

Mechanism Design and Payments*

Thorsten Koepl

Queen's University

Kingston, Ontario

thor@qed.econ.queensu.ca

Cyril Monnet

European Central Bank

Frankfurt, Germany

cyril.monnet@ecb.int

Ted Temzelides

University of Pittsburgh

Pittsburgh, PA 15260

tedt@pitt.edu

Preliminary draft. First version: July 2004

June 11, 2005

Abstract

We introduce a model in which agents face random and unobservable needs to transact over time. We characterize the optimal setup for incentive compatible transactions, the *optimal payment system*. The existence of an equilibrium in which agents transact through a payment system might require certain caps on short-term borrowing. Networks that have knowledge of agents' histories support efficient transactions among their members. If agents transact frequently outside their network, however, incentive constraints imply that inter-network transactions can take place only at the cost of fewer intra-network transactions. Periodic settlement rounds are shown to be beneficial. Finally, the optimal payment system gives rise to a Friedman rule.

*We thank audiences at the Bank of Portugal, CEMFI, the Cleveland Fed and the European Central Bank. We especially thank Ed Green, Benjamin Lester, and Neil Wallace for several comments and suggestions. The views expressed in the paper are not necessarily those of the ECB.

1 Introduction

One of the features of the economy that the Walrasian model abstracts from is the mechanism through which transactions for goods and services take place. While for the study of certain questions this abstraction is one of the main strengths of the Walrasian model, this feature also makes it an inappropriate tool for the study of questions related to payments. Thus, new models are needed to study payments, and our goal is to develop such models using mechanism design.¹

There are several questions about the optimal structure of payment systems that motivate our work. For example, should transactions take place through a centralized system or through small networks? Should there be binding limits or “caps” imposed on the short-term borrowing by participants? What are the effects of reputation through repeated interactions with the system? What is the role of private information and imperfect monitoring in answering the above questions? Can local networks improve allocations through better information on local participants? Our approach emphasizes the role of private information. We are motivated by recent work by Kocherlakota (2004) who extends the model of Mirrlees (1971) and studies optimal taxation under private information. The design of optimal payment systems (and, more generally, of monetary policy) under imperfect monitoring is subject to a similar private information problem since participants’ liquidity position is not directly observable. Importantly, some of the questions that optimal payment system design poses are inherently *dynamic* and, therefore, very hard or impossible to study within the existing literature, which is almost exclusively static.²

We model the participants of the system as agents who face random needs for liquidity and random opportunities to build balances that can be used to meet these needs. To perform either of these activities, they need to interact through a network consisting of a large number of participants. We employ a version of the model of exchange developed by Kiyotaki and Wright

¹To get some idea of the magnitudes in actual systems involving bank payments, the value of the transactions processed through TARGET, the main public payment system in Europe, during March 2004 was over 40 Euro billions, with a daily average of about 1.7 Euro billions. In the United States, the average daily value of transfers through FEDWIRE, the US equivalent of TARGET, during the first quarter of 2004 was 1,683,265 \$ millions.

²See Kahn and Roberds (1998) for a well-cited paper in this literature.

(1989,1993). This model is appropriate for our study for several reasons. First, it is a model in which transactions are explicit, and where it is natural to study the implications of lack of commitment, imperfect monitoring, and reputation. Second, the random matching shocks that agents are subject to in this model are a tractable way of modelling random needs for liquidity. They simply capture the fact that certain participants need to transfer resources to other participants, and that these needs are subject to randomness. Third, the model is consistent with the fact that actual transactions *are* bilateral. Finally, this setup naturally lends itself to mechanism design. Unlike the standard monetary theory approach, our model involves *no currency*.

Our approach is related to the dynamic contracting literature.³ Agents in our model participate in two kinds of transactions. Within network transactions can be perfectly monitored. On the other hand, transactions between agents belonging to different networks are subject to imperfect monitoring. Thus, during such transactions, the payment system needs to induce truthful revelation. This is accomplished through adjustments in agents' balances, which can play a similar role to promised utility adjustments in Green's model.

Our findings suggest that a positive volume of transactions might require that certain caps are imposed on short-term borrowing. Put differently, in order for liquidity to be of value, it needs to be sufficiently scarce. Importantly, the introduction of caps implies a welfare loss, as it rules out certain welfare improving transactions that would take place under full information. This is an example of the common trade-off between efficiency and incentives, since truthful revelation comes at some cost. To avoid an unnecessary loss of certain transactions, caps should be set at the maximum value consistent with incentive compatibility. We find that local networks that have detailed knowledge of the members' histories can support efficient patterns of transactions among their members. However, if agents need to transact frequently outside their network, inter-network transactions might require that some intra-network transactions are sacrificed. Whether the efficient arrangement involves a centralized system or local networks depends on the relative frequency of transactions across networks.

Unlike the standard setup in the dynamic contracting literature, agents in our model transact bilaterally. Thus, in general, payoffs depend on the

³See, for example, Green (1987). Other classic references include Spear and Srivastava (1987) and Atkeson and Lucas (1993).

agents' histories, as well as on the histories of their current and future trading partners. Keeping track of such histories poses a problem that is hard to study analytically. Indeed, this is similar to the well-known problem of dealing with the distribution of money holdings in models of monetary exchange. This difficulty can be overcome in our setup by incorporating to the model a periodic *settlement stage*. We model this stage as a centralized Walrasian market in which agents can trade in order to optimize their balances.⁴

From the mechanism design point of view, our model studies optimal allocations under imperfect monitoring and no commitment. A novelty of our approach is that it involves non-cooperative implementation together with a Walrasian equilibrium aspect. We demonstrate that, under quasi-linear utility, the settlement stage implies that all agents exit the market with the same level of reserves. This greatly simplifies the contracting environment. More importantly, the introduction of periodic settlement rounds are shown to be beneficial. The full information first best allocation, in which the efficient volume of transactions take place both within and outside networks, can be supported under settlement. We demonstrate that the payment system that accomplishes this obeys a form of a Friedman rule.

The paper proceeds as follows. Section 2 describes the basic economic environment subject to indivisibility restrictions. The main topic of the paper, concerning optimal settlement under private information, is studied in Section 3. A conclusion follows.

2 Preliminaries

In this section we introduce the basic economic setup and discuss some specific examples. For expository purposes, in this section we study the case where the size of transactions is exogenous. Several findings turn out to provide useful intuition for the general environment that we study in the next section.

