

Leinonen, Harry (Ed.)

Book

Liquidity, risks and speed in payment and settlement systems: a simulation approach

Bank of Finland Studies, No. E:31

Provided in Cooperation with:

Bank of Finland, Helsinki

Suggested Citation: Leinonen, Harry (Ed.) (2005) : Liquidity, risks and speed in payment and settlement systems: a simulation approach, Bank of Finland Studies, No. E:31, ISBN 952-462-194-0, Bank of Finland, Helsinki,
<https://nbn-resolving.de/urn:nbn:fi:bof-201408071701>

This Version is available at:

<https://hdl.handle.net/10419/212964>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

Harry Leinonen (ed.)

Liquidity, risks and speed in payment and settlement systems – a simulation approach



Bank of Finland Studies
E:31 · 2005

Harry Leinonen (ed.)

**Liquidity, risks and speed in
payment and settlement
systems – a simulation
approach**



Bank of Finland Studies
E:31 · 2005

The views expressed in this study are those of the authors and do not necessarily reflect the views of the Bank of Finland or the respective institutions of the authors.

ISBN 952-462-194-0
ISSN 1238-1691
(print)

ISBN 952-462-195-9
ISSN 1456-5951
(online)

Edita Prima Oy
Helsinki 2005

Abstract

This publication consists of eleven separate studies on payment and settlement systems conducted using simulation techniques. Most have been carried out using the payment and settlement system simulators BoF-PSS1 or BoF-PSS2 provided by the Bank of Finland and presented at the simulator seminars arranged by the Bank. The main focus in the analyses is on liquidity requirements, settlement speed, gridlock situations, gridlock resolving methods, liquidity economising, systemic risk, and the impact of shocks on system performance. The studies look at systems in several countries and cover both RTGS and netting systems as well as securities settlement systems.

Keywords: simulation, payment and settlement system, liquidity, gridlock, systemic risk, counterparty risk

Tiivistelmä

Tämä julkaisu koostuu yhdestätoista erillisestä maksu- ja selvitys-järjestelmää koskevasta tutkimuksesta, jotka on suoritettu simulointi-menetelmiä käyttäen. Useimmat näistä tutkimuksista on tehty käyttäen Suomen Pankin maksu- ja selvitysjärjestelmäsimulaattoreita BoF-PSS1 ja BoF-PSS2. Tutkimukset on esitelty Suomen Pankin järjestämien simulaattoriseminaarien yhteydessä. Pääpaino tutkimuksissa on ollut likviditeettitarpeiden selvittämisessä, katteensiirron nopeudessa, lukkiutumistilanteissa ja niiden avaamiseen liittyvissä metodeissa, likviditeetin käytön tehostamisessa, systeimiriskeissä sekä poikkeus-tilanne- ja vastapuoliriskeissä. Tutkimukset koskevat eri maissa toimivia järjestelmiä, ja niiden joukossa on RTGS-järjestelmiä ja nettoutus-periaatteella toimivia maksujärjestelmiä sekä arvopaperikauppojen selvitysjärjestelmiä.

Asiasanat: simulointi, maksu- ja selvitysjärjestelmä, likviditeetti, lukkiutumistilanne, systeimiriski, vastapuoliriski

Preface

Payment systems are an integral part of an efficient modern economy. With continuous progress in globalisation and rapid technological advances, payment systems are, like other aspects of the economy, in a state of transition. International interdependencies and system linkages are increasing and real time settlement of payment by these systems is becoming commonplace. Liquidity demands will change in the future, as will the tools for liquidity management. The risks in payment and settlement systems are also changing, bringing new kinds of challenges. While counterparty credit risks are generally decreasing, liquidity risks are becoming more challenging. It is important to gain a deep understanding of the systems and their interdependencies in order to increase their efficiency under normal circumstances and their resilience in any abnormal situations that could arise.

Payment and settlement systems have proven to be a complicated area into which simulation techniques provide a good way to penetrate sufficiently deeply. It is possible to build models that closely replicate the real operating environment. These laboratory environments can be used for testing scenarios that cannot be observed in real operating environments.

The Bank of Finland has a long tradition of economic research, and modern payment and settlement systems have been one of the areas on which it has focused. This has resulted over the years in several research publications and seminars. The Bank decided in 1997 to start to develop a special PC-based tool, a simulator, for payment system analysis. The keen external interest in the Bank's simulation studies and the simulator itself led to the decision to create a public simulator version for research purposes (BoF-PSS2), which was ready for general distribution by spring 2004. BoF-PSS2 is currently used by over 30 institutions worldwide, mainly central banks.

I would like to thank all the authors for their contributions to this publication, which I trust will provide a good introduction to the simulation analysis of payment and settlement systems and will stimulate further research to enhance our understanding and improve the models and methodologies in the years ahead.

For the finalisation of the publication we are indebted to Päivi Nietosvaara for the text editing, Brian Fleming for revising the language of some of the contributions and Teresa Magi for printing administration. We are also indebted to Esa Jokivuolle, Harry Leinonen and Jouko Vilmunen, who have served as the editorial board

for the project. Kari Nihtilä at the Finnish Bankers' Association deserves our gratitude for the original idea of using simulation techniques for analysing the impact of EMU on Finnish payment systems. So does Kimmo Soramäki for designing and programming the first Bank of Finland simulator (BoF-PSS1) and for assisting in designing and testing the second one (BoF-PSS2). The European Central Bank warrants a special acknowledgement for making it possible to use Kimmo Soramäki's know-how. The second simulator was built by MSG Software Oy, and a special thank you goes to project manager Leena Tyni and chief designer Ville Ruoppi. The simulator projects have also benefited from the support and contributions of Marianne Palva and Heikki Koskenkylä. Harry Leinonen has been the project leader, and as such has played a central role in creating the simulators.

Finally, I would like to thank all other contributors, sponsors, commentators and users of the simulator for all the help they have provided during different periods of work on and with the simulator. The Bank of Canada, Bank of England and Federal Reserve Bank of New York deserve special acknowledgement for their concrete sponsorship of new features for the simulator. A more complete list of acknowledgements can be found on the simulator web site and in the user manual of the simulator.

I hope users of the simulator will continue to be active and that the simulator will attract new users and sponsors. It is a great pleasure for me to present via this publication the fruits of this productive cooperation between central banks.

Helsinki, May 2005
Matti Louekoski

Contents

Chapter 1

Harry Leinonen – Kimmo Soramäki

Introduction 9

Chapter 2

Harry Leinonen – Kimmo Soramäki

Simulating interbank payment and securities settlement mechanisms with the BoF-PSS2 simulator 15

Chapter 3

Risto Koponen – Kimmo Soramäki

Intraday liquidity needs in a modern interbank payment system – a simulation approach..... 71

Chapter 4

Harry Leinonen – Kimmo Soramäki

Optimising liquidity usage and settlement speed in payment systems 115

Chapter 5

Morten L. Bech – Kimmo Soramäki

Gridlock resolution and bank failures in interbank payment systems 149

Chapter 6

Kurt Johnson – James J. McAndrews – Kimmo Soramäki

Economising liquidity with deferred settlement mechanisms 177

Chapter 7

Johan Pettersson

Simulation of liquidity levels and delays in the Swedish RIX systems..... 217

Chapter 8

James J. McAndrews – George Wasilyew

Simulations of failure in a payment system..... 227

Chapter 9	
<i>Paul Bedford – Stephen Millard – Jing Yang</i>	
Analysing the impact of operational incidents in large-value payment systems: a simulation approach	247
Chapter 10	
<i>Morten L. Bech – Kimmo Soramäki</i>	
Systemic risk in a netting system revisited	275
Chapter 11	
<i>Emmanuel Mazars – Guy Woelfel</i>	
Analysis, by simulation, of the impact of a technical default of a payment system participant.....	297
Chapter 12	
<i>Matti Hellqvist – Jenni Koskinen</i>	
Stress testing securities clearing and settlement systems using simulations	321

Chapter 1

Introduction

Harry Leinonen – Kimmo Soramäki

1	Introduction	10
---	--------------------	----

1 Introduction

The last 10 or 15 years have seen a growing interest in research and development in interbank payment and settlement systems. Central banks have allocated effort and resources to this special area. Among the factors behind the movement in this area are international cooperation, technological advances and a growing dependency on payment systems that function smoothly under all circumstances. Cooperation has been a particularly important factor in the euro area, which has seen extensive linking of payment systems in order to give the area a totally integrated interbank payment system. The key official fora have been the payment system committees of G10 central banks and of the European System of Central Banks, and working groups of the European Commission.

There has been a long tradition at the Bank of Finland to use modelling tools such as simulation methods in its research activities. The research using simulation models for payment systems was initiated around the time Finland was joining the Economic and Monetary Union as it became necessary to examine the impact that the new environment for transferring payments and covering funds would have on Finnish payment systems. Out of this concern evolved the construction of the first payment system simulator (BoF-PSS1). The simulator proved to be an excellent tool for studying liquidity needs and system risks. Although the simulator was not originally intended for external use, other central banks made use of it with the help of the Bank of Finland.

Experience with the program and feedback from other central banks prompted the Bank of Finland to proceed with the development of a new and more diversified simulator designed especially for external use and international distribution. The new simulator (BoF-PSS2) was completed in spring 2004 and is available for research purposes free of charge.

Both the old and the new payment and settlement system simulator attracted international interest and a variety of research and studies in different central banks. The Bank of Finland arranged two international payment and settlement seminars and workshops in 2003 and 2004, and these seem now to have become annual events. The main goals of the seminars and workshops are to stimulate simulation-based payments and securities settlement research, to share research and experiences among the user community and to receive ideas and feedback on simulator development needs.

Payment system simulations seek to replicate the real world, but often in miniature or by incorporating only the most essential factors or dimensions relevant for the problem at hand. Simulations help us understand the dynamics of the payment system better, they can be used to analyse different scenarios and they may also be used in optimising the system studied.

Payment system simulations started to attract interest in the 1990s and were to a large extent used to discover how new settlement conventions would affect the participants and the overall system. This was the background to the Bank of Finland simulations and those conducted at the Bank of England on CHAPS. Extensive performance simulations of abnormal situations were a regulatory requirement for the start of the Continuous Linked Settlement (CLS) system. RTGS services were introduced in the 1990s, but other improvements to interbank settlement were also made, with a major impact on liquidity, risk issues and the management of these. Simulations proved to be a good tool for estimating the impact of these changes in increasingly complicated payment system structures and processes.

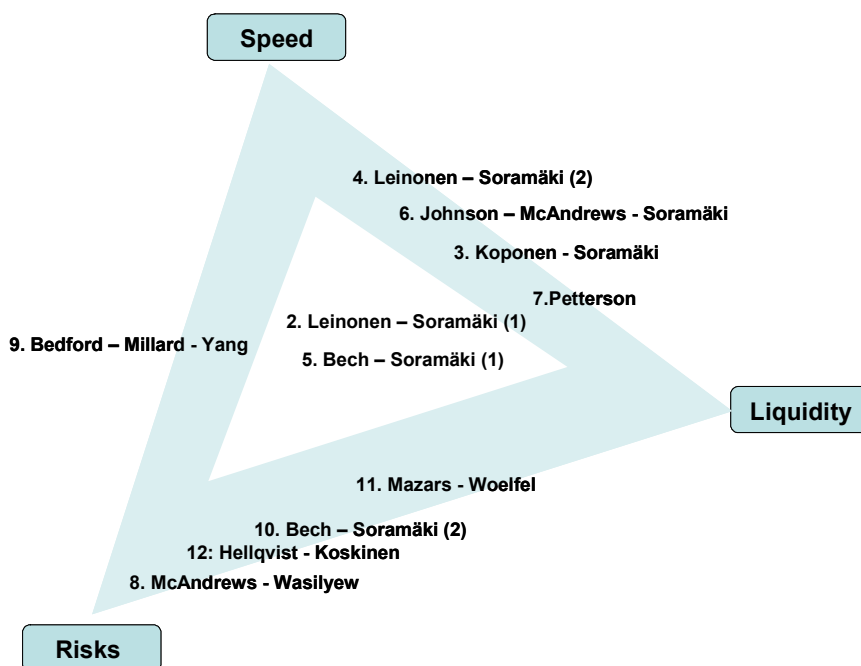
The input data can be of different types: historical, stochastic or behavioural and adaptive. Historical data has mainly been used for ‘what if’ types of study within a given system environment in which parameter relationships are fixed or predictable. Stochastic input has been used for theoretical studies, when real data is not available or there is an interest to compare stochastic situations with real-world observations. Most of the studies in this publication can be considered as static as they do not incorporate in the models the reactions of the participants to changing payment patterns or system features.

As the title of the book suggests, there are three key factors in payment and settlement systems: liquidity requirements, settlement speed and credit risks. The aim of studies in this area is generally to discover a good or even optimal balance between these factors. At one point it became apparent that the net settlement systems being employed contained substantial credit risks and authorities began to require a reduction in the risk positions through real-time settlement. Real-time settlement changes credit positions into liquidity requirements. The credit risks had often been created in the first place by faster payment processing without a simultaneous improvement in the speed of interbank settlement processes. By putting price tags on these basic factors the cost efficiency of different structures can be assessed. The articles of this book provide a broad view of the different aspects and interrelationships between liquidity requirements, risks and settlement speed.

The aim of the book is to bring together in a single publication the simulation studies and research conducted in the area of payment and settlement systems over the past few years. It is mainly based on the papers presented at the simulator seminars. However, some earlier papers have been included in order to illustrate the path of development in payment simulation modelling.

The individual chapters of the book cover three main factors in payment and settlement systems: liquidity, risks and speed. The content of the chapters could also be described according to their focus on these dimensions, as illustrated in Figure 1.1. Some of the chapters focus mainly on a single factor, while others analyse the trade-off between two or all three.

Figure 1.1. **The main themes of the articles**



Chapter 2 provides an overview of the dimensions analysed in the simulations, while Chapter 5 analyses these dimensions in Danish and Finnish payment systems. Chapters 8, 10 and 12 are mostly risk focused, while Chapter 11 adds liquidity issues to the risk analyses. Chapter 9 presents an analysis of operational risks and settlement delays, while chapters 3, 4, 6 and 7 analyse the trade-off between liquidity usage and settlement speed.

Of the many alternatives, the chapters in the publication are ordered according to two of the main themes of payment system analysis: liquidity and risk issues. Where possible, a chronological order is followed, making it possible to follow how analysis has evolved over the years. Chapters 3 to 7 focus mainly on liquidity and gridlock issues, while chapters 8 to 12 focus on different types of risk issues. Each chapter provides an individual stand-alone analysis, but some clearly build on earlier analyses. Each chapter is contributed by named authors.

Chapter 2 (Leinonen and Soramäki) provides an overview of the simulator (BoF-PSS2) and a general description of how the simulation technique can be applied to payment and settlement systems. It describes the features of a modern simulator model programmed closely to production systems and can use both real data and artificial stochastic input.

Chapter 3 (Koponen and Soramäki) begins the articles on liquidity issues. This study was conducted in 1997–1998 using the Bank of Finland's first simulator (BoF-PSS1). The main objective was to discover what kind of impact the introduction of the European Monetary Union would have on domestic systems and their liquidity needs. The study found that the situation for Finnish banks was reassuring and provided at the same time new insights into liquidity dependencies in RTGS and other payment system structures.

Chapter 4 (Leinonen and Soramäki) is a continuation to the Koponen and Soramäki study and assesses the efficiency of different queuing and gridlock resolution features as well as the impact of cost factors on liquidity and delay. It, too, employed the first simulator and the results were first published in 1999.

Chapter 5 (Bech and Soramäki) is a study of gridlocks and their resolution under both normal operating conditions and in situations of a bank failure. It provides a general presentation of earlier studies carried out between 2001 and 2003 in which both BoF-PSS1 and BoF-PSS2 simulators were used and with different data sets. It presents an optimal partial net settlement algorithm to solve gridlocks and assesses its efficiency in different liquidity situations.

Chapter 6 (Johnson, McAndrews and Soramäki) presents the possibilities of speeding up settlement and still economising on liquidity by using a receipt-reactive model in which the liquidity of incoming payments is used efficiently. This also makes it possible to reduce intraday overdrafts.

Chapter 7 (Pettersson) presents the liquidity situation in the Swedish RTGS system and how peak demands could be reduced. This chapter points to the possibility for increased international

benchmarking and learning experiences. Using the same assessment methods and key values, RTGS systems could be compared, and efficient solutions found in other countries could be tried out and used to balance liquidity demands in domestic systems.

Chapter 8 (McAndrews and Wasilyew) begins the articles on risk issues and presents a high-level statistical simulation model with artificial input data based on probability distributions. It focuses on systemic risk in payment systems. From 1995, it is the earliest article in the publication and provides the starting point for the analysis of systemic risk situations and simulation models. It also shows the need for using simulation models to analyse problems that cannot be tackled using calculus.

Chapter 9 (Bedford, Millard and Yang) analyses the impact of operational incidents in the UK RTGS system and the extent to which contingency arrangements can cope with such situations. Operational incidents are typical situations in which simulations provide an opportunity to test in advance different kinds of contingency arrangements.

Chapter 10 (Bech and Soramäki) focuses on systemic risks in netting systems by using a simulator designed for the study of cascading failures. This chapter also points towards the possibility of using different kinds of network models for describing the processes within payment systems.

Chapter 11 (Mazars and Woelfel) analyses the impact of technical defaults in the French net settlement system. This study was conducted using a simulation model developed by Banque de France and describes different methods to reduce the negative impacts of such events.

Chapter 12 (Hellqvist and Koskinen) analyses the impact of operative disturbances in securities settlement systems. The DVP (delivery-versus-payment) requirement in securities settlement systems can result in complicated interdependencies that can expand considerably the impact of even small disturbances. The risk dimensions in securities settlement systems have as yet been little studied, and this is an area that requires more in-depth studies.

The overall result of the studies can be characterised as a deeper knowledge of payment systems and of the internal and external factors and parameters that affects them. The articles also show the learning curve through international research cooperation that has helped the design of more efficient and stable systems. The new features tested via simulations have been implemented in operative systems. However, as can be seen from the papers, payment systems are used in varying environments with different emphasis on risk, liquidity and

speed factors, which results in preferences for different kinds of national and international setups for the payment volumes and conventions in question.

The Bank of Finland intends to continue to develop the simulator software, arrange simulator seminars and stimulate payment systems research with a view to publishing new articles presenting new issues and themes. A fairly promising research topic is the use of behavioural and adaptive input models in which the transaction flow and or processing patterns change over the simulation period due to behavioural and adaptive responses among the agents in the model. Participants in payment systems respond to external and internal stimuli, eg incoming and outgoing payment flows, responses by other participants, increasing risk positions, technical problems and different materialised risk scenarios. A payment system is a network of cooperating participants and payment systems are currently forming links between each other, resulting in a multilayered network structure. Network analysis techniques will probably also become interesting new research topics. Also securities settlement systems have been rather scarcely studied using simulation techniques. The years ahead will likely see more analysis on the interrelationship between funds and securities settlement. We hope that this publication will stimulate new studies in this multidimensional business area.

Chapter 2

Simulating interbank payment and securities settlement mechanisms with the BoF-PSS2 simulator

Harry Leinonen – Kimmo Soramäki

2	Simulating interbank payment and securities settlement mechanism with the BoF-PSS2 simulator.....	19
	Abstract.....	19
2.1	Introduction.....	19
2.2	General features of payment and securities settlement systems.....	21
2.3	Interdependencies in the payment and settlement process.....	24
2.3.1	System hierarchy.....	24
2.3.2	Transaction types and interdependencies.....	27
2.4	The general process of interbank settlement.....	28
2.4.1	Steps in payment processing.....	28
2.4.2	Transaction flows.....	30
2.4.3	Controlling transaction flows.....	32
2.4.4	Gridlock resolution features.....	33
2.4.5	The need for liquidity.....	36
2.4.6	System configuration elements.....	39
2.5	The costs of settlement.....	40
2.5.1	Cost of liquidity.....	41
2.5.2	Cost of financial risks.....	43
2.5.2.1	Credit risk.....	43
2.5.2.2	Liquidity risk.....	44
2.5.2.3	Settlement risk.....	45
2.5.2.4	Systemic risk.....	45
2.5.3	Cost of delayed payments.....	46
2.5.4	Combined costs.....	47

Appendix 1	An overview of the BoF-PSS2 payment and settlement system simulator	49
Appendix 2	BoF-PSS2 screen shots.....	58
References	67

2 Simulating interbank payment and securities settlement mechanisms with the BoF-PSS2 simulator

Abstract

The simulation technique provides a new means for analysing complex interdependencies in payment and securities settlement processing. The Bank of Finland has developed a payment and settlement system simulator (BoF-PSS2) that can be used for constructing simulation models of payment and securities settlement systems.

This chapter describes the main elements of payment and settlement systems (system structures, interdependencies, processing steps, liquidity consumption, cost and risk dimensions) and how these can be treated in simulation studies. It also gives examples of how these elements have been incorporated in the simulator, plus an overview of the structure and features of the simulator itself.

2.1 Introduction

Interbank payment and securities settlement mechanisms are the main facilities for transferring monetary claims and assets between financial institutions. These systems transfer many times the value transferred by cash instruments or retail payments.

The infrastructure has gradually grown into a complicated interactive network of systems that transfer claims and assets at the domestic and international level. Integration of these systems has resulted in critical interdependencies. The configurations found around the world have evolved to address local needs, customs and process-organisation patterns. Technical solutions depend on when the systems were implemented. This has resulted in a wide variety of configurations; with some configurations and system features being better suited to processing specific transaction flows. Suitability and efficiency can be assessed against objectives defined for financial systems, which also vary over time and region. Typical system design

objectives include low counterparty risk, quick throughput, low liquidity consumption and low settlement costs.¹

The characteristics of different payment and securities settlement systems are difficult to analyse with traditional econometric tools, and econometric models are often too general to describe systems at the level of detail needed to capture the differences that arise from various design parameters.

Simulations, in contrast, provide the opportunity to get closer to reality and make detailed analyses. Simulations can use actual production transaction flows and exactly mimic the specific features of each system, thereby yielding more precise and policy-relevant results for the specific environment of interest. They can also be used to provide empirical data on rare events (such as bank failures) or imagined system designs and structures. Of course, simulations have certain limitations in optimisation analysis. Generally, only a ‘what if’ type of enumeration is possible, which always leaves the possibility for undiscovered better solutions.

The aim of this chapter is to describe:

- the general elements present in payment and securities settlement systems;
- the most interesting aspects to analyse in these systems; and
- the possibilities simulations provide for studying these dimensions.

This chapter also serves as a background document for the new version of the Bank of Finland payment and securities settlement simulator (BoF-PSS2). In presenting the general elements of payment and securities settlement systems, it simultaneously describes in general terms the structure and features of the simulator. An overall description of the simulator can be found in Appendix 1. Detailed documentation is posted at www.bof.fi/sc/bof-pss. The simulator is freely available to central banks and research institutions.

The chapter is organised as follows. Section 2.2 presents the general features and structures found in payment and settlement systems. Section 2.3 describes the interdependencies between and within systems. Section 2.4 discusses the payment and settlement process. Section 2.5 describes the cost and liquidity aspects of settlement systems.

¹ For information and studies on different arrangements see Borio et al (1992), BIS (1990 and 1997), and ECB (2001).

2.2 General features of payment and securities settlement systems

Payment and settlement systems can be categorised according to the transaction types they process, ie customer transfers (payment systems), interbank settlement transfers (settlement systems) or a combination of both. Payment systems process different types of transfer (credit transfers, direct debits, cheques, etc) between customer accounts. Pure settlement systems are used solely for interbank settlements and no end-customer information is conveyed. Interbank claims usually originate from payment systems and securities settlement systems representing settlement for batches of individual payments. In a mixed system, customer payments and interbank settlements are processed in parallel.

Securities settlement systems process customer transfers of securities, mainly in book-entry format. While payment and securities settlement systems today are clearly separated, the technical process for transfers related to monetary currencies and book-entry securities are essentially the same. In both types of system, accounts representing funds or securities are credited and debited.

Payment systems may also be categorised as retail or large-value payment systems. This distinction has been important because of the different risks involved and (at least in the past) differences in service speed and efficiency. Most retail payment systems are currently settled on a net basis using very simple algorithms. In these systems, the liquidity impact and settlement risks are generally low, and therefore no sophisticated liquidity and risk management tools are warranted. The opposite is true for large-value payment systems, which often contain sophisticated risk and liquidity management features.

The traditional approach to processing payments was end-of-day net batch processing, whereby payments were collected by the banks in daily batches and handed over to payment systems that cleared them over the following days. Interbank settlement for such payments typically took one to three days. Today, batch systems operate with settlement cycles as short as every 30 minutes. Such systems are called deferred net settlement (DNS) systems.

Thanks to real-time processing capabilities, payments can now be processed individually and immediately. Real-time processing is mainly used for large-value transfers; the bulk of retail payments are still made in deferred batches. Real-time processing should gradually expand to all kinds of payments in response to customer service

requirements and the growth of e-commerce. Real-time payment systems fall into two groups: the real-time gross settlement (RTGS) systems of central banks and private continuous net settlement (CNS) systems. Interbank settlement transfers in RTGS systems are directly booked on central bank accounts: ie payments and settlements are processed simultaneously. In CNS systems, payments are booked immediately, while final settlement, eg with central bank money, is typically delayed until the end of the day.

In true real-time processing, the liquidity need is fixed by the processed payment flow so it cannot be influenced. In fact, the liquidity need can be smoothed by deferring payments (eg queuing) and by netting queued payments between banks with opposing queued payment flows. This also makes it possible to save interbank settlement liquidity, as all payments do not require immediate processing. This has resulted in the emergence of a third group of systems, hybrid systems, which combine features from real-time and deferred net settlement systems. Most large-value payment systems currently operated by central banks are RTGS systems, but they are continually acquiring an increasing number of hybrid features for preserving liquidity, optimising the use of liquidity and resolving gridlock situations.

Gross-based real-time processing is the stated goal of securities processing systems. In most cases, however, such systems actually only deliver a type of deferred real-time processing, which takes place several days after the securities trade was agreed. Thus, the transaction processing of securities settlement systems is typically T+3, although even T+5 systems can be found. Given that securities settlement systems involve so much risk, the current trend is to move from deferred net settlement to deferred real-time settlement. Internationally, the yet-to-be-achieved objective has for some years been to move to T+1 real-time settlement. Limiting risks, increasing settlement speed and removing barriers to efficient cross-border transfers have been very topical issues regarding securities settlement systems.² Securities settlement should eventually move to true real-time (T+0) processing (ie settlement immediately when the deal is made).³ Indeed, a true real-time system can already be found in the Czech Republic.⁴ Some systems also permit securities lending in real

² See BIS (1995), European Commission (2001, 2002 and 2003) and Group of Thirty (2003).

³ See Leinonen (2003).

⁴ See ECB (2002).

time, which can be seen as the beginning of an expanded approach to real-time processing.

Systems may also be categorised as public or private systems depending on the settlement institution. The settlement institution is the institution across whose books transfers between participants take place to achieve settlement within the settlement system⁵ In most cases, the central bank is the principal settlement institution in domestic payment systems. The settlement asset in such systems is central bank money, ie claims against the central bank. For large value payments, the most common settlement asset is central bank money.⁶

The settlement institution in a payment or securities settlement system can also be a private entity such as a commercial bank or a financial institution specifically created to act as a settlement institution (eg a limited purpose bank). In such systems, the settlement asset is commercial bank money, ie claims against private financial institutions. These systems generally have an added risk, as the private settlement institution may go bankrupt. The bankruptcy risk of a central bank, in comparison, is almost nonexistent. Reducing the level of risk that the settlement institution takes on can reduce its probability of failure and the associated risk.

Some countries have a two-tiered settlement hierarchy, whereby a small number of large banks settle on the books of the central bank, while a large number of small and intermediary banks settle with the facilities of the larger banks. Savings and cooperative banks may also maintain an internal settlement bank as an intra-group settlement institution and as the external gateway to other banks or payment systems. A multi-tiered structure adds new risk and processing layers. Modern technology supports flat network-based structures with direct ITC contacts between all parties.⁷

In securities settlement systems, central securities depositories (CSDs), central counterparty clearing houses (CCPs), central banks and private banks can function as settlement institutions. In international securities processing, in particular, private settlement institutions currently play a major role.

The increasing flow of cross-border payments has resulted in a growing demand for multi-currency processing and settlement possibilities. Until recently, most international transfers were settled through the correspondent networks of international banks. Today, there is a clear move towards international clearing and settlement

⁵ BIS (2001).

⁶ See BIS (2003).

⁷ See Leinonen et al (2002).

systems. There is also a distinct preference for central bank settlement of systemically critical systems.

The finality of payments and settlement has become more important with increasing cross-border payment flows over several jurisdictions. Customers and participants need clear rules defining when a transfer is final and irrevocable. The directive on settlement finality⁸ has harmonised the EU legal rules on this issue.

The BoF-PSS2 simulator can be used to study the impact of many of the developments in payment and settlement systems described above. The simulator contains the basic features for RTGS, DNS and CNS processing. It can process large transaction volumes (several million) and the number of participants can be high (tens of thousands). A participant can further have several accounts that can be used for identifying different currencies or types of book-entry security.

2.3 Interdependencies in the payment and settlement process

2.3.1 System hierarchy

Payment and settlement systems generally take on a hierarchical structure where different types of transactions are handled in different systems. Obligations arising from these systems are settled in interbank settlement systems. To reduce liquidity need, interbank settlement can be concentrated in a single settlement institution, typically the central bank and its RTGS system. There are several reasons for this. A central bank is neutral and provides a common settlement institution for all participants. Credit risk and reserve requirements are eliminated, standing facilities and lender-of-last resort support are available, and the central bank can enforce regulations/recommendations on systemically important payment and settlement systems.

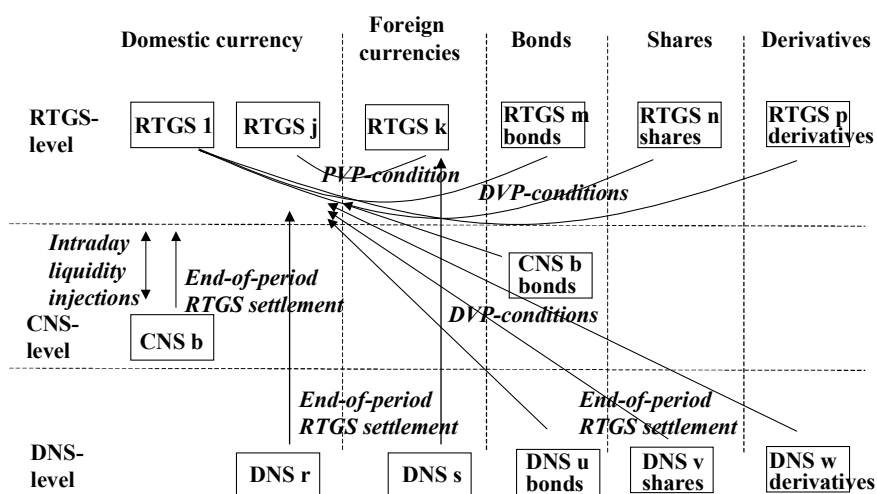
As regards retail payments, cheques are often handled separately from card payments. Credit transfers and direct debits may also be processed in separate systems. Domestic payments are normally processed apart from foreign payments. Securities settlement

⁸ Directive 98/26/EC of the European Parliament and of the Council of 19 May 1998 on settlement finality in payment and securities settlement systems.

transactions are processed in special systems, which are often differentiated into systems for interest-bearing and equity instruments. The result varies from country to country depending on the number of factors. In some countries, the private sector has been eager to build payment system infrastructure; in others, this has been a public task. Sometimes new developments have resulted in completely new systems, while in other cases old systems have been enlarged and upgraded.

The final end-of-day settlement typically occurs in public RTGS systems. This is especially the case for high-value payments, but private settlement institutions are occasionally used. The settlement positions from the other systems are transferred by different means from the other (ancillary) systems to the RTGS system by the end of the day. Figure 2.1 describes the general structural possibilities.

Figure 2.1 **The hierarchical structure of payment and securities settlement systems**



The RTGS level is generally the highest level, as most other systems eventually settle on this level. RTGS systems can be used solely for interbank settlement or also for transmission of individual payment transactions.

CNS systems can be considered private RTGS systems that normally settle at end of day in RTGS systems. However, these systems often use various kinds of swap or liquidity injection methods to reduce their internal risk positions (which implies interfaces with the RTGS systems). Sometimes private systems can ‘autonomously’

settle using central bank RTGS systems. In such cases, the private system transfers central bank liquidity into a separate account held by the central bank or the system itself on behalf of the clearing parties. All transactions are then booked on these accounts. Examples of CNS systems with RTGS interfaces are the international CLS, the French PNS and the US-based CHIPS systems. In all these systems liquidity is transferred into special accounts to reach full collateralisation. Examples of partly collateralised and limit-based systems are the European Euro1 and the POPS system in Finland. Participants in these systems use intraday liquidity swaps to free credit caps and an end-of-day settlement mechanism to square the positions in the RTGS systems.

DNS systems are typically different types of Automated Clearing Houses (ACH) and thereby private ancillary systems using RTGS accounts for end-of-day settlement. The ACH routes individual payments and calculates net positions for end-of-day settlements.

There can also be a number of systems operating in parallel at the same level with intersystem transaction flows. RTGS systems and securities settlement systems can have direct links with similar systems. The European TARGET system is a good example of interacting RTGS systems.

A hierarchy of interdependent systems involves a number of synchronisation problems. Transactions generated in one system have to be processed in the other system. The liquidity is shared by the systems. It should be able to flow between these systems and should be available according to the needs of the different systems. Swap and liquidity injections can be used as methods of liquidity transfer. Interdependence also increases contamination risks. Problems in one system can affect the other system: for example, in a crisis scenario, RTGS problems in one system can hinder processing in the international CLS system, which will in turn affect the other RTGS systems.

It is possible to build many types of system hierarchy with the BoF-PSS2 simulator. The transactions sent from one system can be received by any participant (account) in any other system in the same simulation. There can be a large number of systems in the same simulation. Intersystem accounts are automatically generated for maintaining the intersystem balances. Specific features and settlement algorithms can be specified separately for each system. The simulator can therefore contain eg TARGET-type structures with a large number of interconnected RTGS systems. It can also contain typical domestic structures with one RTGS system as the main system and a

number of ancillary systems processing customer payments, securities, etc.

Simulations can also be used to assess different settlement modes, eg if the system moves from net settlement to gross settlement, or when additional settlement cycles are introduced. Such changes are likely to have an impact on the liquidity requirements of the participants, but also on their submission patterns of payments for processing. Unless the submission patterns are also simulated, this will need to be controlled when the simulation results are assessed. The impact on payment queues can also be studied when systems are simulated with varying levels of liquidity available to participants.

2.3.2 Transaction types and interdependencies

A basic payment consists of a debit and a credit booking. A basic free-of-payment (FOP) securities book-entry transaction consists of an asset debit and an asset credit booking. In both cases, the systems are closed loops and the total amount of funds on participant accounts in a system remains the same.

Central banks and CSDs are in the position to increase the total available balance of liquidity and assets. Other settlement institutions can also increase the credit balances by extending more loans. From the system participants' point of view these are external transactions, but technically they are often carried out through special settlement institution accounts that resemble normal participant accounts.

To ensure simultaneous and dependent deliveries, PVP, DVP and DVD (payment-versus-payment, delivery-versus-payment and delivery-versus-delivery) processes have been established. The PVP process is used when settling currency deals: ie the two payment transactions in different currencies are processed only if both can be made final simultaneously. The DVP process is used for settling securities deals by requiring simultaneous settlement of the payment and asset legs. The DVD process can be used for ensuring securities credits by making the delivery of borrowed securities and their collateral dependent on each other. Most countries and most securities settlement systems require DVP-based settlement to reduce risks.

The simulator provides the possibility for PVP, DVP and DVD transfers within an RTGS system, between different RTGS systems and within a DNS system. PVP, DVP and DVD transactions are identified and matched based on an explicit link code that is provided by the user for each transaction. Simulations can be used to quantify the exposures that arise in unsynchronised settlement of currencies or

securities. Likewise, they can be used to quantify the added liquidity requirements (or delays) when the two legs of the transactions need to be effected simultaneously.

2.4 The general process of interbank settlement

2.4.1 Steps in payment processing

All payment and securities settlement transactions consist of a debit leg and a credit leg. Funds are moved from one interbank settlement account to another, and the customer liabilities are booked on customer accounts in the banks' systems. As a consequence of the dematerialisation of securities certificates, the transfer of securities consists merely of debits from and credits to securities accounts. However, settlement of a security deal requires that both the ownership of the security and the payment are transferred.

System participants perceive the system as a flow of outgoing and incoming transactions that result in a settlement balance. The actions of one participant affect the flow of his outgoing transactions, while the actions of other participants and the system design affect the flow of his incoming transactions. Consequently, there are continuous changes in settlement balances. Rules and requirements set by the settlement institution or agreed among the participants thus impact on settlement balance, credit availability, etc. The settlement features of the system are generally developed to support efficient settlement of transactions.

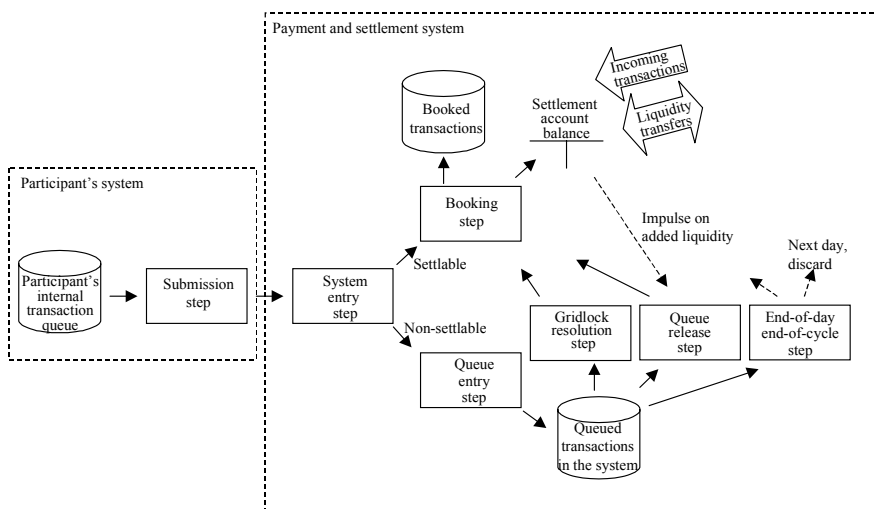
The process within a payment and securities settlement system can be separated into general steps or processes (see also Figure 2.2):

- a. **Submission**, whereby the participant sends a new transaction for processing to the system, possibly from an internal transaction queue.
- b. **Entry**, whereby the system determines whether the transaction in question can be directly booked either completely or partly (step c), queued for deferred settlement (step d), or rejected and resubmitted later.
- c. **Booking**, whereby transactions eligible for booking are booked on settlement accounts as debits from senders and credits to receivers.
- d. **Queue entry**, whereby transactions that are ineligible for booking are transferred to the waiting queue, and where the instruction may be modified (eg to split the transaction into several smaller ones).

- e. **Gridlock resolution**, whereby algorithms for simultaneous settlement of multiple queued transactions are applied to identify a subset of transactions from the queues that can be booked without breaching risk management or other constraints set for the system. Transactions can further be netted bilaterally or multilaterally, and for a sub-group or for all participants.
- f. **Queue release**, whereby queued transactions are released as soon as they become eligible for booking (eg due to added liquidity provided by incoming transactions from other participants or liquidity injections).
- g. **End-of-settlement cycle**, whereby the handling of transactions that will remain ineligible for booking until the end of the settlement cycle is determined.

Figure 2.2

The processes/steps in payment and settlement systems



Payment and securities settlement systems display varying degrees of complexity. The basic elements of submission, entry and booking steps are available in all systems, while queuing-based functions depend on the availability of a queuing mechanism and related sub-functions.

In the case of securities or foreign exchange settlement systems, mechanisms for synchronising the transfer of the security and the payment, or two currencies may additionally be present (delivery-versus-payment or payment-versus-payment processing).

The process within the simulator follows the structure described in Figure 2.2. The simulator design is modular and separate algorithms are used for the different steps. The user can select among several algorithms for each step. The algorithms provided as part of the software should cover the most common settlement conventions. Users can also construct their own custom modules with special features not covered by the provided algorithms or for testing new solutions. The library of simulator algorithms should increase in the coming years.

2.4.2 Transaction flows

The structure of the payment flows has a considerable impact on liquidity requirements and credit positions. Most of the settlement features of modern systems have the general objective of reducing liquidity requirements or risk exposures or increasing settlement speed by rearranging the settlement order of the transaction flow. Payment flows with large intraday variations generally consume more resources than synchronised flows in both directions. Participants can smooth flows by changing submission patterns.

Figure 2.3 **Examples of alternative transaction processing orders and liquidity impacts**

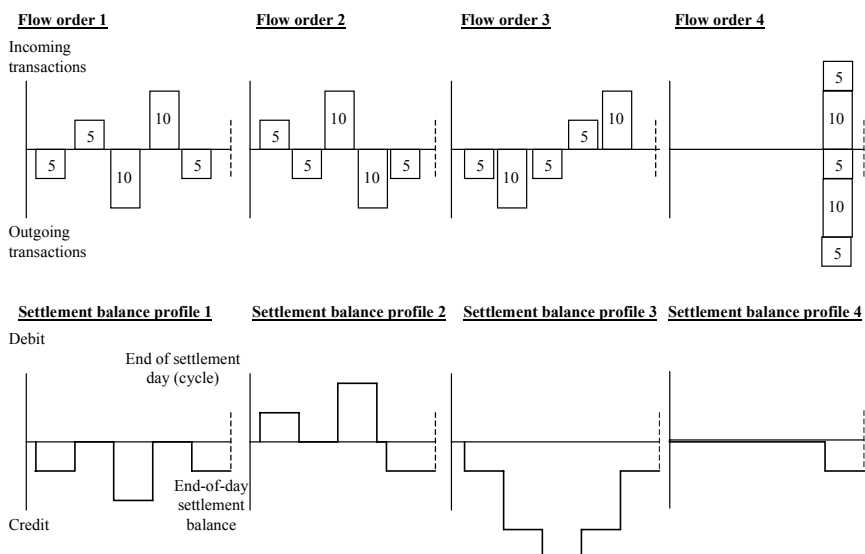


Figure 2.3 gives examples of the impacts of alternative flow orders on the liquidity needs of a bank. All four examples share the same transaction flow, which results in an end-of-day balance of minus 5 units. In example 1, the bank continuously runs a deficit or zero position towards other banks (or the settlement institution). In example 2, the bank can make its intraday position positive or at least zero for most of the day by rearranging the outgoing transfers, eg by queuing payments that would cause the balance of the bank to go below -5 units. Only the final transaction creates a deficit position. In example 3, incoming payments are delayed (for instance, due to technical problems somewhere in the system), while the bank in question submits its outgoing payments early in the day. Here, the bank is subject to a large deficit throughout the day. In case 4, all transactions are delayed to the end of the day and processed simultaneously.⁹

Currently, there is a clear difference between overnight and intraday liquidity costs. Overnight delays in delivery are more costly than intraday delivery differences, so participants are motivated to ensure that their anticipated and planned overnight positions are reached. There is greater flexibility in planning overnight positions when the bulk of transactions are processed early in the day, so the market can choose to introduce certain market practices to be followed that require early submission of transactions (especially small and medium-sized transactions).

Large-value transactions and ancillary system settlements consume liquidity. To benefit from the continuous off-setting in RTGS systems, the market participants might agree that certain transactions are submitted within a specific time interval to help liquidity planning.

Simulations can be used to assess the impact of altered payment flows by simply modifying the submission times of payments in the underlying data. Such changes might be caused by policy changes in the payment system (eg differential pricing or the cost of intraday credit), altered market practices or disturbances in the markets (eg delays in payment submission by the banks due to uncertainties in the market or technical failures). Simulations can be used to determine the impact on payment queues and liquidity requirements due to the altered payment flow.

Simulations can also be used to find processing timetables and patterns that would preserve liquidity. The probable effects of new

⁹ For analysis of intraday payment patterns in Fedwire see McAndrews and Samira (2002), and for TARGET see ECB (2003).

market conventions can be studied via simulations by reformatting the transaction flows according to the new conventions.

2.4.3 Controlling transaction flows

The transaction flow has a considerable impact on liquidity positions and credit risk exposures. Settlement systems therefore often provide procedures by which the transaction flow can be monitored and controlled with greater synchronicity of funds received and sent to preserve liquidity.

The first control point lies within the sending participant's internal system, where the decision on when to submit the transaction to the central payment and securities settlement system is made. The central system may have a pre-submission storage facility to which participants' internal systems can send instructions in advance. The transactions can be released from pre-submission storage based on parameters and rules employed for this purpose, eg a submission time in the instruction.

The most common processing order in payment systems is FIFO (= first in, first out). FIFO is used to release payments from queues that have built up, eg in response to a lack of liquidity. Because transactions have different levels of urgency, instructions are often prioritised to allow more important transactions to bypass the FIFO order. Other queuing orders are also possible. For example, releasing payments in order of size starting from small transactions is likely to reduce the average queuing time per transaction (although this will still probably delay large transactions). In some systems, participants may also reorder the transaction queues according to their preferences (eg the UK's CHAPS system).

Splitting transactions allows more efficient use of the available liquidity. This can be done using two main conventions: by defining a maximum transaction size according to which larger transactions are split or by using the available liquidity in full to create a part of the current transaction that could be settled.

Settlement systems also exhibit hoarding behaviour. Participants may delay transactions to reduce their own liquidity needs, which in turn can cause congestion at the end of the day if other participants also delay their transactions. To control fair reciprocity, multilateral or bilateral sending limits can be used. If bilateral limits are used, a participant will only release new payments to counterparties that have released the anticipated flow of transactions.

In the simulator, transaction flows must be defined separately for each system. Historical payment data are generally used for these transaction flows. When the characteristics of the system are changed (eg optimisation methods are tested), participant behaviour is also likely to change. Simulations can, however, also model participant behaviour. The submission algorithm determines the transaction flow to the system. The only submission algorithm currently available submits the payments to the settlement systems in chronological order. In order to bring greater dynamism to the model, it might, therefore, be worth developing submission algorithms that decide for each participant which payments should be submitted, and when. Alternatively, the rules for participant behaviour in existing systems might be studied.

For the control of transaction flows within the system, the simulator provides queuing based on pure FIFO or FIFO with priority. The FIFO order can also be bypassed. The user can also use user-defined fields for creating custom queuing orders. To ensure an unambiguous queuing order the last sorting field is always the transaction ID. Transactions can be split by using a fixed maximum transaction size or by splitting according to available liquidity. Parallel accounts can be used for introducing hoarding behaviour and keeping liquidity on separate accounts.

2.4.4 Gridlock resolution features

We use the definition for gridlock presented in Bech and Soramäki (2001 and 2002), where gridlock in a settlement system is defined as a situation where there is insufficient liquidity to settle given transactions one by one in a specific order, but there is enough liquidity to settle them simultaneously by a netting procedure. There are several algorithms for solving gridlock situations.

Splitting transactions was described as a transaction control feature, but it can also be seen as a gridlock resolution feature. For example, if two participants have transactions with each other queuing and one of the participants has even a small amount of liquidity available, splitting according to available liquidity can process part of the original payment. This liquidity inflow to the receiver of the payment may trigger settlement of other payments.

Bilateral offsetting is a bilateral process that can solve the gridlock between two parties by netting transactions in the waiting queue. Offsetting can take place in a different order, eg by FIFO, priority or size. Offsetting algorithms that work in a specific order are

undemanding from the computational standpoint, while an algorithm that tries to determine the maximum transaction value to offset irrespective of transaction order may become computationally very complex in the case of many queued transactions, which will have a very large number of possible alternatives.

Full multilateral netting is a very simple case to solve: all transactions in the queue from all participants are netted and the net balance is booked on the settlement account. However, if there is a lack of liquidity to cover the negative balances of participants, full multilateral netting has to be discarded and partial multilateral netting attempted instead.

Partial multilateral netting means that some transactions or participants are removed from the netting procedure to identify a subset of transactions that can be settled in accordance with the system's risk management rules and other rules. As in bilateral offsetting, transactions can be removed in a given queue order, which makes the computational task easier. Transactions are removed one by one from the queues of participants that lack liquidity until a solution is found that satisfies the liquidity constraints. In the case of many participants and long queues, it is a computationally non-trivial task to find an efficient algorithm to solve the problem without a queue order constraint. One possibility to improve the netting rate might be to try out a number of partial netting solutions with different queue orders in succession, eg FIFO first and then order of size.¹⁰

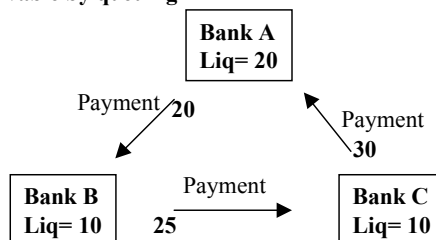
The need for and effects of gridlock resolution depend on the available liquidity and the urgency of settlement. There are hardly ever transactions in queues when participants have ample liquidity. Hence, gridlock situations are rare and there is little need for gridlock resolution algorithms. Netting always requires the retention of payments in queues for a while to let sufficient material for netting accumulate. When all transactions are so urgent that they cannot be allowed to queue, participants have no other option than to ensure sufficient liquidity for immediate settlement.

¹⁰ See Günzter et al (1998) and Ganz et al (1998) for a discussion on partial netting algorithms for payments and securities settlement respectively.

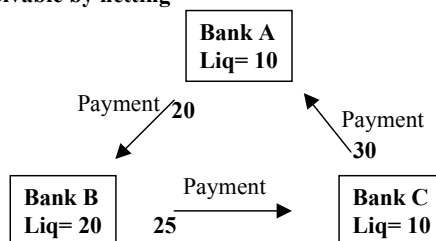
Figure 2.4

Examples of the impact of liquidity distribution (eg gridlock and deadlock)

Liquidity distribution 1:
Circular, resolvable by queuing



Liquidity distribution 2:
Gridlock, resolvable by netting



Liquidity distribution 3:
Deadlock, resolvable only by adding liquidity to Bank C

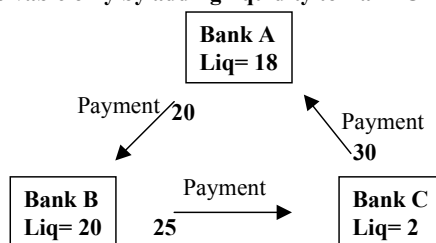


Figure 2.4 shows three typical situations. In example 1, the situation can be resolved by Bank A paying Bank B, which is paying Bank C, which in turn is paying Bank A. Here, Bank A has so much liquidity that it can initiate the circular process. In example 2, a typical gridlock situation, no bank has enough liquidity to make a payment by itself, but when incoming and outgoing payments are netted and the available liquidity is used, the payments can be settled. In the last example, a deadlock, Bank C has no way to make its payment because the liquidity from the incoming payment is still insufficient to cover

the outgoing payment. This, in turn, makes it impossible for A to make its payment and B to make its payment.

The simulator contains a full multilateral netting algorithm and a partial netting algorithm as described in Bech and Soramäki (2001). For the partial netting algorithms, different types of queue order can be defined. The output tables contain detailed statistics on every netting session.

Simulations can be used to test and compare the various optimisation methods. They can be helpful in selecting the best gridlock resolution, splitting, or bilateral offsetting algorithm for the system and in assessing its effects in terms of reduced liquidity requirements and delays in payments. Such simulations cannot only be performed to see the impact at the day-to-day level, but also to assess how such features can remedy abnormal situations such as delays caused by market uncertainty or the effects of participants being unable to fulfil their obligations in the system.

Simulations can be used to quantify the need for any overnight funding or the value and number of transactions that remain unsettled (eg in the event of a participant failure) if funding is not available. They can also be used to test the effectiveness of various gridlock resolution algorithms in their ability to clear queues at the end of the day.

2.4.5 The need for liquidity

Payment processing requires liquidity, ie assets to pay for interbank claims arising from payment transfers. One form of liquidity is the possibility to have negative positions, ie credit limits, whereby the settlement institution grants the necessary credit. This, of course, creates credit risks. Much attention has been devoted to reducing credit risks in payment and settlement systems. Most systems currently have strict limits on intraday credits. Participants can only have negative positions on the settlement account (towards the system or towards the settlement institution) up to a given limit, which is decided based on credit risk evaluations. When the central bank is the settlement institution, banks are often allowed to use the reserve deposits at the central bank as settlement assets.

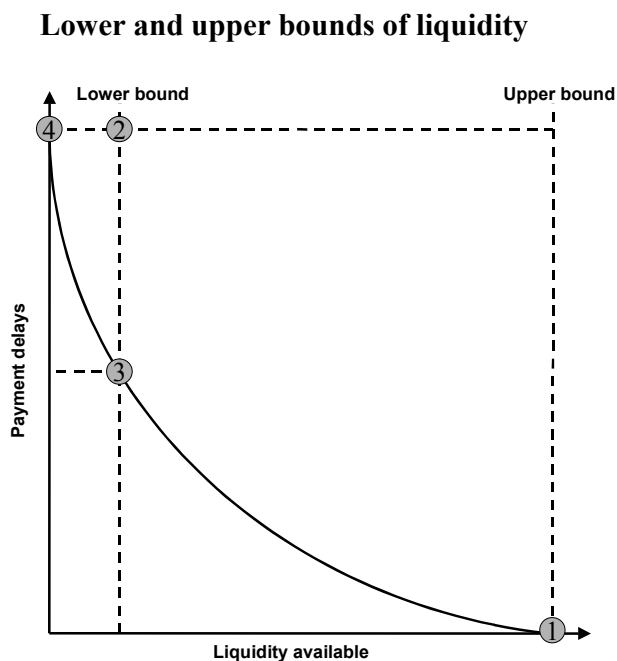
Participants requesting intraday credit often have to put up collateral to secure their limits. These limits can be fully or partly collateralised. Some private systems demand full collateralisation in central bank money and a squaring of accounts when the end-of-day settlement is performed. Some systems, such as the US Fedwire,

charge for intraday overdrafts instead of requiring collateralisation of the positions.

If processing stops when limits are reached, some systems employ liquidity injections or swaps to free the limits. Liquidity can be transferred from another system or another account to the account with insufficient liquidity. Sometimes injections are automated, eg between TBF (the French RTGS system) and RGV (Euroclear France, the securities settlement system).¹¹

To preserve liquidity and/or limit counterparty risk, bilateral limits or filters can be used to reduce the outflow of liquidity. Such a system prioritises payments towards participants depending on the flow in the opposite direction. Maintaining bilateral limits in a system with a large number of participants creates some overhead (the general dimension of the table would be $N*(N-1)$ individual limits).

Figure 2.5



There must be sufficient liquidity in the system to process payments. If every participant has enough funds to settle all transactions when submitted, the liquidity available is at or above the upper bound (point 1). The upper bound is defined as the amount of liquidity above which any additional liquidity will remain unused. When some or all

¹¹ See ECB (2001).

participants have less liquidity than the upper bound, payments cannot be settled immediately and must be delayed. Where this is done by an automatic queuing facility, the trade-off in terms of liquidity and delay is a convex curve as shown in Figure 2.5 above.¹² The more liquidity is reduced, the longer the delays. This is explained by the increase of gridlocks and deadlocks in the system when liquidity is reduced.¹³

The minimum amount of liquidity required to settle all payments submitted during a day is called the lower bound of liquidity. This equals the net amount of incoming and outgoing payments for a bank, or zero if the inflow of funds is higher than the outflow. If all participants in the system have this amount of liquidity available, the system is at point 2 in Figure 2.5 for end-of-day net settlement systems or at point 3 for RTGS systems with an automatic queuing facility. In an end-of-day settlement system, liquidity usage is minimised and delay maximised. In an RTGS system with an automatic queuing facility, the relationship between liquidity needs and delays can be improved, and thus the payment delays for point 3 are significantly lower than for point 2. The RTGS system is, however, very likely to be gridlocked at the end of the day, with some payments remaining unsettled if gridlock resolution is not applied.

If one or more participants have less than the lower bound of liquidity, some payments will necessarily remain unsettled. When liquidity is reduced to zero (point 4), all payments are delayed until the end of the day and remain unsettled unless participants have offsetting positions that can be settled by a partial net settlement algorithm.

The upper bound of liquidity for a system can be determined in the simulator by conducting an RTGS simulation in which all participants have been granted unrestricted intraday credit. The resulting minimum balances will be the upper bound of liquidity required for immediate processing without queuing. The end-of-day balances in this simulation will be the lower bound of liquidity, provided that a netting procedure is applied at the end of the day.

¹² For a more detailed discussion on the trade-off curve, see Koponen and Soramäki (1998) and Leinonen and Soramäki (1999).

¹³ See Bech and Soramäki (2001).

2.4.6 System configuration elements

The features and decision parameters of a particular system configuration can be categorised as follows:

1. Processing and settlement mode (eg real-time or deferred, bilateral or multilateral and gross, hybrid or net settlement).
2. Connections or relationships to other systems (eg ancillary system interfaces).
3. Limits on accounts and net positions.
4. Transaction control features (eg prioritising, queuing, reordering or splitting).
5. Gridlock resolution (eg bilateral off-setting, partial netting or multilateral netting).
6. DVD or PVP support.
7. End-of-day procedures (eg processing of non-settled transactions).
8. Tariff and pricing parameters (eg transaction pricing, fixed tariffs, intraday interest rate and delay costs).

In addition to features described in earlier sections, end-of-day procedures, too, are important to system participants. At the end of the day or the end of a settlement cycle, some transactions may remain in the queues that cannot be settled with the available liquidity. In such cases, there are three options available: discard them, transfer them to the next day or grant/inject additional liquidity to make processing possible. In some cases, the opening hours of the system may be extended to give participants or service providers more time to resolve the situation by getting more liquidity, submitting missing transactions, etc.

In a net settlement system, the lack of liquidity or a credit limit restriction may lead to postponement of the entire net settlement session or implementation of some sort of partial netting algorithm, which can imply unwinding of some of the transaction flow. This, in turn, has implications for the finality of transfers.

In order to cover costs, the services of payment systems are generally priced according to a mix of transaction-based, fixed and start-up tariffs. The settlement service provider can affect the behaviour of participants and the choice of system. For example, if transaction tariffs are high, bunching and the use of alternative routes become relevant issues. High fixed tariffs can motivate smaller banks to pool their resources and cooperate to share fixed costs. These cost elements affect the settlement system structure and market shares over the long run.

Service providers can also price the use of credit. If credit is needed and is costly, then participants have an interest in delaying payments to avoid negative balances for as long as possible. This, however, results in end-of-day congestion when there are no processing rules or customer agreements stipulating faster transfers. To avoid congestion at certain moments or time intervals, the service provider may apply a time-dependent tariff policy.

The simulation technique provides a good opportunity to test the effects of changes in system parameters and see what impact they might have on the liquidity needs of participants, cost components, processing speed, etc.

2.5 The costs of settlement

Existing payment and securities settlement systems involve varying transaction costs depending on the efficiency of the software and hardware systems used, and the available economies of scale. Modern technologies are usually more efficient than old solutions (eg current server hardware outperforms old mainframes). Reduction of manual routines and increased automation also play an important role in reducing operational costs. Operational costs, however, tend to be only a minor part of the total costs of payment and securities settlement system structures. Moreover, operational costs tend to be very similar for different system structures as long as the systems employ efficient technology and are efficiently organised. The interesting cost elements, therefore, are those dependent on the configuration of the system, ie

- a. the cost of liquidity,
- b. the cost of financial risks, and
- c. the costs of delayed transactions.

These cost parameters determine which kind of system design will minimise overall costs, and have an impact on how the participants will submit transactions and at what speed they will be processed in the system. All three types of costs are interrelated and can be traded off against each other by choosing different system designs.

A major determinant of the above costs is the settlement asset or medium used in the system. This can take the following form in a funds transfer system:

- (i) deposits with the settlement institution,
- (ii) credit from the settlement institution, or
- (iii) credits/deposits with other participants.

In the first case, transactions in the system can only be made if enough deposits are available to fund the transaction. In the second alternative, transactions in the system can only be made if the credit lines agreed with the settlement institution facilitate the transfer. In the third option, transactions can be made as long as the credit lines agreed bilaterally with other members of the system are not exceeded. Naturally, a system can use a combination of these options or operate without restrictions on the credits/deposits allowed.

The cost of the settlement medium depends on the alternative investment opportunities. For instance, the cost of liquidity is close to zero in the case of mandatory reserve requirements without alternative investment opportunities. For intraday credits requiring first-class collateral, the costs could be seen as the losses made on keeping low-risk and low-return assets compared to investments with higher yields. In the case of overdrafts with credit risk, there is a need to calculate the costs of taking these credit risks.

In the following sections, we discuss the costs of liquidity, financial risks and delayed transactions associated with the three options for the settlement asset.

2.5.1 Cost of liquidity

Liquidity, in conjunction with payment and securities settlement systems, can be understood as the ability to fulfil one's obligations at a reasonable cost. Liquid assets are assets or claims on other assets that are generally accepted as payment (or assets that can easily be converted into such). Sight deposits, either at the central bank or at commercial banks, are normally the most liquid asset form.

A division is generally made between intraday and overnight liquidity. This division stems from the fact that interest is calculated on the basis of value dates rather than continuously. The cost of both intraday liquidity and overnight liquidity is determined by the central bank. EU central banks currently provide intraday liquidity free of charge, while eg the Federal Reserve charges approximately 0.36%.¹⁴

¹⁴ Federal Reserve Policy Statement on Payments System Risk, January 4, 1999.

Both intraday and overnight credit must be collateralised, which carries an additional cost.

If the system operates on deposits, the cost of intraday liquidity is the opportunity cost of income that would be received by investing the funds held on settlement accounts in assets of equal risk. Such liquidity costs can be close to zero, eg when ample central bank required reserves that would otherwise be idle can be used for settlement purposes. To increase deposits at the settlement institution, participants can borrow funds from each other. While an intraday market could give a market price for intraday funds, no intraday market has yet evolved (probably because intraday liquidity is provided free of interest or at a very low cost by central banks).¹⁵

If the system operates on credit from the settlement institution, the liquidity costs depend on the remuneration and collateralisation required by the settlement institution. While the remuneration cost is direct, collateralisation requirements also pose opportunity costs for the assets pledged as collateral. The laws of supply and demand govern the opportunity costs of collateral. The costs are increased if the list of collateral accepted is short (short supply), and if there are plenty of opportunities to use the collateral to generate profits (eg securities lending) or they are needed elsewhere (increased demand). Further costs of collateralisation may arise if banks are forced to hold inferior portfolios than those that would result from free choice. Moreover, if the list of securities eligible as collateral is short, those on the list may generate lower returns due to their increased liquidity.

If the system operates exclusively on uncollateralised debt relations between the participants, the liquidity costs are zero.

The availability of liquidity can vary, which means the cost of liquidity can also vary. Liquidity costs may also vary seasonally and as the result of general market conditions. Liquidity may become scarce at the end of the day, at the end of the reserve maintenance period or due to exceptional circumstances, when everybody begins to hoard liquidity.

¹⁵ For a discussion on the link between intraday and overnight credit, see Rossi (1995) and Dale and Rossi (1996). For analysis on intraday credit policies by the central bank see Humphrey (1990 and 1996), Furfine and Stehm (1997) and Kahn and Roberds (1998b).

2.5.2 Cost of financial risks

The two main financial risks in payment and securities settlement systems are credit risk and liquidity risk. Credit risk is the risk that a counterparty will never settle an obligation for full value. The party expecting to receive the funds makes up the loss to the principal amount of the transaction.

Liquidity risk can be understood two ways. Firstly, it can be understood from the receiving institution's perspective as the risk that its counterparty fails to settle its obligation for full value when due, but does settle eventually. Secondly, it can be understood from the paying institution's perspective as the risk associated with difficulties in finding the required liquidity on economic terms.

2.5.2.1 Credit risk

If the system operates on deposits at the settlement institution, system participants face the risk that the settlement institution fails with outstanding claims on it.

If the system operates on credit lines from the settlement institution, the situation is reversed, ie the settlement institution faces the risk that a participant fails while in a credit position vis-à-vis the settlement institution. Central banks and other settlement institutions can be exposed to considerable credit risks depending on their credit and collateral policies. The costs of interbank credit risks perceived by banks also depend on official policies towards banks in crises. Currently, most payment and securities settlement systems have internal features limiting maximum credit risks as recommended by the Lamfalussy standards and subsequently the Core Principles for Systemically Important Payment Systems and Recommendations for Securities Settlement Systems.¹⁶

In theory, a system based on credit from the settlement institution could run continuously without periodic settlement of the resulting positions with other claims. Generally, however, systems based on credit from the settlement institution contain a final end-of-day/end-of-settlement-cycle settlement of the credit positions in some safer claim such as central bank money. Many systems have settlement and clearing institutions that can partly take up credit exposures or control collateral or settlement asset accounting.

