Queen's Economics Department Working Paper No. 1053

A Dynamic Model of Settlement

Thorsten Koeppl
Department of Economics, Queen's University

Cyril Monnet
DG Research, European Central Bank

Ted Temzelides
Department of Economics, University of Pittsburgh

Department of Economics
Queen's University
94 University Avenue
Kingston, Ontario, Canada
K7L 3N6

2-2006

A Dynamic Model of Settlement*

Thorsten Koeppl
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

First draft: December 2005

March 17, 2006

*We thank participants at the Cornell University Macro Workshop as well as Narayana Kocherlakota and Ricardo Lagos for comments and suggestions on an earlier draft. We are grateful to the ECB working paper series for helpful comments regarding exposition. The views expressed in the paper are not necessarily those of the ECB. The third author gratefully acknowledges support from the NSF through grant SES 0517862.

Abstract

We investigate the role of settlement in a dynamic model of a payment system where the ability of participants to perform certain welfare-improving transactions is subject to random and unobservable shocks. In the absence of settlement, the full information first-best allocation cannot be supported due to incentive constraints. In contrast, this allocation is supportable if settlement is introduced. This, however, requires that settlement takes place with a sufficiently high frequency.

1 Introduction

A distinguishing feature of payment systems is *settlement*, or the discharge of past obligations through the transfer of an asset.[1] Actual settlement has three defining properties. First, it is not a welfare-improving activity by itself. Rather, settlement involves a mere transfer of an asset between participants in order to fulfill the obligations created by previous transactions. Second, it takes place periodically. For instance, credit card transactions are generally settled on a monthly basis, while settlement of interbank transactions generally takes place daily. Finally, settlement gives the opportunity to all participants in the system to start afresh since, after settling their obligations, they are no longer liable to the system.

In this paper we employ the dynamic model of a payment system developed in Koeppl, Monnet, and Temzelides (KMT, 2005) in order to study the role of settlement. Our main finding is that settlement is an essential part of an optimal payment system as it enables agents to engage in beneficial transactions that would otherwise not be realized. As in KMT, we employ a version of the model of exchange developed by Kiyotaki and Wright (1989, 1993). Our model, however, departs from monetary economics and, instead, emphasizes the role of private information in a way related to the dynamic contracting literature.[2] The use of a dynamic framework is essential since 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.[3]

Agents in our model need to engage in transactions that are subject to a private information friction. Thus, incentives are needed in order to induce truthful revelation. This implies that, in the absence of settlement, it is impossible to support the full information first-best allocation. We introduce a periodic pattern in which each transaction stage, consisting of a finite number of bilateral trades, is followed by a centralized stage in which a planner may reallocate a general good that is produced using a linear technology. We demonstrate that the full information first-best allocation is

[1]For references to different notions of settlement, we refer the reader to BIS (2003).

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

[3]See Kahn and Roberds (1998, 2004) for two papers in this literature. Aiyagari and Williamson (1999, 2000) and Temzelides and Williamson (2001) investigate some of the issues studied here. Their model and conclusions, however, are very different from ours.

attainable in this setup.

As in KMT, we decentralize this allocation via a payment system. This involves assigning individual balances and optimally adjusting these balances given the agents' histories of transactions. Agents can trade balances in a centralized settlement stage which we decentralize as a competitive market.[4] The settlement stage is essential since the first-best allocation cannot be supported in its absence.

Our model displays the following properties. First, the introduction of periodic settlement rounds does not increase welfare by itself. The increase in welfare is accomplished indirectly through the interplay between settlement and intertemporal incentives. Second, the first-best is supportable only if settlement takes place with a sufficiently high frequency. Finally, payment system participants must exit the settlement stage with identical balances. In this sense, agents start afresh as their history does not affect their future transactions.

The reason for settlement being essential in achieving the first-best can be summarized as follows. In the presence of private information, in order for any transactions to take place, the payment system must provide intertemporal incentives. This is costly and, in the absence of settlement, these costs can only be borne directly, creating a distortion during the bilateral transactions stage. Settlement allows accumulating and shifting these costs to the centralized stage. This is efficient since incentives in bilateral transactions can then be set properly. More precisely, under a linearity assumption, balance adjustments in the settlement round do not create direct welfare gains or losses on average. In addition, periodic settlement limits the obligations an agent can accumulate over time. Hence, when settlement occurs frequently enough, and the net value of future transactions is high enough, agents will choose to participate in the system.

