



Shadow-Prices in Payment Systems

Rodrigo Peñaloza

Universidade de Brasília
University of California at Los Angeles

Texto nº 241
Brasília, Setembro de 2002

**UNIVERSIDADE DE BRASÍLIA
DEPARTAMENTO DE ECONOMIA**

TEXTO PARA DISCUSSÃO Nº 241

Shadow-Prices in Payment Systems

Rodrigo A. de Souza Peñaloza

Universidade de Brasília
e
University of California at Los Angeles

Brasília, 30 de agosto de 2002

ã Rodrigo Peñaloza, 2002

UNIVERSIDADE DE BRASÍLIA
DEPARTAMENTO DE ECONOMIA
Campus Universitário Darcy Ribeiro
Instituto Central de Ciências
Caixa Postal 04302, 70910-900 Brasília, DF, Brasil
Tel.: (55-61) 3072498, 2723548
Fax: (55-61) 3402311
E-mail: econ@unb.br
<http://www.unb.br/ih/eco>

Chefe do Departamento

Prof. Flávio Rabelo Versiani

Sub-Chefe do Departamento

Prof. Jorge Madeira Nogueira

Coordenador de Pós-Graduação

Prof. André Rossi de Oliveira

Coordenador de Pesquisa e Extensão

Prof. Maurício Soares Bugarin

Coordenador de Graduação

Prof. Jorge Madeira Nogueira

SÉRIE DE TEXTOS PARA DISCUSSÃO

Comissão Editorial, mandato junho de 2001 a setembro de 2002

André Rossi de Oliveira

Bernardo Mueller

Flávio Versiani

Jorge Nogueira

Maurício Bugarin (editor)

Mauro Boianovsky

Apoio: CESPE UnB

Shadow-Prices in Payment Systems

Rodrigo Andrés de Souza Peñaloza*

University of California at Los Angeles (UCLA)

and Universidade de Brasília (UnB)

e-mail: rodrigo@ucla.edu

July, 2002

Abstract

We study the issue of the optimal design of payment systems. Our model describes modern real-time gross settlement systems without or with centralized queueing facilities as networks in which settlement rules and interbank payments are linked by a dual relationship. This duality approach brings the complicated nature of payment systems down to a very simple, geometrically intuitive and useful language. We model the problem of aggregate liquidity management by the Central Bank as an infinite-dimensional linear programming problem, parameterized by different intraday credit policies. We determine the dual problem associated with the primal systemic liquidity management problem and show how shadow-prices of individual banks that participate in the payment system can be used to set intraday monetary policies so as to reduce systemic liquidity and make the payment system more efficient.

JEL CLASSIFICATION: C61, E58, G28.

KEY-WORDS: central bank, real-time gross settlement systems, queueing, shadow-prices, infinite-dimensional linear programming problem.

1 Introduction

The *Bank for International Settlements* (BIS) has been on the vanguard of payment systems' reforms since the 1990's. Its reports on the safety of payment systems have been the primary source of recommendations followed by all the countries willing to improve their large-value

*I greatly thank Joe Ostroy for encouraging me from the very beginning, for his several useful insights, suggestions, and for our conversations. I also thank David Levine (UCLA and UC Berkeley), Alberto Bennardo, Hans Schollhammer (Anderson School at UCLA), David Rahman, the members of the Economic Theory Center at UCLA, and the participants of the Theory Proseminar. This paper is part of my Ph.D. dissertation at UCLA. This research was supported by a fellowship from CAPES. Comments are welcome and of course all the erros are mine.

payment systems. To give a taste of the impact of the BIS on the real world, it suffices to mention that all European Union member countries must adopt real-time gross settlement (RTGS) systems linked to the TARGET system, a supranational network of payment systems.

In a deferred net settlement (DNS) system, interbank large-value payments are settled at the end of the day on a net basis. Clearly this system economizes on liquidity needs, but makes the system prone to contagion. Since interbank large-value payments have increased substantially since the 1970's, the BIS has recommended the use of RTGS systems, which do not economize on liquidity, but make the system considerably less vulnerable to contagion. One problem inherent to RTGS systems is the possibility of *gridlocks*, a situation in which the flow of outgoing payments stops. Gridlocks can occur for two reasons. First, it just might happen that the pattern of interbank payments is such that, given the settlement rule in place, payments cannot go out because payments cannot come in either. Second, it might be a coordination problem. Thinking of free riding on liquidity from other banks, a bank might want to delay its payments to the end of the day. If all the banks think the same way, delaying will be a Nash equilibrium and, as a result, a gridlock occurs.

Since RTGS systems seem to be the rule from now on everywhere, it is important to make it more efficient. Efficiency here refers to the minimization of liquidity needs, that is, to making it as close as possible to DNS systems without getting rid of its (contagion-killing) constraints.

In an RTGS system (with centrally located queue), banks send payment messages to the central bank continuously during the day. Once a payment message arrives, the central bank checks whether the sending bank has sufficient funds to cover the payment. If it does, the payment is settled immediately. Otherwise, it is queued on a centrally located queue and will be settled as soon as the sending bank gets sufficient funds.

According to the BIS, an individual bank's net intraday liquidity at certain point in time is given by its initial balance held at its central bank account plus net transfers from other banks up to that time minus the total value of outgoing queued payments. Since the sum of individual banks' net transfers is zero, the aggregate net intraday liquidity at any point in time is just the sum of initial balances minus the sum of individual banks' total values of outgoing queued payments at that time. Actually, the distribution of net intraday liquidity also matters, but we will not address this issue here.

Since holding liquidity at the central bank is costly, the objective of the central bank is to minimize aggregate net intraday liquidity. Usually, this is done by extending intraday credit to temporarily illiquid banks. For instance,

To ease the shortage of settlement liquidity under an RTGS system, many central banks provide intraday liquidity with certain restrictions. That is, instead of waiting for the arrival of incoming funds to cover outgoing payments, a sending bank without a sufficient account balance can make a payment by borrowing from the central bank

during the day and paying it back before the end of the day. [Zhou (2000, p.32)].

The use of different queueing arrangements also enables the central bank to speed up the flow of outgoing queued payments, which is equivalent to minimizing aggregate net intraday liquidity. Indeed,

[T]he design and operation of payment queues can play an important role in ensuring that available liquidity is used efficiently. For example, a queue based simply on the principle of first in first out might cause large payments to create unnecessary delays to the system's throughput [BIS (2001), core principles].

Among the several queueing arrangements in place, the most commonly used are two variations of the first-in-first-out (FIFO) rule, for instance, the bypass FIFO, and FIFO with prioritization. Central banks also use optimization routines, which consist of reordering the queued messages and netting them out whenever the risk of gridlock is considerably high.

Another useful procedure is the splitting of payments. Instead of settling a payment either in full or not settling it at all, the queueing arrangement might break it down into smaller pieces, so that the flow of outgoing payments does not get stuck by the large-value of a single payment.

Breaking down transactions enables nearly full usage of system liquidity for settlement purposes at all times. This means that liquidity is circulating rapidly from bank to bank and that the system is economizing on its liquidity. Technically, this increases the number of transactions processed in the system. It may also aid in unwinding a gridlock if there is some unused liquidity in the system. [Leinonen (1998)]

Surprisingly enough, little has been done regarding the design of RTGS systems with queueing. For instance,

[V]ery little has been done to investigate the properties of payment systems that combine RTGS [real-time gross settlement] with queueing. (...) [T]he fundamental question that needs to be addressed is in what sense RTGS systems with queueing offer improvements over net settlement or RTGS without queueing, in terms of relevant trade-offs [Roberds (1999, p.4)]

Freixas & Parigi (1998) use, for the first time, the Diamond-Dybvig framework to make comparative welfare analysis of net settlement versus gross settlement in a two-period model with a liquidity shock and – the novelty of their model - a locational shock. After introducing uncertainty and asymmetric information, they show that two inefficient Nash equilibria arise: a bank-run and an equilibrium with potential contagion. Bech and Garrat (2001), in a very

interesting model, use a Bayesian game to analyze the strategic behavior of banks under the real-time gross settlement system. They show that intraday credit policies play an important role in the banks' decision to delay payments. Building on the literature on precautionary demand for reserves, Angelini (1998) shows that priced intraday liquidity creates an incentive for banks to delay payments and that the outcome will not be socially optimal.

However, the models above do not focus on the design of payment systems properly, but rather on the influence of its design on the behavior of banks or the behavior of their depositors. Angelini (1998) and Bech & Garrat (2001) ingeniously analyze the issue of intraday credit policies, and Freixas & Parigi (1998) highlight the issue of contagion, but none of them address the issue of queueing and the system's liquidity.

On the other hand, some applied economists have made significant contributions to the design of queueing arrangements and optimal liquidity usage in payment systems. Koponen & Soramäki (1998) conclude, from several simulations they undertook under different scenarios, that optimization routines, such as splitting of payments and netting enhance the system's liquidity. Leinonen & Soramäki (1999) also make simulations and conclude, among other things, that an RTGS system with queueing is always more efficient than a net settlement system with batch processing. Bech & Soramäki (2001) present a *model of gridlock resolution*, in which the objective is to maximize the flow of outgoing queued payments subject to liquidity constraints faced by banks under real-time gross settlement systems. Building on the *bank clearing problem* modelled by Güntzer, Jungnickel & Leclerc (1998), they introduce an additional constraint, called the *sequence constraint*, which states that payments have to be settled in a predefined order, usually the FIFO rule. The general principles presented in Bech & Soramäki (2001) make their model the only one that is somehow related to ours¹.

From all the papers cited above we conclude that there has been no unique framework for the payment system. In this paper we offer a framework that is different from anything else that has been done so far. Based on the idea that payment systems are a big network with interbank payments flowing around continuously during the day, we constructed a model that gives the payment system a truly geometric structure. In our model, interbank payments and settlement rules are intimately linked by a dual relationship. This duality enables us to write the total outflow of interbank payments as a linear functional, the aggregate net intraday liquidity as an affine functional, and the liquidity constraints faced by banks as a set of fixed points parameterized by intraday credit policies. The problem of the optimal design of settlement rules then reduces to an infinite-dimensional linear programming problem. Basically, we offer a new tool. This tool has been widely used in general equilibrium analysis and other allocational problems, but not in the payment systems literature. In order to give queueing arrangements a role in the flow of payments, we modify the notion of aggregate net intraday liquidity [BIS

¹It is interesting to note that these papers have been published by the Bank of Finland.

(1997), report on RTGS systems]. We interpret the standard notion as potential liquidity and ours as actual liquidity. Indeed, the value of outgoing payments that really matters is the value of payments that *can* be settled according to the settlement rule (i.e., the queueing arrangements).

The construction of our model is entirely based on the reports by the BIS. It provides a framework for the problem in the following quotation:

From the viewpoint of reducing the need for liquidity, the more the system centre can intervene in the queues by reordering and/or using optimization routines, the more efficient the queueing should in principle be because the centre can observe the queued transfers of all banks and thus maximize all the available information to rearrange the transfers in the queue in order that minimises liquidity needs. To the extent that it succeeds in reducing the number and value of queued transfers, such centralized management can contribute to efficiency and the realisation of early settlement in RTGS systems. [BIS, report on RTGS systems (1997, p. 29-30)]

Since DNS systems economize on systemic liquidity, but are prone to contagion, and since RTGS systems eliminate contagion, but require excessive liquidity and are vulnerable to gridlocks, the problem that the design of an RTGS system must address is the minimization of liquidity requirements (in order to approximate it to a DNS system) and to find a settlement rule that minimizes the negative effects of possible gridlocks. How can we design the settlement rule in an RTGS system so as to make it as close as possible to a DNS system without giving up on its contagion-reducing constraints? In this paper we present a model that solves both problems simultaneously.

We do not need to assume that the path of payments is known *ex ante* by the Central Bank. This is obviously not coherent with the real world environment. There is no way for the Central Bank to know the entire path of payments between every pair of banks throughout the day. It is reasonable however to assume that it knows the general pattern of payments, or, more precisely, the distribution of payments. For instance:

The timing of payments across the FEDWIRE Funds Transfer service exhibits a regular pattern over the course of the day, with payment activity peaking in the late afternoon. (McAndrews & Rajan (2000), p.17)

In order to model this feature, we introduce an element of randomness into our framework.