¹⁶ See BIS (2001) and BIS/IOSCO (2001a, 2001b and 2002) respectively.

If, on the other hand, the system is based on interbank debt relations, the receiving bank has to accept increased liabilities for each transfer received. The liability depends on the transaction finality point towards the customer and towards the other bank.

Interbank liabilities can be distributed among the participants through a loss-sharing agreement. Overall limits can be used to keep the overall credit risks within acceptable bounds. The credit risks can still vary greatly for individual participants during the system's opening hours. A loss-sharing agreement can be used for evening out random peaks closer to average positions. A loss-sharing agreement can be seen as a joint insurance scheme in which everybody takes part in covering the losses using predefined distribution keys.

The payment and securities settlement systems can also be based on a combination of liquidity transfers and credit caps. For instance, a liquidity transfer may be required when the upper limit of the credit cap is reached. An example of such a system is the Finnish POPS system for large-value cheques and express transfers.¹⁷

2.5.2.2 Liquidity risk

In systems that operate on the basis of deposits with (or credit from) the settlement institution, liquidity risks can arise if a participant runs out of funds and is unable to settle transactions as planned. Liquidity risk is therefore closely related to the costs of payment delays and the cost of liquidity. In such a situation, the bank faces the choice of delaying payment until it receives sufficient funds to settle or acquiring the funds from the settlement institution or the market.

There is a further dimension creating liquidity risk that stems from the incentives of system participants to free-ride on the liquidity of other participants. The treasurer of a bank would wish to receive payments early and send them late in order to save liquidity. However, although a delay in settlement reduces the sender's liquidity costs, it increases both its delay costs and the receiver's liquidity costs, as the latter needs to finance its outgoing payments using other means. This creates a dead-weight loss at system level. In an extreme case the system might end up in a situation where the number of payments

¹⁷ See ECB (2001).

submitted is strongly reduced as the participants each await liquidity from others.¹⁸

2.5.2.3 Settlement risk

Settlement risk is a type of credit risk independent of the settlement asset used. Settlement risk in securities and foreign exchange settlement is generally defined as the risk that one party to a transaction will provide the asset it sold but not receive the asset it bought. Settlement risk has both a credit risk and a liquidity risk dimension.

Settlement risk can arise in payment systems if the receiving bank credits the funds on the customer accounts in anticipation of receiving them from the sending bank. If the sending bank goes bankrupt, the credit risk is borne by the receiving bank. Such settlement risks can arise in systems with any of the three types of settlement asset.

In securities and foreign exchange trading, the party that has first made the payment for one leg of the transaction faces the possibility that its counterparty may not deliver the other asset when due. If this happens, it must finance the shortfall until the counterparty honours its obligation (liquidity risk). The party paying first also faces a risk that the counterparty may fail to complete the second leg of the transaction. Thus, it is exposed to liability for the full amount of the transaction (credit risk).

Settlement risk can also arise from legal uncertainty (eg a situation where it is unclear if the rules of netting will be accepted by the courts).

2.5.2.4 Systemic risk

In the context of payment and securities settlement systems, systemic risk refers to the risk that the failure of one participant in a system to meet its required obligations will cause other participants to be unable

¹⁸ On bank incentives in payment systems see Kahn and Roberds (1998a). A game theoretical model on bank behaviour under different credit policies of the central bank has been developed in Bech and Garrat (2003). For a discussion on the deadweight losses and other externalities in payment systems see Schoenmaker (1993) and Angelini (1998). For a model on total costs (liquidity and risk) in net and gross systems see Schoenmaker (1995).

to meet their obligations when due. Notably, similar systemic risks can be caused by the failure of the settlement institution itself.¹⁹

The cost of a systemic disturbance can be high. The chain reaction may expand into an overall systemic crisis and can jeopardise the operation of the entire financial system and ultimately the real economy. Central banks have been concerned to limit systemic risk, eg by issuing the Core Principles for Systemically Important Payment Systems.²⁰

Because of efforts to reduce risks in interbank payment systems, the likelihood of a chain reaction caused by exposures in these systems seems currently to be relatively low. This at least is the outcome of studies for Finnish, Danish and Canadian interbank payment systems.²¹ Blåvarg and Nimander (2002) find similar results for payment system exposures in Sweden, but point out that the systemic risk comes mainly from foreign exchange exposures. The introduction of CLS substantially reduced globally the systemic risks stemming from foreign exchange settlement.

2.5.3 Costs of delayed payments

The speed of transaction settlement and processing is a critical element in settlement system costs. If payments could be postponed without costs, nobody would have an interest in settlement or in providing liquidity. The delay costs are generally determined based on a given time limit. Exceeding the time limit implies such high costs that, in most cases, parties have an incentive to avoid exceeding the time limit. These costs may be explicit fines for delays, but it is often intangible values (eg expected service quality) that make up delay costs. Because there is generally no special gain in paying/settling too early, transactions are often transmitted close to the time limit. This is especially the case in systems where settlement funds are costly. In systems with idle and low-cost settlement assets, continuous transaction flows are often preferred to avoid operational congestion. Interbank payment and securities settlement systems have in the past

¹⁹ For general studies on systemic risk stemming from netting systems see BIS (1989), McAndrews and Wasilyew (1995), Angelini et al (1996) and Borio and Van den Bergh (1993). For a comparison on gross and net settlement systems see Freixas and Parigi (1998). For a survey on different concepts of systemic risk see de Bandt and Hartmann (2000).

²⁰ See BIS (2001).

²¹ See Kuussaari (1996), Bech et al (2002) and Northcott (2002) respectively.

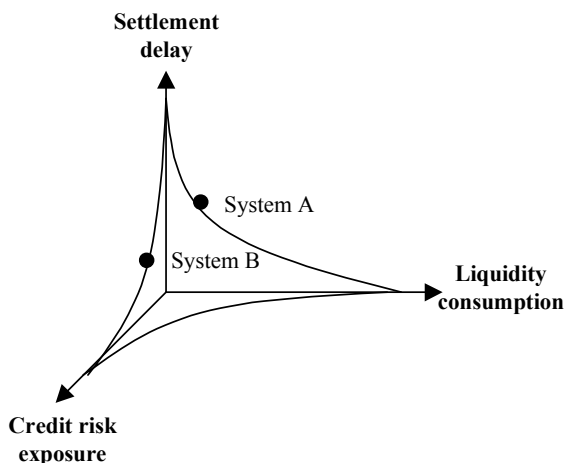
generally shown day-based timetables in the form of T+1, T+2 or even T+5 settlement. The speed is currently improving with more true real-time based systems and batch systems processing with many settlement cycles during the same day. The priority/importance of the transaction also affects the desired processing speed. Additional liquidity costs are acceptable for urgent payments, while the processing of less urgent transactions can be postponed until later. The trend towards enhanced speed also increases the share of immediately settled and processed transactions.

2.5.4 Combined costs

Liquidity requirements, financial risks and payment delays both mirror the transaction flow, but from opposite angles.

By choosing payment and securities settlement system structures and processing features, system participants/providers can attempt to identify the optimal point as a balance between the three main objective variables as depicted in Figure 2.6.

Figure 2.6 **The variables in the objective function**



Generally, the system structure is more efficient when it is closer to the origin. Different participants may have different utility functions and different views on the relative weights of the variables. In general, increasing liquidity consumption, credit risk exposure, or both, can increase settlement speed. Different methods for smoothing and

ordering payment flows, gridlock resolution and netting algorithms can be used for reducing liquidity consumption and/or credit risk exposure with a moderate amount of increased settlement delay. Liquidity consumption and credit risk exposures can be fixed if the transaction flows have to be settled at once when transactions are submitted without an opportunity to affect the flow and depart from the FIFO processing order. When transactions can be rearranged and a certain amount of delay is allowed, the transaction flow can be smoothed, thereby reducing liquidity or credit risk variations. The objective of netting, splitting, or reordering by prioritising features is to smooth transaction flows.

Payment and settlement systems contain varying degrees of risk. Simulations can be used to ascertain risk information unavailable in current system statistics. For example, intraday exposures by individual participants and/or the entire system can be studied. The systemic impacts of failures of large participants in the system can be evaluated and the consequences of large breakdowns examined. In complex environments, it is particularly difficult to foresee all the consequences without simulating possible situations.

Appendix 1

An overview of the BoF-PSS2 payment and settlement system simulator

Background

In the mid-1990s, a payment system simulator was developed within the Bank of Finland. It was found to be a valuable tool for studying the probable effects of the introduction of EMU on Finnish payment systems. Some other countries also used this first BoF simulator for other types of analysis, such as queuing and algorithm studies, and clarification of risk and liquidity issues. Due to its popularity, a more efficient, user-friendly and comprehensive version was built in 2003. BoF-PSS2 is now available for research purposes at no charge and with complete documentation.

Basic features of the simulator

BoF-PSS2 is intended for independent use by payment and settlement systems designers, administrators, analysts at central banks and financial institutions, research institutions and academics. The simulation software is downloaded to the user's computing environment (typically a PC) for local processing. The download includes user documentation. The Bank of Finland will arrange, from time to time, simulation seminars and will provide limited start-up support. User support, if needed, is available from a software company. This support is priced according to said company's normal pricing policy.

The BoF-PSS2 simulator is a tool for conducting a variety of payment system analyses. The basic principle is that given payment flows are processed in a given model of an existing or contemplated payment and settlement system structure. The simulator thus models settlement processes for a specified payment system environment. The simulation results are account bookings and account balances made according to the rules defined for the payment system environment. The transaction processing output can then be analysed with an included analyser tool or exported into other programs such as Excel. Typical areas of interest include intersystem credit risk, liquidity

consumption and risks, settlement speed, gridlock resolution and settlement efficiency.

The simulation process starts with the creation of the payment or transaction flow(s) to be processed and defining of the payment/settlement system environment and rules (eg systems, participants, limits and settlement rules). Next the simulation is run. The results are then compared against other simulation runs or real-life observations. Using the simulation tool is essentially an iterative learning process, ie earlier simulation runs become the basis for improving and refining subsequent simulation runs. Thus, the simulator is not a deterministic econometric optimisation model, but rather a heuristic tool for analysing systems that are too complex for deterministic models. The BoF-PSS2 can process in one simulation several million transactions to be booked on several thousand accounts in several interlinked systems. In other words, optimisation analysis is mainly done via a 'trial-and-error' or enumeration process. The simulation runs are repeated for different values on the decision-making parameters and the resulting objective values compared for the different combinations. For instance, a central bank might want to determine the bank-specific minimum liquidity required for guaranteeing that in 99.5% of all cases continuous settlement during the day is maintained so that transactions are not queued for longer than ten minutes. To find the answer, the bank would model various levels of liquidity to see which levels fulfil the objective. The model is thus a workbench for testing alternatives.

The simulator's basic features are:

1. **Input data import and export tool.** All input data must be presented in comma-separated values (CSV) format. The importer supports free ordering of the data fields and ensures that the imported data are formatted correctly. Basic data validations are carried out and key fields matched against each other. For example, participant keys in transaction data must match those for participant (account) data.
2. **Transaction submission algorithms** determine which transactions are submitted to the system and when they are submitted. The submission algorithm can be used to introduce rule-based user behaviour into a model (eg early submission of large volumes of low-priority payments only allowed for participants from which reciprocal payments have been received).

3. **Transaction processing algorithms** for simulating real-time gross settlement (RTGS) systems with gridlock resolution and other optimisation features, deferred net settlement (DNS) systems and continuous net settlement (CNS) systems.
4. **Linking of different systems:** ie payments debited from one system can be credited to an account in another system, or settlement totals as calculated in an ancillary system can be settled in another system. The first option gives the possibility to simulate a network of interlinked systems (eg TARGET links together 16 different RTGS systems). The second option gives the user the possibility to model payment structures consisting of main and ancillary systems.
5. **Handling of multiple accounts per participant and multiple currencies in each system** can be simulated, as well as simulation of multi-currency and securities settlement systems. In the simulator, settlement of each transaction results in book entries on accounts. There is no difference in the basic process with respect to currency or securities type. The meaningfulness of transactions must thus be ensured in the payment data (eg so that a debit is not made to a EUR account when the corresponding credit goes to a GBP account).
6. **Settling two-leg transactions** where settlement of one leg is conditional upon the settlement of the other leg. This enables simulations of payment-versus-payment (PVP) systems in foreign exchange settlement and delivery-versus-payment (DVP) systems in securities settlement.
7. **An output exporter and analyser** for a given set of basic statistical parameters and an output exporter for transferring the results to external analysis programs (eg Excel) for detailed analysis.

BoF-PSS2 contains basic submission and transaction processing algorithms adequate for most common simulation needs. The model also allows for user customisation via external algorithm interfaces. Users or third parties can program additional algorithms to incorporate in the simulator. Custom algorithms may be submitted to the BoF-PSS2 general library, from where they can be retrieved by other users.

The main difference between the simulator and an actual payment system is that the time function is event-driven and not real-time.

Processing is performed transaction by transaction. The simulator generally operates faster than the real world, but processing speed depends on the volumes, processing complexity and available processing capacity. In RTGS test cases, the simulator has processed and booked about two million transactions in one hour. In any case, the end result is the same as if the simulator had run in real time.

Simulations can be carried out for separate days or for sequences of days. In the latter case, closing balances and unsettled transactions can be carried over to the next day or next settlement occasion/period. The simulator uses the standard calendar and assumes all days (even weekends) are banking days. However, if there are no transactions for a given day the account balances will remain stable until payment flows are detected on a subsequent banking day.

For the sake of manageability, each feature has been implemented in a standardised way (in real systems, features can be implemented in many ways). Thus, the user must transform the input data according to the convention used in the simulator (eg DVP and PVP transactions are matched based on a given individual code field and a code has to be given by the user for such transactions if the matching is performed using other conventions). BoF-PSS2 is limited purely to payment and transaction processing, and the pledge type of collateral aspects of payment systems are not explicitly included. However, repo-type collateral processing could be included through DVP repo transactions, especially if the simulation in question contains a securities settlement system. Both repo- and pool-based collateral requirements and processes might also be introduced into the simulator by defining special collateral systems, which would store the collateral balances and would have DVP links with the RTGS system(s).

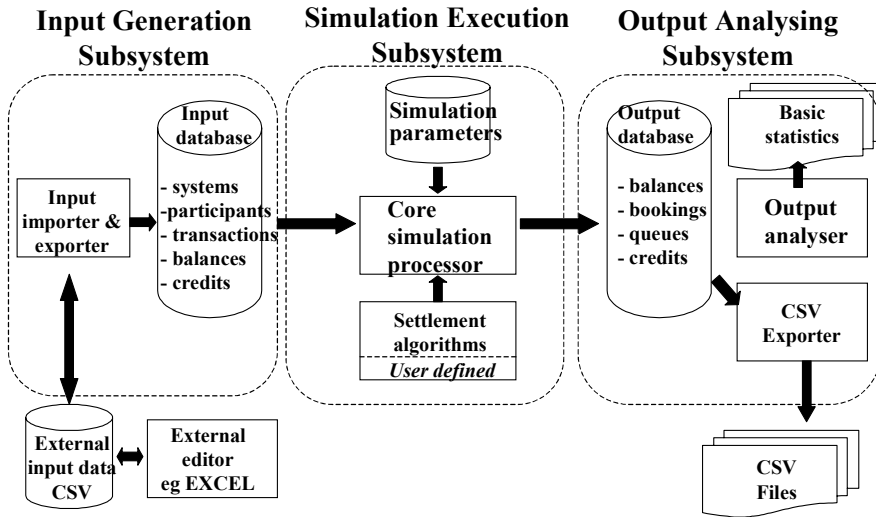
System structure

The BoF-PSS2 system structure consists of three main subsystems:

- a) Input Generation Subsystem,
- b) Simulation Execution Subsystem, and
- c) Output Processing Subsystem.

Figure A2.1

BoF-PSS2 system structure



The Input Generation Subsystem contains tools for importing and validating transaction data, participant data, as well as data on daily account balances and credit limits. All data are stored in database files. The importer checks that the input data are formally valid and transfers them into system database structures. The correctness of the input data is vital. Account numbers in the transaction file must correspond with those in the participant (account) data. While all data must be presented in CSV format, it can be entered in any user-defined order. The user defines templates to describe the CSV files and match the input files with the database structure. Data in the database can also be edited by exporting them as CSV files into eg Excel and then importing them after the changes are made. The simulator does not include a proprietary editor for this purpose.

The Simulation Execution Subsystem includes tools for configuring and running simulations. It also contains the actual simulation and settlement logic. This subsystem keeps a log of all events and bookings and makes reports and statistics on the simulation runs. A configurator/executor tool facilitates configuring and execution of simulation runs.

The Output Processing Subsystem includes the functionality for reporting basic statistics on the most common result parameters. The output database contains the raw data of the booking order of the transactions and the balances on the settlement accounts. The input database contains the transaction flow, while the output database

contains the settlement flow (the settlement order and timing of the submitted transactions). The analyser tool is used for generating additional reports and transferring selected statistics to CSV files for later use. Users typically perform many simulations and then compare the results. The analyser tool provides some basic reports and comparison possibilities, but advanced analyses need to be made by exporting CSV files into eg Excel for further analysis. Before running a simulation, it is advisable to create a structure for the simulation runs and determine how the results will be stored in databases for further analysis. Databases can be overwhelmed when transaction volumes are very high and all transaction-level events are retained in the databases.

Input data

System data defines the systems in the simulation. A large number of systems, each with individual properties, can be included. For example, many RTGS, CNS and DNS systems can run in parallel in the same simulation. There are no specific limitations on the number of systems. The system data contains the information on the features implemented in these systems (eg queuing method, end-of-day conventions, netting algorithms used, opening hours and net settlement timing). System-level input data are defined, due to their complexity, through a separate input screen and not through the CSV format as used for the other input data tables.

Participant (account) data contains information about the participants (accounts) in each simulated system.

Balance data contains the daily initial account balances for each account in the systems. Another way to introduce initial balances is through payments from central bank accounts to the participants' accounts. When simulations for a sequence of days are performed, end-of-day balances for the previous day can be carried forward to the next day or defined separately for each day.

Intraday credit limit data contains information about the intraday credit limits of the accounts and changes to them during the day. In simulations with a sequence of days, credit limit changes remain in force until the next change.

Transaction data includes all information about individual transactions. Such transactions can be payments or transfers of securities, and their properties include the value of the transaction, sender and sending account, and receiver and receiving account. Five user-defined fields in transaction data can be used to carry information

that can be used in user-defined submission and transaction processing algorithms (eg priority codes).

Output data

Simulation results are written to an output database organised by levels (simulation, system, participant and event levels).

The simulation level contains the general data of the simulation, eg date and time, input database and a description given by the user.

The system level contains the overall statistics for each system and transaction and account totals.

The participant level gives totals and averages for each participant (account). These can be used to check the impact of simulated ‘what if’ scenarios on individual participants or participant groups.

The event level data make up the main bulk of the output database and constitute a transaction level log of everything the simulator processed. Submission, booking and queuing events are reported separately. Gridlock resolution events or violation events (eg overdrawn limits) are reported separately. User-defined modules can also write comments for later analysis. To limit the output data, the user can select data to be retained or reported for each simulation run.

There are ready-made basic statistics reports on the system and participant (account) levels. There is also a basic reporting tool for comparing simulations. Because the output data are so vast and users have such diverse needs, a general output exporter has been created. The user can select interesting output data and export them as CSV files for additional analysis in Excel and other applications that support CSV files. **The user defines templates to describe the output CSV files, and selection criteria to select the data content.**

Overview of settlement capabilities

The simulator identifies three general types of system: RTGS (real-time gross settlement), CNS (continuous net settlement), and DNS (deferred net settlement). Table A2.1 provides an overview of the current features available for the various systems. The palette will probably expand as users and the Bank of Finland itself create new modules.

Table A2.1

Current available settlement palette

	RTGS	CNS	DNS
Queuing	Based on available liquidity	Based on available liquidity	Based on credit limits
Gross settlement	FIFO or Bypass-FIFO and priority	FIFO or Bypass-FIFO and priority	
Bilateral offsetting of reciprocal payment	As long as FIFO rule is applied	As long as FIFO rule is applied	
Net settlement	Bilateral, partial or multilateral	Bilateral, partial or multilateral	Bilateral, partial or multilateral
Splitting	Maximum value or available liquidity	Maximum value or available liquidity	
Forced EOD settlement	At specified settlement occasions	At specified settlement occasions	At specified settlement occasions
Transfers to next day	Queued payments and balances	Queued payments	Queued payments
Liquidity injections	Given amount or as % of limit	Given amount or as % of limit	
DvP and PvP	Within and between systems	Within and between systems	Within systems

The user can introduce user-defined modules that contain eg settlement conventions that are not currently supplied as ready-made algorithms.

Hardware and system requirements

The simulator software is distributed online and contains an automated installation package.

The minimum hardware requirements are a PC with an Intel Pentium 4 class processor and at least 256 MB of RAM (main memory). For large simulations, at least 512 MB of RAM is recommended.

The system can be installed and run on Windows NT/2000 or Windows XP. Although untested, it should work with modifications on Linux operating systems with a compliant installation of Sun Microsystems's Java Runtime Environment (JRE) version 1.3.

The simulator requires installation of the MySQL database. It is also recommended that Microsoft Excel be available.

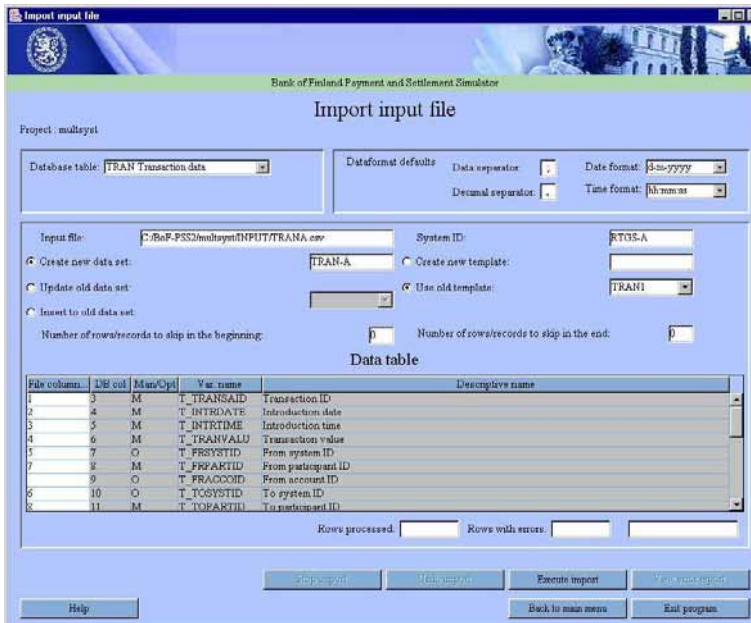
Ordering and further information

Further information (PowerPoint presentations, user manual, installation guide, database descriptions, etc) on the simulator can be found at www.bof.fi/sc/bof-pss.

The simulator can be ordered by sending a fax to the Bank of Finland as described on the web page.

Appendix 2

BoF-PSS2 screen shots



Define system data

Bank of Finland Payment and Settlement Simulator

System control data specification/modification

Project: p_test

System ID: RTGS-M

Modify old system data set Create new system data set BAS1 Copy from old system data set

System full name: Main RTGS Opens (hh:mm 24h): 0700 System description: Basic version with queuing and netting every 30 min
 System type: RTGS Closes: 1900
 System acronym:

System description may be used for output purposes (optional).

Transfer of balances
 Transfer balances to next day

Intraday credit availability
 Credits according to limit table, or
 No credits available, or
 Credit available without limits

Handling of unsettled transactions
 Transfer unsettled transactions to next day/settlement occasion or
 Delete unsettled transactions or
 Force end-of-day settlement of unsettled transactions

Potential algorithms: select one and add it to the attached algorithms by pressing the 'Add algorithm' button. ENT and END algorithms are mandatory.

MNSETTLI MNS interval:Minutes interval (1-60)

Attached algorithms: Changes of order by dragging/dropping. Deletion by striking keyboard 'delete' button for active algorithms.

Algorithm	Algorithm type	Double click row to insert parameter values
ENBASIC1	ENT	
SEBASIC1	SET	
ENDRTGS1	END	
QUFIPOPR	QUE	
MNSETTLI	MNS	interval=30,Minutes interval (1-60)

View data sets

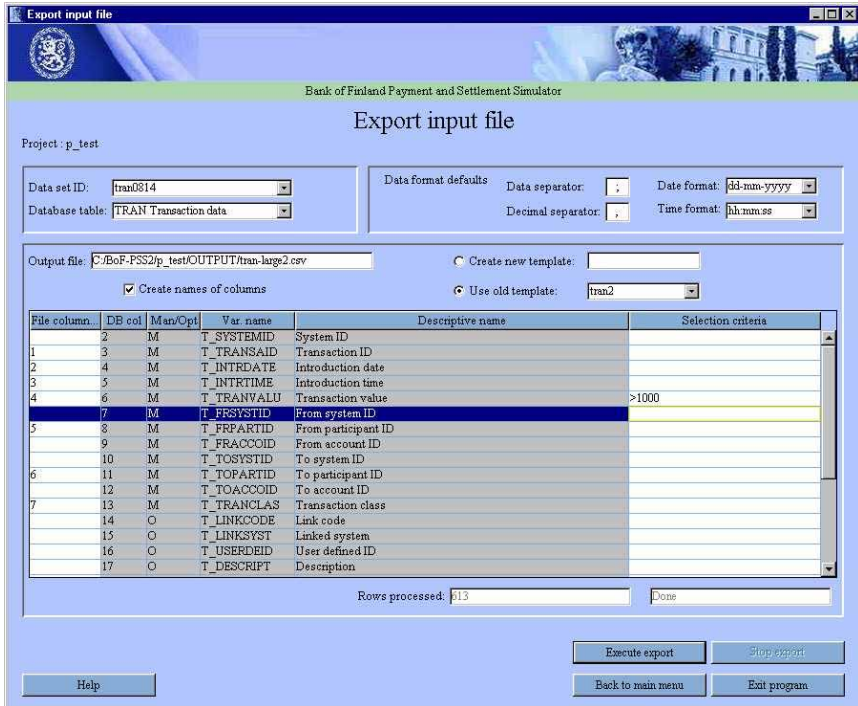
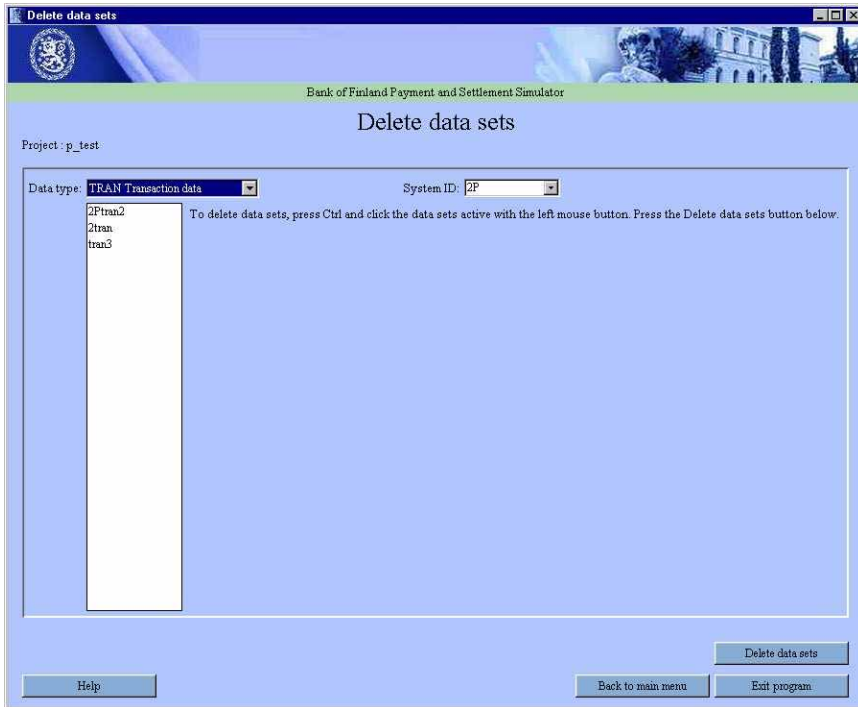
Bank of Finland Payment and Settlement Simulator

View data sets

Project: p_test

Data type: TRAN Transaction data System ID: S-RTGS Data set ID: tran0814

Transaction ID	Introduction date	Introduction time	Transaction value	From system ID	From participant ID	From account ID	To sy:
101	2003-05-12	07:01:00	4880384,39	S-RTGS	1		S-RTGS
102	2003-05-12	07:01:00	683826,60	S-RTGS	1		S-RTGS
103	2003-05-12	07:01:00	2639630,73	S-RTGS	1		S-RTGS
104	2003-05-12	07:01:00	46064,41	S-RTGS	1		S-RTGS
105	2003-05-12	07:01:00	633795,99	S-RTGS	1		S-RTGS
106	2003-05-12	07:01:00	79702,04	S-RTGS	1		S-RTGS
107	2003-05-12	07:01:00	2786,33	S-RTGS	1		S-RTGS
108	2003-05-12	07:01:00	8000,03	S-RTGS	1		S-RTGS
109	2003-05-12	07:01:00	153492,54	S-RTGS	5		S-RTGS
110	2003-05-12	07:01:00	83989,24	S-RTGS	5		S-RTGS
111	2003-05-12	07:01:00	779471,41	S-RTGS	5		S-RTGS
112	2003-05-12	07:01:00	6678,92	S-RTGS	5		S-RTGS
113	2003-05-12	07:01:00	49921,70	S-RTGS	5		S-RTGS
114	2003-05-12	07:01:00	17337,06	S-RTGS	5		S-RTGS
115	2003-05-12	07:01:00	300182,44	S-RTGS	5		S-RTGS
116	2003-05-12	07:01:00	18950,49	S-RTGS	5		S-RTGS
117	2003-05-12	07:01:00	1634545,51	S-RTGS	6		S-RTGS
118	2003-05-12	07:01:00	408639,34	S-RTGS	6		S-RTGS
119	2003-05-12	07:01:00	1856551,99	S-RTGS	6		S-RTGS
120	2003-05-12	07:01:00	184378,62	S-RTGS	6		S-RTGS
121	2003-05-12	07:01:00	16723,85	S-RTGS	6		S-RTGS
122	2003-05-12	07:01:00	12084,40	S-RTGS	6		S-RTGS
123	2003-05-12	07:01:00	57960,37	S-RTGS	6		S-RTGS
124	2003-05-12	07:01:00	4920425,27	S-RTGS	14		S-RTGS
125	2003-05-12	07:01:00	1826540,68	S-RTGS	14		S-RTGS



Simulation configuration

Bank of Finland Payment and Settlement Simulator

Simulation configuration

Project : multisyst

Modify old simulation ID
 Create new simulation ID
 Copy from old simulation ID

Simulation name:

Description:

Sub algorithm:

For defining data sets for the simulation select first the right system ID in the corresponding drop-down box above the table and then desired data sets after which pressing the add/update data set button will move the new data to the data table. A system can be deleted from the table by selecting the system ID and pressing Possible parameters for the submission algorithm.

Data sets to simulation:

System id	System data set	Transaction data set	Participant data set	Credit limit data set	Balance data set
A	ISYS	TRAN-A	ABANKS		
B	ISYS	TRAN-B	BBANKS		
C	ISYS	TRAN-C	CBANKS		

Number of errors:

Simulation execution

Bank of Finland Payment and Settlement Simulator

Simulation execution

Project : p_test

Modify old simulation batch ID
 Create new simulation batch ID
 Copy from old simulation batch ID

Select a simulation ID and add it to the table by pressing the add simulation ID button. A simulation can be deleted from the batch by selecting the simulation ID and by pressing thereafter the Delete button.

Simulation ID

Simulation id	Simulation name	System level stat. SYLS	Account stat. ACST	Transact event stat. TEST	Netting event stat. NEST	Account violation stat. AVST	Booing event stat. BEST	Unsett. transect. stat. UNST	Submitt. transect. stat. SUST	Queued transect. stat. QUST	Comment transect. stat. CTST	Comment intraday credit stat. CCST
2Pjono		<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2Pque1		<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2Pque3		<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

View simulation logs

Bank of Finland Payment and Settlement Simulator

View simulation logs

Project: p_test

Simulation ID	Simulation name	Date/time	Duration	SVLS	ACST	TEST	NEST	AVST	BEST
2Fjono		14-07-2003/11 21:35	00:00:00	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2Fairall		14-07-2003/01 45:11	00:00:00	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2Fairallc		14-07-2003/02:03:53	00:00:00	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2Fque1		15-07-2003/12:21:09	00:00:00	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2Fque3		15-07-2003/01 12:05	00:00:00	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2Fjono2	toinen yrittämä	14-07-2003/11 24:24	00:00:00	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Help Back to main menu Exit program

Basic statistics reports

Bank of Finland Payment and Settlement Simulator

Basic statistics reports

Project: p_test

Simulation ID: 2Fjono2 Simulation name: toinen yrittämä Output file: C:/BoF-FSS2/p_test/OUTPUT_REPORTS/Systemstatistics_2Fjono2.csv

Choose one of the report options:

System statistics report

Account statistics report System ID:

System time series report System ID: Reporting interval per minutes

Account time series report System ID: Reporting interval per minutes

Participant ID: Account ID:

Account ID:

Save in csv-file Save and open csv-file

Help Back to main menu Exit program

Account comparison analyser

Bank of Finland Payment and Settlement Simulator

Account comparison analyser

Project: p_test

Modify old comparison view
 Create new comparison view
 Copy from old comparison view

Output file:
 Difference from the first selected simulation Show percentage Show absolute

Select from the drop-down list the simulation ID and add it to the comparison by pressing the add to comparison button. When all simulations are selected, select a system ID from the system ID drop-down list.

2Pque3	Add to comparison	2P
2Pbalnc		
2Pque1	Delete	

Select from the drop-down list the account statistics field to report and add it to the comparison by pressing the add to comparison button.

Queue stop time	Add to comparison
Average queue value	
Maximum queue value	Delete
Average queue length	
Average time of settlement	
Queue stop time	

System comparison analyser

Bank of Finland Payment and Settlement Simulator

System comparison analyser

Project: p_test

Modify old comparison view
 Create new comparison view
 Copy from old comparison view

Output file:
 Difference from the first selected simulation Show percentage Show absolute

Select from the drop-down list the simulation ID and add it to the comparison by pressing the add to comparison button. When all simulations are selected, select a system ID to the comparison.

2Pque3	Add to comparison	2P	Add to comparison
2Pbalnc			
2Pque1	Delete	2P	Delete
2Pque3			

Select from the drop-down list the account statistics field to report and add it to the comparison by pressing the add to comparison button.

Value settled	Add to comparison
Liquidity available	
Absolute credit limit usage	Delete
Average credit limit	
Value settled	

Bank of Finland Payment and Settlement Simulator

Delete output data

Project: p_test

To delete output data tables, click them active and then press the Delete output data button.

Remove entire simulation	Simulation...	Simulation name	System level stat. SYLS	Account stat. ACST	Transact. event stat. TEST	Netting event stat. NEST	Account violation stat. AVST	Booking event stat. BEST	Unsettled transact. stat. UNST	Submitted transact. stat. SUST	Queued transact. stat. QUST	Comment transact. stat. CTST	Comment intraday credit stat. CCST
<input type="checkbox"/>	2Piono		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	2Piono2	toinen yritysää	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	2Painall		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	2Pbalnc		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	2Pque1		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	2Pque3		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Buttons: Help, Delete output data, Back to main menu, Exit program.

Bank of Finland Payment and Settlement Simulator

Export output file

Project: p_test

Simulation ID: 2Painall
 Database table: TEST Transaction event statistics

Data format defaults: Data separator: . Date format: dd-mm-yyyy
 Decimal separator: , Time format: hh.mm.ss

Output file: C:/BoF-PSS2/p_test/OUTPUT/inban-trans.csv
 Create names of columns
 Create new template:
 Use old template: basevent

File column	DB col	Var name	Descriptive name	Selection criteria
2	E_SYSTEMID	System ID		
1	3	E_TRANSID	Transaction ID	
4	4	E_DATSETID	Data set ID	
3	5	E_INTRDATE	Introduction date	
4	6	E_INTRTIME	Introduction time	
5	7	E_TRANVALU	Transaction value	
6	8	E_FRSYSTID	From system ID	
7	9	E_FRPARTID	From participant ID	=12=14
10	E_FRACCOID	From account ID		
8	11	E_TOGYSYID	To system ID	
9	12	E_TOPARTID	To participant ID	
13	E_TOACCOID	To account ID		
14	E_TRANCLAS	Transaction class		
15	E_LINKCODE	Link code		
16	E_LINKSYST	Linked system		
17	E_USERDEID	User defined ID		

Rows processed:

Buttons: Help, Execute export, Stop export, Back to main menu, Exit program.

Initial specifications

Bank of Finland Payment and Settlement Simulator

Version: 0.9.0 / 2003-06-18
License: Bank of Finland
Distribution:

Initial specifications

Create new project:
 Modify old project:

Databases

The simulator will create a new input and output database for each project in the indicated default directory. The user can change directory and also specify an already existing input and output database when desired. However, there can only be one input and output database assigned to the same project at one time.

Input database:
 Output database:

Default directories

The simulator will use indicated default directories for the different file types. Each project is assumed to have its own default directories. The user can change the defaults and also use common directories.

Default directory for input files:
 Default directory for error list:
 Default directory for output files:
 Default directory for output reports:

User module definitions

Bank of Finland Payment and Settlement Simulator

User module definition

Project: p_test

User module name:
 User module file:
 User module type:

Parameters:
 Checking rules:

BoF-PSS2 Help

Hide Back Forward Print

Contents | Index | Search | Favorites

BoF-PSS2 User Manual

- 1. Introduction
 - 1.1 General overview
 - 1.2 Supported system structures and...
- 2. Installation
- 3. Operating the BoF-PSS2 simulator
 - 3.1 Short description of BoF-PSS2
 - 3.2 Main menu
 - 3.3 Working with projects
 - 3.4 Importing data
 - 3.5 Setting up a payment and settle...
 - 3.6 Viewing input data
 - 3.7 Deleting input data
 - 3.8 Exporting input data
 - 3.9 Setting up simulation runs
 - 3.10 Executing simulations
 - 3.11 Analysing results
- 4. Algorithms and user modules
- 5. Data content and databases
- 6. Application screens
- 7. Technical documentation
- 8. Trouble shooting guide

1.1 General overview

The BoF-PSS2 system structure contains three main subsystems:

- a) [Input generation subsystem](#)
- b) [Simulation execution subsystem](#)
- c) [Output analysing subsystem](#)

The architecture of the PSS2 program is pictured below:

The diagram illustrates the architecture of the BoF-PSS2 Simulator on a Workstation. It is divided into three main subsystems: Input, Execution, and Output. The Input subsystem receives 'Import data' (Exported CSV-files and Input data CSV-files) and stores it in an 'Input database'. The Execution subsystem uses the 'Input database' and an 'Output database' to perform simulations. The Output subsystem then processes the results into 'Output analysing data' (Excel files and Simulation output CSV-files). Below the main subsystems, there are two management data sections: 'Project management data' (Error files, Output reports, System database) and 'System management data' (Temp directory, User module directory). Bidirectional arrows indicate the flow of data between the subsystems and the management data sections.

References

- Angelini, P, Maresca, G and Russo, D (1996) **Systemic Risk in the Netting System**. *Journal of Banking and Finance* 20, 853–868.
- Angelini, P (1998) **An Analysis of Competitive Externalities in Gross Settlement Systems**. *Journal of Banking and Finance* 22, 1–18.
- Bech, M and Garrat, R (2003) **The Intraday Liquidity Management Game**. *Journal of Economic Theory* 109, No. 2 (April): 198-219.
- Bech, M, Madsen, B and Natorp, L (2002) **Systemic Risk in the Danish Interbank Netting System**. Working Paper 8/2002, Danmarks Nationalbank.
- Bech, M and Soramäki, K (2001) **Gridlock Resolution in Payment Systems**. *Monetary Review* 4/2001, Danmarks Nationalbank.
- Bech, M and Soramäki, K (2002) **Liquidity, Gridlocks and Bank Failures in Large Value Payment Systems**. *E-Money and Payment Systems Review*, Central Banking Publications, 111–126.
- BIS (1989), **Report on Netting Schemes (Angell Report)**, Group of Experts on Payment Systems of the Central Banks of the Group of Ten Countries, Bank for International Settlements, Basle.
- BIS (1990) **Large-value Funds Transfer Systems in the Group of Ten Countries**. Group of Experts on Payment Systems of the Central Banks of the Group of Ten Countries, Bank for International Settlements, Basle.
- BIS (1995) **Cross-border Securities Settlement**. Committee of Payments and Settlement Systems of the Central Banks of the Group of Ten Countries, Bank for International Settlements, Basle.
- BIS (1997) **Real-time Gross Settlement Systems**. Prepared by the Committee on Payment and Settlement Systems of the Central Banks of the Group of Ten Countries, Bank for International Settlements, Basle.

- BIS (2001) **Core Principles for Systemically Important Payment Systems**. Committee on Payment and Settlement Systems of the Central Banks of the Group of Ten Countries, Bank for International Settlements, Basle.
- BIS/IOSCO (2001) **Securities lending transactions: market development and implications**. Committee on Payment and Settlement Systems, Technical Committee of the International Organisation of Securities Commissions, Basle.
- BIS/IOSCO (2001) **Recommendations for securities settlement systems**. Committee on Payment and Settlement Systems, Technical Committee of the International Organisation of Securities Commissions, Basle.
- BIS/IOSCO (2002) **Assessment methodology for 'Recommendations for Securities Settlement Systems'**. Committee of Payments and Settlement Systems, Technical Committee of the International Organisation of Securities Commissions, Basle.
- BIS (2003) **The role of central bank money in payment systems**. Committee on Payment and Settlement Systems, Technical Committee of the International Organisation of Securities Commissions, Basle.
- Blåvarg, M and Nimander, P (2002) **Interbank exposures and systemic risk**. Penning och valutapolitik, No. 2/2002, Sveriges Riksbank.
- Borio, C, Russo, D and Van den Bergh, P (1992) **Payment System Arrangements and Related Policy Issues: A Cross Country Comparison**. SUERF Papers on Monetary Policy and Financial Systems 13, Société Universitaire Européenne de Recherches Financières, Tilburg.
- Borio, C and Van den Bergh, P (1993) **The nature and management of payment system risks: an international perspective**. Economic Papers 36, Bank for International Settlements.
- Bowman, M. (1995) **Simulation Modelling of Large Value Transfer Systems**. Paper prepared for Eurobanking 1995, Association for Payment Clearing Services.

- Dale, S and Rossi, M (1996) **A Market for Intra-day Funds: Does it have Implications for Monetary Policy?** Bank of England.
- De Bandt, O and Hartmann, P (2000) **Systemic risk – a Survey.** Discussion Paper No. 2634, Centre for Economic Policy Research.
- ECB (2001) **Payment and securities settlement systems in the European Union.** (the Blue Book), European Central Bank.
- ECB (2002) **Payment and securities settlement systems in accession countries.** European Central Bank.
- ECB (2003) **TARGET Annual Report 2002.** European Central Bank.
- EU Commission (2001) **The Giovanni Group report Cross-Border Clearing and Settlement Arrangements in the European Union.** November 2001, Brussels.
- EU Commission (2002) communication, **Clearing and settlement in the European Union – Main policy issues and future challenges.** May 28, 2002, Brussels.
- EU Commission (2003) **The Giovanni Group Second Report on EU Clearing and Settlement Arrangements.** April 2003, Brussels.
- Freixas, X and Parigi, B (1998) **Contagion and Efficiency in Gross and Net Interbank Payment Systems.** Journal of Financial Intermediation 7, 3–31.
- Furfine, C and Stehm, J (1997) **Analysing Alternative Intraday Credit Policies in Real-Time Gross Settlement Systems.** Finance and Economics Discussion Series 1997-40, Board of Governors of the Federal Reserve System (US).
- Ganz, M, Günzter, M M, Jungnickel, D and Noltsch, G (1998) **Effizientes Real-Time Gross Settlement in der Wertpapierabwicklung.** Die Bank 2, 94–97.
- Group of Thirty (2003) **Global Clearing and Settlement – a Plan of Action.** Washington DC.

- Güntzer, M M, Jungnickel, D and Leclerc, M (1998) **Efficient Algorithms for the Clearing of Interbank Payments**. European Journal of Operational Research 106, 212–219.
- Horowitz, E and Sartaj, S (1978) **Fundamentals of Computer Algorithms**. Computer Science Press, Inc.
- Humphrey, D (1986) **Payments Finality and the Risk of Settlement Failure**. In Saunders, A and White, L C (eds.) Technology and the Regulation of Financial Markets, Lexington Books, D C Heath and Company, 97–120.
- Humphrey, D (1990) **Pricing Intraday Overdrafts**. In Haraf, E S and Cagan, P (eds.) Monetary Policy for a Changing Financial Environment, AIE Press, Lanham, 108–131.
- Humphrey, D (1996) **Comment on Intraday Bank Reserve Management: The Effects of Caps and Fees on Daylight Overdrafts**. Journal of Money, Credit and Banking 28, 909–913.
- Kahn, C and Roberds, W (1998a) **Payment System Settlement and Bank Incentives**. Review of Financial Studies, Vol. 1, No. 4, 845–870.
- Kahn, C and Roberds, W (1998b) **Real-Time Gross Settlement and the Cost of Immediacy**. Working Paper 98-21, Federal Reserve Bank of Atlanta.
- Koponen, R and Soramäki, K (1998) **Intraday Liquidity Needs in a Modern Interbank Payment System – A Simulation Approach**. Studies in Economics and Finance E:14, Bank of Finland.
- Kuussaari, H (1996) **Systemic Risk in the Finnish Payment System: An Empirical Investigation**. Discussion Paper 3/1996, Bank of Finland.
- Leinonen, H (1998) **Interbank Funds Transfer Systems: Liquidity Needs, Counterparty Risks and Collateral**. Discussion Paper 16/1998, Bank of Finland.
- Leinonen, H and Soramäki, K (1999) **Optimising Liquidity Usage and Settlement Speed in Payment Systems**. Discussion Paper 16/1999, Bank of Finland.

- Leinonen, H, Lumiala, V-M and Sarlin, R (2002) **Settlement in modern network-based payment infrastructures – description and prototype of the E-settlement model**. Discussion Paper 23/2002, Bank of Finland.
- Leinonen, H (2003) **Restructuring securities systems processing – a blue print proposal for real-time/t+0 processing**. Discussion Paper 7/2003, Bank of Finland.
- McAndrews, J J and Samira, R (2002) **Timing and Funding of Fedwire Funds Transfers**. FRBNY Economic Policy Review, July 2000.
- McAndrews, J J and Wasilyew, G (1995) **Simulations of Failure in a Payment System**. Economic Research Division Working Paper 95-19, Federal Reserve Bank of Philadelphia.
- Northcott, C A (2002) **Systemic Risk, Designation, and the ACSS**. Financial System Review, December 2002, Bank of Canada.
- Rossi, M (1995) **Pricing Intraday Credit in Real-Time Gross Settlement Systems**. Discussion Paper 211, Financial Markets Group, London School of Economics.
- Schoenmaker, D (1993) **Externalities in Payment Systems: Issues for Europe**. Financial Markets Group, London School of Economics.
- Schoenmaker, D (1995) **A Comparison of Alternative Interbank Settlement Systems**. Discussion Paper 204, Financial Markets Group, London School of Economics.

Chapter 3

Intraday liquidity needs in a modern interbank payment system – a simulation approach

Risto Koponen – Kimmo Soramäki

3	Intraday liquidity needs in an modern interbank payment – A simulation approach	74
	Abstract	74
3.1	Introduction.....	74
3.2	Liquidity and gridlock risks in payment systems	76
3.2.1	Sources of liquidity.....	76
3.2.2	Liquidity, gridlock and other risks in payment systems.....	77
3.2.3	Theoretical boundaries for liquidity needs.....	78
3.2.4	Real boundaries for liquidity needs.....	82
3.3	Payment systems simulator.....	83
3.4	Simulations and payment data	85
3.4.1	Liquidity optimisation methods.....	85
3.4.2	Pre-selected payment system structures	86
3.4.3	The data used in the simulations	88
3.5	Indicators for liquidity and settlement delay	90
3.5.1	Calculation of boundaries for liquidity need.....	90
3.5.2	Settlement delay indicator	92
3.5.3	Liquidity usage indicator	94
3.6	Results from the simulations.....	95
3.6.1	Upper and lower bounds for liquidity	95
3.6.2	Simulations of pre-selected payment system structures	100
3.6.3	Total daily value of payments and liquidity need ..	105
3.7	Conclusions.....	112
	References	114

3 Intraday liquidity needs in a modern interbank payment system – a simulation approach

Abstract

In this chapter three topics in particular are analysed. First, the adequacy of intraday credit limits in the existing and planned Finnish interbank funds transfers systems are studied. Secondly, the efficiency of two actual and two hypothetical payment settlement systems is analysed.

The results show that the existing intraday credit limits of the Finnish banks will be sufficient in the settlement system to be used in Stage Three of Economic and Monetary Union. Systems employing more real-time gross settlement were found to be more efficient in terms of speed of settlement and usage of liquidity per value of payments settled. An RTGS system operating with the same liquidity as an end-of-day net settlement system results in substantially faster settlement of payments.

3.1 Introduction

Over the past few decades, financial activity has increased significantly. This is mainly due to technological advance, deregulation of financial markets and innovations in financial instruments. We have also witnessed an increase in awareness of the credit, liquidity and systemic risks inherent in funds transfer systems.

In trying to reduce risk, interbank payment systems are shifting to a greater reliance on gross settlement, in which payments are settled individually. Final settlements take place sooner and settlement risks are reduced, but it is also the case that more liquidity is usually required than in a net settlement system.

Liquidity usually involves costs, which motivates banks to minimise their liquidity usage. There are several factors that affect banks' liquidity needs such as the relative importance of net and gross settlements and the specific liquidity optimisation methods used.

To study the effects of different concentrations of net and gross settlement and different optimisation methods on a payment system's

liquidity needs, liquidity usage and settlement delay, the Bank of Finland developed a payment system simulator (BoF-PSS1). The simulator handles a wide range of settlement systems, banking sector structures and optimisation methods. The effects of these features can be analysed separately in respect of individual banks or the banking sector as a whole.

This chapter uses the above mentioned simulation approach and it has two main objectives:

- 1 to examine the adequacy of current intraday credit limits of banks operating in Finland under certain pre-selected payment systems
- 2 to examine the efficiency of the pre-selected payment systems from the standpoint of liquidity needs and settlement delay

Both actual Finnish payment data as well as generated data based on the actual data were used in the simulations.

Four pre-selected payment systems were simulated, two actual and two hypothetical. The system used in Finland at the time the payment data were gathered is labelled here the RTGS-with-subnetting structure. The system that will be launched at the start of the third stage of EMU is referred to as the Hybrid structure. The latter payment system was simulated with and without estimated transactions going through the TARGET network. The hypothetical Advanced Hybrid structure entails even more extensive use of gross settlement than the Hybrid structure.

This chapter is divided into seven sections. The next section, section 3.2 sets out the concepts and theory on the handling of intraday liquidity and the risks inherent in liquidity provision. Section 3.3 gives a condensed description of the simulator. Section 3.4 focuses on the simulations and the payment data used. Section 4.5 presents the methods of calculating liquidity boundaries used in the simulations and explains the indicators of liquidity usage and settlement delay. Section 3.6 is devoted to a presentation of the findings of this study. Finally, the most important results are summed up in section 3.7 and some interesting, but still unresolved, questions that have arisen during this project are presented.

3.2 Liquidity and gridlock risks in payment systems

3.2.1 Sources of liquidity

Banks need liquidity in settling their payments. This liquidity can be provided by the central bank or by the banks themselves in the money market. The central bank, depending on its policy preferences, can provide intraday liquidity to the banking sector by allowing banks to use their required and excess reserves for settlement purposes, allowing banks to overdraw on their settlement accounts, or arranging intraday repos.

Many central banks use reserve requirements as a means of conducting monetary policy. A central bank may allow banks to use their required reserves and any excess reserves held at the central bank for settling payments. If required reserves are used for payment settlement, the average amount of liquidity in the settlement account must meet the requirement during the reserve maintenance period. The central bank may also allow settlement system participants to overdraw on their settlement accounts, with or without interest charges. Partial or full collateralisation of overdrafts may be required.

Several types of costs are associated with systems in which collateral requirements are attached to central bank credit facilities. Securities tied up as collateral give rise to opportunity costs because they are no longer available for trading and other purposes during the day. Because of this, the banks may be forced to hold inferior portfolios compared to those that would result from free choice.

In practice there are always some (implicit or explicit) cost factors inherent in liquidity. This makes liquidity scarce and creates an incentive for banks to optimise their use of liquidity. The interest rate in the money market and the opportunity costs of collateral are determined by the markets. The interest rate charged on central bank credit is determined administratively according to risk and monetary policy factors. Thus these cost factors cannot be readily influenced by the banking sector.

There is one important free source of liquidity. This is provided by the payment system itself in the form of incoming payments. The faster liquidity circulates among the banks, the less the aggregate liquidity needed in the system. The more efficient the procedures and technical features, the less the system's need for liquidity injections from the outside.

Besides liquidity costs, there are also costs associated with settlement delays. At least some of the payments are likely to be time-critical, which means that any delay in settlement will generate costs to the sending or receiving bank. These costs may be implicit, in the form of a deterioration in customer service, or explicit, in the form of agreed sanctions.

3.2.2 Liquidity, gridlock and other risks in payment systems

Regarding payment system participants' intraday liquidity needs, the major risks inherent in the system are credit and counterparty risks, liquidity and gridlock risks, and systemic risk. Especially liquidity and gridlock risks are discussed in more detail in subsequent sections.

Liquidity risk is the risk of a loss that arises when a bank's liquid assets or its immediate access to credit are insufficient to cover its payment obligations. Types of liquidity risk include variation risk, availability risk and gridlock risk.

Variation risk arises because of wide variations in a bank's liquidity needs, which means that at times it is unable to forward payments it has undertaken and must postpone the transaction.

Availability risk arises when a market condition or a bank's impaired financial condition reduces the amount of liquidity that the bank can obtain from the market to the extent that it has difficulty in making payments for which it is irrevocably committed. Poor liquidity management may lead to repeated payment delays, compensation claims and, if prolonged, to a loss of customers to rivals.¹

A type of liquidity risk that is associated particularly with queuing arrangements is gridlock risk. Gridlock has been variously defined, eg as a situation in which the failure of one bank to execute transfers prevents a substantial number of other participants' transfers from being executed.²

It should be noted that a queuing system itself does not cause liquidity risks or gridlock. Gridlock is a result of insufficient liquidity on the part of one or more participants. There are also various procedures that can be incorporated in the queuing system that will solve or prevent the formation of gridlocks. These procedures include

¹ Leinonen and Saarinen, p. 36.

² BIS 1997, p. 17.

splitting of payments and netting of queues, both of which are discussed in section 3.4.1.

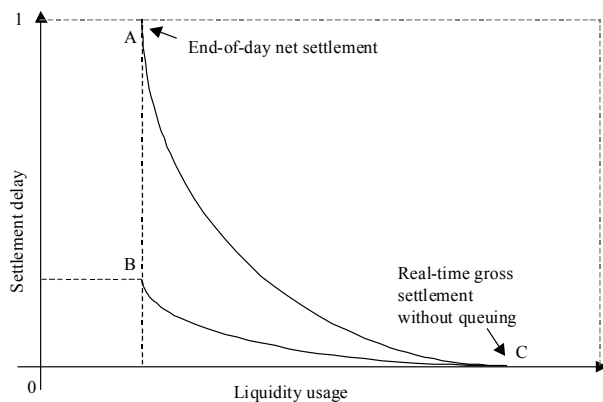
Another type of gridlock is ‘self imposed gridlock’. This type results from the behaviour of the participants. Each participant relies on incoming payments as its only source of liquidity for settling its outgoing payments. Thus, in the extreme case in which each bank delays the sending of its payments, no payments are settled. These types of situations are commonly referred to as prisoners’ dilemma situations, as optimal behaviour by each participant leads to an inferior outcome for all.

3.2.3 Theoretical boundaries for liquidity needs

The relationship between liquidity need and settlement delay in different payment settlement systems is analysed in this study within the framework depicted in figure 3.1. The liquidity used by the settlement system (x axis) consists of settlement resources such as reserves held at, or intraday credit received from, the central bank. The corresponding delay in settlement (y axis) is the time span between receipt of a payment order by the bank and final and irrevocable settlement of the payment.

Figure 3.1

Relationship between liquidity usage and settlement delay in RTGS and TDNS systems without counterparty risk



BC: amount of reserves and intraday credit in RTGS system with queuing

AC: number of daily net settlements in TDNS system without counterparty risk

Liquidity need and settlement delay in payment systems without counterparty risk are considered first.

Case 1. RTGS system with queuing

In a system with queuing, the banks need not have sufficient funds to settle their payments until the end of the day. In this case, the minimum liquidity needed for successful settlement of all of a bank's payments is equal to the excess value of outgoing over incoming payments (absent gridlock³). This is illustrated in equation 3.1 and represented by point B in figure 3.1.

$$LB_{t(\text{theoretical})} = \min \left[0, \left(\sum_{t=0}^T P_t^I - \sum_{t=0}^T P_t^O \right) \right] \quad (3.1)$$

Equation 3.1 Theoretical lower bound for a bank's daily liquidity need in an RTGS system with queuing (P^I = value of incoming payment, P^O = value of outgoing payment)

At point B, settlement delay is at its maximum. A bank can reduce the delay in settling its payments by increasing its liquidity. As a bank increases its liquidity, it eventually reaches point C, which represents the level of liquidity needed for its payments to be settled immediately. The minimum liquidity that a bank needs for immediate payments settlement equals the absolute value of its daily minimum cumulative net amount of incoming and outgoing payments. If the bank's net liquidity position is positive throughout the day, its external liquidity need is zero, since it receives sufficient liquidity in the form of incoming payments. If its net liquidity position is negative, the bank needs to acquire enough liquidity to cover the shortfall in order to settle its payments without delay. This is illustrated by equation 3.2 and represented by point C in figure 3.1.

$$UB_{t(\text{theoretical})} = \min \left[0, \min_t \sum_{i=0}^t (P_i^I - P_i^O) \forall t \in [0, T] \right] \quad (3.2)$$

³ An end-of-day gridlock can be solved by netting the queues and hence the same minimum would hold. It is also possible to solve a gridlock by splitting payments, but it may be necessary to have a splitting system that splits the payments into the smallest currency unit available.

Equation 3.2 Theoretical upper bound for a bank's daily liquidity need in an RTGS system with queuing (P_i^O = value of outgoing payments at time i , P_i^I = value of incoming payments at time i , T = end of day)

The curve segment BC shows the tradeoff between liquidity usage and settlement delay. Liquidity must remain at least at the level of point B if all payments during the day are to be settled. Additional liquidity beyond point C is unnecessary because all payments get settled immediately.

Banks can theoretically choose any point on curve segment BC, according to their preferences. If a bank perceives the cost of liquidity to be high relative to that of settlement delay, it chooses a point near B and vice versa.

Case 2. RTGS system without queuing

In an RTGS system without queuing, all the banks must have adequate liquidity to settle their payments immediately.⁴ A bank's liquidity need equals the upper bound for its liquidity in an RTGS system with queuing, ie the bank's minimum cumulative net amount of incoming and outgoing payments throughout the day. Because this amount of liquidity is needed for immediate payment settlement and any additional liquidity is unnecessary (since all payments are settled immediately), it represents both the lower bound and upper bound for the bank's liquidity in an RTGS system without queuing.

Case 3. TDNS system without counterparty risk

A bank's liquidity need in a time designated net settlement system with end-of-day settlement equals that of point A in figure 3.1. The liquidity need is the same as in an RTGS system with queuing, but the total delay in settlement is at its maximum. If the number of net settlements during the day is increased, settlement delay can be traded off for greater liquidity needs. The curve segment AC shows this tradeoff. If the number of settlements is increased to the point where net settlement is executed after each transaction, the system becomes

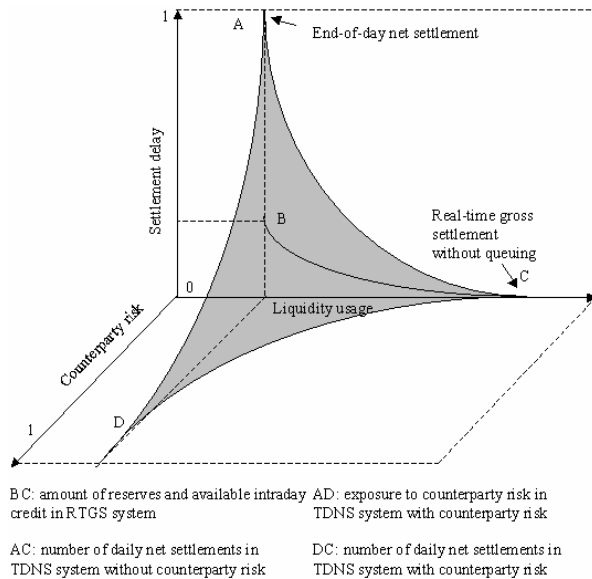
⁴ It is assumed that there is no internal queuing within the banks, as explained in the assumptions for this framework.

in effect a real-time gross settlement system without queuing. This is shown as point C in figure 3.1.

So far we have discussed only systems without counterparty risk. Properly designed real-time gross settlement systems are free of counterparty risk. Depending on the design of the system, a net settlement system can operate with or without counterparty risk. The z axis in figure 3.2 gives the degree of counterparty risk in the settlement of payments. This risk encompasses the risks inherent in the implicit debt relations between system participants.

Figure 3.2

Relationship between a bank's liquidity usage, settlement delay and counterparty risk in an RTGS or TDNS system with counterparty risk



Case 4. TDNS system with counterparty risk

In figure 3.2, risk is introduced into the relationship between settlement delay and liquidity usage. A time designated net settlement system that operates with counterparty risk rather than liquidity is illustrated by the curve AD. By crediting customers' accounts before final settlement, the total settlement delay can be reduced. If all transfers are credited before final settlement, delay is eliminated and counterparty risk is at its maximum, as illustrated by point D in the figure.

The curve AD representing the tradeoff between settlement delay and risk is concave. By crediting the payments of participants representing the smallest counterparty risk, delay in settlement can be reduced with less risk than if payments of the riskier participants are credited before final settlement. The shape of curve AC reflects the assumption of diminishing returns to increases in the number of net settlements during the day.

The area ACD in figure 3.2 represents the possible combinations of the number of net settlements during the day, the amount of risk a bank is willing to take, and the amount of liquidity used for settlements.

In this study, only structures in the xy plane are simulated. This means that all the systems studied have the same level of counterparty risk (zero), which enables efficiency comparisons.

3.2.4 Real boundaries for liquidity needs

If time-critical transfers and payment prioritisation are added to the system, the upper and lower liquidity bounds will change. In this study a bank's bounds within a system with these features are referred to as its real upper and lower bounds (UB_r and LB_r). In actual payment systems at least some payments are likely to be time-critical.

Simulations were necessary for quantifying real bounds. In calculating the real lower bound, an account holder was assumed to hold the smallest possible amount of liquidity for successful settlement of its payments. This amount is equal to its theoretical lower bound of liquidity (LB_t), ie the net amount of all its incoming and outgoing payments throughout the day. The daily limits were raised as needed for timely settlement of time-critical transfers. The resulting maximum liquidity usage for each account holder during the day became its real lower bound for liquidity need.

The theoretical upper bound (UB_t) is calculated as the minimum cumulative net amount of incoming and outgoing payments throughout the day. For the real upper bound (UB_r), the prioritisation and time-criticalness of payments were introduced in the simulations. In principle, these upper bounds should not be affected, as no queuing takes place at this liquidity level. However, there are technical features in the following simulated payment system structures that cause these bounds to change (see section 3.6.1)

In a system without time-critical payments, the theoretical lower bound (LB_t) is always lower than or equal to the theoretical upper bound (UB_t). However, this may not be the case if time-criticalness of

payments is introduced, ie some payments must be settled immediately upon entry into the system or within a specified period of time, as eg those in settlement of net positions of a net settlement system.

