ISSN 1518-3548

BANCO CENTRAL DO BRASIL

Working Paper Series 71

On Shadow-Prices of Banks in Real-Time Gross Settlement Systems

Rodrigo Andrés de Souza Peñaloza

April, 2003

ISSN 1518-3548
CGC 00.038.166/0001-05

Working Paper Series	Brasília	n. 71	Apr	2003	P. 1-49

Working Paper Series

Edited by:

Research Department (Depep)

(E-mail: workingpaper@bcb.gov.br)

Reproduction permitted only if source is stated as follows: Working Paper Series n. 71.

Authorized by Ilan Goldfajn (Deputy Governor for Economic Policy).

General Control of Publications

Banco Central do Brasil

Secre/Surel/Dimep

SBS – Quadra 3 – Bloco B – Edifício-Sede – M1

Caixa Postal 8.670

70074-900 Brasília – DF – Brazil

Phones: (5561) 414-3710 and 414-3567

Fax: (5561) 414-3626

E-mail: editor@bcb.gov.br

The views expressed in this work are those of the authors and do not necessarily reflect those of the Banco Central or its members.

Although these Working Papers often represent preliminary work, citation of source is required when used or reproduced.

As opiniões expressas neste trabalho são exclusivamente do(s) autor(es) e não refletem, necessariamente, a visão do Banco Central do Brasil.

Ainda que este artigo represente trabalho preliminar, citação da fonte é requerida mesmo quando reproduzido parcialmente.

Banco Central do Brasil Information Bureau

Address:	Secre/Surel/Diate
	SBS – Quadra 3 – Bloco B
	Edifício-Sede – 2º subsolo
	70074-900 Brasília – DF – Brazil
Phones:	(5561) 414 (....) 2401, 2402, 2403, 2404, 2405, 2406
DDG:	0800 992345
Fax:	(5561) 321-9453
Internet:	http://www.bcb.gov.br
E-mails:	cap.secre@bcb.gov.br
	dinfo.secre@bcb.gov.br

On Shadow-Prices of Banks in Real-Time Gross Settlement Systems*

Rodrigo Andrés de Souza Peñaloza†

Abstract

I model the functioning of real-time gross settlement systems for large-value interbank transfers as a linear programming problem in which queueing arrangements, splitting of payments, Lombard loans, and interbank credit exposures arise as primal solutions.

Then I use the dual programming problem associated with the maximization of the total flow of payments in order to determine the shadow-prices of banks in the payment system. We use these shadow-prices to set personalized intraday monetary policies such as reserve requirements, availability of Central Bank credit to temporarilly illiquid banks, extension of intraday interbank credit exposures, etc., so as to make the payment system more efficient and less costly in terms of systemic liquidity.

The dual approach shows us how to make banks correctly internalize the intraday network externalities they create in the real-time gross settlement system and provides an objective standard for the daily microprudential surveillance of the payment system.

Keywords: Central banking, systemic risk, monetary policies, payment system, linear programming, duality, shadow-prices.

JEL Classification: C61, E51, E52, E58.

*This paper is part of my Ph.D. dissertation at the University of California at Los Angeles (UCLA). I am grateful to Joe Ostroy for his encouragement and for his profound insights on this subject. I also thank David Levine (UCLA), Alberto Bennardo, and Hans Schollhammer (Anderson School, UCLA). Financial support from CAPES and the University of Brasilia is kindly acknowledged. Comments are welcome and of course errors are solely mine.

†Department of Economics, University of Brasilia (UnB). E-mail address: *penaloza@unb.br*.

1 Introduction

The 1990's witnessed a worldwide change in the design of payment systems for large-value transfers. The increase of systemic risk in deferred net settlement systems due to the increasing value of interbank transfers has been a constant concern for monetary authorities. The Bank for International Settlements has then recommended the adoption of real-time gross settlement (RTGS) systems for large-value transfers. In an RTGS system, interbank payments are settled as they are sent by its gross amount. In other words, no bank can be illiquid any time. This clearly reduces the time lag between delivery of payment messages and final settlement, hence reducing systemic risk. However, the holding of reserve money becomes a cost for banks. Indeed, since no illiquidity is allowed during the day, banks have to hold too much liquidity for settlement purposes. Any miscalculation obliges the bank to obtain liquidity from other sources, such as Lombard loans from the Central Bank. In order to facilitate the flow of payments and to reduce the opportunity cost of reserve money and the risk of gridlocks, some systems allow for queueing mechanisms and splitting of payments. A payment message that is not covered by sufficient funds when it is sent is then queued (and/or split) until the sending bank receives liquidity from other banks in the system.

The literature has focused on the influence of payment systems design on the behavior of banks [see DeBANDT & HARTMANN (2000) for a survey]. We focus instead on the Central Bank itself. Given the pattern of interbank transfers [see McANDREWS & RAJAN (2000) for the pattern of transfers and an analysis of it], what is the best design of real-time gross settlement systems? Recent research has given attention to this question but our approach is quite different [see, for instance, ANGELINI (1998) and ROBERDS (1999)]. The only literature whose approach to this problem somehow resembles ours is the one on gridlock resolution [GÜNTZER, JUNGNICKEL & LECLERC (1998), LEINONEN & SORAMÄKI (1999), and BECH & SORAMÄKI (2001)]. However, our model is much more general in that queueing arises endogenously and not as an ad hoc constraint.

In addition, the novelty of our approach is the focus on duality theory in order to determine the shadow-prices of banks and optimal intraday monetary policies.

The question that comes to the mind of everyone who is interested in central banking and specifically in the current trend towards the safety of payment systems is, How to make an RTGS system less costly in terms of liquidity? This question is actually too general and hides a series of other relevant questions, some of them dating back even to Bagehot:

- 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?
- Can an intraday interbank money market replace the Central Bank's role as a provider of intraday liquidity?
- How does the failure of an individual bank affect the overall flow of payments?

These questions are too relevant to be put aside. We want to provide a framework to answering questions as these, one that is simple enough to be applied to real world policy concerns but economically meaningful. Our answer is, *shadow-prices.*

We model the problem of optimal systemic liquidity in real-time gross settlement systems as a linear programming problem. The primal solution gives the optimal queueing arrangement and splitting of payments. The extension of intraday credit to illiquid banks is also optimally determined. The relevant contribution of our model lies however on the dual problem associated with the Central Bank's liquidity management problem. The dual solution gives the shadow-prices of banks. These shadow-prices can be used by monetary authorities to calculate the effect of personalized intraday monetary policies, such as reserve requirements, provision of Lombard loans, net debit caps, the extension of intraday interbank credit exposure, etc. In addition, shadow-prices help determine intraday monetary policies so as to bring systemic liquidity down to zero. Therefore, reserve money can be fully used for settlement purposes, avoiding waste of systemic liquidity. Section 2 presents the linear programming framework. Section 3 presents the dual program. Section 4 shows how to use shadow-prices and the no-gap theorem to determine liquidity-efficient intraday monetary policies. In section 5 we suppose that the Central Bank withdraws its role of the sole provider of intraday credit to illiquid banks and extend the model to the case of intraday interbank markets. Section 6 presents some examples and section 7 concludes the paper.

2 Real-Time gross settlement systems

Let $\mathbf{B} = \{1, ..., n\}$ be the set of participants in the RTGS payment system, which we call generically by *banks*. Let B_o^i be the initial balance of bank $i \in \mathbf{B}$ on its central bank account. Denote by $B_o = \{B_o^i : i \in \mathbf{B}\}$ the array of initial balances.

The day is divided into a large finite sequence of periods, $\mathbf{T} = \{t_o < t_1 < \cdots < t_K \equiv T\}$, where t_o denotes the beginning of the day and $t_K = T$ denotes the close of business. Whenever we say that settlement occurs at period $t = t_k$, it indeed occurs at time t_k, though a payment message said to be sent at period $t = t_k$ actually means a payment message sent within the time interval $\Delta t_k = (t_{k-1}, t_k]$. We interpret this interval as the time spent on operational procedures for settlement purposes.

Let $\{x_{ij}(\tau) : j \in \mathbf{B} \backslash \{i\}, \tau \leqslant t\}$ be the bank i's array of outgoing payments at time t, where $x_{ij}(\tau)$ is the payment message sent by bank i to the Central Bank at period τ requesting a transfer of $x_{ij}(\tau)$ units of reserve money from bank i's account to bank j's account. Notice that $x_{ii}(t) \equiv 0$, $\forall i \in \mathbf{B}$, $\forall t \in \mathbf{T}$, for self-transfers are nil by definition. Let $m = n^2$ denote the number of pairs of banks. Once we disregard self-transfers, we set $m = n(n-1)$.

According to the Bank for International Settlements [see BIS (1997), report on RTGS), the measure of intraday liquidity at time t from the individual bank's perspective is defined by its initial balance plus net receipt of transfers up to time t minus outgoing transfers at time t:

$$L^i(t) = B_o^i + \sum_{\tau<t} \left[\sum_{j\in\mathbf{B}} x_{ji}(\tau) - \sum_{j\in\mathbf{B}} x_{ij}(\tau) \right] - \sum_{j\in\mathbf{B}} x_{ij}(t)$$

Bank i is said to be illiquid at time t if $L^i(t) < 0$. If $L^i(t) \geqslant 0$, then it is said to be liquid at time t.

The aggregate net intraday liquidity at time t is given by total initial balances held on the Central Bank at the beginning of the day minus the total value of queued outgoing payments at time t, that is, $L(t) = \sum_{i\in\mathbf{B}} B_o^i - \sum_{i\in\mathbf{B}} \sum_{j\in\mathbf{B}} x_{ij}(t)$. Notice that $L(t) = \sum_{i\in\mathbf{B}} L^i(t)$. Indeed, the sum of net transfers over the set of banks is identically zero. At any time t, the aggregate net intraday liquidity level depends only on the initial balances and the current queues of outgoing payments. It does not depend on the history of queues during the day.

The problem with this definition is that, even if the aggregate net intraday liquidity is high, the smooth flow of payments *does* depend strongly on the queueing arrangement. For instance, under FIFO (*first-in first-out*), whenever balance is not enough to cover a funds transfer, all remaining queued payments get stuck. Thus, the *actual* aggregate net intraday liquidity is lower. The above definition does not take this into account. In other words, by not recognizing the role played by queueing arrangements, the standard definition of aggregate net intraday liquidity overestimates aggregate liquidity needs and hence impinges an even higher cost on individual banks. That means that the standard definition downestimates the settlement capacity of the system. It is not able to recognize that the system can work even better than imagined. We interpret the standard definition as potential aggregate net intraday liquidity. That would be the overall liquidity level should every queued payment be settled. Ours is the actual aggregate net intraday liquidity (or simply *systemic liquidity*).

In order to set a definition of aggregate net intraday liquidity that takes queueing arrangements into account, we have to introduce some notation.

Consider the payment $x_{ij}(\tau)$ from bank i to bank j at time τ. Denote by $\upsilon_{ij}(\tau, t)$ the fraction of $x_{ij}(\tau)$ that is settled at time $t \geqslant \tau$. Call it a settlement function, or simply a *settlement*. If $\upsilon_{ij}(\tau, \tau) = 1$, then payment $x_{ij}(\tau)$ is immediately settled with finality. If $\upsilon_{ij}(\tau, \tau) = 0$ and $\upsilon_{ij}(\tau, t) = 1$ for some $t > \tau$, then payment $x_{ij}(\tau)$ is queued for settlement at time t. An obvious restriction on the settlements is that, $\forall i, j \in \mathbf{B}$:

(a) $\sum_{t \geqslant \tau} \upsilon_{ij}(\tau, t) \geqslant 0$, $\forall \tau \in \mathbf{T}$

(b) $\upsilon_{ij}(\tau, t) \equiv 0$, $\forall \tau > t$, $\forall t \in \mathbf{T}$

(c) $\sum_{t \geqslant \tau} \upsilon_{ij}(\tau, t) \leqslant 1$, $\forall \tau \in \mathbf{T}$

Condition (a) says that some portion of the payment $x_{ij}(\tau)$ has to be settled by the end of the day. Condition (b) says that a payment cannot be settled if it has not been sent yet. Condition (c) says that any payment has to be at most fully settled by the end of the day.

When the RTGS system has no centralized queueing facilities, as FEDWIRE, then the following restriction is added:

(d) $\upsilon_{ij}(\tau, t) \equiv 0$, $\forall \tau < t$, $\forall t \in \mathbf{T}$

Condition (d) above says that a payment cannot be queued and has to be either rejected or settled at full at the moment it is sent. Denote by $\mathbf{V}$ the set of settlements.

There are further restrictions that could be imposed on the settlements. Let $\mathbf{S}_{ij} \subset \mathbb{R}$ be a nonempty real set. Assume that $\upsilon_{ij}(\tau, t) \in \mathbf{S}_{ij}$, $\forall i, j \in \mathbf{B}$ and $\forall \tau, t \in \mathbf{T}$. Now consider the following cases:

(i) $\mathbf{S}_{ij} = \{0, 1\}$

(ii) $\mathbf{S}_{ij} = [0, 1]$

(iii) $\mathbf{S}_{ij} = [a_{ij}, b_{ij}]$, where $a_{ij} < 0$ and $b_{ij} > 1$

In case (i), the payment x_{ij} is either settled at full (at whatever time) or not settled at all. In case (ii), the payment x_{ij} can be fractioned. That splitting of payments is an improvement over $0-1$ settlement is obvious. One of the possible causes of gridlock is the fact that payments are indivisible. For instance:

> 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)]

In case (iii) banks i and j are allowed to make intraday bilateral loans. In other words, case (iii) refers to an intraday interbank market. Of course, this institution only makes sense if the Central Bank prevents itself from being the lender of last resort. The numbers $a_{ij} < 0$ and $b_{ij} > 1$ refer to the extension of intraday bilateral debt and credit exposures, respectively. These bounds can be set by the Central Bank as regulatory devices based on whatever measure of soundness.