Time is discrete, t , measured over the positive integers. There is a $[0, 1]$ continuum of infinitely lived agents. To generate transactions, we assume

⁴This approach follows recent developments in monetary theory (see, for example, Lagos and Wright (2004) and references within). If we think of the agents in our model as banks, the centralized market has a meaningful interpretation that fits well with the institutional reality of payment systems: it corresponds to the periodic clearing at the end of a specified period of time (usually a day).

that in any given period, agents are randomly matched bilaterally. Randomness in payments is captured by assuming that an agent needs to transact with the agent he is matched with as a producer (consumer) with probability γ . Consumption gives utility u , while production gives disutility e , with $u > e$.⁵ Agents have a common discount factor, $\beta \in (0, 1)$.

To keep track of production opportunities, we introduce a random variable $s \in \{0, 1\}$, which equals 1 if a meeting is a *trade meeting* (an event of probability 2γ), and 0 otherwise. Throughout the paper we will assume that production is only possible in trade meetings. In such a meeting, we let p denote the probability with which an agent chooses to produce. Thus, $p(s) \in [0, 1]$ denotes the outcome in a meeting of type s . If $s = 1$ and $p(s) = 1$, we will assume that automatically, the consumer receives utility u , and the producer receives disutility e . An *allocation* within a match is a function $p : s \rightarrow [0, 1]$. Throughout, we will study symmetric stationary allocations that can be supported by strategies that constitute perfect equilibria. We will refer to such allocations as *incentive feasible allocations (IFAs)*.

To familiarize the reader with the setup consider the following two extremes. First, assume that agents are *anonymous* and that there are no assets in the economy. An additional important feature of this environment is that there is *no commitment*. The above assumptions rule out reputation effects, as well as any type of trade using currency or any other asset. Clearly, the only IFA in the above economy is autarky, i.e., $p(s) = 0$ for all s . Next, consider the other extreme. Assume the existence of a perfect monitoring and record-keeping technology that allows for the types and actions of all agents to be perfectly observed and recorded in every period. In this case, a “credit” equilibrium can be sustained through a standard reputation argument. In other words, provided that β is sufficiently high, $p(s) = 1$, if $s = 1$, constitutes an IFA, i.e., a transaction takes place whenever production is desired.

The interesting cases concern situations in between these two extremes. To begin studying such cases, consider the following example. Assume that a

⁵The reader might wish to interpret the payment system participants as banks. In that case, we could think of each bank as being associated with a client. When a client of one bank produces for a client of another, some payment needs to be transferred to the producing party. In our setup the client-bank pair is a single economic unit. Thus, when a client produces, the client-pair bank suffers disutility, and similarly for consumption. This allows us to concentrate on the incentives within the payment system without modelling the bank-client relationship itself.

perfect monitoring technology is not available. Instead, each agent is endowed with the ability to costlessly record his past consumption (production) with a *clearinghouse*. The clearinghouse cannot verify whether agents have a production opportunity within a given period. However, when it takes place, production can be verified. In that case, if, say, i produces for agent j , the “account” of i with the clearinghouse is credited by “+1,” while that of j receives an entry of “−1.” Under this scheme agents have the technology to record-keep their transactions by building balances with the clearinghouse. We denote an agent’s balance by d (an integer, not restricted in sign, no upper or lower bound). We will refer to d as the agent’s *state*. The fact that there is no bound on d can be interpreted as a clearinghouse policy that imposes no caps on individual borrowing. We now index the allocation (production decision) in a match involving an agent with balance d by $p_d(s)$.

Clearly, no agent has an incentive to build up a balance as he can always claim that he did not have a production opportunity. At the same time, since there is unlimited borrowing offered by the clearinghouse, declining to increase ones balance does not by itself decrease the probability of consuming in the future. Therefore, for any β , the only IFA is autarky. This is interesting because it suggests that caps might be necessary in order for transactions to take place. Put differently, in order for liquidity to be valuable it has to be scarce. In the next section we will study the effects of caps on agents’ borrowing.

2.1 Stationary Incentive Feasible Allocations

Here we introduce caps on agents’ borrowing. No penalty will be imposed on agents that hit the cap other than that they cannot borrow further unless they produce in order to improve their balance. We will demonstrate that, provided that the discount factor is sufficiently high, this policy can support a positive volume of transactions. Intuitively, the existence of a cap, C , implies that liquidity now becomes sufficiently scarce to be of value. In addition, sufficiently patient agents will produce in order to avoid being in a transaction in which the lack of liquidity prevents them from enjoying consumption. This identifies an interesting trade-off. If the clearinghouse provides little or no liquidity, by setting the cap close to or equal to zero, some welfare improving transactions will not be realized. In the context of the model, this occurs if an agent faced with a consumption opportunity has hit the borrowing cap. To minimize the frequency of such inefficient instances,

the clearinghouse should set the borrowing cap as high as possible. This, however, might result in the non-existence of an equilibrium with a positive volume of transactions.

We now turn to a characterization of IFAs. Let $\mathbf{p} = [p_{-C}, \dots, p_0, p_1, \dots]$ denote the vector of allocations and let $\mathbf{x} = [x_{-C}, \dots, x_0, x_1, \dots]$ denote the distribution of agents (both in the population and per type) across states. In an IFA, associated with a positive cap, C , we have the following value functions for an agent in state d :

$$v_{-C} = \gamma(1 - x_{-C})[p_{-C}(-e + \beta v_{-C+1}) + (1 - p_{-C})\beta v_{-C}] + [1 - \gamma(1 - x_{-C})]\beta v_{-C}, \quad (1)$$

$$v_{d > -C} = \gamma \mathbf{p} \mathbf{x} [u + \beta v_{d-1}] + \gamma(1 - x_{-C})[p_d(-e + \beta v_{d+1}) + (1 - p_d)\beta v_d] + [1 - \gamma p x - \gamma(1 - x_{-C})]\beta v_d. \quad (2)$$

The difference in the two value functions captures the fact that an agent that has hit the borrowing cap cannot consume. For the allocation to be incentive feasible, agents must be better off when they choose strategies which result in this allocation. That is, for all d such that $p_d = 1$, we must have

$$\begin{aligned} -e + \beta v_{d+1} &\geq \beta v_d, \quad \forall d \geq -C, \text{ and} \\ u + \beta v_{d-1} &\geq \beta v_d, \quad \forall d > -C. \end{aligned} \quad (3)$$

