit, since type 1 consumes before type 2. To have a notion of collateral, assume good 2 is produced in the first subperiod although it is consumed in the second and, hence, must be stored across subperiods. Another key friction is that type 2 agents cannot store good 2 for themselves, because this would eliminate temporal separation and any interesting discussion of credit. Thus, only a producer of good 2 can store it. A unit of good 2 stored in the first subperiod returns $1+\rho$ units in the second; we set $\rho = 0$ for now and consider $\rho \neq 0$ below. Although goods can be stored across subperiods, they fully depreciate across periods.

Collateral here works as follows: We can potentially get type 1 agents to produce good 2 by offering them good 1 in the first subperiod and then ask them to deliver good 2 in the second subperiod, after the production cost has been sunk. In a real sense, they are getting a loan to consume good 1, with a promise to deliver good 2 later, backed by storage. To make this imperfect, however, we let type 1 get liquidation utility λx_2 if they consume x_2 out of storage; and if $\lambda > 0$, there is an opportunity cost to delivering the goods, even if the production cost is sunk.[8] We allow a given type 1 agent to derive liquidation utility from any good 2, even if it was produced by another agent, including one from a different sector. However, only goods produced

[8] The assumption of linear liquidation value λx_2 is merely to ease the presentation; we could use $U^1(x_1, y_1, x_2)$, but it adds little other than notation.

within a sector enter $U^i(x_i, y_i)$, which means that any transfers across sectors here will be due to incentive considerations and not the usual gains for trade. Also, we assume that $U^1(x_1, y_1) + \lambda y_1 \leq U^1(x_1, 0)$ for any x_1, so that it is never efficient for type 1 to produce good 2 for their own (liquidation) consumption.

We focus on symmetric and stationary allocations, given by vectors $(x_1^i, y_1^i, x_2^i, y_2^i)$ for each sector i, plus, as we discuss later, descriptions of cross-sector transfers, storage, and liquidation. The planner, or mechanism, collects all production and allocates it to consumers. Therefore, agents deal directly with the planner, but this is relevant only to the extent that we do not restrict them to bilateral trade: The frictions here do not concern the search for a trading partner, but whether a given trader can be trusted to honor a deal. Note that the planner does not store good 2: It is stored by its producer in subperiod 1, then delivered to the planner in subperiod 2, who can pass it on to a consumer (our planner's job is only to organize trade, not to engage any any form of storage or other production). Assuming for now there are no transfers across sectors or liquidation, an allocation is resource feasible if $\gamma_2 y_2 = \gamma_1 x_1$ and $\gamma_1 y_1 = \gamma_2 x_2$ and hence can be summarized by $x = (x_{1,} x_2)$. To reduce notation, without affecting the interesting results, we set $\gamma_1 = \gamma_2 = \gamma$ and $\gamma_0 = 1 - 2\gamma$.

Let $\theta = \pi\gamma\beta/(1-\beta)$, where β is the discount factor across periods and π represents a monitoring technology in the following sense. Our planner, or mechanism, makes recommendations for production, trade and consumption. Agents are free to go along with these recommendations or to deviate from them, potentially facing the risk of punishment. Any deviation by an agent from a recommendation is detected, or monitored, with probability π.[9] When a deviation is monitored, the agent is punished

[9] Imperfect monitoring has obviously been studied by many people in different areas of economics, and we cannot survey them all here. In theories of money and banking, Kocherlakota (1998) makes clear the critical function of monitoring (what he calls memory) but only studies the extremes of perfect or no monitoring. Less extreme versions are studied by Kocherlakota and Wallace (1998), who assume actions can be monitored with a lag, and Cavalcanti and Wallace (1999a, 1999b), who assume some agents can be monitored while others cannot. Our version is that some deviations can be monitored probabilistically, while others cannot.

with autarchy, which means permanent banishment from future market production and consumption (one could consider weaker punishments, but permanent autarchy is obviously the most effective). Notice θ measures the risk of deviating: the probability of being detected π, times one's connection to the market γ, times the weight one puts on future participation $\beta/(1-\beta)$. This conveniently makes θ the critical parameter in the incentive conditions discussed below.

This completes the basic environment, but some special cases are perhaps worth mentioning. First, U^i may be additively separable, in which case we can make it quasi-linear without loss in generality by an appropriate choice of units, say $U^i = u^i(x_i) - x_j$. This means that U^1 is linear in one good and U^2 in the other. It would be different to make them both linear in the same good, say $U^1 = x_1 - v(x_2)$ and $U^2 = u(x_2) - x_1$, meaning that all the gains from trade accrue from x_2 (thus x_1 is only relevant for incentive reasons, if at all). We can even eliminate x_1 altogether and assume, again without loss in generality, that x_2 enters one function linearly, say $U^1 = -x_2$ and $U^2 = u(x_2)$. In this case, there is nothing one can do to reward type 1 for producing within a period, so incentives involve only future promises, while in general both intertemporal and intratemporal incentives matter. Since it nests all these cases, we study the general specification $U^i(x_i, x_j)$, although separable and quasi-linear examples are sometimes useful, mainly for reducing notation.[10]

3 A One-Sector Economy

For now, cross-sector transfers are impossible. We are interested in the set of (stationary and symmetric) IF allocations in a sector. The mechanism recommends an allocation, summarized by $x = (x_1, x_2)$, since consumption by type 1 equals production by type 2 and vice versa. A recommendation x can be implemented if it is IF,

[10] In particular, we do *not* use quasi-linear utility for the major simplication in Lagos and Wright (2005), despite a superficial resemblance in the applications and in the environments (with the multiple subperiods, a double coincidence problem, etc.).

which means that no one wants to deviate. Although we focus on the case in which agents cannot commit to future actions and therefore may deviate whenever they like, we begin by describing what might happen if they could commit to some degree.

3.1 Benchmark Allocations

When agents can fully commit, stationary allocations are constrained only by one participation condition

$$S(x_1, x_2) \geq 0, \tag{1}$$

where $S(x_1, x_2) \equiv U^1(x_1, x_2) + U^2(x_1, x_2)$ is the total surplus per period. In this case, IF allocations only have to generate a greater surplus than autarchy. In any period, we can have $U^i < 0$ for one i, as long as S is positive. Although we generally want to characterize the entire IF set, we can consider the ex ante Pareto optimal, or PO, allocation $x^o = (x_1^o, x_2^o)$ that maximizes S. It is easy to see that, with full commitment, x^o is always IF.

The above benchmark is essentially static: Interesting notions of credit and intertemporal incentives do not arise. In fact, full commitment here is equivalent to commitment for one period, *before* types are realized, even if agents cannot commit across periods. Suppose now that agents can commit only within a period *after* types are realized. We then have to ensure that both types want to participate each period,

$$U^i\left(x_i, x_j\right) + \frac{\gamma\beta}{1-\beta} S(x_1, x_2) \geq \frac{\left(1-\pi\right)\gamma\beta}{1-\beta} S(x_1, x_2),\ i = 1, 2,\ j \neq i.$$

The left side is the payoff to type i who follows the recommendation; the right is the payoff to a deviator, who gets detected and punished with probability π and gets away with it with probability $1-\pi$.[11] Using $\theta = \pi\gamma\beta/(1-\beta)$, the above inequalities

[11] A deviation here means neither producing nor consuming during that period; alternatives, like consuming but not producing, which can be interpreted as a change in the timing, can be analyzed similarly.

reduce to what we call the dynamic participation, or DP, conditions[12]

$$U^i\left(x_i, x_j\right) + \theta S(x_1, x_2) \geq 0,\ i = 1, 2,\ j \neq i. \tag{2}$$

Notice (2) implies (1) but not vice versa: With full commitment, we can always have $U^i < 0$ as long as $S \geq 0$, while now we can have $U^i < 0$ only if future rewards are sufficiently great, and this depends on θ. Also, now we may not be able to support the ex ante PO allocation x^o. Consider $U^1 = -x_2$ and $U^2 = u(x_2)$, so that x_2^o solves $\partial u/\partial x_2 = 1$ while x_1 is irrelevant. In this case, (2) for type 1 becomes $-x_2^o + \theta\left[u(x_2^o) - x_2^o\right] \geq 0$, which holds iff θ is large enough. This illustrates how intertemporal incentives matter. One can interpret this as a model of credit, but it misses some of what we want. In particular, while we must give agent 1 an incentive to produce for agent 2, once he agrees he cannot renege. That is, once agent 1 agrees to produce good 2 in the second subperiod in exchange for good 1 in the first subperiod, he is committed to honor this obligation.

To relax commitment, one could consider what happens when x_2 is actually produced in the second subperiod. Now to get type 1 to produce good 2, after he has already consumed good 1, we need

$$U^1\left(x_1, x_2\right) + \frac{\gamma\beta}{1-\beta} S(x_1, x_2) \geq U^1\left(x_1, 0\right) + \frac{\left(1-\pi\right)\gamma\beta}{1-\beta} S(x_1, x_2),$$

which reduces to

$$U^1\left(x_1, x_2\right) - U^1\left(x_1, 0\right) + \theta S(x_1, x_2) \geq 0. \tag{3}$$

Now IF allocations satisfy (2) for type 2 and (3) for type 1. This captures a notion of credit without commitment and (3) can be interpreted as a repayment constraint, saying we must provide type 1 with the incentive to honor his obligations ex post. In what follows we focus on a different model, however, where type 1 produces good 2 in the first subperiod and stores it but may or may not deliver it in the second

[12] This is a dynamic participation condition because agents can decide to participate every period after types have been revealed, as well as ex-ante, as captured by (1).

subperiod, because this is a useful way to think about collateral considerations and ultimately deposit banking.

3.2 The Baseline Model

If type 1 produces good 2 in the first subperiod and stores it to the second subperiod, he can derive liquidation value λx_2 by consuming instead of delivering it. Thus, he delivers the goods only if

$$U^1(x_1,x_2)+\frac{\gamma\beta}{1-\beta}S(x_1,x_2)\geq U^1(x_1,x_2)+\lambda x_2+\frac{(1-\pi)\gamma\beta}{1-\beta}S(x_1,x_2),$$

where the payoff to a deviator on the right involves consuming x_1 and producing x_2 in the first subperiod, then liquidating in the second, whence with probability π he is punished and with probability $1-\pi$ he is not. This can be simplified to what we call the repayment constraint

$$-\lambda x_2+\theta S(x_1,x_2)\geq 0. \tag{4}$$

If $\lambda=0$, (4) is implied by (1), so IF allocations are constrained only by (2). Intuitively, $\lambda=0$ implies the production cost is sunk when it comes time to deliver x_2, so collateral works very well – in fact it gets us back to full commitment.

If $\lambda>0$, however, in general the set of IF allocations is given by the DP constraint (2) for both types and the repayment constraint (4) for type 1. Let $\mathcal{F}_0$ denote the set of IF allocations. It is easy to show that $\mathcal{F}_0$ is convex. Also $\mathcal{F}_0$ is trivially nonempty as $(0,0)\in\mathcal{F}_0$. It contains more points, as long as we make some mild assumptions that imply that there are gains from trade.[13] Figure 1 shows $\mathcal{F}_0$ delimited by three

[13] Sufficient conditions for this are: i) λ is not too big, and ii) the slope of the curve defined by C_1 in the text below has a greater slope than the curve C_2 for type 2 at $(0,0)$, which follows from standard Inada conditions.

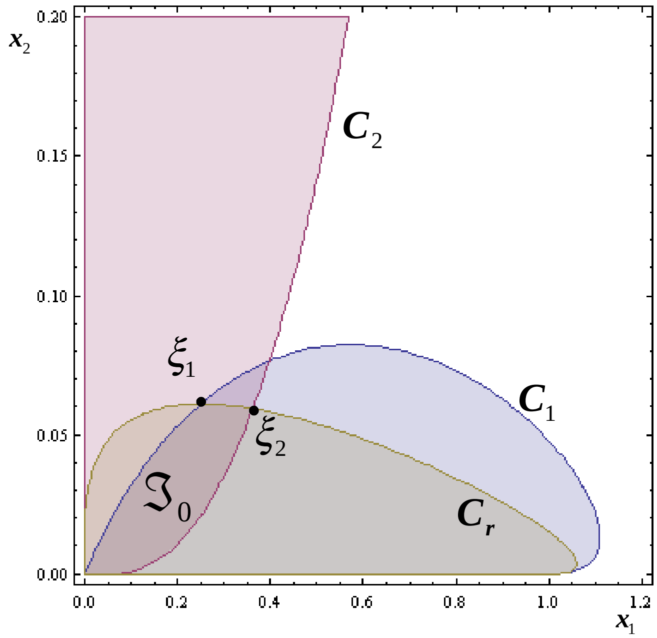

Figure 1: Incentive constraints and $\mathcal{F}_0$.

curves, defined by the three incentive conditions at equality:

$$
\begin{aligned}
C_1 &\equiv \left\{x : U^1(x_1, x_2) + \theta S(x_1, x_2) = 0\right\} \\
C_2 &\equiv \left\{x : U^2(x_2, x_1) + \theta S(x_1, x_2) = 0\right\} \\
C_r &\equiv \{x : -\lambda x_2 + \theta S(x_1, x_2) = 0\}
\end{aligned}
$$

Let $\xi_j \neq (0,0)$ be the point where C_r intersects C_j. We assume ξ_1 and ξ_2 are unique (a single-crossing condition), as shown in Figure 1.