Obviously some payments are more time-critical than others. This is specially true about payments related to foreign exchange transactions:

> *Sequencing transfers* is a way of controlling intraday payment flows by scheduling the timing of outgoing transfers according to the supply of liquidity provided by incoming transfers. Importantly, to the extent that incoming and outgoing transfers are successfully sequenced, it could generate virtual offsetting effects on RTGS payments and hence contribute

to substantially reducing the necessary liquidity. The most common way of sequencing is to use queueing arrangements [...] (italics in the original)

Another way of sequencing transfers may involve message codes indicating the time of day that an individual outgoing transfer should be settled. (BIS report on RTGS, p. 18, italics added).

Time-critically is easily dealt with by our model. If a payment transfer sent at period τ_* has to be settled before the end of the day, say, at period t^* with $\tau_* \leqslant t^* < T$, then the constraint (c) is replaced by $\sum_{t=\tau_*}^{t^*} \upsilon_{ij}(\tau, t) \leqslant 1$. This however will make no serious difference to the nature of our model. Therefore, we will assume that no payment is time-critical in that it should be fully settled until the end of the day but not necessarily before it.

Also for simplicity we will assume that $\mathbf{S}_{ij} = \mathbf{S}$, for some $\mathbf{S}$ as in the cases considered above, $\forall i, j \in \mathbf{B}$.

We propose a modified concept of aggregate net intraday liquidity. Our concept takes the role of queueing in an explicit way. The *actual aggregate net intraday liquidity level at period t* is given by:

$$\ell(t) = \sum_{i \in \mathbf{B}} B_o^i - \sum_{i \in \mathbf{B}} \sum_{j \in \mathbf{B}} \sum_{\tau \leqslant t} x_{ij}(\tau) \upsilon_{ij}(\tau, t)$$

Contrary to the standard definition, the history of settled payments up to time t *does* matter for liquidity purposes. Unsettled payments are part of payments that have to be settled at time t, even though they were sent earlier and queued. Even in payment systems without centralized queueing facilities, rejected payments have to be re-sent later. Thus the amount $x_{ij}(\tau)$ includes both payments being sent at period τ *and* rejected payments sent before τ. The only difference is that in the later case, the Central Bank cannot distinguish between these two types of payments, whereas in the former case it can.

Notice that in payment systems without queueing, the above definition reduces to:

$$\ell(t) = \sum_{i \in \mathbf{B}} B_o^i - \sum_{i \in \mathbf{B}} \sum_{j \in \mathbf{B}} x_{ij}(t) \upsilon_{ij}(t,t)$$

We want to make a further modification to the definition above, in order to take the dynamics of management liquidity by the Central Bank into account. The *systemic liquidity* of an RTGS system is given by its total initial balances at the beginning of the day minus the total value of settleable queued outgoing payments averaged over time, that is:

$$\Lambda = \frac{1}{T} \sum_{t \in \mathbf{T}} \ell(t) = \sum_{i \in \mathbf{B}} B_o^i - \frac{1}{T} \sum_{i \in \mathbf{B}} \sum_{j \in \mathbf{B}} \sum_{t \in \mathbf{T}} \sum_{\tau \leqslant t} x_{ij}(\tau) \upsilon_{ij}(\tau, t)$$

Our definition of systemic liquidity makes explicit the role of queueing arrangements in the aggregate net intraday liquidity management by the Central Bank. The most obvious and reasonable criterion of optimization is that the optimal queue has to minimize the value of unsettled payments:

> The objective of a queueing facility is to optimize queues according to how time-critical the payments are and to evening out payment flows over time. [Leinonen (1997)]

The objective function of the Central Bank is to minimize systemic liquidity needs without violating liquidity constraints of the RTGS system. The lower the systemic liquidity, the less costly will it be for banks to hold reserves on the Central Bank in order to stick to the liquidity constraints imposed by real-time gross settlement systems.

The liquidity constraint of an RTGS system is that no bank can be illiquid any time.

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 *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.

In order to write down the liquidity constraints, it is important to decompose $\ell(t)$ into $\ell(t) = \sum_{i \in \mathbf{B}} \ell_i(t)$ by setting:

$$\ell_i(t) = B_o^i + \sum_{s<t} \sum_{\tau<s} [\sum_{j \in \mathbf{B}} x_{ji}(\tau) \upsilon_{ji}(\tau, s) - \sum_{j \in \mathbf{B}} x_{ij}(\tau) \upsilon_{ij}(\tau, s)] - \sum_{j \in \mathbf{B}} \sum_{\tau \leqslant t} x_{ij}(\tau) \upsilon_{ij}(\tau, t)$$

The amount $\ell_i(t)$ is the liquidity level of bank i at period t. Thus the liquidity constraint faced by bank i at period t is $\ell_i(t) \geqslant 0$, which means that no overdraft is allowed.

In its simplest form, liquidity constraints are given by $\ell_i(t) \geqslant 0, \forall t \in \mathbf{T}$, where:

(L1) $\ell_i(t_o) = B_o^i - \sum_{j \in \mathbf{B}} \upsilon_{ij}(t_o, t_o) x_{ij}(t_o)$

(L2) $\ell_i(t) = B_o^i + \sum_{s<t} \sum_{\tau<s} [\sum_{j \in \mathbf{B}} \upsilon_{ji}(\tau, s) x_{ji}(\tau) - \sum_{j \in \mathbf{B}} \upsilon_{ij}(\tau, s) x_{ij}(\tau)] -$

$- \sum_{j \in \mathbf{B}} \sum_{\tau \leqslant t} \upsilon_{ij}(\tau, t) x_{ij}(\tau)$, for $t \in \mathbf{T} \backslash \{t_o\}$

Thus liquidity constraints simply say that at any time:

total transfer = current transfer + (a fraction of) previously unsettled transfers

Systemic liquidity is always nonnegative, since no bank can be illiquid in an RTGS system. The best situation arises when systemic liquidity is zero, for total reserves will then be sufficient to cover total outflow throughout the day (possibly through transfers across periods, that is, through queueing and Lombard loans). We define such a situation as liquidity-efficiency.

Definition 1 *An RTGS system is said to be liquidity-efficient (or simply* Λ*-efficient) if systemic liquidity is zero:* $\Lambda = 0$.

Whenever an RTGS is not Λ-efficient, some liquidity is being unused, that is, some payments are not being settled even though there is sufficient liquidity in the system for settlement purposes. Λ-efficiency requires that total reserves be high enough to cover the average flow of payments, for $\Lambda = 0$ means $\sum_{i \in \mathbf{B}} B_o^i = \frac{1}{T} \sum_{i \in \mathbf{B}} \sum_{j \in \mathbf{B}} \sum_{t \in \mathbf{T}} \sum_{\tau \leqslant t} x_{ij}(\tau) \upsilon_{ij}(\tau, t)$.

Since the array of initial balance is given, the minimization of systemic liquidity needs is equivalent to the maximization of the cumulated flow of payments:

$$\sum_{i \in \mathbf{B}} \sum_{j \in \mathbf{B}} \sum_{t \in \mathbf{T}} \sum_{\tau \leqslant t} x_{ij}(\tau) \upsilon_{ij}(\tau, t)$$

We will show that we can always define appropriate vectors x and υ, for each different type of RTGS system, such that systemic liquidity is $x \cdot \upsilon$.

Our definition of systemic liquidity shows that it is not correct to say that a safe payment system is one with total reserves high enough to cover the total cumulated flow of interbank payments throughout the day. In order to cover total payments, it suffices that total reserves cover the average total outflow (over time). The rest of the work is done by the settlement function, that is, by the queueing and settlement arrangement.

A bank that participates in an RTGS system has four different sources of funds: balances on its Central Bank account, incoming transfers from other participants, collateralized intraday credit extensions from the Central Bank (also called Lombard loans), and interbank money market funds (such as overnight and term loans). So far we have considered only the first two sources.

If $\pi_i(t)$ is a Lombard loan given to bank i at period t, then its liquidity constraint at time t will be:

$$\sum_{j \in \mathbf{B}} \sum_{\tau \leqslant t} \upsilon_{ij}(\tau, t) x_{ij}(\tau) \leqslant B_o^i + \sum_{s < t} \sum_{\tau < s} [\sum_{j \in \mathbf{B}} \upsilon_{ji}(\tau, s) x_{ji}(\tau) - \sum_{j \in \mathbf{B}} \upsilon_{ij}(\tau, s) x_{ij}(\tau)] + \pi_i(t)$$

if $t < T$, since $\pi_i(t)$ is an extra source of liquidity, and:

$$\begin{aligned}\sum_{j\in\mathbf{B}}\sum_{\tau\leqslant T}\upsilon_{ij}(\tau,T)x_{ij}(\tau)+\pi_i(T) &\leqslant B_o^i+\\ &+\sum_{s<T}\sum_{\tau<s}[\sum_{j\in\mathbf{B}}\upsilon_{ji}(\tau,s)x_{ji}(\tau)-\sum_{j\in\mathbf{B}}\upsilon_{ij}(\tau,s)x_{ij}(\tau)]\end{aligned}$$

if $t = T$. Here $\pi_i(T) \geqslant \sum_{t<T}(1 + r_i(t))\pi_i(t)$, since in the end of the day individual banks have to pay back the loan plus some margin given by interest rates or haircuts. Thus $\pi_i(T)$ is actually a debt.

Since these are interest rates on loans made against collateral in the form of repurchase agreements, we call them *repo interest rates*. For each $t \in \mathbf{T}\backslash\{T\}$, define $r(t) = (r_1(t), ..., r_n(t))$ and let $r = \{r(t) : t \in \mathbf{T}\backslash\{T\}\}$ be the array of repo interest rates.

Assume that the Central Bank has a fixed amount $M > 0$ of money that can be lent to temporarily illiquid banks against collateral throughout the day. Let $M_i \geqslant 0$ be the total amount of money that the Central Bank is able to lend to bank $i \in \mathbf{B}$. Obviously $\sum_{i\in\mathbf{B}} M_i = M$. If we qualify π by banks and period, then we might assume that the *credit constraints* for each bank are given by:

(C1) $0 \leqslant \pi_i(t_o) \leqslant M_i$

(C2) $0 \leqslant \pi_i(t_k) \leqslant M_i - \sum_{\ell=0}^{k-1}\pi_i(t_\ell)$, $1 \leqslant k \leqslant K - 1$

(C3) $\pi_i(t_K) \geqslant \sum_{\ell=0}^{K-1}(1 + r_i(t))\pi_i(t_\ell)$.

Let $\pi = (\pi(t_o), ..., \pi(t_K)) \in \mathbb{R}^{n(K+1)}$ be the vector of Lombard loans from the Central Bank, where $\pi(t) = (\pi_1(t), ..., \pi_n(t)) \in \mathbb{R}^n$, $\forall t \in \mathbf{T}$. Denote by Π the set of vectors satisfying the credit constraints.

Consider the case $\mathbf{S} = [0, 1]$. The Central Bank's problem is to minimize systemic liquidity subject to liquidity and credit constraints. Equivalently, it seeks to maximize the total cumulated flow of interbank payments subject to those constraints.

The most general problem can be stated as follows:

$$\begin{cases} max_{\{\upsilon_{ij}(t,\tau),\pi_i(t)\}} & \sum_{i\in\mathbf{B}}\sum_{j\in\mathbf{B}}\sum_{t\in\mathbf{T}}\sum_{\tau\leqslant t} x_{ij}(\tau)\upsilon_{ij}(\tau,t) \\ s.t. & \sum_j x_{ij}(t_o)\upsilon_{ij}(t_o,t_o) \leqslant B_o^i \\ & \sum_j x_{ij}(t)\upsilon_{ij}(t,t) + \sum_{\tau<t}\sum_j x_{ij}(\tau)\upsilon_{ij}(\tau,t) \leqslant B_o^i \\ & +\sum_{\tau<t}\sum_{s\leqslant\tau}[\sum_j x_{ji}(s)\upsilon_{ji}(s,\tau) - \sum_j x_{ij}(s)\upsilon_{ij}(s,\tau)] \\ & +\alpha_t\pi_i(t),\ t\in\mathbf{T}\backslash\{t_o\} \\ & \alpha_t = \begin{cases} +1, & \forall t\in T\backslash\{t_K\} \\ -1, & t=t_K \end{cases} \\ & 0\leqslant\sum_{\tau\geqslant t}\upsilon_{ij}(t,\tau)\leqslant 1, \forall t\in\mathbf{T}, \forall i,j\in\mathbf{B} \\ & 0\leqslant\upsilon_{ij}(\tau,t)\leqslant 1 \\ & 0\leqslant\pi_i(t_o)\leqslant M_i \\ & \pi_i(t_k)\leqslant M_i - \sum_{\ell=0}^{k-1}\pi_i(t_\ell), \quad 1\leqslant k\leqslant K-1 \\ & \pi_i(t_K)\geqslant\sum_{\ell=0}^{K-1}(1+r_i(t_\ell))\pi_i(t_\ell) \\ & \pi_i(t)\geqslant 0, \forall i\in\mathbf{B}, \forall t\in\mathbf{T} \end{cases}$$

Let $0_n = (0, ..., 0)$ be the n-vector of zeroes. For any $i \in \mathbf{B}$, consider the following n-matrix:

$$X_i(t) = \begin{bmatrix} 0_n \\ \vdots \\ 0_n \\ x_i(t) \\ 0_n \\ \vdots \\ 0_n \end{bmatrix} = \begin{bmatrix} 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \\ x_{i1}(t) & \cdots & x_{in}(t) \\ 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{bmatrix}_{n\times n} \qquad \longleftarrow i^{th} row$$