In addition, for all d such that $p_d = 0$, we must have

$$-e + \beta v_{d+1} \leq \beta v_d. \quad (4)$$

Discounting implies that consuming today is better than consuming at a later date. Hence, only the producer's incentive constraint is binding. In addition, note that \mathbf{p} affects the law of motion of the distribution of agents across states. We concentrate on IFAs for which $x_d > 0$ for some $d > 0$. It can be shown that, under suitable parameter restrictions, the set of IFAs where $p_d = 1$, for some d , is non-empty. Autarky is always incentive feasible. The next proposition asserts that all IFAs involving a positive volume of transactions have the property that agents increase their balances up to a point. If their balance becomes sufficiently high, agents will decline opportunities to increase it further. This results in some welfare loss, since production does not occur in some meetings when it would otherwise be desirable.

Proposition 1 *Assume β is sufficiently high. All stationary IFAs have the property that either $p_d = 0$ for all d , or there exists $\bar{D} \geq 0$ such that (a) $p_d = 1, \forall d \leq \bar{D}$, and (b) $p_d = 0, \forall d > \bar{D}$.*

Proof. Clearly, if β becomes arbitrarily small it is not possible to implement production even in the presence of caps. For part (a) we assume, without loss of generality, that $C = 0$. Let p^* denote a stationary allocation and x^* denote the stationary distribution of agents across states. Then

$$\begin{aligned} v_{d+1} - v_d &= \gamma \mathbf{x}^* \mathbf{p}^* \beta (v_d - v_{d-1}) + (1 - \gamma \mathbf{x}^* \mathbf{p}^*) \beta (v_{d+1} - v_d) \\ &\quad + \gamma (1 - x_0^*) p_{d+1}^* [\beta (v_{d+2} - v_{d+1}) - e] \\ &\quad - \gamma (1 - x_0^*) p_d^* [\beta (v_{d+1} - v_d) - e]. \end{aligned} \quad (5)$$

The proof of part (a) then follows from the following Lemma.

Lemma 2 *Suppose that $d' > d$. There is no IFA with $d \geq -C$ such that $p_d^* = 0$ and $p_{d'}^* = 1$.*

Proof. We proceed by contradiction. Without loss of generality, let $d' = d + 1$. As $p_d^* = 0$ and $p_{d+1}^* = 1$, we can use the expression for $v_{d+2} - v_{d+1}$ to get

$$\begin{aligned} \beta (v_{d+1} - v_d) &= \beta (v_{d+2} - v_{d+1}) + \frac{1 - \beta}{\gamma \mathbf{x}^* \mathbf{p}^*} (v_{d+2} - v_{d+1}) \\ &\quad + \gamma (1 - x_0^*) \beta [(v_{d+2} - v_{d+1}) - (v_{d+3} - v_{d+2})]. \end{aligned} \quad (6)$$

Since $\beta (v_{d+2} - v_{d+1}) + \frac{1 - \beta}{\gamma \mathbf{x}^* \mathbf{p}^*} (v_{d+2} - v_{d+1}) > e$ and $p_d^* = 0$, it must be that $(v_{d+3} - v_{d+2}) > (v_{d+2} - v_{d+1})$. For p^* to be incentive feasible it must then be that $p_{d+2}^* = 1$. Re-writing the expression for $v_{d+3} - v_{d+2}$, we have that $v_{d+4} - v_{d+3} > v_{d+3} - v_{d+2}$. Proceeding by induction we obtain that $v_{d+n+1} - v_{d+n} > v_{d+n} - v_{d+n-1}$, for all $n > 0$, and $p_{d+n+1}^* = 1$, for all n . Hence, for some n large enough, we must have $\beta (v_{d+n+1} - v_{d+n}) > u$, which contradicts incentive feasibility. Turning to part (b), it can be shown that, for β sufficiently large, in all IFAs the expression $v_{d+1} - v_d$ is strictly decreasing in d . It is easy to see that, given that $v_{d+1} - v_d$ is monotonically decreasing, we have that $v_{d+1} - v_d < \beta (v_d - v_{d-1})$. This, in turn, implies that $v_{d+1} - v_d < \beta^d (v_1 - v_0) \leq \beta^{d-1} u$, where the last inequality follows from the fact that $\beta (v_1 - v_0) < u$. Hence, there is \bar{D} sufficiently large, so that $\beta (v_{\bar{D}+1} - v_{\bar{D}}) < e$ and $p_d^* = 0$ for all $d \geq \bar{D}$. ■ ■

The previous discussion indicates that in order for a positive volume of transactions to take place it must be that the cap is set sufficiently low. We now turn to the question of existence of the stationary distribution, \mathbf{x}^* , of agents across states. For any given distribution we have shown that there exists \bar{D} such that $p_d^* = 1$, for all $d < \bar{D}$, and $p_d^* = 0$, otherwise. Using the normalization $C = 0$ we have the following.

Proposition 3 *If β is sufficiently high, there exists a stationary IFA in which $p_d = 1$, for some d . The IFA gives rise to a uniform stationary distribution, \mathbf{x}^* .⁶*

Proof. Let \bar{D} be such that $p_d^* = 1$, for all $d < \bar{D}$. In a stationary distribution, \bar{D} does not change. The law of motion for \mathbf{x} is then characterized by

$$\begin{aligned} x'_0 &= x_0(1 - \gamma(1 - x_0)) + x_1\gamma(1 - x_{\bar{D}}) \\ x'_d &= x_{d-1}\gamma(1 - x_k) + x_d(1 - \gamma(2 - x_0 - x_{\bar{D}})), \\ &\quad + x_{d+1}\gamma(1 - x_0), \quad \forall 1 < d < \bar{D}, \\ x'_{\bar{D}} &= x_{\bar{D}-1}\gamma(1 - x_0) + x_{\bar{D}}[1 - \gamma(1 - x_{\bar{D}})]. \end{aligned} \quad (7)$$

In a stationary distribution we have $x'_d = x_d$ for all d . Suppose $x_0 = x_{\bar{D}}$. The law of motion for x implies that $x_0 = x_1$ and $2x_d = x_{d+1} + x_{d-1}$. Hence, $x_d = x_{d'}$ for all $0 < d, d' \leq \bar{D}$ is a solution to these equations. In other words, the uniform distribution is stationary. ■

The following proposition suggests that the clearinghouse should set the cap at the maximum level that is consistent with the existence of an equilibrium with trade.