The paper proceeds as follows. In Section 2, we present the model and discuss optimal allocations. In Section 3, we study decentralization through a payment system. A brief conclusion follows.

[4] A novelty of this approach is that it involves non-cooperative implementation together with a Walrasian equilibrium aspect.

2 The Model

Time, t, is discrete and measured over the positive integers. There is a $[0, 1]$ continuum of infinitely lived agents. The common discount factor is $\beta \in (0, 1)$. To generate a role for transactions, we assume 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 or as consumer, each with probability γ. More precisely, in each period an agent is in a *trade meeting* with probability 2γ. In this case, he is a potential producer or a potential consumer with equal probability. With probability $(1 - 2\gamma)$ the agent is in a no-trade meeting.

We 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 $e'(q) > 0$, $e''(q) \geq 0$, $\lim_{q \to 0} e'(q) = 0$, and $\lim_{q \to \infty} e'(q) = \infty$. In addition, we assume that $u'(q) > 0$, $u''(q) \leq 0$, $\lim_{q \to 0} u'(q) = \infty$, and $\lim_{q \to \infty} u'(q) = 0$. Thus, there exists a unique q^* such that $u^{'}(q^*) = e^{'}(q^*)$. The quantity q^* uniquely maximizes the joint surplus created in a transaction.[5] Since we will concentrate on this quantity, in order to simplify notation, we will hence denote $u(q^*)$ by u, and $e(q^*)$ by e.

Our environment is subject to private information and commitment frictions. More precisely, we will assume that whether a meeting is a trade meeting; i.e., whether the consumer likes what the producer can offer, is not observable outside the meeting. In addition, agents cannot pre-commit to producing in such meetings. While the opportunity to produce cannot be verified, we assume that production itself, when it takes place, is verifiable. For simplicity, we assume that consumption is not verifiable.[6]

Our efficiency benchmark is the full information first-best allocation, in which the efficient level of production takes place in *all* trade meetings. The difficulty in supporting this allocation lies in the fact that, due to the private information friction, the planner cannot verify whether a trade meeting has taken place. For example, consider a distinguished agent who, say, for the k-

[5] Assuming that $u(\bullet)$ and $e(\bullet)$ are linear would not affect our results provided that the quantity q is restricted to belonging to a compact set, say $[0, \overline{q}]$. In that case, $q^* = \overline{q}$.

[6] Our results are strengthened if a positive fraction of the transactions are assumed to be fully monitored and, thus, not subject to a private information problem. Similarly, assuming that consumption is verifiable would change the structure of the optimal payment system but would not affect our main result regarding the efficiency of settlement. We refer to KMT for details.

th time in a row, reports that he did not produce since he had k consecutive no-trade meetings. Given the information structure, the planner can verify that the agent did not produce in any of the last k meetings. What the planner 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 $(1-\gamma)^k$).

A moment's reflection should convince the reader that the first-best allocation is not supportable. Indeed, if an agent always receives q^* independent of his history of reports, there is no incentive for him to ever produce. Thus, an agent must receive a quantity less than q^* at least on some occasions.

2.1 Periodic Centralized Rounds

We proceed by imposing a periodic pattern in which each *transaction stage*, consisting of a finite number of bilateral transaction rounds, is followed by a *centralized round*. In this round agents can consume or produce a general, non-storable good. We assume that the (dis)utility from this good is linear. More precisely, we assume a periodic pattern of length $n+1$. The first n periods of each cycle involve bilateral transactions. This is followed by one centralized round in which producing ℓ units of a general good implies disutility $-\ell$, while consuming ℓ units gives utility ℓ.

We can imagine that agents report their state (producer p, consumer c, neither n) in each round of the transactions stage to a planner who instructs them how much to produce (consume) in each bilateral transaction. In addition, the planner instructs agents how much to produce (consume) in the subsequent centralized round. The planner's recommendations will, in general, depend on the agents' past history of reports. An allocation, therefore, specifies the quantity produced (consumed) in each transaction as well as in the centralized round. An allocation is *supportable* if it respects incentive, ex ante participation and resource feasibility constraints. Throughout, we restrict attention to outcomes that are stationary and symmetric across agents. We want to characterize the best allocation that the planner can support. We term this allocation (ex ante) efficient. It turns out that this corresponds to the full information first-best allocation, in which q^* is produced in each trade meeting. We only consider here the case where $n=1$. That is, we impose a periodic pattern in which each bilateral transactions round is followed by centralized settlement. Later on we discuss how to decentralize ex-ante efficient allocations for the general case where $n>1$. We have the following.