Now consider the following partitioned $n \times n^2$-matrix:

$$X(t) = \left[\ X_1(t) \mid \ \cdots \ \mid X_n(t)\ \right]_{n\times n^2}$$

Consider:

$$Y_i(t) = \begin{bmatrix} x_{i1}(t) & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & x_{in}(t) \end{bmatrix}$$

What $Y_i(t)$ does is to diagonalize the vector of outgoing payments from bank i at time t. Now, define the following partitioned $n \times n^2$-matrix:

$$Y(t) = \left[\ Y_1(t) \mid \ \cdots \ \mid Y_n(t)\ \right]_{n\times n^2}$$

The matrix of coefficients of the liquidity constraints in an RTGS system with queueing is given by the following $n(K+1) \times n^2 \frac{1}{2}(K+2)(K+1)$-matrix[1]:

$$\mathbf{Q} = \begin{bmatrix} X(t_o) & \mathbf{0}_n & \mathbf{0}_n & \cdots & \mathbf{0}_n & \cdots & \mathbf{0}_n \\ X(t_o)-Y(t_o) & X(t_o) & X(t_1) & \cdots & \mathbf{0}_n & \cdots & \mathbf{0}_n \\ X(t_o)-Y(t_o) & X(t_o)-Y(t_o) & X(t_1)-Y(t_1) & \cdots & \mathbf{0}_n & \cdots & \mathbf{0}_n \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ X(t_o)-Y(t_o) & X(t_o)-Y(t_o) & X(t_1)-Y(t_1) & \cdots & \mathbf{0}_n & \cdots & \mathbf{0}_n \\ X(t_o)-Y(t_o) & X(t_o)-Y(t_o) & X(t_1)-Y(t_1) & \cdots & X(t_o) & \cdots & X(t_K) \end{bmatrix}$$

Then matrix $\mathbf{Q}$ (the matrix of interbank payments with queueing) is a block lower triangular matrix defined by $\mathbf{Q} = [Q_{\alpha\beta}]$, where each $Q_{\alpha\beta}$ is a $n \times n^2$-submatrix[2], $1 \leqslant \alpha, \beta \leqslant K+1$, and, for any $k = 0, 1, 2, ..., K$:

$$Q_{\alpha\beta} = \begin{cases} X(t_\ell) & \text{if } \alpha = k+1 \text{ and } \beta = \frac{(k+\ell+1)(k+\ell+2)}{2}, \quad 0 \leqslant \ell \leqslant k \\ X(t_\ell) - Y(t_\ell) & \text{if } \alpha \geqslant k+2 \text{ and } \beta = \frac{(k+\ell+1)(k+\ell+2)}{2}, \quad 0 \leqslant \ell \leqslant k \\ \mathbf{0}_n & \text{otherwise} \end{cases}$$

Thus, given $k = 0, 1, ..., K$, we can find, for any chosen $0 \leqslant \ell \leqslant k$, the submatrix $Q_{\alpha\beta}$, for any[3] $1 \leqslant \alpha, \beta \leqslant K+1$.

For any $i \in \mathbf{B}$ and any $t \in \mathbf{T}$, let $\upsilon_i(\tau, t) = (\upsilon_{i1}(\tau, t), ..., \upsilon_{in}(\tau, t))$ be the portion of the large-value transfers $x_i(t) = (x_{i1}(t), ..., x_{1n}(t)) \in \mathbb{R}^n$ from bank i to every other bank sent at time τ and settled at time t, $\forall t_o \leqslant \tau \leqslant t, \forall t \in \mathbf{T}$. Define:

$$\upsilon(t_k) = ((\upsilon_i(t_o, t_k))_{1 \leqslant i \leqslant n}, (\upsilon_i(t_1, t_k))_{1 \leqslant i \leqslant n}, ..., (\upsilon_i(t_k, t_k))_{1 \leqslant i \leqslant n}) \in \mathbb{R}^{n+2n+3n+...+(k+1)n}$$

Now consider the vector:

$$\upsilon = (\upsilon(t_o), ..., \upsilon(t_K)) \in \mathbb{R}^{n^2 \frac{1}{2}(K+2)(K+1)}$$

[1]The number of columns in matrix $\mathbf{Q}$ with queueing is $n^2 \frac{1}{2}(K+2)(K+1)$. Compare this with matrix A without queueing, where the number of columns is $n^2(K+1)$. This is because, with queueing, in each period, the central bank settles (fractions of) payments sent at that same period plus (fractions of) unsettled payments carried over from previous periods. Thus, if at t_o, the central bank settles n payments, then at t_1 it settles $2n$ (n from t_1 plus n from t_o), at t_2 it settles $3n$, and so on. Since there are $K+1$ periods, the number of columns with queueing is $n^2 + 2n^2 + \cdots + (K+1)n^2 = n^2 \frac{1}{2}(K+2)(K+1)$. If we disregard self-transfers, the actual dimension of $\mathbf{Q}$ is $n(K+1) \times n(n-1)\frac{1}{2}(K+2)(K+1)$.

[2]If we disregard self-transfers, the actual dimension of $Q_{\alpha\beta}$ is $n \times n(n-1)$.

[3]The number β is given by $\beta = 1 + (1+2+3+...+k+\ell) = \frac{(k+\ell+1)(k+\ell+2)}{2}$, $0 \leqslant \ell \leqslant k$.

Given $x_i(t)$, define $x(t) = (x_1(t), ..., x_n(t)) \in \mathbb{R}^{n^2}$ and set:

$$\overleftarrow{x}(t_k) = (x(t_o), ..., x(t_k)) = (x(t_\ell))_{t_o \leqslant \ell \leqslant t_k} \in \mathbb{R}^{(k+1)n^2}, 1 \leqslant k \leqslant K$$

Define:

$$x = (\overleftarrow{x}(t_o), \overleftarrow{x}(t_1), ..., \overleftarrow{x}(t_K)) \in \mathbb{R}^{n^2 \frac{1}{2}(K+2)(K+1)}$$

Let $B_o = (B_o^1, ..., B_o^n)$ be the vector of initial balances and define:

$$b = \begin{bmatrix} B_o \\ \vdots \\ B_o \end{bmatrix}_{n(K+1)}$$

Now it is easy to see that:

$$x \cdot \upsilon = \sum_{i \in \mathbf{B}} \sum_{j \in \mathbf{B}} \sum_{t \in \mathbf{T}} \sum_{\tau \leqslant t} x_{ij}(\tau) \upsilon_{ij}(\tau, t)$$

and, if no Lombard loans were available, that liquidity constraints would be represented by $\mathbf{Q}\upsilon \leqslant b$.

The following $n^2(K+1) \times n^2 \frac{1}{2}(K+2)(K+1)$-matrix is the matrix of consistency constraints[4]:

$$\mathbf{J} = \begin{bmatrix} I_{n^2} & I_{n^2} & 0 & I_{n^2} & 0 & 0 & \cdots & I_{n^2} & 0 & 0 & 0 & 0 \\ 0 & 0 & I_{n^2} & 0 & I_{n^2} & 0 & \cdots & 0 & I_{n^2} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & I_{n^2} & \cdots & 0 & 0 & I_{n^2} & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & I_{n^2} \end{bmatrix}$$

where I_{n^2} is the n^2-identity matrix[5]. A compact notation for matrix $\mathbf{J}$ is:

$$\mathbf{J} = \left[\begin{array}{c|c|c|c|c} I_{n^2} & I_{2n^2} & & I_{Kn^2} & \\ & & \cdots & & I_{(K+1)n^2} \\ 0_{Kn^2 \times n^2} & 0_{(K-1)n^2 \times 2n^2} & & 0_{n^2 \times Kn^2} & \end{array} \right]$$

Denote by $\mathbf{1}$ the $n^2 \frac{1}{2}(K+2)(K+1)$-vector of 1's. Therefore, we have $\mathbf{J}\upsilon \leqslant \mathbf{1}$.

Notice that we considered self-transfers $x_{ii}(t)$, $\forall i \in \mathbf{B}$, $\forall t \in \mathbf{T}$, for the sake of simplicity only. In practice, the dimension of the matrices above is reduced once

[4]If we disregard self-transfers, the actual dimension of $\mathbf{J}$ is $n(n-1)(K+1) \times n(n-1)\frac{1}{2}(K+2)(K+1)$.

[5]The dimension of the submatrix 0 is obvious.

we disregard self-transfers, $x_{ii}(t)$, and $\upsilon_{ii}(\tau, t)$. In this case, we just replace n^2 by $n^2 - n$, whenever such replacement is applicable.

In order to express the problem above in matrix form in the proper way, we still have to find the matrix representation of Lombard loans and credit constraints included. Define the following $n(K+1) \times n(K+1)$-matrix:

$$\mathbf{C} = \begin{bmatrix} -I_n & 0 & \cdots & 0 & 0 \\ 0 & -I_n & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & -I_n & 0 \\ 0 & 0 & \cdots & 0 & I_n \end{bmatrix}$$

Define the $n(K+1) \times n(K+1)$-matrix:

$$\mathbf{H} = \begin{bmatrix} I_n & 0 & \cdots & 0 & 0 \\ I_n & I_n & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ I_n & I_n & \cdots & I_n & 0 \\ I_n & I_n & \cdots & I_n & -I_n \end{bmatrix}$$

Matrix $\mathbf{H}$ describes the coefficients of the credit constraints if no repo rates are charged. Given the vector $M = (M_1, ..., M_n)$, consider:

$$m^o = \underbrace{(M, ..., M)}_{K\ times} \in \mathbb{R}^{nK}$$

and define $\bar{m} = (m^o, 0) \in \mathbb{R}^{n(K+1)}$, where $0 \in \mathbb{R}^n$. Then the credit constraints can be written as:

$$\mathbf{H}\pi \leqslant \bar{m} \equiv \begin{bmatrix} m^o \\ 0 \end{bmatrix}$$

Therefore, in matrix form, we can write the Central Bank's primal problem as:

$$(\mathcal{P}) \begin{cases} max_{(\upsilon,\pi)\in\mathbf{V}\times\Pi} & x \cdot \upsilon \\ s.t. & \begin{bmatrix} \mathbf{Q} & \mathbf{C} \\ \mathbf{J} & 0 \\ 0 & \mathbf{H} \end{bmatrix} \begin{bmatrix} \upsilon \\ \pi \end{bmatrix} \leqslant \begin{bmatrix} b \\ \mathbf{1} \\ \bar{m} \end{bmatrix} \\ & (\upsilon, \pi) \geqslant 0 \end{cases}$$

It is easy to check that the whole matrix of coefficients:

$$\Psi = \begin{bmatrix} \mathbf{Q} & \mathbf{C} \\ \mathbf{J} & 0 \\ 0 & \mathbf{H} \end{bmatrix}$$

has dimension $n(2+n)(K+1)\times\frac{1}{2}n(K+1)(nK+2+2n)$ and that:

$$rank(\Psi)\leqslant\begin{cases}K+1 & \text{if } n=1 \text{ and } K\geqslant 2 \text{ (trivial case)}\\ 10 & \text{if } n=2 \text{ and } k=2\\ n(n+1)(K+1) & \text{otherwise}\end{cases}$$

When credit is priced, matrix $\mathbf{H}$ requires a slight modification. Let $R(t)$ be a diagonal n-matrix defined by:

$$R(t)=\begin{bmatrix} r_1(t) & 0 & \cdots & 0\\ 0 & r_2(t) & \cdots & 0\\ \vdots & \vdots & \ddots & 0\\ 0 & 0 & \cdots & r_n(t)\end{bmatrix}$$

Define the $n(K+1)\times n(K+1)$-matrix:

$$\mathbf{R}=\begin{bmatrix} I_n & 0 & \cdots & 0 & 0\\ I_n & I_n & \cdots & 0 & 0\\ \vdots & \vdots & \ddots & \vdots & \vdots\\ I_n & I_n & \cdots & I_n & 0\\ I_n+R(t_o) & I_n+R(t_1) & \cdots & I_n+R(t_{K-1}) & -I_n\end{bmatrix}$$

The primal problem with repo rates is:

$$(\mathcal{P})\begin{cases} max_{(\upsilon,\pi)\in\mathbf{V}\times\Pi} & x\cdot\upsilon\\ s.t. & \begin{bmatrix}\mathbf{Q} & \mathbf{C}\\ \mathbf{J} & 0\\ 0 & \mathbf{R}\end{bmatrix}\begin{bmatrix}\upsilon\\ \pi\end{bmatrix}\leqslant\begin{bmatrix} b\\ \mathbf{1}\\ \bar{m}\end{bmatrix}\\ & (\upsilon,\pi)\geqslant 0\end{cases}$$

Definition 2 *A optimizing real-time gross settlement system (or simply an optimizing RTGS system) is a collection* $\{(B_o,x,\upsilon,\mathbf{S}),(\mathcal{P})\}$ *of parameters* $(B_o,x,\upsilon,\mathbf{S})$ *paired with the primal problem* $(\mathcal{P})$ *above.*

Definition 3 *Let* $\{(B_o,x,\upsilon,\mathbf{S}),(\mathcal{P})\}$ *be an optimizing RTGS system. A systemic monetary policy for* $\{(B_o,x,\upsilon,\mathbf{S}),(\mathcal{P})\}$ *is a vector* (B_o,M,π,r)*, where* $B_o=\{B_o^i : i\in\mathbf{B}\}$ *is the array of initial balances,* M *is the vector of credit lines,* π *is a vector of Lombard loans, and* r *is the vector of repo interest rates. A systemic monetary policy* (B_o^*,M^*,π^*,r^*) *is* Λ*-efficient if it makes the RTGS system* Λ*-efficient.*

3 Shadow-prices of banks

Given the matrix formulation of the primal problem, we can easily get the dual problem associated with it. The solution to the dual will give us the shadow-prices of banks in an RTGS system.