Proposition 4 *Consider two IFAs that are supported by a uniform distribution of reserves and respective caps C and C' , with $C > C'$. Welfare is higher in the allocation resulting from the greater cap, C .*

Proof. Given the agents' policy rules, and given that the distribution of money holdings is uniform, each agent will set $p_d^* = 1$ for $d < C(C')$, and $p_d^* = 0$, otherwise. In a cap y -allocation, under a uniform distribution, the probability of either consuming or producing is $y/(y + 1)$. Welfare is given

⁶We have not established that the stationary distribution is either unique or stable. Standard results on Markov chains cannot be directly applied as the transition matrix is state dependent.

by $\gamma(y/(y+1))^2$, which is clearly increasing in y . Hence, welfare is higher in the C -allocation. ■

We end this section by briefly discussing the case in which the CH can condition its policy on reports by *both* agents involved in any given transaction. More precisely, suppose that the clearinghouse can observe that i and j are in a match during a given period and consider the following rule. If both agents report that one produced for the other, say i for j , then the CH assigns a “+1” and a “−1” to the producer and the consumer, respectively. If the consumer reports that he did *not* receive production, the producer is punished to permanent autarky, which gives a lifetime utility of 0. Assume that β is sufficiently high. Then, under the above policy, the allocation in which production takes place in each trade meeting is incentive feasible. Thus, the first best allocation can be supported and there is no need for caps. Note, however, that this scheme relies on rather strong informational requirements. For example, it is necessary that the CH can verify that i and j are in a meeting in which i is the potential producer. If this assumption is withdrawn then i can claim (falsely) that it was j that did not produce for him, etc. In what follows, we shall restrict ourselves to environments in which the CH does not possess such information and in which certain transactions are subject to a form of imperfect monitoring.⁷

2.2 Local Clearinghouses

One case that is of particular interest is when agents transact through *local networks* which have specific information about their members. To this end, we now assume that agents are equally divided into two networks, I and II . The distribution of agents across types is symmetric across the two networks. We assume that an agent needs to transact within his network with probability α and outside the network with probability $1 - \alpha$. In order to isolate the effects of local information within a network from the effects related to the frequency of transactions, we may assume that any two agents are matched with equal probability regardless of their respective networks, i.e., $\alpha = 1/2$. In that case, the only asymmetry between the two networks

⁷Even in such cases, the CH can accomplish more by reverting to collective punishments. For example, the CH could punish *both* i and j , say, to permanent exclusion, unless their reports were mutually consistent. We shall ignore collective punishments in what follows.

concerns the flow of information of the agents' histories of interactions with the payment system.

We first consider two benchmark cases regarding the information structure. In one case, agents in each network are connected to a Local Clearinghouse (LCH). In order to capture the fact that LCHs possess more detailed information about their members we assume that they can monitor the types and actions of all agents in every meeting in which *both* agents belong to the network. There is no record-keeping regarding meetings across networks. In this environment, a standard reputation argument establishes that a “local credit” equilibrium can be sustained within each network. On the other hand, no inter-network trade is possible. The following proposition characterizes IFAs for this case. It can easily be demonstrated using arguments from the previous subsection.

Proposition 5 *If β is sufficiently high, the allocation where production occurs in each trade meeting within a network is incentive feasible under the threat of permanent exclusion. In all IFAs there are no transactions between agents belonging to different networks.*

The second benchmark concerns the case where such LCHs are absent. Instead, like before, we assume that each agent records his balance in each period with a *Central Clearinghouse (CH)*. The CH does not have detailed information about its members. We capture this by assuming that it cannot verify production opportunities. However, if agent i produces for j then i 's account is credited and j 's account is debited. Thus, agents record-keep their transactions, independent of the network of their trading partner, by building a balance with the CH. The second benchmark is identical to that of the previous section. Therefore, we have the following.

Proposition 6 *The IFA are characterized by the same conditions as in the previous section. The optimal cap policy is also the same.*

Now consider the case in which an agent meets with another agent from the same network with probability $\alpha \in (0, 1)$. We have the following.

Proposition 7 *The first benchmark, when it exists, dominates the second if α is sufficiently high. Otherwise, the second benchmark dominates the first.*

The first assertion is true since the first arrangement involves no caps. Thus, a transaction occurs in each trade meeting within the same network. The second assertion is true since the first benchmark will involve a negligible volume of transactions. This is because meetings between agents belonging to different networks always result in autarky. It is clear that, given that consumption outside an agent's network is not verifiable, no IFAs exist in which the first best transactions occur within a network and, at the same time, some transactions occur between networks. Thus, the full information first best allocation cannot be supported in this environment. We now turn to the main part of the paper, in which we study the role of settlement in supporting the first best in the presence of private information.

3 Optimal Payments under Periodic Settlement

In this section we study optimal transactions under private information, no commitment, and imperfect monitoring. We retain the assumption that the population of agents is divided into two symmetric networks, I and II , and that agents transact within the same network with probability α . Agents in each network are connected to a local clearinghouse (*LCH*). We shall think of LCHs as tools of a central planner under the constraint that there is no communication between clearinghouses.⁸ We want to capture the feature that, in some occasions, agents' transactions are observable to the LCH, while in other occasions an agent's ability to perform a transaction is private information and incentives will be needed in order to induce truthful revelation. To this end, we assume that a LCH can observe types and actions only in meetings in which *both* agents belong to the network, while there is no monitoring or record-keeping regarding agents belonging to different networks.