Simulations on the liquidity bounds showed that the real lower bound (LB_r) can be higher than the real upper bound (UB_r). This may be due the inclusion of time-criticalness and prioritisation of payments and the effects of queuing on a bank's liquidity. If a bank has payments queued, the receivers of these payments do not get the funds and so may have to raise their intraday credit limits in order to obtain liquidity needed for settling any time-critical payments. If a bank has to raise its credit limit by more than it can substitute liquidity for settlement delay during the day, its real lower bound will be higher than its real upper bound. At the latter liquidity level, no queuing is needed and the bank gets its payments settled immediately without needing to raise its credit limit during the day, as the liquidity circulates efficiently in the system.

This means that if the banks select adequate liquidity levels, they can settle their payments smoothly and with smaller liquidity needs. A bank may however select a liquidity level that produces a suboptimal result for all participants in the system.

3.3 Payment systems simulator

The simulation runs for this study were done using the payment systems simulator developed by the Bank of Finland, BoF-PSS1. The simulator is an explanatory model of payment settlement systems. It includes procedures for handling payments of actual payment systems and hence it produces exactly the same outcomes as an actual system with the same properties using the same input data. But the simulator enables the study of the effects of different technical and policy features of a payment settlement system.

It should be noted that the simulator is not an optimisation model. No constraints are set on the results of model simulation and no cost calculations are included.

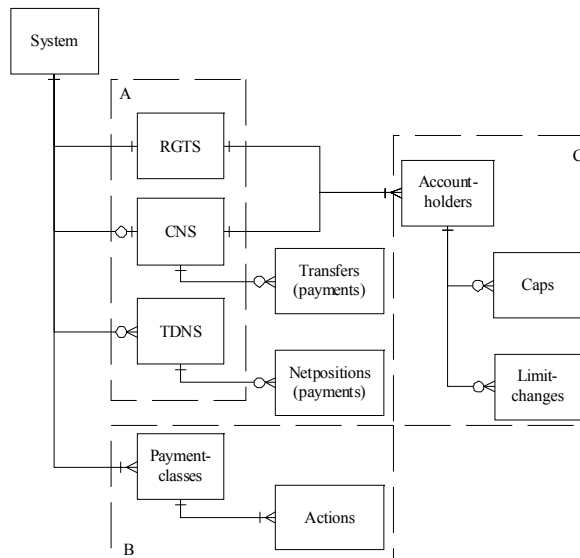
The payment system is organised in the settlement simulator as depicted in figure 3.3, which presents the object model of the simulator. The scenarios drawn in the figure with dashed lines and marked A, B, C are respectively the systems, settlement and account holder scenario. A combination of scenarios selected at the start of a simulation run is referred to as a settlement structure.

The system object in the object model controls the other objects and their interaction according to the property settings. For example, as payments (Transfers or Netposition objects) are generated by the CNS or TDNS objects, they are settled in the RTGS object and the balance property of the Accountholder object is changed. The logic of the settlement resides in the Paymentclasses object. Each payment class can be settled by any of the three methods (RTGS, CNS, and TDNS) or any combination of these during the day.

The settlement structure always includes one RTGS system and may include one CNS system; zero or several TDNS systems can be included. The account holders of the CNS system are a subset of those of the RTGS system. An account holder must participate in the RTGS system in order to participate in the CNS system but not vice versa. Each account holder may have zero or several caps, depending on whether it is a CNS participant. Each account holder may also have zero or several changes in intraday credit limits during the simulation period. At least one payment class and one corresponding settlement procedure must be defined. The simulator imposes no maximum numbers for these.

Figure 3.3

Payment systems simulator object model



3.4 Simulations and payment data

With the simulator, a model resembling the Finnish payment system environment was constructed containing the processing features of the main systems and the transactions of the main participants. The main systems are the RTGS-system, PMJ for retail transactions and POPS system for urgent transfers. In addition the financial market transactions and loro payments were added. Different payment system processing parameters and liquidity optimisation methods was analysed.

3.4.1 Liquidity optimisation methods

The optimisation methods present in the simulated payment system structures include queuing of payments, net settlement of queued payments and splitting of payments.

Queuing of payments

Each participant in an RTGS system holds a settlement account at the central bank, to which debit and credit entries are made. Payments without covering funds are not settled but rejected and returned to the sender for later input or entered into a centrally managed queue.

Different centralised RTGS queuing systems may have different rules for payments settlement. The Finnish RTGS system works on a ‘first in, first out’ (FIFO) basis. Payments that are more time-critical than others can be given higher priority.

Net settlement of queued payments

One way to solve a gridlock is to execute a net settlement of all the queued payments. If each bank has enough liquidity to settle its net amount of queued incoming and outgoing payments, the queues are cleared and each bank’s account appropriately debited or credited. A system is in gridlock if equation 3.3 holds for every bank but not all of the queued payments get settled.

$$\sum_{j=1}^N P_j^{L,i} - \sum_{k=1}^M P_k^{O,i} < L_i \text{ for each bank } i \quad (3.3)$$

Equation 3.3: Definition for gridlock ($P_j^{i,i}$ = value of incoming payment in a queue for bank i , $P_k^{O,i}$ = value of outgoing payment in a queue for bank i , L_i = bank i 's liquidity, N = Number of incoming payments for bank i , M = number of outgoing payments for bank i)

The netting of queues can reduce (in some cases substantially) the system's liquidity needs because the net position is by definition the minimum amount of liquidity that ensures the settlement of all payments. If queued payments are settled individually, a participant's liquidity need depends on the order of settlement. However, at the end of the day, there is no difference between net and gross settlement-with-queuing systems as regards the amount of liquidity needed to settle the day's payments, but only on settlement delay.

Splitting of payments

Another way to make an RTGS system work more smoothly and to avoid gridlocks is to split large payments into several smaller ones. These smaller transfers then represent a source of liquidity to receiving participants. Without this feature, receivers must wait until the paying bank has accumulated enough liquidity, eg via incoming payments. This might in turn prevent a receiving bank from executing its own queued outgoing transfers. Such situations can lead to gridlocks that could have been prevented by the splitting of payments.

The effectiveness of splitting of payments for solving gridlocks depends on the technical features of the splitting. If the splitting is done to the smallest unit of account or payments are split so that all the available liquidity of every bank is used, this is as liquidity-efficient a way of solving gridlocks as the netting of queues. The splitting of payments also requires tenable legal arrangements binding on all parties.

3.4.2 Pre-selected payment system structures

The simulated payment system structures and policies are RTGS with subnetting, Hybrid, Advanced Hybrid and RTGS with queuing. The RTGS-with-subnetting structure refers to the settlement policy and structure used in May 1997. The Hybrid structure reflects the situation as of the start of 1999, and the Advanced Hybrid structure is a

hypothetical structure with even more extensive use of gross settlement.

The characteristics of the settlement scenarios used in the simulations are shown in table 3.1 and the shares of the value of payments settled via the three types of settlement systems are summed up in table 3.2. The PMJ payments include retail payments between banks such as debit transfers, ATM withdrawals, debit card payments and recurrent payments. POPS payments are mainly large-value customer payments that comprise express transfers or cheques. Loro payments are markka-denominated foreign payments. All other payments were settled in real-time on a gross basis in all scenarios.

Table 3.1 Settlement scenarios

	RTGS with subnetting	Hybrid	Advanced Hybrid	RTGS with queuing
PMJ payments	Net settlement at 15:45	Net settlement at 01:00 and 15:45	Net settlement at 01:00 and 15:45	RTGS (bilateral positions)
POPS payments	Within PMJ net settlement	Over limit: RTGS Under limit: CNS	RTGS	RTGS
Loro payments	Net settlement at 14:30	≥ FIM 50 000: (ECU 8300) RTGS < FIM 50 000: within PMJ net settlements	≥ FIM 50 000: (ECU 8300) RTGS < FIM 50 000: within PMJ net settlements	RTGS
Financial markets transactions	Net settlement at 13:00	RTGS	RTGS	RTGS

Table 3.2 Shares of value settled using different settlement systems in the pre-selected settlement structures, %

	Settled by		
	RTGS	TDNS	CNS
RTGS with subnetting	34.6	65.4	0.0
Hybrid	88.4	6.3	5.3
Advanced Hybrid	93.7	6.3	0.0
RTGS with queuing	100.0	0.0	0.0

The results concerning the differences between the net-based RTGS and Hybrid structures are of great importance because both are applied in existing payment systems. The former reflects the situation where payment data are gathered and the latter is the EMU-compatible payment settlement environment as at the start of Stage Three of EMU.

These simulations enable determination of whether the current liquidity reserves possessed by the payment system participants are sufficient also for the EMU-compatible payment system environment. If liquidity shortfalls exist, an injection of liquidity, eg in the form of an increase in the system account balances or intraday overdraft limits might be required to ensure smooth operation of the payment system.

The simulation runs for different structures are presented in table 3.3.

Table 3.3 **Pre-selected payment system structures**

	RTGS with subnetting	Hybrid	Advanced Hybrid	RTGS with queuing
Account holder scenario	The Finnish banking system ¹			
Systems scenario	– no optimisations	– queuing of payments – netting of queues every 20 minutes – prioritisation of payments	– queuing of payments – netting of queues every 20 minutes – splitting of payments worth over ECU 16.6 mill. – prioritisation of payments	– queuing of payments – prioritisation of payments
Settlement scenario	RTGS with subnetting	Hybrid	Advanced Hybrid	RTGS with queuing
Intraday credit limits	Existing limits and 10%- point intervals between theoretical lower and upper bound of liquidity			
Simulation period	4 days of actual payment data and 100 days of generated data			

¹Only banks participating in the simulation project were included; these accounted for over 90 per cent of payments in terms of value and number.

3.4.3 The data used in the simulations

The payment data used in the simulations were provided by eight of the major banks operating in Finland, which constitute over 90 per

cent of total transactions in the BoF-RTGS system, in terms of value or number of transactions.

The time period of the payments is from 13–16 May 1997, which are the business days of a whole week excluding Monday. Although the four-day period is quite short, the week was characterised by most of the banks as representative of their normal payment patterns. In addition to the payments reported by the commercial banks, the payments of the Bank of Finland and certain important nonbank entities were included. Data on their payments was extrapolated from their settlement accounts at the Bank of Finland.

Table 3.4 gives the value and number shares of individual payment classes. The value and number breakdowns over different payment classes are summed up in table 3.5.

Table 3.4 **Outgoing payments (13–15 May 1997)**

	Total value, mill. ECU	Total number of payments	Daily average value, mill. ECU	Daily average number of payments	Average value of payment, mill. ECU	Largest value of payment, mill. ECU
PMJ payments	2,389	2 467,979	597	616,995	0.001	n/a
POPS express transfers	3,685	1,905	921	476	1.9	84
POPS/PMJ cheques and bank drafts	1,812	9,877	453	2,469	0.2	140
Loro payments	18,707	3,711	4,677	928	5.0	585
Financial market transactions	7,363	344	1,841	86	21.4	938
Interbank transfers	6,475	251	1,619	63	25.8	316
Total	40,432	2 484,067				

Table 3.5 **Number and value shares of payment classes (13–16 May 1997)**

Payment class	% of value	% of number
POPS express transfers	10	12
POPS/PMJ cheques and bank drafts	5	61
Loro payments	49	23
Financial market transactions	19	2
Interbank transfers	17	2

Because a simulation period of four days is rather short for drawing conclusions about the effects of changes in the systems, a procedure was developed for extrapolating for additional days by using a random process, which was based on the statistical distributions of the actual data.

The data on future TARGET transactions had also to be estimated and is derived from balance of payment data, with certain assumptions added. Portions of both cross-border and loro payments will be settled in TARGET in Stage Three of EMU. Loro payments to/from outside of the euro area are considered to remain as they are. For loro payments within the euro area, two scenarios were used. In the short-term scenario, 80 per cent of euro-area loro payments were converted into domestic payments and settled as POPS payments, 10 per cent were settled via TARGET and 10 per cent vanish. The numbers of domestic payments were assumed to remain unchanged.

3.5 Indicators for liquidity and settlement delay

3.5.1 Calculation of boundaries for liquidity need

The behaviour of the liquidity position of a hypothetical bank in an RTGS system during a day is illustrated in figure 3.4. Within this context, the bank begins the day with a zero liquidity position and an unlimited credit extension. The flow of payments during the day is quite uneven. The end-of-day liquidity need, point B in the figure, represents the net amount of incoming and outgoing payments during the day. This point is the theoretical lower bound (LB_t) for a bank's liquidity in an RTGS system with queuing.

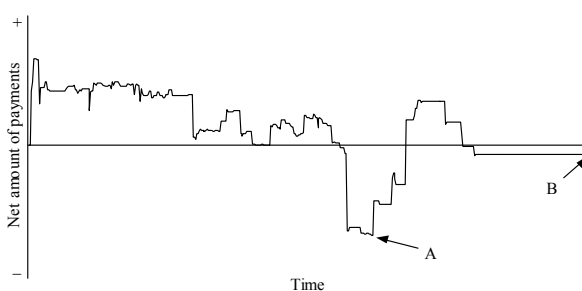
However, this lower bound holds only if none of the payments settled are time-critical and hence liquidity need not be available for settlement until the end of the day. In the Finnish payment settlement systems, net positions originating from net settlement systems must be settled immediately and some payments within an hour after entry into the system.⁵ Thus it was necessary to simulate also a bank's real lower bounds for settlement systems with time-critical payments (LBr). In these simulations each account holder was assigned a limit equivalent to its net position for all incoming and outgoing payments during the day. These limits were then raised as needed for timely settlement of

⁵ Viz POPS-RTGS payments and POPS buckets.

time-critical transfers to find the real lower bound for its liquidity need.

The upper bound for liquidity need is relevant if all payments are settled without queuing. An accountholder's minimum liquidity position during the day then represents the theoretical upper bound for its liquidity need (UB_t). This was calculated as the minimum of the cumulative net positions of incoming and outgoing payments at all points of time during the day. This amount is represented by point A in figure 3.4.

Figure 3.4 **Intraday liquidity usage by a hypothetical bank in an RTGS system**



Because queuing of payments takes place only between the lower and upper bounds, only liquidity levels between these bounds are of interest in this study. In the simulations, eleven different liquidity levels between the bounds were used. These levels are represented in figure 3.5 as points on the line ranging from a liquidity level of 0 per cent to 100 per cent.

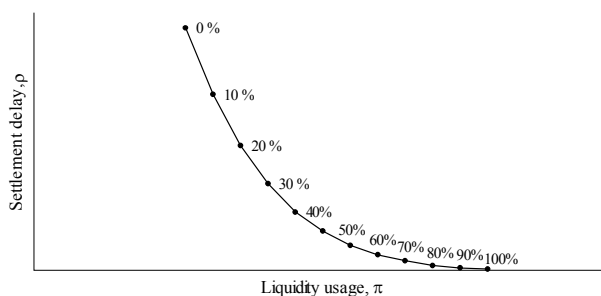
The amount of liquidity available for any account holder i is calculated as shown in equation 3.4. Liquidity available for each bank at a particular liquidity level is the sum of the lower bound and the corresponding liquidity level multiplied by the difference between the bounds. The lower bound for liquidity need is the 0 per cent liquidity level and the upper bound the 100 per cent.

$$LA_i = LB_i + LL * (UB_{t,i} - LB_{t,i}) \quad (3.4)$$

Equation 3.4: Liquidity available, LA_i , for account holder i at a given liquidity level, LL (LB_t = theoretical lower bound, UB_t = theoretical upper bound)

In calculating system liquidity need, the system upper bound, and system lower bound; the corresponding values for each account holder are simply added up. It should be noted that the liquidity must be optimally distributed in order for the system bounds to hold. The curve in figure 3.5 shows the points where the liquidity is optimally distributed across system participants. A reduction of any participant's liquidity would cause extra delay in settlement.

Figure 3.5 **Relationship between a bank's settlement delay and liquidity usage in a payment system with various liquidity levels**



The different liquidity levels are presented as narrowing settlement delay intervals between the points on the curve. However, the simulations showed that distances between neighbouring points representing different combinations of liquidity usage and settlement delay can vary substantially. A small reduction in available liquidity may imply a big change in settlement delay and vice versa.

3.5.2 Settlement delay indicator

The indicator used for settlement delay in this study is called ρ (rho). The values of ρ range from zero to one and it is calculated for each account holder as shown in equation 3.5.

$$\rho = \frac{\sum_{i=1}^T Q_i}{\sum_{t=1}^T \sum_{i=1}^t V_i} \quad (3.5)$$

Equation 3.5: Indicator of settlement delay, ρ , for an account holder (Q_i = value of queue at time i , V_i = value of outgoing payments at time i)

The numerator in equation 3.5 represents the sum of queues, ie the sum of the values of queued payments over each minute of the day. The denominator represents the sum of the cumulative values of outgoing payments over each minute of the day, and ρ is the ratio of the two.

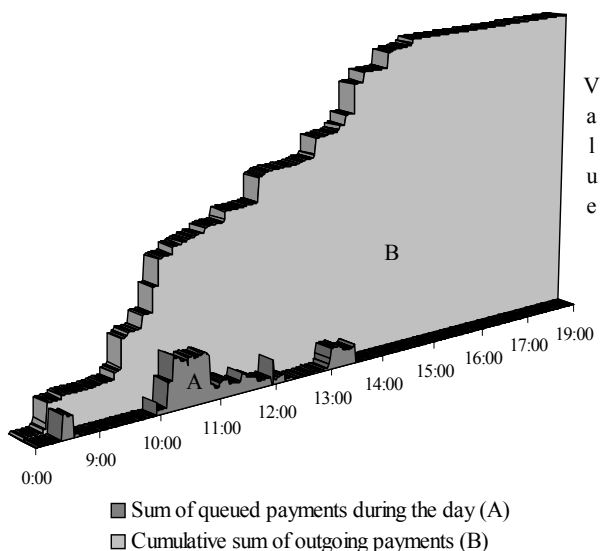
If a bank does not have any liquidity at the start of the day and does not receive any in the form of incoming payments, all transfers remain queued and are not settled at all or only at the end of the day. In this case, ρ equals one. On the other hand, if the bank has an abundance of liquidity, all payments get settled immediately and ρ is zero.

The calculation of ρ is illustrated in figure 3.6. The height of the curve defining the dark grey area (A) represents the total value of a bank's queued payments at each minute. The light grey area (B) represents the bank's cumulative value of all outgoing payments settled at each minute during the day. The settlement delay indicator, ρ , is the ratio of A to B.

In calculating the system ρ , the numerator and denominator in equation 3.5 are summed up over all account holders in the system. The system ρ is thus a weighted average of individual account holders' ρ s, where the weights are corresponding shares of the account holders in the total value of payments.

Figure 3.6

Settlement delay indicator $\rho = A/B$



By using such an indicator, the settlement delay in various systems can be measured in a standardised manner. ρ takes into account the value and queuing times of delayed transfers as well as their importance in the total value of payments.

3.5.3 Liquidity usage indicator

In this study, liquidity usage is calculated as the sum of the peak usage of intraday limits plus the peak usage of the starting liquidity position. The corresponding indicator, denoted π (π_i), is calculated for each bank as the ratio of its liquidity usage to the total value of its outgoing payments during the day. π ranges from zero to one. A π of zero means that there is no need for liquidity from outside the system, and if π equals one, liquidity is needed in the amount of the gross value of outgoing payments. π can also be understood as the reciprocal of the turnover ratio.

$$\pi = \frac{LU}{\sum_{t=0}^T V_t} \quad (3.6)$$

Equation 3.6: Liquidity usage indicator, π , for an account holder (LU = peak use of starting liquidity position + peak use of credits extended, V_t = value of payments sent at time t)

In calculating the system π , the liquidity usages of individual account holders are summed and divided by the total value of payments during the day. This is equivalent to the weighted average of the banks' π s, where the weights are the banks' respective shares of the total value of payments.

3.6 Results from the simulations

3.6.1 Upper and lower bounds for liquidity

The theoretical and real upper bounds (UB_t , UB_r)⁶ and lower bounds (LB_t , LB_r)⁷ for liquidity presented in the following are the averages of the bounds for each of the account holders over the whole time period studied. The lower and upper bounds are calculated for all four preselected payment settlement structures.

For the Hybrid, Advanced Hybrid and RTGS-with-queuing structures, the theoretical lower bound (LB_t) and upper bound (UB_t) the same (see table 3.6). The difference between RTGS-with-subnetting and the other structures is due to the slight difference in the payment data used (opening hours).

The difference between the theoretical (LB_t) and real (LB_r) lower bounds is due to the introduction of time-criticalness and prioritisation for some payments. The intraday credit limits were raised for some banks during the day, due to time-critical payments, which resulted in greater liquidity usage. The prioritisation of payments changes the order in which the payments are settled and it can affect the liquidity usage negatively or positively for individual banks.

The difference between theoretical (UB_t) and real upper bound (UB_r) is due to differences between the structures and technical features of the systems. In the pre-selected systems, the order in which payments are settled varies. The CNS system imposes liquidity needs

⁶ The upper bound represents the amount of liquidity needed for immediate execution of all outgoing payments.

⁷ The lower bound represents the amount of liquidity needed for successful settlement of all outgoing payments at the end of the day.

on the RTGS system at different points of time during the day, depending on how many payments are settled. The TDNS system imposes liquidity needs on the RTGS system at the times when net settlements are due. These difference in the timing of liquidity needs causes the differences in the theoretical (UB_t) and real (UB_r) upper bounds.

Table 3.6 **Upper and lower liquidity bounds for pre-selected payment system structures, mill. ECU, actual payment data for 4 days, per bank average for n=8 banks**

	RTGS with subnetting	Hybrid	Advanced Hybrid	RTGS with queuing
Theoretical lower bound (LB_t)	53.4	53.1	53.1	53.1
Real lower bound (LB_r)	89.8	86.1	86.5	59.9
Theoretical upper bound (UB_t)	225.4	225.4	225.4	225.4
Real upper bound (UB_r)	158.2	229.7	229.7	222.1

Table 3.7 **Ratio of real to theoretical bounds, %, actual payment data for 4 days**

	RTGS with subnetting	Hybrid	Advanced Hybrid	RTGS with queuing
Theoretical lower bound (LB_t)	100.0	100.0	100.0	100.0
Real lower bound (LB_r)	168.1	162.2	163.0	112.9
Theoretical upper bound (UB_t)	100.0	100.0	100.0	100.0
Real upper bound (UB_r)	70.2	101.9	102.3	98.5

In terms of real lower bounds (LB_r), the RTGS-with-queuing structure requires the least liquidity, and there are no significant differences between other structures in this respect. The RTGS-with-queuing structure is the only one of the pre-selected structures not including time-critical payments. The introduction of time-criticalness and prioritisation of some payments⁸ seems to significantly increase a bank's lower bound of liquidity need in all the structures.

In terms of real upper bounds (UB_r), the RTGS-with-subnetting structure requires the least liquidity. This suggests that this structure is

⁸ viz POPS payments, POPS buckets and net settlement transactions.

the most efficient in terms of liquidity needs for settling payments without queuing in an RTGS system. However, it should be noted that all structures other than RTGS-with-queuing included one or more net settlements and that settlement delay in these TDNS systems is not taken into account in the calculations. As the payments are collected for the net settlement, their final settlement is postponed until the net positions are settled between the banks.

Moreover, the differences between lower and upper bounds are quite significant, both in theoretical and practical terms. The simulations show that if the banks choose to settle their payments immediately without queuing, they will need more liquidity than that required for end-of-day settlement of net positions. The RTGS-with-subnetting structure required only 1.8 times the amount of liquidity that is required for immediate settlement, compared to a ratio of 2.7 for both hybrid structures and about 3.7 for the RTGS-with-queuing structure. Corresponding indicators for liquidity usage and settlement delay are summarised in table 3.8.

Table 3.8 Settlement delay and liquidity usage for real lower and upper liquidity bounds, %, actual payment data for 4 days

	Real lower bound (LB _r)		Real upper bound (UB _r)	
	Liquidity usage, π	Settlement delay, ρ	Liquidity usage, π	Settlement delay, ρ
RTGS with subnetting	27	19	37	0
Hybrid	9	18	25	0
Advanced Hybrid	9	17	25	0
RTGS with queuing	6	29	21	0

The RTGS-with-queuing structure requires the least liquidity if there is no queuing. The RTGS-with-subnetting structure is the least efficient structure in this respect. The Hybrid and Advanced Hybrid structures are between the former structures, assuming equal liquidity usage.

The RTGS-with-queuing structure is also superior in terms of liquidity usage in situations with maximal queuing. The cost of the reduction in liquidity usage is more settlement delay. The Hybrid and Advanced Hybrid structures use slightly more liquidity but have substantially shorter settlement delays. Liquidity usage with the RTGS-with-subnetting structure is reduced by only 10 percentage points compared to the situation with no settlement delays. This

structure uses the largest amount of liquidity relative to the value of payments settled.

It can be concluded that the RTGS-with-queuing structure would be the superior option for the banks in terms of liquidity needs and liquidity usage. However, within the lower levels of liquidity (ie at the lower bound level), the RTGS-with-queuing structure results in long delays in settlement. The Hybrid and Advanced Hybrid structures can be seen as good compromises between liquidity usage and settlement delay. In both structures, some of the payments were settled via netting, and the delay in settling these payments is not taken into account here. The system in which the majority of payments was settled via netting was not as good as the other pre-selected systems in terms of efficiency. The circulation speed of liquidity in the RTGS system was very low, only 2.7 at the upper bound of liquidity and 3.7 at the lower bound.⁹

The results were quite stable over the variations on settlement order and number of daily transactions. The simulations with the generated payment data supported the conclusion drawn on the basis of four days of actual payment data. Even though the RTGS-with-queuing structure may be the superior structure at higher levels of liquidity, the Hybrid and Advanced Hybrid structures are good compromises between liquidity usage and settlement delay. The simulations with generated data resulted in only minor differences between the different structures.

The lower bounds were higher with the generated payment data than with the actual data. This suggests that the variation in the value of incoming transfers is greater than the value of outgoing transfers. Also the upper bounds are higher, indicating some deterioration in payments synchronisation during the day. However, the results are in line with the actual payment data.

⁹ The circulation speed of liquidity is the reciprocal of system π , ie the ratio of total value of payments to liquidity usage during a day.

Table 3.9

Upper and lower liquidity bounds for pre-selected payment system structures, mill. ECU, generated payment data, average for n=8 banks

	RTGS with subnetting	Hybrid	Advanced Hybrid	RTGS with queuing
Theoretical lower bound (LB _t)	134.3	134.6	134.6	134.6
Real lower bound (LB _r)	151.7	156.3	156.1	139.4
Theoretical upper bound (UB _t)	236.4	236.3	236.3	236.3
Real upper bound (UB _r)	208.0	237.5	237.4	236.3

Table 3.10

Ratios of real bounds to theoretical bounds, generated payment data, 100 days, %

	RTGS with subnetting	Hybrid	Advanced Hybrid	RTGS with queuing
Theoretical lower bound (LB _t)	100.0	100.0	100.0	100.0
Real lower bound (LB _r)	112.8	116.2	116.0	103.6
Theoretical upper bound (UB _t)	100.0	100.0	100.0	100.0
Real upper bound (UB _r)	88.0	100.5	100.4	100.0

The values of the liquidity usage indicator, π are significantly higher with the generated payment data and the settlement delay indicator, ρ is lower in all pre-selected structures. However, relative to each other, the differences between the structures are not substantial. The Advanced Hybrid structure is superior in terms of settlement delay and the RTGS-with-queuing structure in terms of liquidity usage. The difference between the Advanced Hybrid and Hybrid structures is only 1 percentage point by both indicators.

Table 3.11

**System settlement delay and liquidity usage
at the real lower and upper liquidity
bounds, generated payment data,
100 days, %**

Structure	Real lower bound		Real upper bound	
	Liquidity usage, π	Settlement delay, ρ	Liquidity usage, π	Settlement delay, ρ
RTGS with subnetting	42	12	58	0
Hybrid	18	6	27	0
Advanced Hybrid	17	5	26	0
RTGS with queuing	15	7	25	0

The liquidity need increased for all the systems when RTGS transfers were settled immediately without queuing instead of by netting at the end of the day. The tradeoffs between liquidity usage and settlement delay were found to be fairly restricted. In terms of liquidity usage alone, the RTGS-with-queuing structure appeared to be superior. However, at low levels of liquidity, the RTGS-with-queuing structure entailed longer settlement delays than the Hybrid systems.

3.6.2 Simulations of pre-selected payment system structures

RTGS-with-subnetting vs Hybrid structure

The standard deviation of banks' intraday net balances is larger for the Hybrid structure than for the RTGS-with-subnetting structure. The larger standard deviation suggests that the time-distribution of incoming and outgoing payments is less balanced with the Hybrid structure than with the RTGS-with-subnetting structure.

With the Hybrid structure, the need for intraday credit is significantly greater for the whole banking sector as well as for some individual banks. The extent of queuing is also somewhat greater with the Hybrid structure. On average, 10 payments were queued daily for an average time of 45 minutes. The longest queuing time was about 3 hours. This suggests that on average the existing intraday credit limits are sufficient but that in some cases extra intraday credit or other extra liquidity might be needed.

The average value of a queued payment in the Hybrid structure was however relatively low, ie about ECU 21.4 million, and the average aggregate value of a payment queue was about ECU 43.8 million, suggesting little need for extra liquidity. Moreover, the number of queued payments was very low, ie 2.2 on average for times when there were queues. The most payments in a queue at one time was 15, with a total value of ECU 204.5 million. With the RTGS-with-subnetting structure, which does not entail queuing, only two payments could not be settled immediately and were re-entered into the system at the earliest possible settlement time.

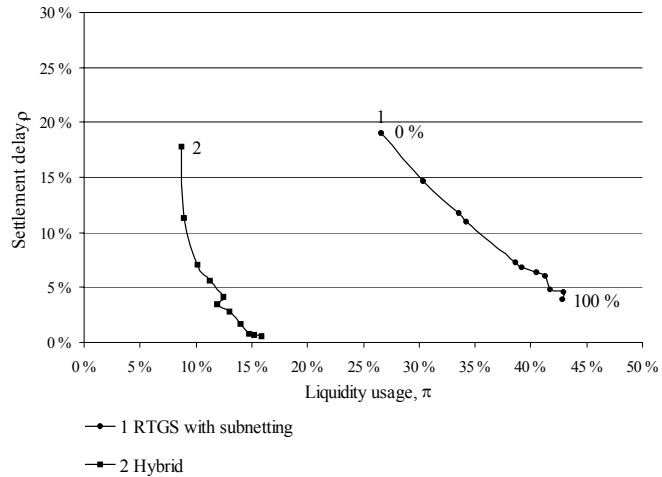
The relative efficiency of RTGS settlement as between the RTGS-with-subnetting and Hybrid structures is shown in figure 3.7. The curves are based on the different available liquidity levels (0 to 100 per cent, ie from theoretical lower to upper bound) and show the liquidity usage relative to the value of outgoing payments (π) and corresponding settlement delay (ρ). The Hybrid structure uses only about a third as much liquidity for a given value of payments and given settlement delay time compared to the RTGS-with-subnetting structure, and is thus much more efficient.

The theoretical system upper bound (UB_t) for the RTGS-with-subnetting structure (ie 100 per cent of the level of liquidity) results in liquidity usage amounting to about 43 per cent of total value sent. The real upper bound (UB_r) results in liquidity usage of about 37 per cent, as was shown in table 3.10. For the Hybrid structure, the corresponding figures are about 16 per cent and 22 per cent.

The curve for the Hybrid structure is concave between available liquidity levels of 30 per cent and 50 per cent, which means that reductions in available liquidity result in queuing at critical points of time during the day and the banks are forced to raise their intraday credit limits in order to settle time-critical transfers. The rise in the limits is greater on average than the reduction in liquidity usage due to the use of lower intraday credit limits. The same kind of concavity can be seen with the RTGS-with-subnetting structure between the liquidity levels of 80 per cent and 100 per cent. This topic was discussed earlier in section 3.3.2.

Figure 3.7

Relationship between system settlement delay (ρ) and liquidity usage (π) in RTGS-with-subnetting and Hybrid structures, actual payment data, 4 days



Another interesting result is that the curve for the Hybrid structure is almost vertical between the two lowest levels of available liquidity (0 and 10 per cent), and between other levels of liquidity the curve is quite steep. This suggests that the selection of a lower liquidity level will not significantly reduce liquidity usage but will result in a significant increase in settlement delay.

The differences in the features of the two structures can explain the significant difference in performance. In the RTGS-with-subnetting structure, the number of transactions is low but their average value is high. The transactions consist mainly of large-value transfers settled via RTGS and settlements of net positions originating from net settlement systems. Because the individual transfers are large and the number of payments small, the liquidity need relative to the value of payments flowing through the system is larger with the RTGS-with-subnetting structure than with the Hybrid structure.

There is no queuing or optimisation routine in the RTGS-with-subnetting structure. In the Hybrid structure the netting of queues reduces liquidity usage by solving gridlocks during the day and by possibly clearing at the start of the day any queues awaiting the opening of the system.

Hybrid vs Advanced Hybrid structure

Differences between the Hybrid and Advanced Hybrid structures (structures 2 and 3) are very small. Differences in standard deviation of balances and average limit usage are insignificant.

With the Advanced Hybrid structure, average queuing time was lower (in two of the three days with queuing) and aggregate value of queued payments slightly lower, whereas the average number of queued payments was 4.5 compared to 2.2 for the Hybrid structure. This is due to the payments splitting in the Advanced Hybrid structure. The peak queuing time is the same for the Hybrid and Advanced Hybrid structures.

The netting of the queues every 20 minutes had only minor effects on liquidity needs and settlement delay, as it succeeded only once, at the start of day. The splitting of payments enhanced the use of existing liquidity to some extent and reduced the average value of queued payments while increasing the number of queued payments.

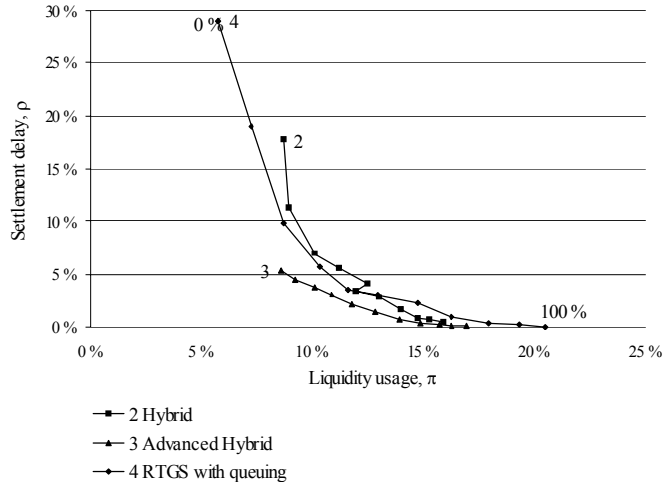
The RTGS-with-queuing structure does not differ much from the other two structures. In terms of average usage of intraday credit limits and queuing times, this structure was superior on two days and inferior on the two other days. However, the differences were small. In terms of number and average value of queued payments, the performance of the RTGS-with-queuing structure was in between the Hybrid and Advanced Hybrid structures. On the other hand, the queues for the RTGS system represented higher value than for the other structures. The average value of a queue was about three to four times that for the Hybrid or Advanced Hybrid structure. Also, the peak value of a queue during the simulation period was about twice as high, ECU 459.5 million.

The relative efficiencies of the Hybrid, Advanced Hybrid and RTGS-with-queuing structures are illustrated in figure 3.8.

The Advanced Hybrid structure is superior from the banks perspective at all liquidity levels vs the other two structures. This suggests that, compared to the Hybrid structure, settling within-limit POPS payments on a gross basis instead of netting them continuously increases system liquidity. However, the differences are small and hence it is safer to say that executing all POPS payments on a gross basis will not cause additional liquidity restraints vs the Hybrid structure, at least when payment splitting is used as a liquidity optimisation method. The splitting of payments apparently reduced settlement delay significantly.

Figure 3.8

Relationship between system settlement delay (π) and liquidity usage (ρ) for RTGS-with-queuing, Hybrid and Advanced Hybrid structures, actual payment data, 4 days



On the other hand, the two PMJ clearings in the Advanced Hybrid structure seem to level off the peaks in interbank payment flows and thus reduce liquidity usage at high available liquidity levels vs the RTGS-with-queuing structure.

At lower levels of available liquidity the RTGS-with-queuing structure uses the least liquidity of all the pre-selected payment system structures but at the cost of having clearly the most settlement delay. The RTGS-with-queuing structure also uses the most liquidity at the nonqueuing liquidity level.

Otherwise the differences between these structures are relatively small. It is noteworthy that the curves for the Hybrid and RTGS-with-queuing structures cross between liquidity levels of 50 per cent and 60 per cent. This suggests that if the banks select higher available liquidity levels, the Hybrid structure will be more cost-effective, whereas if they choose lower levels, the RTGS-with-queuing structure will be more cost-effective.

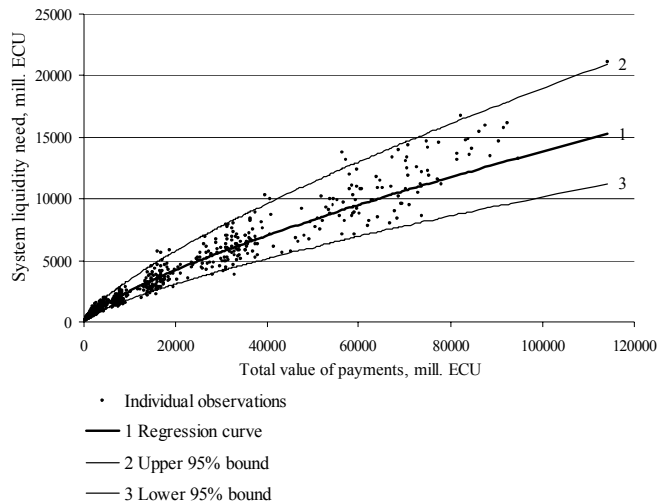
3.6.3 Total daily value of payments and liquidity need

In order to study the relationship between the volumes of payments and the liquidity need generated payment data was used. The generated data was sampled from the pool of actual payments – only the number of transaction was higher than in the actual data set.

The scatter diagram in figure 3.9 shows the daily values of payments settled and corresponding liquidity needed for immediate settlement in individual simulation runs using RTGS with queuing. The two variables are very highly correlated (correlation coefficient approx. 98 per cent). The relationship is slightly loglinear, especially at low aggregate settlement values. At higher values of payments, the relationship can be considered as approximately linear. With large daily payment values, the law of large numbers begins to have an effect and the marginal liquidity need remains more stable.

Figure 3.9

Relationship between system upper bound of liquidity need and daily value of payments in RTGS-with-queuing structure, n=700

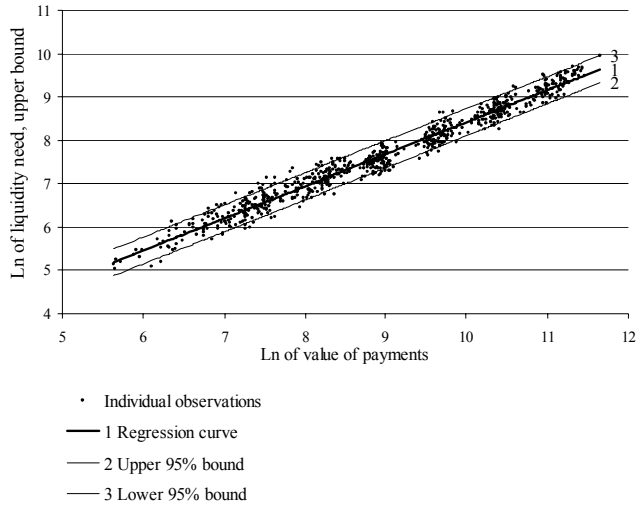


The scatter diagram in figure 3.9 indicates some heteroscedasticity, ie the variance of the distribution of liquidity need tends to increase as daily payments value increases. The heteroscedasticity disappears when logs are used. The transformed scatter diagram with the

estimated regression line and 95 per cent confidence intervals are shown in figure 3.10.

Figure 3.10

Relationship between system liquidity need and value of payments, logarithmic scale, n=700



The estimated regression line in its linear form is presented in equation 3.1 and in its exponential form in equation 3.2. The fit is exceptionally good; the value of R-squared, ie the percentage of variation in the liquidity need that can be explained by the variation in the daily value of payments, is 97 per cent. The R.M.S error of the regression line is 0.19 and, as the residuals are normally distributed, the 95 per cent confidence intervals (shown in figures 3.10 and 3.11) can be calculated.

$$\ln(UB_r) = b * \ln(V) + I + \varepsilon \quad (3.7)$$

Equation 3.7: Regression curve in linear form for the system upper bound of liquidity in the RTGS-with-queuing structure as a function of daily value of payments settled (UB_r = upper bound of liquidity, V = value of payments, I = intersect, ε = error term)

$$UB_r = V^b * e^{1+\varepsilon} = V^{0,739} * e^{1,025+\varepsilon} \quad (3.8)$$

Equation 3.8: Regression curve in exponential form for the system upper bound of liquidity in the RTGS-with-queuing structure as a function of daily value of payments settled

The relationship between π and daily value of payments is depicted in figure 3.11 (π represents the amount of liquidity used for settlement as a share of the total value of payments). The value and variance of π decrease as the daily value of payments increases. As daily payments value increases, the incoming and outgoing transfers more evenly offset each other during the day and so less liquidity per value sent is needed. Moreover, the ‘shock effect’ of large payments on liquidity requirements is reduced, as their share in the total value of payments diminishes.

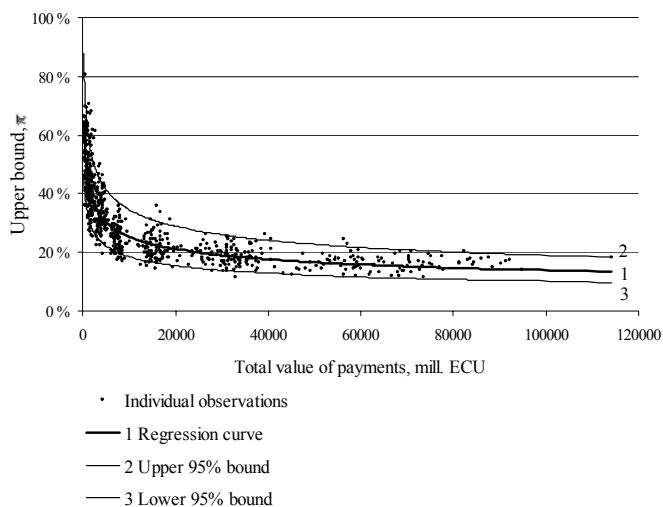
The mathematical formulation of the regression curve is illustrated in equation 3.7. From the equation we can see that π approaches zero asymptotically as the value of payments approaches infinity. The slope of the curve is very modest at high values of payments. The total value of payments should be 30 times bigger than the prevailing value of payments in October 1998 in order for π to be below 10 per cent on average.

$$\pi = \frac{UB_r}{V} = \frac{V^b * e^{1+\varepsilon}}{V} = V^{-0,261} * e^{1,025+\varepsilon} \quad (3.9)$$

Equation 3.9: The regression curve for predicting system from the value of payments settled

Figure 3.11

Relationship between system and daily value of payments, n=700, mill. ECU



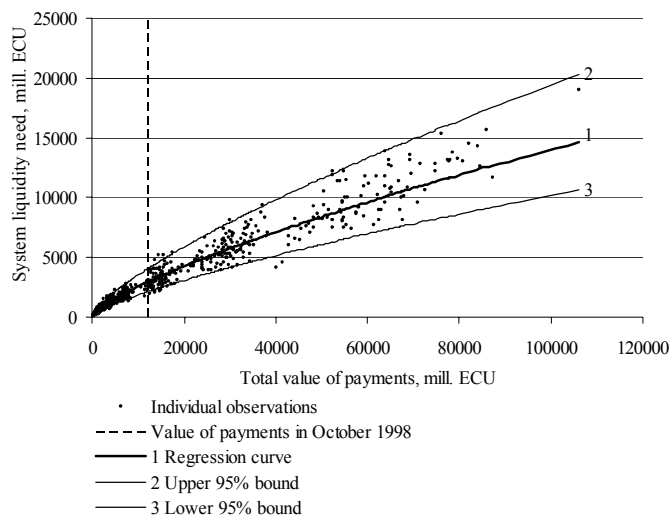
At low values of payments, chance plays a bigger role and the variation in the liquidity need is greater than with higher values of payments. At very low payments values, the liquidity need of the system can reach up to 90 per cent of the gross value of payments. At very high values of payments, the system liquidity need is more predictable and much lower in relative terms.

The Hybrid structure was simulated with the same payment data. Because payments in the Hybrid structure are settled not only in the RTGS system but also in the POPS system and in two net settlements, the payment flow in the RTGS system is somewhat different than in the case of RTGS with queuing. In this structure, prioritisation of payments is also accounted for, and some transfers are considered to be time-critical. In this section the Hybrid structure is analysed for its liquidity needs in the RTGS system where net settlement transactions from other payment systems are included.

The results on the Hybrid structure are very similar to those from the simulations on the RTGS-with-queuing structure. The relationship between liquidity need for immediate settlement and value of payments settled is shown in figure 3.12.

Figure 3.12

Relationship between daily value of payments and system upper bound of liquidity need, Hybrid structure, n=700

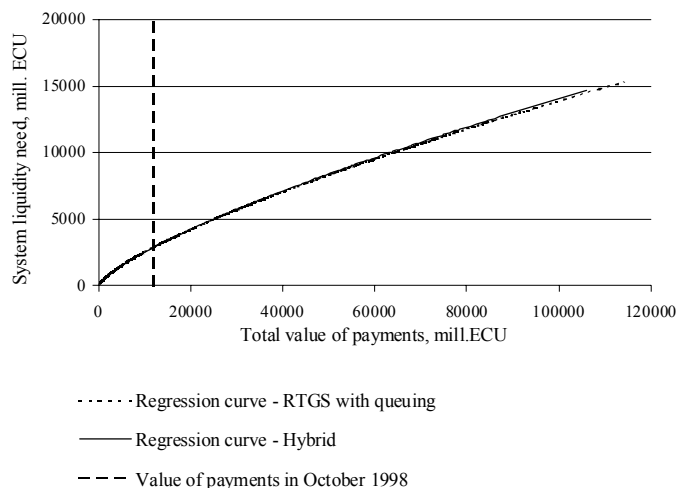


R-squared value for the regression curve was as good as in the case of RTGS with queuing, approximately 97 per cent. The R.M.S error of the regression line was slightly higher, approximately 0.20, and thus the confidence interval was slightly wider.

The regression curves for both systems are shown in figure 3.13. The curve representing the liquidity need of the RTGS-with-queuing structure is lower at payment values in excess of ECU 1 824 million whereas the Hybrid structure uses less liquidity at lower values of payments settled. However, the differences are very marginal. In the Hybrid structure, the time-criticalness of some payments increased the system liquidity need, but the liquidity optimisation feature of netting the queued transfers was also used. As the value of payments flowing through the RTGS system was lower in the Hybrid structure, its curve in figure 3.13 is shorter. The mathematical formulation of regression curve for the Hybrid structure is defined in equation 3.8.

Figure 3.13

Relationship between value of payments and system upper bound of liquidity need RTGS-with-queuing and Hybrid structures



$$UB = V^b * e^{I+\varepsilon} = V^{0,742} * e^{1,002+\varepsilon} \quad (3.10)$$

Equation 3.10: System upper bound for liquidity in the Hybrid structure as a function of the value of payments settled

According to the regression curve for the Hybrid structure, the liquidity needed for immediate settlement of this amount would be ECU 2.89 billion on average and, on 95 per cent of the days, the need would be less than ECU 3.95 billion. The actual and estimated liquidity needs are summarised in table 3.12. The estimates assume that the liquidity is optimally distributed among the participants.

In October 1998 intraday credit extensions for the banks participating in the BoF-RTGS totalled ECU 3.16 billion, the required reserves available for settlement purposes ECU 0.86 billion and excess reserves ECU 0.05 billion, giving a total of ECU 4.07 billion. If all liquidity available for the banks were optimally distributed, queuing would take place only on a few days out of a hundred, according to the estimate.

Table 3.12

**Actual and estimated liquidity needs,
bill. ECU**

Liquidity in October 1998		
Intraday credit limits		3.16
Required reserves		0.86
Excess reserves		0.05
Total		4.07
Estimated liquidity need		
	RTGS with queuing	Hybrid
Average	2.89	2.90
95% of days under	3.95	4.01
99% of days under	4.49	4.58

The liquidity needs of the Finnish banking sector seem to increase somewhat with the shift from the RTGS-with-subnetting structure to the Hybrid structure at the start of 1999. The simulations suggest that on average the existing intraday credit limits are sufficient, albeit some banks may need extra intraday credit or other extra liquidity. Although the liquidity need increases, the value of payments going through the system increases much more, ie the system works more efficiently. The result is that the Hybrid structure uses only slightly more than a third as much liquidity for a given amount of payment volume and settlement delay as the RTGS-with-subnetting structure and is thus much more efficient.

Differences between the Hybrid and Advanced Hybrid structures were found to be very small. This suggests that making all POPS payments on a gross basis will not cause additional liquidity restraints, at least when payment splitting is used as a liquidity optimisation method.

When TARGET transactions are introduced, liquidity needs increase slightly (as anticipated) as the value of payments increases. Liquidity usage decreases slightly and the change in settlement delay is insignificant. Thus one can say that, according to these simulations, inclusion of TARGET transactions with the Hybrid structure does not impose additional liquidity restraints on banks operating in the Finnish interbank payment system.

The real time gross settlement of all payments did not cause additional liquidity restraints or settlement delay. Only the occurrence of gridlocks was found to be greater for the RTGS-with-queuing structure than for the other systems. The probability of gridlock in the

pre-selected scenarios was generally found to be fairly small. The introduction of optimisation features like netting of queues (Hybrid and Advanced Hybrid structures) and splitting of payments (Advanced Hybrid structure) helped resolve gridlock situations.

3.7 Conclusions

The liquidity need of the Finnish banking sector seems to increase with a changeover from the RTGS-with-subnetting payment system structure to the Hybrid structure. Although the total value of payments flowing through the RTGS system was higher on average in the Hybrid structure, the current intraday credit limits of the banks were sufficient. However, some banks experienced delays in settlement, and any additional liquidity would reduce such delays. The upside of the shift from net to gross settlement is of course the greatly reduced overall settlement delay and settlement risks.

No highly significant differences were found between the Hybrid, Advanced Hybrid, and RTGS-with-queuing structures. However, total time of system gridlock was greater (albeit still quite small) for the RTGS-with-queuing structure than for the other systems. The RTGS-with-queuing structure did not include any optimisation features, and the inclusion of such features in the Hybrid and Advanced Hybrid structures helped resolve most of the gridlock situations fairly quickly.

The introduction of TARGET transactions had only a minor influence on banks' liquidity needs and settlement delay. Thus it can be said that according to these simulations the inclusion of TARGET transactions in the Hybrid model does not impose additional liquidity restraints on banks operating in the Finnish interbank payment system, although the absolute values of liquidity needs are somewhat higher.

Liquidity needs increase considerably for all payment system structures studied if the banks settle their payments immediately without queuing instead of settling net positions at the end of the day. It should however be recalled that an RTGS system is capable of operating at the same liquidity level as a net settlement system. In an RTGS system, settlement delay is explicit in the queuing of payments, and in a TDNS with the same risk characteristics, fund transfers become final when the net positions are settled between the settling banks. In the simulations, the RTGS-with-queuing structure was about three and a half times faster in settling payments than an end-of-day net settlement system with the same amount of liquidity.

Focusing solely on liquidity usage, the RTGS-with-queuing structure seems to be the superior system structure for the banks. However, at the lower levels of liquidity, this structure results in extensive, perhaps intolerable, settlement delays. This suggests that the RTGS-with-queuing structure may be the best choice only if the banks operate at high liquidity levels. Thus both the Hybrid and Advanced Hybrid structures can be seen as good compromises between liquidity usage and settlement delay at all liquidity levels. The results using both generated and actual payment data pointed in the same direction.

This study concentrated on the settlement and system scenarios of the simulator in assessing liquidity needs and corresponding settlement delays, but there are other factors that affect banks' liquidity needs. In this study the banking structure and payment characteristics were kept the same.

It might be useful to pursue further study of the effects of different banking structures. The liquidity needs of a system with equal-size banks might differ from a system with banks of differing sizes but with the same total value of payments. In this study also the daily value distribution of the payment data was approximately the same over the 100-day period. With a different structure of small and large payments, liquidity need and settlement delay might differ substantially.

The optimisation methods tested were the splitting of payments and netting of queues. Further studies could be done on payment splitting that uses all available liquidity or routines for bilateral netting of queues. Also the effects of different queuing arrangements and algorithms not based on the FIFO principle might provide interesting topics for study.

Risk considerations were not addressed in this study although the simulator can simulate bank failures and settlement delays. One could assess eg the systemic risk inherent in the Finnish payment system or other systems.

The results concerning the adequacy of liquidity and the bounds for liquidity in the EMU structure with TARGET transactions can be refined as actual data on TARGET transactions becomes available. After some time, as historical data in the context of the EMU becomes available, one will be able to more accurately measure the effects of optimisation routines using a longer period of actual payment data.

References

- Angelini, P (1994) **About the Level of Daylight Credit, Speed of Settlement and Reserves in Electronic Payment Systems.** Banca D'Italia Discussion Paper No. 229.
- Angelini, P and Giannini, C (1993) **On the Economics of Interbank Payment Systems.** Banca D'Italia Discussion Paper No. 193.
- Bank of Finland (1992) **Suomen maksu- ja selvitysjärjestelmät.** Bank of Finland Studies A:84. Helsinki.
- Bank for International Settlements (1989) **Report on netting schemes (Angell report).** Prepared by the Group of Experts on Payment Systems of the central banks of the Group of Ten countries. Basle.
- Bank for International Settlements (1990) **Large-value Funds Transfer Systems in the Group of Ten Countries.** Prepared by the Group of Experts on Payment Systems of the central banks of the Group of Ten countries. Basle.
- 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).** Prepared by the Group of Experts on Payment Systems of the central banks of the Group of Ten countries. Basle.
- Bank for International Settlements (1993) **Central Bank Payment and Settlement Services with Respect to Cross-border and Multi-currency Transactions.** Prepared by the Committee on Payment and Settlement Systems of the central banks of the Group of Ten Countries. Basle.
- Bank for International Settlements (1994) **Payment Systems in Finland.** Prepared by the Bank of Finland and the Committee on Payment and Settlement Systems of the central banks of the Group of Ten countries. Basle.
- Bank for International Settlements (1997) **Real-time Gross Settlement Systems.** Prepared by the Committee on Payment and Settlement Systems of the central banks of the Group of Ten Countries. Basle.

- Berger, A N, Hancock, D and Marquardt, J C (1996) **A Framework for Analysing Efficiency.** Risks, Costs, and Innovations in the Payment System. Journal of Money, Credit, and Banking. Vol. 28 No. 4.
- Boeschoten, W C (1989) **In Search of an Efficient Interbank Settlement System.** Paper prepared for Eurobanking 1989. Economic Research and Special Studies Department. De Nederlandsche Bank.
- Borio, C, Russo, D and Van den Bergh, P (1992) **Payment System Arrangements and Related Policy Issues: A Cross Country Comparison.** SUERF Papers on Monetary Policy and Financial Systems No. 13. Société Universitaire Européenne de Recherches Financières. Tilburg.
- Bossel, H (1994) **Modeling and Simulation.** Univerität Gesamthochschule Kassel. Verlag Vieweg. Wiesbaden.
- Calomiris, C W and Kahn, C M (1996) **The Efficiency of Self-Regulated Payment Systems: Learning from the Suffolk System.** Journal of Money, Credit, and Banking. Vol. 28, No. 4.
- European Monetary Institute (1996) **Payment Systems in the European Union.** Frankfurt am Main.
- Flannery, M J (1996) **Financial Crisis; Payment System Problems, and Discount Window Lending.** Journal of Money, Credit, and Banking. Vol. 28, No. 4.
- Folkerts-Landau, D (1997) **Wholesale Payments and Financial Discipline, Efficiency, and Liquidity.** IMF Working Paper WP/97/154. International Monetary Fund.
- Folkerts-Landau, D, Garber, P and Schoenmaker, D (1990) **The Reform of Wholesale Payment Systems and its Impact on Financial markets.** Group of Thirty Occasional Paper 51. Washington DC.

- Furfine, C and Stehm, J (1997) **Analysing alternative intraday credit policies in real-time gross settlement systems**. Board of Governors of the Federal Reserve System (U.S.). Finance and Economics Discussion Series 1997–1940.
- Kobaykawa, S (1997) **The Comparative Analysis of Settlement Systems**. CEPR Discussion Paper No. 1667. Centre for Economic Policy Research. London, July.
- Kuussaari, H (1997) **Systemic Risk in the Finnish Payment System: an Empirical Investigation**. Bank of Finland discussion paper No. 3/97. Helsinki.
- Leinonen, H (1998) **Interbank funds transfer systems: liquidity needs, counterparty risks and collateral**. Bank of Finland Discussion Paper No. 16/98. Helsinki.
- Leinonen, H and Saarinen, V (1998) **Payment system risks in Finland and the need for regulation and supervision**. Bank of Finland studies A:101. Helsinki.
- Marviala, J and Palva, M (1995) **The Bank of Finland's Real-time Gross Settlement System**. Bank of Finland Bulletin, No. 5. Helsinki.
- Schoenmaker, D (1993) **Externalities in Payment Systems: Issues for Europe**. Financial Markets Group. London School of Economics. London.
- Schoenmaker, D (1995) **A Comparison of Alternative Interbank Settlement Systems**. LSE Financial Markets Group discussion paper No. 204. ESRC Research Centre. London.
- Working Group on EU Payment Systems (1995) **Report to the Council of the European Monetary Institute on the TARGET System**.

Chapter 4

Optimising liquidity usage and settlement speed in payment systems

Harry Leinonen – Kimmo Soramäki

4	Optimising liquidity usage and settlement speed in payment systems.....	118
	Abstract.....	118
4.1	Introduction.....	118
4.2	Interbank payment settlement designs.....	120
4.3	Previous research on payment systems.....	122
4.4	Simulation framework.....	124
4.4.1	Payment and settlement delay.....	124
4.4.2	Concepts of liquidity.....	125
4.4.3	Cost components in settlement systems.....	128
4.4.4	Simulated settlement systems.....	130
4.4.5	Data.....	132
4.5	Simulation results.....	133
4.5.1	Relationship between liquidity usage and settlement delay.....	133
4.5.2	Combined cost of liquidity and delay.....	135
4.5.3	Gridlocks in RTGS systems.....	138
4.5.4	Optimising an RTGS system.....	141
4.6	Conclusions.....	145
	References.....	147

4 Optimising liquidity usage and settlement speed in payment systems

Abstract

We quantify the relationship between liquidity usage and settlement delay in net settlement systems, real-time gross settlement systems and hybrid systems, as well as the combined costs of these. We analyse ways of reducing costs via optimisation features such as netting of queues, offsetting of payments and splitting of payments. The empirical component uses a payment system simulator developed at the Bank of Finland. The data used covers 100 days of actual payments processed in the Finnish BoF-RTGS system. We find that a system with queuing facilities where settlement takes place continuously in real time is more efficient from the perspective of liquidity and risks than a net settlement system with batch processing. Real-time processing enables a reduction in payment delay and risks without necessarily increasing liquidity needs. Participants will operate under immediate payment/settlement if liquidity costs are low enough relative to delay costs and if the liquidity arrangements are sufficiently flexible. The central bank can therefore support risk reduction and payment speed objectives by providing low-cost intraday liquidity as well as more flexible ways for participants to add or withdraw liquidity from the system. Optimising and gridlock-solving features were found to be effective at very low levels of liquidity. Gains from netting schemes with multiple daily netting cycles were found to be somewhat more limited.

4.1 Introduction

The operating speed of a payment system has always depended on the technological level of the system's communication and information processing environment. The use of information technology in batch processing has enabled a reduction in payment processing time from several days in the foregone manual era to typically one day. Frequent intraday processing cycles, and especially real-time processing, have introduced new means of further speeding up payment processing and

settlement. With increasing payment volumes, there is now a need to reassess risk management in existing payment settlement systems. One way to reduce interbank settlement risks is to speed up the settlement process.

In a real-time environment, banks face new challenges in liquidity management. They need to plan for intraday as well as interday fluctuations in liquidity. Not all payments in a real-time environment require immediate processing. This gives system participants the opportunity to employ different types of hybrid settlement structure, which enables the evening out of intraday fluctuations in liquidity demand. The aim of this study is to develop a framework for analysing fluctuations in liquidity demand and assessing the efficiency of different settlement systems, particularly from the viewpoint of liquidity and settlement speed.

We try to determine system-level cost-optimal liquidity usage for various settlement systems and compare the systems as they operate with optimal amounts of liquidity. We quantify the relationship between liquidity usage and settlement delay in net settlement systems, real-time gross settlement systems and hybrid systems, as well as the combined costs of liquidity and delay in these systems.

At the time of writing, most systems do not employ advanced optimisation mechanisms for the settlement of payments. To investigate potential improvements for interbank payment systems we analyse by means of simulations ways of reducing costs by netting queues, offsetting payments and splitting payments. Whether the system-level cost-optimum is reached through independent decisions by system participants is not discussed here. Our major findings relate to risk reduction via real-time settlement, the effects of optimisation routines in hybrid systems, and the effects of liquidity costs on banks' choice of settlement speed. Whether the optimal settlement mode is real-time or delayed depends on the cost relationships between these factors. The study focuses on the optimisation of settlement processes under normal operating conditions. Exceptional situations relating to participant failures and general system crises are beyond the scope of the present analysis. However, the risk reduction measures studied would also be helpful in these kinds of exceptional situation.

In the empirical component of our study, we employ a payment system simulator developed at the Bank of Finland. For a description of the simulator the reader is referred to Section 4.4 as well as Koponen and Soramäki (1998). The simulator enables us to evaluate the impact of changes in system parameters, and thus to compare the effects of alternative settlement schemes with given payment flows. The random and structural characteristics of payment flows have a

major impact on the results. A payment system's liquidity need, assuming a given processing speed, is determined by the payment flows and by the specific settlement structure. The study is organised as follows. The next two sections provide a brief overview of alternative settlement designs and a summary of related research. In Section 4.4, the terminology, methods and data are described. In Section 4.5, we present our results concerning settlement speed, liquidity usage, risk and cost components as well as gridlock issues and optimisation possibilities. Section 4.6 concludes with a summary of the key findings and a discussion of related limitations and applicability.

4.2 Interbank payment settlement designs

A major distinction between different interbank payment systems is whether a system is operating on a net or gross basis and whether payments are processed individually or in batches. The most common three pure implementations of these principles are real-time gross settlement, time-designated net settlement and continuous or secured net settlement. Systems combining net and gross settlement or individual and batch processing are often referred to as hybrid systems.

A pure real-time gross settlement (RTGS) system is defined as a system in which, for each transaction, delivery of payment information and final settlement in central bank money take place simultaneously and continuously. Transfers are settled individually during the day without netting debits against credits. An RTGS system provides continuous intraday finality for the processed transfers (BIS 1997).

In a time-designated net settlement (TDNS) system, the settlement of payments occurs on a net basis at predefined points of time during the day or at the end of the day. The net position, ie the sum of payments the bank has received up to the end of the settlement period minus the payments it has sent, can be calculated on either a bilateral or a multilateral basis. Although many G10 countries have introduced real-time gross settlement systems, according to Folkerts-Landau et al (1996) most of the payment volume in industrialised countries is settled by multilateral netting.

After the interbank payment transfer is initiated and the customer's account debited, the settlement process is generally separated into two sub-processes: final customer crediting and final interbank settlement.

Except in pure RTGS systems, the payment information and settlement are processed in separate systems and there is a need for synchronisation at certain points in time. A counterparty or settlement risk will arise if the customer payment becomes final before the interbank settlement. In such case, settlement of the payment is delayed and the counterparty risk will include the total or net amount of outstanding delayed settlements, depending on the legal rules for netting claims. In order to reduce the risk, the receiving bank can delay payment to the customer. When this is the general practice, final settlements are processed simultaneously or prior to final crediting of the customer. In a system with rapid payment information processing, this will entail a delay in payments in order to allow the slower settlement process to catch up with the payment process. With the same payment flows and settlement times, the amount of delayed payments in a system without settlement risk will equal the total settlement delays in a system with settlement risk. We will refer to settlement delay and payment delay respectively in distinguishing between systems with and without settlement risk.

In continuous net settlement (CNS) systems, payments are credited individually and immediately to receivers' accounts, but final settlements occur periodically or at the end of the day. These systems entail settlement delay, and the amount of risk depends on the total or net value of delayed settlements. To control the risks, participants usually have credit limits, which may be partly or completely collateralised. An RTGS system with queuing arrangements operating with central bank money is analytically equivalent to a secured CNS system (SNS) with fully collateralised debt limits. In both designs, payments are credited to receivers' accounts only if there are sufficient covering funds or collateral. If a bank does not have sufficient liquidity to settle a payment, the payment is delayed and credited to the customer's account only after settlement. When we refer later to a CNS system, we will assume that it entails counterparty risk, ie that settlements are delayed.