In terms of economics, the key point is that the liquidation option implies an opportunity cost of delivering the goods, making collateral imperfect: A promise to deliver x_2 from storage may not be more credible than a promise to produce it. Given λ, collateralized credit only works if θ is big, similar to other forms of credit. So credit works better – that is, we can sustain a bigger and better set of allocations (x_1, x_2) – when agents are patient (β big), have a sizable connection to the market (γ is big), and are more easily monitored (π is big). Given θ, collateralized credit works better when λ is smaller.[14]

We now define the PO set after types are realized as the solution to maximizing per capita welfare

$$\max_{(x_1,x_2)} \mathcal{W}(x_1,x_2) = \omega_1 U^1(x_1,x_2) + \omega_2 U^2(x_2,x_1) + \frac{\beta}{1-\beta} S(x_1,x_2), \tag{5}$$

where ω_j is the Pareto weight on type $j = 0, 1, 2$ with $\omega_1\gamma + \omega_2\gamma + \omega_0(1-2\gamma) = 1$. For given ω_j, let x^* solve (5). This contrasts with the ex ante PO allocation x^o that maximizes $S(x_1, x_2)$. The FOCs for (5) are:

$$\frac{\partial U^1(x_1,x_2)}{\partial x_1} = -\frac{\partial U^2(x_2,x_1)}{\partial x_1} \frac{\omega_2(1-\beta)+\beta}{\omega_1(1-\beta)+\beta} \tag{6}$$

$$\frac{\partial U^2(x_2,x_1)}{\partial x_2} = -\frac{\partial U^1(x_1,x_2)}{\partial x_2} \frac{\omega_1(1-\beta)+\beta}{\omega_2(1-\beta)+\beta} \tag{7}$$

Thus, the PO set is given by

$$\mathcal{P} = \left\{ x \mid \frac{\partial U^1(x_1,x_2)}{\partial x_1} \frac{\partial U^2(x_2,x_1)}{\partial x_2} = \frac{\partial U^2(x_2,x_1)}{\partial x_1} \frac{\partial U^1(x_1,x_2)}{\partial x_2} \right\}. \tag{8}$$

Note that if $x \in \mathcal{P}$, then x_2 is an implicit function of x_1, with

$$\frac{dx_2}{dx_1} = -\frac{U_1^1(U_{22}^2 U_1^2/U_2^2 - U_{12}^2) - U_2^2(U_{11}^1 U_2^1/U_1^1 - U_{21}^1)}{U_1^2(U_{22}^1 U_1^1/U_2^1 - U_{12}^1) - U_2^1(U_{11}^2 U_2^2/U_1^2 - U_{12}^2)}, \tag{9}$$

where U_j^i is the partial of U^i with respect to argument j. In what follows, we assume that goods are normal, that is, for type 1 agents, consumption x_1 is increasing and production x_2 is decreasing in wealth, while for type 2 agents, consumption x_2 is increasing and production x_1 is decreasing, in a standard utility maximization problem. Normal goods implies all terms in parentheses in (9) are positive, and hence implies the curve $\mathcal{P}$ slopes downward in (x_1, x_2) space:

Lemma 1 *$\mathcal{P}$ is downward sloping in the sense that* $\dfrac{dx_2}{dx_1}_{|\mathcal{P}} < 0$.

[14]When $U^1 = u^1(x_1) - x_2$, for example, the repayment condition (4) reduces to $-\lambda x_2 + \theta S(x_1,x_2) \geq 0$ while (2) reduces to $u^1(x_1) - x_2 + \theta S(x_1,x_2) \geq 0$, and the former is more stringent iff type 1's surplus $u^1(x_1) - x_2$ exceeds the opportunity cost of liquidation $-\lambda x_2$. This is particularly transparent when $U^1 = -x_2$, in which case the repayment condition is more stringent iff the marginal opportunity cost of liquidation λ exceeds the marginal production cost.

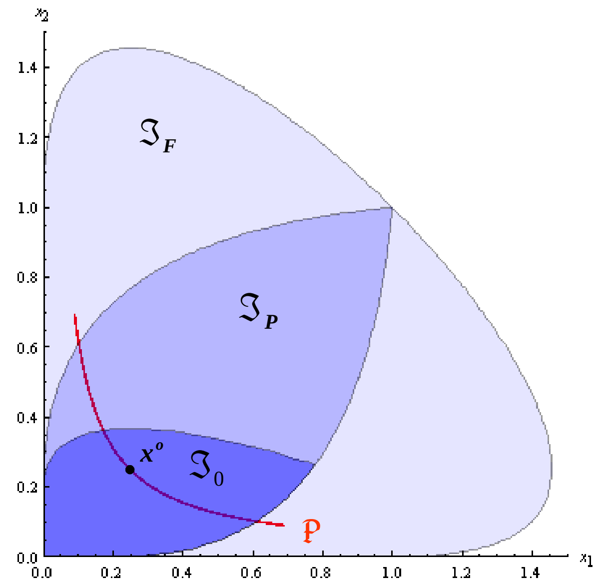

Figure 2: Pareto set and incentive feasible sets under full, partial, and no commitment.

Figure 2 shows $\mathcal{F}_0$ and $\mathcal{P}$ for some parametric examples ($U^i(x, y) = \sqrt{x} - y$, $\beta = 3/4, \gamma = \pi = 1/2$ and $\lambda = 1$). Note that $\mathcal{P}$ is downward sloping (Lemma 1). Also, $\mathcal{P}$ may or may not intersect $\mathcal{F}_0$, depending on θ and λ. Similarly, the ex-ante PO allocation x^o is in $\mathcal{P}$ but may or may not be in $\mathcal{F}_0$. For the sake of illustration, Figure 2 also shows the IF set when we have full commitment $\mathcal{F}_F$ and the IF set when we have partial commitment $\mathcal{F}_P$, with full and partial commitment as defined in Section 3.1.

Figure 3 shows what happens when we change β, π, or γ so that θ increases from θ^a to θ^b. Since this realxes all the incentive conditions, it shifts C_1 up, C_2 down, and C_r out, expanding $\mathcal{F}_0$. Notice also that both ξ_1 and ξ_2 shift to the northeast in this example. The following lemma tells us that this is a general result when C_r cuts C_1 from above (the proof is in the Appendix). Moreover, since decreasing λ affects only the repayment constraint, it shifts C_r out and does not shift C_1 and C_2. Therefore, decreasing λ also shifts ξ_1 and ξ_2 northeast.

Lemma 2 *If $\theta^a \leq \theta^b$ and $\lambda^a \geq \lambda^b$, then when C_r cuts C_1 from above, ξ_i^b lies northeast*

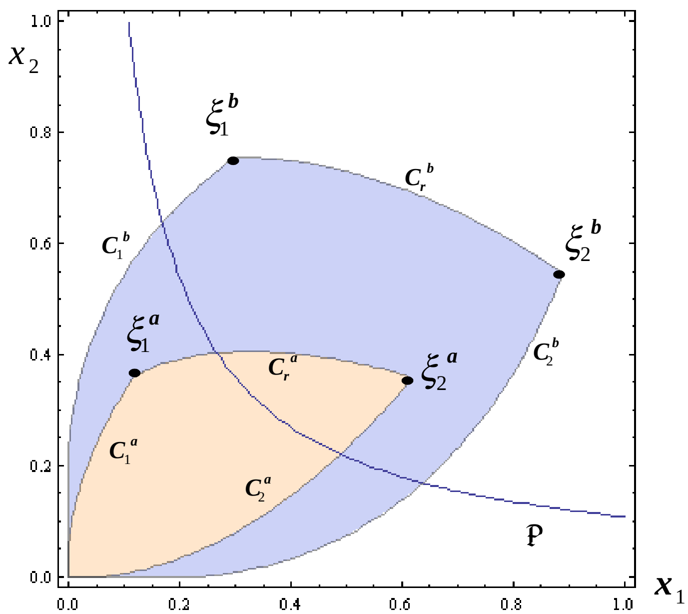

Figure 3: ξ_i, $i = 1, 2$, move northeast as λ decreases or θ increases, as C_r cuts C_1 from above.

of ξ_i^a *in* (x_1, x_2) *space. Otherwise,* ξ_1^b *lies southwest of* ξ_1^a *and* ξ_2^b *lies northeast of* ξ_2^a *in* (x_1, x_2) *space.*

4 A Multi-Sector Economy

It will suffice to focus on $N = 2$. Label the sectors a and b, and assume the following: Each period in sector $i = a, b$, with probability γ^i agents are type 1^i or 2^i traders, producing and consuming goods 1^i and 2^i; they are nontraders with probability $1-\gamma^i$; we can detect deviations with probability π^i; and type 1^i agents have a liquidation value λ^i.[15] Type 1^i can store and can liquidate either good 2^i produced in sector i or good 2^j produced in sector $j \neq i$. There are no gains from trade across sectors for pure mercantile reasons, because goods produced in sector i do not enter U^j and vice versa. But suppose t units of good 2 are produced by all type 1^b agents and

[15] We also assume that each sector contains the same population in the sense that the cardinality of $\mathcal{A}$ is the same. Later we relax this assumption.

transferred to type 1^a in the other sector, and the latter liquidate it. Assuming no one liquidates their own production, which is never efficient, utility for type 1^b is then $U^1(x_1^b, x_2^b + t)$, since they produce x_2^b for type 2^b in their own sector and t for type 1^b in the other sector, and utility for type 1^a is $U^1(x_1^a, x_2^a) + \lambda^a \gamma^b t/\gamma^a$, given the total transfer $\gamma^b t$ is shared equally among a measure γ^a of agents.

These transfers reduce total surplus. Still, they can have incentive effects, which we need to analyze for the following reason. We are ultimately interested in the case in which some good 2 is produced by type 1^b and transferred to type 1^a for storage, and the latter do not liquidate but transfer it back to (2^b agents in) sector b. This delegated storage, we think, captures some features of banking and is essential in the sense that it can change the IF set for the better. Now, direct transfers also change the IF set, but we claim that delegated storage can do better. To make this case, we must first analyze transfers. Without loss of generality, we focus on the case in which γ^i and π^i are such that $\theta^b \geq \theta^a$, which means type 1^b agents have less of a commitment problem than type 1^a, because they value future rewards more than the gain from short-term opportunism. If $\mathcal{F}_0^j$ denotes the IF set in sector j, then clearly $\mathcal{F}_0^a \subset \mathcal{F}_0^b$, as shown in Figure 3.

We also assume $\lambda^a \geq \lambda^b$, which implies that type 1 agents in sector a have a greater ability to liquidate or get greater utility from liquidating.

4.1 Direct Transfers

With transfers from sector b to a,[16] the surpluses in sectors a and b are

$$\begin{aligned} S^a\left(x_1^a, x_2^a, t\right) &= U^1\left(x_1^a, x_2^a\right) + U^2\left(x_2^a, x_1^a\right) + \lambda^a \gamma^b t/\gamma^a \\ S^b\left(x_1^b, x_2^b, t\right) &= U^1\left(x_1^b, x_2^b + t\right) + U^2\left(x_2^b, x_1^b\right). \end{aligned}$$

[16] We typically consider a transfer t in only one direction, for example from sector b to a, since transfers in the other direction are symmetric, and it is never useful to have simultaneous transfers in both directions. Also, note that one can think of t as a tax, but it is not compulsory: We can ask 1^b to voluntarily yield t, but they agree iff it satisfies the relevant incentive conditions.

Obviously we need $S^i(x_1^i, x_2^i, t) \geq 0$, but this is not binding, given the DP conditions: for type 2^i,

$$U^2\left(x_2^i, x_1^i\right) + \theta^i S^i\left(x_1^i, x_2^i, t\right) \geq 0,\ i = a, b; \tag{10}$$

and for type 1^i,

$$\begin{aligned} U^1\left(x_1^a, x_2^a\right) + \lambda^a \gamma^b t/\gamma^a + \theta^a S^a\left(x_1^a, x_2^a, t\right) &\geq 0 & (11)\\ U^1\left(x_1^b, x_2^b + t\right) + \theta^b S^b\left(x_1^b, x_2^b, t\right) &\geq 0. & (12)\end{aligned}$$

We also have the repayment constraints

$$-\lambda^i x_2^i + \theta^i S^i\left(x_1^i, x_2^i, t\right) \geq 0,\ i = a, b. \tag{13}$$

The IF set for the two-sector model with transfer t satisfies (10)-(13).

We can use $t > 0$ to relax the constraints in sector a, at the cost of tightening them in sector b. Notice that t affects the repayment constraint (13) only through the future surplus $S^i(x_1^i, x_2^i, t)$ and affects the DP conditions (11) and (12) directly. Whenever the constraints are binding in sector a but not b, $t \neq 0$ can expand the IF set. Therefore, transfers are essential in the technical sense used in the Introduction. To see how much one can accomplish with direct transfers, consider the allocation $\left(x_1^b, x_2^b\right)$ that entails the biggest transfer t from sector b to a, subject to (10)-(13). This is a standard maximization problem with a unique solution $\left(\tilde{x}_1^b, \tilde{x}_2^b, \tilde{t}\right)$. Since the left sides of the constraints are increasing in θ^b, so is $\tilde{t}$: When agents are more patient, more connected to the market, or more frequently monitored, we can extract more from them.

Some examples illustrate how this works. First, suppose $U^1(x_1, x_2) = x_1 - x_2$ and $U^2(x_2, x_1) = u(x_2) - x_1$. Further assume $\lambda^b = 0$, to make the example stark. Then IF allocations in sector b solve

$$\begin{aligned} u\left(x_2^b\right) - x_1^b + \theta^b\left[u\left(x_2^b\right) - x_2^b - t\right] &\geq 0 & (14)\\ x_1^b - u\left(x_2^b\right) + \left(1 + \theta^b\right)\left[u\left(x_2^b\right) - x_2^b - t\right] &\geq 0 & (15)\\ u\left(x_2^b\right) - x_2^b - t &\geq 0. & (16)\end{aligned}$$

The maximum transfer entails $\tilde{x}_2^b = x_2^*$, $\tilde{x}_1^b = u(x_2^*)$, and $\tilde{t} = u(x_2^*) - x_2^*$, where x_2^* solves $u'(x_2) = 1$. Notice $S^b(\tilde{x}_1^b, \tilde{x}_2^b, \tilde{t}) = 0$, and (14)-(16) hold with equality. Thus, we can get the agents to produce the x_2^* that maximizes the surplus and then tax away the entire surplus with $\tilde{t}$; because $\lambda^b = 0$, we do not have to worry about repayment. In this extreme case, θ^b is irrelevant, since $S^b(\tilde{x}_1^b, \tilde{x}_2^b, \tilde{t}) = 0$.