In the last section we established that, if β is sufficiently high, an equilibrium can be sustained in which a transaction takes place in each trade meeting within a network, under the threat that deviating agents are punished to exclusion. While this equilibrium supports efficient exchange within each network, it implies no transactions across networks, with the resulting efficiency loss. An obvious question is whether the resulting allocation can

⁸This is consistent with viewing each LCH as acting on behalf of the grand coalition of agents prior to each agent learning which network he is assigned to.

be improved upon. The difficulty lies in that LCHs cannot verify whether a trade meeting has taken place in meetings across networks. To see this, consider a distinguished agent who, say, for the k -th time in a row reports to his LCH that he could not produce since he did not have such a trade meeting. Given the information structure, the LCH can verify that the agent had k consecutive meetings outside the network (this is an event of probability $(1 - \alpha)^k$). It can also verify that the agent did not produce in any of these meetings. What the LCH cannot verify, however, is whether the agent had an opportunity to produce and simply declined or whether he did not have any trade meetings (an event of probability $(1 - \gamma)^k$).⁹

In this section we will assume that production of goods is perfectly divisible. Producing q units implies disutility $-e(q)$, while consumption of q units gives utility $u(q)$. We assume that $e'(q) > 0$, $e''(q) \geq 0$, $\lim_{q \rightarrow 0} e'(q) = 0$, and $\lim_{q \rightarrow \infty} e'(q) = \infty$. In addition, we assume $u'(q) > 0$, $u''(q) \leq 0$, $\lim_{q \rightarrow 0} u'(q) = \infty$, and $\lim_{q \rightarrow \infty} u'(q) = 0$. We will restrict ourselves to mechanisms in which the state of an agent consists of an one-dimensional real variable. We shall refer to such mechanisms as *payment mechanisms*. More precisely, our implementation assumes that each agent has a balance, $d \in \mathbb{R}$, with his LCH. For simplicity, we assume that agents' balances are perfectly observable. LCHs adjust balances overtime as a result of agents' reports about their transactions.

A number of actual transactions involve periodic *clearing* or *settlement* rounds. To study the effect of settlement in our model, we introduce a periodic pattern of length $n + 1$. The first n periods of each cycle involve, as previously, bilateral transactions. This is followed by one centralized settlement round, which we model as a Walrasian market. In that market agents from both networks can trade balances for a general, non-storable good. We assume that employing one unit of effort, l , leads to production of one unit of the general good, y , while the payoff from consuming y units of the general good is $U(y)$. Effectively, the Walrasian market corresponds to a settlement round in which agents that are "low" can increase their balances by producing, while those having excessive balances end up as consumers. We assume that $U'(y) > 0$, $U''(y) \leq 0$, and $\exists y^* \in (0, \infty)$, with $U(y^*) > y^*$, such that $U'(y^*) = 1$. The price, p , at which balances are purchased is determined by market clearing conditions.

⁹Recall that since meetings between different networks are anonymous, the distinguished agent cannot be identified outside his network.

In each of the first n periods of the cycle an agent may transact or engage in no transactions within his network. In that case, the LCH monitors the transaction perfectly, observing both the type of the meeting and the quantity consumed or produced. Like before, it is only when an agent transacts outside his network that the LCH cannot observe whether the agent is in a trade meeting or not. If an agent ends up producing outside his network, production can be verified. On the other hand, consumption outside the network is not verifiable. Throughout, we restrict attention to outcomes that are stationary and symmetric across agents and across networks. It is useful to keep in mind the full information first best allocation in this setup. We have the following.

Proposition 8 *The first best allocation involves consumption and production of q^* in each bilateral trade meeting both within and outside a network. In the first best each agent consumes y^* in the settlement round.*

Note that, while aggregate production in the settlement round equals y^* , the optimal production allocation is indeterminate due to the linearity in production costs. Clearly, if α , the probability of transacting within the network, is equal to 1, the first best allocation can be implemented provided that β is sufficiently high. In what follows, we will concentrate on the case where $\alpha \in [0, 1)$. Then, under the maintained assumption that β is sufficiently high, the optimal allocation involves some transactions between networks. The choice of n , which refers to the frequency of the settlement rounds, is interesting for policy purposes. However, for now, we will simply impose that $n = 1$ and proceed to study the LCHs' problem for the above setup.

We start by writing down the agent's problem in the settlement round. Let $V(d, p)$ denote the value function of an agent that enters with a balance of d . Let $v(\hat{d}, \Psi)$ denote the value when he exits the market with \hat{d} , given that the resulting distribution of balances is given by Ψ . Agents solve the following problem in the settlement round.

$$\begin{aligned} V(d, p) &= \max_{y, l, \hat{d}} \{U(y) - l + \beta E v(\hat{d}, \Psi)\} \\ \text{s.t. } p\hat{d} &= pd + l - y. \end{aligned} \tag{8}$$

Agents take as given the price, p , and the distribution of balances at the end of the market, Ψ . Below we argue that, as a consequence of quasi-linearity, this distribution is degenerate. That is, all agents exit with the

same balance. This feature greatly improves the tractability of the problem. In addition, it fits well with how actual payment systems operate, requiring centralized settlement at the end of a specified period, usually a day.

Next, we describe the problem faced by each LCH during bilateral transactions. Recall that the LCHs are considered to be tools of a central planner who values the (ex ante) welfare of all agents equally. However, the LCHs face monitoring frictions. They are able to observe certain characteristics of transactions and they are not able to communicate with each other.¹⁰ More precisely, recall that if an agent has an opportunity to produce in a trade meeting outside his network, he can (falsely) report to his LCH that he was not in a trade meeting. Like before, if production takes place, the LCH can verify the quantity produced in transactions that involve at least one agent belonging to the LCH. Consumption outside the network is, however, not verifiable. Each LCH maximizes a welfare function that involves all matches weighted by the frequency as implied by Ψ .

Let $\mathbf{d} \in \mathbb{R}^2$ denote the vector of balances of two agents that are involved in a bilateral transaction during a given period. For simplicity, we assume that balances are observable both within and outside a network. We shall think of agents as making reports to their respective LCH about the type of the meeting that they are in. Agents that report a trade meeting as producers receive instructions from their LCH about how much to produce. Consumers report to their LCH the quantity they consumed. The LCHs subsequently make balance adjustments that depend on these reports. Like before, production and delivery of goods outside the network is verifiable while consumption is not. In addition, agents that refuse to produce within their network are punished to permanent exclusion.