Hybrid settlement systems try to combine elements from several of the above systems in order to achieve better performance than with pure systems. Hybrid systems are designed to optimise with respect to the different settlement delay and liquidity costs associated with the pure systems. For example, an RTGS system with queuing can be considered a hybrid system.

4.3 Previous research on payment systems

Previous research on payment systems has generally been policy-oriented, much of it being conducted by central banks around the world, mainly in Italy, the United Kingdom and the United States. Most research on payment settlement systems can be included in one of the following three categories: descriptions of current arrangements, analysis of the risks associated in these systems and central bank policy issues, and comparisons of net and gross settlement systems.

The settlement systems in use in different countries have evolved for the most part independently of each other. Until recently there has been little effort to harmonise or standardise these systems. The first category consists of studies that attempt to describe systems in use in different countries. The goal has been to find a common structure and the key similarities and differences between these systems. The study by Borio et al (1992) provides an overview of the common structure underlying different settlement systems. Another earlier study is the survey by the Committee on Payment and Settlement Systems (CPSS)¹ on large-value funds transfer systems in the G10 countries (BIS 1990). The report by CPSS on real-time gross settlement (BIS 1997) is a good review of the concepts relating to RTGS systems.

BIS (1989), Borio and Van den Bergh (1993) and Angelini et al (1996) provide good analyses of issues concerning systemic risk in payment systems, the emphasis being on netting systems. From the viewpoint of this study, the most interesting papers are those on systemic risk that simulate situations in which one or more system participants fail to settle their obligations. The data used in Humphrey (1986) are from the Clearinghouse Interbank Payment System (CHIPS)². McAndrews and Wasilyew (1995) build on Humphrey (1986) and use generated data to study factors that affect the systemic risk in a payment system. In Kuussaari (1996) the extent and effects of a systemic crisis in Finland are analysed empirically by using simulated bank failures. These studies use the same methodology, ie they use ex post settlement data to assess the impacts of bank failures on other system participants. The simulation model used in this study

¹ The Committee on Payment and Settlement Systems (CPSS) is a working group under the Bank for International Settlement (BIS). It was established by the governors of the central banks of the G10 countries to monitor and analyse developments in payment and settlement systems.

² CHIPS is an interbank settlement and payment transfer system organised by the New York Clearing House Association, a group of the largest banks in the city.

differs from the model used in those studies in that it simulates the operation of a settlement system rather than specific events occurring in a system. In our simulations, a bank may fail to settle, but this would be a simulation output rather than input.

The central bank policy issues studied include inter alia the central bank's intraday credit policy (eg Humphrey 1990, Furfine and Stehm 1997), the possible emergence of private intraday money markets (Rossi 1995), and the effects of these on monetary policy (Dale and Rossi 1996). The issue of externalities in payment systems, ie third party effects not internalised in the payment system (eg the effects of payment delays), have been studied by Schoenmaker (1993). The studies by Angelini (1998) and Kahn and Roberds (1998b) also deal with externalities; these studies analyse the effects on banks' settlement behaviour of insufficient or costly liquidity in an RTGS system. The common goal of this research has been to find methods by which the central bank can ensure the stability and smooth functioning of the payment system.

The third body of literature comprises studies that compare the efficiency of net settlement and real-time gross settlement systems. However, efficiency comparisons between different implementations of net or real-time gross settlement systems with queuing facilities are rare. These studies are generally analytical. Schoenmaker (1995) uses an analytical model to compare pure RTGS systems with net settlement systems with caps and loss sharing rules. Kahn and Roberds (1998a) compare the merits of net and gross systems in a framework of bank incentives and moral hazard problems. Kobayakawa (1997) probes whether there is a rationale for gross and net settlement systems to coexist in the same economy. A recent study by Freixas and Parigi (1998) analyses the trade-off between risks and efficiency in net and gross settlement.

Simulations that compare and quantify the efficiency of different implementations of real-time gross settlement systems or net settlement systems have been very rare, perhaps due to data security issues and the demanding processing requirements of simulations. However, some studies have been done. Günzter et al (1997) present several heuristic algorithms for bilateral and multilateral netting of payments in reference to the German payment settlement system Elektronische Abrechnung mit Filetransfer 2 (EAF-2)³. Ganz et al (1998) simulate the efficiency of these algorithms in securities

³ The Elektronische Abrechnung mit Filetransfer 2 (EAF-2, currently Euro Access Frankfurt) is the largest interbank settlement system in Germany and is operated by the Hessian branch of the Bundesbank.

settlement. An earlier study by Boeschoten (1989) uses simulations to assess the impacts of different queuing mechanisms on the Dutch interbank payment system, which operates on the basis of real-time gross settlement. Non-academic simulation exercises in the planning phase of a new system or in assessing the efficiency of existing systems have been done in several countries. The Swiss National Bank has used simulations in assessing the queuing mechanism in SIC⁴ (Vital and Mengle 1988, Vital 1990 and 1994), the Association for Payment Clearing Services (APACS)⁵ in refining the CHAPS⁶ system (Bowman 1995), and recently the Banque de France in introducing its new RTGS system. For the Continuous Linked Settlement (CLS) system, which will be operational in mid-2000, simulations have been run with different settlement volumes in order to assess its operational capacity and the planned paying schedules.

4.4 Simulation framework

4.4.1 Payment and settlement delay

The unit of analysis in this study is a settlement system comprising banks and non-bank entities, which are referred to as system participants. The participants are continuously sending and receiving payments during the day, and the settlement of these payments is recorded, according to established settlement rules, in the settlement accounts.

In this study, the payment or settlement delay for each payment is calculated as the time difference between payment origination by the sending bank and final and irrevocable settlement of the payment. The cost of delay is likely to depend on the value of the delayed payment as well as the duration of the delay.

Let the individual payments be indexed by k . Let s be the value of each payment, t' the time of payment origination, t the time of final and irrevocable settlement of the payment, and T the time all

⁴ SIC (Swiss Interbank Clearing) is a payment settlement system operated by Telekurs SIC AG under the authority of the Swiss National Bank (SNB).

⁵ APACS (Association for Payment Clearing Services) is the industry body for the UK's major banks and building societies and has 29 members. It oversees money transmission and has responsibility for the cooperative aspects of the UK payments industry.

⁶ CHAPS (Clearing House Automated Payment System) is an electronic credit transfer system for sending payments between its members in the UK. Each payment is settled in real time across settlement accounts at the Bank of England.

payments submitted must be settled. The delay indicator, ρ , for the system, is then calculated as follows

$$\rho = \frac{\sum_k (t_k - t'_k) s_k}{\sum_k (T - t'_k) s_k} \quad (4.1)$$

If all payments are settled immediately upon entry into the system, $\rho = 0$ as $t'_k = t_k$ for every k ; if all payments are delayed until the end of the day, $\rho = 1$ as $t_k = T$ for every k . For an individual participant, the delay indicator ρ_i can be calculated similarly by including only those payments that are submitted and settled by participant i .

4.4.2 Concepts of liquidity

In modern payment systems where payments are processed in real time or in batches during the day, liquidity not only has end-of-day value but also intraday value. We analyse a settlement period of one day, during which the liquidity available to the participants is either employed for settlement purposes or remains idle on participants' settlement accounts. The amount of liquidity being employed by a participant i at time t when payments are settled immediately upon entry to the system depends on the flow of incoming and outgoing payments.

We are here interested in the receiver, the sender and the timing of payments, and denote the value of a payment from participant i to participant j at a (for each payment) unique time t as $s_{i,j}^t$. The amount of liquidity employed by participant i at any time \hat{t} during the day equals the excess value of outgoing payments, $s_{i,j}^t$, as compared to incoming payments, $s_{i,j}^t$ up until time \hat{t}

$$LE_{i,\hat{t}} = \max \left(0, \sum_{t=0}^{\hat{t}} \sum_{j \neq i} (s_{ij}^t - s_{ji}^t) \right) \quad (4.2)$$

One of the main decision variables in the study is the amount of liquidity available to the participants initially at the start of the day, for which we use the theoretical bounds of liquidity as calculated in Koponen and Soramäki (1998). To calculate these bounds, it is

assumed that the external liquidity source for any participant during the day is limited to its initial liquidity, ie the initial balance on its central bank settlement account and a possible intraday overdraft. Because the delaying of payments and settlements are analytically equivalent in this context, we use payment delay to describe both situations where payments are delayed to allow the settlement process to catch up and situations where only the final settlements are delayed. If account holders are required only to have sufficient funds at the end of the day ($t = T$), they can delay their payments until the end of the day and need only enough liquidity to cover the net amount of all incoming and outgoing payments at the end of the day, ie the lower bound (LB) of liquidity

$$LB_i = LE_{i,T} \quad (4.3)$$

If during the day sufficient liquidity is not available, payments are delayed. At the lower bound of liquidity, payment delay is at its maximum and any reduction in liquidity would mean a failure to settle one or more payments during the day. The liquidity need for an end-of-day net settlement system equals the lower bound of liquidity. In fact, the lower bound of liquidity marks the minimum external liquidity needed in all settlement systems where payments can be delayed eg by queuing, whether operating on a net or a gross basis.

The upper bound (UB) of liquidity equals the amount of liquidity that must be available to the participants for immediate settlement throughout the day. Any additional liquidity above the upper bound remains idle on participants' settlement accounts for the whole day. If payments are settled immediately upon entry to the system, the upper bound equals the maximum amount of liquidity employed at any time during the day

$$UB_i = \max(LE_{i,t}), \quad t \in \{0, \dots, T\} \quad (4.4)$$

In pure real-time gross settlement systems without the possibility of delaying payments, participants need liquidity at least equal to the upper bound in order to completely avoid settlement failure of any payment during the day.

Settlement system participants can acquire liquidity from many sources. For simplicity, we assume that all liquidity is in the form of funds available in central bank settlement accounts. These funds can be made available through different types of credit and reserve arrangements. For the designated liquidity indicator, π^d , we assume

that a pre-specified amount of funds is deposited at the central bank at time $t = 0$ and these cannot be withdrawn or increased during the day. π^d is calculated for each system participant as the ratio of its available liquidity (L) to the total value of its outgoing payments during the day and is equivalent to the inverse of the turnover ratio

$$\pi_i^d = \frac{L_i}{\sum_{t=0}^T \sum_{j \neq i} S_{ij}^t} \quad (4.5)$$

For the continuous liquidity indicator, π^c , we assume that participants can increase and reduce the external funds in the system (eg by changing the intraday overdraft limit) without friction. The opportunity cost of having liquidity in this flexible regime equals the amount of liquidity actually employed for settlement at each point of time. The advantage of the flexible liquidity regime lies in the fact that peaks in liquidity need during the day can be met by temporarily increasing liquidity. If alternative profitable investment opportunities for these funds or the collateral employed exist, liquidity costs can be reduced. The indicator for continuous liquidity usage, π^c is calculated as follows

$$\pi_i^c = \frac{\sum_{t=0}^T LE_{i,t}}{\sum_{t=0}^T \sum_{t'=0}^t \sum_{j \neq i} S_{ij}^{t'}} \quad (4.6)$$

The nominator equals the cumulative sum of liquidity employed at each point of time during the day. The denominator equals the cumulative sum of payments sent by the bank over the day.⁷

Both indicators of liquidity usage range from zero to one. If $\pi = 0$, a participant's external liquidity need is zero and it receives all the liquidity it needs for immediate payment settlement from incoming payments. If $\pi = 1$, the participant's liquidity usage equals the gross value of its payments.

⁷ The denominator gives the same result as the denominator in equation (5.1). In the former the sum is over payments, and in the latter over the time.

4.4.3 Cost components in settlement systems

In the model, we distinguish between the costs of obtaining liquidity and costs related to delaying payments or settlements. The cost differences in the settlement systems studied depend on the amount of liquidity used, on the resulting payment or settlement delay, and on the relative costs of these. We assume that the operating costs for the different systems are equal. Traditionally, netting systems have been considered operationally more efficient, and this is probably still true for batch-based retail payment systems. Real-time processing and the continuous risk management features required in these systems have evened out the operational efficiency differences between gross and net settlement in large-value payment and settlement systems. As no fundamental difference exists between large-value and retail payment systems, except for the number of transactions processed, it is foreseeable that with further advances in computer and communication technologies the efficiency differences between the two types of systems will also diminish.

The cost of obtaining liquidity depends on a number of market-specific factors. Liquidity obtained from the central bank entails explicit interest costs and/or implicit opportunity costs. Reserve requirements and collateral requirements for overdrafts impose opportunity costs, whereas intraday overdrafts can be explicitly priced. Central bank overdraft and collateral policy will therefore affect settlement costs and the choice of settlement scheme. Because of the inverse relationship between liquidity and delay costs, central banks can promote risk reduction and settlement speed by providing low-cost liquidity. A versatile and broad collateral base reduces opportunity costs. A flexible collateral management process that enables rapid and inexpensive transfer of collateral during the day makes it possible for central bank counterparties to use their trading portfolios as collateral. In the case of an intraday interbank money market, the available funds will normally be explicitly priced.

We calculate the cost of liquidity as $\pi \cdot r$, where π^d is used in the case of a rigid liquidity regime and π^c in the case of a flexible liquidity regime. We assume that the interest costs in both regimes are equal and use an interest rate of 2.5%. This is roughly the average Eonia rate for the period studied.

Costs related to settlement delays consist of credit risk and possible opportunity costs. When the receiving bank agrees to credit the customer finally before receiving the interbank settlement, it implicitly agrees to accept an interbank credit risk. If the receiving

bank has income-generating investment opportunities for the delayed funds, the delay will also mean forgone income possibilities. In our model we assume for simplicity that in payment systems operating with settlement delay the costs can be estimated as a linear function of delayed value and total delay. The cost of settlement delay will thereby take the form of an interest cost. The probability that settlement risks will be realised is not considered here.

Costs related to payment delays arise when the receiving bank interrupts customer payment processing until it receives the interbank settlement. Some customer payments are likely to be time-critical and any delays are likely to generate costs to the receiving and/or sending bank. To the receiving bank, these costs may be implicit, in the form of a deterioration in customer service; to the sending bank, they may be explicit, in the form of sanctions governing payment services. The interest cost of payment delay is assumed to increase linearly with the amount of payments delayed, but the other costs are more likely to rise at an increasing rate (Schoenmaker 1995, Angelini 1998). There is often a threshold beyond which delays are sanctioned more strongly. However, in our model the cost of the delay is calculated as $\rho \cdot r$ for payment systems operating with payment delay.

In both gross and net settlement systems, there is a clear relationship between liquidity usage and settlement/payment delay. The more liquidity used, the speedier the final settlements, up to the upper liquidity bound. If the cost of liquidity and delay are equal, the cost-optimal level of liquidity is likely to be that for which no payments are delayed. A delay in settlement reduces the sender's liquidity costs but increases both its delay costs and the receiver's liquidity costs. This creates a dead-weight loss at system level (Angelini 1998). On the other hand, in a rigid liquidity regime some participants may be able to delay their payments and use the receiver's idle liquidity without incurring costs. The amount of dead-weight loss at system level depends on the relative costs of liquidity and delay. If the sum of the delay costs to the sender and the average delay imposed on the receiver is smaller than the sender's cost of liquidity, the system-level cost-optimum is something other than the zero-delay level of liquidity. In terms of this study, dead-weight loss occurs at system level only if the combined cost of the system liquidity and delay increases as system liquidity is reduced. If the total cost decreases, the banks are able to utilise their idle liquidity and more efficient settlement of payments is achieved by delaying payments.

4.4.4 Simulated settlement systems

The settlement systems studied here are net settlement with different netting intervals, real-time gross settlement with and without queuing, and hybrid systems with optimisation features (netting of queues, offsetting of payments and splitting of payments).

The net settlement systems simulated are summarised in Table 4.1. These systems are simulated with settlement cycles only at the end of the day, three times a day (every 6 hours), 5 times a day (every 3 hours) and 12 times a day (every hour). In the simulations each participant has just enough liquidity to settle all its net positions during the day.

Table 4.1 **Simulations of net settlement systems**

Settlement periods	End-of-day, every 6h, every 3h, every hour
Settlement algorithm	Full
Intraday credit limits	Enough to succeed
Simulation period	100 days: 4 Jan – 21 May 1999

Real-time gross settlement is simulated with and without queuing of payments. The pure RTGS system without queuing is simulated only with the upper bound of liquidity because not all payments can be settled immediately with less liquidity. In the simulations with queuing in RTGS, we assume that banks do not queue their payments internally but enter them into the system immediately upon receipt of a payment order, and that there is no special prioritisation of payments. In order to assess the relationship between liquidity usage and payment delay, the simulations are run with eleven different available liquidity levels (in 10 %-point intervals) for each account holder, between the pre-calculated lower and upper liquidity bounds. The queuing algorithm used is based on the first-in-first-out (FIFO) principle. After settlement of a payment, the algorithm checks whether payments from the receiver of the payment can be settled with the liquidity received with the incoming payment. The procedure is continued until no payments are queued or no payments can be settled with the available liquidity. The latter happens if all participants with queued payments are illiquid or if the system is gridlocked.

In hybrid systems, features from net and gross settlement are combined to improve settlement performance. The optimisation methods analysed are the netting of queued payments and splitting of payments. The algorithms used for netting queued payments are full

and partial multilateral netting and bilateral netting of offsetting payments. The full multilateral netting algorithm fails and all payments are left queued if there is insufficient liquidity available for settling the net positions of all account holders. When the partial netting algorithm is applied, payments are inactivated until the system finds a feasible subset of payments that can be settled with the available liquidity. Inactivated payments are left queued and settled normally when cover becomes available or when they are included in the next netting. The partial net settlement used is a part of the multilateral netting algorithm used in EAF-2 and corresponds to the algorithm inactivation run with criterion two in Güntzer et al (1998). In the bilateral offsetting of payments each queued payment is checked for offsetting payments in the receiver's queue. The offsetting payments are removed from the receiver's queue, and the value of the processed payment is reduced by the value of the offsetting payments. The total value of offsetting payments can exceed the value of the processed payment so long as the sender of the offsetting payments has enough liquidity for immediate settlement.

Table 4.2 **Simulations of real-time gross settlement systems and hybrid systems**

Optimisation methods	Top 10%, 5%, 1% and 1‰ of payments split Netting of queues every 20 and 5 minutes and bilateral offsetting of payments
Settlement scenarios	RTGS with queuing and without payment prioritisation
Intraday credit limits	10 %-point intervals between lower and upper bounds of liquidity
Simulation period	100 days: 4 Jan – 21 May 1999

In payment splitting, a payment is split if it cannot be settled immediately and its value exceeds the splitting minimum. The processed payment is split into the minimum number of payments such that all subpayments are of equal value and do not exceed the splitting limit. The calculation of settlement delay in this case is based on the time span between payment initiation and transferral of full cover of the complete original payment.

4.4.5 Data

The data used in the simulations were collected from the Bank of Finland's settlement account system. They cover all payments transferred in the Bank of Finland's real-time gross settlement system (BoF-RTGS) between 4 January and 21 May 1999. The data cover the payments of 14 account holders and a dummy bank representing participants of the TARGET network. During the analysed period, the daily average number of payments was 1,790 and the average value was EUR 18,265 million. There are two peak times for payments in the system: one in the morning and one before the system closes. Between these peaks, the arrival rate of payments is rather stable, at 20–30 payments per minute. The settlement process starts each day at 8 am Finnish time and all payments entered into the system must be settled by the end of the day, at 7 pm. The characteristics of the data are summarised in table 4.3.

The large variations in number and settlement value of daily payments are due to the extended TARGET opening days. On national holidays, the domestic volumes are negligible and the volume of international transactions is also small. This also partly explains the variation in liquidity bounds. During the simulation period, the liquidity employed in BoF-RTGS averaged EUR 7,224 million and consisted of fully collateralised credit extended by the Bank of Finland worth EUR 5,664 million plus compulsory reserve requirements worth EUR 1,560 million. The actual liquidity available to the participants was EUR 1,382 million more than the average upper bound, which indicates the existence of idle liquidity on some participants' accounts. This resulted in quick settlements. The relatively large amount of liquidity used is due to the characteristics of the national settlement process. The very small number of participants combined with large differences in size and specialisation of system participants resulted in large fluctuations in payment flows. Some net settlement arrangements also 'froze' liquidity for the period for which the net settlement is calculated. The security settlement process in particular demands a large amount of liquidity for several hours during the day.

Table 4.3 **Payment statistics**

	Minimum	Maximum	Average
Value of payments (EUR m)	0.001	6,000.0	12.3
Daily payment flow (EUR m)	4,836.4	49,421.7	19,057.6
Daily payment flow (number)	200	2172	1,603.9
System upper bound of liquidity (EUR m)	2,224.6	38,467.0	5,842.9
System lower bound of liquidity (EUR m)	130.5	3,399.2	846.2
System upper bound as % of payment flow (ie π^d at upper bound)	18.8	77.8	30.2
System lower bound as % of payment flow	0.7	21.4	4.6

To some extent the Finnish figures are overstated, as some participants use the system as a liquidity source for euro outside Finnish markets and Finnish sectors of international markets.

The banks' choice of payments to be settled at each moment depends inter alia on their available liquidity. In the data used, the choice of payments and the timing of settlement were recorded, as was the case for the actual settlement balances and intraday credit limits. If an account holders' liquidity changes, it is likely to settle its payments differently. The choice of settlement process is also likely to affect participants' behaviour. However, for the purpose of this study we considered the payment data to be exogenously determined and not affected by the settlement structure or the liquidity available.

4.5 Simulation results

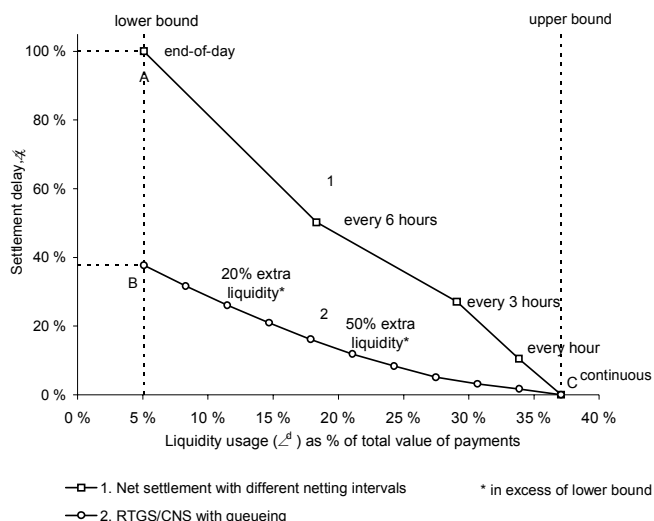
4.5.1 Relationship between liquidity usage and settlement delay

The relationship between liquidity usage and settlement delay is analysed for regimes with rigid and flexible liquidity management systems. Figure 4.1 shows the relationship for a regime with rigid liquidity management for net and real-time gross settlement systems. A bank's liquidity need in a multilateral net settlement system with end-of-day settlement is given by point A in the figure. In the simulations, the system π^d , ie participants' total liquidity need as a share of total value of payments, was on average 5.1%. If the number

of net settlements during the day is increased, payment delay can be reduced by raising the liquidity need. Curve segment AC shows this trade-off. If the number of settlements is increased to the point where net settlement is executed after each transaction (against earlier netted payments), the liquidity usage and settlement delay correspond to those of a real-time gross settlement system without queuing as well as to a CNS system with fully collateralised debt limits. This is shown as point C in figure 4.1, where π^d equals 37.1%.

Figure 4.1

Relationship between liquidity usage and settlement delay, rigid liquidity regime



In an RTGS system with a queuing facility, banks that are willing to accept more delay in their payments can reduce their liquidity usage by moving northwest on the trade-off curve. This applies to CNS systems as well. The system-level trade-off curve, constructed by summing participants' trade-off curves, is shown as segment CB in Figure 4.1. Liquidity must remain at least at the lower bound of liquidity (point B) if all payments during the day are to be settled. A bank can theoretically choose any point on its curve segment BC according to its preferences and estimate of relative costs of liquidity and delay. A bank that weights the cost of liquidity high relative to that of settlement delay will choose a point near B and vice versa. The degree to which a system participant knows the shape of the curve before the end of the day will depend on its information as to incoming and outgoing payments during the day. The range at which

liquidity can be substituted for settlement delay is rather wide in our case, at 32% of the daily value of payments, on average. The curve is convex, with its slope being steeper than -1 up to the 70% liquidity level and between -0.5 and -0.6 thereafter. Up to the 70% liquidity level, an increase in liquidity usage is compensated by a larger relative reduction in settlement delay, after which the trade-off becomes more gradual.

The curve representing different netting intervals in a net settlement system lies above the curve representing an RTGS system with queuing at all points between the upper and lower bounds of liquidity. An RTGS (CNS) system can operate on the same amount of liquidity as any of the netting systems, with reduced payment (settlement) delay. In comparing end-of-day net and real-time gross settlement (both systems operating without counterparty risk), payment delay was found to be 62% less for the RTGS system. Analogously, in comparing CNS systems, risks could be reduced by at least 62%. The relative reduction in payment or settlement delay was even larger in our comparison of an RTGS system operating on the same liquidity as netting systems with more frequent settlement intervals. At the six-hour interval the reduction amounted to 69%, and at the three- and one-hour intervals 84%.

The relationship between liquidity usage and payment delay in a system with flexible liquidity management is very similar to that in the rigid regime. In the RTGS/CNS system, liquidity usage at the upper bound of liquidity was reduced from 37.1% to 27.5%, and at the lower bound from 5.1% to 4.5%. This is explained by the absence of idle liquidity in the flexible regime. The delays in payments are identical in both regimes. For net settlement systems the notion of flexible and rigid liquidity management was not considered relevant.

4.5.2 Combined cost of liquidity and delay

In Figure 4.2, equal interest charges are applied to liquidity usage and to settlement (or payment) delay, and the total cost is plotted as a function of available liquidity. The reason for equal charges is that we would argue that the value of liquidity should be the same for the receiving and sending banks. An increase in the relative price of liquidity bends the curves up on the right-hand side, and an increase in the relative cost of delay bends them up on the left-hand side. If the relative price of liquidity (delayed payments) is high enough, the system-level cost-optimal liquidity level is the lower (upper) bound of liquidity. An increase in either liquidity or delay costs will shift the

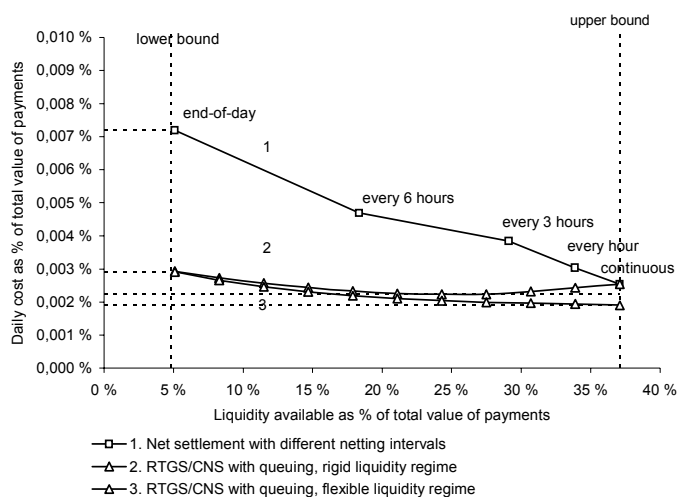
curve up, and a decrease will shift it down. An equal change in the cost elements does not affect the optimal allocation of liquidity.

The combined cost of liquidity and delay in multilateral net settlement systems is shown as curve 1 in the figure. At equal costs for liquidity and delay, the combined cost is at its maximum with end-of-day net settlement, in which case all payments are delayed until the end of the day and liquidity usage equals the lower bound of liquidity. Reducing the interval of net settlements reduces the combined cost, since the cost increase on liquidity usage is less than the cost reduction on settlement delay. The combined costs are thus minimised when net settlements are executed as often as possible, ie after each transaction, against previously settled payments.

The combined cost for an RTGS system with queuing in a rigid liquidity regime is shown as curve 2 in Figure 4.2. At the lower bound of liquidity, delay costs are maximised and liquidity costs minimised. At the upper bound, costs stem solely from liquidity usage. System costs are minimised when the participants post liquidity worth 27% of the total value of payments and liquidity is optimally allocated among participants. Below this level, dead-weight losses raise the combined costs; above it, the burden of increased idle liquidity reduces the costs.

Figure 4.2

Combined daily cost of liquidity and settlement delay with equal interest charges, share of daily value of payments



The relative cost reduction in our case (with equal interest charges) as between real-time gross settlement and multilateral net settlement

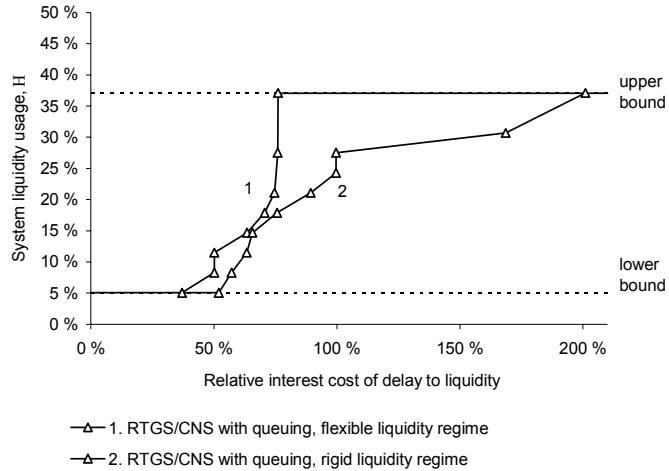
ranges from 59% for end-of-day net settlement to 20% for netting at one-hour intervals. The only cost level at which the costs for net and gross settlement are equal is where payment delay does not entail any cost, ie where there is no need for intraday delivery. In this case, end-of-day net settlement minimises the costs for netting systems and the use of lower bound liquidity for real-time gross settlement systems.

The cost curves for RTGS systems operating under flexible and rigid liquidity regimes are shown as curves 2 and 3 in the figure. The relative cost advantage of the flexible liquidity regime compared to a rigid regime ranges from only 1% at the lower bound of liquidity to 25% at the upper bound. At the lower bound, the participants have very little idle liquidity on their settlement accounts; the amount of idle liquidity increases as more liquidity is made available. Introducing flexibility when banks are operating at minimum liquidity has only a minor impact on costs. It is presumably also the case that when banks operate with ample liquidity the reason is that they must have relatively low liquidity costs. Flexibility is a feature that can help banks overcome wider variations in liquidity demands at certain times of the day.

In systems where either payments or settlements can be delayed, participants have the possibility to adapt to changes in relative costs of delay and liquidity. As the relative costs change, the cost curve shifts and a new optimal allocation of liquidity emerges. In net settlement systems based on end-of-day netting or pure RTGS systems without queuing, both liquidity need and settlement delay are determined solely by payment flows and hence the combined cost of settlement cannot be influenced by participants' decisions. System-level cost-optimal liquidity usage for an RTGS system with queuing as a function of relative delay costs vs liquidity costs is shown in Figure 4.3, for both flexible and rigid liquidity regimes. When the relative cost of settlement delay increases, banks find that reducing delay is more favourable and thus increase their liquidity holdings. In a flexible liquidity regime when the cost of delay is equal to or less than 37% of the cost of liquidity, the optimal allocation of liquidity is the lower bound. The cost of delay is small enough for participants to minimise their costs by simply posting just enough liquidity to enable settlement of all payments up to the end of the day. Between relative costs of 37% and 75%, participants will choose to delay a portion of their payments: the more, the less costly the payment delays relative to liquidity. If the cost of payment delays exceeds 76% of the costs of liquidity, the optimum is to settle all payments immediately, in which case liquidity usage equals the upper bound.

Figure 4.3.

Liquidity usage in an RTGS system with queuing as a function of relative interest costs of delay to liquidity



In a rigid liquidity regime, participants will operate at the lower bound of liquidity if the interest cost of delay is less than 52% of the interest cost of liquidity. If the delay costs increase, banks will use more liquidity, which will reduce the delays. If delay costs are about twice as high as liquidity costs, the system-level cost optimum is for each participant to post enough liquidity to settle its payments immediately. In a rigid liquidity regime, participants will tend to delay payments and thus incur higher delay costs because part of the delay cost is absorbed by the use of idle liquidity. In a flexible liquidity regime, the full costs are experienced by the receiver when payments are delayed, and, with more than 76% of delay costs to liquidity costs, the system-level cost optimum is not to delay payments.

4.5.3 Gridlocks in RTGS systems

Systems that operate with queuing face the risk of gridlock. In systems with centrally managed queues, the gridlocks are visible, whereas in systems with decentralised queue management they are normally hidden. We define gridlock as a situation where each participant of the settlement system is liquid, but not all payments can be settled immediately via the queuing algorithm due to the settlement order of payments. In our model, a participant is considered liquid during the

analysed time period if its net liquidity position (the sum of incoming and outgoing queued payments combined with the available liquidity) is positive. If at least one of the system participants is illiquid, ie its net liquidity position is negative, the cause of any delays in settlement is not considered to be gridlock but rather insufficient funds. Our intention is to analyse how the system can be designed to prevent or solve gridlocks that occur despite the presence of sufficient liquidity.

The amount of gridlocks with the studied payment flows in an RTGS system with queuing is shown in Table 4.4. Finnish banks operate at or above the upper bound of liquidity, so actual gridlocks were rare. As liquidity was reduced in the system, more gridlocks occurred, but even at the lowest levels of liquidity gridlocks were rather rare.

Table 4.4 **Gridlock durations at different levels of liquidity in RTGS with queuing**

	Lower bound (0%)	Low levels (10%–30%)	Middle levels (40%–60%)	High levels (70%–90%)	Upper bound (100%)
Days experiencing gridlocks (%)	28.0	7.3	5.3	5.0	0.0
Maximum time in gridlock (tot. days, min)	35	35	70	73	0.0
Average time in gridlock (of tot. days, min)	4.2	1.3	1.9	1.1	0.0
Average time in gridlock (of days with gridlocks, min)	15.0	17.4	35.1	21.3	0.0

The average total duration of daily gridlocks was only 4.2 minutes at the lower bound of liquidity and 1.1 minutes at high levels of liquidity. On the worst day of the simulation period, with low levels of liquidity, the system was gridlocked for 73 minutes. With low levels of liquidity, the total duration of daily gridlocks was smaller but gridlocks occurred on more days, whereas at higher levels of liquidity gridlocks occurred only on a few days although the duration was longer. Thus the overall benefit of employing gridlock-solving algorithms is quite limited here.

Gridlocks can be solved or prevented by using optimisation methods like netting queued payments or splitting large payments into several smaller ones. The effects of these optimisation methods on gridlocks are summarised in Table 4.5.

Table 4.5

**Changes in daily gridlock duration with
different optimisation methods
in RTGS systems**

	Liquidity available					
	Lower Bound (0%)	Low levels (10%–30%)	Middle levels (40%–60%)	High levels (70%–90%)	Upper bound (100%)	Upper bound (100)
	%	%	%	%	%	%
Netting of queued payments						
Full, every 20 min.	-41	-41	-73	-64	0	-57
Full, every 5 min.	-80	-83	-92	-84	0	-86
Partial, every 20 min.	-83	-94	-96	-88	0	-92
Partial, every 5 min.	-97	-98	-98	-91	0	-96
Offsetting payments	-60	-75	-73	-100	0	-80
Splitting of payments						
Top 1‰ split	0	-3	0	0	0	-1
Top 1% split	-25	-82	-78	-55	0	-67
Top 5% split	-52	-100	-100	-98	0	-95
Top 10% split	-64	-100	-100	-98	0	-96

The netting of queued payments was simulated with full and partial netting and with 5- and 20-minute netting intervals. With full netting of queues every 20 minutes, the reduction in gridlocks amounted on average to 57% and was maximised at medium and high levels of liquidity. If the netting was executed every 5 minutes, the reduction averaged 86% and the level of liquidity had less effect. The use of the partial netting algorithm reduced gridlocks still further. At the 20-minute netting interval, the average reduction was 92%, and at the 5-minute interval 96%. Bilateral offsetting of payments solved all gridlocks at the high levels of liquidity, but was not as effective in reducing gridlocks when less liquidity was available to the participants. At the low and medium levels of liquidity, the reductions in daily duration of gridlocks were 60% and 75% respectively. Both full and partial netting substantially reduced the maximum daily time the system was gridlocked. The resulting total duration of gridlocks was reduced to approximately the time interval between the nettings. Only with bilateral offsetting of payments did the maximum total duration of daily gridlocks remain greater than a half hour at low levels of liquidity.

Splitting large payments into several smaller ones helps prevent the formation of gridlocks. When the size of the payments is smaller, the amount of idle liquidity is less and a greater share of the available liquidity is employed for settlement. The effects of payment splitting

on gridlocks are presented in Table 4.5. Because only queued payments were split, the actual number of payments split varied according to the amount of liquidity available to the participants. At the lower bound of liquidity, the number of payments split was at its maximum, ie about half of the payments eligible for splitting. At the upper bound of liquidity, payments were settled immediately and no splitting was necessary. The splitting of payments at all splitting limits was most effective at the low and medium levels of liquidity, but differences in efficiency were substantial. Splitting only the very large top 1% of payments (16 payments eligible for splitting during the 100-day period) had almost no effect on gridlocks, but splitting the top largest 1% (1,733) already provided good results. Splitting the top 5% (8,948) of payments was almost as good as splitting the top 10% (17,906), even though the number of payments eligible for splitting more than doubled. With splitting of the top 5% and top 10% of payments, all gridlocks were eliminated at low and medium levels of liquidity, and at high of levels of liquidity the total duration of gridlocks was reduced by 98%. This suggests the splitting of payments is technically a very efficient method of alleviating the gridlock problem. Here, it was in many cases more effective than the netting algorithms, especially when over 5% of transactions were split and when more than the lower bound of liquidity was available to the participants.

4.5.4 Optimising an RTGS system

There are basically two ways to optimise settlement in real-time gross settlement systems: to clear gridlocks (and thus reduce delay costs) or to minimise the amount of idle liquidity. The latter applies to rigid liquidity regimes only, because in flexible regimes banks can withdraw the idle liquidity from the system. For optimisation methods such as payment splitting and netting queued payments, the curve representing the trade-off between liquidity usage and settlement delay (Figure 4.1) can be shifted toward the origin and each participant can effect faster settlements with a given amount of liquidity.

The simultaneous settlement of several queued payments via netting clears gridlocks and thereby reduces settlement delay. However, in a real-time environment there are some technical processing problems in applying netting algorithms, especially if these require prolonged calculation times. As payments are entered continuously into the system, the liquidity position of a participant

might change as a result of payments settled while the net position is being calculated, if the system accepts new payments during the netting process. On the other hand, if the real-time process is suspended while netting is in process, there is a risk that payments/settlements will be delayed. This requires rapid netting algorithms, especially in systems with large volumes and many participants.

Table 4.6 summarises the changes in the total cost of liquidity and delay with full or partial net settlement, 5- and 20-minute netting intervals, and bilateral offsetting of payments. The results are calculated for both rigid and flexible liquidity regimes. The cost reduction was found to be similar for the two regimes, albeit the flexible regime resulted in a slight cost advantage. At the upper bound of liquidity, no queuing took place and the optimisation methods had no effect. The effectiveness of full netting algorithms was very modest at all levels of liquidity: A reduction of only 0.1% compared to non-netting was achieved at the 20-minute interval, and a reduction of 0.2% at the 5-minute interval. By using the partial netting algorithm for netting queued payments, costs were reduced substantially at low levels of liquidity. The reduction was in the range of 4% to 4.6% for the non-netting scenario at the 20-minute interval, and 4.7% to 5.3% at the 5-minute interval, depending on the liquidity regime. The bilateral netting of offsetting payments reduced the costs most at the lower bound of liquidity, but was not as efficient as the partial netting algorithm at low and medium liquidity levels. At high levels of liquidity and at the upper bound of liquidity, the costs actually rose marginally. This can be ascribed to a change in the settlement order of payments that increased the liquidity needs.

Table 4.6

**Relative change in total costs of liquidity
and delay with netting of queued payments**

	Liquidity available					
	Lower bound (0%)	Low levels (10%–30%)	Medium levels (40%–60%)	High levels (70%–90%)	Upper bound (100%)	Average
	%	%	%	%	%	%
Rigid liquidity regime						
Full, 20 min.	-0.1	-0.1	-0.1	0.0	0.0	-0.1
Full, 5 min.	-0.2	-0.2	-0.2	-0.1	0.0	-0.1
Partial, 20 min.	-3.3	-4.0	-1.3	-0.2	0.0	-1.8
Partial, 5 min.	-4.7	-4.7	-1.6	-0.2	0.0	-2.2
Offsetting payments	-5.3	-3.3	-0.4	0.3	0.1	-1.4
Flexible liquidity regime						
Full, 20 min.	-0.1	-0.1	-0.2	-0.1	0.0	-0.1
Full, 5 min.	-0.2	-0.2	-0.2	-0.1	0.0	-0.2
Partial, 20 min.	-3.3	-4.6	-1.8	-0.3	0.0	-2.1
Partial, 5 min.	-4.8	-5.3	-2.0	-0.4	0.0	-2.5
Offsetting payments	-5.5	-3.8	-0.6	0.3	0.1	-1.6

The splitting of payments prevents the formation of gridlocks and reduces the upper bound of liquidity, as the maximum liquidity constraint imposed by a payment is reduced from its original value to the value of the splitting limit. Settlement delay is reduced as gridlocks are prevented, but also because a part of the cover is available sooner to counterparties for settling their payments. Because, in the flexible regime, interest is calculated on the actual amount of liquidity employed for settlement, the liquidity costs increase due to the fact that some of the liquidity is used sooner. The changes in the total cost of liquidity and delay are summarised in Table 4.7.

Table 4.7

**Relative change in total costs of liquidity
and delay with payment splitting, compared
to a scenario with no optimisation**

	Liquidity available					
	Lower bound (0%)	Low levels (10%–30%)	Medium levels (40%–60%)	High levels (70%–90%)	Upper bound (100%)	Average
	%	%	%	%	%	%
Rigid liquidity regime						
Top 0.1% split	-0.5	-0.7	-0.2	0.0	0.0	-0.3
Top 1% split	-4.6	-4.5	-1.6	-0.3	0.0	-2.1
Top 5% split	-9.3	-7.2	-2.7	-0.6	0.0	-3.7
Top 10% split	-10.7	-8.1	-3.1	-0.7	0.0	-4.2
Flexible liquidity regime						
Top 1‰ split	-0.5	-0.8	-0.1	0.0	0.0	-0.3
Top 1% split	-3.7	-0.5	3.2	1.6	0.0	0.8
Top 5% split	-6.6	1.2	5.6	2.9	0.0	2.0
Top 10% split	-7.2	1.8	6.2	3.3	0.0	2.4

In the rigid regime, costs are reduced more when a larger share of payments is split and when less liquidity is available to the participants. At the upper bound of liquidity all payments were settled immediately and no splitting was necessary. At high levels of liquidity the effects of all splitting limits were marginal, under 1%. Splitting the top 1‰ of payments also had only minor effects on the costs. In contrast, splitting the top 1% of payments already provided noticeable results and a reduction in total cost of up to 4.6% at the lower bound of liquidity, compared to a scenario without splitting.

From Table 4.7 we can also see that in the flexible liquidity regime the total costs decline with payment splitting at low levels of liquidity but increase when more liquidity is available to the participants. Earlier use of liquidity increases liquidity costs and, because the funds are available sooner to the receivers, reduces delay time and delay costs. At higher levels of liquidity the increase in liquidity costs is not fully compensated by the reduction in settlement delay costs and hence total costs are higher. With a flexible liquidity regime, a cost reduction of 7.2% could be achieved by splitting the top 10% of payments, but a cost increase of 6.2% occurred at medium levels of liquidity. When less payments were split, both the reduction in costs at lower levels of liquidity and the increase in costs at medium and higher levels of liquidity were more modest.

Payment splitting entails some problems. Splitting can be achieved through an agreed upper limit for payment size (Humphrey 1996). It

can also be implemented via an implicit and automated function within the settlement system and can be designed to be more or less transparent to the participants. For a system with payment splitting, the legal rules must take into account special situations in which payment processing is halted when only some parts of individual transactions have been processed.

4.6 Conclusions

The pattern of payment flows affects liquidity usage and the efficiency of the optimisation methods. Variations over time in payment flows can also lead to different results for different time periods. The characteristics of payment flows depend very much on market conventions and customer behaviour.

In addition to the simulations reported here, we have run simulations using Finnish payment data from other time periods and with artificially constructed payment flows. All of our key findings were supported by the results from these simulations. The absolute amounts of delay and liquidity usage varied, but in terms of efficiency the rankings of different systems remained the same. Simulation is the best tool for determining the exact properties of special types of settlement system.

A further limitation is the exogenous character of payment flows. If the features of the settlement system were changed as drastically as in our models, it is likely that this would affect customer and system participant behaviour, which in turn would affect payment flows. We have not attempted to include these kinds of consideration in the present study. Nonetheless, we feel that our approach is able to offer suggestions as to what kinds of enhancement to a given settlement system are worthy of deeper analysis.

The costs involved in payment systems are difficult to measure, especially delay costs, costs associated with settlement risk and, to an extent, liquidity costs. We have used simple linear cost functions, but in fact cost functions are probably highly nonlinear in some cases. However, even the rough cost estimates obtained here can be helpful in assessing the relative importance of different parameters. It is not a problem to include more realistic data in the model when these are available.

Our main findings are as follows:

- A real-time system with queuing (ie RTGS/CNS or a hybrid system) is always more efficient than a net settlement system with batch processing. For a given level of liquidity, the payment or settlement delay is always reduced by real-time processing, as compared to a time-designated net settlement system.
- If liquidity costs are low enough relative to delay costs in an RTGS system, participants operate under immediate payment/settlement. Within a narrow range of relative costs of liquidity and delay, participants switch from maximum delay to immediate settlement. The central bank can support risk reduction and rapid payment targets by providing low cost intraday liquidity as well as more flexible ways for participants to add or withdraw liquidity from the system. Because the Finnish banking system is operating at or above the upper bound of liquidity, the current liquidity scheme adopted by the European System of Central Banks is providing efficient levels of liquidity in terms of risk reduction.
- Optimising and gridlock solving features are needed and effective only when liquidity is scarce. In the Finnish case, a splitting of the top 5–10% of payments or partial netting of queues would be very efficient. The impact of different netting schemes is somewhat more limited. The available netting scheme is very rarely applied because of the abundance of liquidity.
- The costs of the settlement process can be reduced by applying optimisation methods in situations with limited amounts of available liquidity. At very low levels of liquidity, these methods can be effective. The overall benefit depends of course also on the extra development and processing costs. The efficiency of the different optimisation methods for settlement systems are affected by the actual flow of payments processed.

Our impression is that a simulation model similar to the one developed at the Bank of Finland is a useful tool for evaluating the properties of different settlement systems and processes as well as the impacts of changes in important system characteristics such as payment flows, settlement processes, cost and risk parameters and optimisation methods.

References

- Angelini, P, Maresca, G and Russo, D (1996) **Systemic Risk in the Netting System**. *Journal of Banking and Finance* 20, 853–868.
- Angelini, P (1998) **An Analysis of Competitive Externalities in Gross Settlement Systems**. *Journal of Banking & Finance* 22, 1–18.
- BIS (1989) **Report on Netting Schemes (Angell report)**. Prepared by the Group of Experts on Payment Systems of the central banks of the Group of Ten countries, Bank for International Settlements, Basle.
- BIS (1990) **Large-value Funds Transfer Systems in the Group of Ten Countries**. Prepared by the Group of Experts on Payment Systems of the central banks of the Group of Ten countries, Bank for International Settlements, Basle.
- BIS (1997) **Real-time Gross Settlement Systems**. Prepared by the Committee on Payment and Settlement Systems of the central banks of the Group of Ten Countries, Bank for International Settlements, Basle.
- Boeschoten, W C (1989) **In Search of an Efficient Interbank Settlement System**. Paper prepared for Eurobanking 1989. Economic Research and Special Studies Department, De Nederlandsche Bank.
- Borio, C, Russo, D and Van den Bergh, P (1992) **Payment System Arrangements and Related Policy Issues: A Cross Country Comparison**. SUERF Papers on Monetary Policy and Financial Systems 13, Société Universitaire Européenne de Recherches Financières, Tilburg.
- Borio, C and Van den Bergh, P (1993) **The Nature and Management of Payment System Risks: An international perspective**. *Economic Papers* 36, Bank for International Settlements, Basle.
- Bowman, M (1995) **Simulation Modelling of Large Value Transfer Systems**. Paper prepared for Eurobanking 1995, Association for Payment Clearing Services.

- Dale, S and Rossi, M (1996) **A Market for Intra-day Funds: Does it have Implications for Monetary Policy?** Bank of England.
- Folkerts-Landau, D, Garber, P and Schoenmaker, D (1996) **The Reform of Wholesale Payment Systems and Its Impact on Financial Markets.** Group of Thirty Occasional Paper 51, Washington DC.
- Freixas, X and Parigi, B (1998) **Contagion and Efficiency in Gross and Net Interbank Payment Systems.** Journal of Financial Intermediation 7, 3–31.
- Furfine, C and Stehm, J (1997) **Analysing Alternative Intraday Credit Policies in Real-Time Gross Settlement Systems.** Finance and Economics Discussion Series 1997-40, Board of Governors of the Federal Reserve System (US).
- Ganz, M, Günstzer, M M, Jungnickel, D and Noltsch, G. (1998) **Effizientes Real-Time Gross Settlement in der Wertpapierabwicklung.** Die Bank 2, 94–97.
- Güntzer, M M, Jungnickel, D and Leclerc, M (1998) **Efficient Algorithms for the Clearing of Interbank Payments.** European Journal of Operational Research 106, 212–219.
- Humphrey, D (1986) **Payments Finality and the Risk of Settlement Failure.** In: Saunders, A and White, L C (eds.) Technology and the Regulation of Financial Markets. Lexington Books, D C Heath and Company, 97–120.
- Humphrey, D (1990) **Pricing Intraday Overdrafts.** In: Haraf, E S and Cagan, P (eds.) Monetary Policy for a Changing Financial Environment. AIE Press, Lanham, 108–131.
- Humphrey, D (1996) **Comment on Intraday Bank Reserve Management: The Effects of Caps and Fees on Daylight Overdrafts.** Journal of Money, Credit, and Banking 28, 909–913.
- Kahn, C and Roberds, W (1998a) **Payment System Settlement and Bank Incentives.** Review of Financial Studies, Vol. 1, No. 4, 845–870.

- Kahn, C and Roberds, W (1998b) **Real-Time Gross Settlement and the Cost of Immediacy**. Working Paper 98-21, Federal Reserve Bank of Atlanta.
- Kobayakawa, S (1997) **The Comparative Analysis of Settlement Systems**. Discussion Paper 1667, Centre for Economic Policy Research.
- Koponen, R and Soramäki, K (1998) **Intraday Liquidity Needs in a Modern Interbank Payment System – A Simulation Approach**. Studies in Economics and Finance E:14, Bank of Finland.
- Kuussaari, H (1996) **Systemic Risk in the Finnish Payment System: An Empirical Investigation**. Discussion Paper 3/96, Bank of Finland.
- Leinonen, H (1998) **Interbank Funds Transfer Systems: Liquidity Needs, Counterparty Risks and Collateral**. Discussion Paper 16/98, Bank of Finland.
- McAndrews, J J and Wasilyew, G (1995) **Simulations of Failure in a Payment System**. Economic Research Division Working Paper 95-19, Federal Reserve Bank of Philadelphia.
- Rossi, M (1995) **Pricing Intraday Credit in Real-Time Gross Settlement Systems**. Discussion Paper 211, London School of Economics Financial Markets Group.
- Schoenmaker, D (1993) **Externalities in Payment Systems: Issues for Europe**. Financial Markets Group, London School of Economics.
- Schoenmaker, D. (1995) **A Comparison of Alternative Interbank Settlement Systems**. Discussion Paper 204, Financial Markets Group, London School of Economics.
- Vital, C and Mengle, D L (1988) **SIC: Switzerland's New Electronic Interbank Payment System**. Economic Review, Federal Reserve Bank of Richmond.
- Vital, C (1990) **Swiss Interbank Clearing: Further Experience with a Queuing Mechanism to Control Payment System Risk**. Working paper, Bank for International Settlements.

Vital, C (1994) **A Central Bank Appraisal of the Swiss Interbank Clearing system.** Payment Systems Worldwide Vol. 5, No. 1, 4–9.

Chapter 5

Gridlock resolution and bank failures in interbank payment systems

Morten L. Bech – Kimmo Soramäki

5	Gridlock resolution and bank failures in interbank payment systems.....	152
	Abstract.....	152
5.1	Introduction.....	152
5.2	Gridlocks in RTGS systems.....	154
5.2.1	The cost of gridlocks	154
5.2.2	Problem formulation.....	156
5.3	Gridlock resolution	158
5.3.1	The problem.....	158
5.3.2	Solution algorithm	160
5.4	Simulations.....	161
5.4.1	Simulation setting.....	161
5.4.2	Payment data.....	162
5.5	Results.....	163
5.5.1	The lower and upper bound.....	164
5.5.2	The trade-off between liquidity and delay	164
5.5.3	Gridlocks experienced in the system.....	167
5.5.4	Resolution under failure scenarios	171
5.6	Conclusions.....	173
	References	176

5 Gridlock resolution and bank failures in interbank payment systems

Abstract

This chapter analyses the severity of gridlocks in interbank payment systems operating on a real-time basis and evaluates by means of simulations the merits of a gridlock resolution algorithm. Performance of the algorithm is studied under both normal operating conditions and failure scenarios. The data used in the simulations consist of actual payments settled in the Danish and Finnish RTGS systems. The algorithm is found to be applicable to a real-time environment and effective in reducing queuing in the systems at all levels of liquidity, but in particular when intraday liquidity is scarce. It can also alleviate settlement delays caused by the failure of a bank to participate in settlement.

5.1 Introduction

The settlement of interbank liabilities has traditionally taken place in interbank payment systems operated by central banks. Many central banks in the industrialised countries have during the last decade built real-time gross settlement (RTGS)¹ systems for the settlement of interbank funds transfers. In an RTGS system payments are processed individually with finality throughout the day, and, due to the low settlement risk, these systems are primarily used for time critical and/or large value payments.²

The volume of interbank payments increased dramatically throughout the 1980s and 1990s as a result of rapid financial innovation and globalisation of financial markets. At the same time the developments in information technology made gross settlement in

¹ For descriptions of RTGS systems please refer to BIS (1997).

² For ease of exposition the terms bank and central bank are used as shorthand for a participant in and the operator of the interbank payment system, respectively. In many countries, participants include non-banks and the RTGS system is designed and operated in close cooperation between the central bank and the banking community.

real time a technologically feasible alternative. Historically, interbank payments have been settled via end of day netting systems, but as the volumes of funds transfers increased, central banks became concerned about the systemic risk³ stemming from netting systems. In netting systems where payments are credited to customer accounts before final interbank settlement, a failure of one participant may cause the failure of other participants if proper risk controls are not in place. Gross settlement in real time eliminated settlement risk, reduced systemic risk and became the prevalent option chosen by most central banks.

However, the risks were traded off against an increased need for intraday liquidity (ie funds available for settlement). Central banks have found that the provision of free uncollateralised intraday liquidity is not a viable solution to the liquidity need, since it merely results in a transfer of credit risk to the central bank. As a result, intraday liquidity is provided by central banks through an overdraft facility either subject to interest or backed up by collateral. Liquidity is thus costly either in the form of an explicit fee or implicitly in the form of the opportunity cost of the pledged collateral.

BIS (1993) defines a gridlock as a 'situation that can arise in a funds or securities transfer system in which the failure of some transfer instructions to be executed (because the necessary funds or securities balances are unavailable) prevents a substantial number of other instructions from other participants from being executed'. There are, in fact, three possible causes of the failure of transfer instructions to be executed, ie gridlocks. Gridlocks are accordingly classified into three types according to their cause.

Type 1: Gridlocks caused by the processing rules for settlement. Banks submit payments to the system and enough liquidity for their settlement would be available. The system is, however, gridlocked, eg due to the requirement that payments be settled one by one and/or in a predefined order.

Type 2: Gridlocks caused by a lack of funds. Banks submit payments to the system, but these payments cannot be settled by any means due to a constraint on the liquidity available for settlement. In this paper these types of gridlock are also referred to as *deadlocks*.

Type 3: Gridlocks caused by banks withholding payments. Banks do not submit payments for processing, eg because they are

³ Systemic risk is the risk that a failure by one participant in a transfer system to meet its obligations will cause other participants to be unable to meet their obligations, possibly threatening the stability of the financial system as a whole (BIS 1993).

economising on liquidity usage, or because of concerns over counterparty performance.

Type 1 gridlocks can be solved by a gridlock resolution algorithm or by additional liquidity. Type 2 gridlocks can only be solved by additional liquidity. Type 3 gridlocks can be solved by giving the participants the right incentives to submit payments early to the system. Even if liquidity is available, the gridlock may persist.

In this chapter we analyse the severity of gridlocks of type 1 and 2 in an interbank payments system operating on a real-time basis, describe ways to avoid and resolve type 1 gridlocks and analyse by means of simulations the merits of a specific gridlock resolution algorithm applicable to an RTGS environment.

The chapter is organised as follows. In Section 5.2 we model the problem and provide operational definitions for gridlocks. In Section 5.3 we present an algorithm applicable to a real-time environment and discuss its properties. In Section 5.4 we define the simulation setting and the indicators that are used to analyse the effects of the algorithm, and describe the payment data used in the simulations. In Section 5.5 we present our results. Section 5.6 concludes the chapter.

5.2 Gridlocks in RTGS systems

5.2.1 The cost of gridlocks

Banks manage their liquidity throughout the day in order to minimise the cost of settling obligations on behalf of their customers or obligations resulting from their own treasury operations. Depending on the time-criticality of the payments, banks will at least occasionally have an incentive to hold less liquidity on their settlement accounts than is needed for immediate settlement of all obligations. In many cases banks also delay payments in their internal systems, instead of forwarding them into the central queuing facility. Because incoming payments are the source of liquidity with the lowest cost, banks have an incentive to settle their outgoing payments only after they have received liquidity from incoming payments. While it may be optimal for individual banks to hold less liquidity than is needed for immediate settlement of all payments or to withhold payments in their internal systems, this is not necessarily optimal at the system level.

The delay in payment settlement caused by gridlocks has a cost for both the sender and the receiver. For the sender these costs may be implicit, in the form of deterioration in customer service, or explicit,

in the form of sanctions. For the receiver, the cost is either the cost of not receiving the liquidity (which might force it to acquire more costly liquidity in order to settle its pending payments) or the cost of having to delay its own payments. If the receiver has to delay its payments it faces the same costs as the sender of the first payment. Likewise the receiver of this second payment faces the same types of costs as the receiver of the first payment. This way the costs are cumulated forward in the system until a bank acquires enough liquidity to settle its pending payment. This negative externality creates a dead-weight loss at system level (Angelini 1998).

Dead-weight losses occur when payments are delayed, either in a central queue or in the banks' internal systems for the purpose of saving liquidity. When payments are delayed, the system faces the risk of gridlock, which can add substantially to the dead-weight losses experienced by the system.

Currently the cost of liquidity, at least in Europe, is relatively low, and as a consequence payments are settled smoothly. Intraday liquidity is provided by the central banks against full collateralisation and no fees or interest is charged on the amount used. The result is that queues are not a major issue on the daily level. However, should money market disturbances increase the opportunity cost of collateral or disrupt the availability of collateral in general, features that optimise on the liquidity used for settlement of payments might be appreciated. For systems where the lack of liquidity is an ongoing concern, gridlock resolution naturally provides the greatest benefits.

Central bank policy goals in payment systems mainly address the smooth functioning of the system, including efficiency, and the control of risks, especially systemic risk. Efficient resolution of gridlocks enhances the smooth functioning of the system and reduces the liquidity risk and cost by effecting faster settlement of payments. Analogously, it will reduce the costs of settlement at a given level of delays by enabling the banks to hold lower balances on the settlement account and/or incur smaller overdrafts. In order to discourage banks from using incoming payments as their only or main source of liquidity (and thus to decrease the risk of gridlocks), agreements among participants to process a certain share of payments prior to a specified time of day are in place in many countries. Also, differential pricing of payment processing can create an incentive structure that contributes to a smoother settlement of payments. Incentives in payment settlement have been covered inter alia by Angelini (1998), Kahn and Roberds (1998), McAndrews and Rajan (2000) and Bech and Garratt (2003). In this chapter we concentrate on resolving gridlocks in an environment where payments are sent to the central

queue, ie where information on all queued transfers and their preferable order of settlement is available centrally.

5.2.2 Problem formulation

We study a system with a centrally located queue. Banks transfer funds to each other continuously throughout the day and the settlement of these funds takes place when the sending bank's account is debited and the receiving bank's account credited. As long as the sending bank's account balance (including any possible overdraft limit) is equal to or exceeds the value of a payment, settlement takes place immediately. If the balance of the account is not sufficient to cover the payment, the payment is put in the centrally located queue. Queued payments are released according to a scheme predefined by the bank itself. The first pending payment in the queue is released as soon as the bank has accumulated enough liquidity to cover the payment.⁴

Even though it might not be possible to individually settle the first pending payment of any bank due to a lack of funds, it might be possible to simultaneously settle a subset or all of the payments queued. Currently, a range of different measures have been put in place to ensure the smooth settlement of payments, mainly based on the netting of queued transfers, which is either invoked by the central bank when needed, or carried out at predefined intervals. A necessary requirement for all gridlock resolution features is that the central bank has information concerning all incoming and outgoing payments. In practice, this will require a central queuing mechanism in the RTGS system.

In order to formally define a gridlock and the gridlock resolution problem we must introduce some notation. Assume that we have n banks indexed by i . Let Q_i be the set of queued payments of bank i and let $Q = \cup_{i=1}^n Q_i$ be the set of all queued payments. Similarly, let X_i denote the subset of queued payments of bank i to be settled simultaneously and $X = \cup_{i=1}^n X_i$. The ex ante balance and the ex post

⁴ The source of liquidity is not explicitly modelled here. In general, there are four sources of funds: balances maintained with the central bank, credit extensions based on pledge or repo transaction with the central bank, operations with other banks through the money market, and incoming payments from other banks not related to money market operations. In most RTGS systems the incoming payments are a major source of liquidity; eg in TARGET as a whole these account for some 70% of the gross liquidity need.

balance of bank i is given by \bar{B}_i and $B_i(\cdot)$ respectively. The total amount of outgoing payments from the queue of bank i is $S(X_i)$ and the total amount of incoming payments to bank i from the queues of all other banks is $R(X_{-i})$, where $-i$ denotes all banks except bank i .

We define a gridlock in the present context as follows:

Definition 1 (*Type 1 gridlock*)

A type 1 gridlock is a situation where $Q \neq \emptyset$ and there exists a nonempty $X \subseteq Q$ such that if the payments in X were settled simultaneously then

$$B_i(\bar{B}_i, X) = \bar{B}_i - S(X_i) + R(X_{-i}) \geq 0, \quad \text{for } i = 1, \dots, n \quad (6.1)$$

This *liquidity constraint* stipulates that if the payments in X were simultaneously settled then the ex post balance $B_i(\cdot)$ of each bank has to be non-negative. The ex post balance is equal to the ex ante balance, \bar{B}_i (including any intraday credit line from the central bank), minus the total amount of payments sent by the bank, $S(X_i)$, plus the amount of payments received, $R(X_{-i})$.

Further, we define type 2 gridlocks, ie deadlocks as follows:

Definition 2 (*Type 2 gridlock, ie deadlock*)

A deadlock is a situation where $Q \neq \emptyset$ and X (as defined in definition 1) is empty, ie $X = \emptyset$.

A deadlock is a stalemate of payments between banks where the payments cannot be settled by any means without infringing upon the sequence constraint. A deadlock is only resolved by the addition of adequate liquidity in the system or by the addition of payments in the queues, so that the inclusion of these payments turns the situation into a type 1 gridlock.

Finally, we define type 3 gridlocks as follows:

Definition 3 (*Type 3 gridlock*)

A type 3 gridlock is a situation where $Q \neq \emptyset$ and X (as defined in definition 1) is empty, ie $X = \emptyset$ and banks withhold payments in their internal queues.

In a type 3 gridlock the payment queue is empty, as banks withhold their payments in their internal queues. On the basis of information available from the system alone it is not possible to determine whether banks would have payments to submit to the system, but do not do so.

5.3 Gridlock resolution

5.3.1 The problem

The objective of gridlock resolution is to identify and simultaneously settle the largest possible subset of the payments in queue subject to the liquidity available. A mechanism to resolve gridlocks should have the following desirable attributes if it is to be implemented in practice.