Proposition 1 *Suppose that $n = 1$. The allocation where consumption of q^* occurs in all bilateral trade meetings and where $E[\ell^*] = 0$ is supportable if and only if $\beta u \geq e$.*

In other words, provided that β is sufficiently high, the full information first-best is incentive feasible. This is easy to demonstrate by showing that the full information first-best can be supported by repeating a "static" allocation.

Notice that, following the decentralized stage, there are two possible relevant histories. Either the agent was a producer, in which case we assume that the planner instructs him to produce ℓ_p during the centralized stage, or the agent was a consumer, in which case the planner instructs him to produce ℓ_c. Since consumption is not verifiable, agents who were in a no-trade meeting will also be instructed to produce ℓ_c; i.e., $\ell_n = \ell_c$.

Consider a planner who maximizes ex ante expected utility subject to resource, participation, and incentive constraints. The planner's problem at the beginning of a transactions round is

$$\max_{q,\ell_p,\ell_c,\ell_n} \frac{1}{1-\beta}[\gamma(u(q) - e(q)) - \gamma\ell_p - \gamma\ell_c - (1-2\gamma)\ell_n] \tag{1}$$

subject to

$$\gamma\ell_p + \gamma\ell_c + (1-2\gamma)\ell_n = 0 \tag{2}$$

$$-e(q) - \ell_p + \frac{\beta}{1-\beta}[\gamma(u(q) - e(q)) - \gamma\ell_p - \gamma\ell_c - (1-2\gamma)\ell_n] \geq 0 \tag{3}$$

$$u(q) - \ell_c + \frac{\beta}{1-\beta}[\gamma(u(q) - e(q)) - \gamma\ell_p - \gamma\ell_c - (1-2\gamma)\ell_n] \geq 0 \tag{4}$$

$$-\ell_n + \frac{\beta}{1-\beta}[\gamma(u(q) - e(q)) - \gamma\ell_p - \gamma\ell_c - (1-2\gamma)\ell_n] \geq 0 \tag{5}$$

$$-e(q) - \ell_p \geq -\ell_n. \tag{6}$$

The first constraint captures aggregate feasibility. The second, third, and fourth constraints are participation constraints for a producer, a consumer, and a no-trader, respectively. Finally, since the planner cannot observe whether or not an agent has an opportunity to produce during the transaction round, the last constraint is an incentive compatibility constraint. It is easy to verify that $q = q^*$, $\ell_p = -(1-\gamma)e$, and $\ell_c = \ell_n = \gamma e$ solve the above problem provided that $\beta u \geq e$. In addition, the resulting allocation satisfies the ex-ante participation constraint that $\frac{1}{1-\beta}[\gamma(u(q) - e(q)) - \gamma\ell_p - (1-\gamma)\ell_c] \geq 0$.

A few remarks are in order. First, note that the average production of the general good equals zero, which implies that, as mentioned earlier, there is no ex ante direct gain from introducing the centralized round. Second, $\beta u \geq e$ is necessary in order to support a positive volume of transactions even under full information. Thus, it is not possible to improve over the conditions of the above Proposition, in the sense of supporting the first-best allocation for a wider range of parameter values.

To summarize, when settlement is introduced, and provided that it involves costs that enter the agents' expected utility in a linear fashion, the full information first-best allocation is supportable in our model even though the planner is subject to a private information friction. Next we demonstrate how the efficient allocation for this environment can be supported via a payment system. This involves assigning balances to individual participants and specifying rules for how these balances are updated in order to satisfy incentive and participation constraints. In addition, we will decentralize the settlement stage by assuming that it operates as a competitive market.

3 Payment Systems

In the previous analysis we assumed that, subject to a participation constraint, the planner has the ability to re-allocate the output produced in the centralized stage. Here we will discuss how the first-best allocation can be decentralized. As in KMT, we assume that agents face a payment system (PS) which assigns *balances* to participants. The PS specifies rules for how the balances are updated given the histories of reports regarding bilateral transactions. As before, settlement implies that participants can periodically trade balances for the general, non-storable good. Here the settlement stage is modelled as a competitive market in which agents that are "low" can increase their balances by producing, while those with high balances end up as consumers. The price, p, at which balances are traded, is determined by market clearing conditions.