As another example, let $U^1(x_1, x_2) = -x_2$ and $U^2(x_2, x_1) = u(x_2)$, and set $\lambda^b = 1$. Now the IF allocations solve

$$u\left(x_2^b\right) + \theta^b\left[u\left(x_2^b\right) - x_2^b - t\right] \geq 0 \tag{17}$$

$$-x_2^b - t + \theta^b\left[u\left(x_2^b\right) - x_2^b - t\right] \geq 0 \tag{18}$$

$$-x_2^b + \theta^b\left[u\left(x_2^b\right) - x_2^b - t\right] \geq 0. \tag{19}$$

Clearly (17) and (19) are not binding, given $t \geq 0$, so we need to worry about only (18). The maximum tax is

$$\tilde{t} = \frac{\theta^b}{1+\theta^b} u\left(\tilde{x}_2^b\right) - \tilde{x}_2^b, \tag{20}$$

where $u'\left(\tilde{x}_2^b\right) = \left(1+\theta^b\right)/\theta^b$. Notice $\tilde{x}_2^b < x_2^*$. Also, notice $\partial \tilde{x}_2^b / \partial \theta^b > 0$ and $\partial \tilde{t} / \partial \theta^b > 0$. We have to provide type 1 with the incentive $\theta^b\left[u\left(x_2^b\right) - x_2^b - t\right]$ to cover the cost of producing both for type 2 in his sector plus the transfer to the other sector, $x_2^b + t$. And then, when he is supposed to deliver x_2 out of storage, we have to provide him with the incentive to cover the opportunity cost of liquidation $\lambda^b x_2^b = x_2^b$, but this is not binding when $t \geq 0$.

The point is that we can use transfers to extract $t > 0$ from 1^b in sector b and subsidize type 1^a agents, thereby relaxing constraints in sector a. Again, it is no surprise that taxes and transfers can increase the IF set; we present these results only to confirm that deposits can do even more than transfers.

4.2 Deposits: Delegated Storage

Again, the planner collects production and redistributes goods for consumption and storage in both sectors, but now in addition to transfers, the planner deposits $d \geq 0$

units of good 1^a with type 1^b. We still face the DP constraints (10) for type 2^i and (11)-(12) for type 1^a and 1^b, respectively. With deposits, type 1^a only stores $x_2^a - d$, so their repayment constraint is

$$-\lambda^a \left(x_2^a - d\right) + \theta^a S^a(x_1^a, x_2^a, t) \geq 0. \tag{21}$$

Similarly, type 1^b stores $x_2^b + \gamma^a d/\gamma^b$, so their repayment constraint is

$$-\lambda^b \left(x_2^b + \gamma^a d/\gamma^b\right) + \theta^b S\left(x_1^b, x_2^b, t\right) \geq 0. \tag{22}$$

If $d > 0$, then (22) implies (13) in sector b, while (13) implies (21) in sector a. Notice that deposits only interact with transfers in the liquidation value, not the continuation payoff. Finally, the planner faces the resource constraint[17]

$$0 \leq d \leq x_2^a. \tag{23}$$

With delegated storage, the IF set $\mathcal{F}_d$ is given by allocations $\left(x_1^a, x_2^a, x_1^b, x_2^b\right)$ with transfers t and deposits d such that (10)-(12) and (21)-(23) hold. We now prove one of the main results: Delegated storage $d > 0$ is essential.

Proposition 1 *$\mathcal{F}_0 \subseteq \mathcal{F}_d$ for any feasible $d > 0$ and for some parameters $\mathcal{F}_d \backslash \mathcal{F}_0 \neq \varnothing$.*

One can prove this as follows. First obviously any IF allocation with $d = 0$ is feasible when deposits are allowed. Then, to show that more allocations may be feasible, it suffices to give an example. Suppose that $\lambda^b = 0$. We claim that there are some allocations in sector a that are only feasible when $d > 0$. Set $t = \tilde{t}$, to maximize the transfer from sector b to a. Given $\left(x_1^b, x_2^b, t\right) = \left(\tilde{x}_1^b, \tilde{x}_2^b, \tilde{t}\right)$, all incentive constraints are satisfied in sector b. Now in sector a, the relevant conditions (10), (11), and (21) become

$$\begin{aligned} U^2\left(x_2^a, x_1^a\right) + \theta^a S^a\left(x_1^a, x_2^a, t\right) &\geq 0 \\ -U^2\left(x_2^a, x_1^a\right) + \left(1 + \theta^a\right) S^a\left(x_1^a, x_2^a, t\right) &\geq 0 \\ -\lambda^a\left(x_2^a - d\right) + \theta^a S^a\left(x_1^a, x_2^a, t\right) &\geq 0. \end{aligned}$$

[17] We can also allow agents in sector b to deposit goods in sector a by $d < 0$, which makes the resource constraint $-x_2^b \leq d \leq x_2^a$, but for now we focus on $d \geq 0$.

For any allocation such that

$$\lambda^a x_2^a \geq -U^2\left(x_2^a, x_1^a\right) \text{ and } \lambda^a x_2^a \geq U^2\left(x_2^a, x_1^a\right)\frac{\theta^a}{(1+\theta^a)}$$

increasing deposits will relax the repayment constraint and, therefore, expand the IF set. In fact, if we set $d = x_2^a$, the repayment constraint becomes $S^a\left(x_1^a, x_2^a, t\right) \geq 0$, which is redundant. Thus, $d > 0$ can expand $\mathcal{F}_0$ to $\mathcal{F}_d$. We will give some more detailed examples below, but first, we want to discuss what we think is the key economic idea. Suppose you want consumption now and pledge to later deliver something in return, out of stored inventory. When the time to make good rolls around, you are faced with a temptation to renege and liquidate the inventory. By depositing resources with a third party, this temptation is relaxed. Of course, one also has to consider temptation for the third party, but suppose for the sake of argument that this is not binding. Then deposits allow you to get more consumption now than private promises. One can interpret the third party as a bank with deposits as liabilities. And one can interpret these liabilities as helping to facilitate transactions: You are able to get more consumption because future payments to the producer are backed by deposits – by the banker's good name, so to speak. It is as if you offered a deposit receipt, like people exchanged bank notes historically or checks, debit cards, etc., in more modern times, as a means of payment.

We are not primarily concerned with details regarding how one might implement IF allocations: Our planner can in principle simply tell agents what to do and monitor them (albeit randomly) to verify compliance. To make this work, however, some record keeping is obviously required. For instance, we can use a spreadsheet to keep track over time of the agents who are supposed to deliver and receive goods. But suppose keeping records in this way is costly. Then it would be more efficient to adopt the following scheme: In the first subperiod, a type 1^a agent gives a type 1^b agent, his banker, goods in exchange for a deposit receipt; 1^a then gives the receipt to a type 2^a agent in exchange for goods; the banker 1^b stores the deposits, while the producer 2^a holds the asset, until the second subperiod; then 2^a redeems the asset with 1^b in exchange for goods. By showing the receipt to the planner, the banker 1^b can verify that he made good by clearing the payment. Although our model is abstract, this clearly resembles a banking arrangement.[18]

To illustrate how $d > 0$ expands the IF set in more detail, suppose $U^1(x_1, x_2) = x_1 - x_2$, $U^2(x_2, x_1) = u(x_2) - x_1$, $\gamma^a = \gamma^b$, and $\lambda^a > \lambda^b = 0$. The IF set in sector a is given by

$$u(x_2^a) - x_1^a + \theta^a \left[u(x_2^a) - x_2^a + \lambda^a t\right] \geq 0 \quad (24)$$

$$x_1^a - u(x_2^a) + (1 + \theta^a)\left[u(x_2^a) - x_2^a + \lambda^a t\right] \geq 0 \quad (25)$$

$$-\lambda^a (x_2^a - d) + \theta^a \left[u(x_2^a) - x_2^a + \lambda^a t\right] \geq 0. \quad (26)$$

Let $\tilde{t} = u(x_2^*) - x_2^*$ be the maximum feasible transfer. Increasing t from zero to $\tilde{t}$ relaxes all constraints. Increasing d from 0 relaxes the repayment constraint (26). Because the utilities of both agents are linearly separable in x_1^a, the repayment constraint is independent of x_1^a and it, therefore, defines an upper bound $\bar{x}_2^a$ for x_2^a.[19] Increasing d makes the upper bound bigger and expands the IF set. Figure 4 shows the IF set in sector a for three cases: $t = 0$ and $d = 0$ in red; $t = \tilde{t}$ and $d = 0$ in red and green; and $t = \tilde{t}$ and $d = x_2^a$ in red, green, and blue. Note that the IF set in sector b is independent of d.

The next example shows we do not need $\lambda^b = 0$. Consider $U^1(x_1, x_2) = -x_2$,

[18] As Selgin (2006) puts it, "Genuine banks are distinguished from other kinds of financial intermediaries by the readily transferable or 'spendable' nature of their IOUs, which allows those IOUs to serve as a means of exchange, that is, money... Commercial bank money today consists mainly of deposit balances that can be transferred either by means of paper orders known as checks or electronically using plastic 'debit' cards." We have more to say about these issues below, but we want to mention that many regard the English goldsmiths as the original modern bankers, precisely because their receipts circulated in payment (see Section 6 for references). These receipts were the first incarnation of banknotes. Shortly thereafter, they also allowed deposits to be transfered by "drawn note" or check. Originally, the depositors were specifically interested in safe keeping, which would be a consideration if we introduced theft; see He at al (2005, 2008) or Sanchez and Williamson (2008). Relatedly, one could also introduce counterfeiting considerations; see Nosal and Wallace (2007) and the references therein.

[19] Thus, $\bar{x}_2^a$ is an upper bound as by concavity of the utility function, $u(x_2^a) - x_2^a$ is decreasing beyond $x_2^a = x^*$.

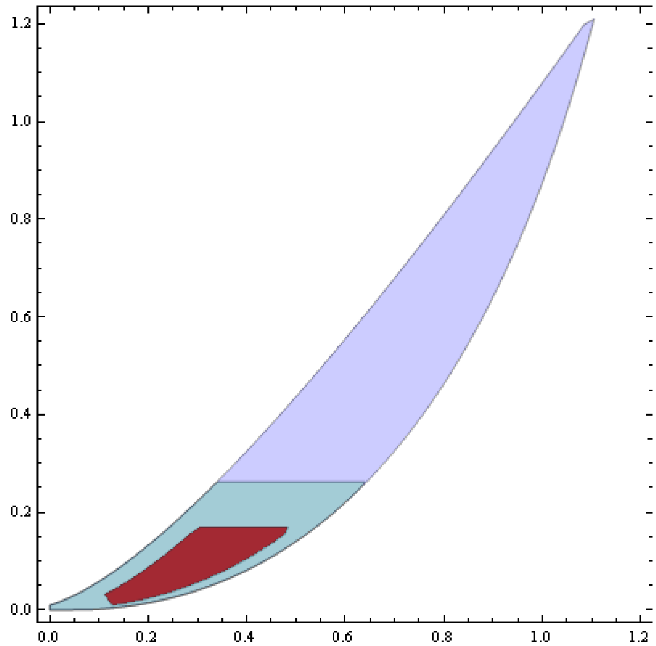

Figure 4: An example in which deposits are essential.

$U^2(x_2, x_1) = u(x_2)$, and $\lambda^a = \lambda^b = 1$; the only difference between sectors may be θ. Also, since x_1 does not enter U^1 or U^2 in this example, we can represent the IF set as $\left(x_2^a, x_2^b\right) \in \mathbb{R}^2$. For the moment, assume the planner seeks to expand the set of IF allocations in sector a. The DP constraints for types 2^a and 2^b are not binding. The relevant constraints in sector b are

$$-x_2^b - t + \theta^b \left[u\left(x_2^b\right) - x_2^b - t\right] \geq 0 \tag{27}$$

$$-x_2^b - d + \theta^b \left[u\left(x_2^b\right) - x_2^b - t\right] \geq 0. \tag{28}$$

Given $t = \tilde{t}$, we obtain

$$-\tilde{x}_2^b - \tilde{t} + \theta^b \left[u\left(\tilde{x}_2^b\right) - \tilde{x}_2^b - \tilde{t}\right] = 0 \tag{29}$$

$$-\tilde{x}_2^b - d + \theta^b \left[u\left(\tilde{x}_2^b\right) - \tilde{x}_2^b - \tilde{t}\right] \geq 0 \tag{30}$$

where $\tilde{t}$ and $\tilde{x}_2^b$ are defined by (20).