In section 2 we present the model of stochastic shadow-prices of banks in RTGS systems with queueing. We introduce the main concepts, present the primal problem and its dual. Duality in this section is given by the duality bracket $\langle \mathcal{C}(\cdot), \mathcal{BV}(\cdot) \rangle$ between the space of continuous functions on compacta and the space of functions of bounded variation, all paired with a sample space. We describe the functioning of RTGS systems as a linear programming problem in

which the objective function is the maximization of total outflow of payments subject to liquidity constraints, i.e., banks cannot be illiquid anytime. Queueing is important here and arises endogenously. We determine the dual problem to get the shadow-prices of each bank associated with its liquidity constraints and the splitting and queueing of its interbank outgoing payments. In section 3 we present the model without queueing in order to apply our framework to FED-WIRE, the U.S. large-value payment system. In this section, duality is given by the duality bracket $\langle \mathcal{L}_1(\cdot), \mathcal{L}_\infty(\cdot) \rangle$ between the space of essentially bounded functions and integrable functions. Section 4 presents some policy implications for intraday monetary management. Section 5 concludes the paper.

The advantage of our model is twofold. First, our formulation of the payment systems is original. We know of no other model that describes the payment systems as we do. Second, our duality approach brings the complicated nature of payment systems down to a very simple, geometrically intuitive and useful language.

From a practical perspective, the monetary policy implemented by the Central Bank through the payment system gives rise to a series of relevant questions. It is these questions that our model will provide answers to. A sample of such policy questions is:

- When is it that an increase on initial balances enhances the flow of payments?
- Is it a good idea to extend free intraday credit to illiquid banks?
- Is it worth for the payment system to allow for an overnight loan between two banks?
- Does the extension of Lombard (collateralized) loans to certain banks really enhance the flow of payments?
- Can Lombard loans be allocated optimally or else should the Central Bank provide banks with liquidity whenever it is requested to do so?
- Is there an optimal queueing mechanism that helps minimize aggregate liquidity needs? If yes, how can we implement it?
- Can an intraday interbank money market replace the Central Bank's role as a provider of intraday liquidity?
- How can we price Lombard loans to illiquid banks when their probability of failure is positive?
- How does the failure of an individual bank affect the overall flow of payments (contagion through the payment systems)?

By modelling the functioning of real-time gross settlement systems as a linear program, it is possible to find the marginal prices of all such policies by analyzing its corresponding dual problem.

The sketch of our model is the following:

- Bank i , the sender, wires a payment message to the Central Bank at time t requesting a payment transfer from its account to the account of bank j , the receiver.
- If bank i 's account has enough funds, the payment is settled and the transfer becomes final. Otherwise, it is queued and its settlement is postponed until bank i receives more funds.
- At no time can bank i , and every other bank, be illiquid, so when a payment cannot be covered by bank i 's funds, there are two alternative procedures to queueing: the extension of collateralized loans from the Central Bank (repurchase agreements), and the allowance of net debit caps. In section 2 queueing is the only procedure. We consider the alternative procedures in the section 3 where an RTGS system without queueing is presented.
- In both cases (queueing and no queueing), the Central Bank can split payments.
- The primal problem of the Central Bank is to find an optimal way to split and queue payment messages throughout the day so as to minimize systemic liquidity subjected to the constraint that no bank can be illiquid (liquidity constraints) and payments have to be settled at most at full by the end of the day (consistency constraint).
- The dual problem of the Central Bank gives us two important pieces of information: the shadow-prices of banks associated to both liquidity and consistency constraints, and the dual value function. If strong duality holds, we can plug the dual value function into the definition of systemic liquidity. Once this is done, the Central Bank can control intraday monetary policies (the setting of initial reserve requirements, the amount of credit available to each bank, haircuts and interest rates on each Lombard loans, the allowance of overnight loans or the adoption of interbank intraday money market, etc.) by means of shadow-prices in such a way to bring systemic liquidity down to zero or as close to zero as possible in order to approximate the RTGS system to a DNS system with little contagion, since the RTGS constraints continue to hold.

2 Stochastic shadow-prices with queueing

In this section we present the model with queueing. The set $\mathbf{B} = \{1, \dots, n\}$ denotes the set of banks. By banks we mean the financial intermediaries that participate in the real-time gross

settlement system. Obviously some of them are not precisely banks, but for the purposes of the functioning of RTGS systems, such a qualification makes little difference.

Let $\mathbf{T} = [0, T]$ denote the business day, $\mathcal{B}(\mathbf{T})$ the Borel- σ -field over \mathbf{T} , and ν any regular bounded Borel measure on $\mathcal{B}(\mathbf{T})$. The measure space $(\mathbf{T}, \mathcal{B}(\mathbf{T}), \nu)$ represents the time component of our model.

Uncertainty is described by a probability space (S, \mathcal{F}, p) , assumed to be atomless.

The interbank payment from bank $i \in \mathbf{B}$ to bank $j \in \mathbf{B}$ at time $t \in \mathbf{T}$ and state of the world $\omega \in S$ is denoted by $x_{ij}(\omega, t)$. Let $m = n^2$ denote the number of pairs of banks. Once we disregard self-transfers, $m = n(n - 1)$. Given the probability space, (S, \mathcal{F}, p) , let $\mathcal{L}_\infty(S)^m$ be the space of p -essentially bounded \mathbb{R}^m -valued functions on S .

We may define a random path of interbank payments as an array $x = \{x_{ij}(\omega, t) : \omega \in S, t \in \mathbf{T}, i, j \in \mathbf{B}\}$, where $x_{ij}(\omega, t) \geq 0$ and $x_{ii}(\omega, t) \equiv 0, \forall i, j \in \mathbf{B}, \forall t \in \mathbf{T}, \forall \omega \in S$, and such that: (i) $x(\cdot, t)$ is p -measurable, $\forall t \in \mathbf{T}$, and (ii) $x(\omega, \cdot)$ is continuous on \mathbf{T} , p -a.e. Then the random path of interbank payments is a Carathéodory function on $S \times \mathbf{T}$. If this is the case, let $\mathcal{C}(\mathbf{T})^m$ be the space of \mathbb{R}^m -valued continuous functions on \mathbf{T} . The condition $x_{ii}(\omega, t) \equiv 0$ means that a bank cannot send a payment to itself. The condition that an interbank payment be a Carathéodory function is for simplification only. In real life, payment messages arrive at discrete random times, so the right modelling would be to consider interbank payments as left-continuous step functions with random discontinuity points. But if we assume that the time interval between two consecutive payments is small, sometimes to the scale of seconds, then we could just assume that they are left-continuous functions. Thus the assumption of continuity is a good approximation. Besides, there is a well developed theory of Carathéodory functions that can be freely used.

We have to consider the random path of interbank payments as being actually defined on a larger domain, one that will allow us to say that the space of random interbank payments and the space of settlement functions are topologically paired, that is, the later space will be the topological dual of the former space. In order to do that, we first have to properly define the space the settlement functions belong to.

The bilateral settlement function between banks i and j is given $v_{ij}(\omega, \tau, t)$. At state $\omega \in S$, $v_{ij}(\omega, \tau, t)$ is the fraction of payment $x_{ij}(\omega, \tau)$ sent at time τ and settled at time $t \geq \tau$. Therefore, any payment may be fractioned and queued. The time domain of the bilateral settlement function is the triangle $A = \{(\tau, t) \in \mathbf{T} \times \mathbf{T} : t \geq \tau\}$. Obviously, $0 \leq v_{ij}(\omega, \tau, t) \leq 1$.

Given the measure space $(\mathbf{T}, \mathcal{B}(\mathbf{T}), \nu)$, let \mathcal{G} the product-Borel- σ -algebra $\mathcal{B}(\mathbf{T}) \times \mathcal{B}(\mathbf{T})$ restricted to the compact set $A \subset \mathbf{T} \times \mathbf{T}$, i.e., $\mathcal{G} = \mathcal{B}(\mathbf{T}) \times \mathcal{B}(\mathbf{T})|_A$.² Let μ be a measure³ on \mathcal{G} . Given the time domain, $A = \{(\tau, t) \in \mathbf{T} \times \mathbf{T} : t \geq \tau\}$, consider the measure space (A, \mathcal{G}, μ) . Let $\mathcal{L}_1(A)^m$ be the space of μ -integrable \mathbb{R}^m -valued functions on A and let $\mathcal{L}_\infty(A)^m$ be the space of

²This means that a set $S \in \mathcal{G}$ if there exists a set $R \in \mathcal{B}(T) \times \mathcal{B}(T)$ such that $S = R \cap A$.

³We may take $\mu = (\nu \times \nu)|_A$.

μ -essentially bounded \mathbb{R}^m -valued functions on A . Finally, let $\mathcal{C}(A)^m$ denote the Banach space of \mathbb{R}^m -valued continuous functions on A , and let $\mathcal{BV}(A)^m$ denote the Banach space of \mathbb{R}^m -valued functions of bounded variation on A , with the total variation as the norm.

Recall that a function of bounded variation can be decomposed into the sum of its absolutely continuous part and its singular part:

$$y(\cdot, a) = y^a(\cdot, a) + y^s(\cdot, a)$$

where $y^a(\cdot, \cdot) \ll \mu$ is the absolutely continuous part and $y^s(\cdot, a) \perp \mu$ is the singular part (its derivative vanishes almost everywhere). Moreover, there exists a constant $M > 0$ such that $|y_{ij}^s(\cdot, a)| \leq M, \forall i, j \in \mathbf{B}$. We set $M = 1$.

The space of random paths of interbank payments can trivially be immersed into the space of \mathbb{R}^m -valued Carathéodory functions β on $S \times A$, i.e., functions satisfying: (i) $\beta(\cdot, \tau, t)$ is p -measurable, $\forall(\tau, t) \in A$, and (ii) $x(\omega, \cdot)$ is continuous on A , $p - a.e.$ Indeed, it suffices to set $\beta_{ij}(\omega, \tau, t) = x_{ij}(\omega, \tau), \forall(\tau, t) \in A$. This identification allows us to consider the random path of interbank payments and the settlement functions as being defined on the same domain, namely $S \times A$.

That being told, we may redefine the random path of interbank payments in a more appropriate way.

Definition 1 A random path of interbank payments is a function $x : S \times A \rightarrow \mathbb{R}^m$ such that:

- (i) $x(\omega, \tau, t) = (x_{ij}(\omega, \tau, t) : i, j \in \mathbf{B})$
- (ii) $x_{ij}(\omega, \tau, t) \geq 0$ and $x_{ii}(\omega, \tau, t) \equiv 0, \forall i, j \in \mathbf{B}, \forall(\tau, t) \in A, p - a.e.,$
- (iii) $x_{ij}(\omega, \tau, t) = x_{ij}(\omega, \tau, \tau), \forall i, j \in \mathbf{B}, \forall t \geq \tau$
- (iv) $x(\cdot, \tau, t) \in \mathcal{L}_\infty(S)^m$, i.e., $x(\cdot, \tau, t)$ is p -essentially bounded on $S, \forall(\tau, t) \in A$
- (v) $x(\omega, \cdot) \in \mathcal{C}(A)^m, p - a.e.,$ i.e., $x(\omega, \cdot)$ is continuous on $A, p - a.e.$

Definition 2 A stochastic settlement density is any function $v : S \times A \rightarrow [0, 1]^m$ such that:

- (i) $v(\omega, \tau, t) = (v_{ij}(\omega, \tau, t) : i, j \in \mathbf{B}),$ with $v_{ij} : S \times A \rightarrow [0, 1]$ being called the bilateral stochastic settlement density between banks i and j .
- (ii) $v_{ij}(\omega, \cdot) \in \mathcal{L}_1(A)$, i.e., $v_{ij}(\omega, \cdot)$ is μ -integrable on $A, \forall i, j \in \mathbf{B}, p - a.e.$
- (iii) $\int_\tau^T v_{ij}(\omega, \tau, s) d\mu_2(s) \leq 1, \forall i, j \in \mathbf{B}, \forall \tau \in [0, T], p - a.e.,$ where μ_2 is the second marginal distribution of μ . (consistency condition)
- (iv) $v_{ij}(\cdot, \tau, t) \in \mathcal{L}_\infty(S)$, i.e., is p -essentially bounded on $S, \forall i, j \in \mathbf{B}, \forall(\tau, t) \in A$.