In each of the first n periods of the cycle, agents engage in bilateral transactions. Recall that whether an agent is a potential producer in a trade meeting is not observable to the PS. We begin by assuming that $n = 1$ and proceed to study the optimal PS given the above setup. In what follows, we analyze a generic period, t, and work backwards, first considering the agent's problem in the settlement stage, and then moving on to the transactions stage.

Let $V(d,p)$ denote the value function of an agent that exits the transaction round with balance d, given that the anticipated price in the following settlement round is p. Let $v(\widehat{d},\Psi)$ denote the value of an agent who exits the settlement round with balance $\widehat{d}$, given that the resulting distribution of balances is denoted by Ψ.[7] Given p and Ψ, agents at the beginning of the settlement round solve the following:

$$\begin{aligned} V(d,p) &= \max_{\ell,\widehat{d}}\{-\ell+\beta Ev(\widehat{d},\Psi)\} && (7) \\ \text{s.t.} \quad p\widehat{d} &= pd+\ell. && (8) \end{aligned}$$

We now turn to the problem faced by the agents during the transactions round. We assume that balances of any two agents that are in a meeting during the transactions round are observable to the PS. As before, agents make reports to the PS about the type of the meeting that they are in (c, p, or n). Those that report a trade meeting as producers receive instructions from the PS on how much to produce. Consumers report the quantity they consumed. The PS subsequently makes balance adjustments depending on these reports.[8]

Not taking into account the agents' balances, there are three possibilities. An agent can be in a consumption, a production, or a no-trade meeting. The vector of policy rules (L_t, K_t, B_t, q_t) determines the respective balance adjustments and the quantity produced in a trade meeting, q_t. These functions in general may depend on the agents' histories of transactions, as summarized by their current balances, as well as on the distribution of balances, Ψ. More precisely, $L_t(K_t)$ is the adjustment for an agent who consumes (produces), 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 during trade meetings is restricted to be positive. After each transaction stage, agents enter the settlement round knowing their new balances.

We will concentrate on arrangements that satisfy certain incentive and participation constraints in the transaction stage. Incentive constraints re-

[7] We describe $E[v(\widehat{d},\Psi)]$ in detail below.

[8] Note that we rule out joint reports by the consumer and the producer in a meeting. See KMT (2005) for a discussion.

quire that the following inequalities hold:

$$\begin{aligned} V(d+L,p) &= V(d+B,p), && (9)\\ -e(q)+V(d+K,p) &\geq V(d+B,p). && (10) \end{aligned}$$

The first constraint equates the continuation utility of a consumer to that of an agent in a no-trade meeting. The equality captures the fact that consumption is not verifiable. The second constraint ensures that agents truthfully report a production meeting. In addition, participation constraints require that producers, consumers, and agents in no-trade meetings, respectively, are better off staying in the system; i.e.,

$$\begin{aligned} -e(q)+V(d+K,p) &\geq 0, && (11)\\ u(q)+V(d+L,p) &\geq 0, && (12)\\ V(d+B,p) &\geq 0. && (13) \end{aligned}$$

We are now ready to formally define a Payment System.

Definition 2 *A Payment System,* $\mathbf{S}$*, is an array of functions* $\mathbf{S} = \{L_t, K_t, B_t, q_t\}$*.* $\mathbf{S}$ *is incentive feasible if it satisfies the incentive and participation constraints.* $\mathbf{S}$ *is simple if balance adjustments do not depend on the agents' current balances. An incentive feasible* $\mathbf{S}$ *is optimal if it supports the efficient allocation.*[9]

Given an incentive feasible payment system, the value function of an agent with balances d at the end of the settlement round, $E[v(d,\Psi)]$, is then given by

$$\begin{aligned} E[v(d,\Psi)] = \int_{d'} \{&\gamma[u(q(d,d'))+V(d+L(d,d'),p)]+\\ &\gamma[e(q(d,d'))+V(d+K(d,d'),p)]+(1-2\gamma)V(d+B(d,d'),p)\}\,d\Psi. \end{aligned} \quad (14)$$

Note that the balance adjustments are in general functions of the agent's own balance, d, and the balance of his trading partner, d'. We next investigate properties of an optimal payment system.

[9] To simplify notation, we suppress the dependence of $\mathbf{S}$ on t.

3.1 Optimal Payment Systems

We let $q_t(\widehat{d}_{t-1}, \widehat{d}'_{t-1})$ stand for the quantity that the PS requires to be produced in a trade meeting, as a function of the balances of the two trading partners. The next Proposition presents a necessary and sufficient condition for an optimal PS to exist in the case where each transaction round is followed by settlement ($n = 1$).