We get an upper bound on d by using the first equation to replace $\theta^b \left[u\left(\tilde{x}_2^b\right) - \tilde{x}_2^b - \tilde{t}\right]$

in the second,

$$d \leq \tilde{t} = \frac{\theta^b}{1+\theta^b} u\left(\tilde{x}_2^b\right) - \tilde{x}_2^b.$$

Hence $d = \tilde{t}$ is the biggest deposit type 1^b can accept given they transfer $\tilde{t}$ to sector a. At $d = \tilde{t}$, the constraints in sector b are tight, but those in sector a are most relaxed. The IF set in sector a is defined by

$$-x_2^a + \tilde{t} + \theta^a \left[u\left(x_2^a\right) - x_2^a + \tilde{t}\right] \geq 0 \tag{31}$$

$$-x_2^a + d + \theta^a \left[u\left(x_2^a\right) - x_2^a + \tilde{t}\right] \geq 0. \tag{32}$$

Since $d \leq \tilde{t}$, only (32) is binding. Therefore, increasing d from 0 to $d > 0$ expands the IF set in sector a. Symmetrically, we can also expand the IF set in sector b and note that $\theta^b \geq \theta^a$ implies $\tilde{t}^b \geq \tilde{t}^a$.[20]

For the example, $U^1(x_1, x_2) = -x_2$ and $U^2(x_2, x_1) = \sqrt{x_2}$. Figure 5 plots $\mathcal{F}_0^i$ in $\left(x_2^a, x_2^b\right)$ space for several cases[21]: $t^i = d^i = 0$ in dark red; $t^i > 0$ and $d^i = 0$ for $i = a$ in dark and medium red, and $i = b$ in dark red plus medium blue. For $t^i, d^i > 0$, $\mathcal{F}_0^b$ is given by the projection of the areas that are dark, medium, and light red on the x_2^b space, while $\mathcal{F}_0^a$ is given by the projection of the areas in dark red and in medium and light blue on the x_2^a space. One panel shows $\theta^a = \theta^b$, and the other shows $\theta^a < \theta^b$. Clearly, the IF set expands more in the sector with lower θ^j. When only transfers are used (dark red plus medium red or medium blue zones), there are two reasons for this. First, $\theta^b > \theta^a$ implies more resources can be extracted from sector b to transfer to a. Second, because utility is concave, transfers have different effects across sectors: Since agents in sector b can sustain better outcomes on their own, they do not gain as much from transfers. The intuition is different when it comes to deposits. Since

[20] Indeed we get

$$\frac{\partial \hat{t}}{\partial \theta} = \frac{1}{(1+\theta)^2} u\left(\hat{x}_2\right) + \frac{\theta}{1+\theta} u'\left(\hat{x}_2\right) - 1 = \frac{1}{(1+\theta)^2} u\left(\hat{x}_2\right) > 0,$$

where the last equality follows from the definition of $\hat{x}_2$ as the value for which $u'\left(\hat{x}_2\right) = (1+\theta)/\theta$.

[21] To be precise, given x_2^j, $\mathcal{F}_0^i$ is the projection of the area shown on the x_2^i space.

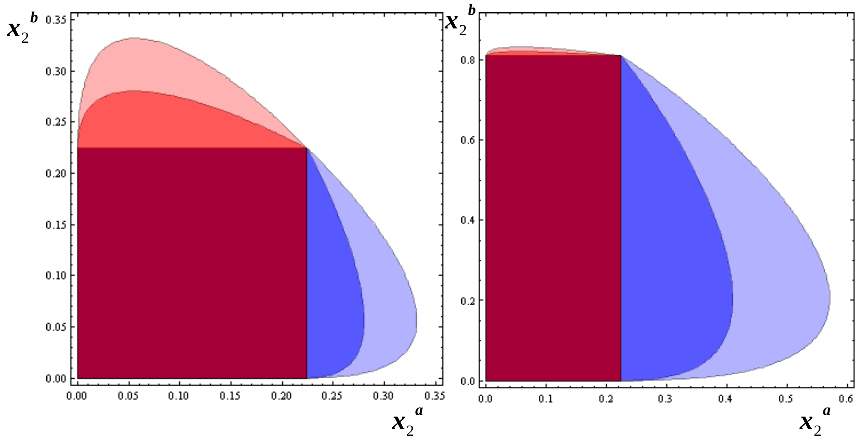

Figure 5: Left: $\theta^i = 0.9$ for $i = a, b$. Right: $\theta^a = 0.9$ and $\theta^b = 9$.

$\theta^b > \theta^a$, agents in sector b have more to lose if they liquidate storage. Hence, we can store more in sector b than a. We will expand on this in the next section.[22]

5 Extensions and Applications

5.1 Who Should Be a Banker?

Having shown that deposits are essential, we now ask who should hold them. To address this, suppose that the two sectors have the same preferences, but potentially different θ^i or γ^i. In terms of θ^i, this can be due to different π^i or γ^i. Even with equal monitoring probabilities, e.g. those with a higher γ^i are less inclined to deviate opportunistically: consistent with experience, as discussed below, individuals with a greater connection to the market are better suited to play the role of bankers, because they have more to lose by absconding with the deposits.

[22] Note that in this example, deposits alone do not expand the IF set. Suppose $t = 0$. Then (31) and (32) tell us the DP constraint (31) always implies the repayment constraint (32) for any $d \geq 0$: absent transfers, 1^a agents are not compensated for their production in subperiod 1, except through future payoffs. Deposits without transfers do not relax this constraint. This is, however, an artifact of the example, in which type 1 agents do not derive utility from good 1. In the appendix **??**, we provide examples in which deposits are essential even when $t = 0$.

Let $\hat{x}^i$ be the best IF allocation in sector i, i.e. it solves

$$\max_{x^i} \mathcal{W}^i\left(x^i\right) \text{ s.t. } x^i \in \mathcal{F}_0^i.$$

At $\hat{x}^i$, there is no IF allocation in sector i that makes one type better off without hurting the other type. The allocation $\hat{x}^i$ has the following property.

Lemma 3 *If the repayment constraint is not binding at $\hat{x}^i$, then $\hat{x}^i \in \mathcal{P}$.*

Proof: Let η^i and φ be the Lagrange multipliers on the DP constraint for type $i = 1, 2$, and the repayment constraint respectively. Suppose the repayment constraint is not binding so that $\varphi = 0$. Then the FOCs to the problem above are,

$$\begin{aligned}
\gamma\omega_1 U_1^1 + \gamma\omega_2 U_1^2 + \frac{\theta}{\pi} S_1 + \eta^1\left(U_1^1 + \theta S_1\right) + \eta^2\left(U_1^2 + \theta S_1\right) &= 0 \\
\gamma\omega_1 U_2^1 + \gamma\omega_2 U_2^2 + \frac{\theta}{\pi} S_2 + \eta^1\left(U_2^1 + \theta S_2\right) + \eta^2\left(U_2^2 + \theta S_2\right) &= 0
\end{aligned}$$

Using $S_i = \partial U^1/\partial x_i + \partial U^2/\partial x_i$, we obtain

$$\frac{\partial U^1/\partial x_1}{\partial U^1/\partial x_2} = \frac{\partial U^2/\partial x_1}{\partial U^2/\partial x_2}.$$

This implies $\hat{x} \in \mathcal{P}$. ■

Now we allow transfers and deposits and ask, given $\hat{x}^a$ and $\hat{x}^b$, can we make agents in sector j better off without hurting agents in sector i? Transfers alone clearly cannot help: A sector making a transfer is always worse off.[23] If deposits can help, we say deposits in sector j are Pareto essential, or PE.[24] More precisely, deposits in sector i are PE if there is an allocation $\left(\tilde{x}^a, \tilde{x}^b\right)$ and deposits $d > 0$ such that $\mathcal{W}^j\left(\tilde{x}^j\right) \geq \mathcal{W}^j\left(\hat{x}^j\right)$ for $j = a, b$ with at least one strict inequality. Notice that a necessary condition for PE in sector i is that the repayment constraint does not bind at $\hat{x}^i$. Otherwise, strictly positive deposits in sector i will make the repayment constraints tighter, thus shrinking the IF set.

[23] Here we do not ask whether t and d together could help, and we only study deposits when we impose $t = 0$.

[24] The term "essential" means that the set of IF allocation becomes bigger or better. Here we mean better, which is why we use the term "Pareto essential". Notice also that if one sector can be made better off, then both sectors can be made better off with a small transfer.

Proposition 2 *Suppose $\theta^b > \theta^a$ and $\lambda^a \geq \lambda^b$. Then only deposits in sector b are PE.*

Proof. The assumption $\theta^b > \theta^a$ and $\lambda^a \geq \lambda^b$ implies $\mathcal{F}_0^a \subset \mathcal{F}_0^b$. We consider different cases.

Case 1: Suppose $x^* \in \mathcal{F}_0^i, i = a, b$, then $d \neq 0$ is not PE.

Case 2: Suppose $x^* \in \mathcal{F}_0^b$ and $x^* \notin \mathcal{F}_0^a$. $d < 0$ can only hurt in sector a as it tightens the repayment constraint and cannot help in sector b. On the other hand, $d > 0$ does not hurt sector b and helps in sector a, if and only if the repayment constraint binds.

Case 3: Suppose $x^* \notin \mathcal{F}_0^b$ and $x^* \notin \mathcal{F}_0^a$. As seen in Figure 6, C_r^j cuts C_1^j and C_2^j, $j = a, b$, exactly once for $x \neq (0, 0)$, by assumption. We now show that $d < 0$ is not PE. To verify this, notice with no transfers that $d < 0$ is PE only if the repayment constraint is not binding at $\hat{x}^a$. Suppose this is true. Since $x^* \notin \mathcal{F}_0^a$, one DP constraint binds, so either $\hat{x}^a \in C_1^a$ or $\hat{x}^a \in C_2^a$. Suppose $\hat{x}^a \in C_1^a$ (the other case is similar). We know by Lemma 3 that $\hat{x}^a \in \mathcal{P}$, as shown in Figure 6. Also, x^* is northwest of $\hat{x}^a$, as moving along $\mathcal{P}$ toward x^* increases $\mathcal{W}$.

Since in case 3, $x^* \notin \mathcal{F}_0^b$, either $\hat{x}^b \in C_1^b$ or $\hat{x}^b \in C_r^b$. From Lemma 2, ξ_1^b lies northeast of ξ_1^a if C_r^i crosses C_1^i from above, and to the southwest otherwise. By Lemma 1, starting at $\hat{x}^a$, moving along $\mathcal{P}$ toward x^*, we hit C_1^b before we hit C_r^b, whether C_r^i cuts C_1^i from above or below. Hence, as shown $\hat{x}^b \in C_1^b$ and $\hat{x}^b \notin C_r^b$. This implies $d < 0$ cannot help, since the repayment constraint is not binding in sector b.

Finally, we show $d > 0$ may be PE. To see this, simply consider any economy in which the repayment constraint binds in sector a but not in sector b. It is easy to build examples like this (e.g., $\theta^a = 0$ and $\lambda^b = 0$). See Figure 7 for another example.

■

Figure 8 illustrates Proposition 2 for the case in which x^* is feasible in sector b but not in sector a. When deposits are not used, $\mathcal{F}_0^a$ is shown in orange, while $\mathcal{F}_0^b$ is shown in light red and blue. Since $x^* \in \mathcal{F}_0^b$, it solves the planner's problem, with $d = 0$.

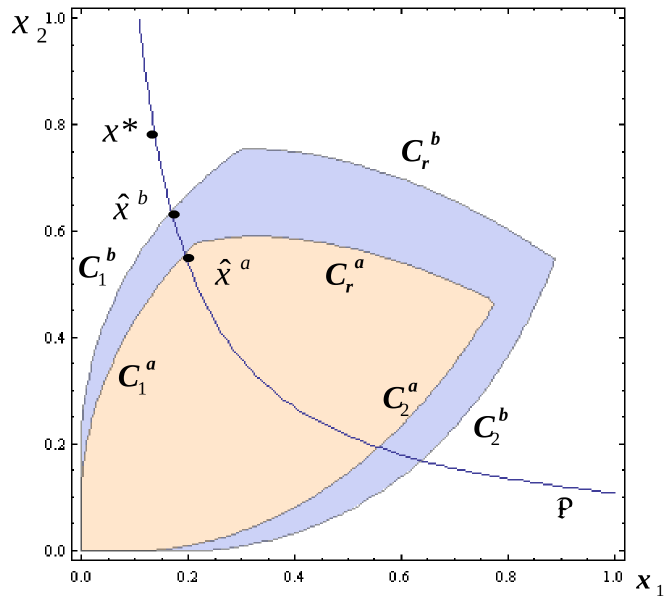

Figure 6: Case 3 - $d < 0$ is not PE.

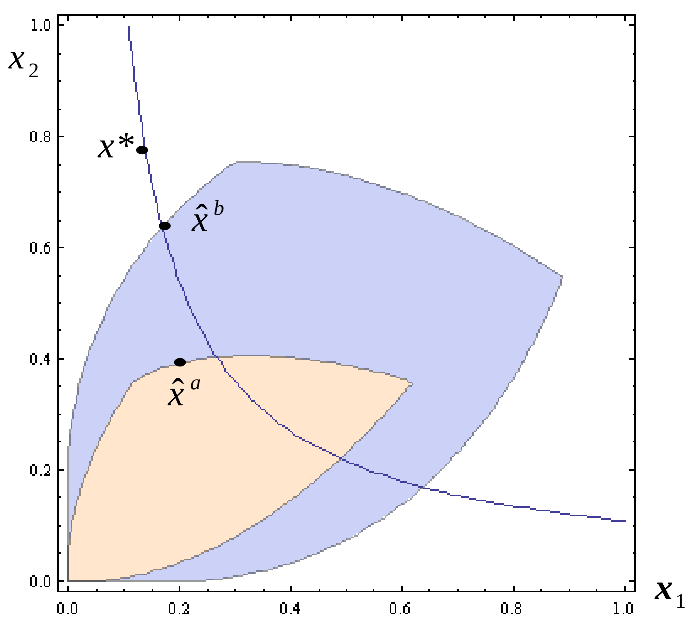

Figure 7: Case 3 - $d > 0$ is PE.