- Optimality: The mechanism should be able to find preferably the optimal solution or at minimum one that is close to it.
- Fairness: The gridlock resolution mechanism should be fair in the sense that the solution does not favour one or more banks relative to others.
- Computational Efficiency: In a real-time environment where the settlement of payments cannot be suspended for an extended period of time; the time needed to find a solution should be very low.
- Minimal Risk: The gridlock resolution mechanism should not expose the system operator or the participants to any significant risks.

The problem of gridlock resolution is mathematically a discrete optimisation problem. A solution specifies which payments to be settled and maximises typically either the value or volume thereof. Guntzer et al (1998) find that the problem can potentially be very difficult to solve, particularly if the number of payments in queue is large. In fact, the computation time grows exponentially in relation to the number of payments. Guntzer et al propose an algorithm to approximately solve the problem.⁵ However, this algorithm is not guaranteed to find the optimal solution; consequently it is not necessarily fair, as one or more participants may ex post prefer another solution. Moreover, intraday liquidity management in a real-time environment is a highly complicated task and many banks are thus reluctant to let a third party control the sequence in which their payments are settled. Furthermore, central banks are unlikely to be willing to accept such a role due to the various legal issues that might arise if a time critical payment was not settled in due time because of the gridlock resolution mechanism.

⁵ Shafransky and Doudkin (2005) also propose an algorithm for the bank clearing problem.

Here, we study a case where the participants specify a strict order in which they want their queued payments to be settled: eg first in – first out. This greatly simplifies finding a solution to the problem and ensures that the solution is optimal, fair and efficient. However, the additional constraint implies that fewer payments tend to get settled compared to what would be possible without the constraint.

With the imposed ordering, the solution can conveniently be stated in terms of the number of payments, l_i , to be picked from the queue of each bank and a unique solution is easily found using the algorithm presented below. In addition, the solution is independent of whether the objective is value or volume and the attributes above are generally fulfilled.

Let m_i be the number of payments in the queue of bank i . Moreover, let the payments in the queue of a bank be ordered from 1 to m_i according to time criticality, with 1 being the most urgent payment. Furthermore, let each payment in the queue be characterised by two elements: i) the amount, $a_{i,k} \in A \subset \mathbb{R}_+$ and ii) the receiver of the payment, $r_{i,k} \in \{1, 2, \dots, n\} \setminus \{i\}$.

Using the same functional symbols as above we can write the problem as follows:

$$\begin{aligned} & \max_{0 \leq l_i \leq m_i} \sum_{i=1}^n l_i \\ & \text{s.t.} \\ & B_i(\bar{B}_i, l) = \bar{B}_i - S(l_i) + R(l_{-i}) \geq 0, \quad \text{for } i = 1, \dots, n \end{aligned} \tag{6.2}$$

$$\text{where } S(l_i) = \sum_{k=1}^{l_i} a_{i,k}, \quad R(l_{-i}) = \sum_{j=1}^n \sum_{k=1}^{l_j} a_{j,k} I(r_{j,k} = i), \quad l = (l_1, l_2, \dots, l_n)$$

and $m = (m_1, m_2, \dots, m_n)$.

For further discussion of the gridlock resolution problem see Bech and Soramäki (2001).

5.3.2 Solution algorithm

A simple algorithm can solve the gridlock resolution problem given by equation (5.2).⁶ The algorithm starts by including all payments in the solution and removes payments one by one from banks with deficient balances until either all banks have a positive ex post balance or all payments have been removed from the initial solution. Formally, the algorithm goes as follows:

Step 1: Include all queued payments in the solution, ie

$$l_i^* = m_i \forall i = 1, \dots, n.$$

Step 2: Calculate $B_i(\bar{B}_i, l^*) \forall i = 1, \dots, n$

If $B_j(\bar{B}_j, l^*) < 0$ for some j then execute step 3.

If $B_i(\bar{B}_i, l^*) > 0$ for all $i = 1, \dots, n$ then stop.

Step 3: Choose any j such that $B_j(\bar{B}_j, l^*) < 0$ and remove the last payment in queue for this bank from the solution. That is

$$l_j^* = l_j^* - 1. \text{ Repeat step 2.}^7$$

The algorithm always finds the optimal solution (which might be empty). The algorithm is fast, as the number of iterations is at most equal to the total number of payments in queue. Moreover, it is fair in the sense that the choice of bank from which to remove a payment in step 3 does not matter for the ultimate solution. Hence, no bank is favoured by the algorithm. In addition, the fixed ordering of the queues minimises several risks by avoiding cases where time critical payments are not settled but less time critical payments are. However, some other solution not respecting the fixed ordering may potentially be able to settle a larger subset of payments in queue given the liquidity available. The properties of the algorithm and the solution found are discussed in more detail in Bech and Soramäki (2001).

⁶ The algorithm was developed by Danmarks Nationalbank in cooperation with the Department of Mathematical Modelling at the Technical University of Denmark as part of the KRONOS project. KRONOS is the new Danish RTGS system.

⁷ A perhaps somewhat counterintuitive fact is that the choice of which of the deficient banks to remove a payment from does not influence the final solution (see appendix in Bech and Soramäki (2001)).

5.4 Simulations

5.4.1 Simulation setting

The simulations consist of two scenarios: one where no gridlock resolution is used, and one where the algorithm presented in the previous section is applied.⁸ We calculate two indicators for both scenarios: the amount of liquidity used for settlement, and an indicator that measures the delays in settlement on the system level.

The amount of liquidity available in the system affects the number and duration of gridlocks. If enough liquidity for each participant to settle their payments immediately is available, naturally no gridlocks occur. We will refer to this amount of liquidity for each bank i as the upper bound UB_i . At the other extreme all banks might have just enough liquidity to settle all their payments at the end of the day via multilateral netting. We shall refer to this amount of liquidity as the lower bound of liquidity LB_i . If the value of payments received during the day is larger than the value of payments sent, then a bank only needs to use the liquidity it receives in the form of incoming payments for settling its own payments. In such a case the lower bound equals zero. On the other hand, if the value of payments sent exceeds the value of payments received for a bank, then the lower bound is equal to the difference.

All simulations with and without application of the gridlock resolution algorithm were run with six different levels of liquidity,

$$L_i = LB_i + \alpha(UB_i - LB_i) \quad (5.3)$$

where $\alpha = \{0, 0.2, 0.4, 0.6, 0.8, 1\}$. It was assumed that banks could not post or withdraw liquidity from the system during the day.

The settlement delay for each payment was calculated as the time difference between payment initiation and final settlement of the payment. For each bank a value-weighted time average of the ratio of payments in queue to payments sent was calculated, ρ_i (see Leinonen and Soramäki (1999) for details). If all payments are settled immediately upon entry into the system, then $\rho_i = 0$, and if all payments are delayed until the end of the day, then $\rho_i = 1$. A system-level delay indicator, ρ , was calculated as a weighted average of the

⁸ The simulations were carried out with version 1 of the Bank of Finland Payment and Securities Settlement Simulator (BoF-PSS1).

individual bank level ρ_i s, with the weights being the value of payments settled by each bank during the day.

5.4.2 Payment data

The simulations were run with data from both the Finnish BoF-RTGS system and the Danish Inquiry and Transfer System. BoF-RTGS is the Finnish national RTGS system of TARGET and operates in euro. The Danish Inquiry and Transfer System used to be the central bank RTGS system for the Danish krona and was not linked to TARGET. The system was replaced by the KRONOS system in November 2001. The Danish data consists of 64 days of transactions processed in the system during the last three months of 1999. The data extrapolated from the Finnish BoF-RTGS system consists of the last 100 days of year 2000. Key figures concerning both systems are summarised in Table 5.2 below. The systems setup in the simulations was, with the exception of operating hours, identical.

Table 5.1 **Key figures concerning payment flows in both systems (EUR million)**

	DN Inquiry and Transfer System (DK)			BoF-RTGS (FI)		
	Minimum	Maximum	Average	Minimum	Maximum	Average
Individual payment value	0.001	1,227	10	0.001	2,098	10
Daily payment flow (value)	1,358	13,783	9,352	4,638	32,718	15,045
Daily payment flow (number)	490	2,342	925	558	1,872	1,428

DN Inquiry and Transfer System

During the analysed period 146 account holders sent or received payments in the DN Inquiry and Transfer System. The daily turnover of the system ranged from 10 to 103 billion kronor (1.4 to 13.8 billion euro), with an average of 63 billion kronor (9.4 billion euro) per day. The number of payments processed in the system ranged between 490 and 2,342, with an average of 925 payments per day. The average number of payments sent on a day by a single participant was only 12, and the average value processed 869 million kronor (121 million euro). However, the system is highly concentrated: in terms of value

the three largest banks account for almost half, and the ten largest banks for almost 90% of all payments processed in the system. The low turnover days in both systems can be attributed to low market activities during Christmas and New Year.

BoF-RTGS

The BoF-RTGS system had 13 participants during the simulated period, and thus in terms of account holders the system is substantially smaller than the Danish system. In Finland, the Central Association of Finnish Cooperative Banks and the Finnish Savings Banks' Association function as central credit institutions for their member banks, which reduces the number of direct participants.

The daily turnover of the system ranged from 4.6 billion euro to 32.7 billion euro with an average of 15 billion euro per day. Some 32% of the value of payments represented incoming TARGET transfers. The number of payments processed ranged between 558 and 1,872, with an average of 1,428 payments per day. The value of payments settled in BoF-RTGS was somewhat higher than the value of payments settled in the DN Inquiry and Transfer System. As in the Danish case, the payment flows in BoF-RTGS are very concentrated; in terms of value the three largest banks account for almost two thirds of payments processed in the system.

5.5 Results

For both systems simulated, the data were extracted from the central bank accounting systems and hence the timing of payments does not necessarily reflect the timing of payments in another environment. The timing decisions made by banks are bound to be different when their liquidity holdings are reduced or increased or when the system characteristics are changed. The simulations presented in this section do not take the possible responses to system and liquidity changes into account and merely show the effects of changes in liquidity and the use of the gridlock resolution algorithm in an environment where all other aspects are kept as they are. Thus the results might overestimate queuing when liquidity is reduced in the system, both in simulations with and without the use of gridlock resolution.

5.5.1 The lower and upper bound

The systems were simulated at six different levels of liquidity ranging from the lower to the upper bound. The lower bound is the liquidity needed to settle all the payments at the end of the day, while the upper bound is the liquidity needed to settle all payments immediately. On average, the lower bounds of liquidity were 10.7% and 4.3% of the total value of payments for the Danish and Finnish data respectively. The upper bounds were 37.2% and 27.4% respectively.

Table 5.2 **Summary of liquidity requirements in both systems (EUR million)**

	DN Inquiry and Transfer System (DK)			BoF-RTGS (FI)		
	Minimum	Maximum	Average	Minimum	Maximum	Average
System UB of liquidity	634	4,925	3,421	639	5,957	2,746
System LB of liquidity	269	2,276	958	11	3,233	423
UB as % of payment flow	29.2	50.7	37.2	15.9	48.9	27.4
LB as % of payment flow	4.1	24.0	10.7	0.1	26.6	4.3

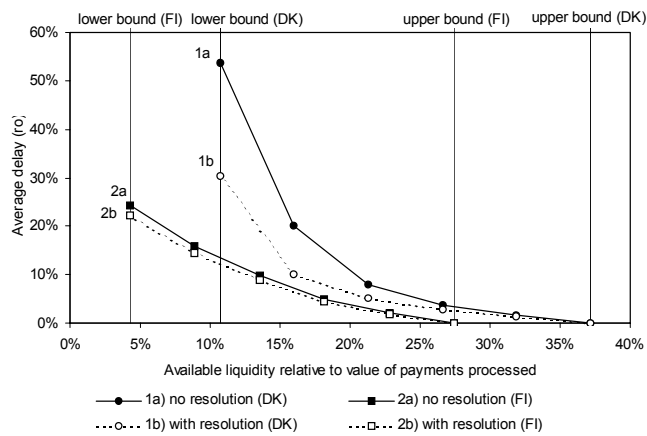
The smaller lower and upper bounds for the Finnish system shown in Table 5.2 can be explained by the fact that the number of participants in the Finnish system is smaller and the payment flows are more homogenous. The netting effect is thus higher and the intraday payment flows more balanced.

5.5.2 The trade-off between liquidity and delay

The trade-off between liquidity and delay for the simulated systems with and without the gridlock resolution mechanism is shown in Figure 5.1. The horizontal axis shows the amount of liquidity available in the system relative to the total value of payments processed in the system. The vertical axis shows the delay indicator p discussed in Section 5.4.1 above.

Figure 5.1

Trade-off between liquidity and delay



The range at which liquidity can be substituted for settlement delay at system level is rather wide in both cases; on average some 23% of the daily value of payments with Danish data and 26% with Finnish data. Both curves are convex, the curve representing simulations on Danish data to a larger extent than the curve with Finnish data. The convexity of the curves can be explained by the increase in gridlocks and the resulting increase in deadweight losses that are experienced when the available liquidity in the system is reduced.

As can be seen in the figure, the settlement delay is reduced by the proposed gridlock resolution mechanism at all levels of liquidity below the upper bound. The reduction in delay is larger at the lower levels of liquidity in both systems. Moreover, the gridlock resolution mechanism is significantly more effective with the Danish data than the Finnish data.

As shown in Figure 5.1, the delays in the system increase as the liquidity available is reduced. In addition, the variability of the settlement delay also increases. The 95% interpercentile ranges are shown for both the Danish and Finnish systems with and without gridlock resolution in Figures 5.2 and 5.3 respectively. As can be seen in Figure 5.2, the 95% interpercentile range without gridlock resolution (the area between curves 1 and 3) for the delay indicator grows from zero to almost 80% in the Danish case when the liquidity available is reduced from the upper to the lower bound. The comparable range is up to 35% with the simulation on Finnish data. Use of the gridlock resolution mechanism also reduces the variability of the settlement delay on a daily level (the area between curves 2

and 4). In particular, the gridlock resolution mechanism reduces the probability of days with severe settlement delays, since it reduces the 97.5% percentile substantially (curves 1 and 2). The reduction in the 2.5% percentile (curves 3 and 4) is not affected as much by the gridlock resolution in either set of data.

Figure 5.2 **95% interpercentile range for the delay indicator (Danish data)**

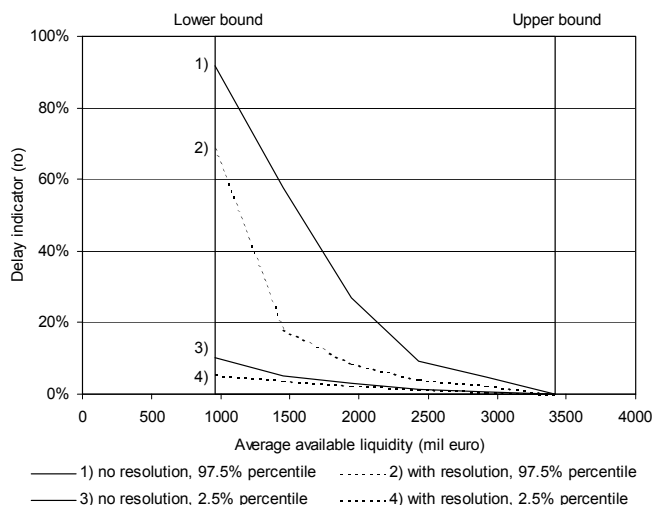
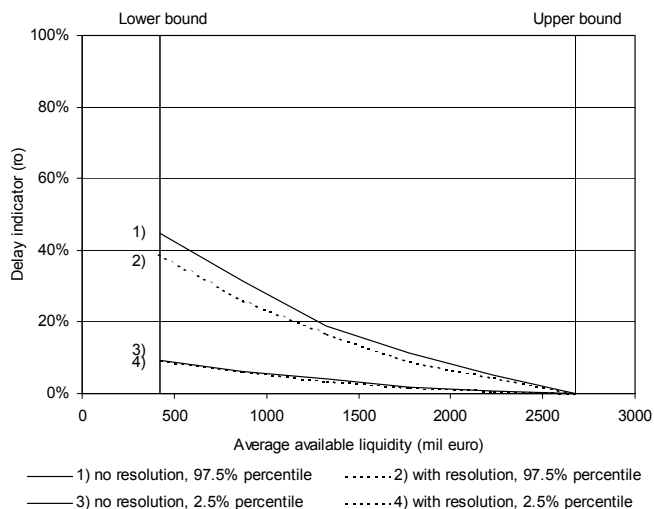


Figure 5.3 **95% interpercentile range for the delay indicator (Finnish data)**



5.5.3 Gridlocks experienced in the system

In the previous section we described how the trade-off between liquidity and delay was affected by the implementation of a gridlock resolution mechanism. In this section we look at the state of the queue and the actual solutions to the gridlock resolution problem in more detail.

The state of the queue on an average day as a function of available liquidity is shown for the two simulated sets of data in Figures 5.4a and 5.4b. The simulations show that even a modest reduction in the liquidity available leads to a substantial reduction in the fraction of the day (a total of 450 and 660 minutes, for Danish and Finnish data respectively) where no payments are queued. Reducing the level of liquidity by one fifth resulted in the formation of queues 82% of the time on an average day with Danish data, and 66% with Finnish data. At the lower bound of liquidity, queues were a permanent feature with both sets of data. In fact, on average, the systems experienced the state of ‘no payments queued’ for less than 15 minutes per day when operating at the lower bound of liquidity.

Figure 5.4a

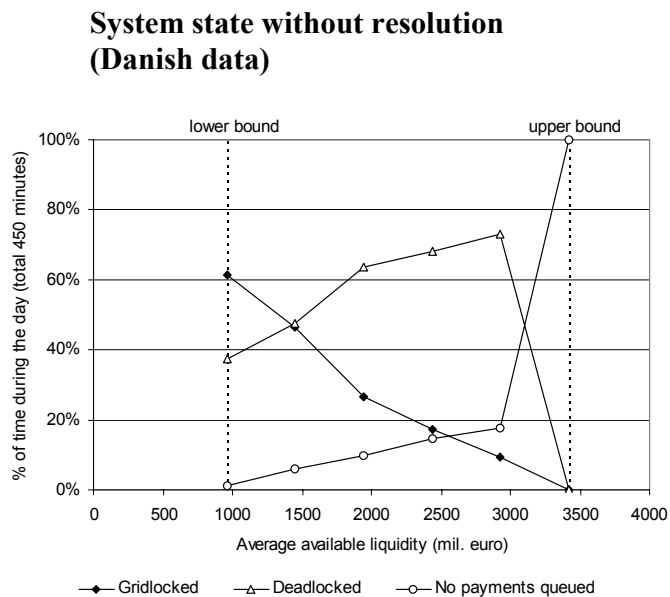
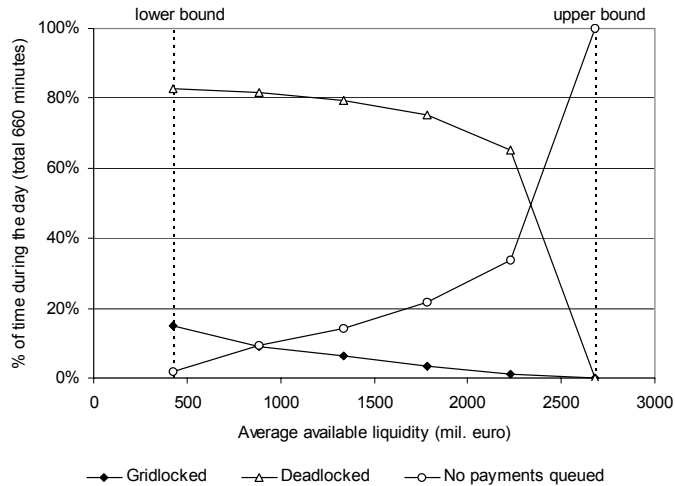


Figure 5.4b

**System state without resolution
(Finnish data)**



Furthermore, the simulations show that the fraction of the day when the queue is gridlocked increases as the liquidity available is reduced. This relationship is almost linear in both cases. In the simulations with Danish data a reduction of the available liquidity by DKK 66 million (EUR 8.9 million) in the system leads to one additional gridlock minute on average. With the Finnish data one more minute is gridlocked for every 23 million euro reduction in available liquidity for the system as a whole. The sensitivity of the system to gridlocks was thus much higher with the Danish data than the Finnish data. At the lower bound of liquidity the system simulated with Danish data was on average gridlocked 62% of the time during an average day, while with Finnish data there were gridlocks in the system only some 15% of the time.

The smaller degree of queuing and gridlocks in BoF-RTGS can in part be explained by the fact that, in the simulations, payments from participants outside Finland (ie incoming TARGET payments) were not queued. The rationale for this was that the source of these payments was another RTGS system with possibly other liquidity optimisation and management features than BoF-RTGS. As liquidity by the Finnish banks in the simulations is reduced, the effects are only partly experienced by the banks themselves, and some of the effects leak to the other TARGET RTGS systems.

The effect of gridlock resolution on the state of the system can be seen in Figures 5.5a and 5.5b. In both cases the implementation of

gridlock resolution increases the fraction of the day where no payments are queued. In the Danish case the number of minutes without any queues is considerably increased when the gridlock resolution algorithm is applied. At low levels of liquidity the increase is around half an hour and at higher levels of liquidity between 10 and 20 minutes. With Finnish data the reduction is not as large, but still noticeable. At the lower bound of liquidity the reduction was on average 20 minutes, at higher levels of liquidity between 3 and 5 minutes.

Figure 5.5a

**The effects of resolution on system state
(Danish data)**

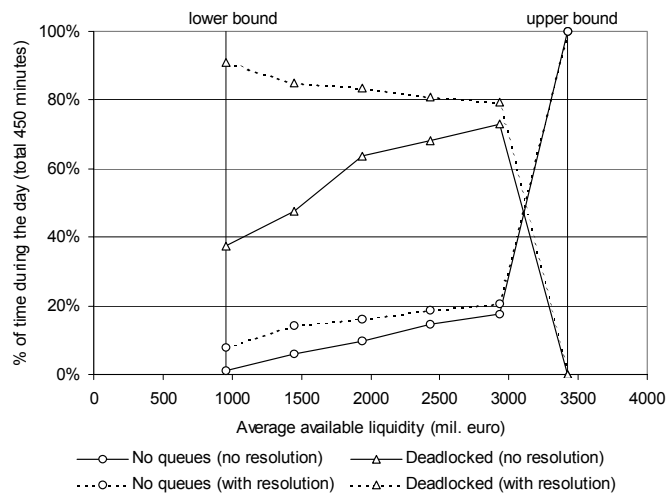
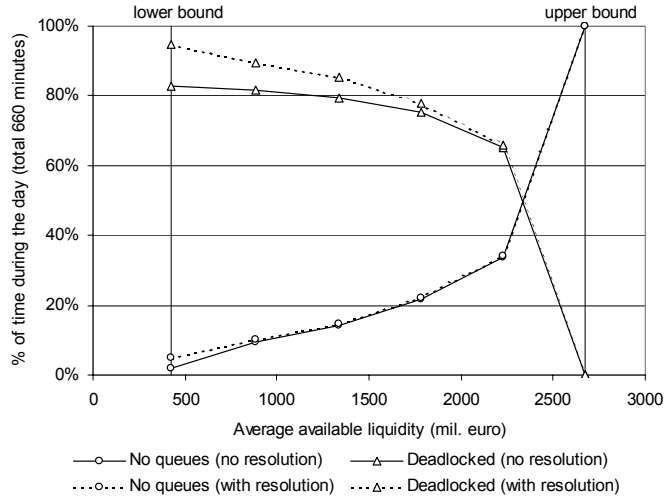


Figure 5.5b

**The effects of resolution on system state
(Finnish data)**



However, most of the gridlocked queues could only be partially resolved, and simultaneous settlement of all queued payments was possible only rather rarely. Full resolution was more likely at high levels of liquidity than low levels of liquidity. At the lower bound of liquidity only some 11% (DK) and 20% (FI) of gridlocks were fully solvable in the respective systems. Full gridlock resolution was successful with Danish data on the average every second day, and with Finnish data every third day. In contrast, the partial gridlock resolution feature was successful on the average 5 and 2 times a day with the Danish and Finnish data respectively. The fact that most of the resolutions were only partial explains why the fraction of the day where the systems have a deadlocked queue also increased as a result of implementing a gridlock resolution mechanism. Recall that a deadlocked queue can only be settled by the infusion of additional liquidity by one or more participants with queued payments. When gridlocks are removed, either all queued payments are settled or the remainder of the queue is left deadlocked. For any given amount of liquidity an efficient system is one with no queues or with queues that deadlocked.

5.5.4 Resolution under failure scenarios

The simulations so far have studied the relationship between liquidity and delay and the effects of gridlock resolution on delays experienced in the system under 'normal circumstances'. In order to test whether gridlock resolution would also be helpful in crisis situations two failure scenarios were modelled: a financial failure scenario and an operational failure scenario. The simulations were carried out with Danish data only.

In the *financial failure* scenario we assume that at the beginning of the day the largest bank in the system (in terms of value of payments sent) goes bankrupt and is removed from the system. Hence, it does not send or receive payments during the day.

In the *operational failure* scenario we assume that the largest bank is unable to connect to the RTGS system and thus unable to send payments until the last opening hour of the system. We look at two variations of this scenario. In the first variation we assume that other banks are not aware of the failure and continue sending payments to the failing bank. In the second variation we loosen up the assumption of other banks being unaware of the operational failure of the largest bank. The treasurer of the bank would be hesitant to keep on sending payments to a bank from which he has not received any payments for the entire day.

We model this behavioural response by giving a lower priority to payments to the failing bank. This could enable settlement to continue between non-failing banks to a greater extent. In all the failure scenarios, the liquidity available for banks was the upper bound of liquidity ($\alpha = 1$), ie the amount of liquidity by which all payments could be settled without delays under normal circumstances.

The effects of both a financial and an operational failure of the largest bank on delays in the Danish system are shown in Table 5.3. In the financial failure scenario the delays were rather modest on average, but on some days not insignificant. An operational failure in the system was more severe in terms of delays than the financial failure of the largest bank. This stems from the fact that in the operational failure scenario the recycling of liquidity in the system was severely slowed, as a large proportion of the liquidity was trapped on the failing bank's account until the last hour.

The average value of the delay indicator in the operational failure scenario (0.12) is by and large comparable to the delay experienced by the system when the liquidity available was halved across the board in the previous simulations with no failures. This is presumably due to the high degree of concentration in the system, as noted above.

Banks are however able to reduce settlement delays by undertaking counter measures when faced with the operational problems of other banks in the system. We model this by assigning a low priority to payments to the failing bank. However, the simulations show the effect of this particular response to be modest, as the delay indicator is reduced by only 1 basis point (see Table 5.3).

Table 5.3 **The effect of failure on settlement delay**

	Financial failure	Operational failure	Operational failure with prioritisation
Average delay	0.04	0.12	0.11
Maximum delay	0.26	0.31	0.30

Note: If all payments are settled immediately, the delay indicator is $\rho=0$, whereas $\rho=1$ if all payments wait until the end of the day.

In the financial failure scenario not all payments could be settled during the day with the liquidity available to the participants. This was due to the changed end-of-day liquidity requirements of the banks resulting from the removal of the payments of the largest bank.

On average some 4.5% of the total value of payments remained unsettled. However, on the worst of the simulated days unsettled payments amounted to 27% of all payments processed. The gridlock resolution mechanism was able to reduce the value of unsettled payments only slightly.

This implies that the cause was mainly a lack of liquidity and not a gridlock at the end of the day. In the operational risk scenario, payments remained unsettled only on one day, and by employing the algorithm all payments could be settled during that day.

When looking at gridlocks experienced in the system under the failure scenarios (Table 5.4) we notice that in the financial failure scenario gridlocks occurred only on six days out of the total of 64 days. Thus, for the majority of days the presence of gridlock resolution had no effect. The gridlocks that occurred were, however, quite persistent, as the system was gridlocked on these days for almost a third of the day on average.

Table 5.4

Gridlocks and effect of resolution under failure scenarios

	Financial failure	Operational failure	Operational failure with prioritisation
Number of days gridlocks occurred	6	24	25
Average share of day gridlocked	30%	19%	17%
Maximum share of day gridlocked	62%	60%	52%
Average decrease in settlement delay by gridlock resolution	31%	11%	10%
Maximum decrease in settlement delay by gridlock resolution	63%	43%	39%

Note: The total number of days was 64. The numbers reported are on days when gridlock occurred.

By resolving these gridlocks the settlement delay during these six days was decreased by 31% on average on those days.

In the operational failure scenarios, gridlocks were much more common. Without prioritisation of payments between non-failing banks, the system was gridlocked on 24 of the simulated 64 days. However, only an 11% reduction in the settlement delay could be achieved by resolving these. By assigning lower priority to payments to the failing bank, the system performance could not be significantly improved. When interpreting the results of the simulations it is important to bear in mind that data reflects the banks' choices as to eg the timing of payments on a given day, taking into account the liquidity available. It is therefore likely that these choices would be different under other circumstances.

5.6 Conclusions

When banks do not hold sufficient liquidity to settle all obligations immediately the system will experience gridlocks that could be solved by an appropriate algorithm. For both simulations with Danish and Finnish data we found an almost linear relationship of increasing gridlocks as the liquidity in the system was reduced.

For the Danish data the system was up to 62% of the time in gridlock when liquidity was scarce. The potential benefits from a resolution of gridlocks are thus clear. By applying the algorithm,

queuing could be substantially reduced, by up to 50% at low levels of liquidity for the Danish data. At higher levels of liquidity, the effects were more modest, but still concrete. With the Finnish data the effects were relatively modest, mainly due to the fact that the system experienced much fewer gridlocks. This could possibly be explained by the fact that payments arriving from other RTGS systems in TARGET were not queued and represented an extra source of liquidity for the banks. The liquidity received from these payments could be used for settling outgoing payments irrespective of the liquidity used in the simulations. Thus the scope for optimisation was smaller in the Finnish case, which was simulated as an 'open' system in contrast to the 'closed' Danish system.

The proposed algorithm was also found to reduce the risk of severe settlement delays, with the Danish data to a larger extent than with the Finnish data. This represents a further improvement in the system and should enable banks to operate on lower levels of liquidity.

The use of the algorithm can be seen as an improvement to the current system designs, as all other aspects of settlement can remain as they are. The banks can continue to be sure that the payments are released in their preferred order, as any sequence constraint can be used in solving the Gridlock Resolution Problem, reducing the legal risk involved. Due to this fact, the algorithm is also invariant whether prioritisation of payments or multiple payment classes are used. In these cases the underlying set of payments is just organised differently. Further on, the calculation time of the algorithm increases only linearly with the size of the problem, resulting in fast calculation times even with large numbers of queued payments. Thus the ongoing RTGS settlement does not have to be suspended for a significant period of time.

Overall, we find that gridlock resolution, from the central bank perspective, can be seen as a tool to ensure a smoother functioning of the system when liquidity holdings of the participants are low, or when the system is experiencing a temporal shortage of liquidity. From the participants' perspective, it reduces the costs associated in settlement, either through decreased delay costs or through their ability to hold less liquidity for the settlement of payments without incurring more delays in payments. We conclude that the algorithm explained in this paper is of practical importance, efficient, and easy to implement in the design of existing RTGS systems.

Furthermore, the gridlock resolution might on some occasions be able to remedy liquidity squeezes caused by financial or operational failures of banks participating in payment systems. However it cannot

be seen as a substitute for sufficient prudential supervision and payment system oversight that reduces the likelihood of these events.

References

- Angelini, P (1998) **An Analysis of Competitive Externalities in Gross Settlement Systems.** Journal of Banking and Finance 22, 1–18.
- BIS (1993) **Payment systems in the Group of Ten countries.** Prepared by the Committee on Payment and Settlement Systems of the central banks of the Group of Ten countries, Bank for International Settlements, Basle.
- BIS (1997) **Real-time gross settlement systems.** Prepared by the Committee on Payment and Settlement Systems of the central banks of the Group of Ten countries, Bank for International Settlements, Basle.
- Bech, M L and Garratt, R (2003) **The Intraday Liquidity Management Game.** Journal of Economic Theory 109, 198–219.
- Bech, M L and Soramäki, K (2001) **Gridlock Resolution in Interbank Payment Systems.** Discussion Paper 9/2001, Bank of Finland.
- Bech, M L and Soramäki, K (2002) **Liquidity, Gridlocks and Bank Failures in Large Value Payment Systems.** Robert Pringle and Matthew Robinson (eds.) E-money and Payment Systems Review, Central Banking Publications Ltd, 111–125.
- Güntzer, M M – Jungnickel, D and Leclerc, M (1998) **Efficient algorithms for the clearing of interbank payments.** European Journal of Operational Research 106, 212–219.
- Horowitz, E and Sartaj S (1978) **Fundamentals of Computer Algorithms.** Computer Science Press, Inc.
- Kahn, C and Roberds, W (1998) **Payment System Settlement and Bank Incentives.** The Review of Financial Studies 11, 845–870.
- Koponen, R and Soramäki, K (1998) **Intraday Liquidity Needs in a Modern Interbank Payment System – A Simulation Approach.** Studies in Economics and Finance E:14, Bank of Finland.

Leinonen, H and Soramäki, K (1999) **Optimising Liquidity Usage and Settlement Speed in Payment Systems**. Discussion Paper 16/99, Bank of Finland.

McAndrews, J and Rajan, S (2000) **The Timing and Funding of Fedwire Funds Transfers**. FRBNY Economic Policy Review, Vol. 6, No. 2, July.

Shafransky, Y M and Doudkin, A A (in press) **An optimisation algorithm for the clearing of interbank payments**. European Journal of Operational Research.

Chapter 6

Economising liquidity with deferred settlement mechanisms^{*}

Kurt Johnson – James J. McAndrews – Kimmo Soramäki

6	Economising liquidity with deferred settlement mechanisms	180
	Abstract.....	180
6.1	Introduction.....	180
6.2	Characteristics of deferred settlement	182
6.3	System designs.....	183
6.3.1	Receipt-reactive gross settlement system.....	184
6.3.2	One-hour and six-hour netting system	185
6.4	Simulations of the three systems	186
6.4.1	Daylight overdrafts and delay indicators.....	187
6.5	Sensitivity analysis.....	189
6.5.1	Liquidity.....	191
6.6	End-of-day queue analysis.....	193
6.7	Discussion of simulation results and likely behavioral responses	195
6.7.1	Bank participation and risk.....	196
6.8	Conclusion	199
Appendix 6.1	Detailed dynamics of the receipt-reactive gross settlement system	200
Appendix 6.2	Simulation results and analysis	202
Appendix 6.3	Funds overdraft accounting procedure	213
References	215

^{*} This is a reprint from the *Economic Policy Review*, Volume 10, Number 3, December 2004, Federal Reserve Bank of New York. The publisher gratefully acknowledges the permission to reprint this article. The views expressed do not necessarily reflect the views of the Federal Reserve Bank of New York, or of the Federal Reserve System, or of the European Central Bank.

6 Economising liquidity with deferred settlement mechanisms

Abstract

Real-time gross settlement (RTGS) systems such as the Federal Reserve's Fedwire Funds Service enable participating banks to settle payments immediately and in the full amount; however, the high level of liquidity inherent in the systems requires large intraday credit extensions.

An examination of several deferred settlement mechanisms that could potentially complement RTGS systems includes a novel mechanism – a receipt-reactive gross settlement system – that bases the settlement of a bank's payments on the value of receipts over a given time rather than on the bank's balance.

The results indicate that the RRGs mechanism can significantly reduce intraday credit extensions while modestly delaying the average time of payment settlement. In addition, this mechanism provides good incentives for banks to submit payments earlier in the day.

6.1 Introduction

On a typical day, the total value of payments settled by the Federal Reserve's Fedwire Funds Service exceeds \$1.8 trillion. On average, credit extended to banks using Fedwire is about \$30 billion over the course of the day, while the peak intraday amount reaches \$86 billion. Given this high level of credit extensions, it is worthwhile asking whether payment settlements could be managed with a lower level of outstanding credit, thus allowing system operators to economise liquidity.¹

Fedwire operates as a real-time gross settlement (RTGS) system. RTGS systems transfer the full amount of payment orders between commercial bank participants immediately upon receipt, thus avoiding short-term debt obligations between participants. This is a desirable feature that has prompted many central banks worldwide to implement these systems over the past decade. However, because payment

¹ On average, overnight deposits made by commercial banks are worth about \$15 billion.

transfers between participants are made immediately in the full amount, and because of the asynchronous timing of payments by participants, maintaining the liquidity needs of RTGS systems can be costly. Indeed, some system operators have altered their RTGS systems in recent years to economise on the funds needed to complete settlements.

One way to reduce a system's liquidity needs is by using deferred settlement mechanisms such as netting. In netting systems, payment orders are deferred until some designated time – usually late in the day – when the participants exchange only the net amounts they owe or are owed. If all participants successfully submit these net amounts, the system settles all the payments accumulated during the day with the least amount of funds possible – that is, just the net amounts. To achieve this economy in funds use, a netting system delays the settlement of payments so that all orders remain pending until the net settlement payments are completed successfully. This delay feature creates distinct liquidity and risk management characteristics.

Another type of deferred settlement mechanism queues payments as they enter the system. Some European RTGS systems use these 'queue-augmented RTGS systems,' or hybrid systems.² Such systems save on liquidity – as in a netting system – but with less delay than end-of-day netting imposes.

In this article, we propose alternative ways of settling payments submitted to the Fedwire Funds Service that would result in lower intraday credit extensions. We analyse the effects of complementing an RTGS system with various deferred settlement mechanisms by performing simulations on historical Fedwire data. Although others have studied the effects of such modifications on payments systems, this is the first examination in the context of Fedwire.

One function of a payments system design could be to minimise the combined cost of delaying payments and the risk of extending intraday credit to commercial banks – that is, the credit risk that a central bank assumes by providing intraday credit. We do not use an explicit, objective function to evaluate the various alternatives to an RTGS system because we do not know banks' preferences regarding delays or specific default risks. We can, however, evaluate those designs that reduce both delays and credit extensions as preferable to others. In short, some modifications may clearly be more effective than others but none compares easily with a pure RTGS system, which

² See McAndrews and Trundle (2001) for a description of several new designs that have been put into use in various large payment systems and for a discussion of some specific policy issues that are associated with the novel designs.

by definition eliminates delays. Our results suggest that, compared with RTGS systems, alternative settlement designs could significantly reduce credit extensions while modestly delaying the average time of settlement of payments.

This article is organised as follows. In the next section, we discuss the basics of deferred settlement systems, and in Section 6.3 we describe the systems used in our simulations. In Section 6.4, we describe the performance metrics and results of the basic simulations, and in Section 6.5 we conduct a sensitivity analysis of the results obtained from alternative levels of queued payments. Sections 6.6 and 6.7 present more detailed analyses of liquidity use and the level of end-of-day queues. In Section 6.8, we discuss our results and the likely behavioral responses by a bank participant to the availability of the simulated systems. We conclude with a discussion of our results in light of the previous literature.

6.2 Characteristics of deferred settlement

As a baseline case, a pure RTGS system is one in which no payments are deferred for settlement – all payments are released upon receipt by the system operator as long as the participant has adequate funds to settle the payments. If not, the payment is rejected. Deferred settlement can work in conjunction with RTGS systems. In practice, deferred settlement mechanisms can operate in many ways, but all require certain criteria by which payment orders are entered, ranked, and settled. In addition, criteria for the end-of-day closing or the emptying of queues are required in a queuing system.

The *entry criterion* in a deferred settlement mechanism determines whether payments are deferred or whether they are settled immediately by pure RTGS. This criterion can be based either on decisions made by participating banks or on an automatic feature created by prespecified criteria. In many European RTGS systems, deferment is automatic: Rather than reject payment orders outright, the systems automatically place payments in a queue if RTGS settlement of the payment would breach the credit limit of the participant.

The *order criterion* defines the ranking or ordering of payment messages that are queued. Most contemporary RTGS systems adhere to the first-in, first-out (FIFO) principle – that is, payments that enter the queue earlier have priority over payments that enter the queue later. The FIFO principle is easy to implement. Assuming liquidity

constraints, FIFO performs reasonably well when a system's smaller payments are submitted generally earlier in the day. Because early, larger payments can obstruct FIFO-ordered queues, some RTGS systems now have 'bypass-FIFO' algorithms that allow participants to reorder and prioritise queued payments. Some, such as the CHAPS Clearing Company in the United Kingdom, allow the participants to select the order of the queued payments according to other criteria, for example, by the value of the payments.

The *settlement criterion* defines the rules by which payment messages are released from the queue, triggering the flow of money from the payer's account to the payee's account. Payments can be released from the queue either individually or in groups. In most queuing arrangements, payments are released individually, but some systems employ 'gridlock-resolution' algorithms that allow multiple payments to be settled from the queue simultaneously if the release of the payments on an individual basis is not possible (see Bech and Soramäki (2001, 2002)). In previously described queuing systems, the settlement criterion has been based on the balance of the participant – that is, payments are released as soon as the participant's balance is high enough to cover the payment's settlement. The release of payments from the queue can be based on other criteria as well. In this article, we propose a novel deferred settlement mechanism that bases the settlement of queued payments on the value of incoming payments rather than on a participant's account balance.

The last important element of the design of a queuing system is *end-of-day close of queues*, or how to 'empty' the queue of payment orders at the close of business. One method used in some European systems is to settle all messages remaining in the queue after a certain length of time through an exchange of the net amounts of the payments. An alternative method of closing queues is to return unreleased payments to banks before the RTGS system closes so banks can redirect the unreleased messages to the RTGS system.

6.3 System designs

In our simulations, we design a system in which payments go through one of two alternative channels – deferred settlement or real-time gross settlement. These two main channels recognise that some payments are more time-critical than others. Banks would likely want to settle their time-critical payments through the immediate RTGS channel while the less time-critical payments could go through the

deferred settlement channel to save liquidity. The deferred settlement channel allows payments to take one of three possible paths: a one-hour netting system, a six-hour netting system, or a unique type of system that we call a receipt-reactive gross settlement (RRGS) system.

6.3.1 Receipt-reactive gross settlement system

In the receipt-reactive gross settlement system that we simulate, a portion of banks' payments is randomly selected to be settled through RTGS, while the remaining set of payments is placed in a queue for deferred settlement. Queued payment messages are ranked on a first-in, first-out basis. For the release criterion, a payment message's release from the queue is triggered by the arrival of incoming funds received by the bank within a specified period of time. In our simulations, this time period is one calendar minute.³ The system will release within one minute as many payments from the front of the queue as possible to offset – but not exceed – the amount of incoming funds. In the simulations, this process continues throughout the day until 5:30 p.m., when the queue closes. The exhibit below shows the effects of a receipt-reactive queue on a bank's balance during a minute in which the bank both sends and receives RTGS payments (see Appendix 6.1 for a detailed description of this queuing system).

An important design element of a queuing system is how to 'empty' the queue of payment messages at the end of the day. In our simulations, we randomly reassign the unreleased payments a settlement time between 5:30 p.m. and 6:00 p.m. This is our approximation to returning, at 5:30 p.m., the unreleased (that is, unsettled) payments to the banks that submitted them, with the banks subsequently resubmitting the payments through the RTGS system over the next half-hour.

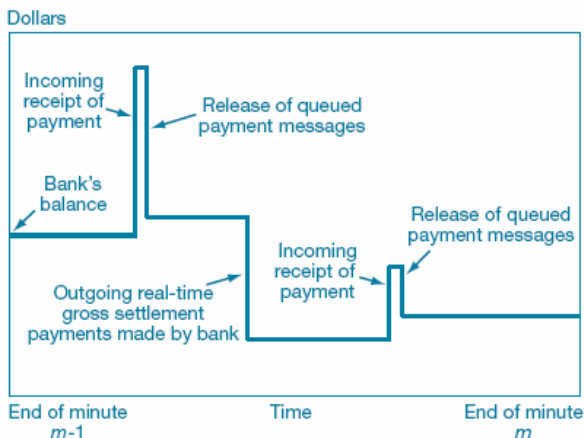
This feature of our queue design was chosen for a number of reasons. The end-of-day return of the queued messages is consistent with the basic design of queuing systems used as adjuncts to RTGS systems. The receipt-reactive queuing system is intended to encourage banks to submit to the queue those payments that can be settled with incoming funds. If particular payments cannot be settled that way, then the banks would likely submit the payments through the RTGS

³ We choose one calendar minute as our time period because banks' daylight overdrafts are calculated as of the end of each calendar minute. By choosing a one-minute period within which receipts can be set against the release of payments from the queue, we prevent the release of those payments from causing an overdraft at the end of the minute.

system, which is designed to settle payments immediately with available balances. Therefore, returning banks' payments that remain in the queue to the respective banks near the end of the day is consistent with the intended use of the receipt-reactive queue and of the RTGS system.

Figure 6.1

Dynamics of a bank's balance under receipt-reactive gross settlement



6.3.2 One-hour and six-hour netting system

We simulate the performance of two simple netting systems. In both systems, a portion of all payments is settled by RTGS while a complementary set of payments is put in a queue. The queued payments are cumulated for a certain period of time, netted, then settled – even if they cause an overdraft in the banks' account balances. In the first simulation, the payments are cumulated, netted, then settled after each hour of the operating day. In the second simulation, cumulated payments are netted and settled after six hours. (See Table 6.1 for a summary of the simulations.)

Table 6.1

Summary of simulated systems

System	Entry criterion	Order criterion	Release criterion	End-of-day close of queue
One-hour netting	Randomly selected 50 per cent of payment orders. In conducting our sensitivity analysis, we randomly selected either 20 per cent or 50 per cent of payments	Not applicable	All payment orders for all banks in queues are netted and released at <i>one-hour</i> intervals.	All payment orders in queue are netted at 6:30 p.m., the end of the Fedwire day.
Six-hour netting			All payment orders for all banks in queues are netted and released at <i>six-hour</i> intervals	
Receipt-reactive gross settlement		First-in, first-out	A payment order at the front of a bank's queue is released from queue as receipts for the bank, within a calendar minute, exceed the value of the payment order to be released	Any payment orders remaining in queues at 5:30 p.m. are randomly and uniformly assigned and settled by real-time gross settlement over the next thirty minutes.

6.4 Simulations of the three systems

To gauge the usefulness of these three complements to RTGS systems – one-hour netting, six-hour netting, and receipt-reactive gross settlement – we simulate their performance using a program developed by the Bank of Finland. The simulator is described in detail in Leinonen and Soramäki (2003). (See also the Bank of Finland's website: <http://www.bof.fi/sc/bof-pss>.)

In the simulations, we initially assume that banks would submit half of all individual payments to the queue for deferred settlement and the other half of their payments to the RTGS system. In the simulations, we include all Fedwire funds transfers for a randomly selected set of ten days between October 1999 and February 2000. We perform a sensitivity analysis (see Section 6.5) and present robustness checks to gauge the effects resulting from different levels of participation in the three deferred settlement arrangements. Using three days of data, we reproduce the simulations with either 20 per cent or 80 per cent of all payments assigned to the deferred settlement mechanism. The simulations are conducted on historical payment transactions similar to the generalised example below (figure 6.2)

Figure 6.2

Payment transaction example

Sender Account number	Receiver Account number	Value sent (dollars)	Submit time (hour:minutes)	Routing flag
02100xxxx	02100yyyy	100.50	10:20	1 or 0

In this example, Bank A (with American Banking Association account number 02100xxxx) sends Bank B (02100yyyy) \$100.50 at 10:20 a.m. through the Fedwire service. If the routing flag is one, the payment is routed to the deferred settlement mechanism. If the flag is zero, the payment is routed to the RTGS system. The routing flag randomly assigns a one or a zero according to a predetermined level of participation in the deferred settlement mechanism (for example, 20 per cent, 50 per cent, or 80 per cent of the day's payments have routing flags equal to one). For any given day, several hundred thousand transactions are routed, one by one in the order of their time stamps, to the settlement mechanism assigned. If routed to RTGS, a payment will be settled immediately. If routed to the deferred settlement mechanism, settlement could be delayed. We report the detailed results of these simulations in Appendix 6.2.

Our primary performance metrics focus on the system's impact on daylight overdrafts and on the delay in the time of payment – that is, the difference between the time the payment was submitted and the time it was settled.

6.4.1 Daylight overdrafts and delay indicators

A bank incurs a daylight overdraft when its balance falls below \$0. The Federal Reserve measures daylight overdrafts outstanding at the end of each minute of the Fedwire operating day and reports the aggregate peak and average overdrafts for all banks. The aggregate *peak overdraft* occurs in a specific minute in which the aggregate overdraft has the highest value of all the minutes in the Fedwire operating day; the *average overdraft* is the sum of all the banks' overdrafts for all minutes of the day divided by the number of minutes in the Fedwire operating day. We focus on the average overdraft because that is the basic measure used by the Federal Reserve to calculate the fees it charges banks for their credit use, which is inherent in daylight overdrafts (see Appendix 6.3 for a description).

The basic indicator of delay that we consider is the *average time of settlement* across all payments and banks. It is the average time at which payments settle – where the time of actual settlement is weighted by the value of the payment.

The *delay statistic* is a standardised indicator that may take values between zero and one. RTGS, with its immediate settlement, results in a zero-delay statistic. End-of-day netting results in a delay statistic equal to one. In comparison with the average time of settlement

statistic, the delay statistic tends to weight more heavily payments that are entered early in the day, even if the payments are small in value.

We compare the three simulated queue-augmented RTGS systems with Fedwire’s historical performance using these liquidity and delay metrics (Table 6.2). We find that only the receipt-reactive gross settlement system reduces the use of overdraft liquidity by a statistically significant amount. While the two netting systems affect the average overdraft, neither of these differences is statistically significant (see Appendix 6.2).

Table 6.2 Averages from simulations with 50 per cent participation

Treatment	Average overdraft (billions of dollars)	Percentage change from RTGS	Peak overdraft (billions of dollars)	Percentage change from RTGS	Average time of settlement	Delay statistic (per cent)
Real-time gross settlement (RTGS)	20.29	–	66.28	–	14:35	0
One-hour netting	20.41	0.58	64.69	–2.40	14:51	7.53
Six-hour netting	19.45	–4.10	60.15	–9.25	15:46	34.35
Receipt-reactive gross settlement	17.52	–13.64	72.66	9.62	15:18	13.74

Source: Authors’ calculations, based on data from Federal Reserve Bank of New York, Fedwire Funds Service.

The one-hour netting mechanism shows no statistically significant change in the average overdraft relative to RTGS. Payments are delayed the least under one-hour netting, with the average time of settlement of payments moving to 2:51 p.m. from the RTGS average time of 2:35 p.m. For six-hour netting, the decrease in overdrafts is higher (although it is not statistically different from no change), but so is the delay in payments, as the average payment time moves to 3:46 p.m. – the latest time of all the alternatives. The receipt-reactive gross settlement system reduces average overdrafts by about 14 per cent – the largest amount among the alternatives considered here. The time of payment is 3:18 p.m., slightly later than for the one-hour netting alternative. The average time of settlement increases by seventy-one minutes for six-hour netting while it increases by forty-three minutes for the receipt-reactive queuing system – a 65 per cent difference – while six-hour netting shows a 150 per cent increase in the delay statistic when compared with the receipt-reactive system. Our results suggest that the receipt-reactive system performs markedly better in settling early-morning and early-afternoon payments, as well as smaller ones.

The receipt-reactive queuing system’s increase in peak overdrafts, in conjunction with its decrease in average overdrafts, is a bit puzzling (Table 6.2). Why do peak overdrafts increase if payments remain in

the queue until sufficient receipts arrive for the bank? The answer is related to the current timing of payments on Fedwire. As described in McAndrews and Rajan (2000), a high proportion of the value transferred over Fedwire occurs in the late afternoon. During that period, banks have a large number of payment receipts, as well as payment outflows. Under the receipt-reactive gross settlement system, the high level of receipts that occurs during the late afternoon (from the 50 per cent of payments that continue to be settled through the RTGS system) begins to trigger the release of a large number of queued payments. That process cascades as the payments released from the queue are receipts for other banks, which triggers further releases from the queues. For many banks – primarily large banks – the outflow from the queue ‘absorbs’ all of the bank’s receipts. Few, if any, of its receipts at that time of day add to its balance; instead, the receipts facilitate the release of payments from queues. The combination of receipts being dedicated to the release of payments from the queue and the submission of many RTGS payments at that time of day drives many accounts deeply into overdraft. Smaller banks, not having the same heavy outflow of payments, enjoy significantly more positive balances at that time, as the cascade of payments occurs. On balance, this process further concentrates payment activity in time, and reduces the use of overdrafts on average, even though the peak overdraft is increased. Of course, we would expect that banks would alter their behavior if they expected such a cascade of payments to be triggered.

6.5 Sensitivity analysis

To test the robustness of our results, we examine two simulations in which 20 per cent and then 80 per cent of payments are randomly selected for deferred payment. These simulations were conducted on three days of data. Because an analysis of only three days yields such a small sample, we do not consider statistical significance, but simply report averages.⁴

One potential problem with a receipt-reactive gross settlement system is that banks may make an excessive number of submissions to that system and very few to the RTGS system. If all banks find it

⁴ A single day’s simulation for the receipt-reactive system with 80 per cent of payments placed in the queue required more than two months to complete, using the single computer we employed for the simulations.

convenient and economical to submit payments to the queue for deferred settlement, then a paucity of RTGS payments might cause widespread payment delays.⁵ By varying the number of payments submitted for deferred settlement, we can evaluate how each of the alternative systems would perform, assuming different behavioral patterns (see Table 6.3 for the results of these simulations). The numbers for the RTGS system are calculated using only the three days pertinent to the alternate level simulations and thus differ from those in Table 6.2.

Table 6.3 System averages comparing alternative levels of participation

Treatment	Average overdraft (billions of dollars)	Percentage change from RTGS	Peak overdraft (billions of dollars)	Percentage change of RTGS	Average time of settlement	Delay statistic (per cent)
Real-time gross settlement (RTGS)	20.52	-	68.49	-	14:32	0
One-hour netting						
20 per cent	20.37	-0.74	67.16	-1.93	14:37	2.52
50 per cent	20.55	0.12	65.27	-4.69	14:47	7.31
80 per cent	19.83	-3.38	64.66	-5.59	14:53	9.93
Six-hour netting						
20 per cent	20.53	0.04	64.42	-5.94	15:01	13.97
50 per cent	19.87	-3.16	62.50	-8.74	15:45	34.89
80 per cent	17.32	-15.59	50.77	-25.87	16:28	55.44
Receipt-reactive gross settlement						
20 per cent	20.88	1.75	71.47	4.36	14:48	3.35
50 per cent	17.38	-15.33	70.40	2.80	15:18	14.78
80 per cent	11.49	-44.05	55.92	-12.52	16:29	22.50

Source: Authors' calculations, based on data from Federal Reserve Bank of New York, Fedwire Funds Service.

These simulations suggest that our results are relatively robust at different levels of participation.

The level of liquidity savings and the length of delays in settlement increase as more payments are submitted to the queues. Of interest is the result that the RRGs system maintains comparable or higher levels

⁵ In such a case, a gridlock-resolution method of settling payments from queues might be useful. See Bech and Soramäki (2001) for a discussion.

of liquidity savings for a given delay than does the six-hour netting system, at all levels of submission to the queue. In addition, even with 80 per cent of payments submitted to the queue, the RRGs system imposes an average settlement delay that is no greater than the delay with six-hour netting. However, the important difference is that RRGs does this with much greater liquidity savings than six-hour netting does. The delay statistic provides some evidence for why this occurred. The average time it takes to settle a payment for the 80 per cent receipt-reactive system is one minute more than the average time for six-hour netting, suggesting that the receipt-reactive system holds more large payments in its queue until the end of the day. The delay statistic for six-hour netting is about 2.5 times higher than the statistic for the receipt-reactive system, indicating that the RRGs queue outperforms the six-hour netting system in overdraft management

6.5.1 Liquidity

To provide more insight into the performance of these deferred settlement systems, we need to understand how the mechanisms use liquidity throughout the day. While our initial analysis focused on the mechanisms' effects on overdrafts and delays, we now look at four measures of liquidity, assess their levels in each system, then compare the levels with those found in the RTGS system (see Appendix 6.2 for further details).

The *average funds transfer* statistic measures the average level of liquidity, or funds, that must be transferred from an individual account from one minute to the next, across the minutes of the day, to complete the payments for the day. For example, in the RTGS system, the average funds transfer for an account is \$226,000 per minute. To make this transfer, the bank must have sufficient funds in its account or receive sufficient funds from other banks. For any bank in the system, the average funds transfer may take values between zero and the per-minute gross value of payments sent by the bank.

In general, netting exhibits lower average liquidity usage over the day as indicated by the average funds transfer measure compared with an RTGS or RRGs system. Once queued, payments must stay in the queue until the next net settlement time and do not cause any balance changes in the interim. This liquidity conservation increases with longer netting times and larger participation in the netting systems (Table 6.4).

Table 6.4

Liquidity, correlation, and skewness results

Treatment	Average funds transfer (thousands of dollars)	Maximum funds transfer (thousands of dollars)	Correlation with RTGS balances (per cent)	Skewness
Real-time gross settlement (RTGS)	226	2,206	100.0	–
One-hour netting				
20 per cent	209	5,521	99.5	13.9
50 per cent	170	10,808	98.0	18.8
80 per cent	130	13,901	97.2	23.4
Six-hour netting				
20 per cent	201	10,944	96.2	20.3
50 per cent	145	12,122	90.5	23.2
80 per cent	86	17,356	84.7	19.0
Receipt-reactive gross settlement				
20 per cent	209	1,827	94.8	13.0
50 per cent	165	1,975	86.6	26.2
80 per cent	140	3,315	76.0	24.2

Source: Authors' calculations, based on data from Federal Reserve Bank of New York, Fedwire Funds Service.

If we look at the minute with the highest average funds transfer, we obtain the *maximum funds transfer*. Using this indicator, we see that the netting systems demonstrate very high liquidity requirements during their net settlement periods – up to \$17 million. These one-minute maximums occur near the end of the day when the Federal Reserve's RTGS system experiences its usual peak in volume. The receipt-reactive gross settlement system, however, produces lower maximum liquidity outflows, demonstrating its ability to smooth the liquidity usage over time. This smoothing – a feature of the more dynamic receipt-reactive queuing mechanism – results in greater average liquidity usage than the netting alternatives, but also produces significantly smaller maximum requirements. The lower maximum requirements are important as it is generally less costly for banks to make small payments that are distributed throughout the day than to make large payments. The smoothing effect of the RRGs system is thus a desirable feature.

The *correlation with the RTGS* measure represents the degree of correlation and hence the exhibited independence between the end-of-minute balances of RTGS and the other simulated systems. The one- and six-hour netting mechanisms exhibit a greater correlation between their balances and the original balances created by the RTGS system, despite the netting systems' lower average figures of liquidity use.

The lower correlation for RRGs indicates that, while it results in higher balance transfers than the netting systems throughout the day, these balances are circulated more rapidly, allowing earlier settlement of payments and a greater divergence in the pattern of balances from the original RTGS balances.

The *skewness* measures the positive differences in a bank's balances from an RTGS system across the minutes of the day. This measure is obtained by calculating the skewness of the balance differences between RTGS and the simulated system. When compared with the netting systems, the receipt-reactive gross settlement system's distribution of the differences in balances is more positively skewed, suggesting that more participants maintain positive balances – and hence a reduced need for overdrafts on average. The receipt-reactive queue acts to smooth the liquidity usage across the payments system participants and across the day, generating larger balances and smaller overdrafts from the system's liquidity.

6.6 End-of-day queue analysis

When looking at the simulated receipt-reactive gross settlement system, it is important to determine how the end-of-day settling of payments influenced our results. The RRGs system closed at 5:30 p.m. and the remaining payments were evenly disbursed over the next thirty minutes. Table 6.5 gives the percentages of all payments submitted to the queue that settled by 5:30 p.m. In addition, the table presents the average value of payments in the queue at 5:30 p.m. to offer a sense of the role that size played in the settlement of these queued payments.

The one-hour netting arrangement quite naturally yields the best results here because the payments in the queue only start accumulating from 4:30 p.m. Because this time of day experiences the largest value payments (McAndrews and Rajan 2000), the average payment values are twice as high as those produced by six-hour netting and RRGs. The six-hour netting system shows that payments entered as early as 12:30 p.m. remain in the queue, which results in a low percentage of settled queued payments by 5:30 p.m. In terms of overall settlement by 5:30 p.m., the RRGs system performs similarly to a one-hour netting system at the 20 per cent and 50 per cent participation levels and more like a six-hour netting system at the 80 per cent level. This result indicates that when 80 per cent of payments are deferred to

RRGS, many never get settled and are returned to the banks that sent them.

Table 6.5 **Queue characteristics at 5:30 p.m.**

Treatment	Queued payments settled (per cent)		Average payment value in queue (thousands of dollars)
	Value	Volume	
One-hour netting			
20 per cent	73	91	10,234
50 per cent	69	90	10,653
80 per cent	73	91	10,063
Six-hour netting			
20 per cent	25	41	4,338
50 per cent	26	40	4,291
80 per cent	25	41	4,310
Receipt-reactive gross settlement			
20 per cent	64	73	4,533
50 per cent	72	80	4,890
80 per cent	28	48	5,184

Source: Authors' calculations, based on data from Federal Reserve Bank of New York, Fedwire Funds Service.

A feature of the simulated RRGs mechanism is that the FIFO rule was strictly adhered to when settling queued payments. The FIFO rule results in situations where large payments in front of the queue block smaller payments from being settled. The average value of queued payments at 5:30 p.m. was considerably higher for all levels of payments queued. The value of payments remaining in the queue rose as the participation in the queue increased. At the 80 per cent level, the average value of payments that remained queued was 50 per cent higher than the average value of payments in the original RTGS system (\$3.45 million). This result suggests that the performance of RRGs systems could likely be improved by using a different order criterion, such as allowing smaller payments to go first, splitting the large transactions into several smaller ones, or using gridlock-resolution mechanisms.

6.7 Discussion of simulation results and likely behavioral responses

Netting has long been acknowledged as an efficient way to reduce the liquidity needs of a payments system. Roberds (1993) shows that the netting of payments achieves the theoretical minimum of liquidity use for the settlement of a specific set of payments. Another way to reduce liquidity needs is to use various queuing arrangements in RTGS systems.

Many simulation studies of the behavior of queue-augmented RTGS systems have been conducted recently. The research examines the mechanical reordering of payments that is possible with a queuing system and then compares this feature with other alternative designs to see how each handles liquidity use and payment delays. Koponen and Soramäki (1998) and Leinonen and Soramäki (1999) use the simulation approach to measure the liquidity savings and settlement delays for a number of alternative queuing, netting, and payment-splitting techniques with an RTGS system using data from the Bank of Finland's RTGS system. Both studies conclude that queuing systems can contribute substantial improvements compared with pure netting systems. Bech and Soramäki (2001) analyse the effects of a simultaneous settlement algorithm based on FIFO to settle a subset of payments from queues.

A report on RTGS systems by the Bank for International Settlements (1997) discusses how to sequence payments so that incoming transfers from other banks can be used to fund payments. A pure RTGS system relies solely on the participants (in a decentralised way) to time their payments in a manner that best utilises their incoming funds.

Our simulation results show that netting in conjunction with RTGS settlement is not very desirable. Netting every six hours yields relatively modest but statistically insignificant reductions in daylight overdrafts (3.2 per cent) compared with a significant delay in payment timing (one hour and thirteen minutes in average settlement time). Compared with one-hour netting, the settlement delays in six-hour netting are much greater. The simulated RRGS mechanism outperforms the netting systems when both liquidity savings and payment delays are taken into account. The receipt-reactive system saves significant liquidity (13.6 per cent) and results in relatively modest payment delays (forty-three minutes) when compared with six-hour netting. These results are in line with those of the studies cited above.

6.7.1 Bank participation and risk

One question to ask is whether banks would actually use deferred payment mechanisms if they were available. Our simulations cannot predict banks' behavior. However, by relying on how the mechanisms work in a simulated environment and applying theoretical reasoning from the models of payment behavior, we can gain insight into the likely endogenous responses to implementing these complements to an RTGS system.

Consider the netting systems, for example. The simulation shows that a netting system that settles payments each hour of the day, even if it were to attract 50 per cent of the payments on Fedwire, would not create a statistically different level of daylight overdrafts. When we consider the *participation risk* in a netting system, it is unlikely that banks would use such a facility.

Participation risk is present in many netting systems. In the operation of most payments systems, if a bank enters payments into the system, the bank is expected and, in some cases, obligated to settle those payment orders. If the bank expects many offsetting payments to be entered, but in fact only a few are, the bank may face a larger-than-expected settlement obligation. Netting systems tend to be more effective in gathering offsetting payments as more banks participate. Hence, a start-up netting system exposes potential participants to risk. If many banks participate, the bank will have a small expected settlement obligation, but if only few others do, the bank may have a larger-than-expected settlement obligation.