Proposition 3 *There exists a simple optimal PS if and only if $\beta u \geq e$.*

Proof. Let $q_t(\widehat{d}_{t-1}, \widehat{d}'_{t-1})$ be defined by

$$q_t(\widehat{d}_{t-1}, \widehat{d}'_{t-1}) = \begin{cases} q^*, \text{ if } \widehat{d}_{t-1} = \widehat{d}'_{t-1} = 0; \\ 0, \text{ o.w.} \end{cases} \tag{15}$$

Hence, if the balances of both agents are equal to 0, the producer is instructed to produce q^*; otherwise, no production takes place. For all t, define the balance adjustments (K_t, L_t, B_t) by the following three equations:

$$\begin{aligned} B_t &= L_t, & (16) \\ -e + V(\widehat{d}_{t-1} + K_t, p_t) &= V(\widehat{d}_{t-1} + B_t, p_t), & (17) \\ \gamma K_t + \gamma L_t + (1 - 2\gamma) B_t &= 0. & (18) \end{aligned}$$

The first two equations express the incentive constraints (IC). The third equation expresses that, since $E[\ell] = 0$ in each period, aggregate balances remain constant over time. Note that the second equation implies that

$$V(\widehat{d}_{t-1} + K_t, p_t) > V(\widehat{d}_{t-1} + B_t, p_t). \tag{19}$$

We guess that, given balance adjustments (K_t, L_t, B_t) defined above, every agent chooses balances $\hat{d}_t = 0$ in all settlement rounds. To verify that this is consistent with an equilibrium, we need to verify that all participation constraints (PC) hold. These are given by

$$\begin{aligned} V(\hat{d}_{t-1} + B_t, p_t) &\geq 0, & (20) \\ u + V(\hat{d}_{t-1} + B_t, p_t) &\geq 0, & (21) \\ -e + V(\hat{d}_{t-1} + K_t, p_t) &\geq 0. & (22) \end{aligned}$$

Since the second inequality above is satisfied whenever the first one holds, and the third inequality holds whenever the IC conditions hold, it remains to verify the participation constraint which requires that $V(\hat{d}_{t-1} + B_t, p_t) \geq 0$.

Let $X_t \in \{K_t, B_t, L_t\}$ denote the balance adjustment in period t. Stationarity of aggregate equilibrium balances implies that $E[X_{t+1}] = 0$. Using the linearity of V and the stationarity of $\hat{d}$, we then obtain

$$\begin{aligned} V(\hat{d}_{t-1} + B_t, p_t) &= -p_t\hat{d}_t + p_t\hat{d}_{t-1} + p_tB_t + \beta E[v(\hat{d}_t, \Psi)] \\ &= p_tB_t + \beta\gamma(u-e) + \beta E[V(\hat{d}_t + X_{t+1}, p_{t+1})] \\ &= p_tB_t + \beta\gamma(u-e) + \beta V(\hat{d}_t, p_{t+1}) + \beta p_{t+1}E[X_{t+1}] \\ &= p_tB_t + \beta\gamma(u-e) + \beta V(\hat{d}_t, p_{t+1}). \qquad (23) \end{aligned}$$

Since $\hat{d}_{t-1} = \hat{d}_t$, and since V is linear in balances, this yields

$$V(\hat{d}_t, p_t) = \frac{\beta}{1-\beta}\gamma(u-e). \qquad (24)$$

It is easy to check that, given the definition of (K_t, L_t, B_t), setting $K_t = (1-\gamma)\frac{e}{p}$ and $B_t = -\gamma\frac{e}{p}$, for all t, satisfy the IC and the constant-balances constraint. Hence, by the linearity of V,

$$\begin{aligned} V(\hat{d}_{t-1} + B_t, p_t) &= p_tB_t + V(\hat{d}_{t-1}, p_t) \\ &= -\gamma e + \frac{\beta}{1-\beta}\gamma(u-e) \geq 0, \qquad (25) \end{aligned}$$

which holds if $\beta u \geq e$.
For the converse, suppose that $\beta u < e$. Since $B_{t+1} = L_{t+1}$, for all t, the following three conditions must hold for a simple PS to be optimal.