When no repo rate is charged, the dual is:

$$(\mathcal{D}) \begin{cases} min_{(\lambda,\mu,\xi)} & b \cdot \lambda + \mathbf{1} \cdot \mu + \bar{m} \cdot \xi \\ s.t. & \begin{bmatrix} \mathbf{Q}^\top & \mathbf{J}^\top & 0 \\ \mathbf{C}^\top & 0 & \mathbf{H}^\top \end{bmatrix} \begin{bmatrix} \lambda \\ \mu \\ \xi \end{bmatrix} \geqslant \begin{bmatrix} x \\ 0 \end{bmatrix} \\ & (\lambda, \mu, \xi) \geqslant 0 \end{cases}$$

where λ is the vector of shadow-prices associated with liquidity constraints, μ is the vector of shadow-prices associated with consistency constraints, and ξ is the vector of shadow-prices associated with credit constraints. Here, $\mathbf{Q}^\top$ is the transpose of $\mathbf{Q}$, and similarly for the other matrices.

From the first set of dual constraints, $\mathbf{Q}^\top\lambda + \mathbf{J}^\top\mu \geqslant x$, we get, after careful calculations, that, $\forall(i,j) \in \mathbf{B} \times \mathbf{B}$, $i \neq j$:

$$\begin{cases} x_{ij}(t_\ell)\{\lambda_i(t_k) + \sum_{\theta=k+1}^{K}[\lambda_i(t_\theta) - \lambda_j(t_\theta)]\} + \mu_{ij}(t_\ell) \geqslant x_{ij}(t_\ell), & \forall k = 0, ..., K-1, \\ & \forall \ell = 0, ..., k \\ x_{ij}(t_\ell)\lambda_i(t_K) + \mu_{ij}(t_\ell) \geqslant x_{ij}(t_\ell), & k = K, \forall \ell = 0, ..., K \end{cases}$$

Alternatively, provided $x_{ij}(t) > 0$, $\forall t \in \mathbf{T}$, the dual liquidity constraints can be written as:

$$\begin{cases} \lambda_i(t_k) + \sum_{\theta=k+1}^{K}[\lambda_i(t_\theta) - \lambda_j(t_\theta)] + \frac{\mu_{ij}(t_\ell)}{x_{ij}(t_\ell)} \geqslant 1, & \forall k = 0, ..., K-1, \\ & \forall \ell = 0, ..., k \\ \lambda_i(t_K) + \frac{\mu_{ij}(t_\ell)}{x_{ij}(t_\ell)} \geqslant 1, & k = K, \forall \ell = 0, ..., K \end{cases}$$

The dual constraints above are in a more intuitive form, specially because the shadow-price $\mu_{ij}(t_\ell)$ associated to consistency constraints needs to be interpreted more carefully. Usually the settlement function lies in the interval $[0, 1]$, for it is a percentage-type variable. Even when it takes values outside $[0, 1]$, its correct interpretation still is as a percentage-type variable. Thus (omitting the indices i and j and the arguments s and t for simplicity) a unit increase from $\upsilon = 1$ to $\upsilon = 2$

means a 100% increase, that is, a change from a payment of x to $2x$. If we want to find the effect of a dollar increase, we have to divide μ by x. Since $v_{ij}(s,t)$ is a "percentage variable", the effect of a dollar increase on the total amount $x_{ij}(s)$ to be settled upon the maximum flow of payments is $\frac{\mu^*_{ij}(s)}{x_{ij}(s)}$.

From the second set of dual constraints, $\mathbf{C}^\top\lambda + \mathbf{H}^\top\xi \geqslant 0$, we get the following inequalities:

$$\begin{cases} -\lambda_i(t_k) + \sum_{\theta=k}^{K}\xi_i(t_\theta) \geqslant 0, & \forall k = 0, ..., K-1, \forall i \in \mathbf{B} \\ \lambda_i(t_K) - \xi_i(t_K) \geqslant 0, & \text{for } k = K, \forall i \in \mathbf{B} \end{cases}$$

Hence the dual problem can be rewritten as:

$$\begin{cases} min_{(\lambda,\mu,\xi)} & \sum_{i\in\mathbf{B}}\sum_{t\in\mathbf{T}} B_o^i\lambda_i(t) + \sum_{i\in\mathbf{B}}\sum_{j\in\mathbf{B}}\sum_{t\in\mathbf{T}}\mu_{ij}(t) + \sum_{i\in\mathbf{B}}\sum_{t\in\mathbf{T}\backslash\{T\}} M_i\xi_i(t) \\ s.t. & x_{ij}(\tau)\{\lambda_i(t) + \sum_{\theta=t+1}^{T}[\lambda_i(\theta) - \lambda_j(\theta)]\} + \mu_{ij}(\tau) \geqslant x_{ij}(\tau), \forall t < T, \forall\tau \leqslant t \\ & x_{ij}(\tau)\lambda_i(T) + \mu_{ij}(\tau) \geqslant x_{ij}(\tau), \forall\tau \in \mathbf{T} \\ & -\lambda_i(t) + \sum_{\theta=t}^{T}\xi_i(\theta) \geqslant 0, \forall t < T \\ & \lambda_i(T) - \xi_i(T) \geqslant 0, \forall i, j \in \mathbf{B} \\ & (\lambda, \mu, \xi) \geqslant 0 \end{cases}$$

Consider a payment $x_{ij}(\tau)$ from bank i to bank j made at period τ. The period $t \geqslant \tau$ economic value of this payment is decomposed into two values:

- Its face value, $x_{ij}(\tau)$, multiplied by a corrective factor: bank i's current liquidity shadow-price, $\lambda_i(t)$, plus an adjusted future bilateral net price between banks i and j, $\sum_{\theta=t+1}^{T}[\lambda_i(\theta) - \lambda_j(\theta)]$.

- To all this it is added the consistency shadow-price, $\mu_{ij}(\tau)$, associated with the splitting and queueing of the payment $x_{ij}(\tau)$.

Thus the economic value of a payment varies as the day passes by. The economic value of a payment is an affine transformation of its face value. The additive factor is attached to the payment once it is sent, but the multiplicative factor (the slope) varies with time.

The objective function of the dual problem is:

$$\sum_{i\in\mathbf{B}}\sum_{t\in\mathbf{T}} B_o^i\lambda_i(t) + \sum_{i\in\mathbf{B}}\sum_{j\in\mathbf{B}}\sum_{t\in\mathbf{T}}\mu_{ij}(t) + \sum_{i\in\mathbf{B}}\sum_{t\in\mathbf{T}\backslash\{T\}} M_i\xi_i(t)$$

The first term, $\sum_{i\in\mathbf{B}}\sum_{t\in\mathbf{T}} B_o^i\lambda_i(t)$, can be rewritten as $\sum_{i\in\mathbf{B}} B_o^i(\sum_{t\in\mathbf{T}}\lambda_i(t))$. Thus, each initial balance, B_o^i, is multiplied by the cumulated shadow-price, $\sum_{t\in\mathbf{T}}\lambda_i(t)$. Initial reserves are set at the beginning of the day, i.e., before period t_o. It is not possible to change initial balances after that. Then the effect of a dollar increase on initial reserves is the sum of the effects period by period. Intuitively, a dollar change of initial reserves is equivalent to a unit change of the constant on the right hand side of the liquidity constraints period by period. Therefore, the first term reflects the effect of changes in initial reserves.

Usually, the Central Bank set initial reserves so as to conform them with medium and long-run macroeconomic monetary policies, such as inflation targets, and so on. However, for the payment system itself, it is important that monetary authorities consider the consequences of reserve requirements in the very short-run on a day-by-day basis.

If the Central Bank knew the shadow-prices associated with liquidity constraints of individual banks, it would be able to require higher or lower reserves from the right set of banks. Requiring higher reserves from banks with zero shadow-prices has no effect on the flow of payments. The only effect is the withdrawal of liquidity from the economy.

The term $\sum_{i\in\mathbf{B}}\sum_{t\in\mathbf{T}\setminus\{T\}} M_i\xi_i(t)$ can be rewritten as $\sum_{i\in\mathbf{B}} M_i(\sum_{t\in\mathbf{T}\setminus\{T\}}\xi_i(t))$. Imagine that each bank begins the day with a certain amount of money, M_i, that can be lent to it by the Central Bank. The total effect of a dollar increase of credit is given by the sum of the effects period by period. The last period is naturally excluded from this calculation for the very simple reason that at the last period no credit is extended: it is the time to pay back any Lombard loans eventually received. Thus the last term reflects the overall effect of changes in the total amount of credit that the Central Bank makes available to each bank. In other words, it measures the effect of intraday loans *from* the Central Bank *to* individual banks.

The middle term, $\sum_{i\in\mathbf{B}}\sum_{j\in\mathbf{B}}\sum_{t\in\mathbf{T}}\mu_{ij}(t)$, is the total shadow-price of consistency constraints. Notice that its dimension is the same as the dimension of payments. Imagine that a bank sends a payment at some period and that it is fully paid by the end of the day, so that, the corresponding consistency constraint is biding. What is the effect of allowing such bank to transfer an extra dollar to the receiving bank? The answer is given by the corresponding consistency shadow-price. Therefore the middle term measures the effect of intraday loans *between* individual banks.

When repo rates are charged, the dual is:

$$(\mathcal{D}) \begin{cases} min_{(\lambda,\mu,\xi)} & b\cdot\lambda + \mathbf{1}\cdot\mu + \bar{m}\cdot\xi \\ s.t. & \begin{bmatrix} \mathbf{Q}^\top & \mathbf{J}^\top & 0 \\ \mathbf{C}^\top & 0 & \mathbf{R}^\top \end{bmatrix} \begin{bmatrix} \lambda \\ \mu \\ \xi \end{bmatrix} \geqslant \begin{bmatrix} x \\ 0 \end{bmatrix} \\ & (\lambda,\mu,\xi) \geqslant 0 \end{cases}$$

Equivalently:

$$\begin{cases} min_{(\lambda,\mu,\xi)} & \sum_{i\in\mathbf{B}}\sum_{t\in\mathbf{T}} B_o^i\lambda_i(t) + \sum_{i\in\mathbf{B}}\sum_{j\in\mathbf{B}}\sum_{t\in\mathbf{T}}\mu_{ij}(t) + \sum_{i\in\mathbf{B}}\sum_{t\in\mathbf{T}\backslash\{T\}} M_i\xi_i(t) \\ s.t. & x_{ij}(\tau)\{\lambda_i(t) + \sum_{\theta=t+1}^{T}[\lambda_i(\theta)-\lambda_j(\theta)]\} + \mu_{ij}(\tau) \geqslant x_{ij}(\tau), \forall t<T, \forall\tau\leqslant t \\ & x_{ij}(\tau)\lambda_i(T) + \mu_{ij}(\tau) \geqslant x_{ij}(\tau), \forall\tau\in\mathbf{T} \\ & -\lambda_i(t) + \sum_{\theta=t}^{T-1}\xi_i(\theta) + (1+r_i(t))\xi_i(T) \geqslant 0, \forall t<T \\ & \lambda_i(T) - \xi_i(T) \geqslant 0,\ \forall i,j\in\mathbf{B} \\ & (\lambda,\mu,\xi)\geqslant 0 \end{cases}$$

4 Liquidity-efficient systemic monetary policies

Let (λ,μ,ξ) be a solution to the dual problem, that is, the set of shadow-prices. The vector λ gives the shadow-prices of initial reserve requirements. The vector μ gives the shadow-prices of queueing and splitting. Finally, the vector ξ gives the shadow-prices of credit extensions. Now consider the dual value function, i.e., the dual objective function evaluated at shadow-prices:

$$\mathcal{D}(\lambda,\mu,\xi) = \sum_{i\in\mathbf{B}}\sum_{t\in\mathbf{T}} B_o^i\lambda_i(t) + \sum_{i\in\mathbf{B}}\sum_{j\in\mathbf{B}}\sum_{t\in\mathbf{T}}\mu_{ij}(t) + \sum_{i\in\mathbf{B}}\sum_{t\in\mathbf{T}\backslash\{T\}} M_i\xi_i(t)$$

In the definition just set (with some abuse of notation), the control variables are initial reserves (the B_o^i's), the extent of intraday interbank exposures (the 1's), and intraday credit lines available to banks (the M_i's). Repo interest rates are also a control variable, though they appear only in the constraints, not in the dual value function.