If a payments system were to offer a netting arrangement in conjunction with an RTGS system, banks that choose to use the netting mechanism would face liquidity risks. Those banks that use the netting system may end the day needing more liquidity than if they had carefully managed their payments strictly through an RTGS system. RTGS system participants likely submit payments in ways meant to lessen their liquidity demands and risks endogenously. As a result, a netting service might fail to attract participation when offered in conjunction with the current Fedwire service.

Deferred settlement systems provide a means to allow participants in an RTGS system to make some payments *contingent* on the submission of offsetting payments by their counterparties. Allowing a participant's payment submission strategy to be contingent on the payment submissions of others reduces the risk of loss associated with making payments and the risk of having one's counterparties fail to make expected offsetting payments prior to defaulting. The reduction in that risk should lessen the incentive to delay submitting one's

payments, at least in the contingent, deferred settlement payment option.

We expect that banks would face lower participation risk in the receipt-reactive queuing system. The basic operation of the queue is useful to a single bank in isolation, even if no other banks use their options to queue payments. Queuing enables the bank to automatically synchronise outgoing payments with its incoming payments. This is an important advantage in encouraging use of the queue: Banks face no risk of increasing their liquidity demands by using the queue. They can always do at least as well for themselves by using the queue to manage their liquidity demands as they would if they had their own in-house automatic queue management system.

How would the banks change the timing of their payments to the system in the new environment? Models of the timing of payments in an RTGS model have been presented by Angelini (1998, 2000), Buckle and Campbell (2002), Kobayakawa (1997), and Kahn, McAndrews, and Roberds (2003). Some of these studies consider the default risk of counterparties and focus on the possibility of payment delays in an RTGS system, partly because of the risk of sending the gross amount of a payment to a counterparty in advance of receiving an offsetting amount from that counterparty. Because payments made in an RTGS system are not contingent on the submission of payments by counterparties, banks could be reluctant to submit payments in a timely fashion, and as a result, the timing of all payments could be delayed (relative to the time that a central planner would choose to have the payments sent). Buckle and Campbell (2002) consider requirements, chosen jointly by the participants, that commit participants to submit certain percentages of their payments by certain times of the day. Such a requirement is in effect in the U.K. payments system, CHAPS, and, on the European level, guidelines issued by the European Banking Federation govern the timing of payments in the TARGET system.

Recent theoretical work by Bech and Garratt (2003) analyses the incentives that banks have in an RTGS system – in the case in which there is no default risk among counterparties – to coordinate the timing of payments. In their model, banks wish to complete a set of payments while economising their holdings of overnight balances. Bech and Garratt find that banks are expected to synchronise the timing of their payment submissions to take best advantage of incoming transfers, which allows all the banks to economise on their holdings of overnight balances. McAndrews and Rajan (2000) present evidence consistent with that model for the Fedwire system.

The receipt-reactive queue mechanism can be a useful way for a bank to reduce its demand for costly liquidity. In fact, we would also expect banks to enter more of their day's activity earlier in the day, with much of it placed in the queue. We expect this to be the case because in most RTGS systems today, a bank's only method of liquidity management is simply to delay payments. In contrast, by using a receipt-reactive queue, banks would not face any downside risk to their liquidity position by the early entry of payment messages in the queue because the amount released from the queue would always be less than the incoming payment amount.

In the United States, some banks use the Clearing House Interbank Payments System, which has liquidity-saving features. The presence of such a system could serve to satisfy banks' demand for an alternative to RTGS. If that is so, then appending liquidity-saving features to the existing RTGS system may not lead to a great deal of use of the liquidity-saving features.

The primary reason why a receipt-reactive system would generate good incentives for the early submission of payments is that the release of payments from the queue is independent of the timing of a *bank's own* RTGS payments. Because an RTGS system does not rely on a bank's balance, but only a bank's receipts to trigger the release of payments, the history of a bank's submission of RTGS payments does not affect the release of its payments from the queue. As a result, there is no incentive to delay making RTGS payments to allow the release of queued payment messages. This situation results in incentives for earlier entry of RTGS payments, which would endogenously improve the circulation of liquidity, releasing the queued messages of others in a virtuous circle.⁶

Some degree of transparency of the queues might also offer banks information that they cannot gain by using internal queues. With the use of such features, we would tend to expect widespread participation of banks in queuing mechanisms. In this case, gridlock resolution – which requires a central queue – could be used to optimise further the queue's performance.

Finally, centralised queuing in general may be beneficial for smaller banks. While larger banks currently time the entry of their payments into RTGS systems with the aid of internal queues, the operation of the receipt-reactive queue would give more banks the option of automated payment settlement.

⁶ These and other behavioral responses of queuing systems are discussed in McAndrews and Trundle (2001) and Roberds (1999).

6.8 Conclusion

In this article, we simulate deferred settlement mechanisms to understand the liquidity implications of using the mechanisms to complement real-time gross settlement systems. Using historical data on all payments made over ten days on the Fedwire Funds Service, we simulate two different netting systems and a receipt-reactive gross settlement system.

We find that, unlike an RTGS system, both netting and RRGs queuing systems introduce delays to payments. However, both netting and queuing also have the potential to reduce – in some cases, significantly – daylight overdrafts. These results appear to be robust to alternative assumptions about the level at which banks are willing to submit payments to a queue for deferred settlement.

The receipt-reactive gross settlement system we examine is novel in that it releases payments from the queue based on a bank's receipts over a given time rather than on its balance. The simulations in this article indicate that an RRGs system reduces significantly more overdrafts than a six-hour netting system would, with considerably less delay in payments.

Our consideration of the receipt-reactive gross settlement system reveals that it may provide good incentives for banks to submit payments early to the queuing system, as the release of payments from the queue is independent of the submission of the bank's own RTGS payments. This feature is likely to encourage banks to quicken the timing of payments and to reduce the number of daylight overdrafts. As a result, such a system might prove to be a true liquidity-saving complement to an RTGS system.

While simulations provide a good starting point for studying enhancements to RTGS systems, our results suggest that these systems warrant further investigation. For example, how banks would change their behavior when offered these alternatives to payment settlement remains an open question. Going forward, a better theoretical and empirical understanding of banks' payment behavior would help inform policymakers considering enhancements to RTGS systems.

Appendix 6.1

Detailed dynamics of the receipt-reactive gross settlement system

In this appendix, we explain the features of our proposed receipt-reactive gross settlement design. A real-time gross settlement (RTGS) payment from bank i to bank j entered at time t is p_{ij}^t . Similarly, a queued payment message is $q_{ij}^{t'}(r)$, with t' denoting the time of entry in the queue and r denoting the message's rank in the order of its entry. A bank's funds balance at time t is denoted b_i^t . A settled payment is s_{ji}^t , where a settled payment is either an RTGS payment or a queued payment message that has been released at time t .

The basic receipt-reactive gross settlement design operates as follows: By the end of minute m , the payment messages $\{q_{ij}^{t'}(r) : j \neq i, t' < m, r = 1, 2, \dots, k\}$ are released from the queue, where k is the maximum rank that satisfies inequality (A6.1-1)

$$\sum_{(m-1) < t \leq m} \sum_{j \neq i} (s_{ji}^t) \geq \sum_{r=1}^k \sum_{j \neq i} (q_{ij}^{t'}(r)) \quad (\text{A6.1-1})$$

At the beginning of each minute, the rank of queued payments is reset so that the oldest queued payment is assigned rank one, the second-oldest, rank two, and so on. (The actual algorithm releases payments within the minute as soon as sufficient receipts arrive.)

Inequality (A6.1-1) states that the first k queued payment messages of the bank are released in minute m when the value of the bank's receipts in that same minute are greater than or equal to the value of the k payment messages to be released. A bank's balance at the end of minute m will then be equal to

$$b_i^m = b_i^{m-1} - \sum_{(m-1) < t \leq m} \sum_{j \neq i} (p_{ij}^t) + \left(\sum_{(m-1) < t \leq m} \sum_{j \neq i} (s_{ji}^t) - \sum_{r=1}^k \sum_{j \neq i} (q_{ij}^{t'}(r)) \right) \quad (\text{A6.1-2})$$

Equation (A6.1-2) states that a bank begins a minute with its balance of the previous minute, and its balance decreases in the minute by any RTGS payments it makes and increases by the net amount of its receipts, less any release of payments from its queue. By inequality (A6.1-1), the net amount of its receipts must be at least as large as the amount that is released from the queue. If a bank's queued payments were numerous and finely divisible – so that the receipts are approximately equal to the amount released from the queue – then the bank's balance would be approximately equal to its previous balance minus its outgoing RTGS payments: $b_t^m = b_t^{m-1} - \sum_{(m-1)<t \leq m} \sum_{j \neq i} (p_{ij}^t)$.

In a theoretical limit to the use of this queuing system, banks could place in the queue all payments whose value they expect to be offset by incoming payments. Banks' real-time payments would equal the amount of net payment outflows that they would expect during the day – equivalent to the multilateral net debit of a net settlement system. If expectations were fully realised and banks held sufficient balances to fund their payments, the amount of balances held would equal the amount of settlement payments that the banks would need if they settled payments in a multilateral net settlement system. At the same time, banks would have the advantages of real-time release of payments and the associated release of payments from the queue throughout the day. This is the essential theoretical benefit of this design.

No practical implementation of this system is likely to achieve the theoretical maximum in liquidity savings. In fact, the practical implementation of this queuing system is an important aspect of the mechanism. One important element is that the receipt-reactive system relies on some funds flowing among participants to trigger the release of queued payment messages. If all banks were to queue all their payments, all payments would remain queued. In such a case, the system operator could consider using a 'gridlock-resolution' mechanism to break the logjam and release some payments from the queue, as in Bech and Soramäki (2001, 2002).

Appendix 6.2

Simulation results and analysis

To gauge the effects of some of the liquidity enhancements described in our article, we tested various features of these possible enhancements using a simulation program developed by the Bank of Finland. The simulation program is a version of the one described in Leinonen and Soramäki (1999) and Koponen and Soramäki (1998). (A more in-depth description can be found at the Bank of Finland's website: <http://www.bof.fi/sc/bof-pss>.) The simulations first generated baseline output data for both real-time gross settlement (RTGS) and net settlement systems. Further simulations examined an alternative enhancement that releases gross payments against the aggregate amount of incoming funds within each minute – receipt-reactive gross settlement (RRGS).

We randomly selected 20 per cent, 50 per cent, or 80 per cent of all Fedwire Funds Service payments and placed those payments either into a queue for deferred settlement or into a netting system. Those payments not entered into the queue or netting system were settled by the RTGS system. With the exception of the changes in settlement time imposed by the queuing or netting arrangements, we assumed no behavioral changes in Fedwire that would affect the timing of payment entry. The bulk of our analysis was performed with ten days of data, directing 50 per cent of payments to the deferred settlement mechanisms. The 20 per cent and 80 per cent simulations tested the network effect associated with varying degrees of participation in these queuing and netting arrangements. The 80 per cent simulations were particularly time consuming, so the 20 per cent and 80 per cent sensitivity analyses were performed on only three days of data.

Because these simulations focus on reducing both the Federal Reserve's risk exposure in granting intraday credit and the liquidity use by banks, we selected statistics that assess these areas. The analysis involves average and peak overdrafts and settlement delays.

Data

The simulations were performed using ten typical days of funds transfer activity data from the Fedwire Funds Service (Table A6.2-1). The days were randomly selected from the period October 1999

through February 2000. We included all transaction types but eliminated payment transfers of less than \$100. The transaction data only included master accounts. Subordinate account numbers were changed to their related master account numbers before simulation. A uniform random-number generator was used to select the 50 per cent of payments to be queued in each treatment. The rest of the payments were automatically processed by the RTGS system according to their historical timing during the simulations. While not directly involved in the simulations, National Book-Entry System (NBES) securities transaction data were used for the overdraft analysis presented later.

Table A6.2-1 **Summary statistics for Fedwire Funds data**

Date	Number of payments	Sum of payments (billions of dollars)	Number of banks	Average value (millions of dollars)	Standard deviation of payment value (millions of dollars)	Sum of opening balances (billions of dollars)
1/6/2000	369,094	1,181.34	6,289	3.20	28.66	10.03
1/26/2000	373,685	1,318.44	6,191	3.53	31.84	15.61
2/23/2000	406,644	1,421.39	6,339	3.50	31.47	16.90
2/24/2000	404,356	1,470.65	6,293	3.64	32.55	16.42
10/1/1999	541,075	1,840.46	6,767	3.40	32.85	21.50
10/8/1999	406,628	1,377.19	6,400	3.39	31.23	12.76
11/4/1999	373,811	1,329.56	6,228	3.56	32.23	15.38
11/10/1999	395,304	1,354.38	6,159	3.43	31.42	10.87
12/6/1999	400,689	1,394.28	6,234	3.48	30.95	10.88
12/16/1999	413,024	1,383.84	6,295	3.35	30.53	12.71

Source: Authors' calculations, based on data from Federal Reserve Bank of New York, Fedwire Funds Service.

Analytical framework

In the following discussion, let s_{ij}^t represent a settlement of a payment order at time t from i to j . The balance at the end of any minute, m , is equal to the previous balance plus the difference between the cumulative value of outgoing and incoming payments: $i = 1, 2, \dots, D$.

$$b_i^m = b_i^{m-1} - \sum_{(m-1) < t \leq m} \sum_{j \neq i} (s_{ij}^t) + \sum_{(m-1) < t \leq m} \sum_{j \neq i} (s_{ji}^t) \quad (\text{A6.2-1})$$

where $m = 1, 2, \dots, M$.

Overdrafts

Overdraft calculations are based on end-of-minute balances. For overdraft accounting purposes, a bank's balance is affected by several services. We attempt to extract the Fedwire Funds Service's

contribution to the bank's balance, and hence its overdraft, from the bank's other non-Fedwire funds transactions using a method that mimics the Federal Reserve's overdraft accounting procedures.⁷ This involves a comparison of the Fedwire funds and NBES balances to determine the applicable balance, b_i^m , applied in the following formulas. The analysis assumes that same-day NBES and non-Fedwire funds activities remain unchanged (see Appendix 6.3 for further elaboration). Government-sponsored enterprises, the Clearing House Interbank Payments System, Federal Reserve System banks, and government agency accounts were included in the simulations, but removed prior to analysis.

The overdraft during the day for any minute, m , and bank, i , equals the absolute value of a negative balance or zero.

$$OD_i^m = \left| \min(0, b_i^m) \right| \quad (A6.2-2)$$

The average continuous overdraft is

$$\text{Average OD} = \frac{\sum_m \sum_i OD_i^m}{1,081} \quad (A6.2-3)$$

that is, the sum of overdrafts for each bank during the day divided by the number of minutes Fedwire is open (eighteen hours and one minute).⁸

Peak overdraft is

$$\text{PeakOD} = \max_m \left(0, \sum_i OD_i^m \right) \quad (A6.2-4)$$

⁷ These are activities unrelated to the daily Fedwire Funds Service and they are posted to the Daylight Overdraft Reporting and Pricing System at particular times. These include the Automated Clearing House network, checks, currency and coins, and savings bonds.

⁸ Effective 16 May 2004, the Federal Reserve Bank expanded the operating hours of the Fedwire Funds Service from eighteen hours to twenty-one-and-a-half hours. The new hours begin at 9:00 p.m. Eastern Time (ET) on the preceding calendar day (with a cycle date of the following calendar day) and end at 6:30 p.m. ET, regardless of the Bank's location or time zone.

Delay and time statistics

The settlement delay for each payment is calculated as the time difference between payment origination by the sending bank and the final and irrevocable settlement of the payment. The two statistics that we use to measure the delay imposed by the queuing and netting arrangements in our proposed design are the delay statistic and the average time of settlement.

The delay statistic for the system is calculated as

$$\text{Delay} = \frac{\sum_t \sum_i \sum_{j, j \neq i} s_{ij}^{t, t'} (t - t')}{\sum_t \sum_i \sum_{j, j \neq i} s_{ij}^{t, t'} (T - t')} \quad (\text{A6.2-5})$$

In the notation for the delay statistic, we capture, for each settled payment, both its entry time, denoted by t' , and its release time, denoted by t . The delay statistic measures in the numerator the value of the settled payments multiplied by the time they spent delayed in the queue. In the denominator, the value of settled payments is multiplied by the time that payments could have been queued, had their settlement been delayed until the queue is closed.

The delay statistic is a standardised indicator that may take values between zero and one. In an RTGS system, for example, payments spend no time in the queue, and $t' = t$, resulting in a delay statistic equal to zero. In an end-of-day netting system, the settlement of payments is delayed until end of day, and $t = T$, resulting in a delay statistic equal to one.

The delay statistic places greater emphasis than the average time of settlement on the settlement of both early-morning and early-afternoon payments, as well as on smaller payments. Early-morning or early-afternoon payments carry more weight than their nominal value because the delay statistic repeatedly counts these payments for every minute that they remain unsettled.

The average time of settlement (ATOS) is the average time weighted by the value of the payments settled at each minute, t .

$$\text{ATOS} = \frac{\sum_t \left(\sum_i \sum_{j \neq i} s_{ij}^t \right) \times t}{\sum_{i=0}^T \sum_i \sum_{j \neq i} s_{ij}^t} \quad (\text{A6.2-6})$$

These statistics are calculated for all applicable treatments. In addition, we ran a statistical test on the average continuous overdraft statistic to provide added confirmation of a scenario's overdraft savings or loss. Because the same payment data are used for each simulation treatment, the statistics for each treatment can be viewed as different variables of the same group. Furthermore, the small, ten-day sample size and unequal variances across treatments suggest that a standard parametric statistical analysis is inappropriate. Therefore, we use the nonparametric Wilcoxon matched-pairs signed-rank test to determine whether a statistical difference in average continuous overdraft exists between the dependent groups or treatments. All treatments are compared with the baseline performance of the RTGS system. When a treatment shows a nominal increase or decrease in average continuous overdraft when compared with each of these two treatments, we conduct a one-sided test to determine whether a statistically significant increase or decrease can be found versus the null hypothesis that there is no statistical difference. All tests were conducted at the 5 per cent level.

Treatments and results

To properly compare aggregate statistics from the sensitivity analysis with the 50 per cent simulations, we added two lines for the RTGS simulation results, as well as for all 50 per cent simulations. The sensitivity analysis average and sensitivity analysis standard deviation figures represent aggregate statistics calculated using the three days in our 20 per cent and 80 per cent sensitivity analysis simulations: January 6, 2000; January 26, 2000; and November 4, 1999.

Real-time gross settlement

The RTGS simulation provides a benchmark for the analysis of the alternative queuing and netting arrangements (Table A6.2-2). In this simulation, 100 per cent of the payments are settled immediately by the RTGS system. The delay statistic for RTGS is zero by definition and the average time of settlement equals the time when an average dollar was submitted to the system.

Table A6.2-2

Real-time gross settlement simulations

Date	Average overdraft (billions of dollars)	Peak overdraft (billions of dollars)	Time of peak overdraft	Average time of settlement	Delay statistic (per cent)
1/6/2000	19.87	65.54	15:31	14:28	0
1/26/2000	20.57	70.73	15:49	14:32	0
2/23/2000	19.51	68.60	15:54	14:31	0
2/24/2000	21.01	67.58	15:55	14:41	0
10/1/1999	20.22	67.58	14:23	14:39	0
10/8/1999	19.69	59.68	14:01	14:39	0
11/4/1999	21.14	69.20	14:38	14:37	0
11/10/1999	22.91	75.58	15:54	14:35	0
12/6/1999	19.67	59.10	15:47	14:43	0
12/16/1999	18.28	59.20	15:53	14:33	0
Average	20.29	66.28	15:22	14:35	0
Standard deviation	1.23	5.47	0:44	0:04	0
Average for three days	20.52	68.49	15:19	14:32	0
Standard deviation for three days	0.64	2.67	0:36	0:04	0

Source: Authors' calculations, based on data from Federal Reserve Bank of New York, Fedwire Funds Service.

Note: The figures for the average for three days and the standard deviation for three days are calculated for the days for which the sensitivity analysis simulations were conducted.

One-hour net settlement

Net settlement of queued payments occurred every hour, and the net amounts were settled immediately thereafter through the RTGS system (Table A6.2-3). Accounts had unlimited liquidity available. At 18:30, remaining queued transfers were netted and the net balances were transferred between banks.

One-hour net settlement actually increases the average overdraft by 0.6 per cent, although the difference is not statistically different from the RTGS overdraft. One-hour net settlement, like six-hour net settlement, lowers the peak overdraft slightly while generating a 7.5 percentage point addition in the delay indicator. As expected, the average time of settlement is delayed by twenty-one minutes when compared with the RTGS average time of settlement – a logical result of delaying 50 per cent of the payments for up to an hour

Table A6.2-3

One-hour net settlement simulations with alternate levels of participation

Date	Average overdraft (billions of dollars)	Percentage change from real-time gross settlement	Peak overdraft (billions of dollars)	Percentage change from real-time gross settlement	Time of peak overdraft	Average time of settlement	Delay statistic (per cent)
20 per cent							
1/6/2000	19.86	0.0	63.78	-2.7	15:31	14:33	2.49
1/26/2000	20.24	-1.6	70.62	-0.2	14:55	14:36	2.26
11/4/1999	21.01	-0.6	67.09	-3.0	14:38	14:43	2.82
Average	20.37	-0.7	67.16	-1.9	15:01	14:37	2.52
Standard deviation	0.58	0.8	3.42	1.6	00:27	0:04	0.28
50 per cent							
1/6/2000	20.02	0.8	59.08	-9.9	15:31	14:43	7.24
1/26/2000	20.36	-1.0	68.94	-2.5	14:53	14:46	6.86
2/23/2000	19.48	-0.2	66.53	-3.0	15:54	14:46	7.03
2/24/2000	21.36	1.6	68.30	1.1	16:10	14:55	7.38
10/1/1999	20.27	0.2	64.18	-5.0	15:10	14:54	7.72
10/8/1999	19.96	1.4	59.80	0.2	14:01	14:54	7.72
11/4/1999	21.27	0.6	67.80	-2.0	15:23	14:52	7.84
11/10/1999	23.49	2.5	76.83	1.6	16:31	14:50	7.54
12/6/1999	19.65	-0.1	57.01	-3.5	16:07	14:58	8.00
12/16/1999	18.21	-0.4	58.43	-1.3	16:22	14:49	7.97
Average	20.41	0.6	64.69	-2.4	15:36	14:51	7.53
Standard deviation	1.41	1.1	6.19	3.3	0:46	0:04	0.39
Average for three days	20.55	0.1	65.27	-4.7	15:15	14:47	7.31
Standard deviation for three days	0.64	1.0	5.39	4.4	0:20	0:04	0.49
80 per cent							
1/6/2000	19.06	-4.1	56.36	-14.0	15:31	14:49	9.94
1/26/2000	19.77	-3.9	69.67	-1.5	15:49	14:51	9.31
11/4/1999	20.66	-2.3	67.93	-1.8	15:31	14:59	10.55
Average	19.83	-3.4	64.66	-5.6	15:37	14:53	9.93
Standard deviation	0.80	1.0	7.24	7.1	00:10	0:05	0.62

Source: Authors' calculations, based on data from Federal Reserve Bank of New York, Fedwire Funds Service.

Note: The figures for the average for three days and the standard deviation for three days are calculated for the days for which the sensitivity analysis simulations were conducted.

Six-hour net settlement

Net settlement of queued payments occurred every six hours (6:30, 12:30, and 18:30), and the net amounts were settled immediately thereafter through the RTGS system (Table A6.2-4). Accounts had unlimited liquidity available. At 18:30, remaining queued transfers were netted and the net balances were transferred between banks. The average and standard deviation figures represent those statistics for the three days involved in our 20 per cent and 80 per cent simulations: January 6, 2000; January 26, 2000; and November 4, 1999.

Six-hour net settlement produced a modest 3.2 per cent reduction in average overdraft, although, once again, the difference is not

statistically significant. However, this treatment produced the lowest overall peak overdraft. The 35 per cent delay statistic is roughly five times that of one-hour net settlement and, as we will see, is nearly three times that of RRGs. The average time of settlement was one hour and thirteen minutes later than in RTGS.

Table A6.2-4 **Six-hour net settlement simulations with alternate levels of participation**

Date	Average overdraft (billions of dollars)	Percentage change from real-time gross settlement	Peak overdraft (billions of dollars)	Percentage change from real-time gross settlement	Time of peak overdraft	Average time of settlement	Delay statistic (per cent)
20 per cent							
1/6/2000	20.53	3.4	63.55	-3.0	15:31	14:56	13.30
1/26/2000	19.63	-4.6	65.27	-7.7	14:55	15:01	14.04
11/4/1999	21.43	1.4	64.44	-6.9	14:38	15:07	15.59
Average	20.53	0.0	64.42	-5.9	15:01	15:01	13.97
Standard deviation	0.90	4.1	0.86	2.5	00:27	0:05	0.65
50 per cent							
1/6/2000	20.47	3.0	69.99	6.8	15:39	15:42	34.56
1/26/2000	19.27	-6.3	60.67	-14.2	14:53	15:44	34.70
2/23/2000	19.58	0.4	65.72	-4.2	15:54	15:42	33.56
2/24/2000	18.83	-10.4	58.67	-13.2	16:45	15:48	33.64
10/1/1999	17.83	-11.8	59.59	-11.8	14:25	15:45	32.83
10/8/1999	20.14	2.3	56.63	-5.1	16:14	15:51	35.88
11/4/1999	19.88	-5.9	56.85	-17.8	16:36	15:49	35.41
11/10/1999	22.92	0.1	70.41	-6.9	16:31	15:45	34.00
12/6/1999	19.43	-1.2	55.01	-6.9	16:06	15:51	34.56
12/16/1999	16.19	-11.5	47.92	-19.1	15:53	15:44	34.38
Average	19.45	-4.1	60.15	-9.3	15:53	15:46	34.35
Standard deviation	1.75	5.8	6.95	7.7	0:44	0:03	0.90
Average for three days	19.87	-3.2	62.50	-8.7	15:42	15:45	34.89
Standard deviation for three days	0.60	5.3	6.76	13.3	0:51	0:03	0.46
80 per cent							
1/6/2000	18.15	-8.7	51.37	-21.6	14:23	16:25	54.70
1/26/2000	16.20	-21.2	50.26	-28.9	15:57	16:29	55.94
11/4/1999	17.62	-16.6	50.67	-26.8	16:06	16:31	55.67
Average	17.32	-15.6	50.77	-25.9	15:29	16:28	55.44
Standard deviation	1.01	6.4	0.56	3.8	00:57	0:03	0.65

Source: Authors' calculations, based on data from Federal Reserve Bank of New York, Fedwire Funds Service.

Note: The figures for the average for three days and the standard deviation for three days are calculated for the days for which the sensitivity analysis simulations were conducted.

Receipt-reactive gross settlement

The gross amount of payments received during each minute provided the available liquidity for release from the queue (Table A6.2-5). The payments subject to deferral were held in the queue if they were not offset by incoming payments in that minute. The available liquidity

from incoming payments resets to zero at the start of a new minute and does not accumulate past that minute. Payments settled on a first in, first out (FIFO) basis when a bank received sufficient incoming funds. Queued payment messages were transferred to RTGS for settlement. The nondeferred payments were settled immediately by the RTGS system and did not affect the incoming funds accounting. Starting at 17:30, the remaining queued payments were spread evenly over the next thirty minutes and settled by the RTGS system.

In sum, receipt-reactive gross settlement offers a significant overdraft reduction coupled with an increase in the peak overdraft. The RRGs produced statistically significant lower average continuous overdrafts than did the RTGS treatment.

Table A6.2-5 Receipt-reactive gross settlement simulations with alternate levels of participation

Date	Average overdraft (billions of dollars)	Percentage change from real-time gross settlement	Peak overdraft (billions of dollars)	Percentage change from real-time gross settlement	Time of peak overdraft	Average time of settlement	Delay statistic (per cent)
20 per cent							
1/6/2000	21.41	7.8	76.52	16.8	15:45	14:44	3.51
1/26/2000	20.20	-1.8	67.91	-4.0	14:55	14:46	3.01
11/4/1999	21.04	-0.5	69.98	1.1	15:24	14:55	3.53
Average	20.88	1.7	71.47	4.4	15:21	14:48	3.35
Standard deviation	0.62	5.2	4.49	10.8	00:25	0:05	0.30
50 per cent							
1/6/2000	19.18	-3.4	78.12	19.2	16:52	15:19	15.61
1/26/2000	15.06	-26.8	60.52	-14.4	16:48	15:13	12.32
2/23/2000	15.90	-18.5	68.51	-0.1	16:38	15:15	13.38
2/24/2000	17.24	-18.0	70.92	4.9	16:53	15:18	12.81
10/1/1999	14.18	-29.9	64.44	-4.6	16:48	15:22	14.06
10/8/1999	19.05	-3.2	76.50	28.2	16:46	15:22	14.78
11/4/1999	17.89	-15.4	72.58	4.9	16:51	15:24	16.42
11/10/1999	21.97	-4.1	89.96	19.0	16:23	15:17	13.56
12/6/1999	20.82	5.8	78.78	33.3	16:50	15:18	11.03
12/16/1999	13.89	-24.0	66.23	11.9	16:55	15:14	13.42
Average	17.52	-13.6	72.66	9.6	16:46	15:18	13.74
Standard deviation	2.77	11.9	8.55	14.9	0:09	0:03	1.58
Average for three days	17.38	-15.3	70.40	2.8	16:50	15:18	14.78
Standard deviation for three days	2.11	11.7	9.00	16.9	0:02	0:05	2.17
80 per cent							
1/6/2000	12.59	-36.7	67.76	3.4	17:12	16:29	19.02
1/26/2000	8.61	-58.1	45.93	-35.1	17:05	16:29	23.11
11/4/1999	13.28	-37.2	66.07	-4.5	17:02	16:29	25.35
Average	11.49	-44.0	59.92	-12.5	17:06	16:29	22.50
Standard deviation	2.52	12.3	12.15	20.3	00:05	0:00	3.21

Source: Authors' calculations, based on data from Federal Reserve Bank of New York, Fedwire Funds Service.

Note: The figures for the average for three days and the standard deviation for three days are calculated for the days for which the sensitivity analysis simulations were conducted.

Indicators of liquidity use

Overdraft calculations use only the negative balances and therefore measure only the lower end distribution of balances according to a cutoff point of zero. While this is useful from a risk management standpoint, we calculate further statistics regarding the performance of these systems by analysing them in terms of their liquidity usage. Each simulation restructures the timing of the original RTGS payments by changing the release sequence of the payments. By doing so, the different systems generate different liquidity levels and balance distributions than the original RTGS system. When evaluating the different simulations' temporal restructuring of the original RTGS payments, the approaches presented here attempt to indicate efficiency of liquidity use both above and below zero, providing a better overall view of each simulation's effect on liquidity usage. We do this indirectly by measuring both the degree of difference in the simulation balances when compared with RTGS balances and the direction of that difference.

Our first liquidity calculation measures the average absolute change in balances that occurs per minute for each bank. In effect, this calculation gives the amount of money that an average bank would have to move, either in or out of its account, for any given minute of the day. A liquidity-usage measure calculates the extent to which the balance must fluctuate in order to settle the payments. An RTGS system requires the most liquidity.

To measure liquidity, we first calculated the absolute value of a bank's change in balance from one minute to the next and summed this amount across all banks for a given minute. We did this for all 1,080 difference periods, summed the results, and divided by $1,080 \times N$, with N being the number of banks. The b_i^m used in the liquidity equations is the sum of each bank's end-of-minute balance across the three days subject to the alternative levels of payment submission simulations

$$\text{Average funds transfer} = \frac{1}{1,080 \times N} \sum_{m=2}^{1,081} \sum_{i=1}^N |b_i^m - b_i^{m-1}| \quad \forall (b_i^m - b_i^{m-1}) < 0$$

Our second measure is the maximum funds transfer across the minutes

$$\frac{1}{N} \max_m \left(\sum_{i=1}^N |b_i^m - b_i^{m-1}| \right) \quad \forall (b_i^m - b_i^{m-1}) < 0$$

Our third measure of liquidity use in the systems is an indirect one. We calculated the degree of difference in the balances by measuring the independence exhibited between the end-of-minute balances of RTGS and the simulated system. Our independence calculation produced Pearson correlation coefficients for each minute and summed them across the 1,081 minutes of the Fedwire day. We then divided by 1,081 to get the average per minute. For RTGS, the correlation of its end-of-minute balances with itself is one. For the simulated systems, this number is the correlation between end-of-minute balances in the simulated system and those in the RTGS system. Lower numbers demonstrate more independence from the original RTGS payment distribution. The $b_{i,SIM}^m$ and $b_{i,RTGS}^m$, respectively, represent the collection of the pertinent simulated balances and the RTGS balances in minute m . The balances included in this collection are the sum of each bank's end-of-minute balance across the three days subject to the alternative levels of payment submission simulations.

$$\text{Correlation with RTGS} = \frac{1}{1,081} \sum_{m=1}^{1,081} \text{corr}_m(b_{i,SIM}^m - b_{i,RTGS}^m)$$

Our fourth measure, the skewness of the difference in balances, gauges the degree of positive change in balances imposed by the simulated system on the original RTGS payments. The difference in balances between the simulated system and RTGS is calculated for each bank and minute. A skewness statistic is generated for each minute and the average for the day is then calculated. The $b_{i,SIM}^m$ and $b_{i,RTGS}^m$, respectively, represent the sum of each bank's simulated and RTGS end-of-minute balances across the three days subject to the alternative levels of payment submission simulations.

$$\text{Skewness} = \frac{1}{1,081} \sum_{m=1}^{1,081} \text{skewness}_m(b_{i,SIM}^m - b_{i,RTGS}^m)$$

Appendix 6.3

Funds overdraft accounting procedure

An explanation of daylight overdrafts can be found in Coleman (2002). When determining a bank's balance, the Federal Reserve's Daylight Overdraft Reporting and Pricing System (DORPS) accounts for several funds credits and debits that we could not observe. These funds are not processed on the Fedwire Funds Service and are posted to the DORPS system through other means. We call these postings 'extraneous funds' and include such funds primarily from checks, the Automated Clearing House network, return checks, currency and coins, savings bonds, and account deficiency credits and debits. Extraneous funds play a major role in a bank's balance management. To circumvent this problem, we had to employ the following method:

We had access to the following information: the DORPS total end-of-minute balances, National Book-Entry System (NBES) transaction data, and Fedwire funds data. Since the DORPS balance data contain the opening balance for Fedwire funds and the NBES has an opening balance of zero, end-of-minute balances were constructed from the transaction data. The following formulas describe the end-of-minute balance situation for each bank at a particular minute m

$$\text{Totalbal}_m = \text{NBESbal}_m + \text{FedwireFundsbal}_m + \text{XFundsbal}_m \quad (\text{A6.3-1})$$

The extraneous funds balance was extracted according to the following formula

$$\text{XFundsbal}_m = \text{Totalbal}_m - \text{NBESbal}_m - \text{FedwireFundsbal}_m \quad (\text{A6.3-2})$$

The RTGS extraneous funds and NBES balances were then held constant. The Bank of Finland simulator used the opening balance to create new Fedwire funds balances, $\text{FedwireFundsbal}_m^*$. The Fedwire funds balance was then constructed using the RTGS extraneous funds balance

$$\text{Fundsbal}_m = \text{FedwireFundsbal}_m^* + \text{XFundsbal}_m \quad (\text{A6.3-3})$$

The new funds balance, in conjunction with the RTGS book-entry securities balance, resulted in a new total balance

$$\text{Totalbal}_m^* = \text{NBESbal}_m + \text{Fundsbal}_m \quad (\text{A6.3-4})$$

We then compared the three components in equation (A6.3-4) to determine each bank's applicable funds overdraft, FundsOD_m , according to the following DORPS accounting principles

If $\text{Fundsbal}_m \geq 0$, then $\text{FundsOD}_m = 0$.

If $\text{Fundsbal}_m < 0$ and $\text{NBESbal}_m < 0$, then $\text{FundsOD}_m = \text{Fundsbal}_m$.

If $\text{Fundsbal}_m < 0$ and $\text{NBESbal}_m \geq 0$:

If $\text{Totalbal}_m^* \leq 0$, then $\text{FundsOD}_m = \text{Totalbal}_m^*$.

Otherwise, $\text{FundsOD}_m = 0$.

$\text{FundsOD}_m = \text{OD}_i^m$ in appendix 6.2 for each bank i .

References

- Malli: Berger, A N and Smith, D C (2003) **Global Integration in the Banking Industry**. Federal Reserve Bulletin November.
- Angelini, P (1998) **An Analysis of Competitive Externalities in Gross Settlement Systems**. Journal of Banking and Finance, Vol. 22, No. 1, 1–18.
- Angelini, P (2000) **Are Banks Risk Averse? Intraday Timing of Operations in the Interbank Market**. Journal of Money, Credit, and Banking, Vol. 32, No. 1, 54–73.
- Bank for International Settlements (1997) **Real-Time Gross Settlement Systems**. Committee on Payment and Settlement Systems Publications, No. 22.
- Bech, M and Garratt, R (2003) **The Intraday Liquidity Management Game**. Journal of Economic Theory, Vol. 109, No. 2, 198–219.
- Bech, M and Soramäki, K (2001) **Gridlock Resolution in Interbank Payment Systems**. Bank of Finland Discussion Papers, No. 9.
- Bech, M and Soramäki, K (2002) **Liquidity, Gridlocks, and Bank Failures in Large Value Payment Systems**. E-money and Payment Systems Review. London: Central Banking Publications.
- Board of Governors of the Federal Reserve System (2002) **Guide to the Federal Reserve's Payment System Risk Policy**. May 2002. Washington, DC.
- Buckle, S and Campbell, E (2002) **Settlement Bank Behaviour in a Real-Time Gross Settlement Payment System with Collateralised Intraday Credit**. Bank of England Working Papers, No. 209.
- Coleman, S P (2002) **The Evolution of the Federal Reserve's Intraday Credit Policies**. Federal Reserve Bulletin.

- Kahn, C, McAndrews, J J and Roberds, W (2003) **Settlement Risk under Gross and Net Settlement.** Journal of Money, Credit, and Banking 35, No. 4, 591–608.
- Kobayakawa, S (1997) **The Comparative Analysis of Settlement Systems.** Centre for Economic Policy Research Discussion Papers, No. 1667.
- Koponen, R and Soramäki, K (1998) **Intraday Liquidity Needs in a Modern Interbank Payment System – A Simulation Approach.** Bank of Finland Studies in Economics and Finance, series E, No. 14.
- Leinonen, H and Soramäki, K (1999) **Optimising Liquidity Usage and Settlement Speed in Payment Systems.** Bank of Finland Discussion Papers, No. 16.
- Leinonen, H and Soramäki, K (2003) **Simulating Interbank Payment and Securities Settlement Mechanisms with the BoF-PSS2 Simulator.** Bank of Finland Discussion Papers, No. 23.
- McAndrews, J J and Rajan, S (2000) **The Timing and Funding of Fedwire Funds Transfers.** Federal Reserve Bank of New York Economic Policy Review 6, No. 2, 17–32.
- McAndrews, J J and Trundle, J (2001) **New Payment System Design: Causes and Consequences.** Bank of England Financial Stability Review, No. 11, 127–136.
- Roberds, W (1993) **The Rise of Electronic Payments Networks and the Future Role of the Fed with Regard to Payment Finality.** Federal Reserve Bank of Atlanta Economic Review 78, No. 2, 1–22.
- Roberds, W (1999) **The Incentive Effects of Settlement Systems: A Comparison of Gross Settlement, Net Settlement, and Gross Settlement with Queueing.** Unpublished paper, Federal Reserve Bank of Atlanta, August.

Chapter 7

Simulation of liquidity levels and delays in the Swedish RIX system

Johan Pettersson

7	Simulation of liquidity levels and delays in the Swedish RIX sytem.....	218
	Abstract.....	218
7.1	Introduction.....	218
7.2	The RIX system	219
7.3	System qualities	220
7.4	Conclusions.....	225
	References	227

7 Simulation of liquidity levels and delays in the Swedish RIX system

Abstract

This article discusses certain efficiency aspects of RIX, the Swedish system for large-value payments.¹ The aspects addressed are high levels of liquidity and heavy demands on technical capacity during peak periods. The simulations presented in this article show that liquidity needs could be reduced significantly if the system was redesigned. Furthermore, a more even flow of payments would reduce technical queues during peak periods; this could be done through a change in throughput guidelines.

7.1 Introduction

Central banks pay more and more attention to efficiency in large value payment systems. In RIX banks borrow considerable amounts during the day from the Riksbank to make their payments. As these loans are granted against collateral the banks must pledge large quantities of securities, often government bonds or other securities with a high credit rating. Holding securities is costly for the banks, which could entail an efficiency problem. Another concern is that the system experiences overload during peak periods. In my work with RIX , initial analysis indicated that a different setup could improve the system in several aspects. However, the complex flow of transactions and the large amount of data have been difficult to analyse without simulation tools. For the following analysis the Bank of Finland kindly let me make use of their payment system simulator (BoFPSS2). The following simulations aim at guiding the future developments in the of Swedish payment system.

¹ This article is a deeper study of the simulations that where presented in Sveriges Riksbanks report 2003:2.

7.2 The RIX system

In 1984–1985, the Riksbank implemented its first on-line accounting system. This was an integrated accounting and payment system; banks with accounts at the Riksbank were allowed to report their clearing and interbank transactions electronically to the Riksbank. The system had both deferred net settlement and real-time gross settlement functions.

The RIX system was introduced in 1990. It is an on-line real-time gross settlement system (RTGS) system integrated with the accounting system. RIX was further developed in 1997–1998 with the implementation of a queuing system and SWIFT communication (SWIFT FinCopy) which allows for straight through processing. In 1999, the Riksbank implemented E-RIX, which is linked to the ESCB's TARGET system. RIX thus consists of two parallel but separate systems: K-RIX for settlement in Swedish kronor and E-RIX for settlement in euro. This article focuses only on efficiency problems in K-RIX, henceforth simply referred to as RIX.

The majority of the participants in RIX are credit institutions. There are 18 participants, of which 7 are Swedish banks and 5 are branches of foreign banks. The other participants are the three Swedish clearing organisations, the Swedish National Debt Office, CLS bank, and the Riksbank.

Investment firms and remote access participants may be accepted as participants in RIX. A Norwegian credit institution was accepted as a remote access participant at the beginning of 2003. The Riksbank may also accept foreign central banks as participants in RIX.

During 2002, the number of transactions in RIX averaged around 90 thousands a month, with a monthly turnover of about SEK 9,500 billion (EUR 1,041 billion). About 80% of transactions were interbank payments, and the remainder being customer payments.

Communication in RIX is handled by SWIFT. The RIX system uses the SWIFT service SWIFT FinCopy. Participants can keep track of their payments and the status of their Riksbank accounts via an on-line connection (RIX on-line). Information provided for participants includes the balance of accounts, the amount of liquidity reserved for certain types of payments, payments already processed, and the queue of incoming and outgoing payments. Each participant's information is limited to its own account(s) and payments.

A credit facility can be attached to the accounts of participating credit institutions, securities firms or government agencies. The Riksbank grants both intraday and overnight credits in Swedish

kronor. Full collateral is required for both intraday and overnight borrowing. Also remote participants are allowed to have an intraday facility in RIX.

The fee structure in RIX consists of a transaction fee, which is the same for all participants, and an annual fee based on the volume of sent transactions.

The RIX system has been further developed through the inclusion of the Swedish krona in CLS in 2003, the migration to SWIFTNetFIN and the implementation of the updated securities settlement system VPC NewClear in 2004.

Since Sweden is not part of the Euro-area, RIX has to be developed independently of target 2. In 2001, the Riksbank and the major participants in the system established a working group to outline a strategy for a future central settlement system. The working group identified a number of alternative strategies and compiled a list of participants' requirements.

7.3 System qualities

As mentioned above the RIX system is based on the RTGS principle. Payments are settled according to the first-in first-out (FIFO) rule, with the exception of CLS payments which can be prioritised over other payments. If there is a shortage of liquidity in a participant's account the payments are queued until the participant receives extra liquidity from incoming payments or increased intraday credits. Once released from the queue, payments are also settled according to FIFO.

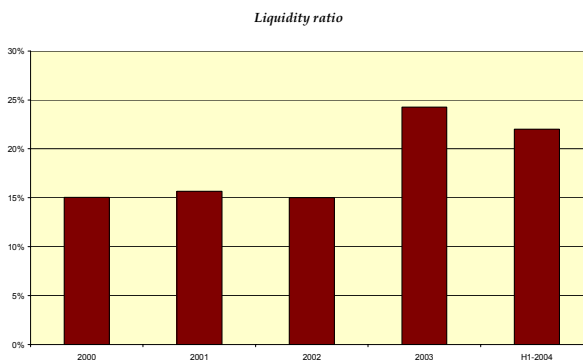
Intraday liquidity against collateral is provided by the Riksbank. Participants are allowed to borrow up to an amount corresponding to the value of the collateral pledged, minus a haircut. An automatic system for pledging allows participants to pledge and release various types of securities in real time, and the liquidity can thereby be adjusted on demand in situations of liquidity shortage.

During 2000–2002, the liquidity ratio for the participants in RIX was 15%. In 2003, the ratio increased to almost 25%, which means that the average amount pledged rose by 40 billion Swedish kronor.²

² Liquidity ratio is defined here as pledged amount / total value settled during the day.

Figure 7.1

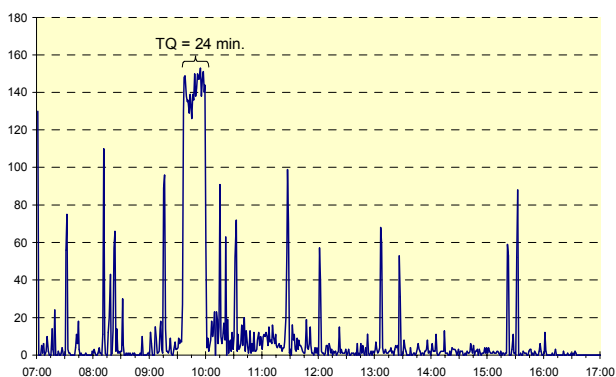
Liquidity ratio in RIX (2000-H12004)



The participating banks in RIX have a schedule for the input of payments that creates an offsetting effect, as it increases the likelihood of meeting payments. However, the schedule concentrates a large volume of payments in a short period of time, which creates a technical queue, due to lack of capacity. This creates a situation where the banks are unable to send time-critical payments and manage their liquidity position during the most critical minutes of the day. Figure 7.2 shows that the technical queue for a representative day is 24 minutes. RIX’s capacity to process payments between 140 and 160 payments a minute, depending on the size of the message attached.

Figure 7.2

Transactions per minute in RIX on a representative day in December 2003



Problem formulation

Liquidity

RTGS mitigates risks, since it eliminates credit exposures between participants. However, the need for liquidity is higher than with netting systems, which creates additional costs for participants.³ In response to this, some countries have systems that provide both RTGS and continuous net settlement (CNS) or offsetting. In Germany, RTGS and offsetting are provided by RTGS+, and in the United States, CHIPS provides CNS and Fedwire provides RTGS. Offsetting and continuous net settlement reduce the need for liquidity without reintroducing the setbacks of netting systems, such as credit risk exposures.

High levels of liquidity raise the cost of liquidity for participants in RIX, since they have to increase their level of collateral. Participants' holdings of collateral can be financed either through borrowing or by refraining from making alternative investments in order to be able to buy eligible collateral. In both cases there is an opportunity cost for the collateral. If a bank chooses to borrow, the cost is the difference between the borrowing rate and the risk-adjusted return on the collateral. If the bank instead refrains from making alternative investments, the cost will comprise the difference between the return on the collateral and the risk-adjusted return on an alternative investment.

Simulating RIX payment flows in a system environment that enables continuous net settlement could give valuable information on the opportunities to save liquidity for participants.

Capacity

In May 2003, the Riksbank increased the technical capacity of RIX, which reduced the technical queue. However, continuously adapting the system to cope with higher payment peaks is costly. However, the system environment has recently changed in a number of respects; one change is that the Swedish krona has been introduced into CLS. This has reduced the amount of transactions during peak hours and

³ For a discussion on the cost of liquidity usage and delays in payment systems, see Leinonen and Soramäki (1999).

increased the importance of a reliable real-time function, since CLS payments are time-critical.

On the other hand, the introduction of SWIFTNetFIN increases the length of the technical queue. This is an improvement in the messaging part of the system, but requires more capacity. In addition the number of payments in RIX is constantly increasing (from 1,500 payments a day in 1999 to today's figure of 5,100).

Hence, it is to soon to say if the increased capacity is enough to give a lasting decrease of the technical queue. Simulating RIX payments flow in a system environment without capacity constraint may give valuable information for assessing whether to invest in further capacity.

Simulation summary

I have used the Bank of Finland's simulator (BoFPSS2) to simulate historical transactions with different system setups. The basis for the analysis has been payment settlement data from a representative day. The trade-off between liquidity requirement and the system efficiency in processing payments is captured by the average settlement time \bar{t} , defined as;

$$\bar{t} = \frac{\sum_{t=0}^{t=T} V_t}{\sum V_t} t \quad (7.1)$$

where V is value settled at settlement time \bar{t} . Naturally, a system with a higher degree of gross settlement tends to have a lower \bar{t} .

The parameters I have changed in the simulations are (1) the form of settlement, ie real-time settlement or settlement with offsetting, and (2) credit restrictions on the participants. The aim was to study how fast payments flow through a system with offsetting functionality, and how liquidity restrictions affect this flow.

In the simulations I used all transactions that passed through the RIX system on September 9, 2003, without alteration.

The following scenarios where simulated:

1. Real-time settlement: In this simulation the simulator's RTGS system setup was used. Payments are queued if there is not enough liquidity. The real-time settlement simulation has two scenarios with different levels of credit. (1) In the first scenario (RTGSP) all

participants have access to unlimited credit. This scenario provides a benchmark for the upper limit of liquidity usage and a benchmark for the lower limit of average settlement time, since RTGS with unlimited credit gives immediate settlement at maximum liquidity usage. (2) In the second scenario (RTGSSB), the credit limit is set individually for each participant at an amount corresponding to its largest payment during the day. This simulation aims at showing how much the average transaction is delayed when lowering the liquidity in the system but retaining RTGS settlement.

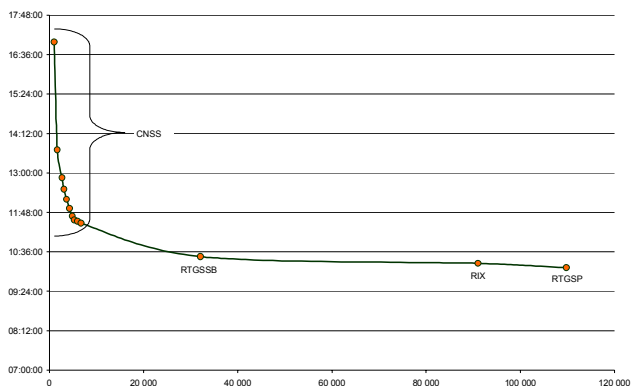
2. *Continuous net settlement:* In these simulations a CNS system setup was used. A payment is settled with finality when the participant has enough liquidity or if the payment can be settled through offsetting with an opposite payment. If neither of these conditions are fulfilled the payments are queued. The simulation was performed with ten different scenarios in which the credit limits per participant varied from 100 to 1000 million kronor. The purpose of the simulation was to give an indication of what liquidity levels could be expected if the RTGS algorithm in RIX was exchanged with an offsetting algorithm.

Simulation results

Figure 7.3 shows liquidity levels and average settlement times for the two RTGS scenarios, the ten CNS scenarios and the actual RIX settlement on September 9, 2003. The figure shows that the RTGSP scenario resulted in a liquidity requirement of SEK 110 billion, or around 23% of the total payment value. This is slightly more than the actual liquidity that was used in RIX that day, which was SEK 91 billion. The lower level of liquidity came at a cost of queuing some of the payments, which is shown by a slight upward shift in the average settlement time (5 minutes). In the RTGSSB scenario the liquidity requirement was as low as SEK 32 billion. However, the lower level of liquidity created liquidity queues and the average settlement time increased 19 minutes. If this delay were acceptable to the participants, the liquidity levels in RIX could be lowered by changing the code of conduct.

Figure 7.3

Simulation results



The CNS scenarios (CNSS) show that the liquidity requirement can be considerably reduced, but at the cost of delay that increases exponentially as liquidity decreases. However, the delay in the CLSS with a credit limit of 800–1000 million kronor where not more than 70–80 minutes. This is a strong indication that there is a lot to gain in liquidity levels from having a system that can settle in real time as well as through continuous net settlement.

Leinonen and Soramäki (1999) run similar simulations on a larger set of data from the Finnish BOF-RTGS system. The result in that study shows a linear trade-off between liquidity and delay, in contrast to the exponential trade-off that the current study on the Swedish data shows. The exponential effect can be explained by the high concentration of payments to a certain point in time. This concentration is perfect for saving liquidity without long delays using a CNS algorithm. However, both studies show that a lot of liquidity can be saved by allowing a small delay, placing limitations on liquidity and using continuous net settlement.

7.4 Conclusions

1. If the Riksbank and the participants are willing to accept a slightly longer delay in settlement, the demand on liquidity could decrease in the order of SEK 60 billion. Although the simulations are a substantial simplification of reality, the liquidity saving is so large that it gives good cause for continued analysis and discussions.

2. The cost of liquidity can vary from bank to bank and over time. A system that offers both offsetting and real-time gross settlement would give banks with a high liquidity cost the option to settle a greater part of their payments through offsetting, while also affording the opportunity to settle payments with RTGS when timing is critical. Combining RTGS and offsetting qualities in one system would give greater freedom of choice and allow for greater dynamics and better optimisation opportunities for participants.
3. Throughput guidelines aim to save liquidity through enhancing the ability of the system to meet payments. The simulations show that a system with continuous net settlement and restrictions on available liquidity can save liquidity without such guidelines. Allowing non-time-critical payments to be settled through continuous net settlement and time-critical payments through a prioritised RTGS track would release capacity and decrease the bottlenecks that create capacity queues in RIX today.
4. The behaviour in RIX, that gives high concentration of payments to a short period during the day, creates perfect conditions for an efficient application of a CNS algorithm. This and the other conclusions indicate that continued studies in this field and discussions between the Riksbank and RIX participants would be useful.

References

Leinonen, H and Soramäki, K (1999) **Optimising Liquidity Usage and Settlement Speed in Payment Systems**. Discussion Paper 16/99, Bank of Finland.

Chapter 8

Simulations of failure in a payment system*

James J. McAndrews – George Wasilyew

8 Simulations of failure in a payment	230
Abstract	230
8.1 Simulations of failure in a payment system.....	230
8.2 Motivation	232
8.3 Simulations and results	234
8.4 Conclusion	239
Tables 8.1–8.4A	241
Figure 8.1	245
Appendix	246
References	247

* This paper first appeared as Working Paper 95-19, Federal Reserve Bank of Philadelphia, June 1995. The views expressed do not reflect the views of the Federal Reserve Bank of Philadelphia, New York, or the Federal Reserve System.

8 Simulations of failure in a payment system

Abstract

We conduct simulations of the performance of a payment settlement system after one of the participants in the system fails to deliver payment. This exercise is meant to supplement Humphrey (1986), who, using actual payment entry data from the Clearinghouse Interbank Payment System (CHIPS), simulated the performance of CHIPS in the event of the failure of the bank with the largest net debit to the system. We follow a similar strategy of simulation, using data we generate, to test the performance of a settlement system as the number of banks, the size of the payments, and the likelihood that any two banks in the system exchange payments vary. Our results suggest that the risk that any given payment fails to be executed varies directly with the number of participants in the settlement system, the variance of the size of payments entered into the system, and the likelihood of interaction among banks in the system.

8.1 Simulations of failure in a payment system

Participation in a network can result in increased risk to an agent. For example, in a payments system network, elementary bilateral payments are carried out multilaterally. The failure to settle accounts by one participant in the network can result in the insolvency of network participants who engaged in no direct trades with the original failing firm. In this way, participation in a network can impart an externality of increased risk at the same time that it may improve the average outcome for the participant.

The phenomenon of a risk externality is modeled in this paper as a compounding of bilateral risks among agents when they interact multilaterally. The model takes the bilateral exchange as basic: two firms trade one unit of a good for money. The exchange is ‘settled’ by transfer of ‘good funds’; there is a risk that the payer bank will not have sufficient funds because of delays in an exogenous (and risky) stream of payments due to it from customers. The payer bank defaults when its balance of good funds falls below the level needed to complete the payment at the time of settlement. In a multilateral

netting payment system, individual trades are batch processed; that is, several payments are bunched and settled at one time. The resulting linkages among the network participants create systemic risk: a bank is exposed to risk that its trading partners will default because of the default of a firm that has not directly exchanged payments with the bank in question. This sort of risk is absent in purely gross payment – continuous process payment arrangements, in which a payment is settled by delivery of the gross amount of the obligation with no netting of offsetting positions, and in which payments are settled sequentially rather than in a batch.

Whether the systemic risk is of great significance or whether it can be controlled are important issues.¹ Because hatching and netting of payments reduces the amount of good funds necessary to settle a given set of payments, the system of multilateral netting-batch processing may well be preferred, even given the presence of risks not seen in a system of bilateral – continuous process payments. Our purpose in this paper, however, is neither to determine the significance of systemic risk nor to devise control mechanisms for systemic risk. Instead, we focus on what elements of a payment system are related to increased or decreased systemic risk.

We've simulated the operations of a particular type of payment system – a multilateral net settlement system *with unwinding* – to gauge the sensitivity of the system operations to changes (treated as exogenous) in important parameters of the model. While unwinding has long been recognised as an unsatisfactory settlement mechanism in large-value payment systems, we've chosen it as a benchmark to show changes in the effects of a participant's failure as the number of participants and other features of the payments themselves change.² We believe that the exercise has implications for systemic risk in other settlement systems as well.

This paper extends the work of David Humphrey (1986), who first simulated the effects of a failure to settle by a large participant in a payment system with unwinding. Our simulations investigate three hypotheses about how systemic risk varies with aspects of a settlement system. We measure systemic risk by studying the effects of a default of the largest net debtor in the payment system. We examine both the number of banks surviving the unwinding of payments, and the total value of payments that are ultimately made after unwinding, as a percentage of the total payments that the participants have entered in

¹ See Boris and Van den Bergh (1993) for an excellent review of the issues involved in systemic risks arising in payment systems.

² See, for example, BIS (1990).

the system at the time that the largest net debtor defaults on its payments. The three hypotheses are: 1) for a given number of participants and distribution of payments, as the probability of interaction among banks (ie, the connectivity of the network, or the ‘distribution of counterparties’) rises, systemic risk rises; 2) for a given distribution of counterparties and distribution of payments, as the number of participants rise, the systemic risk rises; 3) for a given number of participants and distribution of counterparties, as the variance of the size of payments increases, systemic risk rises.

In the next section of the paper we will motivate these hypotheses. In Section 8.3 we will describe the simulations and give the results. Section 8.4 concludes.

8.2 Motivation

First, we introduce some standard notation and terminology. We let Z_{ij} equal the **gross** amount of the payment to be made by bank i to bank j . Then the **bilateral net** payment to be paid by bank i to bank j is given by $B_{ij} = Z_{ij} - Z_{ji}$. The bilateral net payment for the system of payments in a group of banks is

$$Z_B = \sum_{i=1}^N \sum_{j < i}^N |Z_{ij} - Z_{ji}| \quad (8.1)$$

This is simply the sum across banks of each banks net debit to each of its payment counterparties. The **multilateral net** payment for a system of payments in a group of banks is given by

$$F_M = \left(\frac{1}{2} \right) \sum_i^N \left| \sum_{j \neq i}^N Z_{ij} - Z_{ji} \right| \quad (8.2)$$

This sum is the sum across banks of the absolute value of each bank’s net debit to all the other banks; because some banks are net creditors and some are net debtors, the absolute value is halved to determine the total amount that needs to be paid by the net debtors to the net creditors to discharge their obligations. The net debit for an individual bank in a system of multilateral settlement is

$$F_M^i = \sum_{j \neq i}^N Z_{ij} - \sum_{j \neq i}^N Z_{ji} \quad (8.3)$$

For an individual bank, its multilateral net debit is simply the sum of what it owes all other banks minus the sum of what is owed to it by all other banks.

A bank faces a distribution of outcomes when it accepts payment in a multilateral settlement system. One possibility is the payment orders it receives from its customers on a given day exceed the amount of its available good funds. The probability of this happening is $\Pr(F_M^i > r_t)$, where r_t is the payer's good funds balance at the time of settlement, and F_M^i is the bank's net debit. Given a distribution of payments then, banks choose the level of their good funds to limit this occurrence. Various settlement systems also require participants to have access to lines of credit to cover most contingencies. However, the possibility exists that a bank could exhaust these sources of funds and fail to deliver on its payment obligation.

We make the assumption that the probability of default is greater as the variance of the bank's net debit is greater. The primitive in our simulation is the bilateral net debit, $B_{ij} = Z_{ij} - Z_{ji}$. We let it be distributed normally, with mean 0 and variance σ^2 . This implies that the distribution of the underlying gross payments is normal with a finite mean, μ , and variance $[\sqrt{(1/2)}]\sigma^2$. For a settlement system with N banks, the distribution of an individual bank's multilateral net debit, F_M^i , is normal with mean 0 and variance $(k-1)\sigma^2$, where k , $k \leq N$, is the number of banks with which each bank exchanges payments.

This implication of our normal distribution assumption, that the variance of the bank's multilateral net debit increases with the number of counterparties, motivates the hypotheses we investigate in the simulations. First, as the number of banks in the system (with which banks exchange payments) increases, the variance of an individual bank's net debit increases, and hence the hypothesis to be investigated is that the likelihood and consequences of a failure in the system are greater with more banks in the system, all other things equal. The same is true for a group of banks with more interconnections, since this effectively increases the number of banks in the system, and hence the variance of the individual net debits. Finally, the larger the variance of the underlying payments, the greater will be the likelihood and consequences of a failure in the system.

8.3 Simulations and results

The simulations we have performed rely on a highly stylised model of banks. The banks settle their payments through a multilateral net settlement system with unwinding. The size of the payments made by banks and the identity of the payee bank are drawn from a random distribution for each run of the simulation. In each run of the model, we simulate the consequences of the largest net debtor in the settlement system's defaulting. We judge systemic effects by replicating such a failure 500 times for each set of parameters. We focus on three measures of the successful settlement of payments: 1) the probability that the unwinding procedure finds a subset of banks that can successfully settle the payments among themselves; 2) the number of such banks; and 3) the value of the payments that are made (contingent upon finding a stable group of banks that can complete settlement) as a percentage of the *payments outstanding at the time of the default* of the largest net debtor bank. Humphrey (1986) also used this last metric as a measure of the *unexpected* increase in the need for settlement funds; presumably the bank expects and intends to fund the settlement need that is outstanding at the end of the regular settlement cycle, but any default that increases its settlement obligations above that amount are unexpected. In contrast to Humphrey (1986) the data we use are internally generated, which allows us to perturb the parameters to see how the settlement system responds to changes in those key parameters, which is something Humphrey did not do.

Figure 8.1 may help the reader follow the description of the simulation. The banks in the simulations are all of (asset) size 1. The number of banks varies from 5 to 100 in the simulations. A particular bank, bank i , makes a payment to any other bank, bank j , with a certain known probability (fixed at the outset of the simulation): 0.5, 0.7, or 0.9. Given that probability, in each round of the simulation, a draw is made from a binomial probability distribution for each pair of banks. If a particular ordered pair of banks, say banks i and j , draws a 1 in the sampling from the binomial distribution, then a draw is made from a distribution of payments; if 0 is drawn (from the binomial distribution), there is no payment between the two banks. The payment distribution is a normal distribution with mean zero and a standard deviation that varies across the simulations. If the payment drawn is a negative number, then bank i is to make a payment of that amount to bank j ; if it is positive, then i is to receive a payment of that amount from bank j . Once an ordered pair of banks (i,j) draws a one, the ordered pair (j,i) is not sampled. Hence we interpret the payment

that is drawn for the pair (i,j) as the **bilateral net** payment to be exchanged between the two banks.

The simulation proceeds as follows: after drawing payment partners for each bank in the sample, and payment amounts for the banks, the multilateral net payments are calculated for the group of banks and for each bank in the sample, which we denote by $F_M^i(0)$, where 0 is for round zero of the simulation. An assumption of the exercise is that banks **expect** and are prepared to settle (have ready access to good funds) the payment amount $F_M^i(0)$.