$$\gamma K_{t+1} + (1-\gamma)B_{t+1} = d_{t+1} - d_t \qquad (26)$$

$$p_{t+1}K_{t+1} - e \geq p_{t+1}B_{t+1} \qquad (27)$$

$$(d_t - d_{t+1})p_{t+1} + \frac{\beta}{(1-\beta)}\gamma(u-e) \geq -p_{t+1}B_{t+1}. \qquad (28)$$

The first equation is the law of motion on balances implied by market clearing, while the second one expresses incentive compatibility. The last inequality is the participation constraint, $V(d_t+B_{t+1}, p_{t+1}) \geq 0$. We can now replace $p_{t+1}K_{t+1}$ everywhere to obtain two inequalities that involve only B_{t+1}:

$$\begin{aligned} (d_{t+1} - d_t)p_{t+1} - \gamma e &\geq p_{t+1}B_{t+1} \qquad (29) \\ (d_t - d_{t+1})p_{t+1} + \frac{\beta}{(1-\beta)}\gamma(u-e) &\geq -p_{t+1}B_{t+1}. \qquad (30) \end{aligned}$$

Combining these two we obtain

$$\frac{\beta}{(1-\beta)}\gamma(u-e) \geq \gamma e. \tag{31}$$

This is a contradiction. It is easy to check that the above condition, which can be simplified to give $\beta u \geq e$, also ensures that no agent chooses to exit the economy after selling his balances in the settlement round. ■

Note that the above scheme requires that all agents exit the settlement round with $d = 0$, in all periods. Therefore, the PS requires that Ψ is degenerate.[10] While this adds an additional constraint on the agents, the above Proposition demonstrates that this leads to the first best being supported under the same condition as in the full information case.

It is straightforward to adjust this argument to the case where some transactions are monitored. In addition, using the incentive and participation constraints, we can derive additional properties of the optimal $\mathbf{S}$. Since consumption 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., $B = L < 0$, which means that irrespective of whether agents can consume q^* or are in a no-trade meeting, they are penalized with decreasing balances. In addition, it must be that $p_t(K - B) \geq e$ for all t. In other words, agents are rewarded for producing. Finally, the aggregate net production of the general good is zero in the settlement round; i.e., $E[\ell] = 0$ for all agents. It should be clear, however, that the incentive constraints imply that different agents actually end up with different values of ℓ during the settlement round.

3.2 Settlement Frequency

The case where $n = 1$ literally implies that settlement takes place after every transaction. Next, we assume that a settlement round occurs after each transaction stage of length $n > 1$. In particular, we will derive conditions for a PS to support the efficient allocation with respect to $(q^*, E[\ell^*])$ when $n > 1$. We will assume that the agents' discount factor between transactions is $\tilde{\beta}$, but that there is no discounting between the last transaction and the

[10] Lagos and Wright (2004) study a monetary model in which trade is periodically centralized. They assume quasi-linearity and bargaining in order to obtain a degenerate distribution of money holdings. Since we do not impose bargaining in the transactions round, their result does not hold in our model.

settlement round. We consider two cases. In the first case, there is equal discounting within and across transaction stages ($\tilde{\beta} = \beta$). In the second case, there is no discounting within the transaction stage ($\tilde{\beta} = 1$).

Definition 4 *A PS is history-independent if the balance adjustments in any round s, $1 < s \leq n$, of the transaction stage do not depend on the transactions in the previous $s - 1$ rounds.*

In what follows we will concentrate on a particular class of history-independent PS, which we term *simple repeated PS*. Such PS employ identical discounted balance adjustments during each transaction round. Formally, let $X_t \in \{K_t, L_t, B_t\}$ be the balance adjustment for a PS when $n = 1$. A simple repeated PS has balance adjustments equal to

$$X_{t+s} = \frac{X_{t+n}}{\tilde{\beta}^{n-s}} \tag{32}$$

for all $s = 1, \ldots, n$, where $n - s$ represents the number of transaction rounds until the next settlement stage. The next Proposition gives conditions that are necessary and sufficient for an optimal simple repeated PS to exist.

Proposition 5 *Assume that $\tilde{\beta} = \beta < 1$. There exists an optimal simple repeated PS if and only if $\beta^n u \geq e$. Thus, there exists $\bar{n} \geq 1$ such that a simple repeated PS is optimal if and only if $n < \bar{n}$.*

Proof. Assuming a simple repeated PS, let balances within the transaction rounds be given by

$$X_{t+s} = \frac{X}{\beta^{n-s}} \tag{33}$$

for all $s = 1, \ldots, n$, and all t, where $X \in \{K, L, B\}$ is the balance adjustment for the optimal PS when $n = 1$.