Recall that the primal value function (omitting stars representing optimal primal solutions for simplicity) is:

$$\mathcal{P}(\upsilon,\pi) = \sum_{i\in\mathbf{B}}\sum_{j\in\mathbf{B}}\sum_{t\in\mathbf{T}}\sum_{\tau\leqslant t} x_{ij}(\tau)\upsilon_{ij}(\tau,t)$$

Minimum liquidity is given by:

$$\Lambda = \sum_{i\in\mathbf{B}} B_o^i - \frac{1}{T}\sum_{i\in\mathbf{B}}\sum_{j\in\mathbf{B}}\sum_{t\in\mathbf{T}}\sum_{\tau\leqslant t} x_{ij}(\tau)\upsilon_{ij}(\tau,t)$$

Λ-efficiency requires $\Lambda = 0$, i.e., $\sum_{i\in\mathbf{B}} B_o^i = \frac{1}{T}\mathcal{P}(\upsilon,\pi)$. If there is no duality gap, $\mathcal{P}(\upsilon,\pi) = \mathcal{D}(\lambda,\mu,\xi)$, that is Λ-efficiency requires $\sum_{i\in\mathbf{B}} B_o^i = \frac{1}{T}\mathcal{D}(B_o, M)$. Thus Λ-efficiency is represented by the equation:

$$\sum_{i\in\mathbf{B}} B_o^i(\bar{\lambda}_i - 1) + \sum_{i\in\mathbf{B}}\sum_{,j\in\mathbf{B}} \bar{\mu}_{ij} + \sum_{i\in\mathbf{B}} M_i(\bar{\xi}_i - \frac{1}{T}\xi_i(T)) = 0$$

where:

$$\bar{\lambda}_i = \frac{1}{T}\sum_{t\in\mathbf{T}} \lambda_i(t),$$

$$\bar{\mu}_{ij} = \frac{1}{T}\sum_{t\in\mathbf{T}} \mu_{ij}(t)$$

$$\frac{1}{T}\sum_{t\in\mathbf{T}} \xi_i(t)$$

are the average shadow-prices (throughout the day). This is the *fundamental equation for liquidity-efficiency.* It shows how reserve requirements, amounts of intraday credit made available to individual banks, and the extension of interbank exposure can be used to achieve liquidity-efficiency by means of shadow-prices. Notice that intraday interest rates on Lombard loans is also a control variable, even though it does not appear explicitly in the dual value function. It is hidden in the dual constraints, but can certainly be set by the Central Bank so as to achieve liquidity-efficiency as well.

If the Central Bank knew the shadow-prices, it would be able to set intraday monetary policies in the best possible way.

Sometimes the Central Bank decides to increase reserve requirements in order to smooth out the flow of payments or to decrease them in order to economize on unnecessary liquidity. Our model shows that such policy may be costly in terms of liquidity. All depends on shadow-prices. If a bank has zero shadow-price, every extra dollar required from it will not enhance the flow of payments. The opportunity cost of such extra reserve money will represent a deadweight loss for the economy. On the other hand, if a bank has a high shadow-price, any extra dollar required from it will enhance the flow of payments by something more than a dollar. Then it is less costly for the Central Bank to require higher reserves from the right set of banks. In the end, it will get the amount of extra reserves it wants without creating deadweight losses.

Let us first analyze a very simple situation in which no Lombard loans are needed. Usually a specific amount of initial reserves is set one day (for whatever purposes) and is not changed for some time until new initial reserves are required. If the Central Bank knew the pattern of payments and consequently the shadow-prices, then, in the simple case just mentioned, the Central Bank will want to require new reserves in such a way as to get Λ-efficiency. When the Central Bank realizes that $\Lambda > 0$, then some liquidity is not being used for settlement purposes. Our first reaction would be to change reserve requirements by $\frac{\Lambda}{n}$, for every bank. But this is not a good idea. For instance, a bank with zero shadow-price will simply loose the amount $\frac{\Lambda}{n}$, if this is an additional reserve, or the Central Bank will simply loose the same amount if this is a decrease of reserve requirements.

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:

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

This means that some reserve money is not been used for settlement purposes, that is, some payments could be settled, but are not. 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:

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

One possible solution to it is:

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

Here, $\mathbf{B}_1 = \{i : \bar{\lambda}_i > 1\}$ is the set of banks with average liquidity shadow-price strictly above unity and $\mathbf{B}_o = \{i : 0 \leqslant \bar{\lambda}_i \leqslant 1\} = \mathbf{B} \backslash \mathbf{B}_1$ 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*}.

Whenever systemic liquidity has to be reduced, the Central Bank wants to keep initial reserve levels of those banks with average shadow-prices below unity and to change reserve requirements of banks with average shadow-price above unity. The higher the shadow-price of a bank, the lower the change of reserve.

Suppose total reserves have to be decreased. Pick a bank with shadow-price above unity. Then its contribution to aggregate liquidity is its initial balance plus its average marginal contribution from liquidity constraints. What is left over if this bank were not in the system is $\Lambda - B_o^i(\bar{\lambda}_i - 1)$. This amount is split equally among all banks with average shadow-price above unity, and hence each share is weighed according to the inverse of the excess of the respective shadow-price over and above unity, i.e., $\bar{\lambda}_i - 1$. The resulting amount is the first term of the new reserve level,

$\frac{\Lambda - B_o^i(\bar{\lambda}_i - 1)}{(\bar{\lambda}_i - 1)\#\mathbf{B}_1}$. To this it is added the following. Consider the total contribution of every other bank with shadow-price above unity. This is given by their initial reserves plus marginal contribution from liquidity constraints plus marginal contributions from queued and eventually settled payments to everybody else. This amount is split among the banks with shadow-price above unity and again weighed according to the inverse of the excess of shadow-price over and above unity.

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. Let $D_i(t)$ be the extension of net debit cap that the Central Bank allows bank i to incur at time t. It can clearly be reinterpreted as a monetary amount added to initial reserve requirements at each time during the day. Shadow-prices tell us the extent of such net debit caps. Therefore the fundamental equation for liquidity-efficiency becomes:

$$\sum_{i \in \mathbf{B}} B_o^i(\bar{\lambda}_i - 1) + \sum_{i \in \mathbf{B}} \frac{1}{T} \sum_{t \in \mathbf{T}} D_i(t)\lambda_i(t) + \sum_{i \in \mathbf{B}} \sum_{,j \in \mathbf{B}} \bar{\mu}_{ij} + \sum_{i \in \mathbf{B}} M_i(\bar{\xi}_i - \frac{1}{T}\xi_i(T)) = 0$$

Suppose, for example, that there are no Lombard loans and that net debit caps are constant, i.e., banks have the same net debit cap, say $\bar{D}_i(t) = D > 0$. If $\Lambda > 0$, how can the Central Bank set D? Consider systemic liquidity before the introduction of net debit caps:

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

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

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

that is:

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

Since the solution above does not depend on the name of the bank, then either every bank will be granted the same net debit cap or the same tax.

Net debit caps can be personalized. Indeed, suppose we want to set the optimal D_i^*. Then:

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

A possible solution is:

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

Notice that some banks will get net debit caps, which amount to a form of intraday 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 for 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 $\mathbf{B}_1 = \{i : \bar{\lambda}_i > 1\}$, the set of banks with average liquidity shadow-price strictly above unity and $\mathbf{B}_o = \{i : 0 \leqslant \bar{\lambda}_i \leqslant 1\} = \mathbf{B}\backslash\mathbf{B}_1$, the rest of banks. Then another solution is:

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

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 the 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 Intraday interbank market

One of the key roles of the Central Bank is to extend credit lines to banks that become illiquid within the day. Usually such credit extension is a collateralized loan in the form of repurchase agreements: the illiquid bank sells securities to the Central Bank and commits itself to buying them back until the end of the day. This is an intraday money market in which the Central Bank is the sole seller of liquidity and every individual bank is a buyer.

The interbank market, contrariwise, functions on an overnight basis. Individual banks lend to each other from one day to the next. We are not interested in this overnight interbank market, but rather on the possibility of an intraday interbank market. Such money market only makes sense in settlement systems where the Central Bank prevents itself from extending intraday funds to illiquid banks, hence transferring all the costs of liquidity provision to the private sector. Such market did actually exist for some time in the Swiss system (SIC). The Swiss Central Bank has recently nevertheless resumed its role of seller of intraday funds, so that the small-scaled Swiss intraday interbank market had indeed a short life.

The main characteristic of an intraday interbank money market through the RTGS payment system is the bilateral credit exposure between banks. For instance, if bank i has to send x_{ij} dollars to bank j at time t, but actually sends $x_{ij} + y_{ij}$, then y_{ij} can be interpreted as a loan from bank i to bank j. A simpler formulation, and one that fits our notation perfectly, is to allow the settlement function υ to take values beyond the interval $[0, 1]$. If bank i sends a payment x_{ij} to bank j and the settlement function takes the value $\upsilon_{ij} = 1.1$, then bank i settles its payment *and* lends $0.1x_{ij}$ to bank j. If bank j has to send a payment x_{ji} but the settlement function takes the value $\upsilon_{ji} = -0.2$, then bank j does not settle its payment *and* gets a loan of $0.2x_{ji}$ from bank i. We regard the loans $0.1x_{ij}$ and $0.2x_{ji}$ as separate loans from bank i to bank j, even if they occur at the same time and the origins of the flow are not the same. Intraday interbank loans can be optimally determined provided we impose upper and lower bounds on the values taken by the settlement function and the consistency condition.

The interpretation of the bounds mentioned above is straightforward. They represent the maximum credit exposures, in percentage terms, that pairs of individual banks can take bilaterally. These bounds can be set up by the banks involved themselves or by the Central Bank. In either case, the shadow prices associated with such constraints will tell us the worth of a variation of the maximum bilateral credit exposure.

Let $E_{ij} \geqslant 0$ be the maximum credit exposure in the bilateral transaction between banks i and j. Then $-E_{ij} \leqslant \upsilon_{ij} \leqslant 1 + E_{ij}$. Decompose $\upsilon_{ij}(t,\tau)$ into $\upsilon_{ij}(t,\tau) = \upsilon_{ij}^{+}(t,\tau) - \upsilon_{ij}^{-}(t,\tau)$, where $\upsilon_{ij}^{+}(t,\tau) \geqslant 0$ is the positive part and $\upsilon_{ij}^{-}(t,\tau) \geqslant 0$, is the negative part of $\upsilon_{ij}(t,\tau)$, respectively. Then the limited credit exposure is represented by the constraint $-E_{ij} \leqslant \upsilon_{ij}^{+}(t,\tau) - \upsilon_{ij}^{-}(t,\tau) \leqslant 1 + E_{ij}$, that is:

$$\begin{cases} \upsilon_{ij}^{+}(t,\tau) - \upsilon_{ij}^{-}(t,\tau) \leqslant 1 + E_{ij} \\ -\upsilon_{ij}^{+}(t,\tau) + \upsilon_{ij}^{-}(t,\tau) \leqslant E_{ij} \end{cases}$$

If the bounds are not in percentage terms but rather in absolute value, that is, if the maximum credit exposure of bank i to bank j is set up, for instance, by a fraction of capital adequacy level or any other criterion, then we have $-e_{ij} - x_{ij} \leqslant \upsilon_{ij}x_{ij} - x_{ij} \leqslant e_{ij}$. Therefore, provided $x_{ij} > 0$:

$$\frac{-e_{ij}}{x_{ij}} \leqslant \upsilon_{ij} \leqslant 1 + \frac{e_{ij}}{x_{ij}}$$

In this case, we just have to replace E_{ij} by $\frac{e_{ij}}{x_{ij}}$. Hence there is no loss of generality in assuming that credit exposure is in percentage terms.

Therefore, the Central Bank's liquidity management problem in an RTGS system with queueing and intraday interbank money market is:

$$\begin{cases} max_{\{\upsilon_{ij}^{+}(t,\tau),\upsilon_{ij}^{-}(t,\tau)\}} & \sum_{i\in\mathbf{B}}\sum_{j\in\mathbf{B}}\sum_{t\in\mathbf{T}}\sum_{\tau\leqslant t}x_{ij}(\tau)[\upsilon_{ij}^{+}(\tau,t)-\upsilon_{ij}^{-}(\tau,t)] \\ s.t. & \sum_{j}x_{ij}(t)[\upsilon_{ij}^{+}(t,t)-\upsilon_{ij}^{i}(t,t)]+\sum_{\tau=0}^{t-1}\sum_{j}x_{ij}(\tau)[\upsilon_{ij}^{+}(\tau,t)-\upsilon_{ij}^{-}(\tau,t)]\leqslant \\ & B_{o}^{i}+\sum_{\tau=0}^{t-1}\sum_{s=0}^{\tau}\{\sum_{j}x_{ji}(s)[\upsilon_{ji}^{+}(s,\tau)-\upsilon_{ji}^{-}(s,\tau)]- \\ & -\sum_{j}x_{ij}(s)[\upsilon_{ij}^{+}(s,\tau)-\upsilon_{ij}^{-}(s,\tau)]\} \\ & \upsilon_{ij}^{+}(t,\tau)-\upsilon_{ij}^{-}(t,\tau)\leqslant 1+E_{ij} \\ & -\upsilon_{ij}^{+}(t,\tau)+\upsilon_{ij}^{-}(t,\tau)\leqslant E_{ij} \\ & \sum_{\tau\geqslant t}\upsilon_{ij}^{+}(t,\tau)-\sum_{\tau\geqslant t}\upsilon_{ij}^{-}(t,\tau)\leqslant 1 \\ & -\sum_{\tau\geqslant t}\upsilon_{ij}^{+}(t,\tau)+\sum_{\tau\geqslant t}\upsilon_{ij}^{-}(t,\tau)\leqslant 0,\forall t\in\mathbf{T},\forall i,j\in\mathbf{B} \\ & \upsilon_{ij}^{+}(t,\tau),\upsilon_{ij}^{-}(t,\tau)\geqslant 0,\forall t\in\mathbf{T},\forall\tau\in\{t,...,T\},\forall i\in\mathbf{B} \end{cases}$$

We will now find the matrix representation of the problem above. For any $i\in\mathbf{B}$ and any $t\in\mathbf{T}$, let $\upsilon_{i}^{+}(\tau,t)=(\upsilon_{i1}^{+}(\tau,t),...,\upsilon_{in}^{+}(\tau,t))$ be the positive part of the portion of the large-value transfers $x_{i}(t)=(x_{i1}(t),...,x_{1n}(t))$ from bank i to every other bank sent at time τ and settled at time t, $\forall t_{o}\leqslant\tau\leqslant t,\forall t\in\mathbf{T}$. Define:

$$\upsilon^{+}(t_{k})=((\upsilon_{i}^{+}(t_{o},t_{k}))_{1\leqslant i\leqslant n},...,(\upsilon_{i}^{+}(t_{k},t_{k}))_{1\leqslant i\leqslant n})\in\mathbb{R}^{n+2n+3n+...+(k+1)n}$$