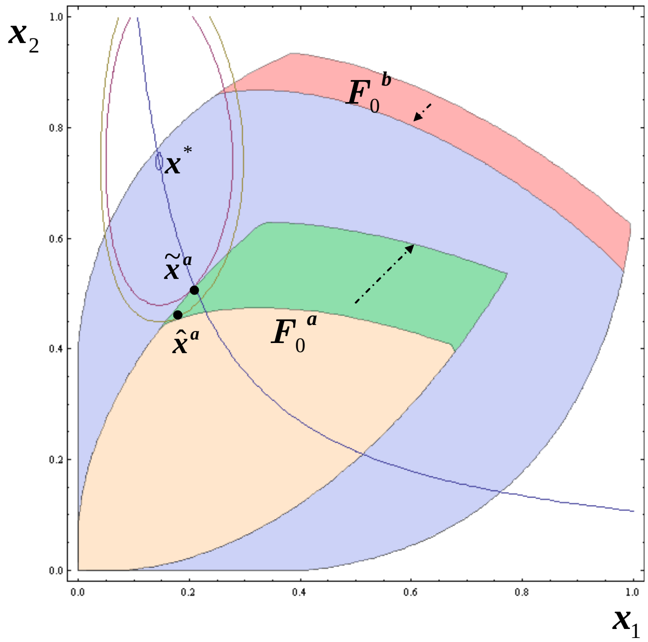

Figure 8: Illustration of Proposition 2.

But $x^* \notin \mathcal{F}_0^a$, and the figure shows indifference curves for the planner, including one through $\hat{x}^a$. Introducing $d > 0$ shifts the repayment constraint in sector b inward and the one in sector a outward, shrinking the IF set in sector b by the red area and increasing it in sector a by the green area. This has no effect on welfare in sector b, since x^* is still feasible, but makes sector a better off, since a better allocation $\tilde{x}^a$ is now feasible.

As shown, $\gamma^b > \gamma^a$ implies we can make sector a better off without hurting sector b by allowing the former to make deposits with the latter. Proposition 2 says we cannot make sector b better off without hurting sector a when $\gamma^b > \gamma^a$. In conclusion, those agents who have more at stake in the market, in the sense of larger γ, are more suited to become bankers. Of course, there is nothing special about γ; what matters is a higher θ, which in principle can be due to higher γ, π, or β. It is efficient to get others to produce goods by offering them consumption then take their output and store it with a banker, with a higher θ, since this means he can more credibly commit to honoring claims on these deposits than the original producer.

5.2 Who Should Be Monitored?

Monitoring is a costly activity, and we now analyze the planner's problem when he incurs a monitoring cost. To simplify the analysis, we distinguish between the probability of monitoring production, which we fix at $\pi \in [0,1]$ in both sectors, and the probability of monitoring repayment, which we denote by $\acute{\pi}^i$ in sector i and determine endogenously. We assume the cost of monitoring with probability $\tilde{\pi}^i$ is $\tilde{\pi}^i k^i$.[25] We now define a new benchmark with $d = 0$ in sector i as the pair $(\acute{x}^i, \acute{\pi}^i)$ solving

$$\max_{(x,\acute{\pi})} \mathcal{W}^i(x) - \acute{\pi} k^i \text{ s.t. } x \in \acute{\mathcal{F}} \text{ and } 0 \leq \acute{\pi} \leq 1, \tag{33}$$

where $\mathcal{W}^i(x)$ and $\acute{\mathcal{F}}_0^i$ are as above, except the repayment constraint is now

$$-\lambda x_2^i + \acute{\theta}^i S\left(x_1^i, x_2^i\right) \geq 0, \tag{34}$$

where $\acute{\theta}^i = \acute{\pi}^i \gamma\beta/(1-\beta)$. Clearly (34) is binding; otherwise, we could reduce monitoring costs.[26]

In this application, we are interested in minimizing total monitoring costs, rather than asking if deposits are PE.[27] Also, in this section, we assume there is exactly one agent in $\mathcal{A}_1$ in each sector at each date. This means that we have exactly one candidate banker in each sector and we simply want to know which one is the better candidate (below we also discuss what happens more generally). Note that we can still have different γ^a and γ^b, if we relax the assumption that the total population in $\mathcal{A}^a$ and $\mathcal{A}^b$ are the same.

[25] Later we relax this assumption and assume that the cost of monitoring is $c = k_0 + \tilde{\pi} k$ to allow for increasing returns to scale.

[26] Also, notice that x^* is never efficient when monitoring is endogenous: Reducing monitoring implies a first order gain, while moving away from x^* entails only a second order loss.

[27] With transferable utility and $\omega_1 = \omega_2 = \omega$, when monitoring is paid out of production, we can use transfer t to compensate type 1^b for the increase in monitoring cost and tax type 1^a for decrease in monitoring cost. Therefore, any decrease in total monitoring cost with transfer is Pareto improving.

If agents from sector j deposit goods with sector i, the planner can reduce $\acute{\pi}^j$. The difficulty is that the repayment constraint always binds in *both* sectors. Therefore, agents 1^i need to be monitored more intensively. Although this is costly, it may be desirable if total monitoring costs are lower. Precisely, deposits d in sector i reduce the overall monitoring costs if $\acute{\pi}^i$ increases to $\acute{\pi}^{i\prime}$ and $\acute{\pi}^j$ decreases to $\acute{\pi}^{j\prime}$ so that $\left(\acute{\pi}^{j\prime} - \acute{\pi}^j\right) k^j < \left(\acute{\pi}^i - \acute{\pi}^{i\prime}\right) k^i$. In the Appendix, we prove the following:

Proposition 3 *Fix* $\left(x^a, x^b\right) = \left(\acute{x}^a, \acute{x}^b\right)$. *If* $\gamma^b \geq \gamma^a$, $\lambda^b \geq \lambda^a$, *and* $k^b \leq k^a$, *then only deposits in sector* b *reduce total monitoring costs. If* $\gamma^b \geq \bar{\gamma}$ *(defined in the proof), then* $\acute{\pi}^a = 0$ *and only agent* 1^b *is monitored for repayment.*

With one agent of type 1 in each sector and $\gamma^b > \gamma^a$, agent 1^b has a bigger stake in the economy. As a consequence, his surplus is higher at the benchmark allocations than of agent 1^a, that is, $S\left(\acute{x}_1^b, \acute{x}_2^b\right) \geq S\left(\acute{x}_1^a, \acute{x}_2^a\right)$. Therefore, a relatively lower increase in monitoring probability is enough to eliminate opportunistic behavior by 1^b. Finally, if γ^b is large enough, then there will no monitoring of repayment in sector a, because all their production can be deposited with sector b.

We now consider an example in which the monitoring cost is borne by the type 1 agent in each sector. We show that deposits in sector b can reduce the total monitoring cost, while agent 1^a transfers resources to agent 1^b as a compensation for increasing the monitoring cost in sector b. We set $U^1\left(x_1, x_2\right) = -x_2$ and $U^2\left(x_2, x_1\right) = u\left(x_2\right)$. For simplicity, we assume the monitoring cost is paid up front by agents 1^a and 1^b and only once. For any allocation x_2^i, the DP and repayment constraints with deposits d and transfers t are

$$\begin{aligned}
-x_2^a - t - \pi_d^a k^a + \theta^a\left[u\left(x_2^a\right) - x_2^a\right] &\geq 0 \\
-\lambda^a\left(x_2^a - d\right) + \theta_d^a\left[u\left(x_2^a\right) - x_2^a\right] &\geq 0\text{, and} \\
-x_2^b - \pi_d^b k^b + \lambda^b t + \theta^b\left[u\left(x_2^b\right) - x_2^b\right] &\geq 0 \\
-\lambda^b\left(x_2^b + d\right) + \theta_d^b\left[u\left(x_2^b\right) - x_2^b\right] &\geq 0
\end{aligned}$$

where θ_d^i has the usual definition. An allocation x_2^i is feasible if there is a $\pi_d^i \leq 1$ such that the DP and repayment constraints hold. In the Appendix, we show conditions for which x_2^i is feasible (β high or λ^i, k^i small). We also show that the repayment constraints bind when $d = t = 0$. This gives us the monitoring probability π_0^i and, therefore, the total monitoring cost $c = \pi_0^a k^a + \pi_0^b k^b$ (as there is only one type 1 agent in each sector). Now consider $d > 0$ and $t > 0$. If d (and t) are small enough, the monitoring probabilities are again given by the repayment constraints holding at equality. Then the total monitoring cost c satisfies

$$\frac{\beta}{1-\beta}c = \frac{\lambda^a k^a \left(x_2^a - d\right)}{\gamma^a \left[u\left(x_2^a\right) - x_2^a\right]} + \frac{\lambda^b k^b \left(x_2^b + d\right)}{\gamma^b \left[u\left(x_2^b\right) - x_2^b\right]}.$$

Clearly, with deposits, the change in the monitoring cost is

$$\Delta c = \frac{\lambda^b k^b}{\gamma^b \left[u\left(x_2^b\right) - x_2^b\right]} - \frac{\lambda^a k^a}{\gamma^a \left[u\left(x_2^a\right) - x_2^a\right]}$$

To make sure agent 1^b is willing to incur the additional monitoring cost, agent 1^a should free up enough resources to compensate him. The resources that are freed up are the savings in the monitoring cost, or

$$\frac{1-\beta}{\beta}\frac{\lambda^a k^a d}{\gamma^a \left[u\left(x_2^a\right) - x_2^a\right]} \tag{35}$$

This is also the maximum transfer t that agent 1^b can expect, and its liquidation value must cover the extra monitoring cost he suffers, or

$$\lambda^b t \geq \frac{1-\beta}{\beta}\frac{\lambda^b k^b d}{\gamma^b \left[u\left(x_2^b\right) - x_2^b\right]} \tag{36}$$

Therefore, by combining (35) and (36), agent 1^b is made better off when taking deposits and receiving a transfer whenever

$$\frac{\lambda^a k^a}{\gamma^a \left[u\left(x_2^a\right) - x_2^a\right]} \geq \frac{k^b}{\gamma^b \left[u\left(x_2^b\right) - x_2^b\right]}.$$

We summarize this discussion in the Proposition that follows.

Proposition 4 *Let $U^1(x_1, x_2) = -x_2$ and $U^2(x_2, x_1) = u(x_2)$. For any feasible allocation $\left(x_2^a, x_2^b\right)$, deposits and transfers are Pareto improving if and only if*

$$\frac{\max\{1, \lambda^b\} k^b}{\gamma^b \left[u\left(x_2^b\right) - x_2^b\right]} \leq \frac{\lambda^a k^a}{\gamma^a \left[u\left(x_2^a\right) - x_2^a\right]}.$$

Again, the banker should be agent 1^b if he has more at stake, $\gamma^b \left[u\left(x_2^b\right) - x_2^b\right] > \gamma^a \left[u\left(x_2^a\right) - x_2^a\right]$; if he has less ability to default, by which we mean a more efficient monitoring techonology $k^b \leq k^a$; or if he has a lower value of liquidation, $\max\{1, \lambda^b\} \leq \lambda^a$.

Finally, notice that there is an interesting trade-off when we consider the efficient number of bankers in general. Fewer bankers reduce the total monitoring cost, other things being equal, but entail more deposits per banker; therefore we must use a higher probability of monitoring to get each one to behave well. Therefore, in general, there is an optimal number of bankers, which depends on the monitoring cost function. In fact, even if there is only one sector – that is, one set of ex ante homogeneous agents – if we were willing to consider asymmetric allocations, it would be desirable in general to designate some subset of them to be banker, and concentrate monitoring efforts on them. The trade-off would again be that fewer bankers reduce costs for a given level of monitoring but entail larger deposits per bank, which increases incentives to renege. For illustration, let us consider our last example with a monitoring cost function $k_0 + \pi k$, where $k_0 > 0$. This cost function implies increasing returns to scale. Suppose there are n agents of type 1^b, so n potential bankers. Given a feasible allocation x_2^b and deposits d, the monitoring probability – if there is a single banker – is given by the binding repayment constraint, so that

$$\pi_1 = \frac{\lambda^b k^b \left(n x_2^b + d\right)}{\gamma^b \left[u\left(x_2^b\right) - x_2^b\right]}.$$

Notice that the banker now has to store his own production x_2^b, the deposit d, as well as the production of all other agents 1^b, $(n-1)\, x_2^b$. The total monitoring cost is then $k_0 + \pi_1 k$. Now suppose we increase the number of bankers to $m \leq n$. Then the monitoring probability is

$$\pi_m = \frac{\lambda^b k^b \left(n x_2^b + d\right)}{m \gamma^b \left[u\left(x_2^b\right) - x_2^b\right]} = \frac{\pi_1}{m}.$$

It is π_1/m, as the total storage can now be split in m equal parts. The total monitoring cost, however, becomes $m\left(k_0 + \pi_2 k\right) = mk_0 + \pi_1 k$. This is clearly higher than the cost of monitoring in the economy with a single banker. Notice that we have assumed that the allocation x_2^b was feasible with a single banker in sector b. This may not be always true, in which case there is a minimum number of bankers necessary to sustain x_2^b. A moment's reflection will convince the reader that the minimum number of bankers m^*, guaranteeing that x_2^b is feasible given d, must satisfy $\pi_{m^*} = 1$. Clearly, this is also the optimal number of bankers in our example, and it is given by

$$m^* = \frac{\lambda^b k^b \left(n x_2^b + d\right)}{\gamma^b \left[u\left(x_2^b\right) - x_2^b\right]}.$$

Everything being equal, there are fewer bankers, therefore, the less they have to store, the more they have at stake or the lower their ability to default.