Definition 3 A stochastic settlement integral is any \mathbb{R}^m -valued function y on $S \times A$ such that:

- (i) $y(\omega, \cdot) \in \mathcal{BV}(A)^m, p - a.e., y(\omega, \tau, t) = (y_{ij}(\omega, \tau, t) : i, j \in \mathbf{B}),$ with $y_{ij} : S \times A \rightarrow [0, 1]$ being called the bilateral stochastic settlement integral between banks i and j .
- (ii) $y_{ij}(\omega, \tau, T) - y_{ij}(\omega, \tau, \tau) \leq 1, \forall i, j \in \mathbf{B}, \forall \tau \in [0, T], p - a.e.$ (consistency condition)
- (iii) $\frac{dy_{ii}^a}{d\mu}(\cdot, \tau, t) \in \mathcal{L}_\infty(S), \forall(\tau, t) \in A, \forall i, j \in \mathbf{B}$

From the above definitions we can easily conclude that a stochastic settlement density can be interpreted as the Radon-Nikodym derivative of a stochastic settlement integral. Condition (ii) in the definition of stochastic settlement integrals can be restated as $\int_t^T v_{ij}(\omega, t, s) d\mu_2(s) \leq 1$, $\forall \tau \in [0, T]$, $p - a.e.$, provided we set the settlement density as the Radon-Nikodym derivative of a stochastic settlement integral, that is, $v_{ij}(\omega, t, s) = \frac{dy_{ij}}{d\mu}(\omega, t, s)$. In this sense:

$$y_{ij}(\omega, \tau, T) - y_{ij}(\omega, \tau, \tau) = \int_{\tau}^T \frac{dy_{ij}}{d\mu}(\omega, \tau, s) d\mu_2(s), \forall \tau \in [0, T], p - a.e.$$

where μ_2 is the second marginal distribution of μ , which will coincide with ν whenever μ is the product-measure $\nu \times \nu$ restricted to A . Similarly, μ_1 is the first marginal distribution of μ .

The consistency condition says that payments keep being partially settled until the end of the day. Moreover, it is possible that some portion of the payment remains unsettled at the end of the day. We chose to impose the weak inequality instead of the equality in order to give some room for overnight interbank loans.

The value $x_{ij}(\omega, \tau, t)v_{ij}(\omega, \tau, t) = x_{ij}(\omega, \tau, \tau)v_{ij}(\omega, \tau, t)$ represents the portion of the interbank transfer $x_{ij}(\omega, \tau, \tau)$ sent at period $\tau \in [0, T]$ and state $\omega \in S$ and settled at period $t \geq \tau$. The total expected cumulated outflow of payments is then given by the linear functional:

$$\sum_{i \in \mathbf{B}} \sum_{j \in \mathbf{B}} \int_S \int_0^T \int_{\tau}^T x_{ij}(\omega, \theta, s) v_{ij}(\omega, \theta, s) d\mu(s, \theta) dp(\omega)$$

In order to stick with our duality formulation, we will say that the total expected cumulated outflow of payments is given instead by the linear functional expressed in terms of settlement measures:

$$\sum_{i \in \mathbf{B}} \sum_{j \in \mathbf{B}} \int_S \int_0^T \int_{\tau}^T x_{ij}(\omega, \theta, s) dy_{ij}(\omega, \theta, s) dp(\omega)$$

Definition 4 *The expected systemic liquidity of a real-time gross settlement system is given by total initial reserves at the beginning of the day minus the expected total value of queued outgoing payments averaged over time, that is:*

$$\Lambda = \sum_{i \in \mathbf{B}} B_o^i - \frac{1}{\nu(\mathbf{T})} \sum_{i \in \mathbf{B}} \sum_{j \in \mathbf{B}} \int_S \int_0^T \int_{\tau}^T x_{ij}(\omega, \theta, s) dy_{ij}(\omega, \theta, s) dp(\omega)$$

The definition of systemic liquidity is a generalization of the standard definition as given by the BIS report on RTGS systems (BIS, 1997). Our definition takes the role of queueing into account and considers the cumulated total outflow of payments throughout the day.

In RTGS systems, banks have to hold enough balance throughout the day to settle interbank payments. Overdrafts are not allowed. At each period t , individual reserves are given by individual initial balances plus net transfers up to time t . Period t is not included. Thus, individual banks face a kind of *cash-in-advance* (or Clower) constraint. According to Clower's (1967) seminal paper:

[T]he total value of goods demanded cannot in any circumstances exceed the amount of money held by the transactor at the outset of the period. [Clower (1967)]
 (Italics in the original)

Though Clower meant goods demanded by consumers in a monetary economy, the nature of such liquidity constraints is the same one faced by banks in an RTGS system, the difference being only that the objective function is linear and that the goods we are dealing with are settlement functions.

Thus in real-time gross settlement systems no bank can be illiquid. Whatever the state of the world, at any period the following liquidity constraint must be satisfied:

$$\sum_{j \in \mathbf{B}} \int_{\tau \in [0, t]} x_{ij}(\omega, \tau, t) y_{ij}^a(\omega, \tau, t) d\mu_1(\tau) \leq B_o^i + \int_{\theta \in [0, t]} \int_{\tau \in [0, \theta]} \sum_{j \in \mathbf{B}} x_{ji}(\omega, \tau, \theta) dy_{ji}(\omega, \tau, \theta) - \int_{\theta \in [0, t]} \int_{\tau \in [0, \theta]} \sum_{j \in \mathbf{B}} x_{ij}(\omega, \tau, \theta) dy_{ij}(\omega, \tau, \theta)$$

The expression on the left-hand side is the total value of queued payments from bank i to every other bank at time t and state of the world ω . The first term on the right-hand side is bank i 's initial reserve. The second term on the right-hand side is the total value of payments that bank i had received from the system up to time t . The third term on the right-hand side is the total value of payments that bank i had sent to the system up to time t . Therefore, the right-hand side is initial balance plus net transfers up to time t .

In each period, a bank can only use its current balance for settlement purposes. It is in accord with the main principle of RTGS systems: banks are not allowed to become illiquid at any time whatever the state of the world. As transfers are made throughout the day, individual balances fluctuate within the interval $[0, \sum_{i \in \mathbf{B}} B_o^i]$. Indeed, the minimum balance is zero, and the maximum balance occurs when an individual bank absorbs all liquidity from the system. This is however an extreme situation. In order to avoid such extreme distribution of liquidity, some systems impose an upper bound on current balances. For instance, in some RTGS systems, an individual balance fluctuates within the interval $[0, 2B_o^i]$, that is, maximum current balance is twice the initial balance. Such upper bounds are indeed important for the safety of payment systems. It is well known that the distribution of liquidity can cause gridlocks. It is important to remark that our model can also deal with the control of distribution of liquidity. In this case however we would have a quadratic component. For instance, we could impose an additional constraint in which the variance of the distribution of liquidity in the system is bounded. The shadow-price associated with such constraint would give us the price of a unit change in the distribution of liquidity. Nevertheless, the model would become quadratic, and we do not want to go beyond the linear programming framework. A stronger constraint of course – and one that would keep the model linear – would be the imposition of upper bounds on current balances, but doing that would make the model less simple without additional benefits.

We will now write the primal problem in matrix form. This will make it easier for us to find out the dual problem. The resulting dual solution will be the stochastic shadow-prices of banks in the real-time gross settlement system.

Notice that we did not include the extension of Lombard loans in our stochastic version. We are here dealing only with the issue of queueing and partial settlement. The stochastic shadow-prices associated with liquidity constraints will give the marginal prices of initial reserve requirements. The stochastic shadow-prices associated with consistency constraints will give the marginal prices of queueing bilateral payments and settling them at full.

Define the following matrix, $\forall i \in \mathbf{B}, \forall(\omega, \tau, t) \in S \times A$:

$$X_i(\omega, \tau, t) = \begin{bmatrix} 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \\ x_{i1}(\omega, \tau, t) & \cdots & x_{in}(\omega, \tau, t) \\ 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{bmatrix}_{n \times n} \quad \leftarrow i^{th} \text{ row}$$

Also define:

$$Y_i(\omega, \tau, t) = \begin{bmatrix} x_{i1}(\omega, \tau, t) & 0 & \cdots & 0 \\ 0 & x_{i2}(\omega, \tau, t) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & x_{in}(\omega, \tau, t) \end{bmatrix}_{n \times n}$$

Consider the partitioned matrices:

$$\begin{aligned} X(\omega, \tau, t) &= \begin{bmatrix} X_1(\omega, \tau, t) & | & \cdots & | & X_n(\omega, \tau, t) \end{bmatrix}_{n \times n^2} \\ Y(\omega, \tau, t) &= \begin{bmatrix} Y_1(\omega, \tau, t) & | & \cdots & | & Y_n(\omega, \tau, t) \end{bmatrix}_{n \times n^2} \end{aligned}$$

Then the expression $\sum_{j \in \mathbf{B}} x_{ji}(\omega, \tau, \theta) dy_{ji}(\omega, \tau, \theta) - \sum_{j \in \mathbf{B}} x_{ij}(\omega, \tau, \theta) dy_{ij}(\omega, \tau, \theta)$, $\forall i \in \mathbf{B}$, can be written in matrix form, $\forall(\tau, t) \in A$, $p - a.e.$, as:

$$Y(\omega, \tau, \theta) dy(\omega, \tau, \theta) - X(\omega, \tau, \theta) dy(\omega, \tau, \theta) = [Y(\omega, \tau, \theta) - X(\omega, \tau, \theta)] dy(\omega, \tau, \theta)$$

The expression $\sum_{j \in \mathbf{B}} \int_{[0,t]} x_{ij}(\omega, \tau, t) y_{ij}^a(\omega, \tau, t) d\mu_1(\tau)$ can be written in matrix form as:

$$\int_{[0,t]} X(\omega, \tau, t) y^a(\omega, \tau, t) d\mu_1(\tau)$$

Clearly the objective function is a linear functional that can be represented as $x \cdot y$, the inner product on $\langle \mathcal{L}_\infty(S)^m \times \mathcal{C}(A)^m, \mathcal{L}_\infty(S)^m \times \mathcal{BV}(A)^m \rangle$. For the time being, however, let us write it as $\int_S \int_A x(\omega, a) dy(\omega, a) dp(\omega)$.

Let I_m be the identity m -matrix and $\mathbf{1}$ the m -vector of 1's. Let $B_o = (B_o^i : i \in \mathbf{B})$ be the vector of initial reserves. The consistency constraint $\int_\tau^T \frac{dy_{ij}}{d\mu}(\omega, \tau, s) d\mu_2(s) \leq 1, \forall i, j \in \mathbf{B}, \forall t \in [0, T], p - a.e.$, can be written as

$$\int_\tau^T I_m \frac{dy}{d\mu}(\omega, \tau, s) d\mu_2(s) \leq \mathbf{1}, \forall \tau \in [0, T], p - a.e.$$

Therefore in matrix notation the primal problem is:

$$\left\{ \begin{array}{l} \text{sup}_y \int_S \int_A x(\omega, a) dy(\omega, a) dp(\omega) \\ \text{s.t.} \int_{[0,t]} X(\omega, \tau, t) y^a(\omega, \tau, t) d\mu_1(\tau) \leq B_o + \int_{A^-(t)} [Y(\omega, a) - X(\omega, a)] dy(\omega, a), p - ae \\ \int_\tau^T I_m \frac{dy}{d\mu}(\omega, \tau, s) d\mu_2(s) \leq \mathbf{1}, \forall \tau \in [0, T], p - ae \\ y(\omega, \cdot) \in \mathcal{BV}(A)_+^m, p - ae \\ \frac{dy^a}{d\mu}(\cdot, \tau, t) \in \mathcal{L}_\infty(S)^m, \forall (\tau, t) \in A. \end{array} \right.$$

In the primal problem above, we have the following:

- (a) $Y - X \in \mathcal{L}_\infty(S)^{nm} \times \mathcal{C}(A)^{nm}$
- (b) $\int_{[0,\cdot]} X(\cdot, \tau, \cdot) y^a(\cdot, \tau, \cdot) d\mu_1(\tau) \in \mathcal{L}_\infty(S)^{nm} \times \mathcal{C}(A)^{nm}$
- (c) $x \in \mathcal{L}_\infty(S)^m \times \mathcal{C}(A)^m$

The conditions above guarantee that the problem is well-posed [Papageorgiou (1982)] and that strong duality holds.