Now consider the vector:

$$\upsilon^{+}=(\upsilon^{+}(t_{o}),...,\upsilon^{+}(t_{K}))\in\mathbb{R}^{n^{2}\frac{1}{2}(K+2)(K+1)}$$

After disregarding self-transfers, the actual dimension of the vector υ^{+} is given by $n(n-1)\frac{1}{2}\left(K+2\right)\left(K+1\right)$. Analogously, define the vector of the negative parts, υ^{-}, as:

$$\upsilon^{-}=(\upsilon^{-}(t_{o}),...,\upsilon^{-}(t_{K}))\in\mathbb{R}^{n^{2}\frac{1}{2}(K+2)(K+1)}$$

Finally, define the vector:

$$\upsilon=\begin{bmatrix}\upsilon^{+}\\ \upsilon^{-}\end{bmatrix}\in\mathbb{R}^{n^{2}(K+2)(K+1)}$$

Notice that the actual dimension of υ is much smaller, $n(n-1)\left(K+2\right)\left(K+1\right)$. Given the bounds E_{ij}, set $E_{i}=(E_{i1},...,E_{in})$ and define $\mathbf{e}=(E_{1},...,E_{n})$. Therefore the matrix representation of the primal problem is:

$$\begin{cases} max_{(\upsilon^+,\upsilon^-)} & x\cdot\upsilon^+ - x\cdot\upsilon^- \\ s.t. & \begin{bmatrix} \mathbf{Q} & -\mathbf{Q} \\ \mathbf{J} & -\mathbf{J} \\ \mathbf{I} & -\mathbf{I} \\ -\mathbf{I} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \upsilon^+ \\ \upsilon^- \end{bmatrix} \leqslant \begin{bmatrix} b \\ \mathbf{1} \\ \mathbf{1}+\mathbf{e} \\ \mathbf{e} \end{bmatrix} \\ & \upsilon^+,\upsilon^- \geqslant 0 \end{cases}$$

Hence we have four sets of constraints: liquidity constraints, consistency constraints, upper-boundedness constraints, and lower-boundedness constraints. For each of these, consider the respective vector of shadow-prices: λ,μ,ζ,η.

The dual is:

$$\begin{cases} min_{(\lambda,\mu,\zeta,\eta)} & b\cdot\lambda + \mathbf{1}\cdot\mu + (\mathbf{1}+\mathbf{e})\cdot\zeta + \mathbf{e}\cdot\eta \\ s.t. & \begin{bmatrix} \mathbf{Q}^\top & \mathbf{J}^\top & \mathbf{I} & -\mathbf{I} \end{bmatrix} \begin{bmatrix} \lambda \\ \mu \\ \zeta \\ \eta \end{bmatrix} = x \\ & (\lambda,\mu,\zeta,\eta) \geqslant 0 \end{cases}$$

The typical row of the dual constraint is:

$$\begin{cases} x_{ij}(t_\ell)\{\lambda_i(t_k) + \sum_{\theta=k+1}^{K}[\lambda_i(t_\theta) - \lambda_j(t_\theta)]\} + \mu_{ij}(t_\ell) + \zeta_{ij}(t_\ell,t_k) - \eta_{ij}(t_\ell,t_k) = x_{ij}(t_\ell), \\ \qquad \forall k = 0,1,...,K-1, \forall \ell = 0,1,...,k \\ x_{ij}(t_\ell)\lambda_i(t_K) + \mu_{ij}(t_\ell) + \zeta_{ij}(t_\ell,t_k) - \eta_{ij}(t_\ell,t_k) = x_{ij}(t_\ell), \qquad k = K, \forall \ell = 0,1,...,K \end{cases}$$

Plugging the dual value into the definition of systemic liquidity yields:

$$\begin{aligned} \Lambda \; = \; & \sum_{i\in\mathbf{B}} B_o^i - \frac{1}{T}\sum_{i\in\mathbf{B}}\sum_{t\in\mathbf{T}} B_o^i\lambda_i(t) - \frac{1}{T}\sum_{i\in\mathbf{B}}\sum_{j\in\mathbf{B}}\sum_{t\in\mathbf{T}}\mu_{ij}(t) \\ & -\frac{1}{T}\sum_{i\in\mathbf{B}}\sum_{j\in\mathbf{B}}\sum_{t\in\mathbf{T}}(1+E_{ij})\zeta_{ij}(t) - \frac{1}{T}\sum_{i\in\mathbf{B}}\sum_{j\in\mathbf{B}}\sum_{t\in\mathbf{T}}E_{ij}\eta_{ij}(t) \end{aligned}$$

Liquidity-efficiency can obviously be implemented through the equation:

$$b\cdot\mathbf{1} - b\cdot\lambda - \mathbf{1}\cdot\mu - (\mathbf{1}+\mathbf{e})\cdot\zeta - \mathbf{e}\cdot\eta = 0$$

Among other things, the shadow-prices show how to control the extent of interbank intraday exposures (the control variable $\mathbf{e}$) so as to implement liquidity-efficiency. Suppose the Central Bank cannot discriminate banks through reserve requirements, and that payments have to be settled at full by the end of the day. Then the only control of systemic liquidity is the extent of interbank intraday exposure. In order to achieve liquidity-efficiency, we have to calculate the extension of exposure. A possible solution is:

$$E_{ij}^{*} = \frac{\frac{1}{n}B_o^i(1-\bar{\lambda}_i) - \bar{\mu}_{ij} - \bar{\zeta}_{ij}}{\bar{\zeta}_{ij} - \bar{\eta}_{ij}}$$

provided $\bar{\zeta}_{ij} \neq \bar{\eta}_{ij}$, $\forall i, j \in \mathbf{B}$.

6 Examples

In this section we give some examples to illustrate our results. Consider the following RTGS system. There are two banks, $\mathbf{B} = \{1, 2\}$, and the business day is divided into three periods, $\mathbf{T} = \{t_o, t_1, t_2\}$, which we call *morning*, *noon*, and *end of the day*, respectively. The Central Bank has no uncertainty as to the amounts that banks will transfer to each other throughout the day. The pattern of transfers is represented by the following matrices:

$$x(t_o) = \begin{bmatrix} 0 & 80 \\ 120 & 0 \end{bmatrix}$$

$$x(t_1) = \begin{bmatrix} 0 & 180 \\ 120 & 0 \end{bmatrix}$$

$$x(t_2) = \begin{bmatrix} 0 & 100 \\ 120 & 0 \end{bmatrix}$$

Each entry $x_{ij}(t)$ is to be read as "bank i transfers $x_{ij}(t)$ dollars to bank j at period t". Clearly diagonal entries are set equal to zero, for banks do not do self-transfers.

Initial balances are given by $B_o^1 = 100$ and $B_o^2 = 120$. These are the reserves banks are required to hold at their Central Bank accounts at the beginning of the day.

The Central Bank has a certain amount of money, say \$130, to lend to temporarily illiquid banks. Suppose that it has $M_1 = 50$ available for lending to bank 1 and $M_2 = 80$ available for lending to bank 2.

The first situation we will analyze is the choice between (a) an RTGS system with no centralized queueing facilities, as FEDWIRE, and in which intraday credit is extended by the Central Bank as a way to smooth out liquidity problems, and (b) an RTGS system in which the Central Bank does not provide intraday liquidity but unfunded payments can be queued and fractioned. We will show that both systems yields the same amount of waste of reserve money. Then we show that (c) when the provision of intraday liquidity is shifted down to individual banks via an intraday interbank market, liquidity-efficiency is achieved and no reserve money is wasted. Finally, we illustrate (d) the setting up of personalized net debit caps.

6.1 Queueing, no Lombard loans

The choice variable is $v_{ij}(s,t)$, which represents the portion of the payment $x_{ij}(s)$ from bank i to bank j at period s that will be settled at period $t \geqslant s$.

The primal problem is:

$$\left\{\begin{array}{ll}
max & 80v_{12}(t_o,t_o) + 80v_{12}(t_o,t_1) + 80v_{12}(t_o,t_2) + 120v_{21}(t_o,t_o) + \\
 & +120v_{21}(t_o,t_1) + 120v_{21}(t_o,t_2) + 180v_{12}(t_1,t_1) + 180v_{12}(t_1,t_2) + \\
 & +120v_{21}(t_1,t_1) + 120v_{21}(t_1,t_2) + 100v_{12}(t_2,t_2) + 120v_{21}(t_2,t_2) \\
s.t. & 80v_{12}(t_o,t_o) \leqslant 100 \\
 & 120v_{21}(t_o,t_o) \leqslant 120 \\
 & 80v_{12}(t_o,t_1) + 180v_{12}(t_1,t_1) \leqslant 100 + 120v_{21}(t_o,t_o) - 80v_{12}(t_o,t_o) \\
 & 120v_{21}(t_o,t_1) + 120v_{21}(t_1,t_1) \leqslant 120 + 80v_{12}(t_o,t_o) - 120v_{21}(t_o,t_o) \\
 & 80v_{12}(t_o,t_2) + 180v_{12}(t_1,t_2) + 100v_{12}(t_2,t_2) \leqslant \\
 & \qquad \leqslant 100 + 120v_{21}(t_o,t_o) + 120v_{21}(t_o,t_1) + 120v_{21}(t_1,t_1) - \\
 & \qquad\qquad -80v_{12}(t_o,t_o) - 80v_{12}(t_o,t_1) - 180v_{12}(t_1,t_1) \\
 & 120v_{21}(t_o,t_2) + 120v_{21}(t_1,t_2) + 120v_{21}(t_2,t_2) \leqslant \\
 & \qquad \leqslant 120 + 80v_{12}(t_o,t_o) + 80v_{12}(t_o,t_1) + 180v_{12}(t_1,t_1) - \\
 & \qquad\qquad -120v_{21}(t_o,t_o) - 120v_{21}(t_o,t_1) - 120v_{21}(t_1,t_1) \\
 & v_{12}(t_o,t_o) + v_{12}(t_o,t_1) + v_{12}(t_o,t_2) \leqslant 1 \\
 & v_{21}(t_o,t_o) + v_{21}(t_o,t_1) + v_{21}(t_o,t_2) \leqslant 1 \\
 & v_{12}(t_1,t_1) + v_{12}(t_1,t_2) \leqslant 1 \\
 & v_{21}(t_1,t_1) + v_{21}(t_1,t_2) \leqslant 1 \\
 & v_{12}(t_2,t_2) \leqslant 1 \\
 & v_{21}(t_2,t_2) \leqslant 1 \\
 & v_{ij}(s,t) \geqslant 0, \, \forall t \geqslant s
\end{array}\right.$$

To understand the constraints (which we call liquidity constraints), take, for instance, the third one:

$$80\upsilon_{12}(t_o, t_1) + 180\upsilon_{12}(t_1, t_1) \leqslant 100 + 120\upsilon_{21}(t_o, t_o) - 80\upsilon_{12}(t_o, t_o)$$

At noon, t_1, bank 1 sends $x_{12}(t_1) = 180$ to bank 2. A fraction $\upsilon_{12}(t_1, t_1)$ of it will be settled at noon. Besides, a fraction $\upsilon_{12}(t_o, t_1)$ of the payment sent in the morning, $x_{12}(t_o) = 80$, is scheduled to be settled at noon. Thus, the value of outgoing payments from bank 1 to bank 2 at noon is the left-hand side of the constraint above. The right-hand side is the sum of two things: initial balance, $B_o^1 = 100$, plus net transfers received *before* noon. That is, in the morning, bank 1 sent $80\upsilon_{12}(t_o, t_o)$ to bank 2 and received $120\upsilon_{21}(t_o, t_o)$ from bank 2, so that net transfers become $120\upsilon_{21}(t_o, t_o) - 80\upsilon_{12}(t_o, t_o)$. The same reasoning applies to both banks every period.

The optimal settlement is:

$$\begin{bmatrix} \upsilon_{12}^*(t_o, t_o) \\ \upsilon_{21}^*(t_o, t_o) \\ \upsilon_{12}^*(t_o, t_1) \\ \upsilon_{21}^*(t_o, t_1) \\ \upsilon_{12}^*(t_o, t_2) \\ \upsilon_{21}^*(t_o, t_2) \\ \upsilon_{12}^*(t_1, t_1) \\ \upsilon_{21}^*(t_1, t_1) \\ \upsilon_{12}^*(t_1, t_2) \\ \upsilon_{21}^*(t_1, t_2) \\ \upsilon_{12}^*(t_2, t_2) \\ \upsilon_{21}^*(t_2, t_2) \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 7/9 \\ 2/3 \\ 0 \\ 1/3 \\ 4/5 \\ 5/6 \end{bmatrix}$$

Thus $\upsilon_{12}^*(t_o, t_o) = 1$ means that the payment $x_{12}(t_o) = 80$ sent by bank 1 to bank 2 in the morning is fully settled in the morning. Look at $\upsilon_{21}^*(t_1, t_1) = \frac{2}{3}$ and $\upsilon_{21}^*(t_1, t_2) = \frac{1}{3}$. These numbers mean that $\frac{2}{3}$ of the payment $x_{21}(t_1) = 120$ sent by bank 2 to bank 1 at noon (that is, \$80) is settled at noon, while the remaining fraction, $\frac{1}{3}$, is settled at the end of the day (that is, \$40).