Note that, ignoring the agents' balances, there are six possibilities regarding meetings. An agent can be in a consumption, a production, or a non-trade meeting either within or outside his network. The vector of policy rules (P_t, R_t, A_t, q_t) determines the new balances as well as the quantity produced, q_t , for agents involved in transactions within the network. These functions depend on the agents' current balances and on the distribution of balances, Ψ . More precisely, $P_t(R_t)$ gives the balance adjustment for an agent who consumes(produces) within his network, while A_t is the adjustment resulting from not trading within the network. Similarly, the vector

¹⁰This lack of communication is an extreme way of capturing the fact that an agent's state might be less transparent outside its country or region of origin.

of policy rules (L_t, K_t, B_t, Q_t) determines the new balances and the quantity produced, Q_t , for agents involved in transactions outside their network. Like before, $L_t(K_t)$ is the adjustment for an agent who consumes(produces) outside his network, while B_t is the adjustment for an agent who does not transact. Recall that balances are represented by real numbers not restricted in sign, while production of goods is restricted to be positive. At the end of the decentralized stage balances are adjusted according to reports and agents enter the settlement round knowing their new balances.

We will concentrate on arrangements that satisfy certain incentive and participation constraints. The incentive constraints require that the following inequalities hold.

$$u(Q) + V(d + L, p) \geq V(d + B, p), \quad (9)$$

$$-e(Q) + V(d + K, p) \geq V(d + B, p). \quad (10)$$

Finally, the participation constraints require that

$$V(d + A, p) \geq 0, \quad (11)$$

$$V(d + B, p) \geq 0, \quad (12)$$

$$-e(q) + V(d + R, p) \geq 0, \quad (13)$$

$$u(q) + V(d + P, p) \geq 0. \quad (14)$$

We are now ready to formally define a payment system.

Definition 9 *A Payment System \mathbf{S} is defined to be an array of functions $\mathbf{S} = \{P_t, R_t, A_t, q_t; L_t, K_t, B_t, Q_t\}$. \mathbf{S} is incentive feasible if the array is chosen so as to satisfy the incentive and participation constraints. \mathbf{S} is simple if the balance adjustments do not depend on the level of the agents' current balances. An incentive feasible \mathbf{S} is optimal if it can support a first best allocation.¹¹*

Let $v(\mathbf{d})$ be the expected value of an agent with balance d in a meeting with an agent with balance \tilde{d} . This is given by

¹¹To simplify notation we shall often suppress the dependence of \mathbf{S} on t .

$$\begin{aligned}
v(\mathbf{d}) &= \alpha\{\gamma[u(q) + V(d + P, p)] + \gamma[-e(q) + V(d + R, p)] \\
&\quad + (1 - 2\gamma)V(d + A, p)\} \\
&\quad + (1 - \alpha)\{\gamma[u(\tilde{Q}) + V(d + L, p)] + \gamma[-e(Q) + V(d + K, p)] \\
&\quad + (1 - 2\gamma)V(d + B, p)\}. \tag{15}
\end{aligned}$$

Recall that the first best allocation involves production of q^* in *all* bilateral trade meetings. The next two propositions establish the existence of an optimal payment system provided that the Friedman rule results in the settlement stage.

Proposition 10 *Assume that the Friedman rule, $p = \beta p_{+1}$, holds. A simple incentive feasible \mathbf{S} can support the first best allocation in the decentralized stage.*

Proof. Let y^* be the solution to the equation $U'(y^*) = 1$. Using the definition of v and symmetry across networks (i.e. $\tilde{Q} = Q$), we can rewrite the agent's problem in the Walrasian market as

$$\begin{aligned}
V(d, p) &= U(y^*) - y^* + pd + \max_{\hat{d} \leq \hat{d} \leq \bar{d}} -p\hat{d} + \beta\{V(\hat{d}, p_{+1}) \\
&\quad + \alpha \int_{d'} [\gamma[u(q(\hat{d}, d')) - e(q(\hat{d}, d')) + p_{+1}(P(\hat{d}, d') + R(\hat{d}, d'))] \\
&\quad + (1 - 2\gamma)p_{+1}A(\hat{d})]d\Psi + \\
&\quad (1 - \alpha) \int_{d'} [\gamma[u(Q(\hat{d}, d')) - e(Q(\hat{d}, d')) + p_{+1}(L(\hat{d}, d') + K(\hat{d}, d'))] \\
&\quad + (1 - 2\gamma)p_{+1}B(\hat{d}, d')]d\Psi\}, \tag{16}
\end{aligned}$$

where p_{+1} denotes next period's price. Linearity of $V(d, p)$ in d implies that

$V(\widehat{d}, p_{+1}) = V(0, p_{+1}) + p_{+1}\widehat{d}$. Hence, we obtain the following expression.

$$\begin{aligned}
V(d, p) &= U(y^*) - y^* + pd + \beta V(0, p_{+1}) + \max_{\underline{d} \leq \widehat{d} \leq \bar{d}} -p\widehat{d} + \beta p_{+1}\widehat{d} \\
&\quad + \beta \left\{ \alpha \int_{d'} [\gamma [u(q(\widehat{d}, d')) - e(q(\widehat{d}, d'))] + p_{+1}(P(\widehat{d}, d') + R(\widehat{d}, d'))] \right. \\
&\quad \left. + (1 - 2\gamma)p_{+1}A(\widehat{d}, d') \right] d\Psi + \\
&\quad (1 - \alpha) \int_{d'} [\gamma [u(Q(\widehat{d}, d')) - e(Q(\widehat{d}, d'))] + p_{+1}(L(\widehat{d}, d') + K(\widehat{d}, d'))] \\
&\quad \left. + (1 - 2\gamma)p_{+1}B(\widehat{d}, d') \right] d\Psi \}. \tag{17}
\end{aligned}$$

The first order condition with respect to \widehat{d} gives

$$\begin{aligned}
&-p + \beta p_{+1} + \beta \left\{ \alpha \int_{d'} [\gamma [u'(q(\widehat{d}, d')) - e'(q(\widehat{d}, d'))]] \frac{\partial q}{\partial \widehat{d}} \right. \\
&\quad \left. + \gamma p_{+1} \left[\frac{\partial (P(\widehat{d}, d') + R(\widehat{d}, d'))}{\partial \widehat{d}} \right] + (1 - 2\gamma)p_{+1} \frac{\partial A(\widehat{d}, d')}{\partial \widehat{d}} \right] d\Psi \\
&\quad (1 - \alpha) \int_{d'} [\gamma [u'(Q(\widehat{d}, d')) - e'(Q(\widehat{d}, d'))]] \frac{\partial Q}{\partial \widehat{d}} + p_{+1} \left[\frac{\partial (L(\widehat{d}, d') + K(\widehat{d}, d'))}{\partial \widehat{d}} \right] \\
&\quad \left. + (1 - 2\gamma)p_{+1} \frac{\partial B(\widehat{d}, d')}{\partial \widehat{d}} \right] d\Psi \} = 0. \tag{18}
\end{aligned}$$