Next, we search for the bank with the largest multilateral net debit. To judge systemic risk, we assume that this bank defaults on its settlement obligation. Round one of the simulation begins the unwinding procedure. First, all payments to and from the defaulting bank are ‘unwound,’ that is, they are removed from the set of payments entered into the system. Next, the multilateral net debit is calculated for the remaining banks and their payments, yielding a set of payment amounts, $F_M^i(1)$. Finally, we calculate the difference $F_M^i(1) - F_M^i(0)$. This difference represents the **unexpected requirements** of settlement caused by the initial default. We establish a threshold level of additional and unexpected funds needed for settlement (which we vary across runs of the simulation). Funds required above the threshold are assumed to be unavailable to the bank (even by accessing its lines of credit), and the bank defaults on its settlement obligation. If no bank exceeds the threshold, the simulation ends with round one, and all the remaining banks proceed with settlement.

Round two of the simulation begins if some bank or banks in round one have an unexpected additional settlement need for funds that exceed the threshold. In that case, those banks are removed from the settlement and their payments are unwound. Next, the $F_M^i(2)$ are calculated; the differences $F_M^i(2) - F_M^i(0)$ are calculated and compared with the threshold, and settlement occurs if no bank’s unexpected additional settlement needs exceed the threshold. If some bank’s additional settlement needs exceeds the threshold, then we move to round three, and so on. The simulation continues until settlement occurs or no bank is left in the settlement system.

The first results concern the size of the payments sent through the system. The size of payment is reflected in two parameters of the simulation: the size of the underlying bilateral net debit (which is influenced directly by the variance of the bilateral net debit, σ^2) and the size of the default threshold. The simulation is run with the

variance of the distribution of the underlying bilateral net payments fixed at 0.2 and 0.4, and for default thresholds of 0.2 and 0.3.

Before examining the specific hypotheses, we wish to point out that the differences among the point estimates reported in the tables are statistically significant, assuming the output of the simulations is normally distributed. This is a result of the large size of the samples produced by the simulations. Specifically, using a two-sample t test, and testing for the equality of two numbers (that are different in size) in any of the tables, we reject the hypothesis that the numbers are equal. The test is stated in an appendix. Consider the entries in the last two columns of the bottom row of Table 8.1 (for which the probability of interaction is 0.9, and the number of banks increases from 75 to 100). The means of the surviving proportion of payments are 0.16 and 0.15, a difference of 0.01; the t statistic for the difference of means is $(.01/.000129) = 77$, with 998 degrees of freedom (where the denominator is calculated as in the appendix). Hence the difference between these means is significant at the .005 percent level.

With two default thresholds, two levels for the variance of the underlying net debit, three levels of the probability of interaction, and five different numbers of banks, there are 60 configurations of parameters possible. Only 56 of the 60 cases were investigated and reported in the tables.

Effect of size of payment

Larger bilateral net debits can arise either from larger underlying gross payment flows or from greater disparity in the size of payments between any two banks. The simulation results, presented in Tables 8.1 through 8.4, suggest that smaller sized payments, that is, a smaller variance in bilateral net debits, result in less disruption to the system than do larger sized payments (again, payment sizes are measured as a percentage of a bank's capital). Specifically, for the three probabilities of interaction for banks (0.5, 0.7, and 0.9), the percentage of the original payments that survive the unwinding is greater when the variance of the bilateral net debit is 0.2 than when it is 0.4. For example, looking at the middle cell in each table, when there are 50 banks participating and the probability of any two banks exchanging a payment is 0.7 and when the variance of the distribution of bilateral net debits is 0.2, about 19 percent of the payments survive the shock of the failure of the largest net debtor and subsequent unwindings (when the threshold for failure is 0.2); when the variance of the

distribution of bilateral net debits is 0.4, 14 percent of the payments survive.

The effects of varying the second measure of size of payment, the threshold at which a bank's payment is assumed to fail, are found by comparing Tables 8.1 and 8.3, and by comparing Tables 8.2 and 8.4. Once again, the larger the relative size of the payment (ie, when the payment threshold is 0.2, rather than 0.3), the more disruption to payments is caused by the initial default and the unwinding procedure. The simulation suggests that smaller payments between banks entail less systemic risk than do larger payments. This result, while quite intuitive, is perhaps not as predictable as one might think. With the lower threshold for default of 0.2, it could possibly be the case that all of the largest net debtors are removed from the system in the first round and that all of the other payments go through with less disruption thereafter. In contrast, the larger threshold of 0.3 might allow too many 1.1 'large' net debits to survive in the system for a few rounds, until they fail in a later round, and then with a relatively larger effect than in the case of the lower default threshold. However, these possibilities typically do not overcome the more direct effect of the lower threshold of 0.2, which is to eliminate more banks from the system, thereby resulting in a greater systemic disturbance than is the case with the more generous threshold for default.

The hypothesis is further investigated by using a different measure of systemic failure. The second measure of systemic failure we use is the number of banks that survive the failure of the largest net debtor and subsequent unwinding of payments. The results of the simulations are shown in Tables 8.1A through 8.4A (where the parameters for Table 3.1 are the same as for Table 3.1A, for example). The results of this measure of systemic risk conform to the results of the first measure of the percentage of payments that survive: the larger the payments exchanged (whether measured by a larger variance of the underlying bilateral net debits or by a lower threshold for failure), the fewer banks survive the systemic shock. For example, comparing the middle cells of Tables 8.1A and 8.2A, about 16 banks out of 54 survive the unwinding when the variance of the bilateral net debit is 0.2, while about 13 banks out of 50 survive when the variance is 0.4. Looking at Table 8.3A, about 20 banks survive when the failure threshold is 0.3, compared to the same cell in Table 8.1A, when 16 survive when the failure threshold is 0.2. Again, the simulation suggests that smaller payments between banks entail less systemic risk than do larger payments.

Effect of probability of exchanging payments

The next hypothesis is that greater levels of interaction among banks result in greater systemic risk. This hypothesis is investigated by varying the probability of interaction between pairs of banks. The probability of interaction is the mean of a binomial distribution from which we draw for each ordered pair of banks in the sample. If the draw is '1,' that pair of banks is assumed to exchange a payment; the simulation program then makes a draw from a normal distribution of the net debit amount. In all the tables, the probabilities of interaction vary between 0.5, 0.7, and 0.9. In most cases, as the probability of interaction rises, the effects of the default and unwinding worsen.³

For example, in Table 8.1, the middle column shows the effects of varying the probability of payments being exchanged between any two of the 50 banks in the system. If the likelihood of exchanging a payment with another bank is one-half, then about 22 percent of the payments outstanding after the default of the bank with the largest net debit survive the process of unwinding. As the probability of exchanging payments rises to 0.7 and then to 0.9, the percentage of surviving payments falls to 19 percent and then 17 percent. This result holds if we use the number of surviving banks as the measure of the systemic effects of the initial settlement failure.

The result suggests that the greater the number of links among participants in a settlement system, the greater the consequences of the failure of any one member of the system. There are fewer 'firebreaks' between any two members in the system, even if those two members do not directly exchange payments among themselves.

Effects of number of participants

The final hypothesis we investigate is that a larger number of banks in a payment between systems with very small numbers of participants and other system results in greater systemic risk. Once again, by examining the tables it is clear that a lower percentage of payments typically survive the unwinding as the number of banks in the system

³ In three out of the 73 cases it, which we can compare the effects of a greater probability of interaction (ie, the first columns of Tables 8.1A, 8.2A, and 8.3A) the number of banks surviving the unwinding is greater when there is a greater probability of interaction. We conjecture that these results, which occur only when there are 5 participating banks, and do not reflect on the value of the payments that survive the unwinding, are a result of the small number of banks in the system, and show a difference.

increases. For example, in Table 8.1 the proportion of payments that survive when there are five banks averages to be .77, while it falls to just .14 as the number of banks in the system increases to 100.

The number of banks that survive the unwinding tells a similar story. Tables 8.1A to 8.4A reveal that fewer banks, as a percentage of the banks in the system, tend to remain to settle their payments as the number of initial banks in the settlement system grows. Referring to Table 8.1A we see that more than half of the banks remain to settle when there are initially five banks in the settlement system (2.6 out of the four left after the bank with the largest net debit drops out), and about one quarter when there are initially 100 banks in the system (27.4 out of 99).

8.4 Conclusion

The simulations we carried out were based on a highly stylised model of payment settlement with unwinding. The banks in this simulation behave mechanically rather than rationally, and, therefore, the results of the simulations should be viewed with some caution. The central question investigated in the paper is how a settlement system based on unwinding would sort through the participants in the system and their obligations and produce a set of banks and payments that can successfully settle, given an initial failure of the bank with the largest debit to the system. Systems of settlement with unwinding are still in place in some settlement systems, and the complexity of the multilateral interactions in a large payment system and the short amount of time (typically overnight) within which the system must achieve a resolution of its payment entries are all significant constraints that make a mechanical model of the process a useful first step in analysing how the system reaches settlement.

The results suggest that payments of smaller size result in a smaller risk to the system of settlement. This result supports the movement of the CHIPS settlement system to move away from relying on unwinding for settlement (since it is a large dollar value payment system), as well as the decision by the Federal Reserve System to allow an unwinding rule in the VISANET settlement system (a small dollar value settlement system), so long as individual payment amounts are capped at \$100,000.

Another result is a greater likelihood that any two banks in the system exchange payments generally results in greater losses to the system. Among other implications, this result suggests that highly

concentrated banking systems (in which all payments are funnelled through few banks raising the likelihood that they exchange payments with one another) may have a higher degree of systemic risk than an unconcentrated system.

The final result is that the greater the number of banks in the system, the greater the systemic risk. This result raises the possibility that a system of correspondent banks may result in a lower risk level. The correspondents net the payments of their respondent banks internally, and enter payments that can't be netted and settled internally into the system. It is not clear whether these subnetworks would lower systemic risk or not. This is a topic for future research.

Tables 8.1–8.4A

Table 8.1

Variance of bilateral net debit = .2; payment default threshold = .2; first table entry is the average percentage of payments (across 504 runs of the procedure) completed in the payment system after the default of the largest net debtor and the subsequent unwinding process; the number in parenthesis is the standard error across the rounds of the simulation.

Probability of interaction	Number of banks				
	5	25	50	75	100
.5	.77 (.31)	.31 (.16)	.22 (.07)	.19 (.05)	.14 (.04)
.7	.68 (.35)	.23 (.11)	.19 (.06)	.17 (.05)	.16 (.04)
.9	.60 (.38)	.20 (.10)	.17 (.06)	.16 (.05)	.15 (.04)

Table 8.2

Variance of bilateral net debit = .4; payment default threshold = .2; first table entry is the average percentage of payments (across 540 runs of the procedure) completed in the payment system after the default of the largest net debtor and the subsequent unwinding process; the number in parenthesis is the standard error across the rounds of the simulation.

Probability of interaction	Number of banks				
	5	25	50	75	100
.5	.69 (.32)	.18 (.10)	.15 (.05)	.15 (.05)	.14 (.04)
.7	.54 (.34)	.14 (.08)	.14 (.05)	.13 (.04)	.13 (.03)
.9	.42 (.30)	.13 (.08)	.12 (.05)	.12 (.04)	.12 (.03)

Table 8.3

Variance of bilateral net debit = .2; payment default threshold = .3; first table entry is the average percentage of payments (across 500 runs of the procedure) completed in the payment system after the default of the largest net debtor and the subsequent unwinding process; the number in parenthesis is the standard error across the rounds of the simulation.

Probability of interaction	Number of banks				
	5	25	50	75	100
.5	.89 (.25)	N.A.	.31 (.14)	.24 (.07)	N.A.
.7	N.A.	.38 (.22)	.25 (.09)	.21 (.07)	.20 (.04)
.9	.78 (.31)	.29 (.15)	.22 (.07)	.20 (.05)	.19 (.04)

Table 8.4

Variance of bilateral net debit = .4; payment default threshold = .3; first table entry is the average percentage of payments (across 500 runs of the procedure) completed in the payment system after the default of the largest net debtor and the subsequent unwinding process; the number in parenthesis is the standard error across the rounds of the simulation.

Probability of interaction	Number of banks				
	5	25	50	75	100
.5	.73 (.34)	.23 (N.A.)	.18 (.06)	.17 (.05)	.15 (.04)
.7	.59 (.35)	.18 (.11)	.16 (.06)	.15 (.04)	.14 (.04)
.9	.51 (.35)	.16 (.09)	.14 (.05)	.14 (.04)	.13 (.03)

Table 8.1A

Variance of bilateral net debit = .2; payment default threshold = .2; table entry is the average number of banks that survive the default of the largest net debtor and the subsequent unwinding process; the number in parenthesis is the standard error across the rounds of the simulation.

Probability of interaction	Number of banks				
	5	25	50	75	100
.5	2.6 (.74)	10.9 (3.3)	17.9 (3.4)	25.1 (4.5)	27.4 (4.5)
.7	1.8 (.62)	9.0 (2.8)	16.2 (3.4)	23 (4.1)	30 (5.2)
.9	2.7 (.72)	8.2 (2.6)	15.1 (3.3)	22.3 (4.2)	29 (4.8)

Table 8.2A

Variance of bilateral net debit = .4; payment default threshold = .2; table entry is the average number of banks that survive the default of the largest net debtor and the subsequent unwinding process; the number in parenthesis is the standard error across the rounds of the simulation.

Probability of interaction	Number of banks				
	5	25	50	75	100
.5	2.3 (.49)	7.6 (2.6)	14.2 (3.2)	21.0 (4.4)	27.4 (4.5)
.7	2.4 (.57)	6.8 (2.3)	13.4 (3.1)	19.3 (3.9)	26.1 (4.5)
.9	2.2 (.48)	6.3 (2.0)	12 (3.1)	18.4 (3.6)	24.2 (4.3)

Table 8.3A

Variance of bilateral net debit =.2; payment default threshold =.3; table entry is the average number of banks that survive the default of the largest net debtor and the subsequent unwinding process; the number in parenthesis is the standard error across the rounds of the simulation.

Probability of interaction	Number of banks				
	5	25	50	75	100
.5	3 (.8)	N.A.	22.7 (5.9)	29.4 (5.3)	N.A.
.7	N.A.	12.4 (4.4)	19.8 (4.2)	26.5 (5.0)	34.5 (5.0)
.9	3.3 (.7)	10.5 (3.3)	17.9 (3.6)	25.6 (4.5)	32.9 (4.9)

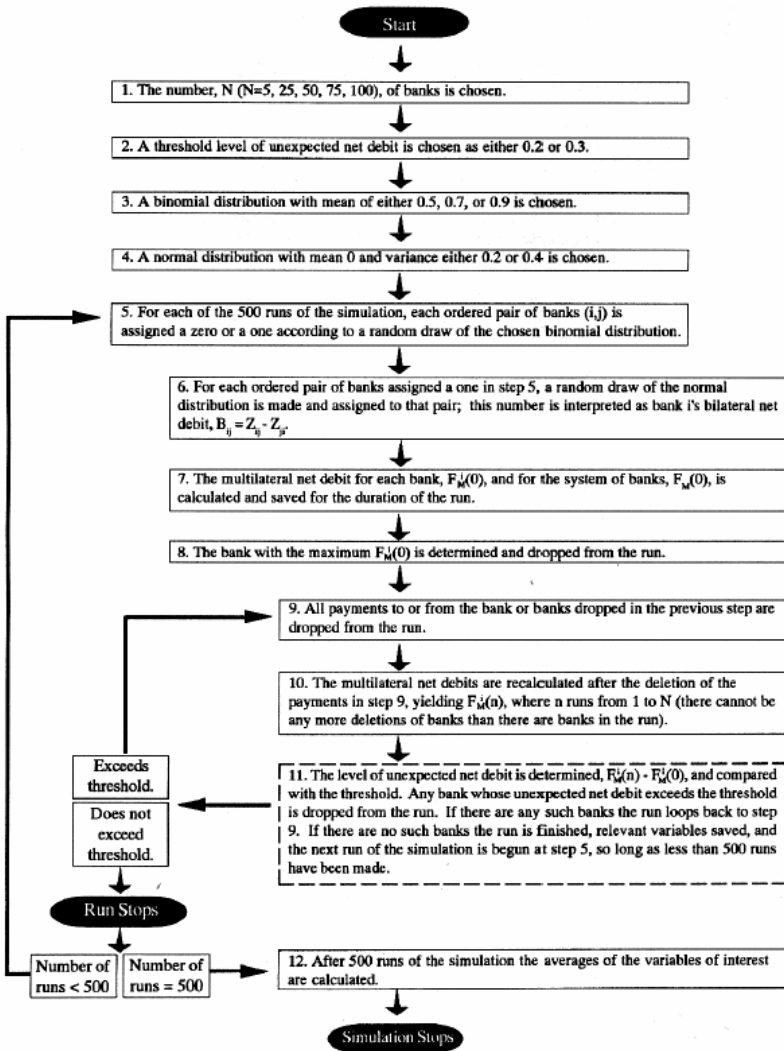
Table 8.4A

Variance of bilateral net debit = .4; payment default threshold = .3; table entry is the average number of banks that survive the default of the largest net debtor and the subsequent unwinding process: the number in parenthesis is the standard error across the rounds of the simulation.

Probability of interaction	Number of banks				
	5	25	50	75	100
.5	2.5 (.65)	N.A.	15.9 (3.5)	22.8 (4.3)	29.4 (4.7)
.7	2.5 (.67)	7.8 (2.6)	14.9 (3.2)	21.3 (4.0)	27.9 (4.7)
.9	2.4 (.61)	7.2 (2.3)	13.6 (3.1)	20.4 (3.7)	26.9 (4.6)

Figure 8.1

Flow chart of the simulation



Appendix

The t test for difference of means is found in Larsen and Marx (1986), p. 352. The data from any two of the simulations are random samples, X_1, X_2, \dots, X_{500} and Y_1, Y_2, \dots, Y_{500} . We assume that they are drawn from a population that is distributed normally (and, independently), with means μ_X , and μ_Y . The sample variances are denoted S_X^2 and S_Y^2 , and the pooled variance is S_p^2 , which is defined as

$$S_p^2 = \frac{(499S_X^2)(499S_Y^2)}{998} \quad (\text{A8.1})$$

because the sample size for both samples is 500. To test the hypothesis ($\mu_X = \mu_Y$), we form

$$T = \frac{\bar{X} - \bar{Y}}{S_p \sqrt{\frac{2}{500}}} \sim T_{998} \quad (\text{A8.2})$$

References

- BIS (1990) **Report of the Committee on Interbank Netting Schemes of the Central Banks of the Group of Ten Countries.** Bank for International Settlements, Basle.
- Boris, C E V and Van den Bergh, P (1993) **The Nature and Management of Payment System Risks: An International Perspective.** Bank for International Settlements, BIS Economic Papers, No. 36.
- Humphrey, D B (1986) **Payments Finality and Risk of Settlement Failure.** In Technology and the Regulation of Financial Markets, eds. Saunders, A and White, L J. Lexington Books, D.C. Heath and Company, 97–120.
- Larsen, R J and Marx, M L (1986) **An Introduction to Mathematical Statistics and Its Applications.** 2nd Edition, Prentice-Hall.

Chapter 9

Analysing the impact of operational incidents in large-value payment systems: a simulation approach

Paul Bedford – Stephen Millard – Jing Yang

9	Analysing the impact of operational incidents in large-value payment systems: A simulation approach.....	250
	Abstract.....	250
9.1	Introduction.....	251
9.2	Types of operational incident	253
9.3	A simulation-based framework for analysing the impact of operational incidents in LVPS.....	256
9.4	Application: Assessing the resilience of CHAPS Sterling.....	265
9.5	Conclusions.....	275
	References	276

9 Analysing the impact of operational incidents in large- value payment systems: A simulation approach

Abstract

Over recent years, central banks have focussed considerable attention on improving the operational resilience of large-value payment systems (LVPS). These systems play a vital role in supporting financial market activity, often standing at the centre of a complex web of infrastructural arrangements for discharging payment obligations and clearing/settling securities trades. Their reliable operation is thus crucial to the continued stability of the financial system and, more generally, to establishing an environment that allows economic agents to exploit valuable opportunities for exchange of goods or assets. It is recognised, however, that the possibility for operational incidents to impair the ability of an LVPS to settle payments cannot be eliminated entirely; an element of residual operational risk will always remain.

In this paper, we propose a simulation-based approach to examining the *impact* of different types of operational incident affecting an LVPS. Our methodology consists of the three discrete stages: the first stage involves identification of a set of worst-case scenarios; the second assesses the impact of the operational incidents associated with these scenarios; and the third and final stage uncovers the empirical distribution of the consequences of operational disruption.

By way of illustration, we present results obtained from applying the proposed methodology to data extracted from CHAPS Sterling (the UK's main large-value payments system). We are able to conclude that the likelihood of an operational failure causing a significant disruption to CHAPS Sterling payments activity is relatively low. In large part, this finding reflects the availability of a range of robust contingency arrangements for CHAPS Sterling and the abundance of liquidity in the system.

9.1 Introduction

During the 1990s, central banks devoted significant resources towards reducing the financial risks banks and other financial institutions can become exposed to through their participation in large-value payment systems (LVPS). Most significantly, many countries implemented new systems based on models of real-time gross settlement (RTGS) that eliminate the financial risks associated with systems that defer settlement in order to allow a netting process to be carried out.¹ For example, the UK's main LVPS – CHAPS Sterling – converted to RTGS in April 1996.

More recently, attention has shifted towards ensuring that key market infrastructures – and LVPS in particular – exhibit sufficiently robust levels of operational resilience.² The value transferred by these systems each day often amounts to 20% or more of the annual GDP of the country concerned, which implies that a disruption to their operation has the potential to impact significantly on the users of these systems and, in extreme cases, undermine the stability of the financial system. To the extent that operational disruption also distorts economic agents' optimal trading decisions, there may additionally be an impact on the wider economy.

A shock to the operation of an LVPS, for example because of a failure of the central payment processing infrastructure, has the potential to compromise one or more of the participants' ability to make payments discharging (possibly very large) settlement obligations. In turn, this increases the likelihood of financial distress at one institution having a knock-on impact on other institutions – that is, operational disruption can be a source of systemic risk. The immediate implication is that the payment system itself is of systemic importance.

The importance of operational resilience is recognised in the *Core Principles for Systemically Important Payment Systems* developed by the G10 central banks (BIS, 2001). In particular, Core Principle VII (CP VII) states that a system should ensure a high degree of security and operational reliability and should have contingency arrangements for timely completion of daily processing. Compliance with CP VII is the minimum central bank overseers typically expect from an LVPS.

¹ More detailed discussions of the properties of RTGS and deferred net settlement (DNS) systems can be found in BIS (1997) or McAndrews and Trundle (2001).

² See Bank of England (2005) for a summary of recent initiatives aimed at strengthening the resilience of UK market infrastructures.

But many such systems, partly motivated by heightened awareness of the possibility for deliberate, large-scale acts of terrorism to disrupt financial institutions and markets, now aim to achieve even higher standards.³ New arrangements and procedures have been introduced to ensure continued operation in all but the most extreme of circumstances.

It is recognised, however, that efforts to improve resilience cannot eliminate entirely the risk that normal operations will be disrupted in some way. Indeed, this is not what CP VII seeks to achieve; rather, the focus is on how operational risk in systemically important payment systems should be mitigated and controlled. This involves analysis of both the likelihood and impact dimensions of the risk.

The focus of this paper is on the *impact* dimension of operational risk in LVPS. In particular, we propose a methodology for determining and evaluating the effects of a set of ‘worst-case’ (in terms of scale of disruption) scenarios. We do not explicitly consider the probability that a particular scenario will occur; the overall level of operational risk remains undetermined. Nevertheless, impact assessments based on simulations of remote probability events – our worst-case scenarios – represent an integral part of the overall risk management process. It is our contention, therefore, that the approach proposed here can serve as a valuable addition to the ‘tool-kit’ available to the operators and overseers of LVPS for evaluating the extent to which a system is exposed to operational risk.

The remainder of the paper is structured as follows: Section 9.2 describes the main types of operational incident that can occur in a large-value payment system; Section 9.3 introduces our proposed methodology for identifying worst-case scenarios and evaluating their impact using simulation techniques; Section 9.4 illustrates the application of the methodology using transactions data captured from the CHAPS Sterling system; and Section 9.5 presents conclusions.

³ In addition to the threat of terrorism, further motivations for improving the operational resilience of LVPS include the risk of natural disasters and catastrophic IT failures (possibly brought about by ‘cyber-attacks’).

9.2 Types of operational incident

Sources of operational disruption

It is a stylised fact that no two large-value payment systems are identical. Indeed, there are many dimensions along which the design and structure of LVPS can differ. The external environment within which a particular system operates may also have an (idiosyncratic) impact on its risk profile. However, at an abstract level, it is possible to characterise all LVPS in terms of four constituent parts:

- A set of settlement banks;
- The provider/operator of the communications network;
- The provider/operator of the payment processing infrastructure; and
- The settlement agent.⁴

For an LVPS that settles payment instructions on a gross (rather than net) basis, the efficient operation of the system will additionally require some form of liquidity generation mechanism. In practice, the settlement agent would typically act as the liquidity provider, extending short maturity (usually intra-day) loans to settlement banks under pre-specified conditions.⁵

For simplicity, we assume that a single entity acts as both the provider/operator of the payment processing infrastructure and as the settlement agent (and liquidity provider). Henceforth, we refer to this entity as ‘the system operator’. This assumption conforms relatively closely with real-world practices: most LVPS currently operating in the G10 countries settle their participants’ payment obligations across accounts held at the local central bank, and it is common (albeit far from universal) for the central bank also to provide the payment processing infrastructure supporting the settlement process. Examples of LVPS that operate in accordance with this characterisation include CHAPS Sterling in the UK and the US Fedwire system.⁶

⁴ The terms ‘settlement agent’ and ‘settlement institution’ are often confused. In this paper, we use the former to refer to the institution providing the accounts across which payments are settled.

⁵ For example, the Bank of England (acting as settlement agent) allows the CHAPS Sterling settlement banks to borrow intraday against eligible collateral.

⁶ In the case of Fedwire, the Federal Reserve also provides the principal communications network.

With this assumption in hand, we are left with three possible sources of operational disruption in an LVPS. More precisely, we define as an ‘operational incident’ any event that compromises the normal functioning of: one or more settlement bank; the provider/operator of the communications network; or the system operator.

Operational disruption involving the *system operator* would typically constitute the unavailability of the core payment processing infrastructure. Possible causes range from an IT (software or hardware) failure to the absence of sufficiently well-trained staff. External events such as natural disasters, power failures or terrorist action have the potential to have similar effects. Unavailability of the core infrastructure would render the LVPS unable to handle payments in the normal way until the problem is resolved. Contingency arrangements that allow a limited number of payments to be processed and settled via alternative means may be available, but it is unlikely that these will be able to replicate fully the service offered by the LVPS under normal operating conditions.

A failure of the *communications network* supporting an LVPS would have a similar impact. New payment instructions would be prevented from reaching the system operator (except by contingency means, if these are available), although payments that have already been received by the system operator could still be processed and settled in the normal way.

Finally, an operational incident could entail the inability of one or more of the *settlement banks* to submit payment instructions to the LVPS. Such a situation would typically arise from a failure of the internal (back-office) systems of the bank or banks concerned, the possible sources of which are similar to those that may cause the core payment processing infrastructure to be unavailable. Unaffected settlement banks would be able to continue to operate as usual, and any payment instructions they submit to the system operator would be processed and settled in the normal way (subject, crucially, to the availability of sufficient liquidity).

Role of LVPS design

The likely impact of any one of the operational incidents described above, expressed in terms of the volume and value of payment instructions affected, will depend, at least in part, on the design of the LVPS. The settlement model – RTGS or deferred net settlement (DNS) – is of particular significance.

In an RTGS system, individual payments are settled with finality on a continuous basis throughout the day. Conditional on the sending settlement bank having sufficient liquidity available, a payment instruction is settled immediately upon its receipt by the system operator. In the event of a liquidity shortfall, payment instructions would typically be held in a central queue to await the arrival (in the account of the sending bank) of additional funds.⁷ Crucially, continuous intraday finality means that all payment instructions successfully settled before an operational incident occurs would be unaffected by the disruption to normal operations.

Where finality is deferred, however, all payment instructions submitted to the LVPS since the last settlement event are 'at risk' from operational disruption. But this is not to say that the consequences of operational incidents are unambiguously greater in a deferred finality DNS system than in an RTGS system. Consider two systems – one RTGS, the other DNS – handling an identical set of payments. The temporary unavailability of any constituent part of the DNS system would be unlikely to have a significant impact if the problem can be resolved in sufficient time to allow all payments to be processed and (net) settlement to take place as planned. By contrast, disruption to the RTGS system would almost certainly result in the final settlement of some payment instructions taking place later in the day than would otherwise have been the case.

The significance of settlement delays in an RTGS system depends on the extent to which the affected payments are time-critical and thus require final settlement at (or before) a certain time. For a large proportion of the payments settled through a typical RTGS system, the precise time of settlement is unlikely to be of major significance. Nevertheless, it is probable that a certain sub-set of payments (for example, those discharging obligations incurred in other market infrastructures) are genuinely time critical.⁸ Delays to the settlement of any of these payments would represent a crystallisation of liquidity risk.

Furthermore, it is important to recognise that there are two conduits through which operational disruption can cause settlement delays in an RTGS system – one direct, the other indirect. Clearly, any

⁷ There are two sources of additional funds: incoming payments (which may represent inter-bank loans) and intraday borrowing from the settlement agent. See BIS (1997) for further discussion on queuing arrangements.

⁸ See Bedford et al (2004) for further discussion on time-critical payments. Unfortunately, it is rarely possible to determine the time-criticality of individual payment from data captured from LVPS.

type of operational incident disruption will have an immediate impact on the ability of at least one settlement bank to settle payments in the usual way. But there could also be significant ‘second-round’ effects. A key feature of any RTGS system is that the liquidity position of each settlement bank is directly influenced, on a continuous basis, by the pattern of payment flows.⁹ Consequently, any disruption to the payment activities of one RTGS settlement bank will have an indirect effect on the liquidity positions of all other settlement banks. In turn, this could force these banks’ to delay the settlement of their own payments.

Beyond the chosen settlement model, a further important aspect of LVPS design concerns the use of dedicated contingency arrangements that aim to reduce the impact and, where possible, the likelihood of episodes of operational disruption. A wide range of alternative measures have been employed, and no two systems use an identical set of arrangements.

9.3 A simulation-based framework for analysing the impact of operational incidents in LVPS

In this section, we outline our proposed framework for assessing the impact of three different types of operational incidents:

- i. The inability of one settlement bank to send payments;
- ii. Similar problems involving multiple settlement banks (simultaneously); and
- iii. The unavailability of the central payment processing infrastructure.

For incidents (i) and (ii), our simulation methodology focuses upon the *indirect* impact of the disruption to normal operations. The immediate, direct consequence of these operational incidents will be that at least one settlement bank is unable to settle payments in the normal way. However, there may also be an impact on the liquidity positions of the other, otherwise unaffected settlement banks; it is

⁹ In a DNS system, by contrast, liquidity is transferred between settlement banks at discrete times; thus the relationship between the pattern of payment flows and liquidity positions is far weaker than in an RTGS system.

these indirect effects that we focus on. Incident (iii), by contrast, impacts upon all settlement banks equally and there is no material distinction between direct and indirect effects.

For each type of operational incident, we propose a three-stage procedure. The first stage involves examining data on payment flows through the LVPS concerned to identify, on the basis of objective criteria, the ‘worst-case’ point in time for an operational incident to occur. Using this information, the second stage of the procedure is to simulate the pattern of payment activity in the LVPS operating normally and under the assumption that the worst-case scenario actually occurs; comparison of the simulation results allows an assessment of the impact of the operational incident to be made. The third and final stage involves repeating the process using different (and non-overlapping) data samples in order to allow the empirical distribution of the impact of each type of incident to be derived.

Our methodology assumes that there are no circumstances under which operational disruption creates doubts regarding the financial soundness of LVPS settlement banks; therefore it does not aim to provide meaningful insights in respect of the management of financial crises. Furthermore, we assume throughout that the operational disruption described is of sufficient severity to preclude a resumption of normal payment processing activities during the same business day.

Stage 1: Identifying worst-case scenarios

The first stage of our proposed methodology involves identifying a set of ‘worst-case’ scenarios that will form the basis for the simulation analysis. For each of the three types of operational incident listed above, the worst-case scenario is defined in terms of the point of time (within the chosen sample period) when the occurrence of the incident entails the largest potential impact to the LVPS concerned. Clearly, the precise way in which the worst-case scenario is identified will depend on the nature of the incident under consideration and criteria used to measure ‘potential impact’.

We consider first the case of an **operational problem involving a single settlement bank**, the immediate impact of which is to prevent the bank concerned (henceforth referred to as the ‘stricken bank’) to submit payments to the LVPS. The direct effect of the incident is that the stricken bank will not be able to complete its payment activities as planned. In addition, and as discussed in the previous section, the incident may also have indirect effects that compromise the ability of unaffected settlement banks to settle their own payments.

In a DNS system, the scale of these indirect effects is likely to be very small. Even if, as assumed, the operational problem persists to the point of settlement, it is probable that the system operator will be able to invoke contingency arrangements in order to complete the settlement process. Consequently, there is no reason to expect that the settlement banks unaffected by the original incident will be unable to settle all of their payments as usual. Given this observation, we restrict our attention to RTGS systems.

An operational incident involving one (or more) RTGS settlement bank has the potential to influence the liquidity positions of other settlement banks not directly affected by the initial shock. In particular, there is a risk that these banks will encounter liquidity shortages as a result of large amounts of available liquidity becoming trapped in the account of the stricken bank. Put differently, the stricken bank could act as a ‘liquidity sink’.

However, it is important to recognise that the likelihood of liquidity shortages materialising depends critically on the behaviour of the settlement banks not directly affected by the initial shock. Indeed, we would expect these banks, upon learning of the operational disruption, to take action aimed at preventing the stricken bank from becoming a liquidity sink.¹⁰ The obvious way of achieving this is to stop sending payments to the stricken bank. Anecdotal evidence from CHAPS Sterling suggests that the time-lag between an individual settlement bank experiencing an operational failure and the flow of payments to that bank slowing significantly is typically of the order of ten minutes.¹¹

The behavioural response of the unaffected settlement banks notwithstanding, the possibility for a liquidity sink to develop remains real. This is most particularly the case in situations where the stricken bank holds a large credit balance on its account with the settlement agent at the time of the initial operational incident. The risk is also greater where the stricken bank is due to receive a large gross value of payments in the few minutes immediately following the operational failure (that is, before other settlement banks have an opportunity to respond to the incident). In the event that a liquidity sink does indeed develop, its likely significance – in terms of potential to cause

¹⁰ McAndrews and Potter (2002) analyse the behavioural response of US settlement banks following the terrorist attacks on 11 September 2001.

¹¹ More generally, the time lag between an operational incident occurring and a significant slowing in the flow of payments to the stricken bank is likely to depend on, *inter alia*, the number of settlement banks and the specific operational procedures in place in the LVPS concerned. Consequently, we would expect the lag to vary across systems.

settlement delays – will be increasing in the volume/value of payment instructions still to be settled on the day concerned.

The potential for an LVPS settlement bank to become a liquidity sink can be assessed using actual payment flow data for a given sample period.¹² In particular, we propose measuring the likelihood of a liquidity sink effect arising by calculating, for each settlement bank at every point in time, a ‘virtual credit balance’. This metric is defined as the actual balance on a bank’s account with the settlement agent plus the gross inflow of payments over the next x minutes.

Building on this analysis, we propose identifying the worst-case date and time for an individual settlement bank to encounter operational problems on the basis of two criteria. First, the incident should occur before a particular time of day. Imposing this ‘time-of-day’ constraint serves to ensure that the volume/value of payments still to be settled after the initial shock is significant. Second, and subject to the time-of-day constraint, the incident should occur at the point of the highest observed virtual credit balance (and involve the settlement bank holding that balance).

We should note two important caveats to this approach to identifying the worst-case scenario for an operational incident affecting a single RTGS settlement bank. First, we have implicitly assumed that overnight balances on accounts held with the settlement agent are close to zero. If, on the other hand, a settlement bank chooses (or is required) to hold a credit balance overnight, then it is necessary to scale the virtual credit balance downwards by this amount. Second, our approach represents just one possible, albeit intuitively appealing, method of identifying the worst-case scenario. It cannot be ruled that there is some alternative scenario that, due to the particular pattern of payment flows, would generate more serious effects. However, the only way to be certain that the ‘true’ worst-case has been identified would be to simulate every possible contingency; clearly this is infeasible.

Turning to the case of an **operational incident affecting multiple settlement banks simultaneously**, we propose a very similar technique for identifying the worst-case scenario. Rather than a single stricken bank, we now have y (>1) stricken banks. Furthermore, it seems reasonable to expect an operational incident affecting several settlement banks simultaneously would be highly visible; therefore we

¹² In purely computational terms, it is often more convenient to use output from the benchmark simulations (see next section) to identify the worst-case scenarios.

assume that the flow of payments to the stricken banks ceases immediately after the initial shock.

There are again two criteria for determining the worst-case scenario: first, the time-of-day constraint should be imposed; second, the operational incident should occur at the point at which the aggregate credit balance of any combination of y settlement banks is maximised (this criterion also identifies the settlement banks affected by the worst-case operational incident).

The final type of operational incident concerns the **failure of the core payment processing infrastructure**. Under the assumption that the LVPS concerned is at least able to open as usual, the worst-case scenario is clearly one in which the incident occurs very early on the day during which the largest volume/value of payments are due to be settled. This applies equally to RTGS and DNS systems.

Recognising the potential for a failure of the core payments processing platform to prevent the settlement of a large volume/value of payments, most LVPS have developed contingency arrangements designed to allow the most urgent (that is, time-critical) payments to be settled. The presence of these back-up arrangements, which typically differ from system to system, can potentially change the nature of the worst-case scenario.

In the main, however, contingency arrangements are not able to cater for the full volume of payments processed under normal operating conditions; the possibility of some transactions remaining unsettled cannot, usually, be ruled out.

To give an example of such contingency arrangements, take CHAPS Sterling. In the event that the RTGS infrastructure at the centre of CHAPS Sterling is inoperable, the system is able to revert to 'RTGS by-pass mode', under which it operates as a (protected) DNS system. The aim of by-pass mode is to allow as many payments as possible to settle on a same-day basis. Section 4 provides further details on by-pass mode and its consequences for the (operational) risk profile of CHAPS Sterling.

Stage 2a: Establishing a benchmark

The second stage of the methodology involves simulating activity in the LVPS concerned under the assumption that the worst-case scenario actually occurs. But to arrive at meaningful results, it is first necessary to establish a set of benchmarks against which the outcome of the simulations of the LVPS operating under stressed conditions can be compared.

Establishing suitable benchmarks is relatively straightforward. The first task is to simulate the LVPS under normal operating conditions using actual payment flow data (for a given sample period). In addition, it is convenient to allow all settlement banks to draw on unlimited amounts of intraday credit. Output from this simulation can then be used to derive two hypothetical liquidity levels useful in analysis of RTGS systems in particular. The ‘upper bound’ of liquidity measures the amount of intraday credit an RTGS settlement bank would need to obtain in order for all its outgoing payments to settle immediately upon their submission to the system (that is, without being placed in the central queue to await the arrival of additional liquidity). The ‘lower bound’ of liquidity, on the other hand, refers to the amount required for the settlement bank just to cover its net outflow of funds across the day as a whole.

Although the total amount of intraday credit actually obtained in an RTGS system may in fact exceed the upper bound of liquidity, the upper and lower bounds nonetheless define a suitable range across which settlement banks’ ability to draw on intraday credit can be varied. Both metrics should be calculated for each settlement bank on each day in the sample period. Further discussion of the concept of upper and lower bound of liquidity concept can be found in, for example, Bech and Sorämaki (2001).

In order to facilitate investigation of the extent to which the amount of liquidity available in an LVPS influences the ability of the system to withstand different types of operational disruption, it is useful to perform additional benchmark simulations. These should be conducted using the same transaction data as previously and with the LVPS operating under normal conditions, but with different amounts of intraday liquidity available to the settlement banks. In particular, the amount of liquidity available should be constrained to a range of levels between the upper bound (UB) and the lower bound (LB), defined on the basis of the following expression (where $0 \leq \alpha \leq 1$):

$$UB - \alpha (UB - LB) \quad (9.1)$$

Stage 2b: Evaluating the impact of operational incidents

With a set of benchmarks established, the next step in the methodology is to simulate payment activity in the LVPS conditional on an assumption that the worst-case scenarios identified in stage one actually occur. As with the benchmark simulations, a series of

experiments should be carried out (for each type of operational incident) using different levels of available liquidity.

In order to measure the impact of operational disruption, we propose using a total of four metrics:

- Total value of unsettled payments;
- Total volume of unsettled payments;
- Average queue value; and
- A delay indicator.

The first two metrics are straightforward, while average queue value is a simple across-time average of the value of payments held in the central queue (awaiting the arrival of additional liquidity). For our final metric, we use the delay indicator introduced by Bech and Soramäki (2001). This statistic is based on the amount of time each individual payment instruction spends in the central queue relative to its maximum possible queuing time. As a value-weighted average of the (relative) delay to all payments, it may be viewed as a measure of the aggregated level of delay in the system. Algebraically, the statistic is defined as defined as:

$$D = \frac{\sum_i (t_{2,i} - t_{1,i}) \cdot a_i}{\sum_i (t_{\text{end}} - t_{1,i}) \cdot a_i} \quad (9.2)$$

where a_i is the value of payment i , $t_{1,i}$ and $t_{2,i}$ are, respectively, the submission and settlement times for payment i and t_{end} is the time for the end of the business day. By construction, $0 \leq D \leq 1$; a value of one shows that every payment has been held in the queue for the maximum possible time (that is, from the point it was first submitted to the system until the end of the business day).

All four of these metrics can be calculated from the output of the benchmark and stressed-conditions simulations. Comparison of the two sets of results reveals the impact of the operational incident concerned. For each type of incident, several sets of results will be obtained – one for each liquidity level (or value of α) considered.

Stage 3: Uncovering empirical distributions for the impact of operational disruption

Applying the first two stages of our proposed methodology produces a range of quantitative measures capturing the consequences of different types of operational disruption in an LVPS. The results represent point estimates of the worst-case impact of an operational incident occurring during a given sample period. Although these findings are useful in their own right, the value of our proposed methodology would be significantly increased if it could also be used to monitor *trends* in the potential impact of operational disruption.

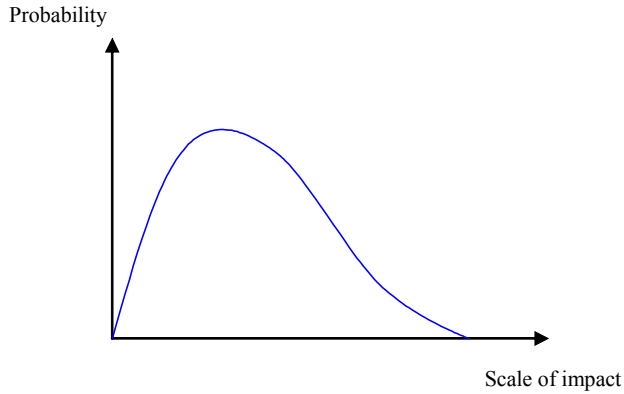
A problem arises, however, from the fact that observed payment flow data is used to identify the worst-case scenarios. This implies that the potential worst-case impact of an operational incident is in fact a random variable and will vary over time. Direct comparison of results obtained using data from two discrete sample periods therefore conveys relatively little meaningful information. Put differently, an observation that the potential impact of operational disruption appears smaller when assessed using payment flow data from February than it did using data from January does not necessarily imply that the system concerned has improved its level of resilience; rather, the results may simply reflect variation in the pattern of payment activity.

The third stage of our methodology aims to overcome this difficulty. More specifically, we propose employing the standard approach to drawing comparisons across realisations of random variables: namely, to derive the empirical distribution of the potential impact of each different type of operational disruption. In practical terms, this involves repeating stages one and two of our methodology as many times as is feasible, using historical payment flow data broken down into (non-overlapping) sample periods of uniform length.

Figure 1a offers a stylised illustration of the distribution of the impact of one particular type of operational incident for a given level of liquidity. In the diagram, the x-axis captures the scale of the disruption (as measured by any one of the four metrics described above).

Figure 9.1a

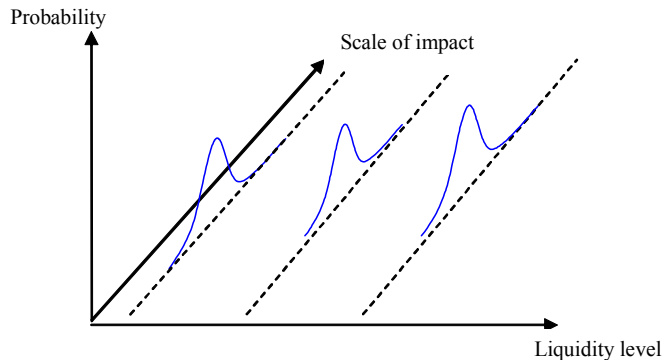
**Impact of an operational incident
(for given liquidity level)**



For each type of operational incident, distributions such as that shown in Figure 9.1a can be constructed for a range of liquidity levels. Given that we would expect the shape of distribution to be related to the amount of liquidity in the system, it is often illustrative to compare the different distributions. Figure 9.1b illustrates how this could be done diagrammatically.

Figure 9.1b

**Impact of an operational incident
(different liquidity levels)**



Once derived, the empirical distributions provide the basis for more informed interpretation of results obtained from applying stages one and two of the methodology to 'new' data extracted from an LVPS on a regular basis. In particular, it would be possible for the operator and/or overseer of the system to determine whether the results imply a

level of vulnerability to operational disruption that is high relative to the historical average. Although this type of finding may give some cause for concern, the way in which the distribution itself evolves over time is likely to be more significant. In general, we would expect to interpret any shift in the mass of the distribution towards having larger ‘impact’ as reflective of increased vulnerability within the system. If the extent of such a shift is sufficiently great, action may then be warranted to improve the resilience of the system (perhaps through a strengthening of contingency arrangements).

There are, however, a number of practical considerations that should be taken into account when deriving the empirical distributions. In particular, it is possible that the underlying payment flow data will be non-stationary; if this is the case, shifts in the distribution(s) of the impact of an operational incident should be interpreted with caution. Furthermore, this concern provides a rationale for ensuring that the length of each sample period is in line with any seasonal patterns observed in the data. For example, in situations where aggregate payment values are always significantly higher on (say) Fridays than on other days, a sample period of at least one week should be used in order to ensure that the worst case scenario over the whole week is identified.

9.4 Application: Assessing the resilience of CHAPS Sterling

In this section, we apply the first two stages of our proposed methodology using payment data extracted from the CHAPS Sterling system (for which the Bank of England acts as system operator). In so doing, we must take due account of the design of the system and procedures for its day-to-day operation. In particular, the likely effect of contingency arrangements should be reflected in the application of the methodology. Bedford et al (2004) provide a detailed description of the mechanisms that have been put in place to control operational risk in CHAPS Sterling; some of these mechanisms, notably the availability of ‘RTGS by-pass mode’, are relevant to our analysis and are discussed at various points in the section.

Data and simulation method

In applying the methodology, we use payment flow data from February 2004. During this particular month, CHAPS Sterling handled approximately 2.1 million payments with a total value in excess of £4 trillion. Put differently, the system settled, on average, over 100,000 payments worth £215 billion per day.¹³

For the sample period selected, we follow the proposed methodology by first identifying the worst-case scenario for each of our three types of operational incident. For this analysis, we set the time-of-day constraint to require that the incident occur before 12noon and assume that, in the case of an incident involving a single settlement bank, there is a delay of ten minutes before the flow of payments to that bank ceases (that is, we set $x = 10$). We then perform benchmark and stress-conditions simulations using five different liquidity levels, defined by $\alpha = 1, 0.75, 0.5, 0.25$ and 0 . This exercise produces a single point in the distribution for each type of incident scenario and each liquidity level. An area for future work is to repeat this exercise for a range of sample periods using historical CHAPS Sterling data going back several years (stage three of the proposed methodology).

The results of our simulations are presented below. All our simulation experiments were carried out using a payment system simulator developed by the Bank of Finland (the BOF-PSS2). Each experiment was conducted using a simulation set-up intended to replicate CHAPS Sterling as closely as possible, including in respect of central queuing arrangements and procedures for gridlock resolution.¹⁴

Operational failure affecting one settlement bank

As discussed previously, this type of operational incident has greatest potential to have an impact on the unaffected settlement banks'

¹³ James (2003) provides a comprehensive statistical summary of payment activity in CHAPS Sterling.

¹⁴ Under operational procedures in place in February 2004, the CHAPS Sterling central queuing mechanism was not used; this implies that settlement banks queued payments for which insufficient liquidity was available within their own back-office systems. The simulations carried out for this article relax this constraint and allow payments to queue centrally. In addition, the simulations employed gridlock resolution procedures at the end of the day only. (Gridlock resolution involves the simultaneous gross settlement of offsetting payment flows).

payment activities when the stricken bank acts as a liquidity sink. In particular, employing stage one of our proposed methodology allowed us to identify a point at which one particular CHAPS Sterling settlement bank held a credit balance of £4.2 billion (on its account with the Bank of England) as the worst-case time for the operational incident to occur.

Over the remainder of the day concerned, this settlement bank was due to be either sender or receiver for some 46,000 payments with an overall value of £45.7 billion. This illustrates a point made by James (2003) – an operational incident of the kind considered here has the potential to have a large *direct* impact on the total volume and value of payments actually settled in CHAPS Sterling.

The focus of our analysis, however, is on the *indirect* effects of the operational incident. That is, we are interested in the extent to which the operational failure creates liquidity shortages in the system as a whole and consequently causes delays to the settlement of payments between the unaffected settlement banks.

Under normal operating conditions, any level of initial liquidity at least equal to the lower bound would be sufficient to allow all CHAPS Sterling payments to settle same-day. This is not necessarily the case, however, following an operational incident; the failure of the unaffected settlement banks to receive payments from the stricken bank may leave them short of liquidity. In fact, the simulation results reported in Table 9.1 reveal that, at most liquidity levels, the disruption did not prevent settlement of a substantial volume and value of payments between the unaffected settlement banks. Indeed, a significant impact (in terms of unsettled payments) was observed only at the lower bound of liquidity.

Table 9.1 **Effect of an operational failure affecting one CHAPS Sterling settlement bank on payments between other settlement banks**

Liquidity level	Value of unsettled payments (£ billions)	Volume of unsettled payments	Average queue value (£ billions)	Delay (D)
$\alpha = 0$ (UB)	0.00	0	0.01	0.00
$\alpha = 0.25$	0.00	0	0.03	0.01
$\alpha = 0.5$	0.01	7	0.07	0.01
$\alpha = 0.75$	0.35	66	0.15	0.04
$\alpha = 1$ (LB)	4.03	4,086	0.23	0.06

In practice, the actual amount of liquidity available in CHAPS Sterling significantly exceeds (by about 50%) the upper bound. The findings reported in Table 9.1 therefore imply that CHAPS Sterling is well-placed to withstand the indirect (liquidity-related) effects of an operational incident affecting one settlement bank.

The foregoing discussion considers the extreme outcome of the operational failure causing liquidity shortages that are of sufficient scale to prevent the settlement of all CHAPS Sterling payments. We are also interested, however, in the extent to which settlement of individual payments (which may be time-critical) is delayed beyond the preferred time – that is, the amount of extra time payments spend in the central queue awaiting the arrival of additional liquidity. Following the methodology outlined in the previous sections, the final two columns of Table 9.1 present two alternative measures of queuing and delay, in each case expressed relative to results obtained from the benchmark simulations of CHAPS Sterling operating under normal conditions (with the appropriate amount of available liquidity).

The results point to the conclusion that an operational failure involving a single settlement bank is unlikely to have a significant impact on the system as a whole.¹⁵ The only exception to this outcome is where the initial level of liquidity is very low. However, the fact that current levels of liquidity in CHAPS Sterling are in fact significantly higher than the *upper* bound implies that such an outcome is improbable.¹⁶ Furthermore, our simulations do not take account of the availability of the CHAPS Sterling ‘stricken bank scheme’ (described in Bedford et al (2004)), the effects of which would be to reduce any liquidity constraints introduced as a result of an operational failure affecting a single settlement bank.

Operational failures affecting multiple settlement banks

For this type of operational incident, we assume that three settlement banks are affected and left unable to submit payments to the system (that is, we set $y = 3$). Again employing the methodology introduced in Section 9.3, examination of payment flow data from February 2004 allowed us to identify a point at which three CHAPS Sterling

¹⁵ More precisely, the conclusion to be drawn is that the impact on the ability of the unaffected settlement banks to make payments to each other is unlikely to be significant.

¹⁶ James and Willison (2004) investigate the factors lying behind the CHAPS Sterling settlement banks’ liquidity (or collateral) posting decisions.

settlement banks (collectively) controlled £4.8 billion of liquidity as the worst-case time for an operational disruption to occur.

The immediate consequence of an operational incident occurring at this time would have been that nearly 51,000 payments with an overall value of £143.4 billion could not have been settled same-day because they involved one of the stricken banks as either payer or payee. As before, it is clear that the direct effects of this type of incident would be material.

Turning to the indirect effects, Table 9.2 reports the outcome of our simulations in terms of the impact on the payment activities of the settlement banks not directly affected by the original incident. The results are again presented relative to the outcome of the benchmark simulations of CHAPS Sterling operating under normal conditions.

A comparison of Tables 9.1 and 9.2 reveals that the indirect effects of operational disruption involving multiple settlement banks are, unsurprisingly, significantly greater than in circumstances where a single settlement bank is affected. Nevertheless, it remains the case that no payments between unaffected settlement banks were left unsettled at the upper bound of liquidity (by implication, this result would also hold at actual levels of liquidity in CHAPS Sterling).

Table 9.2 **Effect of an operational failure affecting three CHAPS Sterling settlement banks on payments between other settlement banks**

Liquidity level	Value of unsettled payments (£ billion)	Volume of unsettled payments	Average queue value (£ billions)	Delay (D)
$\alpha = 0$ (UB)	0.00	0	0.02	0.03
$\alpha = 0.25$	0.68	24	0.05	0.05
$\alpha = 0.5$	2.84	1,078	0.07	0.06
$\alpha = 0.75$	7.58	3,225	0.13	0.20
$\alpha = 1$ (LB)	13.08	6,299	0.02	0.25

It is also clear from Table 9.2 that an operational event affecting three settlement banks is likely to lead to a significant increase in settlement delays, even at upper bound levels of liquidity. A caveat to this finding, however, is that the removal of all payments involving the three stricken banks has a very significant impact on the set of payments upon which the measures of queuing and delay are based (relative to the benchmark simulations). This explains why the

relationship between liquidity level and the average queue value is not monotonic.

In order to understand the intuition behind the non-monotonic result, which is at first glance surprising, consider the following (hypothetical) example. Suppose that an RTGS system consists of four banks: two large banks, A and B, which wish to send payments of £10 million to each other and £5 million to each of the other two smaller banks, C and D. Both bank C and bank D wish to send payments of £5 million to each of the other three banks. To keep things simple, we divide the day into two equal periods and assume that no bank can use payments received in period 1 as liquidity for outgoing payments in that period. Furthermore, we impose the restriction that payments placed in the queue in period 1 cannot be settled until period 2.

Now suppose that, in period 1, both bank A and bank B pay banks C and D £5 million each, and that banks C and D pay £5 million to each other. Figure 2a illustrates the flow of payments in period 1.

In period 2, the remaining payments are settled: banks A and B pay £10 million to each other; and both bank C and bank D pay banks A and B £5 million each. Figure 2b illustrates.

Figure 9.2a

Payment flows in period 1

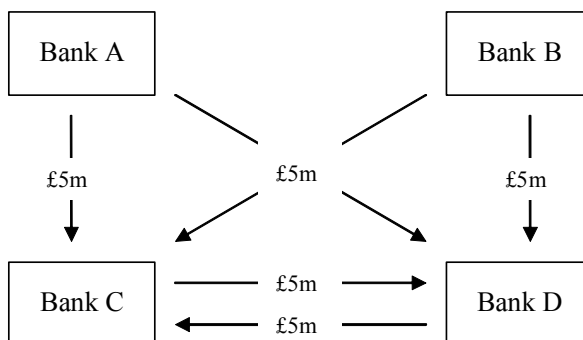
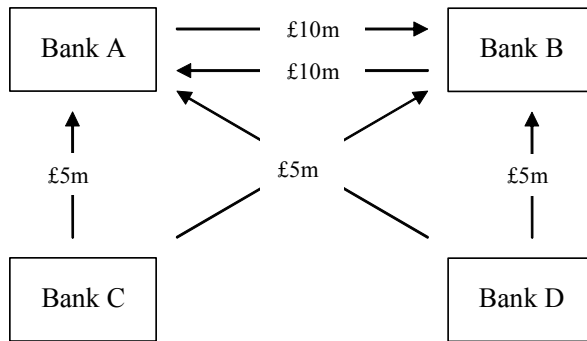


Figure 9.2b

Payment flows in period 2

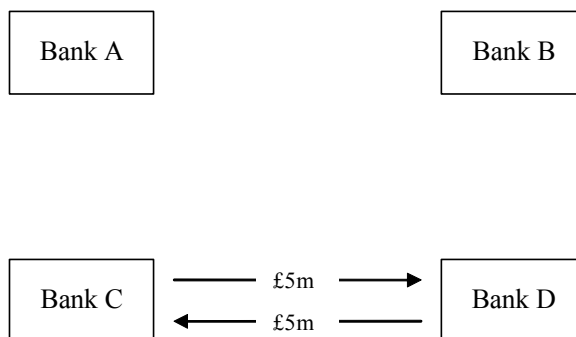


In this example, the upper bound of liquidity will be £20 million for banks A and B and £5 million for banks C and D. With this amount of liquidity in the system, no payments would be placed in the central queue. But suppose that the actual amount of liquidity available is limited to 50% of the upper bound: £10 million for banks A and B; and £2.5 million for banks C and D. With these levels of liquidity, the payments between banks C and D would not be able to settle in period 1; rather, both payments will be placed in the central queue until they can settle in period 2. The queue value at the end of period 1 will thus be £10 million. Moving to period 2, there will be insufficient liquidity to settle the payments between banks A and B. Consequently, the queue value at the end of period 2 will be £20 million and the average queue value (across the day as a whole) will be £15 million.

Now suppose that, at the beginning of the day, both banks A and B experience an operational problem that prevents them from submitting payments to the system for the whole day (ie both periods). Banks C and D respond by not sending payments to either of the stricken banks. Figures 9.2c depicts the modified pattern of payments flows in period 1. In period 2, no payments will be made.

Figure 9.2c

**Payment flows in period 1
(with operational incident)**



At upper bound liquidity levels, the operational incident would not prevent the payments between banks C and D from being settled in period 1. But at the lower liquidity levels described above, both payments would be held in the queue for the whole day. The average queue value would thus be £10 million, actually *lower* than in the case of no operational disruption.

Operational failure of the central payment processing infrastructure

In this final scenario, it is assumed that an operational incident renders the CHAPS Sterling central payment processing infrastructure inoperable. Unless the incident occurs close to the end of the day, the standard response to such a situation (which has never occurred) would be to invoke RTGS by-pass mode, under which CHAPS Sterling reverts to a model of deferred net settlement. Payment information continues to flow between settlement banks, but finality is no longer achieved in real-time. Rather, obligations are settled on a multilateral net basis at the end of the day.

Invoking by-pass mode would allow payment processing to continue, but also creates the potential for additional settlement risks to arise. In particular, the departure from RTGS would result in the accumulation of intraday credit exposures between the CHAPS Sterling settlement banks. The default of a settlement bank holding a net debit position could then cause other settlement banks (and/or their customers) to incur financial losses.

By-pass mode can operate in two ways. Under the preferred variant, all net debit positions are fully backed by liquidity held with

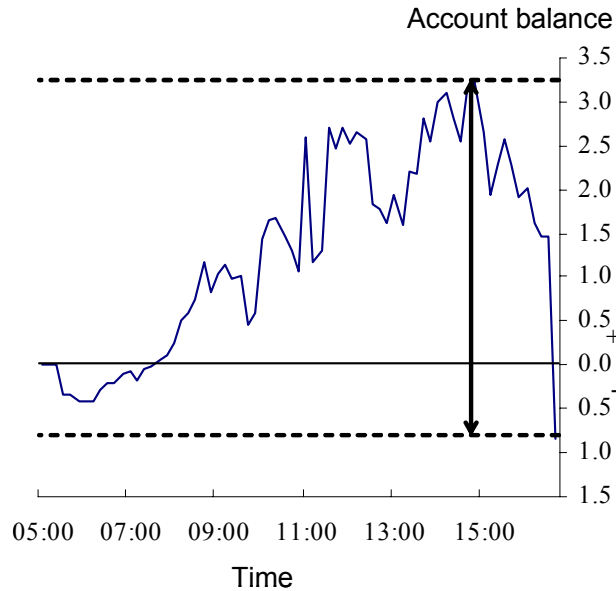
the Bank of England and credit risk is tightly controlled. This approach is only viable, however, when it is possible to establish with certainty the amount of liquidity each settlement bank has available at the time the operational failure occurs. If the required information is not available, it is necessary to implement the second variant of by-pass mode, under which all settlement banks self-impose net sender (debit) caps equal to a maximum of £1 billion. Uncollateralised credit exposures therefore arise to the extent that one or more settlement bank incurs a net debit position in excess of the amount of liquidity it has available at the Bank of England. Our analysis focuses on this second variant of RTGS by-pass mode.

When CHAPS Sterling is operating in by-pass mode, it is possible that the net sender caps would have the effect of preventing the settlement of some payments. In particular, this outcome would arise if the net value of payments an individual settlement bank has to make during a period of by-pass mode operation exceeds £1 billion. In practice, however, the constraint imposed by the sender caps could be relaxed by means of inter-bank loans agreed between the settlement banks, which, when processed by CHAPS Sterling (operating in by-pass mode), would reduce the net debit position of the borrower.

Nevertheless, this analysis implies that one of the potentially most difficult scenarios involves invoking by-pass mode at a point of time when an individual settlement bank would incur a large net debit position during the period CHAPS Sterling is operating in by-pass mode. On any given day, the maximum possible value of this position may be calculated from payment flow data by measuring the difference between each settlement bank's largest intraday net credit balance position and its end-of-day balance. Figure 3 illustrates a randomly-generated settlement account balance over the course of one business day – the vertical distance shown by the arrow represents the maximum possible net debit position the settlement bank concerned could incur were by-pass mode to be invoked intraday.

Figure 9.3

Sample settlement account balance



By examining data from February 2004, we were able to identify the time at which invoking by-pass mode would have led to the largest single net debit position being incurred. We take this to be the worst-case scenario. A simulation experiment was then performed under the assumption that by-pass mode was invoked (with net sender caps set at £1 billion for each settlement bank) at this exact time.

The simulation results indicate that, under the worst-case scenario, 23 payments with a total value of £3.8 billion would have remained unsettled at the end of the day. This finding stems from the fact that one settlement bank needed to make net payments in excess of £1 billion between the time that by-pass mode was invoked and the end of the day, but was unable to do so as a result of the sender cap.

An important aspect of this analysis is that it has assumed that settlement bank behaviour is unchanged. In practice, a settlement bank would be likely to raise additional liquidity by borrowing in the inter-bank market (or elsewhere) in order to ensure it can settle all of its outstanding transactions. This has the effect of transferring, but not eliminating, the credit risk associated with the net debit position; rather than being within CHAPS Sterling, the exposure would then be held outside the system by the lending bank.

9.5 Conclusions

In this paper, we have proposed a three-stage methodology for using simulation techniques to assess the impact of different types of operational incident in large-value payment systems. Our framework is designed to be sufficiently general to apply to any generic LVPS, although it should be recognised that no two systems share an identical design (including in respect of contingency arrangements) and operating environment. It is therefore necessary to ensure that the basic methodology is applied in a manner well-suited to the system concerned. With this caveat, the proposed approach can be useful to the operator and overseers of an LVPS both as a means of assessing the robustness of an LVPS at a specific point in time, and, where sufficient data are available, as the basis for conducting on-going surveillance of the extent to which the system is vulnerable to disruption arising from sources of operational risk.

Employing the methodology to CHAPS Sterling (the UK's main LVPS), we find that the system exhibits a reassuringly high level of resilience. In particular, the results of our simulations (carried out using a payment system simulator developed by the Bank of Finland) show that CHAPS Sterling is well-placed to withstand the indirect effects of a variety of plausible, though low-probability, types of operational disruption, and thus that the likelihood of operational risk in CHAPS Sterling acting as a source of financial instability appears to be reasonably small. Nevertheless, a degree of liquidity risk is still potentially present; this is especially the case in situations where an operational incident affects many settlement banks simultaneously (Table 9.2).