The value of total outflow is \$640. Therefore, minimum liquidity is:

$$
\begin{aligned}
L &= \sum B_o^i - \frac{1}{T}\sum_i \sum_j \sum_s \sum_{t \geqslant s} x_{ij}(s) v_{ij}(s,t) \\
&= 220 - \frac{1}{3}640 \\
&= \frac{20}{3} \\
&= \$6.67
\end{aligned}
$$

Therefore, \$6.67 of reserve money remains unused. This represents a waste of liquidity. In other words, the money is in there, but is not being used.

The dual program gives us the following shadow-prices:

$$
\begin{bmatrix}
\lambda_1^*(t_o) \\
\lambda_2^*(t_o) \\
\lambda_1^*(t_1) \\
\lambda_2^*(t_1) \\
\lambda_1^*(t_2) \\
\lambda_2^*(t_2) \\
\mu_{12}^*(t_o) \\
\mu_{21}^*(t_o) \\
\mu_{12}^*(t_1) \\
\mu_{21}^*(t_1) \\
\mu_{12}^*(t_2) \\
\mu_{21}^*(t_2)
\end{bmatrix}
=
\begin{bmatrix}
0 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 80 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0
\end{bmatrix}
$$

The dual value is 640, so there is no duality gap. Morning payments are fully settled. Bank 2's noon payment is queued and fully settled by the end of the day. Bank 1's noon payment is partially settled. The unsettled amount is $\frac{1}{9} \times \$180 = \$20$. Bank 1's end of the day payment is partially settled as well. The unsettled amount is $\frac{1}{5} \times \$100 = \$20$. Bank 2's end of the day payment is partially settled and the unsettled amount is $\frac{1}{6} \times \$120 = \$20$. Thus bank 1 would have to have its reserve requirements increased by $\$y$ as given by the equation $y \times (\lambda_1^*(t_o) + \lambda_1^*(t_1) + \lambda_1^*(t_2)) = \$40$, that is, \$20. On the other hand, bank 2 would have to have its reserve requirements increased by $\$z$ as given by the equation $z \times (\lambda_2^*(t_o) + \lambda_2^*(t_1) + \lambda_2^*(t_2)) = \$20$, that is, \$10. Therefore, if bank 1's initial reserves were \$120 (instead of \$100) and bank 2's initial reserves were \$130 (instead of \$120), the system would be able to settle all payments in full through queueing and splitting.

6.2 Lombard loans, no queueing

Recall that the amounts $M_1 = 50$ and $M_2 = 80$ are available for lending to banks 1 and 2, respectively. These amounts have to be allocated optimally throughout the day to each bank. Thus if bank 1 needs \$10 in the morning, it will have only \$40 available at noon, and so on.

The primal solution is:

$$\begin{bmatrix} \upsilon_{12}^*(t_o) \\ \upsilon_{21}^*(t_o) \\ \upsilon_{12}^*(t_1) \\ \upsilon_{21}^*(t_1) \\ \upsilon_{12}^*(t_2) \\ \upsilon_{21}^*(t_2) \\ \pi_1^*(t_o) \\ \pi_2^*(t_o) \\ \pi_1^*(t_1) \\ \pi_2^*(t_1) \\ \pi_1^*(t_2) \\ \pi_2^*(t_2) \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 7/9 \\ 3/4 \\ 9/10 \\ 1 \\ 0 \\ 0 \\ 0 \\ 10 \\ 0 \\ 10 \end{bmatrix}$$

Notice that bank 2 and the Central Bank entered into a repurchase agreement at noon. Bank 2 got a Lombard loan of \$10 at noon from the Central Bank and paid it back in the end of the day.

The maximum flow of payments is \$640. Therefore, minimum liquidity is:

$$\begin{aligned} L &= \sum B_o^i - \frac{1}{T}\sum_i \sum_j \sum_s \sum_{t \geqslant s} x_{ij}(s)\upsilon_{ij}(s,t) \\ &= 220 - \frac{1}{3}640 \\ &= \frac{20}{3} \\ &= \$6.67 \end{aligned}$$

This example shows that replacing *queueing/no loans* by *loans/no queueing* did not enhance the flow of payments. It achieved the same minimum liquidity, \$6.67, that queueing allowed the system to achieve.

The dual solution is:

$$\begin{bmatrix} \lambda_1^*(t_o) \\ \lambda_2^*(t_o) \\ \lambda_1^*(t_1) \\ \lambda_2^*(t_1) \\ \lambda_1^*(t_2) \\ \lambda_2^*(t_2) \\ \mu_{12}^*(t_o) \\ \mu_{21}^*(t_o) \\ \mu_{12}^*(t_1) \\ \mu_{21}^*(t_1) \\ \mu_{12}^*(t_2) \\ \mu_{21}^*(t_2) \\ \xi_1^*(t_o) \\ \xi_2^*(t_o) \\ \xi_1^*(t_1) \\ \xi_2^*(t_1) \\ \xi_1^*(t_2) \\ \xi_2^*(t_2) \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 80 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 80 \\ 1 \end{bmatrix}$$

The dual value is 640, so there is no duality gap. The shadow-price of bank 1's credit constraint at the end of the day is $\xi_1^*(t_2) = 80$. For the system to be liquidity-efficient, systemic liquidity has to be reduced by \$6.67. Alternatively, the total flow of payments has to be increased by \$20. Thus the amount of credit available to bank 1 has to be increased by $\$w$, where w solves $w \times \xi_1^*(t_2) = \$20$, that is, $w = \$0.25$.

6.3 intraday interbank market with queueing

Consider an RTGS system with queueing, but assume that the Central Bank provides no intraday liquidity. Then banks have to rely on an interbank intraday money market. For simplicity we assume that such market is interest-rate-free, but we can easily introduce interest rates for intraday loans between banks and even get shadow-prices associated with them.

The choice variable $v_{ij}(s,t)$ is now free. However we keep the constraint $0 \leqslant \sum_{t \geqslant s} v_{ij}(s,t) \leqslant 1$, $\forall s$.

We call the constraint above the consistency constraint. If it holds with equality, it says that payments, however queued, have to be fully settled by the end of the day.

For simplicity we assume it holds with inequality, which means that a portion of payments may remain unsettled by the end of the day, leaving room for an overnight interbank market.

The solution to the primal is:

$$\begin{bmatrix} \upsilon_{12}(t_o,t_o) \\ \upsilon_{21}(t_o,t_o) \\ \upsilon_{12}(t_o,t_1) \\ \upsilon_{21}(t_o,t_1) \\ \upsilon_{12}(t_o,t_2) \\ \upsilon_{21}(t_o,t_2) \\ \upsilon_{12}(t_1,t_1) \\ \upsilon_{21}(t_1,t_1) \\ \upsilon_{12}(t_1,t_2) \\ \upsilon_{21}(t_1,t_2) \\ \upsilon_{12}(t_2,t_2) \\ \upsilon_{21}(t_2,t_2) \end{bmatrix} = \begin{bmatrix} 5/4 \\ 1 \\ 0 \\ 0 \\ -1/4 \\ 0 \\ 2/3 \\ 5/6 \\ 1/3 \\ 1/6 \\ 3/5 \\ 5/6 \end{bmatrix}$$

Bank 1 settles its morning payment, $x_{12}(t_o) = 80$, at full and lends 25% of it, $\frac{1}{4} \times 80 = \$20$, to bank 2 in the morning. Bank 2 pays its morning payment at full. Bank 2 will repay the \$20 loan in the end of the day, as can be seen form the solution $\upsilon_{12}(t_o,t_2) = -\frac{1}{4}$, i.e., bank 1 "pays" $\upsilon_{12}(t_o,t_2)x_{12}(t_o) = -\frac{1}{4} \times 80 = -\$20$ to bank 2 in the end of the day, which means that bank 1 actually receives it from bank 2.

From $\upsilon_{12}(t_1,t_1) = \frac{2}{3}$ and $\upsilon_{12}(t_1,t_2) = \frac{1}{3}$, we get that bank 1 settles $\frac{2}{3}$ of its noon payment and queues $\frac{1}{3}$ for settlement at the end of the day.

Maximum outflow is \$660, so minimum liquidity is zero, that is, the system is liquidity-efficient:

$$\begin{aligned} L &= \sum B_o^i - \frac{1}{T}\sum_i \sum_j \sum_s \sum_{t \geqslant s} x_{ij}(s)\upsilon_{ij}(s,t) \\ &= 220 - \frac{1}{3}660 \\ &= \$0 \end{aligned}$$

Therefore, given total reserves, $\sum_i B_o^i = 220$, the optimal arrangement of queueing and interbank intraday money market led the payments system to operate at its full capacity. All the reserves were used for settlement, no reserve money remained unused.

Notice that payments add up to \$720, but only \$660 were settled. This is because total reserves were \$220 (multiplying it by the number of periods, 3, we get \$660). Here there is no need to determine shadow-prices, since the system is already liquidity-efficient.

6.4 Net debit caps

Given the shadow-prices calculated in the case of queueing, we get average shadow-prices given by $\bar{\lambda}_1^* = \frac{2}{3}$, $\bar{\lambda}_2^* = 1$, $\bar{\mu}_{12}^* = \frac{80}{3}$, and $\bar{\mu}_{21}^* = 0$. Therefore optimal net debit caps are given by:

$$\begin{aligned} D_1^* &= \frac{B_o^1(1-\bar{\lambda}_1) - \bar{\mu}_{12}}{\bar{\lambda}_1} \\ &= \frac{100(1-\frac{2}{3}) - \frac{80}{3}}{\frac{2}{3}} \\ &= 10 \end{aligned}$$

$$\begin{aligned} D_2^* &= \frac{B_o^2(1-\bar{\lambda}_2) - \bar{\mu}_{21}}{\bar{\lambda}_2} \\ &= \frac{120(1-1) - 0}{1} \\ &= 0 \end{aligned}$$

Hence bank 1 should be granted with a \$10 net debit cap whereas bank 2 is not granted any net debit cap.

7 Conclusion

Our model provides a simple framework for the understanding of real-time gross settlement systems. It is well known that RTGS systems reduce systemic risk but are costly for banks in terms of liquidity. Our model shows how to minimize this cost by designing queueing arrangements, Central Bank credit lines, and splitting of payments in an optimal way. The main contribution of our model is the focus on the dual problem. If the Central Bank knew the shadow-prices of banks, it could set liquidity-efficient intraday monetary policies by controlling reserve requirements, intraday interest rates, credit lines, extension of intraday interbank exposures, etc. The fundamental equation for liquidity-efficiency tells us that such policies should be personalized. Each bank would be priced according to its marginal contribution to the total outflow of payments.

Of course in real life the Central Bank does not know *ex ante* the pattern of interbank transfers. After business hours, however, it has the complete data set of interbank transfers during that day. Suppose it uses such data to run the primal problem and to calculate the collection of individual daily shadow-prices. The primal solution will tell the Central Banker how intraday monetary policies should have been set in the first place in order to make the payment system liquidity-efficient that day. The dual solution will give the Central Banker an objective policy standard to pricing banks according to their marginal contribution to the overall stability (or instability) of the payment system. Shadow-prices can be calculated on a day-by-day basis. Besides, our model shows that the relevant shadow-price of an individual bank is the average of its shadow-prices within the day. After collecting a long series of individual daily shadow-prices, the means of individual empirical distributions will give a good measure of the true shadow-prices of individual banks. It is these mean shadow-prices that can be used by monetary authorities to regulate payment systems efficiently. Knowing the initial reserve shadow-price of a bank is close to zero, it makes no sense for the Central Bank to increase its reserve requirements. Knowing that bilateral exposure shadow-prices of a bank is well above unity, it is better for the Central Bank to fetter its bilateral exposures in the intraday interbank market. Our model therefore provides a solid foundation for the *microprudential surveillance* of payment systems[6].

Future tasks include the extension of the model to a many-days environment in which overnight interbank loans would play a prominent role, thus bridging the gap between intraday monetary policy and medium and long-run macroeconomic monetary policy. Our model can be used to determine shadow-prices of banks in

[6] The stochastic version of the model with continuous time is also available.

real world payment systems, provided data on intraday interbank transfers are made available. Indeed, as a by-product, our model shows that the available data on intraday interbank transfers should be disaggregated.

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] BANK FOR INTERNATIONAL SETTLEMENTS (2001): “Collateral in wholesale financial markets: recent trends, risk management and market dynamics”. CGFS, March, Basel, Switzerland.

[4] BANK FOR INTERNATIONAL SETTLEMENTS (2001): “Recommendations for securities settlement systems”. CPSS-IOSCO, January, Basel, Switzerland.

[5] BANK FOR INTERNATIONAL SETTLEMENTS (2001): “Core principles for systemically important payment systems”. CPSS, January Basel, Switzerland.

[6] BANK FOR INTERNATIONAL SETTLEMENTS (1999): “Implications of repo markets for central banks”. CGFS, March, Basel, Switzerland.

[7] BANK FOR INTERNATIONAL SETTLEMENTS (1997): “Real-time gross settlement systems”. CPSS, March, Basle, Switzerland.

[8] 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.

[9] BECH, M. & K. SORAMÄKI (2001): “Gridlock resolution in interbank payment systems”. *Bank of Finland Discussion Paper* 9-2001.

[10] C.H.I.P.S. (2001): *Rules and Operating Procedures*. Clearing House Service Company, January.

[11] CLOWER, R. (1967): "A reconsideration on the microfoundations of monetary theory". *Western Economic Journal*, 6: 1-8.

[12] DeBANDT, O. & P. HARTMANN (2000): "Systemic risk: a survey". European Central Bank, WP #35.

[13] DEUTSCHE BUNDESBANK (2000): "RTGS$^{\text{plus}}$: the Bundesbank's new system for individual payments". Deutsche Bundesbank Monthly Report, June 2000: 59-71.

[14] FREIXAS, X. & B. PARIGI (1998): "Contagion and efficiency in gross and net interbank payment systems". Journal of Financial Intermediation, 7: 3-31.