5.3 Rate Of Return Dominance

In this section, we show that efficient bankers need not have the best storage technology, if they are relatively better at commitment. This implies a simple rate of return dominance result. Consider again the case in which $U^1\left(x_1, x_2\right) = -x_2$ and $U^2\left(x_2, x_1\right) = u\left(x_2\right)$. Also, let $\gamma^a = \gamma^b$ and $\lambda^a = \lambda^b = 1$, and assume that the planner puts equal weights on types 1 and 2, $\omega_1 = \omega_2 = \omega$. Each unit stored in sector i returns $1 + \rho^i$, where for the sake of illustration we assume $\rho^a = \rho > 0 = \rho^b$. We show that, for some parameters, deposits in sector b are PE, despite a better storage technology in sector a. Intuitively, this explains why individuals deposit wealth in their checking accounts, despite the existence of alternative investments with higher returns: The agents holding these deposits can be counted on to make good on their obligations, making their liabilities good payment instruments.

We know from the earlier analysis that for this specification, the dynamic incentive constraints for type 2 do not bind. Hence, absent any interaction between the two sectors and given $\omega_1 = \omega_2$, the planner solves

$$\max_{x_2^i} \quad u(x_2^i) - \frac{x_2^i}{1+\rho^i} \tag{37}$$

$$\text{s.t.} \qquad -\frac{x_2^i}{1+\rho^i} + \theta^i[u(x_2^i) - \frac{x_2^i}{1+\rho^i}] \geq 0 \tag{38}$$

$$-x_2^i + \theta^i[u(x_2^i) - \frac{x_2^i}{1+\rho^i}] \geq 0 \tag{39}$$

Ignoring (38) and (39), the first best allocation $x_2^{*,i}$ solves $u'(x_2^{*,i})\,(1+\rho^i) = 1$. Hence, $\rho > 0$ implies $x_2^{*,a} > x_2^{*,b}$. Notice that as the return increases, type 1 can reduce production and sustain a given x_2^i. Therefore, when $\rho^i > 0$, only the repayment constraint (39) is relevant, and the dynamic participation constraint can be ignored. Define the θ^i for which $x_2^{*,i}$ is feasible by $\overline{\theta}^i$, so that $\theta^i < \overline{\theta}^i$ implies the repayment constraint is violated at $x_2^{*,i}$.

Given $\rho > 0$, we can have $\overline{\theta}^b > \overline{\theta}^a$, so that $x^{*,a}$ is feasible in sector a but $x^{*,b}$ is not feasible in sector b.[28] Here we focus on a particular case, in which deposits in sector b are PE. In the Appendix, we verify the following:

Proposition 5 *Deposits in sector b are PE if $\theta^a < \overline{\theta}^a$ and either: (a) $\theta^b \geq \overline{\theta}^b$ and $\theta^a\rho < 1+\rho$ or (b) $\theta^b < \overline{\theta}^b$ and $\theta^a\rho < \left(1 - \theta^a/\theta^b\right)(1+\rho)$.*

The condition $\theta^a < \overline{\theta}^a$ implies that $x_2^{*,a}$ is not IF in sector a, so that deposits potentially have a role. Then consider the situation in sector b. In the first case (a), agents in sector b do not have a commitment problem because $\theta^b \geq \overline{\theta}^b$, although they do have inferior storage technology. Therefore, making deposits in sector b requires agents in sector a to produce more to make up for the loss in return if they want to sustain a given level of consumption. There is a trade-off between commitment and

[28] In this case, even if we assumed $\theta^b > \theta^a$, the repayment constraint could bind in sector b but not in sector a. This would occur if, for example, $\overline{\theta}^b > \theta^b > \theta^a > \overline{\theta}^a$. There are many interesting possibilities, some of which we analyze in the working paper (Mattesini et al. 2009).

returns. The condition $\theta^a \rho / (1+\rho) < 1$ insures the commitment issue is sufficiently severe so that it is worthwhile for agents in sector a to give up something on the rate of return and deposit resources in sector b.

The second case (b) is similar, except that agents in sector b have a binding repayment constraint when $\theta^b < \overline{\theta}^b$. Therefore, they need to be compensated for taking deposits, to prevent default. A transfer from sector a does just that, but it comes on top of the additional production required from agents in sector a to cover for the loss in return. Hence, in this case, deposits in sector b are PE if $\theta^a \rho / (1+\rho) + \theta^a / \theta^b < 1$, which is a stricter condition than case (a). Also, as the storage technology of sector a improves, their commitment problem must be worse for deposits in sector b to be PE. In any case, the key message here is that bankers are not necessarily agents with the best investment opportunities, and for deposits to be used efficiently in payments they do not necessarily have to have the greatest return.

6 A Brief Digression on History

We have established that deposits are essential: For incentive reasons, it can be beneficial for an agent who wants something from a second agent to deposit goods with a third party – an intermediary – who holds them until they are withdrawn by the second party. The reason is that the third party may be more credible in terms of commitment to honor its obligations, either because this agent is more likely to be monitored or has more at stake if he gets caught. This can be an efficient arrangement, even if the third party does not have access to the best storage technology or, more generally, the highest return investment opportunities. As we said, we think this resembles banking. In this section, we go into a little more detail on banking history. We begin by discussing the fact that our theory of banking involves no outside money, although the deposit receipts discussed above constitute inside money. We then go on to discuss certain other features of banking history in the context of our model.

First, banks were historically institutions that accepted deposits for a variety of reasons, including safekeeping and access to investment opportunities. We already mentioned the English goldsmith bankers.[29] Going back in history, Mueller (1997) describes in considerable detail the medieval Venetian bankers. He distinguishes between regular deposits, which were specific goods that bankers had to deliver on demand, and irregular deposits, involving specie or coins that only had to be repaid with the same value (and not the same specie or coins). The depositor making an irregular deposit tacitly agreed to the investment by the banker of the deposits. One point we want to emphaisze is that deposits of real goods existed long before the invention of coinage (outside money) in Lydia in the 7th century.

In ancient Mesopotamia and Egypt, for example, mainly for security and to economize on transportation costs, goods were often deposited in palaces and temples. As Davies (2002) puts it:

> Grain was the main form of deposits at first, but in the process of time other deposits were commonly taken: other crops, fruit, cattle and agricultural implements, leading eventually and more importantly to deposits of the precious metals. Receipts testifying to these deposits gradually led to transfers to the order not only of depositors but also to a third party. In the course of time private houses also began to carry on such deposit business and probably grew to be of greater importance internally than was the case in contemporary Egypt. The banking operations of the temple and palace based banks preceded coinage by well over a thousand

[29] The story of the goldsmiths which is well known, is described in standard reference books as follows: "the direct ancestors of modern banks were, however, neither the merchants nor the scrivenors but the goldsmiths. At first the goldsmiths accepted deposits merely for safekeeping; but early in the 17th century their deposit receipts were circulating in place of money and so became the first English bank notes" (*Encyclopedia Britannica* 1954, vol. 3, p. 41). "The cheque came in at an early date, the first known to the Institute of Bankers being drawn in 1670, or so" (*Encyclopedia Britannica* 1941, vol. 3, p. 68). For more specialized references, see Joslin (1954) and Quinn (1997). Although many people call the goldsmiths the first modern bankers, some others give this credit to the Templars; see Weatherford (1997) or Sanello (2003) for their story.

years and so did private banking houses by some hundreds of years.

So while it is clearly interesting to analyze banking in monetary economies, we think it is interesting and historically relevant to discuss banking even in nonmonetary economies.

A very early development in the evolution of banking is that deposits were used to facilitate exchange. As in the model, throughout history a second party is more likely to give you something today if you can use the liability of a credible third party. As we said above, notes, checks, debit cards, and related instruments issued by commercial banks have this feature. Returning to Venice, Mueller (1997) explains how deposits "served a function comparable to that of checking accounts today; that is, it was not intended primarily for safekeeping or for earning interest but rather as a means of payment which facilitated the clearance of debts incurred in the process of doing business. In short, the current account constituted 'bank money,' money based on the banker's promise to pay."[30]

In any case, this system obviously can only work if bankers are relatively credible, or trustworthy. Our theory says that the more visiblean agent is or the more he has at stake, the more credible he becomes. The Rialto banks in medieval Venice offer evidence consistent with this idea. "Little capital was needed to institute a bank, perhaps only enough to convince the guarantors to pledge their limited backing and clients to deposit their money, for it was deposits rather than funds invested by partners which provided bankers with investable capital. In the final analysis, it was the visible pratimony of the banker – alone or as part of a fraternal compagnia – and his reputation as an operator on the market place in general which were placed on the balance to offset risk and win trust." (Mueller 1997, p. 97)

[30] According to many, those deposits did not actually *circulate* and transferring funds from one account to another "generally required the presence at the bank of both payer and payee" (Kohn 1999). This is the argument for regarding the goldsmiths, whose liabilities did circulate, as the operators of the first modern banks (see also Quinn 2002). But even if they did not circulate in this sense, the deposits of the earlier bankers cleary still facilitated payments.

Our model also uses random monitoring and makes the assumption that when a banker is caught cheating, he is expelled from the market. The direct evidence for this is scant, but the history of the Venetian bankers makes us think that these features are not far off. "In order to maintain 'public faith,' the Senate in 1467 reminded bankers of their obligation to show their account books to depositors upon request, for the sake of comparing records ... Penalty for noncompliance was set at 1,000 ducats." (Mueller 1997, p. 45). Thus, while it may have been prohibitively costly for depositors to continuously check the books, one can imagine that monitoring was performed every so often. And if caught cheating, the punishment was indeed lifetime banishment from any banking activity in Venice. Apparently, this happened very rarely – consistent with theory.

We also mention that many bankers historically started as merchants, who almost by definition have a greater connection to the market than a typical individual. As Kohn (1999) describes it, for example, the great banking families in Renaissance Italy and Southern Germany in the 16th century were originally merchants, who began lending their own capital and then started collecting deposits from other merchants, nobles, clerics, and small investors. They were not the wealthiest group; wealth then was concentrated in the hands of landowners, who controlled agriculture, forests, and mineral rights. But the merchants arguably had the most to lose from reneging on obligations. Thus, "because commerce involved the constant giving and receiving of credit, much of a merchant's effort was devoted to ensuring that he could fulfill his own obligations and that others would fulfill theirs." (Kohn 1999)

Also, returning again to Venice:

> In the period from about 1330 to 1370, eight to ten bankers operated on the Rialto at a given time. They seem to have been relatively small operators on average... Around 1370, however, the situation changed [and] Venetian noble families began to dominate the marketplace. After the

> banking crisis of the 1370s and the War of Chioggia, the number of banchi di scritta operating at any given time on the Rialto dropped to about four, sometimes as few as three. These banks tended, therefore, to be larger and more important than before. Their organizational form was generally either that of the fraterna or that of the partnership, the latter often concluded between a citizen and a noble. (Mueller 1997, p. 82)

As in our model, there seem to have been interesting issues concerning the efficient number of bankers, and revolving around greater credibility or commitment and larger amounts of deposits per bank.

Finally, what does our theory have to say about banking crisis in general and the recent financial crisis in particular? Gorton (2009) argues that the recent banking panic is a wholesale panic, whereby financial firms ran on other financial firms by not renewing sale and repurchase agreements. This is akin to a retail panic in which the depositors do not renew their demand deposits but choose to withdraw. According to Gorton, "depositors" were firms that lent money in the repo market. The location of subprime risks among their counterparties was unknown. Depositors were confused about which counterparties were really at risk and consequently ran all banks. While our framework is too simple to grasp the intricacies of the recent financial crisis, a small perturbation to our model can highlight exactly this one fundamental mechanism. As a perturbation, we can consider that the probability to be active in the market γ is subject to shocks (this is akin to technology shocks). It is then intuitive that the uncertainty surrounding the occurence of such a shock and its magnitude can induce agents to not renew their deposits (or deposit less) to reestablish the banker's incentives. What are these shocks? They depend on the nature of the firm's business. For example, γ could be affected by the housing market if the firm's business is to originate mortgage loans. More broadly, γ could also be affected by political events, such as a declaration of war, since wars can very easily shut down trade. On this point, let us return to Venice one last time, as Mueller (1997, p. 127) indeed writes "The concatenation of commercial failures with bank failures was a constant feature. (...) War or the threat of war was sufficient to close off a flourishing foreign market suddenly with the result that stocks accumulated locally; gluts brought a fall in prices which, in turn, caused failures."

In summary, although we can only touch on economic history in this paper, we think the key features of our theory and its implications are not inconsistent with the record.

7 Conclusion

This paper studied banking using a mechanism design approach. We began by describing an economic environment, with preferences, technologies, and certain frictions including temporal separation, imperfect monitoring, commitment issues, and costly record keeping. We described the set of IF allocations and optimal allocations. We did not start with assumptions about what banks are, who they are, or what they do. Rather, we looked at the set of IF or efficient allocations and tried to interpret the outcomes in terms of arrangements that resemble banking. In the model, it is efficient for certain agents, chosen endogenously based on their attachment to the market and our ability to monitor them, to accept deposits that will help faciliate exchange. This activity can be part of an efficient arrangement even if these agents do not have the best storage technologies or investment opportunities; if they have an advantage in commitment, this will make them more trustworthy. Of course, other things equal, it is better if bankers have good investment opportunities.

The arrangement generated by the model clearly resembles salient aspects of banking in both modern and historical contexts. We proved that this activity is essential: If we were to rule it out, the set of feasible allocations would be inferior. This was not a foregone conclusion – frictionless models do not have an essential role for banks. We also discussed issues related to who would make a good banker, how many bankers should we have, and who should be monitored when monitoring is costly. We think our approach is novel and complementary with other theories of banking. We also think it is consistent with economic history. Even if some of these results are not too surprising, one can use the theory to identify relatively precisely the relevant effects and the nature of the trade-offs. All of this comes directly out of a mechanism design approach, without primitive assumptions about what is a bank, who is a bank, or what banks do.