The notation above is for simplification purposes only. The more explicit notation is:

$$\left\{ \begin{array}{l} \text{sup}_y \sum_{i \in \mathbf{B}} \sum_{j \in \mathbf{B}} \int_S \int_0^T \int_\tau^T x_{ij}(\omega, \theta, s) dy_{ij}(\omega, \theta, s) dp(\omega) \\ \text{s.t.} \sum_{j \in \mathbf{B}} \int_0^t x_{ij}(\omega, \tau, t) y_{ij}^a(\omega, \tau, t) d\mu_1(\tau) \leq B_o^i + \int_0^t \int_0^\theta \sum_{j \in \mathbf{B}} x_{ji}(\omega, \tau, \theta) dy_{ji}(\omega, \tau, \theta) \\ \quad - \int_0^t \int_0^\theta \sum_{j \in \mathbf{B}} x_{ij}(\omega, \tau, \theta) dy_{ij}(\omega, \tau, \theta) \\ \int_\tau^T \frac{dy_{ij}}{d\mu}(\omega, \tau, s) d\mu_2(s) \leq 1, \forall i, j \in \mathbf{B}, \forall \tau \in [0, T], p - a.e. \\ y(\omega, \cdot) \in \mathcal{BV}(A)_+^m, p - ae \\ \frac{dy_{ij}^a}{d\mu}(\cdot, \tau, t) \in \mathcal{L}_\infty(S), \forall i, j \in \mathbf{B}, \forall (\tau, t) \in A. \end{array} \right.$$

The Central Bank's primal problem is a generalized continuous linear program. The underlying space in the primal problem is $\mathcal{L}_\infty(S)^{nm} \times \mathcal{C}(A)^{nm}$. Its topological dual space is $\mathcal{L}_1(S)^{nm} \times \mathcal{BV}(A)^{nm}$. However, since we imposed the consistency condition, we can restrict it, with no loss of generality, to the linear subspace $\mathcal{L}_\infty(S)^{nm} \times \mathcal{BV}(A)^{nm}$. Then the duality bracket on which our linear programming problem is defined is:

$$\langle \mathcal{L}_\infty(S)^{nm} \times \mathcal{C}(A)^{nm}, \mathcal{L}_\infty(S)^{nm} \times \mathcal{BV}(A)^{nm} \rangle$$

Consider the quadruple $(i, j, \omega, t) \in \mathbf{B} \times \mathbf{B} \times S \times \mathbf{T}$ representing bank i , bank j , time t and a state of the world ω . For each liquidity constraint associated with $(i, i, \omega, t) \in \mathbf{B} \times \mathbf{B} \times S \times \mathbf{T}$, let $\lambda_i(\omega, t)$ be its shadow-price. We call $\lambda_i(\omega, t)$ *bank i 's liquidity shadow-price at state ω and time t* . For each consistency constraint associated with (i, j, ω, t) , let $\xi_{ij}(\omega, t)$ be its shadow-price. We call $\xi_{ij}(\omega, t)$ the *consistency shadow-price of the pair (i, j) at state ω and time t* .

The parameter associated with a particular liquidity constraint is the initial reserve balance held by a particular bank at the beginning of the day. Thus its liquidity shadow-price measures the effect on the outflow of payments of a unit variation in initial reserves. Given the liquidity shadow-price, the Central Bank can draw a better understanding of the effect of changes in initial reserves on the level of aggregate liquidity needs.

The parameter associated with a particular consistency constraint is the number 1. The interpretation of the consistency shadow-price requires further thoughts. When a payment $x_{ij}(\omega, \tau, \tau)$ is sent during the day, say at time τ , the primal solution will determine how much of it should be settled from that time on until the end of the day. At each time $t \geq \tau$, the amount $x_{ij}(\omega, \tau, \tau)v_{ij}(\omega, \tau, t)$ will be settled. By the end of the day, at most 100% of it should have been settled. Thus it is possible that some fraction remains optimally unsettled. So a unit increase on the parameter 1 means a 100% increase in the payment settlement. That is, a bank sends a payment during the day and, by the end of it, twice the payment has been settled. That means that that bank made a payment, settled it by the end of the day, and in addition *lent* to the receiving bank the very same amount it has paid. Therefore the true dollar effect of an extra dollar *lending* from a bank to another is measured by the corresponding consistency shadow-price divided by the amount transferred, $\frac{\xi_{ij}(\omega, \tau)}{x_{ij}(\omega, \tau, \tau)}$. If at the primal solution the consistency constraint holds with equality, then the consistency shadow-price will tell us by how much the flow of payments (or equivalently, minimum aggregate liquidity needs) will change should a bank pay an extra dollar over and above the amount due or should a bank pay a dollar less than the true amount due.

Denote by $\lambda(\omega, t) = (\lambda_i(\omega, t) : i \in \mathbf{B})$ the vector of liquidity shadow-prices at state ω and time t . Let $\xi(\omega, t) = (\xi_{ij}(\omega, t) : i, j \in \mathbf{B})$ be the vector of consistency shadow-prices at state ω and time t .

The dual problem is given by:

$$\left\{ \begin{array}{l} \text{inf}_{(\lambda, \xi)} \quad \int_S \int_A B_o d\lambda(\omega, a) + \int_S \int_A \mathbf{1}_m \cdot y^a(\omega, \tau, t)^\perp X(\omega, \tau, t)^\perp d\lambda(\omega, \tau, t) \\ \quad + \int \int_A \mathbf{1}_m \cdot d\xi(\omega, \tau, t) \\ \text{s.t.} \quad \int_{[0, t]} y^a(\omega, \tau, t)^\perp X(\omega, \tau, t)^\perp \frac{d\lambda(\omega, \tau, t)}{d\mu} d\mu_1(\tau) + \\ \quad + \int_t^T [X(\omega, t, \tau)^\perp - Y(\omega, t, \tau)^\perp] d\lambda(\omega, t, \tau) \geq x(\omega, t, \tau) \\ \lambda(\omega, \cdot) \in \mathcal{BV}(A)_+^n, \quad p - a.e. \\ \frac{d\lambda^a(\cdot, a)}{d\mu} \in \mathcal{L}_\infty(S), \quad \forall a \in A \end{array} \right.$$

We postpone the analysis of the dual to the next section and to the particular case of

FEDWIRE, with no centralized queueing facilities.

3 Stochastic shadow-prices without queueing

In this section we consider the special case of RTGS systems without queueing facilities, such as FEDWIRE. From a topological point of view, this problem is less general, so we can add more geometrical structure to the framework and get some additional insights about the dual relationship between interbank payments and settlement functions.

From a practical point of view, interbank payment messages arrive at the central bank at sporadic discrete times, so that they can be regarded as a sequence (actually, a vector or even a matrix, depending on the way we write it). Since several payment messages may arrive on a minute-by-minute (or even second-by-second) basis during the day, we will assume that time is continuous and that interbank payments are measurable functions on the time interval. Of course, measurability here is understood as Borel-measurability. Measurability is the weakest assumption we can make without losing mathematical structure.

Let $\mathbf{T} = ([0, T], \mathcal{B}([0, T]), \nu)$ be a measure space representing the time interval (a business day), where λ is the Lebesgue measure on the Borel- σ -algebra $\mathcal{B}([0, T])$. Recall that \mathbf{T} is an atomless complete finite measure space⁵. Let $\mathcal{L}_\infty(\mathbf{T})$ be the space of measurable bounded real functions on \mathbf{T} , endowed with the sup-norm.

There is no way for the central bank to know *ex ante* what the pattern of payments will precisely be during the day. This pattern is uncertain. Usually it is concentrated towards the end of the day, but, again, it is random. Therefore, the flow of payments is the sample path of a stochastic process. To model this feature, let (S, \mathfrak{S}, p) be a complete atomless probability space.

$\mathbf{B} = \{1, \dots, n\}$ denotes the set of banks. Let $x_{ij}(t, \omega)$ be the monetary value of the payment from bank i to bank j on the books of the central bank at time t and state ω . Assume that $x_{ij} : \mathbf{T} \times S \rightarrow [0, \infty)$ is a stochastic process such that $x_{ij}(t, \cdot) \in \mathcal{L}_\infty(p)_+$, $\forall t \in \mathbf{T}$, and, $x_{ij}(\cdot, \omega) \in \mathcal{L}_\infty(\nu)_+$, $p - a.e.$ Now define the function $x_i : \mathbf{T} \rightarrow \mathcal{L}_\infty(p)_+^n$, where $\mathcal{L}_\infty(p)_+^n = \mathcal{L}_\infty(p, [0, \infty)^n)$, by $x_i(t) = (x_{ij}(t))_{j \in \mathbf{B}}$, where $x_{ij}(t) \in \mathcal{L}_\infty(p)_+$, $\forall t \in [0, T]$. Finally, define x by $x = (x_i)_{i \in \mathbf{B}}$.

Given $m = n^2$, let $E = \mathcal{L}_1(p, \mathbb{R}^m)$ be the space of integrable functions with values in \mathbb{R}^m , endowed with the weak topology. Recall that E is a locally convex, separable, metrizable linear topological space and that its topological dual is $E^* = \mathcal{L}_\infty(p, \mathbb{R}^m)$. Whenever $f \in E$ (or E^*), we assume that $f = (f_1, \dots, f_n)$ and $f_i = (f_{ij})_{j \in \mathbf{B}}$. It is useful to see $f \in E$ (evaluated at ω) as a matrix:

$$f(\omega) = \begin{bmatrix} f_{11}(\omega) & \cdots & f_{1n}(\omega) \\ \vdots & \ddots & \vdots \\ f_{n1}(\omega) & \cdots & f_{nn}(\omega) \end{bmatrix}$$

⁵Notice that λ may also be any atomless measure other than Lebesgue's.

With this representation in mind, the payments attached to queued messages at certain time $t \in \mathbf{T}$ and state $\omega \in S$ are summarized by:

$$x(t, \omega) = x(t)(\omega) = \begin{bmatrix} x_{11}(t, \omega) & \cdots & x_{1n}(t, \omega) \\ \vdots & \ddots & \vdots \\ x_{n1}(t, \omega) & \cdots & x_{nn}(t, \omega) \end{bmatrix}$$

Thus, by our assumptions, for each $t \in \mathbf{T}$, $x(t) \in E^* = \mathcal{L}_\infty(\mu, \mathbb{R}^m)$ and $x \in \mathcal{L}_\infty(\nu, E^*)$. The space of payments is then the positive cone of the space of all (equivalence classes of) Bochner-integrable functions that are essentially bounded, where the norm is given by $\|x\| = \text{ess sup}\{\|x(t)\| : t \in \mathbf{T}\}$.

Notice that we implicitly assume that there is exactly one message to each bank queued at time t , so that the size of the queue is always n (obviously the payment of an individual bank to itself is zero). This not an oversimplification. Indeed, banks usually send bulk payment messages to another bank for any short period of time, so it is reasonable to assume that the queue has at most one message to each bank for any short period of time. So instead of sending a payment message to bank j and few seconds later sending another one to the same bank, a bank i actually sends both payments together.

We will now construct a definition of settlement function that is dual to the definition of interbank payments function.

Consider the following subset of E :

$$K = \{f \in E : f(\omega) \in [0, 1]^m, p - a.e. \text{ and } f_{ii}(\omega) = 0, \forall \omega \in S, \forall i \in \mathbf{B}\}.$$

Clearly, $K \subset E = \mathcal{L}_1(p, \mathbb{R}^m)$ is convex and nonempty. The set K is weakly compact separable and metrizable for the relative topology induced from $\mathcal{L}_1(p, \mathbb{R}^m)$, and is uniformly integrable.