[15] 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.

[16] GÜNTZER, M., D. JUNGNICKEL & M. LECLERC (1998): "Efficient algorithms for the clearing of interbank payments", *European Journal of Operational Research*, 106: 212-219.

[17] HELLER, D., T. NELLEN, & A. STURM (2000): "The Swiss interbank clearing system". Mimeo, Swiss National Bank, Payment System Subsection.

[18] KAHN, C. & W. ROBERDS (2001): "Real-time gross settlement and the costs of immediacy". *Journal of Monetary Economics*, 47: 299-319.

[19] 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*.

[20] KAHN, C. & W. ROBERDS (1998): "Payment system settlement and bank incentives". *Review of Financial Studies*, 11(4): 845-870.

[21] LEINONEN, H. (1998): “Interbank funds transfer systems: liquidity needs, counterparty risks, and collateral”. *Bank of Finland discussion paper # 16/98.*

[22] LEINONEN, H. & K. SORAMÄKI (1999): “Optimizing liquidity usage and settlement speed in payment systems”. *Bank of Finland discussion paper # 16/99.*

[23] McANDREWS, J. & S. RAJAN (2000): “The timing and funding of FEDWIRE funds transfers”. F*RBNY Economic Policy Review,* July. Federal Reserve Bank of New York.

[24] 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.

[25] ROCHET, J-C & J. TIROLE (1996): “Interbank lending and systemic risk”. *Journal of Money, Credit, and Banking*, 28: 733-762.

[26] ZHOU, R. (2000): “Understanding intraday credit in large-value payment systems”. *Economic Perspectives*, Federal Reserve Bank of Chicago.

Banco Central do Brasil

Trabalhos para Discussão

Os Trabalhos para Discussão podem ser acessados na internet, no formato PDF,
no endereço: *http://www.bc.gov.br*

Working Paper Series

Working Papers in PDF format can be downloaded from: http://www.bc.gov.br

1	**Implementing Inflation Targeting in Brazil** *Joel Bogdanski, Alexandre Antonio Tombini and Sérgio Ribeiro da Costa Werlang*	Jul/2000
2	**Política Monetária e Supervisão do Sistema Financeiro Nacional no Banco Central do Brasil** *Eduardo Lundberg*	Jul/2000
	Monetary Policy and Banking Supervision Functions on the Central Bank *Eduardo Lundberg*	Jul/2000
3	**Private Sector Participation: a Theoretical Justification of the Brazilian Position** *Sérgio Ribeiro da Costa Werlang*	Jul/2000
4	**An Information Theory Approach to the Aggregation of Log-Linear Models** *Pedro H. Albuquerque*	Jul/2000
5	**The Pass-Through from Depreciation to Inflation: a Panel Study** *Ilan Goldfajn and Sérgio Ribeiro da Costa Werlang*	Jul/2000
6	**Optimal Interest Rate Rules in Inflation Targeting Frameworks** *José Alvaro Rodrigues Neto, Fabio Araújo and Marta Baltar J. Moreira*	Jul/2000
7	**Leading Indicators of Inflation for Brazil** *Marcelle Chauvet*	Sep/2000
8	**The Correlation Matrix of the Brazilian Central Bank's Standard Model for Interest Rate Market Risk** *José Alvaro Rodrigues Neto*	Sep/2000
9	**Estimating Exchange Market Pressure and Intervention Activity** *Emanuel-Werner Kohlscheen*	Nov/2000
10	**Análise do Financiamento Externo a uma Pequena Economia Aplicação da Teoria do Prêmio Monetário ao Caso Brasileiro: 1991–1998** *Carlos Hamilton Vasconcelos Araújo e Renato Galvão Flôres Júnior*	Mar/2001
11	**A Note on the Efficient Estimation of Inflation in Brazil** *Michael F. Bryan and Stephen G. Cecchetti*	Mar/2001
12	**A Test of Competition in Brazilian Banking** *Márcio I. Nakane*	Mar/2001

13	**Modelos de Previsão de Insolvência Bancária no Brasil** *Marcio Magalhães Janot*	Mar/2001
14	**Evaluating Core Inflation Measures for Brazil** *Francisco Marcos Rodrigues Figueiredo*	Mar/2001
15	**Is It Worth Tracking Dollar/Real Implied Volatility?** *Sandro Canesso de Andrade and Benjamin Miranda Tabak*	Mar/2001
16	**Avaliação das Projeções do Modelo Estrutural do Banco Central do Brasil para a Taxa de Variação do IPCA** *Sergio Afonso Lago Alves*	Mar/2001
	Evaluation of the Central Bank of Brazil Structural Model's Inflation Forecasts in an Inflation Targeting Framework *Sergio Afonso Lago Alves*	Jul/2001
17	**Estimando o Produto Potencial Brasileiro: uma Abordagem de Função de Produção** *Tito Nícias Teixeira da Silva Filho*	Abr/2001
	Estimating Brazilian Potential Output: a Production Function Approach *Tito Nícias Teixeira da Silva Filho*	Aug/2002
18	**A Simple Model for Inflation Targeting in Brazil** *Paulo Springer de Freitas and Marcelo Kfoury Muinhos*	Apr/2001
19	**Uncovered Interest Parity with Fundamentals: a Brazilian Exchange Rate Forecast Model** *Marcelo Kfoury Muinhos, Paulo Springer de Freitas and Fabio Araújo*	May/2001
20	**Credit Channel without the LM Curve** *Victorio Y. T. Chu and Márcio I. Nakane*	May/2001
21	**Os Impactos Econômicos da CPMF: Teoria e Evidência** *Pedro H. Albuquerque*	Jun/2001
22	**Decentralized Portfolio Management** *Paulo Coutinho and Benjamin Miranda Tabak*	Jun/2001
23	**Os Efeitos da CPMF sobre a Intermediação Financeira** *Sérgio Mikio Koyama e Márcio I. Nakane*	Jul/2001
24	**Inflation Targeting in Brazil: Shocks, Backward-Looking Prices, and IMF Conditionality** *Joel Bogdanski, Paulo Springer de Freitas, Ilan Goldfajn and Alexandre Antonio Tombini*	Aug/2001
25	**Inflation Targeting in Brazil: Reviewing Two Years of Monetary Policy 1999/00** *Pedro Fachada*	Aug/2001
26	**Inflation Targeting in an Open Financially Integrated Emerging Economy: the Case of Brazil** *Marcelo Kfoury Muinhos*	Aug/2001

27	**Complementaridade e Fungibilidade dos Fluxos de Capitais Internacionais** *Carlos Hamilton Vasconcelos Araújo e Renato Galvão Flôres Júnior*	Set/2001
28	**Regras Monetárias e Dinâmica Macroeconômica no Brasil: uma Abordagem de Expectativas Racionais** *Marco Antonio Bonomo e Ricardo D. Brito*	Nov/2001
29	**Using a Money Demand Model to Evaluate Monetary Policies in Brazil** *Pedro H. Albuquerque and Solange Gouvêa*	Nov/2001
30	**Testing the Expectations Hypothesis in the Brazilian Term Structure of Interest Rates** *Benjamin Miranda Tabak and Sandro Canesso de Andrade*	Nov/2001
31	**Algumas Considerações sobre a Sazonalidade no IPCA** *Francisco Marcos R. Figueiredo e Roberta Blass Staub*	Nov/2001
32	**Crises Cambiais e Ataques Especulativos no Brasil** *Mauro Costa Miranda*	Nov/2001
33	**Monetary Policy and Inflation in Brazil (1975-2000): a VAR Estimation** *André Minella*	Nov/2001
34	**Constrained Discretion and Collective Action Problems: Reflections on the Resolution of International Financial Crises** *Arminio Fraga and Daniel Luiz Gleizer*	Nov/2001
35	**Uma Definição Operacional de Estabilidade de Preços** *Tito Nícias Teixeira da Silva Filho*	Dez/2001
36	**Can Emerging Markets Float? Should They Inflation Target?** *Barry Eichengreen*	Feb/2002
37	**Monetary Policy in Brazil: Remarks on the Inflation Targeting Regime, Public Debt Management and Open Market Operations** *Luiz Fernando Figueiredo, Pedro Fachada and Sérgio Goldenstein*	Mar/2002
38	**Volatilidade Implícita e Antecipação de Eventos de *Stress*: um Teste para o Mercado Brasileiro** *Frederico Pechir Gomes*	Mar/2002
39	**Opções sobre Dólar Comercial e Expectativas a Respeito do Comportamento da Taxa de Câmbio** *Paulo Castor de Castro*	Mar/2002
40	**Speculative Attacks on Debts, Dollarization and Optimum Currency Areas** *Aloisio Araujo and Márcia Leon*	Apr/2002
41	**Mudanças de Regime no Câmbio Brasileiro** *Carlos Hamilton V. Araújo e Getúlio B. da Silveira Filho*	Jun/2002
42	**Modelo Estrutural com Setor Externo: Endogenização do Prêmio de Risco e do Câmbio** *Marcelo Kfoury Muinhos, Sérgio Afonso Lago Alves e Gil Riella*	Jun/2002

43	**The Effects of the Brazilian ADRs Program on Domestic Market Efficiency** *Benjamin Miranda Tabak and Eduardo José Araújo Lima*	Jun/2002
44	**Estrutura Competitiva, Produtividade Industrial e Liberação Comercial no Brasil** *Pedro Cavalcanti Ferreira e Osmani Teixeira de Carvalho Guillén*	Jun/2002
45	**Optimal Monetary Policy, Gains from Commitment, and Inflation Persistence** *André Minella*	Aug/2002
46	**The Determinants of Bank Interest Spread in Brazil** *Tarsila Segalla Afanasieff, Priscilla Maria Villa Lhacer and Márcio I. Nakane*	Aug/2002
47	**Indicadores Derivados de Agregados Monetários** *Fernando de Aquino Fonseca Neto e José Albuquerque Júnior*	Set/2002
48	**Should Government Smooth Exchange Rate Risk?** *Ilan Goldfajn and Marcos Antonio Silveira*	Sep/2002
49	**Desenvolvimento do Sistema Financeiro e Crescimento Econômico no Brasil: Evidências de Causalidade** *Orlando Carneiro de Matos*	Set/2002
50	**Macroeconomic Coordination and Inflation Targeting in a Two-Country Model** *Eui Jung Chang, Marcelo Kfoury Muinhos and Joanílio Rodolpho Teixeira*	Sep/2002
51	**Credit Channel with Sovereign Credit Risk: an Empirical Test** *Victorio Yi Tson Chu*	Sep/2002
52	**Generalized Hyperbolic Distributions and Brazilian Data** *José Fajardo and Aquiles Farias*	Sep/2002
53	**Inflation Targeting in Brazil: Lessons and Challenges** *André Minella, Paulo Springer de Freitas, Ilan Goldfajn and Marcelo Kfoury Muinhos*	Nov/2002
54	**Stock Returns and Volatility** *Benjamin Miranda Tabak and Solange Maria Guerra*	Nov/2002
55	**Componentes de Curto e Longo Prazo das Taxas de Juros no Brasil** *Carlos Hamilton Vasconcelos Araújo e Osmani Teixeira de Carvalho de Guillén*	Nov/2002
56	**Causality and Cointegration in Stock Markets: the Case of Latin America** *Benjamin Miranda Tabak and Eduardo José Araújo Lima*	Dec/2002
57	**As Leis de Falência: uma Abordagem Econômica** *Aloisio Araujo*	Dez/2002
58	**The Random Walk Hypothesis and the Behavior of Foreign Capital Portfolio Flows: the Brazilian Stock Market Case** *Benjamin Miranda Tabak*	Dec/2002
59	**Os Preços Administrados e a Inflação no Brasil** *Francisco Marcos R. Figueiredo e Thaís Porto Ferreira*	Dez/2002

60	**Delegated Portfolio Management** *Paulo Coutinho and Benjamin Miranda Tabak*	Dec/2002
61	**O Uso de Dados de Alta Freqüência na Estimação da Volatilidade e do Valor em Risco para o Ibovespa** *João Maurício de Souza Moreira e Eduardo Facó Lemgruber*	Dez/2002
62	**Taxa de Juros e Concentração Bancária no Brasil** *Eduardo Kiyoshi Tonooka e Sérgio Mikio Koyama*	Fev/2003
63	**Optimal Monetary Rules: the Case of Brazil** *Charles Lima de Almeida, Marco Aurélio Peres, Geraldo da Silva e Souza and Benjamin Miranda Tabak*	Feb/2003
64	**Medium-Size Macroeconomic Model for the Brazilian Economy** *Marcelo Kfoury Muinhos and Sergio Afonso Lago Alves*	Feb/2003
65	**On the Information Content of Oil Future Prices** *Benjamin Miranda Tabak*	Feb/2003
66	**A Taxa de Juros de Equilíbrio: uma Abordagem Múltipla** *Pedro Calhman de Miranda e Marcelo Kfoury Muinhos*	Fev/2003
67	**Avaliação de Métodos de Cálculo de Exigência de Capital para Risco de Mercado de Carteiras de Ações no Brasil** *Gustavo S. Araújo, João Maurício S. Moreira e Ricardo S. Maia Clemente*	Fev/2003
68	**Real Balances in the Utility Function: Evidence for Brazil** *Leonardo Soriano de Alencar and Márcio I. Nakane*	Feb/2003
69	**r-filters: a Hodrick-Prescott Filter Generalization** *Fabio Araújo, Marta Baltar Moreira Areosa and José Alvaro Rodrigues Neto*	Feb/2003
70	**Monetary Policy Surprises and the Brazilian Term Structure of Interest Rates** *Benjamin Miranda Tabak*	Feb/2003