8 Appendix

8.1 Proof of Lemma 2

Let $\xi_i^j = \left(\xi_{i1}^j, \xi_{i2}^j\right)$. First notice that ξ_{11}^i is such that $U^1\left(\xi_{11}^i, \xi_{12}^i\right) = -\lambda\xi_{12}^i$ for $i = a, b$. Also, since $U^1(x_1, x_2) - U^1(x_1, 0) \leq -\lambda x_2$, for all (x_1, x_2) we easily get for all x_1

$$\frac{\partial U^1(x_1, x_2)}{\partial x_2} \leq -\lambda.$$

Let $\breve{x}_2$ and x_2' be the solution to:

$$U^1\left(\xi_{11}^i, \breve{x}_2\right) + \theta^i S\left(\xi_{11}^i, \breve{x}_2\right) = 0 \tag{40}$$

$$-\lambda x_2' + \theta^i S\left(\xi_{11}^i, x_2'\right) = 0 \tag{41}$$

To prove the lemma for ξ_1, it is enough to show that C_r^i shifts by more than C_1^i when θ increases. Obviously, $\theta^i = \theta^a$ implies $\breve{x}_2 = x_2' = \xi_{12}^a$. We now show $d\breve{x}_2/d\theta^a < dx_2'/d\theta^a$; this implies that C_r^i shifts by more than C_1^i as θ increases. Totally differentiating (40) while keeping ξ_{11}^a constant,

$$\frac{d\breve{x}_2}{d\theta^a} = \frac{-S}{\frac{\partial U^1}{\partial \breve{x}_2} + \theta^a \frac{\partial S}{\partial \breve{x}_2}} = \frac{\lambda\xi_{12}^a/\theta^a}{-\frac{\partial U^1}{\partial \breve{x}_2} - \theta^a \frac{\partial S}{\partial \breve{x}_2}}$$

where the last equality follows from $\breve{x}_2 = \xi_{12}^a$ at θ^a, so that $-\lambda\xi_{12}^a + \theta^a S\left(\xi_{11}^a, \xi_{12}^a\right) = 0$. The denominator is positive, by definition of C_1^a, so $d\breve{x}_2/d\theta^a > 0$. Now, totally differentiating (41), we obtain

$$\frac{dx_2'}{d\theta^a} = \frac{-S}{-\lambda + \theta^a \frac{\partial S}{\partial x_2}} = \frac{\lambda\xi_{12}^a/\theta^a}{\lambda - \theta^a \frac{\partial S}{\partial \breve{x}_2}}$$

where again the denominator is positive. Since $\lambda \leq -\partial U^1/\partial x_2$, we get that $d\breve{x}_2/d\theta^a < dx_2'/d\theta^a$. This proves the lemma for ξ_1^i.

We use a similar argument to prove the result for ξ_2^i. First, if ξ^i exists, it must be that C_r^i crosses C_2^i from above, because the slope along C_r^i is larger than the slope of C_2^i at $(0, 0)$. Hold ξ_{22}^a constant and define $\breve{x}_1$ and x_1' as the solution to

$$U^2\left(\xi_{22}^a, \breve{x}_1\right) + \theta^a S\left(\breve{x}_1, \xi_{22}^a\right) = 0 \tag{42}$$

$$-\lambda\xi_{22}^a + \theta^a S\left(x_1', \xi_{22}^a\right) = 0 \tag{43}$$

We show $dx_1'/d\theta^a > d\breve{x}_1/d\theta^a$, so that C_r^a shifts again more than C_2^a. Totally differentiating (42) while keeping $\breve{x}_2^a$ constant, we obtain

$$\frac{d\breve{x}_1}{d\theta^a} = \frac{-S}{\frac{\partial U^2}{\partial \breve{x}_1} + \theta^a \frac{\partial S}{\partial \breve{x}_1}} = \frac{\lambda \xi_{22}^a/\theta^a}{-\frac{\partial U^2}{\partial \breve{x}_1} - \theta^a \frac{\partial S}{\partial \breve{x}_1}}$$

where the last equality follows from the fact that at θ^a, $\breve{x}_1 = \xi_{21}^a$ so that $-\lambda\xi_{22}^a + \theta^a S\left(\xi_{21}^a, \xi_{22}^a\right) = 0$. Notice that the denominator is positive, by definition of C_2^a (for any allocation on C_2^a, we have $U^2 + \theta^a S = 0$ and any allocation below C_2^a satisfies $U^2 + \theta^a S < 0$; hence, starting from an allocation on C_2^a and increasing x_1 while leaving x_2 constant brings us below C_2^a so that $\partial U^2/\partial x_1 + \theta^a \partial S/\partial x_1 < 0$.) As a consequence, $d\breve{x}_2/d\theta^a > 0$. Now, totally differentiating (43), we obtain

$$\frac{dx_1'}{d\theta^a} = \frac{-S}{\theta^a \frac{\partial S}{\partial \breve{x}_1}} = \frac{\lambda\xi_{22}^a/\theta^a}{-\theta^a \frac{\partial S}{\partial \breve{x}_1}}$$

where again the denominator is positive. Since $-\partial U^2/\partial \breve{x}_1 \geq 0$, we get that $dx_1'/d\theta^a > d\breve{x}_1/d\theta^a$.■

8.2 Essential Deposits With $t = 0$: An Example

Here we show that deposits are essential, even when transfers are set to zero. We use an example where type 2 are endowed with a unit of good 1, instead of a production technology. We assume that agents of type 1 derive utility $\bar{u}$ from consuming this unit of good 1 so that $U^1(1, x_2) = \bar{u} - x_2$. Also, since there is no production of good 1 we set $U^2(x_2, x_1) = u(x_2)$. Assume the planner always transfers all of the endowment of good 1 to type 1. Also assume $\gamma^a = \gamma^b$, $\lambda^a = \lambda^b = 1$ and $\theta^b \geq \theta^a$. The relevant constraints in sector b are

$$\begin{aligned} \bar{u} - x_2^b - t + \theta^b\left[\bar{u} + u\left(x_2^b\right) - x_2^b - t\right] &\geq 0 \\ -x_2^b - d + \theta^b\left[\bar{u} + u\left(x_2^b\right) - x_2^b - t\right] &\geq 0. \end{aligned}$$

Absent deposits, we extract from sector b at most

$$\hat{t} = \frac{\theta^b}{1+\theta^b} u\left(\hat{x}_2^b\right) + \bar{u} - \hat{x}_2^b,$$

where as before $\hat{x}_2^b$ is the amount of good 2 that maximizes the transfer to sector a. Absent transfers, we can deposit with sector b at most

$$\hat{d} = \theta^b\left[u\left(\hat{x}_2^b\right) + \bar{u}\right] - \left(1+\theta^b\right)\hat{x}_2^b,$$

For ease of exposition, suppose $\bar{u} \geq \hat{d}$. We now consider how the constraints in sector a are relaxed when transfers or deposits alone are used. When only transfers are used, the dynamic participation constraint (11) and the repayment constraint (21) in sector a become

$$\begin{aligned} \bar{u} - x_2^a + \hat{t} + \theta^a \left[\bar{u} + u\left(x_2^a\right) - x_2^a + \hat{t}\right] &\geq 0 && (44) \\ -x_2^a + \theta^a \left[\bar{u} + u\left(x_2^a\right) - x_2^a + \hat{t}\right] &\geq 0 && (45) \end{aligned}$$

Clearly, only the repayment constraint (45) binds, and replacing the value for $\hat{t}$, it becomes

$$-x_2^a + \theta^a \left[\bar{u} + u\left(x_2^a\right) - x_2^a\right] + \theta^a \left[\frac{\theta^b}{1+\theta^b} u\left(\hat{x}_2^b\right) + \bar{u} - \hat{x}_2^b\right] \geq 0 \tag{46}$$

In the case where only deposits are used, the dynamic participation constraint and the repayment constraint become respectively

$$\begin{aligned} \bar{u} - x_2^a + \theta^a \left[\bar{u} + u\left(x_2^a\right) - x_2^a\right] &\geq 0 && (47) \\ -x_2^a + d + \theta^a \left[\bar{u} + u\left(x_2^a\right) - x_2^a\right] &\geq 0 && (48) \end{aligned}$$

Since $d \leq \hat{d} \leq \bar{u}$, the repayment constraint (48) implies the dynamic participation constraint (47). Also, as any feasible deposits satisfy $d \leq x_2^a$ the IF set is expanded most if

$$d = \min\{x_2^a, \hat{d}\}$$

For allocations where $x_2^a < \hat{d}$, (48) is equivalent to IR and hence redundant. In cases where $\hat{d} < x_2^a$, the repayment constraint becomes

$$-x_2^a + \theta^a \left[\bar{u} + u\left(x_2^a\right) - x_2^a\right] + \theta^b \left[u\left(\hat{x}_2^b\right) + \bar{u} - \frac{1+\theta^b}{\theta^b} \hat{x}_2^b\right] \geq 0 \tag{49}$$

If θ^a is low, (46) implies (49) and deposits implement more allocations than transfers. The figure below shows the IF set for cases where: $t = d = 0$ in dark red; $t > 0$ and $d = 0$ in dark and medium red; and $d > 0$ and $t = 0$ in dark, medium, and light red. x_2^a is represented on the x-axis, while x_2^b is on the y-axis. The key point is that, absent transfers, the upper bound $\bar{x}_2^a$ on the IF set is independent of x_2^b. This is not the case in sector b, as deposits from sector a influence incentives directly.[31]

[31] The figure is somewhat deceptive as it is not necessarily the case that the upper bound on the IF set in sector a with only transfers equals the one with only deposits. Also, note that for low $\bar{u}$, some allocations are feasible with only transfers and not with only deposits; indeed, for $\bar{u} = 0$, only transfers work.

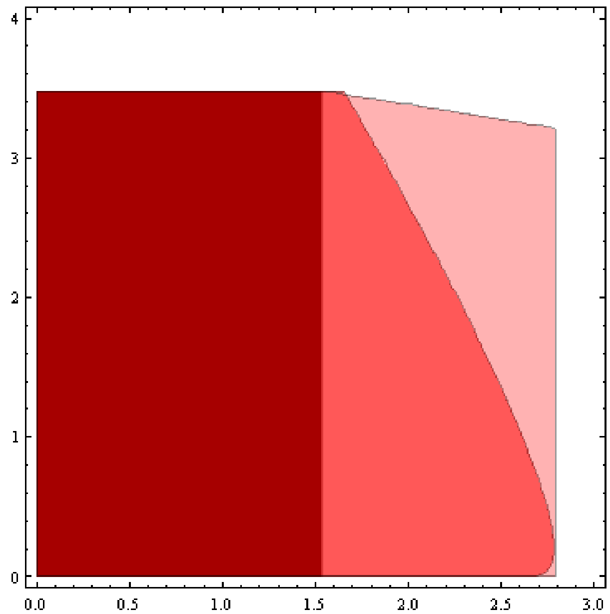

Figure 9: Fixed endowment.

8.3 Proof of Proposition 3

Since $\gamma^b > \gamma^a$, it must be that $S\left(\acute{x}_1^b, \acute{x}_2^b\right) \geq S\left(\acute{x}_1^a, \acute{x}_2^a\right)$. With deposits d, and since there is one candidate banker in each sector, the repayment constraint in sector b becomes $-\lambda^b\left(x_2^b + d\right) + \tilde{\theta}^b S\left(x_1^b, x_2^b\right) = 0$. Therefore, we obtain

$$\frac{\partial \acute{\pi}^b}{\partial d} = \frac{1-\beta}{\beta} \frac{\lambda^b}{\gamma^b S\left(\acute{x}_1^b, \acute{x}_2^b\right)}.$$

The repayment constraint in sector a is $-\lambda^a\left(x_2^a - d\right) + \tilde{\theta}^a S\left(x_1^a, x_2^a\right) = 0$, so that

$$\frac{\partial \acute{\pi}^a}{\partial d} = -\frac{1-\beta}{\beta} \frac{\lambda^a}{\gamma^a S\left(\acute{x}_1^a, \acute{x}_2^a\right)}.$$

Therefore, increasing deposits from sector a to b reduces the overall monitoring cost $\acute{\pi}^a k^a + \acute{\pi}^b k^b$ since

$$\frac{\partial \acute{\pi}^a}{\partial d} k^a + \frac{\partial \acute{\pi}^b}{\partial d} k^b = \frac{1-\beta}{\beta}\left(\frac{\lambda^b k^b}{\gamma^b S\left(\acute{x}_1^b, \acute{x}_2^b\right)} - \frac{\lambda^a k^a}{\gamma^a S\left(\acute{x}_1^a, \acute{x}_2^a\right)}\right) < 0,$$

where the inequality follows from $S\left(\acute{x}_1^a, \acute{x}_2^a\right) \leq S\left(\acute{x}_1^b, \acute{x}_2^b\right)$, $\gamma^a \leq \gamma^b$ and $k^b \leq k^a$. Hence, from $d = 0$, only $d > 0$ can reduce total monitoring cost.