Let $v_i : [0, T] \rightarrow \mathcal{L}_1(p, [0, 1]^n)$, be defined by $v_i(t) = (v_{ij}(t))_{j \in \mathbf{B}}$, where $v_{ij}(t) \in \mathcal{L}_1(p, [0, 1])$, $\forall t \in \mathbf{T}$. Let $m = n^2$. For ease of notation, we will use the isometric isomorphism $\mathcal{L}_1(p, [0, 1]^n)^n \approx \mathcal{L}_1(p, [0, 1]^m)$. Assume that $v : \mathbf{T} \rightarrow K$ is $(\mathcal{B}([0, T]), \mathcal{B}(K))$ -measurable, where $\mathcal{B}(K)$ is the Borel- σ -algebra generated by the relative weak topology on K inherited from E . We call v a *settlement function*. Let $\mathcal{L}_1(\nu, K)$ be the set of Bochner-integrable settlement functions⁶.

Then $v(t, \omega)$ has representation:

$$v(t, \omega) = v(t)(\omega) = \begin{bmatrix} v_{11}(t, \omega) & \cdots & v_{1n}(t, \omega) \\ \vdots & \ddots & \vdots \\ v_{n1}(t, \omega) & \cdots & v_{nn}(t, \omega) \end{bmatrix}$$

We have that, for each $t \in \mathbf{T}$, $v(t) \in E = \mathcal{L}_1(p, \mathbb{R}^m)$ and $v \in \mathcal{L}_1(\nu, E)$.

⁶Given a finite measure space (Θ, F, ν) and a Banach space E , a ν -measurable function $f : \Theta \rightarrow E$ is Bochner-integrable if, and only if, $\int_\Theta \|f\| d\nu < \infty$ [Diestel & Uhl (1977)].

Definition 6 A settlement function is any Bochner-integrable function $v : \mathbf{T} \rightarrow K$, that is, any Bochner-integrable function $v \in \mathcal{L}_1(\nu, E)$ given by $v(t)(\omega) = v(t, \omega) = (v_1(t, \omega), \dots, v_n(t, \omega))$ such that $v_{ij}(t, \omega) \in [0, 1]$, $\forall i, j \in \mathbf{B}$, $v_{ii}(t, \omega) \equiv 0$, $\forall i \in \mathbf{B}, \forall t \in \mathbf{T}$, $p - a.e.$

Settlement functions and interbank payments functions are then dual concepts. To be more specific, on the dual system $\langle \mathbb{R}^m, \mathbb{R}^m \rangle$, given $v(t, \omega) \in \mathbb{R}^m$ and $x(t, \omega) \in \mathbb{R}^m$, the bilinear form $*$: $\mathbb{R}^m \times \mathbb{R}^m \rightarrow \mathbb{R}$ is given by $x(t, \omega) * v(t, \omega) = \sum_{(i,j) \in \mathbf{B} \times \mathbf{B}} x_{ij}(t, \omega) v_{ij}(t, \omega)$. On the dual system $\langle E, E^* \rangle$ the bilinear form $\langle \cdot, \cdot \rangle : E^* \times E \rightarrow \mathbb{R}$ is given by:

$$\langle x(t), v(t) \rangle = \int_S x(t, \omega) * v(t, \omega) dp(\omega) = \int_S \sum_{(i,j) \in \mathbf{B} \times \mathbf{B}} x_{ij}(t, \omega) v_{ij}(t, \omega) dp(\omega)$$

On the dual system $\langle \mathcal{L}_1(\nu, E), \mathcal{L}_\infty(\nu, E^*) \rangle$ consider the bilinear form \cdot : $\mathcal{L}_\infty(\nu, E^*) \times \mathcal{L}_1(\nu, E) \rightarrow \mathbb{R}$ is given by:

$$x \cdot v = \int_0^T \langle x(t), v(t) \rangle d\lambda(t) = \int_0^T \int_S \sum_{(i,j) \in \mathbf{B} \times \mathbf{B}} x_{ij}(t, \omega) v_{ij}(t, \omega) dp(\omega) d\lambda(t)$$

Now it is easy to see the topological duality between payments and settlement. Given the Banach space $E = \mathcal{L}_1(p, \mathbb{R}^m)$ and its dual $E^* = \mathcal{L}_\infty(p, \mathbb{R}^m)$, we have:

- 1 $v \in \mathcal{L}_1(\nu, E)$ denotes the settlement function, with the condition that $v(t) \in K$, $\forall t \in \mathbf{T}$. The space of settlement functions is endowed with the weak topology.
- 2 $x \in \mathcal{L}_\infty(\nu, E^*) = (\mathcal{L}_1(\nu, E))^*$ denotes the payments function, which belongs to the dual space, endowed with the weak-star topology.
- 3 $\langle \mathcal{L}_1(\nu, E), \mathcal{L}_\infty(\nu, E^*) \rangle$ is the topological dual pair denoting the duality between interbank payments and settlements with duality bracket representing the expected total outflow of interbank payments:

$$x \cdot v = \int_0^T \int_S \sum_{(i,j) \in \mathbf{B} \times \mathbf{B}} x_{ij}(t, \omega) v_{ij}(t, \omega) dp(\omega) d\nu(t)$$

Consider the Central Bank's liquidity management stochastic problem:

$$\left\{ \begin{array}{l} \text{sup}_{v \in \mathcal{L}_1(\nu, K)} \int_0^T \int_S \sum_{(i,j) \in \mathbf{B} \times \mathbf{B}} x_{ij}(t, \omega) v_{ij}(t, \omega) dp(\omega) d\nu(t) \\ \text{s.t.} \quad \sum_{\{j:(i,j) \in \mathbf{B} \times \mathbf{B}\}} x_{ij}(t, \omega) v_{ij}(t, \omega) \leq B_o^i + \bar{D}_i(t, \omega) + z_i(t, \omega) \\ \quad + \int_{[0,t)} \left[\sum_{j \in \mathbf{B}} x_{ji}(\tau, \omega) v_{ji}(\tau, \omega) - \sum_{j \in \mathbf{B}} x_{ij}(\tau, \omega) v_{ij}(\tau, \omega) \right] d\nu(\tau) \\ 0 \leq v_{ij}(t, \omega) \leq 1, \forall i \in \mathbf{B}, \forall t \in [0, T], p - a.e. \end{array} \right.$$

Here, B_o^i is the initial balance of bank i at its central bank account. It is the minimum balance that the central bank requires bank i to hold in the beginning of every day.

The function \bar{D}_i denotes the net debit cap for bank i . It is a source of liquidity offered by the central bank to bank i . The overall effect of a net debit cap is that bank i is allowed to incur

temporary intraday overdrafts up to a certain amount. Let $\bar{D} : \mathbf{T} \times S \rightarrow [0, \infty)^n$ be defined by $\bar{D}(t, \omega) = (\bar{D}_1(t, \omega), \dots, \bar{D}_n(t, \omega))$, where $\bar{D}_i(t, \omega)$ is a net debit cap for bank i at time t and in state ω .

Finally,

$$z_i(t, \omega) = \begin{cases} \sum_{\{j:(i,j) \in \mathbf{B} \times \mathbf{B}\}} z_{ij}(t, \omega), & \text{if } t < T \\ - \int_0^T y_i(\tau, \omega) d\nu(\tau) & \text{if } t = T \end{cases}$$

where $z_{ij}(t, \omega)$ is the amount of the collateralized loan received by bank i at time t and state ω in order to fulfil its payment obligations toward bank j , and $y_i(t, \omega) = \sum_{j \in \mathbf{B}} y_{ij}(t, \omega)$ is the total value of the collateral offered at time t and state ω that will be repurchased by bank i from the central bank at the end of the day, so that $z_i(T, \omega) = - \int_0^T y_i(\tau, \omega) d\nu(\tau)$, is the *repo price* to be paid at the end of the day and state ω . Let $z : \mathbf{T} \times S \rightarrow [0, \infty)^n$ be defined by $z(t, \omega) = (z_1(t, \omega), \dots, z_n(t, \omega))$.

We may even suppose that Lombard loans are prices through haircuts, which are, not in name but in essence, interest rates on intraday loans. Each Lombard loan at any time has a time-dependent interest rate. It is reasonable to assume that Lombard loans made earlier in the day have lower prices, while Lombard loans made by the end of the day have higher prices. Let $r_i(t, \omega)$ be the interest rate charged on a Lombard loan to bank i at time t and state ω . Then:

$$z_i(t, \omega) = \begin{cases} \sum_{\{j:(i,j) \in \mathbf{B} \times \mathbf{B}\}} z_{ij}(t, \omega), & \text{if } t < T \\ - \int_0^T (1 + r_i(\tau, \omega)) y_i(\tau, \omega) d\nu(\tau) & \text{if } t = T \end{cases}$$

Let $\pi = (B_o, \bar{D}, z)$ be the parameter and assume that there is a compact set $\Pi \subset [0, \infty)^n \times \mathcal{L}^\infty(\lambda, E^*)^2$ such that $\pi \in \Pi$. We call π a *systemic monetary policy*.

Define the function $g : \mathbf{T} \times S \rightarrow \mathbb{R}_+^n$ by $g(t, \omega) = (g_1(t, \omega), \dots, g_n(t, \omega))$, where $g_i(t, \omega) = B_o^i + \bar{D}_i(t, \omega) + z_i(t, \omega)$, $\forall i \in \mathbf{B}$, $\forall t \in \mathbf{T}$, p -a.e. By assumption, $g \in \mathcal{L}^\infty(\nu, E^*)$, since $(\bar{D}, z) \in \text{proj}_{2,3}\Pi \subset \mathcal{L}^\infty(\nu, E^*)^2$ and $B_o = (B_o^1, \dots, B_o^n)$ is constant. The function g describes the collection of sources of liquidity. It includes both initial reserve requirements and intraday credit policies in the form of net debit caps and Lombard loans.

Recall that $x(t, \omega) \times v(t, \omega) = \sum_{(i,j) \in \mathbf{B} \times \mathbf{B}} x_{ij}(t, \omega) v_{ij}(t, \omega)$, where $x(t, \omega)$ and $v(t, \omega)$ are matrices. Let $\text{vec}[x(t, \omega)]$ and $\text{vec}[v(t, \omega)]$ be the vectorizations of $x(t, \omega)$ and $v(t, \omega)$, respectively. With abuse of notation, but clearly without any confusion, we can still denote such vectorizations by $x(t, \omega)$ and $v(t, \omega)$, respectively, with $x(t, \omega), v(t, \omega) \in \mathbb{R}^m$. In other words, $x(t, \omega)$ would be written as $x(t, \omega) = (x_{11}(t, \omega), \dots, x_{1n}(t, \omega), \dots, x_{n1}(t, \omega), \dots, x_{nn}(t, \omega))$, and similarly for $v(t, \omega)$, i.e., $v(t, \omega) = (v_{11}(t, \omega), \dots, v_{1n}(t, \omega), \dots, v_{n1}(t, \omega), \dots, v_{nn}(t, \omega))$.