Assuming that the Friedman rule holds, and considering a simple \mathbf{S} , the first order condition reduces to

$$\begin{aligned}
&\alpha \int_{d'} \gamma [u'(q(\widehat{d}, d')) - e'(q(\widehat{d}, d'))]] \frac{\partial q}{\partial \widehat{d}} d\Psi \\
&\quad + (1 - \alpha) \int_{d'} [\gamma [u'(Q(\widehat{d}, d')) - e'(Q(\widehat{d}, d'))]] \frac{\partial Q}{\partial \widehat{d}} d\Psi = 0 \tag{19}
\end{aligned}$$

Clearly, any \widehat{d} that satisfies $u'(q(\widehat{d}, d')) = e'(q(\widehat{d}, d'))$ and $u'(Q(\widehat{d}, d')) = e'(Q(\widehat{d}, d'))$ also satisfies the above first order condition. This completes the proof. ■

The following proposition gives a sufficient condition for a simple incentive feasible payment system that supports an optimal allocation to exist.

Proposition 11 *Provided that $(\beta - 1)d_0 > (1 - \alpha)\gamma(e(q^*) - u(q^*))$, there exists a simple optimal \mathbf{S} that is consistent with the Friedman rule and under which agents start each new cycle with the same balances.*

Proof. The proof proceeds by constructing an optimal payment system satisfying the degenerate distribution property. First, we demonstrate that the agents' balance decisions in the settlement stage do not depend on their previous balance. In order for the agents' problem (8) to be well defined, we assume that $d \in [\underline{d}, \bar{d}]$, for some finite bound $\underline{d} < \bar{d}$. It should be clear that any stationary IF allocation with $q > 0$ and $Q > 0$ necessarily has to feature a payment system with $d \leq \bar{d}$, thus, the upper bound never binds. The lower bound requirement is akin to the familiar no Ponzi scheme condition. Later, we argue that the lower bound will also not bind for a particular choice of $[\underline{d}, \bar{d}]$.

Let y^* be the solution to the equation $U'(y^*) = 1$. Using the definition of v , the linearity of V , and symmetry across networks (i.e. $\tilde{Q} = Q$), we can rewrite the agent's problem in the settlement stage as

$$\begin{aligned} V(d, p) = & U(y^*) - y^* + pd + \\ & \max_{\underline{\hat{d}} \leq \hat{d} \leq \bar{\hat{d}}} \beta \{ V(\hat{d}, p) + \alpha [\gamma [u(q) - e(q) + p_{+1}(P + R)] + (1 - 2\gamma)p_{+1}A] + \\ & (1 - \alpha) [\gamma [u(Q) - e(Q) + p_{+1}(L + K) + (1 - 2\gamma)p_{+1}B]] \} - p\hat{d}. \end{aligned} \quad (20)$$

The decision variable \hat{d} is then independent of d due to quasi-linearity. In order to demonstrate that agents choose the same balances in the settlement stage we first construct functions $q(\hat{d}, d')$ and $Q(\hat{d}, d')$ and the corresponding payment system. Let $q(d, d')$ and $Q(d, d')$ be such that $pd = e(q(d, d'))$ and $p\hat{d} = e(Q(\hat{d}, d'))$ for all $d < \bar{d}$ and $q(d, d') = q^*$, $Q(\hat{d}, d') = Q^*$ for $d \geq \bar{d}$. Following Lagos and Wright (2005), we can ensure that these lead to a unique choice of balance, \bar{d} , while implementing q^* and Q^* .

Next, we construct a simple payment system that is consistent with the Friedman rule and show that it can support the first best allocation in the decentralized stage. Stationarity implies $p_t \bar{d}_t = p_{t+1} \bar{d}_{t+1}$. This together with

the Friedman rule implies that for all t , $\beta\bar{d}_t = \bar{d}_{t+1}$. The law of motion of d_{t+1} implies

$$d_{t+1} = d_t + \alpha[\gamma(P_t + R_t) + (1 - 2\gamma)A_t] + (1 - \alpha)[\gamma(L_t + K_t) + (1 - 2\gamma)B_t]. \quad (21)$$

which, using that $\beta\bar{d}_t = \bar{d}_{t+1}$ can be written as

$$0 > (\beta - 1)d_t = \alpha[\gamma(P_t + R_t) + (1 - 2\gamma)A_t] + (1 - \alpha)[\gamma(L_t + K_t) + (1 - 2\gamma)B_t]. \quad (22)$$

In addition, the incentive and participation constraints must hold. By stationarity we have

$$(\beta - 1)p_t d_t = (\beta - 1)p_{t-1} d_{t-1} = \dots = (\beta - 1)p_0 d_0. \quad (23)$$

Setting $B_t = L_t = -\frac{\gamma}{1-\gamma}u(q^*)/p_t$ and $K_t = e(q^*)/p_t$, for all t , satisfies the incentive constraints. In addition, provided that $(\beta - 1)p_0 d_0 > (1 - \alpha)\gamma(e(q^*) - u(q^*))$ and that $U(y^*) - y^*$ is sufficiently high, we can find $R_0 = P_0 = A_0 > 0$ that preserve the participation constraints and such that the following equality holds at $t = 0$.

$$(\beta - 1)p_0 d_0 = (1 - \alpha)\gamma(e(q^*) - u(q^*)) + \alpha p_0[\gamma(P_0 + R_0) + (1 - 2\gamma)A_0]. \quad (24)$$

Next, we construct \mathbf{S}_t recursively. first note that the above equality pins down d_1 . We can then find $R_1 = P_1 = A_1 > 0$ such that a similar equality holds at $t = 1$. This can be repeated for all $t > 1$. Thus, a sufficient condition for the existence of the desired \mathbf{S} is that

$$p_0 d_0 < \frac{(1 - \alpha)\gamma(u(q^*) - e(q^*))}{(1 - \beta)}. \quad (25)$$