First, note that for a simple repeated PS, all participation constraints are satisfied if the participation constraint for the n-th period holds in the worst-case scenario. Hence, all participation constraints are satisfied if the constraint resulting from consuming n times in a row is satisfied, or if

$$(d_t - d_{t+n})p_{t+n} + \frac{\beta}{1 - \beta^n} \left(\sum_{s=0}^{n-1} \beta^s \right) \gamma(u - e) \geq -p_{t+n} \sum_{s=1}^{n} B_{t+s}. \tag{34}$$

Using $B_{t+s} = L_{t+s}$, for all $s = 1, \ldots, n$, the aggregate law of motion on balances is given by

$$\left[\gamma \sum_{s=1}^{n} K_{t+s} + (1-\gamma) \sum_{s=1}^{n} B_{t+s}\right] = E[\sum_{s=1}^{n} X_{t+s}] = d_{t+n} - d_t. \qquad (35)$$

Finally, in every period, the incentive compatibility constraint is given by

$$-e + \beta^{n-s} p_{t+n} K_{t+s} \geq \beta^{n-s} p_{t+n} B_{t+s}, \qquad (36)$$

for all $s = 1, \ldots, n$. Using the definition of X_{t+s}, this can be written as

$$-e + p_{t+n} K_{t+n} \geq p_{t+n} B_{t+n}. \qquad (37)$$

Since we are employing a simple repeated PS, the last equation implies that incentive compatibility constraints are identical for all periods of the cycle. Using the fact that

$$\sum_{s=1}^{n} X_{t+s} = \frac{1-\beta^n}{\beta^{n-1}(1-\beta)} X, \qquad (38)$$

we obtain for the law of motion

$$p_{t+n} K_{t+n} = \frac{1}{\gamma} \frac{\beta^{n-1}(1-\beta)}{1-\beta^n} (d_{t+n} - d_t) p_{t+n} - \frac{1-\gamma}{\gamma} p_{t+n} B_{t+n}. \qquad (39)$$

Replacing $p_{t+n} K_{t+n}$ in the incentive compatibility and the participation constraints, we obtain two inequalities that involve only the term $\sum_{s=1}^{n} B_{t+s}$:

$$\begin{aligned}(d_{t+n} - d_t) p_{t+n} - \gamma e \frac{1-\beta^n}{\beta^{n-1}(1-\beta)} &\geq p_{t+n} \sum_{s=1}^{n} B_{t+s} &\quad (40)\\ (d_t - d_{t+n}) p_{t+n} + \frac{\beta}{(1-\beta)} \gamma (u - e) &\geq -p_{t+n} \sum_{s=1}^{n} B_{t+s}. &\quad (41)\end{aligned}$$

The efficient allocation is supportable if and only if both inequalities are satisfied. This is the case whenever

$$\frac{\beta}{1-\beta} \gamma (u - e) \geq \gamma e \frac{1-\beta^n}{\beta^{n-1}(1-\beta)}, \qquad (42)$$

or, whenever $\beta^n u \geq e$. Finally, note that if n is large, the left hand side of the above condition converges to zero. Thus, for every $\beta < 1$, there exists N such that for $n > N$ an optimal simple repeated PS does not exist. ■

The following discusses the case where there is no discounting within the transaction stage. It easily follows from the above Proposition.

Corollary 6 *Assume that $\tilde{\beta} = 1$. For any $n \in N$, an optimal simple repeated PS exists if and only if there exists a simple optimal PS for $n = 1$.*

Returning to the general case where $\tilde{\beta} < 1$, it should be clear that, if the optimal allocation can be supported for a given $\widetilde{n}$, then it remains supportable for all $n < \widetilde{n}$. Since the converse is not true, our model suggests that a high frequency of settlement can be beneficial. The intuition behind the case where an optimal simple PS does not exist is related to the one obtained in Levine (1991), but comes from a very different model. More precisely, if settlement is sufficiently infrequent, a positive fraction of agents experience a long sequence of no-production opportunities during the transaction round, resulting in a high balance that eventually needs to be settled. Sufficiently impatient agents will choose to exit the economy rather than going through the settlement process; i.e., the participation constraint will fail.

4 Discussion

We studied simple optimal payment systems in a dynamic model in which the ability of agents to perform certain welfare-improving transactions is subject to random and unobservable shocks. In the absence of settlement, incentive constraints imply that the first-best allocation is not incentive feasible. The first-best is supportable, however, if settlement is introduced, provided that it takes place with a sufficiently high frequency.