The Central Bank wants to minimize systemic liquidity:

$$\Lambda = \sum_{i \in \mathbf{B}} B_o^i - \frac{1}{\nu(\mathbf{T})} \int_0^T \int_S x(t, \omega) v(t, \omega) dp(\omega) d\nu(t)$$

The objective function can then be replaced by the total payments flow, that is:

$$F(v) = \int_0^T \int_S x(t, \omega) v(t, \omega) dp(\omega) d\nu(t)$$

Define the following matrices, $\forall i \in \mathbf{B}$, $\forall(t, \omega) \in \mathbf{T} \times S$:

$$X_i(t, \omega) = \begin{bmatrix} 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \\ x_{i1}(t, \omega) & \cdots & x_{in}(t, \omega) \\ 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{bmatrix} \leftarrow i^{th} \text{ row}$$

$$Y_i(t, \omega) = \begin{bmatrix} x_{i1}(t, \omega) & 0 & \cdots & 0 \\ 0 & x_{i2}(t, \omega) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & x_{in}(t, \omega) \end{bmatrix}_{n \times n}$$

Consider the partitioned matrices:

$$X(t, \omega) = \begin{bmatrix} X_1(t, \omega) & | & \cdots & | & X_n(t, \omega) \end{bmatrix}_{n \times n^2}$$

$$Y(t, \omega) = \begin{bmatrix} Y_1(t, \omega) & | & \cdots & | & Y_n(t, \omega) \end{bmatrix}_{n \times n^2}$$

Then the expression $\sum_{j \in \mathbf{B}} x_{ji}(\tau, \omega) v_{ji}(\tau, \omega) - \sum_{j \in \mathbf{B}} x_{ij}(\tau, \omega) v_{ij}(\tau, \omega)$, $\forall i \in \mathbf{B}$, can be written in matrix form as:

$$Y(\tau, \omega) v(\tau, \omega) - X(\tau, \omega) v(\tau, \omega) = [Y(\tau, \omega) - X(\tau, \omega)] v(\tau, \omega), \forall \tau \in \mathbf{T}, p - a.e.$$

In order to write down the dual problem properly, the constraint $v_{ij}(t, \omega) \leq 1$, $\forall t \in \mathbf{T}, p - a.e.$, $\forall i, j \in \mathbf{B}$, must be written explicitly in the feasible set. Relax the condition $v \in \mathcal{L}_1(\nu, K)$ to $v \in \mathcal{L}_1(\nu, E)$ with the added constraint $v \leq \mathbf{1}$, i.e., $v(t, \omega) \leq \mathbf{1}$, $\forall t \in \mathbf{T}, p - a.e.$, where $\mathbf{1}$ is the vector with entries equal to 1. Therefore we have:

$$\begin{cases} \sup_{v \in \mathcal{L}_1(\nu, E)} & \int_0^T \int_S x(t, \omega) v(t, \omega) dp(\omega) d\nu(t) \\ s.t. & X(t, \omega) v(t, \omega) \leq g(t, \omega) + \int_0^t [Y(\tau, \omega) - X(\tau, \omega)] v(\tau, \omega) d\nu(\tau) \\ & 0 \leq v(t, \omega) \leq \mathbf{1}, \forall t \in \mathbf{T}, p - a.e. \end{cases}$$

Define the operator $\Phi : \mathcal{L}_1(\nu, E) \rightarrow \mathcal{L}_\infty(\nu, E)$ by:

$$\Phi(v)(t, \omega) = X(t, \omega) v(t, \omega) - \int_0^t [Y(\tau, \omega) - X(\tau, \omega)] v(\tau, \omega) d\nu(\tau)$$

and let the operator $\Psi : \mathcal{L}_\infty(\nu, E^*) \rightarrow \mathcal{L}_1(\nu, E^*)$ be given by:

$$\Psi(\eta)(t, \omega) = \eta(t, \omega)X(t, \omega) - \int_t^T \eta(\tau, \omega)[Y(\tau, \omega) - X(\tau, \omega)]d\nu(\tau)$$

Notice that we can also write:

$$\Psi(\eta)(t, \omega) = X(t, \omega)^\top \eta(t, \omega) - \int_t^T [Y^\top(\tau, \omega) - X^\top(\tau, \omega)]\eta(\tau, \omega)d\nu(\tau)$$

Proposition 7 *The operators Φ and Ψ satisfy $\langle \eta, \Phi v \rangle = \langle \Psi \eta, v \rangle$, hence $\Psi = \Phi^*$, i.e., Ψ is the adjoint of Φ .*

Proof. Applying Fubini's theorem, we have:

$$\begin{aligned} \langle \eta, \Phi v \rangle &= \eta(t, \omega)X(t, \omega)v(t, \omega) - \int_0^T \eta(t, \omega)\Phi(v)(t, \omega)d\nu(t) \\ &= \eta(t, \omega)X(t, \omega)v(t, \omega) - \int_0^T \eta(t, \omega)\left\{\int_0^t [Y(\tau, \omega) - X(\tau, \omega)]v(\tau, \omega)d\nu(\tau)\right\}d\nu(t) \\ &= \eta(t, \omega)X(t, \omega)v(t, \omega) - \int_0^T \int_0^t \eta(t, \omega)[Y(\tau, \omega) - X(\tau, \omega)]v(\tau, \omega)d\nu(\tau)d\nu(t) \end{aligned}$$

We claim that:

$$\begin{aligned} &\int_0^T \int_0^t \eta(t, \omega)[Y(\tau, \omega) - X(\tau, \omega)]v(\tau, \omega)d\nu(\tau)d\nu(t) \\ &= \int_0^T \int_t^T \eta(t, \omega)[Y(\tau, \omega) - X(\tau, \omega)]d\nu(\tau)v(\tau, \omega)d\nu(t) \end{aligned}$$

Indeed, if we define $K(t, \tau, \omega) = \eta(t, \omega)[Y(\tau, \omega) - X(\tau, \omega)]v(\tau, \omega)$, it amounts to show that, for each fixed $\omega \in \cdot$:

$$\int_0^T \int_0^t K(t, \tau, \omega)d\nu(\tau)d\nu(t) = \int_0^T \int_t^T K(\tau, t, \omega)d\nu(t)d\nu(\tau)$$

It is easy to see that $K(\cdot, \cdot, \omega) \in \mathcal{L}_1([0, T] \times [0, T])$. Indeed, the components of X and Y are measurable and bounded, and v is bounded by $\mathbf{1}$. The integral on the left-hand side is equal to $\int \int_\Delta K(t, \tau, \omega)d\nu(\tau)d\nu(t)$, where $\Delta \subset [0, T] \times [0, T]$ is the triangle $\Delta = \{(t, \tau) \in [0, T] \times [0, T] : 0 \leq \tau \leq t, 0 \leq t \leq T\}$. Simple calculations⁷ show that $\int \int_\Delta K(t, \tau, \omega)d\nu(\tau)d\nu(t) = \int_0^T \int_t^T K(\tau, t, \omega)d\nu(t)d\nu(\tau)$. Thus:

$$\begin{aligned} \langle \eta, \Phi v \rangle &= \eta(t, \omega)X(t, \omega)v(t, \omega) - \int_0^T \int_0^t \eta(t, \omega)[Y(\tau, \omega) - X(\tau, \omega)]v(\tau, \omega)d\lambda(\tau)d\lambda(t) \\ &= \eta(t, \omega)X(t, \omega)v(t, \omega) - \int_0^T \int_t^T \eta(t, \omega)[Y(\tau, \omega) - X(\tau, \omega)]d\lambda(\tau)v(\tau, \omega)d\lambda(t) \\ &= \langle \Psi \eta, v \rangle \end{aligned}$$

Therefore, $\Psi = \Phi^*$, i.e., Ψ is the adjoint operator of Φ , as was to be shown. ■

Thus the paired linear programming problems in an RTGS system are given by:

⁷Solve this integral as an iterated integral in the reverse order and interchange the names of t and τ [see Schechter (1972, lemma 2.2)].

$$\begin{array}{l}
\text{Primal} \left\{ \begin{array}{l} \text{sup}_{\nu} \quad x \cdot v \\ \text{s.t.} \quad \Phi v \leq g \\ \quad \quad Iv \leq \mathbf{1} \\ \quad \quad v \geq 0 \end{array} \right. \quad \text{Dual} \left\{ \begin{array}{l} \text{inf}_{(\eta, \xi)} \quad \lambda \cdot g + \xi \cdot \mathbf{1} \\ \text{s.t.} \quad \Phi^* \lambda + I^* \xi \geq x \\ \quad \quad \eta, \xi \geq 0 \end{array} \right.
\end{array}$$

In other words, the dual problem is:

$$\left\{ \begin{array}{l} \text{inf}_{(\eta, \xi)} \quad \int_0^T \int_S \lambda(t, \omega) g(t, \omega) dp(\omega) d\nu(t) + \int_0^T \int_S \mathbf{1} \cdot \xi(t, \omega) dp(\omega) d\nu(t) \\ \text{s.t.} \quad X(t, \omega)^\top \lambda(t, \omega) + \xi(t, \omega) \geq x(t, \omega) + \int_t^T [Y^\top(t, \omega) - X^\top(t, \omega)] \lambda(\tau, \omega) d\nu(\tau) \\ \quad \quad \lambda(t, \omega), \xi(t, \omega) \geq 0, \forall t \in [0, T], p - a.e. \\ \quad \quad (\lambda, \xi) \in \mathcal{L}_\infty(\nu, K) \times \mathcal{L}_\infty(\nu) \end{array} \right.$$

Putting it in a more explicit way, the dual problem is:

$$\left\{ \begin{array}{l} \text{inf}_{(\eta, \xi)} \quad \int_0^T \int_S \sum_{i \in \mathbf{B}} \lambda_i(t, \omega) g_i(t, \omega) dp(\omega) d\nu(t) + \int_0^T \int_S \sum_{(i, j) \in \mathbf{B} \times \mathbf{B}} \xi_{ij}(t, \omega) dp(\omega) d\nu(t) \\ \text{s.t.} \quad x_{ij}(t, \omega) \lambda_i(t, \omega) + \xi_{ij}(t, \omega) \geq x_{ij}(t, \omega) + x_{ij}(t, \omega) \int_t^T (\lambda_j(\tau, \omega) - \lambda_i(\tau, \omega)) d\nu(\tau) \\ \quad \quad \lambda_i(t, \omega), \xi_{ij}(t, \omega) \geq 0, \forall i, j \in \mathbf{B}, \forall t \in [0, T], p - a.e. \\ \quad \quad (\lambda, \xi) \in \mathcal{L}_\infty(\nu, K) \times \mathcal{L}_\infty(\nu) \end{array} \right.$$

Notice that, even though the space of dual programs is $\mathcal{L}_1(\nu, K) \times \mathcal{L}_1(\nu)$, we want to find solutions in the linear subspace $\mathcal{L}_\infty(\nu, K) \times \mathcal{L}_\infty(\nu)$.

The dual constraints in the continuous case resembles the dual constraints for the discrete case, that is:

$$x_{ij}(t, \omega) \{ \lambda_i(t, \omega) + \int_t^T [\lambda_i(\tau, \omega) - \lambda_j(\tau, \omega)] d\nu(\tau) \} + \xi_{ij}(t, \omega) \geq x_{ij}(t, \omega)$$

The dual constraint above shows that the economic value of interbank payments are different from the face values of the payments. Recall that the dimension of the consistency shadow-prices is the same dimension of interbank payments, so the inequality above is well defined. The economic value of an interbank payment $x_{ij}(t, \omega)$ made at time t and state ω has to be as high as the face value of the payment. The economic value of $x_{ij}(t, \omega)$ has two components:

- $x_{ij}(t, \omega) \{ \lambda_i(t, \omega) + \int_t^T [\lambda_i(\tau, \omega) - \lambda_j(\tau, \omega)] d\nu(\tau) \}$: the face value multiplied by its bilateral net price. The bilateral net price is given by the liquidity shadow-price of bank i at time t and state ω plus the cumulated net future bilateral shadow-prices between banks i and j . This amounts to instant shadow-price plus a bilateral adjustment.
- $\xi_{ij}(t, \omega)$: the consistency shadow-price, which measures the endogenous price of splitting the payment.