To see that the constructed payment system is optimal, note that the expected balance of an agent coming from the decentralized stage in period t is $d_t + \alpha[\gamma(P_t + R_t) + (1 - 2\gamma)A_t] + (1 - \alpha)[\gamma(L_t + K_t) + (1 - 2\gamma)B_t]$. By the law of large numbers, this equals d_{t+1} . Therefore, the expected utility of an agent in the settlement stage at any date t is given by

$$\begin{aligned} & \alpha \quad \{ \gamma[U(y^*) - y^* - p_t(d_{t+1} - d_t - P_t)] \\ & \quad + \gamma[U(y^*) - y^* - p_t(d_{t+1} - d_t - R_t)] \\ & \quad + (1 - 2\gamma)[U(y^*) - y^* - p_t(d_{t+1} - d_t - A_t)] \} \\ + (1 - \alpha) & \quad \{ \gamma[U(y^*) - y^* - p_t(d_{t+1} - d_t - L_t)] \\ & \quad + \gamma[U(y^*) - y^* - p_t(d_{t+1} - d_t - K_t)] \\ & \quad + (1 - 2\gamma)[U(y^*) - y^* - p_t(d_{t+1} - d_t - B_t)] \}, \quad (26) \end{aligned}$$

which can be simplified to give $U(y^*) - y^*$. In the decentralized stage, the payment system implies that q^* and Q^* are produced in all bilateral trade meetings. Hence, the expected welfare under the above payment system is the same as the ex-ante welfare of the full information first best allocation. This completes the proof. ■

Using the incentive and participation constraints, we can derive further properties of \mathbf{S} . First, note that, since consumption outside the network is not verifiable, in order for an agent that consumed to report truthfully, \mathbf{S} needs to treat him the same way as if he reported a no trade meeting, i.e., $B_t = L_t$. In addition, it must be that $p_t(K_t - B_t) \geq e(Q^*)$, $\forall t$. In other words, agents are rewarded for producing outside their network. While the proposed payment system implies that all agents consume y^* in each settlement round, it will clearly not result in each agent producing y^* . To satisfy the incentive constraints, those that consume outside their network must end up producing a higher amount in the settlement round than all other agents.

4 Comments

We introduced an intertemporal model of a payment system and used it in order to investigate the structure of optimal transactions under private information. Using recent developments in monetary economics, we were able to generate a tractable environment, even though the histories of transactions can be fairly complex. An extension that we are currently pursuing involves studying the case where $n > 1$. It is an open question whether the Friedman rule will remain optimal in this case, or whether positive inflation will be beneficial as in Levine (1991). A more general topic involves the study of optimal dynamic contracting in abstract private information environments in which there are periodic “full information rounds.”

Our model could be used to investigate several issues related to payments, we mention only two here. First, we could study the effects of the possibility of default by either a single participant or an entire network. Indeed, the existence of local networks implies that the resulting allocation has to be robust not only to deviations by individual participants, but also to coalitional deviations. In the presence of aggregate risk, an entire network might find it profitable to exit if its position within the system becomes sufficiently unfavorable. This, in turn, could create certain contagion effects. Avoiding this problem will require the study of stronger forms of implementation, say in

coalition proof perfect equilibria. Second, given that we deal with dynamic incentives, we could investigate the time consistency of various clearinghouse policies, a problem that the current analysis abstracts from. This relates to the debate of public versus private provision of payment system services since optimal dynamic schemes might require a high degree of commitment. Related issues will arise in our model once we endogenize the interactions between LCHs, which we currently model as passive tools of a central planner.

References

- [1] Abreu, D., Pearce, D., and E. Stacchetti (1986) "Optimal Cartel Equilibria with Imperfect Monitoring," *Journal of Economic Theory* 39, p. 251-269
- [2] Abreu, D., Pearce, D., and E. Stacchetti (1990) "Toward a Theory of Discounted Repeated Games with Imperfect Monitoring," *Econometrica* 58 (5), p. 1041-1063
- [3] Atkeson A. and R. Lucas Jr. (1993) "On Efficient Distribution with Private Information," *Review of Economic Studies*, vol. 59 (July)
- [4] Cavalcanti R., Erosa A., and T. Temzelides (2004) "Liquidity, Money Creation and Destruction, and the Returns to Banking," Manuscript for the Models of Monetary Economics II, the Next Generation, *International Economic Review*, forthcoming
- [5] Fujiki H., Green E.J., and A Yamazaki (1999) "Sharing the Risk of Settlement Failure," Federal Reserve Bank of Minneapolis Working Paper 594D
- [6] Green E.J. (1987) "Lending and the Smoothing of Uninsurable Income," in *Contractual Arrangements for Intertemporal Trade*, Edward C. Prescott and Neil Wallace, Eds. Minneapolis
- [7] Green, E.J., and R.H. Porter (1984) "Non-Cooperative Collusion Under Imperfect Price Information," *Econometrica* 52 (1), p. 87-100
- [8] Kahn C.M., and W. Roberds (1998) "Payment System Settlement and Bank Incentives," *The Review of Financial Studies* 11(4), p. 845-870
- [9] Kiyotaki N., and R. Wright (1989) "On Money as a Medium of Exchange," *Journal of Political Economy* 97, p. 927-954
- [10] Kiyotaki N., and R. Wright (1993) "A Search-Theoretic Approach to Monetary Economics," *American Economic Review* 83, p. 63-77
- [11] Kocherlakota N. (1998) "Money is Memory," *Journal of Economic Theory* 81(2), p. 232-251

- [12] Kocherlakota N. (2004) "The New Dynamic Public Finance," SED Plenary Session
- [13] Lagos R. and R. Wright (2005) "A Unified Framework for Monetary Theory and Policy Analysis," *Journal of Political Economy*, forthcoming
- [14] Levine D. (1991) "Asset Trading Mechanisms and Expansionary Policy." *Journal of Economic Theory* 54
- [15] Mills D.C. (2003) "Alternative Central Bank Credit Policies for Liquidity Provision in a Model of Payments," Federal Reserve Board of Governors Working Paper
- [16] Mirrlees J. (1971) "An Exploration in the Theory of Optimal Income Taxation," *Review of Economic Studies* 38, p. 175-208
- [17] Myerson R.B. (1991) *Game Theory: Analysis of Conflict*, Cambridge, Mass. Harvard University Press
- [18] Spear S., and S. Srivastava (1987) "On Repeated Moral Hazard with Discounting," *Review of Economic Studies* 54(4)
- [19] Temzelides T. and S. Williamson (2001) "Payments Systems Design in Deterministic and Private Information Environments," *Journal of Economic Theory* 99, p. 297-326