Journal of Mathematical Economics 45 (2009) 679–692

Contents lists available at ScienceDirect

Journal of Mathematical Economics

journal homepage: www.elsevier.com/locate/jmateco

A duality theory of payment systems☆

Rodrigo Peñaloza

Center of Research in Economics and Finance (CIEF), University of Brasilia, 70910-900 Brasilia, Brazil

ARTICLE INFO

Article history:
Received 23 February 2007
Received in revised form 29 February 2008
Accepted 15 May 2008
Available online 29 May 2008

JEL classification:
E58

Keywords:
Central banking
Real-time gross settlement systems
Shadow-prices

ABSTRACT

We model the Central Bank's management of intraday liquidity in modern real-time gross settlement systems as a linear programming problem parameterized by different intraday monetary policies, such as reserve requirements, net debit caps and Lombard loans. We then use duality theory to determine the shadow-prices of constraints of each bank. These shadow-prices can be used by the Central Bank to set personalized intraday monetary policies in order to reduce idleness of money and to give a microfoundation of the too-big-to-fail policy.

1. Introduction

A payment system is a network of banks in which funds transfers are made during the day. There are basically two types of payment systems. In the Deferred Net Settlement (DNS) system, interbank payments are cleared, netted out and settled with finality in the end of the day. Under the DNS system, banks clearly economize on liquidity needs, but are prone to systemic risks. The increase of systemic risk in DNS systems due to the increasing value of interbank transfers has been a constant concern for monetary authorities. The Bank for International Settlements (BIS) has therefore 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 their gross amount. In other words, no bank can be illiquid at any given moment. 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. The holding of excess liquidity for settlement purposes is known to be the main problem of RTGS systems, for a huge amount of money remains idle. We call it *liquidity idleness*. In order to facilitate the flow of payments and to reduce liquidity idleness, some systems allow for queueing of payments, net debit caps, Lombard loans from the Central Bank and splitting of payments. Net debit cap is a controlled overdraft that is granted to a bank. It allows the bank to become temporalily illiquid at some cost. A Lombard loan is a loan from the Central Bank to a particular bank. Usually the Central Bank lends against collateral and requires the bank to buy the collateral back in the end of the day. At the moment of lending, the price of the collateral is slightly reduced. This procedure is called a haircut. Some countries allow for splitting of payments. Instead of paying all or nothing, a bank can settle a fraction of the payment at one moment and settle the remaining fraction later in the day. We

☆ A prior version of this paper has circulated under the title "Shadow-prices in payment systems", Discussion Paper #241, Department of Economics, University of Brasilia.

0304-4068/$ – see front matter
doi:10.1016/j.jmateco.2008.05.002

call all these instruments *intraday monetary policies*. We refer to the RTGS report (BIS, 1997) for further details on payment systems.

In this paper we solve the following problems: (a) *How can the Central Bank set intraday monetary policies so as to minimize liquidity idleness?* (b) *How can the Central Bank measure the marginal effect of a bank's dollar withdrawal from the system on the optimum flow of payments?* Answering these questions is important. Problem (a) is a question that every central banker asks and it deserves attention. Besides, knowing the marginal effect of a dollar withdrawal from the system gives the Central Bank a sharp idea about the consequences of its intraday monetary policies. We can think of it as a measure of marginal contagion.

In order to solve these problems we model the management of payment systems by the Central Bank as an infinite-dimensional linear programming problem. Taking interbank payments as given as well as intraday monetary policies, the Central Bank chooses how to settle payments in order to minimize liquidity idleness. We then use duality theory to find the shadow-prices of liquidity constraints of each bank. The goal of our model is to use such shadow-prices to draw policy implications. The overall conclusion is that intraday monetary policies should be personalized.

The basis for our main contributions arise from the very geometric and functional analitic framework of the model, which is a novelty in the literature on payment systems. First, we show that interbank transfers and settlement of payments are linked by a dual relationship. Then we prove the existence of an optimal way for the Central Bank to settle interbank payments. Finally, we use duality theory to answer questions (a) and (b) above. As a by-product of our duality theory, we provide a shadow-price argument for the personalized treatment of banks in payment systems, which is in line with the economic wisdom that advocates that agents (in our case, banks) be priced according to their marginal values. Our duality theory of payment systems shed some light on the *t* oo-big-to-fail policy in that big banks tend to have very high shadow-prices, which explain their amplified impact on the system. Our model then offer a shadow-price argument in favor of that policy.

The infinite-dimensional linear programming framework has been widely used in general equilibrium analysis and other allocational problems (see, for instance, Anderson and Nash, 1987; Gretsky et al., 1992, 1999, 2002; Makowski and Ostroy, 2000), but not in the payment systems literature. The construction of our model is entirely based on the BIS report on real-time gross settlement systems (BIS, 1997).

The literature has focused on the influence of payment systems design on the behavior of banks (see DeBandt and Hartmann, 2000 for a survey). Freixas and Parigi (1998) used the Diamond-Dybvig framework to make comparative welfare analysis of DNS versus RTGS in a two-period two-island model with a liquidity shock and – the novelty of their model – a locational shock. They showed that, if returns are stochastic, then an RTGS system dominates DNS system whenever it is costly to keep an inefficient bank open, the probability of high return is low and the proportion of consumers that have to consume in another location is low. In other words, RTGS system is preferred when the probability of a bank failure is high. Bech and Garrat (2001), in a very interesting model, use a Bayesian game to analyze the strategic behavior of banks under the real-time gross settlement system. They show that intraday credit policies play an important role in the banks' decision to delay payments. Building on the literature on precautionary demand for reserves, Angelini (1998) shows that priced intraday liquidity creates an incentive for banks to delay payments and that the outcome will not be socially optimal.

Contrarywise to the papers cited above, we focus on the Central Bank itself. Some recent research has given attention to it (see, for instance, Angelini, 1998; Roberds, 1999) but our approach is quite different. The only literature whose approach to problem (a) somehow resembles ours is the one on gridlock resolution. Leinonen and Soramäki (1999) conclude, from simulations, that an RTGS sytem with queueing is always more efficient than a net settlement system with batch processing. Bech and Soramäki (2001) present a *model of gridlock resolution*, in which the objective is to maximize the flow of outgoing queued payments subject to liquidity constraints faced by banks under RTGS systems. Building on the *bank clearing problem* modelled by Güntzer et al. (1998), they introduce an additional constraint, called the *sequence constraint*, which states that payments have to be settled in a predefined order, usually the FIFO (first-in-first-out) rule. The general principles presented in Bech and Soramäki (2001) make their model the only one that is somehow related to ours.

We conclude that there has been no unique framework for the payment system problem. In this paper we offer a framework that is different from what has been done so far, so we hope to shed some light on the linear structure of payment systems and on the importance of shadow-prices as a powerful instrument in the hands of the central banker.

The framework of the model is presented in Section 2. In Section 3 we state a theorem regarding the existence of a solution to the Central Bank's problem. The main contribution of our model lies however on the dual problem associated with the Central Bank's liquidity management problem. The dual problem is presented in Section 4. 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, etc. We show how shadow-prices can help determine intraday monetary policies so as to bring systemic liquidity idleness down to zero. Therefore, reserve money can be fully used for settlement purposes, avoiding waste of systemic liquidity. The only intraday monetary policy that we do not address here is the queueing of payments. Elsewhere we include queueing and extend the model to another duality bracket (Peñaloza, 2002). Section 5 presents some policy implications. Section 6 concludes the paper. The proofs are presented in Appendix A.