If strong duality holds, the primal and dual values coincide:

$$\begin{aligned}
\mathcal{D}(B_o, \mathbf{1}) &= \int_0^T \int_S \sum_{i \in \mathbf{B}} \lambda_i(t, \omega) g_i(t, \omega) dp(\omega) d\nu(t) + \int_0^T \int_S \sum_{(i,j) \in \mathbf{B} \times \mathbf{B}} \xi_{ij}(t, \omega) dp(\omega) d\nu(t) \\
&= \int_0^T \int_S x(t, \omega) v(t, \omega) dp(\omega) d\nu(t) \\
&= \mathcal{P}(x)
\end{aligned}$$

Plugging the dual value into the definition of systemic liquidity, we get:

$$\begin{aligned}
\Lambda &= \sum_{i \in \mathbf{B}} B_o^i - \frac{1}{\lambda(\mathbf{T})} \int_0^T \int_S x(t, \omega) v(t, \omega) dp(\omega) d\nu(t) \\
&= \sum_{i \in \mathbf{B}} B_o^i - \frac{1}{\lambda(\mathbf{T})} \int_0^T \int_S \sum_{i \in \mathbf{B}} \lambda_i(t, \omega) g_i(t, \omega) dp(\omega) d\nu(t) \\
&\quad + \int_0^T \int_S \sum_{(i,j) \in \mathbf{B} \times \mathbf{B}} \xi_{ij}(t, \omega) dp(\omega) d\nu(t)
\end{aligned}$$

Since $g_i(t, \omega) = B_o^i + \bar{D}_i(t, \omega) + z_i(t, \omega)$, we have:

$$\begin{aligned}
\Lambda &= \sum_{i \in \mathbf{B}} B_o^i - \frac{1}{\nu(\mathbf{T})} \int_0^T \int_S \sum_{i \in \mathbf{B}} \lambda_i(t, \omega) (B_o^i + \bar{D}_i(t, \omega) + z_i(t, \omega)) dp(\omega) d\nu(t) \\
&\quad - \int_0^T \int_S \sum_{(i,j) \in \mathbf{B} \times \mathbf{B}} \xi_{ij}(t, \omega) dp(\omega) d\nu(t) \\
&= \sum_{i \in \mathbf{B}} B_o^i - \sum_{i \in \mathbf{B}} B_o^i \frac{1}{\nu(\mathbf{T})} \int_0^T \int_S \lambda_i(t, \omega) dp(\omega) d\nu(t) \\
&\quad - \frac{1}{\nu(\mathbf{T})} \int_0^T \int_S \sum_{i \in \mathbf{B}} \lambda_i(t, \omega) (\bar{D}_i(t, \omega) + z_i(t, \omega)) dp(\omega) d\nu(t) \\
&\quad - \sum_{(i,j) \in \mathbf{B} \times \mathbf{B}} \frac{1}{\nu(\mathbf{T})} \int_0^T \int_S \xi_{ij}(t, \omega) dp(\omega) d\nu(t)
\end{aligned}$$

The Central Bank wants to use shadow-prices to set the control variables (initial balances, net debit caps, and Lombard loans) in such a way to bring systemic liquidity down to its minimum level, $\Lambda = 0$. When systemic liquidity is zero, the design of the payment system allows interbank payments to flow smoothly without any outside money being unused any time.

Definition 8 *A systemic monetary policy $\pi \in \Pi$ is liquidity-efficient if systemic liquidity is zero, $\Lambda = 0$.*

Define the following expected averages of shadow-prices over time:

$$\begin{aligned}\bar{\lambda}_i^e &= \frac{1}{\nu(\mathbf{T})} \int_0^T \int_S \lambda_i(t, \omega) dp(\omega) d\nu(t) \\ \bar{\xi}_{ij}^e &= \frac{1}{\nu(\mathbf{T})} \int_0^T \int_S \xi_{ij}(t, \omega) dp(\omega) d\nu(t)\end{aligned}$$

Then the fundamental equation for liquidity-efficiency becomes:

$$\begin{aligned}0 &= \sum_{i \in \mathbf{B}} B_o^i (1 - \bar{\lambda}_i^e) - \sum_{(i,j) \in \mathbf{B} \times \mathbf{B}} \bar{\xi}_{ij}^e \\ &\quad - \frac{1}{\nu(\mathbf{T})} \int_0^T \int_S \sum_{i \in \mathbf{B}} \lambda_i(t, \omega) (\bar{D}_i(t, \omega) + z_i(t, \omega)) dp(\omega) d\nu(t)\end{aligned}$$

4 Policy implications

In this section we present some policy implications drawn from our dual approach.

Suppose systemic liquidity is positive, which means that some outside money is being unused:

$$\begin{aligned}0 < \Lambda &= \sum_{i \in \mathbf{B}} B_o^i (1 - \bar{\lambda}_i^e) - \sum_{(i,j) \in \mathbf{B} \times \mathbf{B}} \bar{\xi}_{ij}^e \\ &\quad - \frac{1}{\nu(\mathbf{T})} \int_0^T \int_S \sum_{i \in \mathbf{B}} \lambda_i(t, \omega) (\bar{D}_i(t, \omega) + z_i(t, \omega)) dp(\omega) d\nu(t)\end{aligned}$$

How can the Central Bank bring systemic liquidity down to zero, or as close to zero as possible? Increasing reserve requirements uniformly, *ceteris paribus*, is not the answer, though it is what Central Banks usually do. Indeed, the Central Bank should increase reserve requirements of those banks with average expected liquidity shadow-prices above unity, $1 - \bar{\lambda}_i^e < 0$. Some banks may have their reserve requirements increased by as much as everybody else, but this extra outside money may become stuck, for what these banks had were already enough for their settlement purposes. Thus any extra money becomes unused, and this represents a deadweight loss.

Suppose, for instance, that the only source of intraday liquidity, besides net transfers, is the amount of initial reserves. If systemic liquidity is positive, then:

$$0 < \Lambda = \sum_{i \in \mathbf{B}} B_o^i (1 - \bar{\lambda}_i^e) - \sum_{(i,j) \in \mathbf{B} \times \mathbf{B}} \bar{\xi}_{ij}^e$$

How could the Central Bank set reserve requirements so as to reduce systemic liquidity and make the RTGS system as close as possible to a DNS system with no systemic risk? In this simple case we are analyzing, the answer is given by the equation:

$$\sum_{i \in \mathbf{B}} B_o^{i*} (1 - \bar{\lambda}_i^e) - \sum_{(i,j) \in \mathbf{B} \times \mathbf{B}} \bar{\xi}_{ij}^e = 0$$

One possible solution to it is:

$$B_o^{i*} = \begin{cases} \frac{1}{\#\mathbf{B}_1(\bar{\lambda}_i^e - 1)} \sum_{k \in \mathbf{B}_o} \{B_o^k(1 - \bar{\lambda}_k^e) - \sum_{j \in \mathbf{B}} \bar{\xi}_{kj}^e\} - \frac{\sum_{j \in \mathbf{B}} \bar{\xi}_{ij}^e}{\bar{\lambda}_i^e - 1} & \text{if } i \in \mathbf{B}_1 \\ B_o^i & \text{if } i \in \mathbf{B}_o \end{cases}$$

Here, $\mathbf{B}_1 = \{i : \bar{\lambda}_i^e > 1\}$ is the set of banks with expected average liquidity shadow-price strictly above unity and $\mathbf{B}_o = \{i : 0 \leq \bar{\lambda}_i^e \leq 1\} = \mathbf{B} \setminus \mathbf{B}_1$ is the rest of banks. The above solution says that banks with low shadow-prices are required to keep their historical reserve balance, but banks with high shadow-prices are required to change initial reserves to B_o^{i*} .

Among other things, the Central Bank can estimate the effect of different intraday credit policies. For instance, what happens when some bank is allowed to use net debit caps? Suppose the Central Bank could reward some banks with net debit caps. Shadow-prices tell us the extent of such net debit caps. Suppose, for instance, that there are no Lombard loans and that net debit caps are constant and uniform, i.e., banks have the same net debit cap, say $\bar{D}_i(t, \omega) = D > 0$. If $\Lambda > 0$, what is the best D ? Consider the systemic liquidity before the introduction of net debit caps:

$$\Lambda = \sum_{i \in \mathbf{B}} B_o^i(1 - \bar{\lambda}_i^e) - \sum_{(i,j) \in \mathbf{B} \times \mathbf{B}} \bar{\xi}_{ij}^e > 0$$

If the Central Bank introduces a net debit cap, then it want set D so as to get:

$$\sum_{i \in \mathbf{B}} B_o^i(1 - \bar{\lambda}_i^e) - \sum_{(i,j) \in \mathbf{B} \times \mathbf{B}} \bar{\xi}_{ij}^e - D \sum_{i \in \mathbf{B}} \bar{\lambda}_i^e = 0$$

that is:

$$D^* = \frac{\sum_{i \in \mathbf{B}} B_o^i(1 - \bar{\lambda}_i^e) - \sum_{(i,j) \in \mathbf{B} \times \mathbf{B}} \bar{\xi}_{ij}^e}{\sum_{i \in \mathbf{B}} \bar{\lambda}_i^e}$$

Net debit caps can be personalized. Indeed, if we go further into our example, suppose we want to set the optimal D_i^* . Then they solve:

$$\sum_{i \in \mathbf{B}} B_o^i(1 - \bar{\lambda}_i^e) - \sum_{(i,j) \in \mathbf{B} \times \mathbf{B}} \bar{\xi}_{ij}^e - \sum_{i \in \mathbf{B}} D_i^* \bar{\lambda}_i^e = 0$$

This is equivalent to:

$$\sum_{i \in \mathbf{B}} \{B_o^i(1 - \bar{\lambda}_i^e) - \sum_{j \in \mathbf{B}} \bar{\xi}_{ij}^e - D_i^* \bar{\lambda}_i^e\} = 0$$

A possible solution is obtained by setting:

$$B_o^i(1 - \bar{\lambda}_i^e) - \sum_{j \in \mathbf{B}} \bar{\xi}_{ij}^e - D_i \bar{\lambda}_i^e = 0$$

which implies that:

$$D_i^* = \frac{B_o^i(1 - \bar{\lambda}_i^e) - \sum_{j \in \mathbf{B}} \bar{\xi}_{ij}^e}{\bar{\lambda}_i^e}$$

Notice that some banks will get net debit caps, which amount to a form of subsidy from the Central Bank, whereas other banks will have to pay taxes. Indeed, depending on the shadow-prices, D_i^* may be positive or negative. Thus net debit caps should be financed by banks themselves through redistribution of liquidity. While some banks get net debit caps, other banks get a positive lower bound on current balance during the day.

From a political point of view, it is better to give banks with low shadow-prices no debit caps and to set another level of net debit caps to banks with high shadow-prices. Of course, banks with high shadow-prices will have to fully bear the costs of other banks not paying taxes. Again, consider $\mathbf{B}_1 = \{i : \bar{\lambda}_i^e > 1\}$, the set of banks with expected average liquidity shadow-price strictly above unity and $\mathbf{B}_o = \{i : 0 \leq \bar{\lambda}_i^e \leq 1\} = \mathbf{B} \setminus \mathbf{B}_1$, the rest of banks. Then another solution is:

$$D_i^* = \begin{cases} \frac{B_o^i(1 - \bar{\lambda}_i^e) - \sum_{j \in \mathbf{B}} \bar{\xi}_{ij}^e}{\bar{\lambda}_i^e} + \frac{1}{\#\mathbf{B}_1 \bar{\lambda}_i^e} \sum_{k \in \mathbf{B}_o} \{B_o^k(1 - \bar{\lambda}_k^e) - \sum_{j \in \mathbf{B}} \bar{\xi}_{kj}^e\} & \text{if } i \in \mathbf{B}_1 \\ 0 & \text{if } i \in \mathbf{B}_o \end{cases}$$

Comparing this result with the previous one, we see that net debit caps are reduced by the amount given by the second term. The amount of reduction is exactly the amount of taxes not paid by banks with low shadow-prices divided equally among banks with high shadow-prices and weighed by the inverse of expected average liquidity shadow-prices. Thus banks with even higher shadow-prices are rewarded with a lower reduction of net debit caps, whereas banks with not so high shadow-prices are punished with a higher reduction of net debit caps.

5 Conclusion

By modelling the functioning of RTGS systems as a linear programming problem, the Central Bank can use shadow-prices to implement systemic monetary policies optimally, i.e., in such a way to bring systemic liquidity down to as close to zero as possible.

It is common practice for Central Banks to set reserve requirements to fulfil medium and long run macroeconomic policies. Our model shows that this way of setting reserve requirements may yield deadweight losses through the payment system. Reserve requirements should be personalized and priced according to the externalities that individual banks create in the payment system. Though a pricing like that is unthinkable from a political point of view, it is certainly

correct from an economic perspective. On the other hand, should the pricing of reserve requirements be out of questions for whatever reasons, the Central Bank still can use other monetary tools such as net debit caps, Lombard loans and haircuts on such Lombard loans. All these tools can be priced, so each bank would be charged differently, in such a way to reflect its economic value for the system.

The literature on payment systems has focused too much on the incentive effects that the design of payment systems have on individual banks' intraday liquidity management, but has given little attention to the optimal design of payment systems proper. Our model fills this gap out.