The linearity assumption was necessary in order for settlement to affect welfare only indirectly.[11] If the disutility from balance adjustments was non-linear, it would not be possible to achieve the first-best allocation in both the bilateral and the settlement stage. Thus, linearity is necessary in order to support the first-best in our model. One interpretation of this assumption is that settlement costs are linear. For example, if settlement involves the liquidation of certain assets, such as government bonds, such costs are linear in the number of assets liquidated. Alternatively, if settlement uses central bank money, the cost of settlement involves the opportunity cost from holding such money, which corresponds to the interest rate on an equivalent number

[11] Other examples in which some form of linearity is invoked in static mechanism design environments include the classic papers by Clarke (1971) and Groves and Loeb (1975). See also Jarque (2003) for a more recent reference that deals with an intertemporal environment.

of risk-free bonds. A second interpretation is that if settlement involves the need for a sale of indivisible (large-denomination) assets or the supply of extra work hours, linearity captures the fact that it is typically beneficial to use lotteries in order to randomize over the amount of assets that are liquidated, or over the extra work hours supplied.[12]

[12]See Rocheteau, Rupert, Shell and Wright (2005) for a discussion of how the linear structure of lotteries can be beneficial in that regard.

References

[1] Aiyagari R., and S. Williamson (1999) "Credit in a Random Matching Model with Private Information," *Review of Economic Dynamics* 2, 36-64

[2] Aiyagari R., and S. Williamson (2000) "Money and Dynamic Credit Arrangements with Private Information," *Journal of Economic Theory* 91, 248-279

[3] Atkeson A. and R. Lucas Jr. (1993) "On Efficient Distribution with Private Information," *Review of Economic Studies*, vol. 59 (July)

[4] BIS (2003) "Glossary of Terms Used in Payment and Settlement Systems", March.

[5] Clarke E.H. (1971) "Multipart Pricing of Public Goods," *Public Choice* 11, p. 17-33

[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] Groves T. and M. Loeb (1975) "Incentives and Public Inputs," *Journal of Public Economics* 4, p. 211-226

[8] Jarque A. (2003) "Repeated Moral Hazard with Effort Persistence," Manuscript, University of Rochester

[9] Kahn C.M., and W. Roberds (1998) "Payment System Settlement and Bank Incentives," *The Review of Financial Studies* 11(4), p. 845-870

[10] Kahn C.M., and W. Roberds (2004) "Transferability, Finality, and Debt Settlement," Manuscript

[11] Kiyotaki N., and R. Wright (1989) "On Money as a Medium of Exchange," *Journal of Political Economy* 97, p. 927-954

[12] Kiyotaki N., and R. Wright (1993) "A Search-Theoretic Approach to Monetary Economics," *American Economic Review* 83, p. 63-77

[13] Kocherlakota N. (2005a) "Zero Expected Wealth Taxes: A Mirrless Approach to Dynamic Optimal Taxation", *Econometrica* 73(5), p. 1587-1621

[14] Kocherlakota N. (2005b) "Optimal Monetary Policy: What We Know and What We Don't Know" *International Economic Review* 46(2) p. 715-729

[15] Koeppl T., Monnet C., and T. Temzelides (2005) "Payment Systems and Mechanism Design," Manuscript

[16] Lagos R. and R. Wright (2005) "A Unified Framework for Monetary Theory and Policy Analysis," *Journal of Political Economy*, forthcoming

[17] Levine D. (1991) "Asset Trading Mechanisms and Expansionary Policy," *Journal of Economic Theory* 54

[18] Mirrlees J. (1971) "An Exploration in the Theory of Optimal Income Taxation," *Review of Economic Studies* 38, p. 175-208

[19] Rocheteau G., P. Rupert, K. Shell and R. Wright (2005) "A Simple Model of Monetary Exchange Based on Non-Convexities and Sunspots," Manuscript

[20] Shi S. (1997) "A Divisible Search Model of Fiat Money," *Econometrica* 64(1), p. 75-102

[21] Spear S., and S. Srivastava (1987) "On Repeated Moral Hazard with Discounting," *Review of Economic Studies* 54(4)

[22] Temzelides T. and S. Williamson (2001) "Payments Systems Design in Deterministic and Private Information Environments," *Journal of Economic Theory* 99, p. 297-326