To prove the second part of the proposition, let $\breve{x}^a$ solve $\max_x \mathcal{W}^a(x)$, subject to the DP constraints only. If

$$\bar{\pi} \equiv \frac{1-\beta}{\beta}\frac{\lambda^b\left(\acute{x}_2^b + \breve{x}_2^a\right)}{\gamma^b S\left(\acute{x}_1^b, \acute{x}_2^b\right)} \leq 1$$

then it is optimal to set $\breve{\pi}^b = \bar{\pi}$, $d = \breve{x}_2^a$, and $\acute{\pi}^a = 0$. $\bar{\gamma}$ is then defined as

$$\bar{\gamma} \equiv \frac{1-\beta}{\beta}\frac{\lambda^b\left(\acute{x}_2^b + \breve{x}_2^a\right)}{S\left(\acute{x}_1^b, \acute{x}_2^b\right)}$$

■

8.4 Existence of Feasible Allocations (Section 5.2)

With $d = t = 0$, the DP and repayment constraints are

$$\begin{aligned} -x_2^i - \pi_0^i k^i + \theta^i\left[u\left(x_2^i\right) - x_2^i\right] &\geq 0 \\ -\lambda^i x_2^i + \theta_0^i\left[u\left(x_2^i\right) - x_2^i\right] &\geq 0 \end{aligned}$$

Let π_P^i and π_R^i be the monitoring probability such that the DP and repayment constraint bind respectively. The repayment constraint is satisfied only if $\pi_0^i \geq \pi_R^i$. The DP is satisfied only if $\pi_0^i \leq \pi_P^i$. Therefore, given an allocation x_2^i, π_0^i exists only if $\pi_R^i \leq \pi_P^i$. Replacing the expression for π_R^i and π_P^i, we obtain

$$\frac{\lambda^i k^i}{\gamma^i}\frac{(1-\beta)}{\beta} \leq \left[\theta\frac{u\left(x_2^i\right) - x_2^i}{x_2^i} - 1\right]\left[u\left(x_2^i\right) - x_2^i\right]$$

An allocation x_2^i is feasible if this condition holds. Since the planner seeks to minimize the monitoring cost for each allocation x_2^i, he will set $\pi_0^i = \pi_R^i$ for any feasible allocation.

8.5 Proof of Proposition 5

The planner's problem with no interaction between sectors is given by (37). The first best solution is given by x_2^{*i} solving $u'(x_2^{*,i}) = 1/(1+\rho^i)$. Denote by $\underline{x_2^i}$, the level of x_2^i that satisfies the repayment constraint (39) as an equality. Because of concavity, $u(\underline{x_2^i})/\underline{x_2^i}$ is decreasing $\underline{x_2^i}$ and, therefore, we have

$$\underline{x_2^b} \geq \underline{x_2^a} \text{ iff } \theta^b - \theta^a \geq \theta^a\theta^b\frac{\rho}{1+\rho}. \tag{50}$$

Define $\overline{\theta}^i$ by $u(x_2^{*,i})/x_2^{*,i} = 1/\overline{\theta}^i + \overline{\theta}^i / (1+\rho^i)$ as the level of market connection below which the repayment constraint binds in sector i. These next two claims establish Proposition 5.

Claim 1 *Deposits in sector b are PE if*

$$\theta^a < \overline{\theta}^a,\ \theta^b < \overline{\theta}^b,\ \textit{and}\ \theta^a \rho < \left(1 - \frac{\theta^a}{\theta^b}\right)(1+\rho).$$

Proof. When $\theta^b < \overline{\theta}^b$, the solution to (37) in sector b is $\underline{x}_2^b$. Deposits are incentive compatible only if agents 1^a make a transfer τ to agents 1^b. The repayment constraint in sector b with transfer τ and deposits d evaluated at $\underline{x}_2^b$ is $-\underline{x}_2^b - d + \theta^b \left[u(\underline{x}_2^b) - \underline{x}_2^b + \tau\right] \geq 0$. By definition, $\underline{x}_2^b + \theta^a \left[u(\underline{x}_2^b) - \underline{x}_2^b\right] = 0$ and the minimum transfer τ that keeps the constraint satisfied is $\tau = d/\theta^b$. The repayment constraint with transfers in sector a is

$$-\left(x_2^a - d\right) + \theta^a \left[u\left(x_2^a\right) - \frac{x_2^a}{1+\rho} - d\frac{\rho}{1+\rho} - \tau\right] \geq 0. \tag{51}$$

Substituting $\tau = d/\theta^b$, the repayment constraint in sector a gives

$$\theta^a u\left(x_2^a\right) - x_2^a \left(1 + \frac{\theta^a}{1+\rho}\right) + d\left(1 - \frac{\theta^a \rho}{1+\rho} - \frac{\theta^a}{\theta^b}\right) \geq 0,$$

so x_2^a is increasing in d if $\theta^a \theta^b \rho < \left(\theta^b - \theta^a\right)(1+\rho)$.

Notice welfare in sector a is proportional to

$$u\left(x_2^a\right) - \frac{x_2^a}{1+\rho} - \frac{\rho}{1+\rho} d - \tau. \tag{52}$$

Subtituting $\tau = d/\theta^b$ into (52), as well as the value for d such that the repayment constraint is satisfied with equality and maximizing, welfare in sector a is increasing in d iff

$$u'\left(\underline{x}_2^a\right) > 1 + \frac{1}{\theta^b}.$$

This requires that $u'\left(\underline{x}_2^a\right) > 1 + 1/\theta^b = u(\underline{x}_2^b)/\underline{x}_2^b > u'(\underline{x}_2^b)$, the last inequality following from concavity. Therefore, we need $\underline{x}_2^a < \underline{x}_2^b$, or by (50), $\theta^a \theta^b \rho \leq \left(\theta^b - \theta^a\right)(1+\rho)$. Finally, the repayment constraint is still the relevant constraint when d is small enough. This completes the proof. ■

Claim 2 *Deposits in sector* b *are PE if*

$$\theta^a < \overline{\theta}^a,\ \theta^b \geq \overline{\theta}^b \text{ and } \theta^a \rho < 1 + \rho.$$

Proof. Given x_2^a and d, agents 1^a have to produce y such that $x_2^a = (y - d)(1 + \rho) + d$. For small d, only the repayment constraint is relevant,

$$-(x_2^a - d) + \theta^a \left[u(x_2^a) - \frac{x_2^a}{1+\rho} - d\frac{\rho}{1+\rho} \right] \geq 0. \tag{53}$$

To show deposits are PE in sector b, we show that increasing d relaxes the repayment constraint in sector a. The left side of (53) is increasing in d iff $\theta^a \rho < 1 + \rho$. The left side of (53) is decreasing in x_2^a. Therefore, $\theta^a \rho < 1 + \rho$ implies that increasing d allows higher x_2^a.

Finally, we need to show that welfare is increasing in d at $\underline{x}_2^a$. Welfare in sector a is proportional to

$$u(x_2^a) - \frac{x_2^a}{1+\rho} - \frac{\rho}{1+\rho} d.$$

Using (53) to eliminate d and maximizing with respect to x, welfare is increasing in x at $\underline{x}_2^a$ if

$$(1 + \rho) u'(\underline{x}^a) > \frac{1}{1+\rho} - \theta^a \left(\frac{\rho}{1+\rho} \right)^2.$$

Since $u'(\underline{x}^a) > 1/(1 + \rho)$, this condition is always satisfied. ■

References

Andolfatto, D. and E. Nosal (2008) "A Theory of Money and Banking," *Federal Reserve Bank of Cleveland Working Paper* 03-10.

Andolfatto, D., E. Nosal and N. Wallace (2007) "The Role of Independence in the Green-Lin Diamond-Dybvig Model," *Journal of Economic Theory* 137, 709-715.

Boyd, J.H. and E.C. Prescott (1986) "Financial Intermediary Coalitions," *Journal of Economic Theory* 38, 211-232.

Calomiris, C. and C. Kahn (1991) "The Role of Demandable Debt in Structuring Optimal Banking Arrangements," *American Economic Review* 81, 497-513.

Cavalcanti, R. and N. Wallace (1999a) "A Model of Private Bank Note Issue," *Review of Economic Dynamics* 2(1), 104-136.

Cavalcanti, R. and N. Wallace (1999b) "Inside and Outside Money as Alternative Media of Exchange," *Journal of Money Credit and Banking* 31(3), 443-457.

Davies, G. (2002) *A History of Money From Ancient Times to the Present Day*, 3rd. ed. Cardiff: University of Wales Press.

Diamond, D. and P. Dybvig (1983) "Bank Runs, Deposit Insurance, and Liquidity," *Journal of Political Economy* 91, 401-419.

Diamond, D. (1984) "Delegated Monitoring," *Review of Economics Studies* L1, 393-414.

Diamond, D. and R. Rajan (2001) "Liquidity Risk, Liquidity Creation and Financial Fragility: A Theory of Banking," *Journal of Political Economy* 109 (2), 287-327.

Ennis, H. and T. Keister (2008) "Commitment and Equilibrium Bank Runs," *Staff Report* 274, Federal Reserve Bank of New York, January.

Freixas, X. and Rochet, J.C. (2008) *Microeconomics of Banking.* Cambridge: MIT Press.

Gorton, G. and A. Winton (2002) "Financial Intermediation," in *Handbook of the Economics of Finance*, Constantinides, Harris and Stulz, eds. Amsterdam: North Holland.

Gorton, G. (2009) "Slapped in the Face by the Invisible Hand: Banking and the Panic of 2007," mimeo.

Green, E.J. and P. Lin (2003) "Implementing Efficient Allocations in a Model of Financial Intermediation," *Journal of Economic Theory* 109, 1-23.

He, P., L. Huang and R. Wright (2005) "Money and Banking in Search Equilibrium" *International Economic Review* 46, 637-670.

He, P., L. Huang and R. Wright (2008) "Money, Banking and Monetary Policy" *Journal of Monetary Economics* 55, 1013-24.

Huangfu, X. and H. Sun (2008) "Private Money and Bank Runs," Mimeo, Queen's University.

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

Joslin, D.M. (1954) "London Private Bankers, 1720-1785" *Economic History Review* 7, 167-86.

Kocherlakota, N. (1998) "Money is Memory," *Journal of Economic Theory* 81, 232-251.

Kocherlakota, N. and N. Wallace (1998) "Incomplete Record-Keeping and Optimal Payment Arrangements," *Journal of Economic Theory* 81, 272-289.

Kohn, M. (1999) "Early Deposit Banking," Working Paper 99-03, Department of Economics, Dartmouth College.

Kiyotaki, N. and J. Moore (2005) "Financial Deepening," *Journal of the European Economic Association* 3, 701-712.

Kiyotaki, N. and J. Moore (2008) "Liquidity, Business Cycles, and Monetary Policy," Mimeo, Princeton University.

Lagos, R. and R. Wright (2005) "A Unified Framework for Monetary Theory and Policy Analysis," *Journal of Political Economy* 113, 463-484.

Leland, H.E and D.H. Pyle (1977) "Informational Asymmetries, Financial Structure and Financial Intermediation," *Journal of Finance* 32, 371-387.

Mattesini, F., C. Monnet and R. Wright (2009) "Banking: A Mechanism Design Approach" Working Paper, Federal Reserve Bank of Philadelphia.

Mills, D.C. (2008) "Imperfect Monitoring and the Discounting of Inside Money," *International Economic Review* 49.

Mills, D.C. and R.R. Reed, (2008) "Default Risk and Collateral in the Absence of Commitment," Mimeo, Board of Governors.

Mueller, R.C. (1997) *The Venetian Money Market, Banks, Panics, and the Public Debt, 1200-1500*, Baltimore: Johns Hopkins University Press.

Myers, S.C. and R. Rajan (1998) "The Paradox of Liquidity," *Quarterly Journal of Economics*, 113 (3), 733-771.

Nosal, E.and N. Wallace (2007) "A Model of (the Threat of) Counterfeiting," *Journal of Monetary Economics* 54, 994-1001.

Peck, J. and K. Shell (2003) "Equilibrium Bank Runs," *Journal of Political Economy* 111, 103-123.

Quinn, S. (1997) "Goldsmith-Banking: Mutual Acceptance and Interbanker Clearing in Restoration London," *Explorations in Economic History* 34, 411-42.

Quinn, S. (2002) "Money, Finance and Capital Markets," in *The Cambridge Economic History of Britain since 1700, vol. 1, 1700-1860*, 3rd ed., R. Floud and D. N. McCloskey, eds. Cambridge: Cambridge University Press.

Rajan, R. (1998) "The Past and Future of Commercial Banking Viewed through an Incomplete Contract Lens," *Journal of Money, Credit, and Banking* 30 (3), 524-550.

Sanches D. and S. Williamson (2008) "Money and Credit with Limited Commitment and Theft," Mimeo, Washington University in St. Louis.

Sanello, F. (2003) *The Knights Templars: God's Warriors, the Devil's Bankers*. Lanham, MD: Taylor Trade Publishing.

Selgin, G. (2006) *Banking*, mimeo, forthcoming in *Encyclopedia Britannica*.

Townsend, R. (1979) "Optimal Contracts and Competitive Markets with Costly State Verification," *Journal of Economic Theory* 21, 265-93.

Townsend, R. (1988) "Models as Economies," *Economic Journal* 98, 1-24.

Townsend, R. (1989) "Currency and Credit in a Private Information Economy," *Journal of Political Economy* 97, 1323-1345.

Townsend, R. (1987) "Arrow-Debreu Programs as Microfoundations of Macroeconomics," in T.F. Bewley, ed., *Advances in Economic Theory, Fifth World Congress. Econometric Society Monographs Series* 12. Cambridge: Cambridge University Press.

Townsend, R. (1990) *Financial Structure and Economic Organization.* Basil Blackwell.

Wallace, N. (1988) "Another Attempt to Explain an Illiquid Banking System: The Diamond and Dybvig Model with Sequential Service Taken Seriously," *Federal Reserve Bank of Minneapolis Quarterly Review* 12, 3-16.

Wallace, N. (1990) "A Banking Model in which Partial Suspension is Best," *Federal Reserve Bank of Minneapolis Quarterly Review* 14 , 11-23.

Wallace, N. (2005) "From Private Banking to Central Banking: Ingredients of a Welfare Analysis," *International Economic Review* 46, 619-636.

Wallace, N. (2001) "Whither Monetary Economics?," International Economic Review 42, 847-869.

Wallace, N. (2008) "The Mechanism Design Approach to Monetary Economics," Mimeo in progress.

Weatherford, J. (1997) *The History of Money.* New York: Crown Publishers.