We believe that our framework can be used to study the reactions of individual banks to the design of payment systems in a widely different way from the standard models, which are usually based on Diamond-Dybvig economies. To be specific on this point, once the optimal settlement rule is found out, individual banks can manage intraday liquidity in the following way: given beliefs on other banks' behavior, the individual bank chooses the optimal time to send a payment message. This amounts to a simple stopping time problem. Going a step further, this shows that the Central Bank should have added incentive compatibility constraints so as to induce banks not to delay payment messages. The Central Bank should thus introduce incentive compatibility constraints to each bank into the primal problem, knowing that individual banks have beliefs about each others' action.

References

- [1] ANGELINI, P. (1998): "An analysis of competitive externalities in gross settlement systems". *Journal of Banking and Finance*, 22: 1-18.
- [2] BAGEHOT, W. (1873): *Lombard Street: a Description of the Money Market*. London: H. S. King.
- [3] BANCO CENTRAL DO BRASIL (2000): "Projeto de reestruturação do sistema de pagamentos brasileiro", Nota Técnica.
- [4] BANK FOR INTERNATIONAL SETTLEMENTS (2001): "Collateral in wholesale financial markets: recent trends, risk management and market dynamics". CGFS, March, Basel, Switzerland.
- [5] BANK FOR INTERNATIONAL SETTLEMENTS (2001): "Recommendations for securities settlement systems". CPSS-IOSCO, January, Basel, Switzerland.
- [6] BANK FOR INTERNATIONAL SETTLEMENTS (2001): "Core principles for systemically important payment systems". CPSS, January Basel, Switzerland.
- [7] BANK FOR INTERNATIONAL SETTLEMENTS (2001): "Glossary of terms used in payments and settlement systems". CPSS, January, Basel, Switzerland.

- [8] BANK FOR INTERNATIONAL SETTLEMENTS (1999): “Securities lending transactions: market development and implications”. IOSCO-CPSS, July, Basel, Switzerland.
- [9] BANK FOR INTERNATIONAL SETTLEMENTS (1999): “Implications of repo markets for central banks”. CGFS, March, Basel, Switzerland.
- [10] BANK FOR INTERNATIONAL SETTLEMENTS (1997): “Real-time gross settlement systems”. CPSS, March, Basle, Switzerland.
- [11] BANK FOR INTERNATIONAL SETTLEMENTS (1992): “Delivery versus payment in securities settlement systems”. CPSS, Basle, Switzerland.
- [12] BANK FOR INTERNATIONAL SETTLEMENTS (1990): “Report of the committee on interbank netting schemes of the central banks of the group of ten countries” (Lamfalussy report). Basle, Switzerland.
- [13] BECH, M. & K. SORAMÄKI (2001): “Gridlock resolution in interbank payment systems”. *Bank of Finland Discussion Paper 9-2001*.
- [14] BENAVIDES, T., G. ACEDO & H-K. XU (1996): “Random fixed points of set-valued operators”. *Proceedings of the American Mathematical Society*, 124(3): 831-838.
- [15] BOSSONE, B. & M. CIRASINO (2001): “The oversight of the payments systems: a framework for the development and governance of payment systems in emerging economies”. CEMLA/WORLD BANK, payments and securities clearance and settlement systems research series. July 2001.
- [16] C.H.I.P.S. (2001): *Rules and Operating Procedures*. Clearing House Service Company, January.
- [17] CLOWER, R. (1967): “A reconsideration on the microfoundations of monetary theory”, *Western Economic Journal*, 6: 1-8.
- [18] COCO, G. (2000): “On the use of collateral”, *Journal of Economic Surveys*, 14(2): 191-214.
- [19] DAVIES, G. (1996): *A History of Money from Ancient Times to the Present Day*. Rev. ed. Cardiff: University of Wales Press.
- [20] DeBANDT, O. & P. HARTMANN (2000): “Systemic risk: a survey”. European Central Bank, WP #35.
- [21] DEUTSCHE BUNDESBANK (2000): “RTGS^{plus}: the Bundesbank’s new system for individual payments”. Deutsche Bundesbank Monthly Report, June 2000: 59-71.
- [22] EMMONS, W. (1997): “Recent developments in wholesale payment systems”. Federal Reserve Bank of Saint Louis.

- [23] FREEDMAN, S. (2000): “Regulating the modern financial firm: implications of disintermediation and conglomeration”. Universität St. Gallen, Department of Economics, discussion paper #2000-21.
- [24] FREIXAS, X. & B. PARIGI (1998): “Contagion and efficiency in gross and net interbank payment systems”. *Journal of Financial Intermediation*, 7: 3-31.
- [25] FREIXAS, X., C. GIANNINI, G. HOGGARTH & F. SOUSSA (2000): “Lender of last resort: what have we learned since Bagehot?”. *Journal of Financial Services Research*, 18: 63-84.
- [26] FURFINE, C. (1999): “Interbank exposures: quantifying the risk of contagion”, *BIS working paper #70* (June).
- [27] GÜNTZER, M., D. JUNGNICHEL & M. LECLERC (1998): “Efficient algorithms for the clearing of interbank payments”, *European Journal of Operational Research*, 106: 212-219.
- [28] HELLER, D., T. NELLEN, & A. STURM (2000): “The Swiss interbank clearing system”. Mimeo, Swiss National Bank, Payment System Subsection.
- [29] HU, S. & N. PAPAGEORGIOU (1997): *Handbook of Multivalued Analysis, I: Theory*. Kluwer Academic Publishers, Dordrecht, the Netherlands.
- [30] JEVONS, W. (1875): *Money and the Mechanisms of Exchange*. A. Appleton and Company, New York, edition of 1900.
- [31] KAHN, C. & W. ROBERDS (2001): “Real-time gross settlement and the costs of immediacy”. *Journal of Monetary Economics*, 47: 299-319.
- [32] KAHN, C. & W. ROBERDS (2000): “The CLS bank: a solution to the risks of international payments settlement?”. *Federal Reserve Bank of Atlanta, WP# 2000-15*.
- [33] KAHN, C. & W. ROBERDS (1998): “Payment system settlement and bank incentives”. *Review of Financial Studies*, 11(4): 845-870.
- [34] LEINONEN, H. (1998): “Interbank funds transfer systems: liquidity needs, counterparty risks, and collateral”. *Bank of Finland discussion paper # 16/98*.
- [35] LEINONEN, H. & K. SORAMÄKI (1999): “Optimizing liquidity usage and settlement speed in payment systems”. *Bank of Finland discussion paper # 16/99*.
- [36] McANDREWS, J. & S. RAJAN (2000): “The timing and funding of FEDWIRE funds transfers”. *FRBNY Economic Policy Review*, July. Federal Reserve Bank of New York.
- [37] PAYMENTS RISK COMMITTEE (2000): “Intraday liquidity management in the evolving payment system: a study of the impact of the Euro, CLS bank, and CHIPS finality”. Report by the Intraday Liquidity management Task Force, April, FRB New York.

- [38] ROBERDS, W. (1999): “The incentive effects of settlement systems: a comparison of gross settlement, net settlement, and gross settlement with queueing”. IMES discussion paper #99-E-25, Bank of Japan.
- [39] ROCHET, J-C & J. TIROLE (1996): “Interbank lending and systemic risk”. *Journal of Money, Credit, and Banking*, 28: 733-762.
- [40] SANTOS, J. (2001): “Bank capital regulation in contemporary banking theory: a review of the literature”. *Financial Markets, Institutions, and Instruments*, 10: 41-84.
- [41] XU, H-K. (1998): “Measurability of fixed point sets of multivalued random operators”. *Journal of Mathematical Analysis and Applications*, 225: 62-72.
- [42] ZHOU, R. (2000): “Understanding intraday credit in large-value payment systems”. *Economic Perspectives*, Federal Reserve Bank of Chicago.

The ECO/UnB Working Paper Series

The Department of Economics of the University of Brasilia publishes its Working Papers Series since April 1972. On August 30, 2002 the series was renewed with the on-line publication of the papers. All Working Papers may be freely downloaded from the Department site: <http://www.unb.br/ih/eco>.

Working papers published since August 2002:

- 231 Posse de escravos e estrutura da riqueza no agreste e sertão de Pernambuco: 1777-1887. Flávio Rabelo Versiani and José Raimundo O. Vergolino, 30 August 2002, 29p.
- 232 On the natural rates of unemployment and interest: the Robertson connection. Mauro Boianovsky and John R. Presley, 30 August 2002, 34p.
- 233 Contas Nacionais e o meio ambiente: reflexões em torno de uma abordagem para o Brasil. Charles C. Mueller, 30 August 2002, 25p.
- 234 Economics of air pollution: hedonic price model and smell consequences of sewage treatment plants in urban areas. Sérgio A. Batalhone, Jorge M. Nogueira and Bernardo P. M. Mueller, 30 August 2002, 25p.
- 235 The Brazilian depression of the 80s and 90s. Mirta Bugarin, Roberto de G. Ellery Jr., Victor Gomes and Arilton Teixeira, 30 August 2002, 30p.
- 236 Informal employment in Brazil – A choice at the top and segmentation at the bottom: a quantile regression approach. Maria Tannuri-Pianto and Donald M. Pianto, 30 August 2002, 23p.
- 237 False contagion and false convergence clubs in stochastic growth theory. Stephen de Castro and Flávio Gonçalves, 30 August 2002, 20p.
- 238 Spot and contract markets in the Brazilian wholesale energy market. Paulo C. Coutinho and André Rossi de Oliveira, 30 August 2002, 19p.
- 239 Tributação da renda e do consumo no Brasil: uma abordagem macroeconômica. Valter Borges de Araújo Neto e Maria da C. S. de Sousa, 6 September 2002, 31p.
- 240 Vote splitting, reelection and electoral control: towards a unified model. Maurício S. Bugarin. 30 August 2002, 26p.

Forthcoming working papers:
(Subject to change)

- 241 Shadow-prices in payment systems. Rodrigo Peñaloza, 30 August 2002, 30p.
- 242 Welfare implications of the Brazilian social security system. Roberto de G. Ellery Jr. and Mirta N. S. Bugarin, 13 September 2002, 28p.
- 243 Os agentes econômicos em processo de integração regional – Inferências para avaliar os efeitos da ALCA. Renato Baumann and Francisco Galvão Carneiro, 20 September 2002, 33p.
- 244 Leading by example: a simple evolutionary approach. André Rossi de Oliveira and João R. O. de Faria, 27 September 2002, 24p.
- 245 The role of institutions in sustainable development. Bernardo Mueller and Charles Mueller, 4 October 2002, 23p.
- 246 Consórcios intermunicipais de saúde: uma análise à luz da teoria dos jogos. Luciana Teixeira, Maria Cristina MacDowell and Mauricio Bugarin, 11 October 2002, 30p.

- 247 Liquidity constraints and the behavior of aggregate consumption over the Brazilian business cycle. Mirta Bugarin and Roberto de G. Ellery Jr, 18 October 2002, 19p.
- 248 Pricing water and sewage services in urban areas: Evidences of low level equilibrium in a developing economy. Ricardo Coelho de Faria, Jorge M. Nogueira and Bernardo Mueller, 25 October 2002.
- 249 Wrong incentives for growth in the transition from modern slavery to labor markets: Babilon before, Babilon after. Stephen de Castro, 1 November 2002, 23p.
- 250 Vintage capital, distortions and development. Samuel Pessoa and Rafael Rob, 8 November 2002, 40p.
- 251 A note on Armstrong's optimality of exclusion property. Mauricio Bugarin, 15 November 2002, 20p.
- 252 Preços de escravos em Pernambuco no século XIX. Flávio R. Versiani and José Raimundo O. Vergolino, 22 November 2002, 20p.
- 253 A model of capital accumulation and rent seeking. Paulo Barelli and Samuel de Abreu Pessoa, 29 November 2002, 40p.
- 254 Anchors away: the cost and benefits of Brazil's devaluation. Edmund Amann and Werner Baer, 6 December 2002, 19p.
- 255 Um seguro agrícola "eficiente". Aécio S. Cunha, 13 December 2002, 65p.
- 256 Campaign contributions with swing voters. Manfred Dix and Rudy Santore, 27 December 2002, 18p.
- 257 Impostos e a História. Aécio S. Cunha, 3 January 2003, 12p.