To a large extent, our findings in respect of the robustness of CHAPS Sterling reflect the abundance of liquidity available in the system. A further relevant factor is the range of contingency arrangements incorporated into the design of the system (including the possibility to revert to deferred net settlement by invoking RTGS bypass mode). However, operational risk is a constant and evolving threat to the smooth functioning of LVPS, implying that the operators and overseers of such systems should keep the sufficiency of contingency arrangements under constant review; our proposed methodology offers an additional tool for achieving this aim.

References

- Bank of England (2005) **Payment Systems Oversight Report 2004.**
- BIS (1997) **Real-time Gross Settlement Systems.** Bank for International Settlements, Basle.
- BIS (2001) **Core Principles for Systemically Important Payment Systems.** Bank for International Settlements, Basle.
- Bech, M and Soramäki, K (2001) **Gridlock resolution in interbank payment systems.** Bank of Finland Discussion Paper 9/2001.
- Bedford, P, Millard, S and Yang, J (2004) **Assessing operational risk in CHAPS Sterling: a simulation approach.** Bank of England Financial Stability Review, June 2004.
- James, K (2003) **A statistical overview of CHAPS Sterling.** Bank of England Financial Stability Review, June 2004.
- James, K and Willison, M (2004) **Collateral posting decisions in CHAPS Sterling.** Bank of England Financial Stability Review, December 2004.
- McAndrews J and Potter, S (2002) **Liquidity Effects of the Events of September 11, 2001.** Federal Reserve Bank of New York Economic Policy Review, November 2002.
- McAndrews, J and Trundle, J (2001) **New payment system designs: Causes and consequences.** Bank of England Financial Stability Review, December 2001.

Chapter 10

Systemic risk in a netting system revisited

Morten L. Bech – Kimmo Soramäki

10 Systemic risk in a netting system revisited	278
Abstract.....	278
10.1 Introduction.....	278
10.2 Methodology for assessing systemic risk	281
10.3 Payments, positions and exposures.....	282
10.4 Simulation results.....	285
10.4.1 Systemic risk in single bank failures	285
10.4.2 Systemic risk and net debit position of primary failure	287
10.4.3 Systemic risk of multiple bank failures.....	290
10.5 Summary	293
Annex 1/10	294
Annex 2/10	296
References	297

10 Systemic risk in netting system revisited

Abstract

The paper explores the extent of intrinsic systemic risk present in payments originated in the US Fedwire system. We do so by considering the systemic risk if these payments were settled in an unsecured multilateral net settlement system. Our results seem to indicate that systemic consequences in a single bank failure scenario are rather modest. Our second research question related to the common wisdom and the assumption underlying the Lamfalussy standard IV that the largest single net debtor in the system causes the most severe systemic consequences. The simulations showed that for an unsecured net settlement system this does not always hold. This is in fact true for any system where the settlement of all payments is not guaranteed. Finally, our results indicated that the systemic consequences increase surprisingly little as a consequence of multiple simultaneous failures.

10.1 Introduction

Central banks have for the past few decades been concerned about the intraday exposures arising from the settlement of payments. The concerns relate in particular to systemic risk, which in the payment system context can be understood as the impact a failure by a participant in the system may have on other system participants. It is often argued that in a worst case scenario, the losses stemming from exposures taken by participants in the payment system may become contagious and may eventually impede the effective functioning of the payment system or the financial system at large.

During the last decade the risk management techniques used by large-value payment systems have been substantially enhanced, as the risk of contagion in the payment system became better understood. The key developments have been the enhancement of risk management techniques used by the systems themselves, better management of risks by their participants, and the drafting of regulatory standards for payment systems. A major development for payment systems has been the introduction of real-time gross

settlement (RTGS). An RTGS system processes payments in real time on a transaction by transaction basis. Hence, it provides instant finality throughout the business day and thus eliminates intraday credit exposures. Systems that continue to operate on intraday exposures have introduced caps, collateralisation, loss-sharing rules and other risk management techniques to manage and limit intraday exposures. Also payment system participants have enhanced their internal intraday risk management procedures eg by setting bilateral and multilateral limits against their counterparties, and by monitoring their exposures. As a result most systems today employ a wide range of mechanism to mitigate payment system exposures and thereby systemic risk. At the same time the central banks have taken a more active role in overseeing payment systems. The Lamfalussy standards developed by the international Committee for Payment and Securities Settlement (BIS 1990) are a corner stone of risk management standards for payment systems set by central banks. The standard most relevant to systemic risk is number IV. It states that:

“Multilateral netting systems should, at a minimum, be capable of ensuring the timely completion of daily settlements in the event of an inability to settle by the participant with the largest single net-debit position.” (BIS 1990, p. 5)¹

In other words, all systems should be able to withstand the failure of the single largest net debtor without systemic consequences. Central banks generally view this requirement as a minimum standard. Consequently, operators or regulators may impose additional safeguards on a system to ensure that settlement can take place even in the event of multiple failures. Such systems are sometimes referred to as Lamfalussy plus compliant.

The first study on the magnitude of systemic risks stemming from the payments system was conducted by Humphrey (1986). The paper concluded that systemic risk could be a real threat. In the case of a failure of a major participant in CHIPS, the major US private interbank payment system, a high number of other participants would potentially fail. At the time of the paper CHIPS was operating on an unsecured basis, ie in case a participant would fail on its end-of-day

¹ The standard is reiterated with another wording in the more recent Core Principles for Systemically Important Payment System BIS (2001), p. 9. Core Principle V: “A system where multilateral netting takes place should, at minimum, be capable of ensuring timely completion of daily settlements in the event of an inability to settle by the participant with the largest single settlement obligation”.

payment obligation all payments from and to the failing participant would be unwound and the multilateral positions recalculated. More recent studies on systemic risk in payment systems find that systemic risk in payment systems seems to be low. Angelini, Maresca and Russo (1995) followed Humphrey with some modifications and found out that the interbank settlement exposures in the Italian payment system (BI-REL) were much smaller than the results reported by Humphrey on CHIPS. Similar conclusions have been reached by Kuussaari (1996) on the Finnish interbank payment system, Bech et al (2002) for the Danish interbank netting system (PBS) and Northcott (2002) for the Canadian Automated Clearing Settlement System (ACSS).

Surprisingly little work has been carried out to investigate the two key propositions underlying the Lamfalussy standard. First, are the most severe systemic consequences in fact produced by the failure of the single largest net debtor? Second, how severe are the systemic consequences resulting from multiple simultaneous failures, and are they worse than the failure of a single participant?

In this paper we first explore the extent of intrinsic systemic risk present in our data set, payments originated in the US Fedwire system². We do so by considering the systemic risk if these payments were settled in an unsecured multilateral net settlement system. The results of the simulations show the extent of systemic risk that can be removed by using more secure forms of settlement, such as the RTGS mode that is currently employed to settle these payments. We base our results on simulations where one or more banks are set into insolvency at the end of the day and the impact of the failure is propagated through the system. Using the same methodology we go on to tackle both of the above questions. We are particularly interested on the validity of the assumption that the bank with the largest single settlement obligation causes the worst systemic consequences and on the marginal impact of increasing the number of simultaneous bank failures from one.

The paper is organised as follows. Section 10.2 presents our methodology for assessing systemic risk. Section 10.3 explores the data used in the simulations. Section 10.4 presents and section 10.5 discusses and summarises the results.

² Fedwire is the large-value USD interbank payment system operated by the Federal Reserve.

10.2 Methodology for assessing systemic risk

De Bandt and Hartman (2000) define a systemic event as an event where a shock to either a set of financial institutions or markets leads to considerable adverse effects on other financial institutions or markets. A systemic event consists of two parts: the shock and the propagation mechanism. Systemic risk is the possibility of losses following such an event. The degree of systemic risk depends on the likelihood of the event and the consequences of its materialisation.

Here, we take the set of financial institutions to be participants in a hypothetical unsecured end-of-day net settlement system and define a systemic event as the situation where the failure to settle by one or more participants leads to the settlement failure of at least one other participant. We focus exclusively on the impact of a systemic event and do not try to quantify the likelihood of the event. Hence, our results only provide a partial measure of the systemic risk. Our methodology follows Humphrey (1986) and Angelini et al (1996). Nevertheless it differs in several respects and we seek to highlight similarities as well as differences below.

Following Humphrey (1986) and Angelini et al (1996), we consider sudden and unexpected failures of participants. Angelini et al (1996) simulate the failure of every single bank in the system. Humphrey (1986) considers only banks with the largest net credit position on the day of failure. Kuussaari (1996) and Bech et al (2002) simulate the failure of the bank with the largest multilateral net debit position on the particular day. As regards single bank failure scenarios, we consider the failure of each bank with a multilateral net debit position in the system. In contrast to prior studies we also consider simultaneous failures of two and four participants in the system.

In our methodology only banks with a negative revised position vis-à-vis the settlement institution can fail. If a participant has a positive net position vis-à-vis the institution, it does not have a payment obligation and its failure need not affect the settlement process. We assume that the system rules stipulate that whenever a participant fails on its settlement obligation, it is removed from the system. All the payments to and from the failing participant are unwound. After the unwinding of payments, the remaining participants' multilateral net positions are recalculated.

The recalculation of the positions upon the primary failure generally causes some banks' positions to improve and other banks' positions to deteriorate. Humphrey (1986) assumes that a deterioration

in the multilateral net position is a loss to that participant. The amount of these losses depends on the degree by which the participant has permitted its customers to use funds received from the failing participant and the customer's obligation and ability to return these funds upon the failure in the final settlement of the funds. If the customer is obliged, and can return the funds, or if the funds were not available to the customer before final settlement, then the participant does not experience a loss. We take the side of caution and follow Humphrey (1986) in assuming that the full amount is a principal loss to the participant.³ Thus, we implicitly assume a recovery rate of zero for the losses.⁴

If an affected bank has a net debit position and the deterioration of its position exceeds a defined threshold value, we consider that bank to fail on its settlement obligation. It is removed from the system and payments to and from it are unwound. This will again lead to a recalculation of the positions. The process is iterated until no new secondary failures take place. Following Bech et al (2002) we consider a range of different threshold values for the secondary failures. The exact procedure of contagious failures is illustrated via an example in Annex 1.

10.3 Payments, positions and exposures

We use for the simulations interbank payments originated over the Fedwire Funds Service (Fedwire) for January 2003 (21 business days). Participants use Fedwire to handle large-value, time-critical payments, such as payments for the settlement of interbank purchases and sales of federal funds; the purchase, sale, and financing of securities transactions; the disbursement or repayment of loans; and the settlement of real estate transactions. In 2003, an average of 491,158 transfers worth \$1.7 trillion originated over Fedwire per day. The system has more than 9.500 participants.

³ If payments are not irrevocably settled before the failure takes place, the non-failing participants face different demand of liquidity and do not experience a principal loss. In this case banks may still not be able to honor their obligation (ie may fail to settle) because of insufficient liquidity. The same methodology can be applied to such an environment, only the interpretation of the results must be adjusted.

⁴ For comparison, Furfine (2003) uses recovery rates of 60% and 95%. The first rate is reported by James (2001) to be the typical loss in assets of a failing bank. The second rate is the one recovered from the insolvency of Continental Illinois, as reported by Kaufman (1994). Our results thus depict an unlikely but extreme scenario in this respect.

We include in the simulation only transfers, and thus exposures, between depository institutions. Hence, we ignore transfers to, from or on the behalf of the U.S. Government; Federal Reserve Banks; State and local governments; Federally related agencies; and payments and securities settlement institutions such as CHIPS, CLS and DTCC. For computational convenience we limit our sample to the top 1,000 depository institutions in terms of value of transfers originated. These institutions cover 99% of all transfers originated on Fedwire measured in terms of value.

Moreover, we remove all interbank federal funds overnight loans using an approach similar to Furfine (1999). These loans are used by banks mainly to manage their end-of -day positions vis-à-vis the Federal Reserve and hence serve to manage the interbank exposures and end-of-day liquidity positions created by the payment flows.⁵

The depository institutions in our sample exchange payments to the tune of \$1.3 trillion per day. On average, bilateral netting reduces the interbank positions by 76% to \$306 billion and multilateral netting implies a further reduction to \$56 billion or 96%. Both the bilateral and the multilateral netting effects are fairly stable over the period varying within a narrow 3-percentage point range.

Table 10.1 **Descriptive statistics of payment flow, net positions and netting effect**

	Turnover	Daily				Per bank	
		Bilateral Net Position	Multilateral Net Position	Bilateral Netting Effect	Multilateral Netting Effect	Capital	Daily Links
		<u>\$billions</u>				<u>\$billions</u>	
Mean	1,286.1	305.9	56.0	76%	96%	0.6	63
Median	1,259.9	290.5	56.4	76%	96%	0.1	32
Minimum	1,188.9	274.6	41.0	75%	94%	0.001	1
Maximum	1,509.8	366.5	81.3	78%	97%	56.2	893
St. Deviation	91.7	30.1	11.8	1%	1%	3.2	102

Source: Own Calculations

⁵ See also Furfine (2003).

We use Tier 1 capital values from year-end 2002 Call Reports.⁶ For foreign institutions we use a U.S. capital equivalency value.⁷ The average capital per bank is \$600 million but the median is only \$100 million. The largest institution in terms of capital had in excess of \$56 billion whereas the smallest had less than a million.

In principle, we have 499,500 bilateral positions per day.⁸ However, not every pair of depository institutions in Fedwire exchanges payments with each other on a daily basis. In fact, over the sample period there were only between 29,000 and 36,000 non-zero bilateral positions on any given day. The average daily number of non-zero bilateral positions per bank was 63 and the median was 32. However, the distribution of bilateral positions with other banks is highly skewed to the left with 95% of banks having less than 228. The most connected bank had an average of 893 links per day. Distributions of these statistics are provided in Annex 2.

The average (non-zero) bilateral position is \$30,000 and the median is \$300. The distribution of bilateral positions is symmetric around zero and 99% of all positions are \pm \$77 million. However, the largest bilateral position over the sample period was \$14.4 billion. The average bilateral exposure, ie positive bilateral position, is \$10 million but the median is only \$81,000. Less than 5% of bilateral exposures are greater than \$20 million and less than 1% are greater than \$200 million.

The average multilateral net position is by definition zero and the sample median is a mere \$200,000. 99% of the multilateral positions are between $-\$1.2$ and $\$1.1$ billion. The largest position due in the sample is \$44.5 billion and the largest position owed is \$17.5 billion. The average multilateral exposure, ie positive multilateral net position, is \$743 million but the median is only \$43 million. 5% of the multilateral exposures are greater \$3.3 billion and 1% are greater than \$11 billion.

⁶ All banks insured by the Federal Deposits Insurance Corporation (FDIC) are required to file consolidated Reports of Condition and Income (Call Report) as of the close of business on the last day of each calendar quarter.

⁷ If the foreign institution is a financial holding company then the capital amount is 35% of worldwide capital. If the foreign institution has Strength of Support Assessment (SOSA) rating of 1 then the capital amount is 25% of worldwide capital. If the foreign institution has a SOSA rating of 2 then the capital amount is 10% of worldwide capital. If the foreign institution has a SOSA rating of 3 then the capital amount is 5% of worldwide capital.

⁸ $1000 \times (1000 - 1)/2$.

10.4 Simulation results

The results are organised as follows. Section 10.4.1 analyses the magnitude of intrinsic systemic risk. Section 10.4.2 analyses the relationship between the net payment obligation of the primary failure and the resulting systemic risk. Both of these sections are based on simulations where a single bank fails on its payment obligation. Section 10.4.3 studies the impact of multiple simultaneous failures. All results are first presented for a scenario where the failure threshold equals the capital of the banks. Sensitivity of all results is investigated for several lower failure thresholds.

10.4.1 Systemic risk in single bank failures

Our results indicate that systemic risk present in Fedwire payment flows seems to be generally rather low. The vast majority of bank failures did not cause any systemic consequences in the simulations. The number of banks causing secondary failures on a given day ranged between 15 and 30.

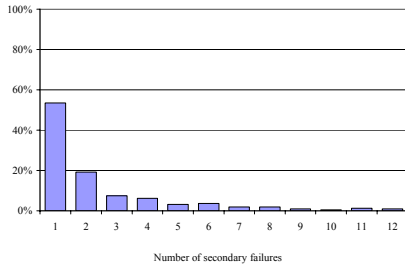
The systemic consequences of the failure of these banks were also generally modest. In over half of the cases the number of contagious failures was limited to a single additional bank. The median total capital of all secondary failures was \$147 million and the median total value of unsettled payments around \$2 billion. In contrast to the total capitalisation of the banks (\$600 billion) and the average daily value of payments (\$1.2 billion) these represent negligible shares. The median losses for all banks were \$100 million. The results are illustrated in Figure 10.1.

On a few occasions, the systemic consequences were higher. In the worst case, the number of secondary failures amounted to twelve banks. While the total capital of secondary failures never exceeded \$5.5 billion (under 1% of total) and the total losses were contained to \$10.4 billion (1.7% of total capitalisation), the value of payments remaining unsettled could grow higher in relative terms. On the worst day the value of these payments was \$143 billion, almost 11% of the day's turnover.

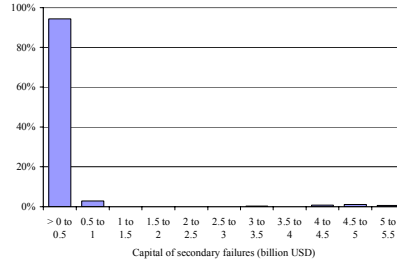
Figure 10.1

Statistics on systemic risk in a single bank failure scenario (% of simulations where systemic consequences – with $\alpha = 1$ – were present)

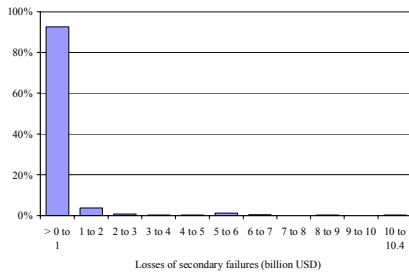
Number of secondary failures



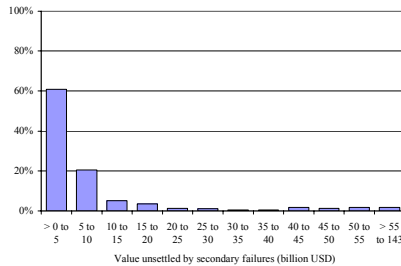
Capital of secondary failures



Losses for secondary failures



Value unsettled



In order to study the sensitivity of the results we carried out several simulations with lower failure thresholds. We calculated the new failure thresholds by multiplying the capital of the participant with a capital scaling factor α taking several discrete values between 0.05 and 1. The scaling factor can also be understood as a measure of the fragility of the system, where the fragility increases with smaller values for α .

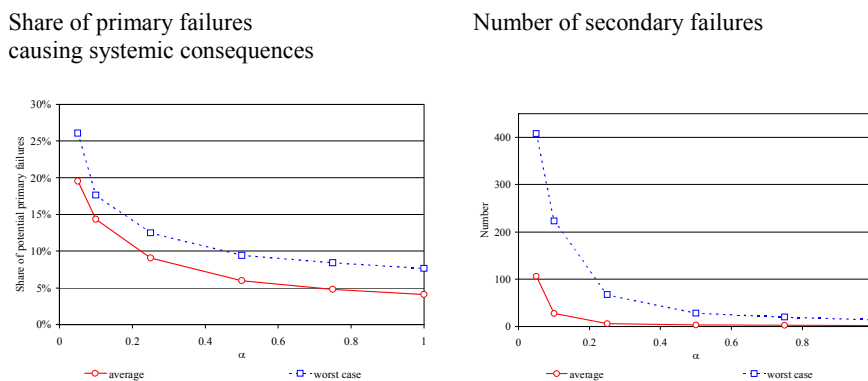
In general, the relationship between the fragility of the system and systemic risk seems to be convex, for both the likelihood of systemic consequences caused by a primary failure (Figure 10.2, left graph) and the resulting number of secondary failures (Figure 10.2, right graph).

When α equaled 0.5 the daily number of primary failures causing systemic consequences ranged between 19 and 37 banks, around double the amount experienced under α of one. Also the magnitude of systemic consequences could grow to be substantially higher. The number of secondary failures was up to 24 banks and the value of

unsettled payments up to \$688 billion. The latter represented almost 46% of the daily turnover. Generally, however, the systemic consequences were not as severe. In two thirds of the simulations the number of secondary failures did not exceed two banks and the value of unsettled payments was in nine out of ten simulations below \$5.5 billion. The simulations, however, show that already at this level of α low probability but high impact scenarios do exist.

The system seems to be rather robust with only slight increases in secondary failures until α equals 0.25. When α is reduced below this ‘tipping point’, both the average and maximum number of secondary failures increase rapidly. When α is reduced to 0.05, systemic consequences are experienced more frequently. Depending on the day, 68 to 139 different primary failures caused secondary failures. The level of systemic consequences was multiplied manifold, and on the worst day over 400 secondary failures were experienced.

Figure 10.2 **Relationship between failure threshold and number of secondary failures**

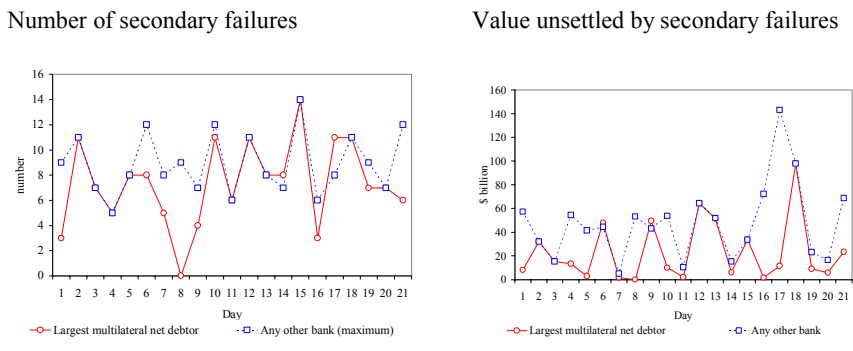


10.4.2 Systemic risk and net debit position of primary failure

In a payment system with unwinding, the initial impact of a failure is driven by the failing participant’s bilateral positions vis-à-vis other participants. Therefore it is not a necessity for the participant with the highest multilateral net debit position to cause the most severe systemic consequences. On 9 out of the 21 days a bank other than the largest multilateral net debtor caused the most secondary failures. In five of these cases the worst impact was caused by the bank with the highest combined bilateral net debit positions. On 3 out of the 21 days

the impact was the same. In terms of unsettled payments, a bank other than the single largest net debtor caused the worst impact on 14 days the. On one of the days, day 8, the failure of the single largest multilateral net debtor did not cause any systemic consequences. The highest number of secondary failures over the whole period was caused both by the single largest net debtor and some other bank. A time series comparing the impact of the failure of the largest net debtor and the highest impact of any other bank is presented in Figure 10.3.

Figure 10.3 **Failure of the single largest multilateral net debtor and systemic risk ($\alpha = 1$)**

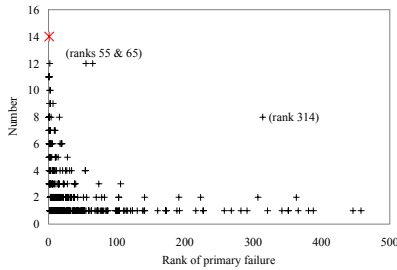


To study the relationship between the banks' multilateral net debit position and systemic consequences, we ranked the banks according to the size of their position on a daily basis. We found that, in general there seems to be a downward sloping relationship between the multilateral net debit position of a bank and the degree of systemic impact, which the failure of a bank causes. Banks with lower multilateral net debit positions are less likely to produce higher systemic consequences. There are, however, important outliers as is evident from Figure 10.4. The failure of banks that were ranked as 2nd, 55th and 65th largest multilateral net debtors in the system produced the second highest number of secondary failures (on days 6, 10 and 21 in Figure 10.3). Also, the failure of a bank with the 314th largest net debit positions on a particular day was still among the top 10 banks causing the worst impact. Looking only at the failure of the largest net debtor might thus not capture the worst case scenario. In our simulations, the primary failure causing the highest value of unsettled payments among the secondary failures was a bank with the fourth largest multilateral net debt position (on day 17 in Figure 10.3).

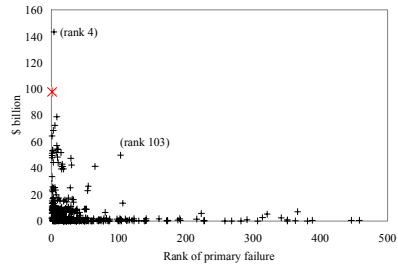
Figure 10.4

Is there a relationship between systemic risk and size of net debit position?

Number of secondary failures



Value unsettled by secondary failures



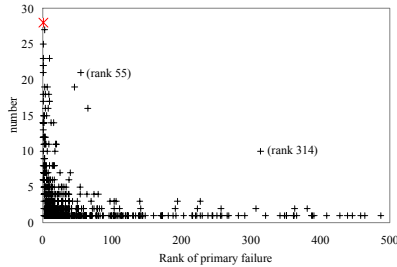
The results are not changed drastically when α is reduced to 0.5. While the failure banks with a higher multilateral net debit positions do not always cause systemic consequences, they are more likely to do so than banks with lower net debit positions. The same is true as regards the value of unsettled payments.

Figure 10.5

Sensitivity of results to lower levels of α

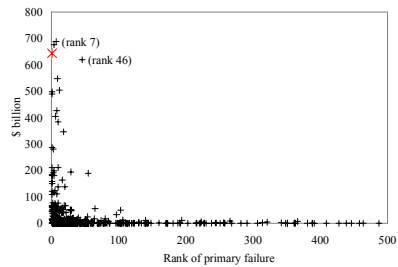
Number of secondary failures

$\alpha = 0.5$

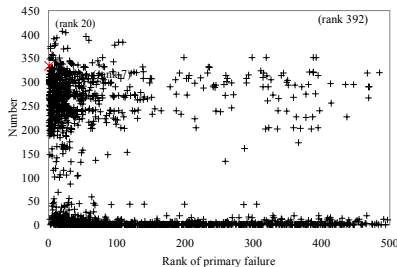


Value of payments unsettled

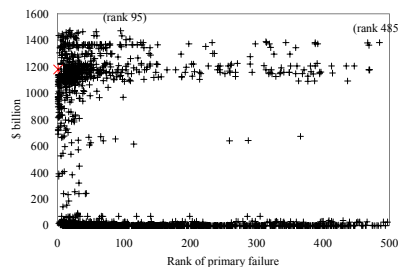
$\alpha = 0.5$



$\alpha = 0.05$



$\alpha = 0.05$



The results change, however, completely when the system becomes extremely fragile ie when α is reduced to 0.05. This is, however, an unrealistically low failure threshold as it implies that losses amounting to over only 5% of the current capital of the banks cause the bank to fail. Under these scenarios virtually the failure of any bank, irrespective of its multilateral net debit position, could cause severe systemic consequences. While banks with higher net debit positions are still more likely to cause a high number of secondary failures, a high number of banks with low net debit positions can do so as well. Also, the system tends to end up in two configurations after the contagion process. Either the systemic consequences are contained to a handful of secondary failures (and a few billion unsettled payments) or around 300 secondary failures (and \$1200 billion of unsettled payments) take place. On rare occasions the number of secondary failures could reach over 400, and almost all of the day's payments could end up being unwound. Moderate systemic consequences, however, were missing.

10.4.3 Systemic risk of multiple bank failures

A sudden and unexpected bank failure is a very rare event. The likelihood of a sudden and unexpected failure of more than one bank is naturally even more remote. It is interesting, however, to analyse how much worse multiple failures can be and what are the dynamics at play when more than one bank is removed from settlement.

We saw in the previous section that the worst systemic consequences on a particular day can be produced by the failure of virtually any bank in the system, especially when the system is very fragile. Likewise virtually any combination of multiple failures is a potential 'worst case' scenario. It is, however, computationally not possible by enumeration to find the set of banks causing the most severe systemic consequences – especially when the number of simultaneous failures exceeds two banks.⁹ To select the combination of banks that produce the worst impact we would therefore need information that is present in the payment and bank data that would tell us which combinations to try. Currently we do not possess such information.

⁹ With 1000 banks the number of combinations for any two banks is approximately half a million, with three banks, 166 million, and with four banks, 41 billion.

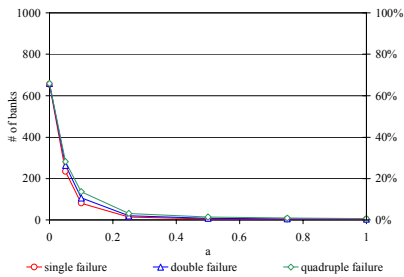
While the bank with the largest multilateral net debit position did not always produce the worst systemic consequences, these banks were more likely to produce severe consequences than banks with lower net debit positions. We therefore select our primary failures from the top 10 banks with the highest multilateral net debit positions. We simulate in addition to a single failure, also the failure of two and four banks. For a given failure threshold this gives us 945 observations (45 combinations for 21 days) for the simultaneous failure of two banks and 4410 observations for the failure of four banks. For comparative reasons the simulations presented here for the single failure scenario are based on the failure of the same 10 banks.

As expected, a higher number of primary failures result in a higher number of secondary failures. The differences are, however, surprisingly small. The curves representing the number of secondary failures at given failure thresholds are very close to each other irrespective of the number of primary failures, especially at low and high levels of α . The number of primary failures seems therefore, to be a less decisive factor for systemic consequences than the failure threshold used. This is true for both the average impact and the worst-case scenario. The value of unsettled payments was, on the other hand, substantially higher in the multiple failure scenarios than in the single failure scenario, especially at high levels of α and in the worst case scenario.

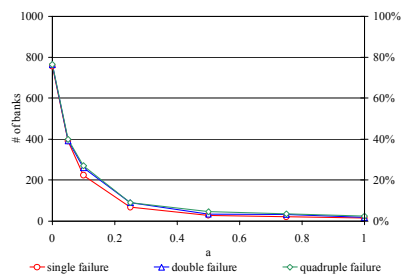
Figure 10.6

The impact of multiple failures

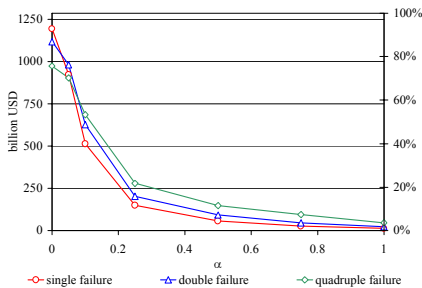
Number of secondary failures (average)



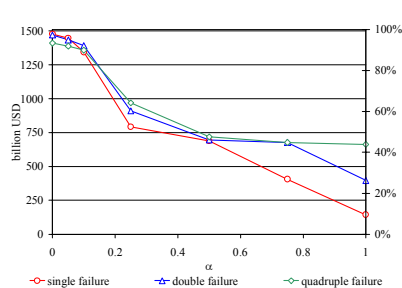
Number of secondary failures (worst case)



Value unsettled by secondary failures (average)



Value unsettled by secondary failures (worst case)



Why do multiple failures not cause more severe systemic consequences than present in the simulations? One possible explanation is that the losses are distributed to a higher number of banks. When looking at the number of banks with a deterioration in their multilateral net debit position caused directly by the primary failure we found that the number of these banks increases substantially when more than one bank fails. While in the single failure scenario an average of 267 banks were affected, the number for double and quadruple failure scenarios was 434 and 465 banks respectively. Another explanation could be that the losses caused by the first failure are offset by gains from a second, third or fourth failure. It could, however, also be that the results are driven by our choice of primary failures. We simulated the failure of the largest banks in terms of their multilateral net debit positions. If these banks transfer heavily payments with each other, a multiple failure could mainly result in exposures between these banks whereas another set of simultaneous failures could be more severe from a systemic risk perspective.

10.5 Summary

Our results seem to indicate that systemic consequences in a single bank failure scenario are rather modest, especially when the failure threshold is set to levels that are more realistic. This does not, however, mean that they can not be severe. Although the results are in line with more recent studies on the topic, results with other data sets or longer time series could be different. Severe contagion may be a low probability but high impact event. In an unsecured net settlement system no limits for exposures exist. Even though the exposures present in our data did not grow high enough to cause widespread disruptions when realised, they can do so – both in other systems and under other operating conditions. The positions between the participants in the system may be more imbalanced under disruptions to the system's or the participants' operations.

Our second research question related to the common wisdom and the assumption underlying the Lamfalussy standard IV that the largest single net debtor in the system causes the most severe systemic consequences. The simulations showed that for an unsecured net settlement system this may not always hold. This is in fact true for any system where the settlement of all payments is not guaranteed. While participants with larger net debit positions are more likely to cause more severe systemic consequences, virtually any other bank can do so as well. These results were very sensitive to the level of failure threshold used. The lower the failure threshold the more this was the case. The assumption underlying the Lamfalussy standard holds, however, for the purpose it was intended to, ie multilateral net settlement systems that guarantee settlement.

Our results indicated that the systemic consequences increase surprisingly little as a consequence of multiple simultaneous failures. As it is not possible via enumeration to simulate all combinations of multiple banks it might be the case that our results do not exhibit the worst-case scenario. In systems without guaranteed settlement the worst initial impact is caused by the bank with the highest combined bilateral net debit positions. The contagion, however, depends on the thresholds for failure by the effected banks and potentially the networks topology – ie the links between the banks in the payment network. One might be able to define better 'fragility' measures both on the system and participant level through this branch of research.

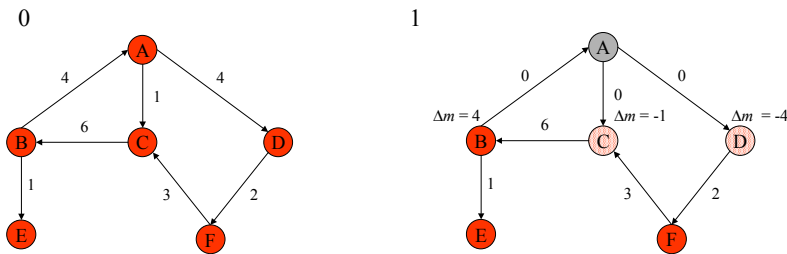
Annex 1/10

Illustration of the methodology

Let us consider the following example to illustrate the methodology and to introduce some terms.

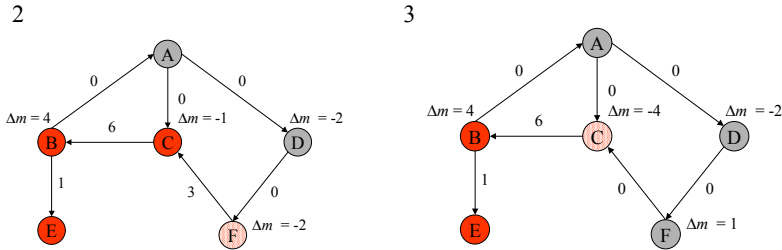
The system consists of six participants that have exchanged payments during the day. The end-of-day bilateral net debit positions vis-à-vis each participant are presented below. The arrows represent the direction of the debt relation, eg Bank A owes four units to Bank D. The threshold values for triggering failures are in the example the following: banks B=C=D=3, and banks E=F=1.

1. We let bank A fail and calculate the changes in multilateral net positions when payments from and to bank A are removed from settlement (unwound). Bank A is the *primary failure*. As this is the first step in the process of contagion, we call this round *generation 1*. As a result of the primary failure, Bank B experiences a positive change (+4) in its multilateral net position, and banks C (-1) and D (-4) a negative change. As discussed before we consider Bank D's loss to equal the negative change in its multilateral net position ($-\Delta m$). Bank C's threshold value for failure (3) is higher than its loss (1) and therefore it continues to participate in the system. The loss of bank D (4), however, exceeds its threshold value (3).



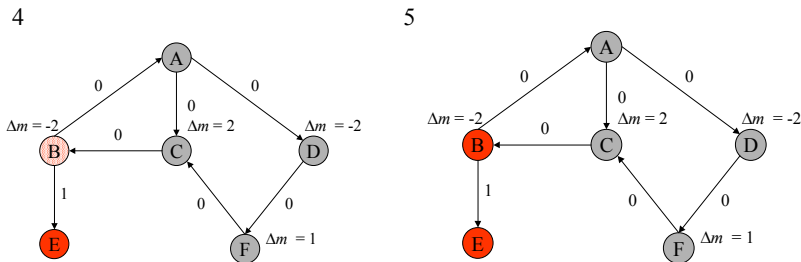
2. As a consequence, Bank D is removed from the system and the multilateral net positions are recalculated. The recalculation of the positions moves the contagion process to the second generation. Bank D is the only direct secondary failure. The failure of bank D causes a negative change (-2) in Bank F's multilateral net debit position.

3. Bank F fails, as its loss is higher than its threshold value for failure ($2 > 1$). Bank F is the first *indirect secondary failure* and the only failure in the third generation.



4. The failure propagates in the system. The deterioration of Bank C's multilateral net debit as a result of the failures of A, D and F is more than its threshold value for failure. Bank C is removed and the positions recalculated for the fourth time. Bank C is the second indirect secondary failure and the third indirect failure in total. The combined deterioration caused by the failure of banks A and C on bank B's positions is two – less than its capital.

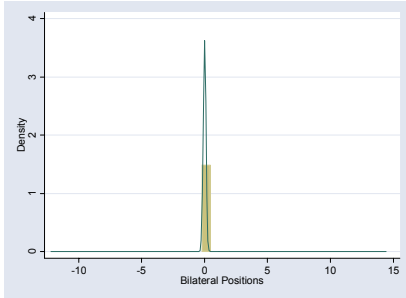
5. The contagion ends at generation five as no new failures take place.



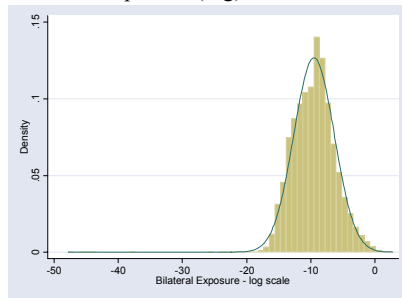
Annex 2/10

Payment statistics (\$ billion)

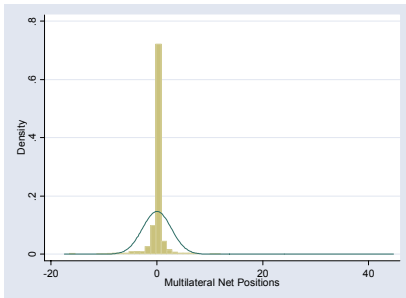
Bilateral Net Positions



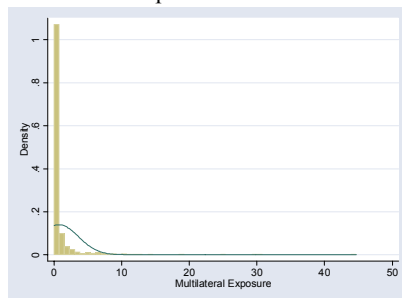
Bilateral Exposure (log)



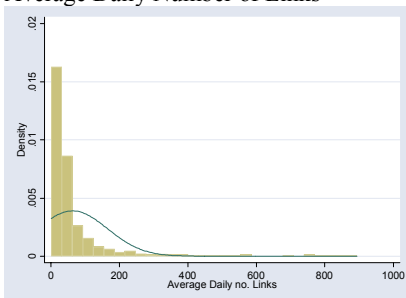
Multilateral Net Positions



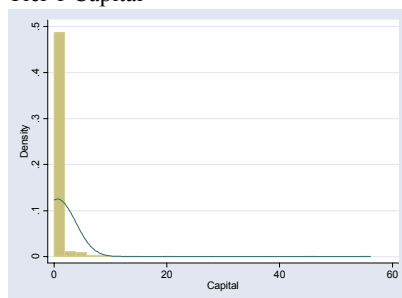
Multilateral Exposures



Average Daily Number of Links



Tier 1 Capital



References

- Angelini, P, Maresca, G and Russo, D (1996) **Systemic risk in the netting system.** *Journal of Banking and Finance* 20, 853–869.
- Bech, M, Madsen, B and Natorp, L (2002) **Systemic Risk in the Danish Interbank Netting System.** Danmarks Nationalbank working paper, No. 8.
- BIS (1990) **Report of the Committee on Interbank Netting Schemes of the Central Banks of the Group of Ten countries.** Prepared by the Committee on Payment and Settlement Systems. No 4. Bank for International Settlements, Basle.
- BIS (2001) **Core Principles for Systemically Important Payment Systems.** Prepared by the Committee on Payment and Settlement Systems. No 43. Bank for International Settlements, Basle.
- De Bandt, O and Hartmann, P (2000) **Systemic Risk: A Survey.** Working Paper No. 35, European Central Bank.
- Furfine, C (2003) **Interbank Exposures: Quantifying the Risk of Contagion.** *Journal of Money, Credit, and Banking* 35. Issue 1, 111–129.
- Humphrey, D (1986) **Payments Finality and Risk of Settlement Failure.** In: Saunders, A, White, L J (Eds), *Technology and the Regulation of Financial Markets: Securities, Futures, and Banking.* Heath, Lexington, MA, 97–120.
- James, C (1991) **The Losses Realised in Bank Failures.** *Journal of Finance* 46, 1223–1242.
- Kaufman, G (1994) **Bank Contagion: A Review of the Theory and Evidence.** *Journal of Financial Services Research* 8, 123–150.
- Kuussaari, H (1996) **Systemic Risk in the Finnish Payment System: An Empirical Investigation.** Discussion Paper 3, Bank of Finland.

Northcott, C A (2002) **Estimating Settlement Risk and the Potential Contagion in Canada's Automated Clearing Settlement System.** Working Paper 2002-41, Bank of Canada.

Chapter 11

Analysis, by simulation, of the impact of a technical default of a payment system participant

Emmanuel Mazars – Guy Woelfel

11	Analysis, by simulation, of the impact of a technical default of a payment system participant	300
	Abstract	300
11.1	Introduction.....	300
11.2	Presentation of PNS	302
11.3	Methodology	304
11.4	Normal functioning of PNS	309
11.5	The default of a participant.....	311
	11.5.1 Without varying bilateral limits	311
	11.5.2 Varying bilateral limits.....	316
11.6	Conclusions.....	321

11 Analysis, by simulation, of the impact of a technical default of a payment system participant

Abstract

Payment systems play a very important role in ensuring the safe and efficient transfer of deposits and financial instruments. Consequently, the failure of these systems may have a destabilising impact. Business continuity plans have thus been developed to ensure their robustness. However, their smooth functioning is also contingent on the capacity of participants to submit their payment orders. The Banque de France, in its role of overseer of the French payment systems, conducted a study with a view to enhancing its understanding of the consequences and the impact of the technical default of a participant in such systems.

This study, carried out using a simulator of the functioning of the Paris Net Settlement (PNS) large-value payment system, operated by the CRI (Centrale des Règlements Interbancaires), shows that the technical default of a participant in this system has negative consequences on the smooth running of the system. Indeed, a situation in which a major participant, in the wake of a technical incident, is unable to submit its payment orders in a normal fashion to its counterparties in PNS, could further exacerbate congestion in the system and result in almost 10% of payments being rejected among non-defaulting participants.

The consequences of a technical default could nevertheless be greatly reduced if the participants set their bilateral sender limits at a lower level than that currently observed and if they reacted rapidly to information indicating a technical default by reducing their bilateral limits with the defaulting participant (defaulter).

11.1 Introduction

Over the past few years, the various parties concerned (supervisory authorities, financial system operators and users) have stepped up efforts to increase the resilience of critical infrastructures to ensure the smooth functioning of systems in the major financial centres. The

different measures taken are part of a single approach that aims notably to ensure that sufficient continuity of service is guaranteed for these infrastructures in the event of a major disruption. Operators of interbank transfer systems (securities settlement systems, payment systems) have undertaken large-scale efforts to improve the soundness of their own infrastructures. However, their smooth functioning is also contingent on the capacity of participants to submit their payments normally.

This study, conducted by the Banque de France, is part of a larger review of operational risk and sets out to better apprehend the consequences and the impact on the functioning of payment systems of the technical default of one of its participants, when it prevents the latter, following an incident (such as the failure of its access platform), from making payments to its counterparties. Analysing such failures is part of the Banque de France's payment systems oversight duties, whose aim is to ensure their security and efficiency in accordance with Article L141-4 of the Monetary and Financial Code.¹ This study focuses on the analysis of the impact of the technical default of a participant in the large-value payment system PNS, using a simulator of the functioning of payment systems developed by the Banque de France. PNS was chosen for this analysis because of its risk management functionalities, which are representative of those of the new generation of large-value payment systems operating on a real-time net settlement basis.

The latter are increasingly being used throughout the world. These type of functionalities can be found in TARGET2, for example, which the Eurosystem has scheduled to go live in 2007, to replace its current network of large-value payment systems linked to TARGET. The results obtained highlighted the fact that the technical default of a major participant in the system would have a substantial impact on transfers between the other participants, but that it would also be possible to reduce these consequences under certain conditions.

Sections 11.2 and 11.3 of this study describe the main characteristics of the PNS system and the methodology used. Section 11.4 illustrates the impact of different parameters on its functioning such as the amount of liquidity submitted by participants or the value of the bilateral limits set by them. The consequences of the technical default of the largest debtor in PNS are analysed in Section 11.5. It shows that the impact of the technical default of a participant may be

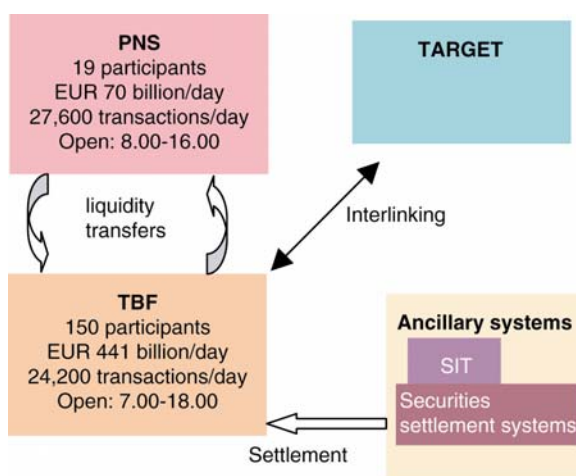
¹ "The Banque de France shall ensure the smooth operation and the security of payment systems within the framework of the tasks of the European System of Central Banks relating to the promotion of the smooth operation of payment systems."

mitigated if the other participants in the system respond appropriately, in particular by setting apposite bilateral limits with the defaulter.

11.2 Presentation of PNS

Paris Net Settlement (PNS) is a large-value payment system operated by the CRI (Centrale des Règlements Interbancaires).² It provides real-time settlement of transactions on central bank money accounts that must always remain in credit.

Figure 11.1 **Paris Net Settlement (PNS) linkages**



*Figures at December 2004

PNS is linked to *Transferts Banque de France* (TBF), the real-time gross settlement system and French component of TARGET operated by the Banque de France. TBF is mainly used for the settlement of monetary policy operations, operations processed by post-market infrastructures (securities clearing and settlement systems), operations processed by SIT (retail payment system), and for urgent payments, and payments to other components of TARGET. PNS is mainly used for the settlement of less urgent large-value domestic payments. Participants start and end the day with zero account balances in PNS.

² The CRI is jointly owned by the Banque de France and eight participating commercial banks.

The link between TBF and PNS enables participants to transfer, without delay, liquidity from their account in TBF to their account in PNS in the way that best suits their cash flow needs: each participant usually makes an initial injection of liquidity into PNS at the beginning of the day and may then add further sums or return liquidity to TBF depending on the nature of its dealings throughout the day. At the end of the day, participants account balances in PNS are automatically transferred back to their accounts in TBF.

The description of the main characteristics of PNS, in particular the terms and conditions of settlement, sheds light on the factors contributing to the fluidity of payments executed via this system, under normal conditions, and on the tools available to participants for managing the effects of a technical default.

Payment orders in PNS are settled in real time if they satisfy a certain number of criteria (balances must remain positive, and the FIFO rule³ and bilateral limits must be observed), or are placed in a queue if these criteria are not met. Queued payments are settled using three different processes that observe the constraints of bilateral limits. The first two also observe the FIFO rule.

- queue scanning, which involves, once an account has been credited, the system checking whether any queued transactions on that account can then be settled,
- bilateral optimisation, which is triggered whenever a payment is added to the queue. The system examines all the queued transactions between the sending participant and the receiving participant, and attempts to process some offsetting transactions simultaneously,
- multilateral optimisation, which is automatically launched twice a day as well as on the initiative of the system operator, and makes it possible to resolve gridlock by checking whether a large number of payments can be simultaneously settled.

By simultaneously settling a number of payments that partially offset each other, the last two mechanisms make it possible to significantly

³ FIFO ('First in First out') means that priority is given to payments according to the order in which they arrive in the system. Nevertheless, a threshold of EUR 1 million exists below which payments bypass the FIFO rule in order to avoid overloading the settlement process. Payments of under EUR 1 million may therefore be settled directly, even if they are submitted after others already in the queue.

reduce the amount of liquidity in central bank money required for the smooth functioning of PNS, compared with a system that only offers gross settlement.

Gridlock

Gridlock is a situation in which several payments cannot be settled individually but can be settled simultaneously. An example of simple gridlock is when three participants A, B, C, all have a liquidity of 10, and A has to make a payment of 15 to B, B a payment of 20 to C and C a payment of 25 to A. In this case, no payment can be settled even though each participant has sufficient liquidity for the simultaneous settlement of all three orders.

Bilateral sender limits are set freely by each participant vis-à-vis its counterparties, thus allowing them to manage liquidity flows and control risks. A bilateral limit is the net amount of money a participant is willing to pay another participant before being paid back. Correct use of such limits enables participants to limit liquidity flows to counterparties withholding payments. There are many reasons why a participant may make late payments. One reason, discussed in this study, would be a technical incident that disrupted the smooth functioning of the participant's access to the system. Another reason could be of a more strategic nature: given that intraday liquidity has a cost, a participant may be tempted to wait to be paid by its counterparties before submitting its payments, so as to benefit from a free ride on the liquidity of others (incoming payments).

11.3 Methodology

This study uses a simulator of TBF and PNS payment systems, developed by the Banque de France in Java script, which almost identically reproduces the functioning of these systems.⁴ The simulations are based on 20 actual days of PNS operation⁵ in January

⁴ In the current version of the simulator it is not possible to change the bilateral limits set by the participants during the day in PNS. These limits are sometimes raised by participants before the closing of the system so that all the queued payments can be settled.

⁵ It should be noted, however, that there are minor differences between the bilateral limits used in the simulation and the real ones, as real limits could not be obtained for all days of transactions.

2004, and are presented in Microsoft Access tables. The simulator reproduces the actual functioning of PNS by processing payments one by one in the same order in which they actually arrive in the days considered. It produces, at the end of each day, a new table showing each payment simulated, the time of settlement (or, as the case may be, whether it was rejected at the end of the day), if the payment was settled in real time or placed in a queue when it entered the system – and, in this case, the reason for it being placed in the queue (exceeding the bilateral limit, insufficient balance, respecting FIFO), as well as the process that enables settlement of the payment (real-time settlement, queue scanning, bilateral optimisation or multilateral optimisation). By choosing a full month of operation, it was possible to smooth the seasonal effects over the month. Indeed, payment flows may differ significantly over the course of the month (in particular in the run-up to the end of the reserve maintenance period), however they remain almost identical from one month to the next. Moreover, January 2004 can be considered to be characteristic of a “normal” month of operation for the system. In total, over 1,200 simulations were carried out, enabling us to test a large number of parameters.

The impact of technical default of one of its participants on the operation of PNS was tested using different scenarios. Technical default is understood to mean the inability of a participant to make its payments to its counterparties in the system in the wake of a technical incident that affects it. It may however continue to receive payments in a normal manner.⁶ There may be many reasons for such an incident including failure of its system access or of upstream applications. In all of the failure scenarios tested, the defaulter could no longer submit payments, but continued to receive them. Consequently, the technical default of a participant is simulated as follows: as of the opening of the system, no payments are submitted by the defaulter but it continues to receive all those made by its counterparties. This is the worst case scenario in terms of the length of the incident and the behaviour of the participants. In terms of length, this scenario assumes that the incident takes place at the opening of the system and is not resolved before the end of the day of transactions. In terms of behaviour, it assumes a maximum “liquidity sink” effect in the system because, in this case, liquidity accumulates on the defaulter’s account while the latter is unable to redistribute it in the system by submitting payments. In practice, we observe that in the rare event of the

⁶ Even though it is likely that this participant would not have tools to view and process the payments made by its counterparties.

technical default of a participant in PNS, the other participants continued to make payments normally to the defaulter, but no incident has ever lasted the whole day. This assumption differs from that of the Bank of England which, in order to study the consequences of the technical default of a participant in its payment system CHAPS Sterling, assumed that payments to defaulter stop ten minutes after the incident occurs.⁷

For each actual day of transactions and each participant, the following values were calculated:

- theoretical lower bound of liquidity (LBL). This is the minimum amount of liquidity that a participant must transfer in PNS for all its payments to be settled. The LBL of participant A is calculated as follows:

$$\sum_X P_{A \rightarrow X} - \sum_X P_{X \rightarrow A}, 0 \quad (11.1)$$

where X represents the counterparties of A and $P_{A \rightarrow X}$ represents all the payments of A to X.

A participant in credit over the whole day therefore has a LBL of zero.

Table 11.1 **Amount of liquidity simulated as a function of α**

α	1	0.75	0.50	0.25	0.10	0
Liquidity (in billions of euro)	7.096	6.732	6.370	6.007	5.789	5.644
Liquidity simulated as a % of AL	100	94.87	89.77	84.65	81.58	79.54

⁷ The Bank of England, Assessing operational risk in CHAPS Sterling: a simulation approach, Financial Stability Review: June 2004. Based on this assumption, the Bank of England set out to determine the point at which (date and time) an incident would have the greatest impact by carrying out simulations using actual transaction data from February 2004.

Table 11.2

Value of bilateral limits simulated as a function of β

α	1	0.75	0.50	0.25	0.15	0.10	0.05	0
Bilateral limits simulated as a % of ABL	100	76.37	52.75	29.12	19.67	14.94	10.22	9.58

- Theoretical lower bound of bilateral limits (LBBL). The LBBL of participant A with participant B is the minimum value of the bilateral limit necessary to settle all payments from A to B:

$$\text{LBBL}(A \rightarrow B) = \max(P_{A \rightarrow B} - P_{B \rightarrow A}, 0) \quad (11.2)$$

Simulations with different levels of liquidity injected into PNS⁸ and bilateral limits were carried out (simulated liquidity levels and bilateral limits were denoted LS and BLS respectively), by varying their values between their “actual” value (ie that observed during the actual days, and denoted AL and ABL respectively) and their theoretical lower bounds. In the rest of this article, the level of liquidity simulated will be represented by an indicator (α), and the level of the bilateral limits simulated by an indicator (β) defined by the following equations:

- level of liquidity simulated (α)

$$\text{LS} = \text{LBL} + \alpha(\text{AL} - \text{LBL}) \text{ with } \alpha \in [0, 1], \quad (11.3)$$

- level of bilateral limits simulated (β)

$$\begin{aligned} \text{BLS} &= \text{LBBL} + \beta(\text{ABL} - \text{LBBL}) \\ &\text{with } \beta \in \sqcup \{\infty\}. \end{aligned} \quad (11.4)$$

Hence, with $\alpha = \beta = 1$, the simulation is based on the functioning of PNS with liquidity levels and bilateral limits equal to their actual

⁸ All liquidity transfers between TBF and PNS were reduced by the same percentage, without changing the times of these transfers. For example if a participant has made the liquidity transfers: EUR 100,000 at 8am and EUR 10,000 at 2pm from TBF to PNS, EUR 20,000 at 3.30pm from PNS to TBF, then a simulation with a ratio of 0.9 will be carried out with the following liquidity transfers: EUR 90,000 at 8am and EUR 9,000 at 2pm from TBF to PNS, EUR 18,000 at 3.30pm from PNS to TBF.

values and with $\alpha = \beta = 0$, the simulation is based on the functioning of PNS with liquidity levels and bilateral limits equal to their theoretical lower bounds. Lastly, $\beta = \infty$ enables us to model the functioning of the system without bilateral limits.

The amount of liquidity simulated in PNS as a function of α is shown in Table 11.1.

The value of the bilateral limits simulated as a function of β is shown in Table 11.2 (as a percentage of the actual bilateral limits).

In order to describe the impact of a default on the overall functioning of PNS, the following indicators were used:

- the delay indicator. The delay indicator δ , which describes the fluidity of the system, is defined as follows:⁹

$$\delta = \frac{\sum_i (t_{\text{settled},i} - t_{\text{sent},i})m_i}{\sum_i (t_{\text{end}} - t_{\text{sent},i})m_i} \quad (11.5)$$

where m_i is the value of the payment i ; $t_{\text{sent},i}$ and $t_{\text{settled},i}$ are the times of submission and settlement of payment i respectively; t_{end} is the time the system closes, ie 4pm. Therefore, as the delay indicator decreases, the fluidity of the system increases: $\delta = 0$ when all the payments are settled immediately, and $\delta = 1$ when all the payments are settled at the end of the day. When modelling a technical default, the delay indicator can be adjusted to solely take account of payments between non-defaulting participants, as we set out to identify the consequences of such a default for the latter. In order to calculate this indicator, the above equation is used, changing only the value range of i , which no longer includes the payments made to the defaulting participants;

- the rejected payments indicator. The number and value of rejected payments provides a better understanding of the risk resulting from the technical default of a participant in the system. A distinction was made between payments rejected between non-defaulting participants and rejections of those made to the defaulter. The number of rejected payments calculated by simulation probably overestimates the outcome of a real case scenario as it is likely that

⁹ This definition of the delay indicator was introduced by Risto Koponen and Kimmo Soramäki in the Article “Intraday Liquidity Needs in a Modern Interbank Payment System. A Simulation Approach”, Bank of Finland, 1998.

some participants would be able to adjust their level of liquidity accordingly, which was not simulated.

A number of other indicators were also used, such as the percentage of payments settled in real time in value and volume terms and the average time payments spend in the queue. Their values were calculated for each day of the month simulated, but were not systematically analysed. These values provide a complementary insight into the underlying factors that affect the main indicators described above.

11.4 Normal functioning of PNS

In order to determine, by comparison, the impact of a default, the characteristics of a standard day of transactions in PNS must be ascertained. So as to better apprehend the normal functioning of the system without taking account of defaulting participants, a number of conditions of liquidity and bilateral limits were simulated, varying α from 0 to 1 and β from 0 to infinity.

Figures 11.2 and 11.3 show the delay indicator (δ) and the rejected payments for different levels of liquidity (α) and bilateral limits (β). In each of the charts, the delay indicator is given on the left-hand scale. The right-hand scale shows the value of the rejected payments as a percentage of the value of all transactions in the system. In addition, the average number of payments rejected daily is given above each point.

The main observation emerging from these simulations is that the system can function smoothly with significantly lower liquidity levels and bilateral limits than their actual values.

This conclusion is supported by the delay indicator, whose upper bound reached $\delta = 0.165$ (for a value of $\delta = 0.09$ with $\alpha = \beta = 1$). The low values obtained for the delay indicator show that PNS is a system that functions broadly in the same way as a gross settlement system. Indeed, even under the most extreme conditions tested, δ remains very far from $\delta = 1$, which would correspond to the functioning of a deferred net settlement system with a settlement at the end of the day.

This observation corroborates the conclusions of a previous study conducted several weeks after the PNS system went live.¹⁰

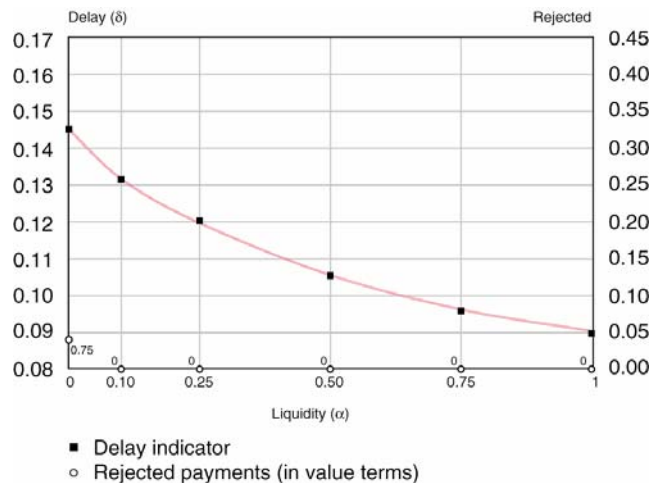
These simulations also show that there were almost no rejected payments, even for low values of α and β . Rejected payments only appeared at values below $\beta = 0.50$; they are characteristic of situations of gridlock that are not resolved by optimisation algorithms (a value of $\beta = 0.50$ should theoretically allow all payments to be settled).

However, participants in the system can change their bilateral limits during the day. We can thus assume that the few payments that were rejected in the simulations would have actually been released in reality by raising the bilateral limits.

Lastly, the use of optimisation mechanisms greatly increases when β declines and more moderately when α declines, which shows that these mechanisms are more responsive to a decline in the bilateral limits than to a decline in liquidity. These mechanisms ensure the smooth functioning of the system in conditions of tighter liquidity or limits.

Figure 11.2

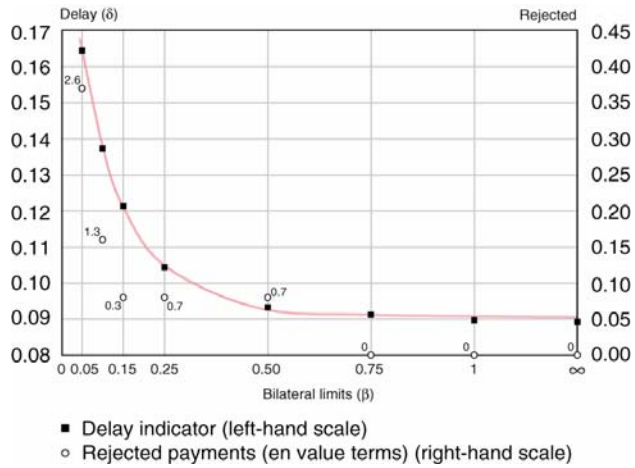
$$\delta = f(\alpha) \text{ with } \beta = 1$$



¹⁰ “Les caractéristiques de fonctionnement des systèmes français de règlement de montant élevé TBF et PNS: quelques enseignements tirés de travaux de modélisation”, Gilles Ryckebusch, Jean-François Ducher and Denis Beau, Banque de France Bulletin No. 71, November 1999.

Figure 11.3

$$\delta = f(\beta) \text{ with } \alpha = 1$$



Figures 11.2 and 11.3 show that almost no payments were rejected (right-hand scale), even at much lower levels of liquidity and limits than their actual values.

11.5 The default of a participant

11.5.1 Without varying bilateral limits

The following analysis focuses on the consequences of the default of the largest debtor in the system, ie the participant with the highest value of total payments over a given day.

The default of the largest debtor was simulated at the start of the day by initially assuming that non-defaulting participants do not change their behaviour vis-à-vis the defaulter. In the model, non-defaulting participants continue to submit payments normally and do not change their bilateral limits with the defaulter. A number of conditions of liquidity and bilateral limits were tested by varying α from 0.5 to 1 and β from 0.05 to infinity.

These value ranges enable us to test all the conditions of bilateral limits in the system, from the lowest to the highest values. A value of

$\beta = \infty$ models the functioning of a system that would not be constrained by the existence of bilateral limits.¹¹

Table 11.3 **Payments settled using optimisation mechanisms** (as a % of queued payments)

	$\beta = 0.05$	$\beta = 0.5$	Actual case	$\alpha = 0.5$	$\alpha = 0$
Bilateral optimisation	60.71	31.00	29.16	31.48	35.56
Multilateral optimisation	8.34	3.50	3.47	3.36	5.57

The value range of α makes it possible to simulate the functioning of the system with an amount of liquidity equal to or below its actual value (ie observed during actual transaction days). It is useful to make simulations with an amount of liquidity below that of actual liquidity because participants generally use the same technical platform to access TBF and PNS. Consequently, it is likely that a participant in technical default in PNS would also be in technical default in TBF, which could result in tighter liquidity conditions in the system and fewer liquidity transfers on the part of non-defaulting participants to PNS and thus a lower liquidity level α .

The impact of a technical default on the functioning of PNS was measured using the three following indicators: the delay indicator δ , the rejected payments between non-defaulting participants (as a percentage of payments between non-defaulting participants, in value and volume terms) and rejected payments made to the defaulter (as a percentage of payments made to the defaulter, in value and volume terms). Figures 11.4, 11.5 and 11.6 show the value of these indicators for the different conditions of liquidity and bilateral limits considered.

In a preliminary analysis, these indicators show that, irrespective of the level of liquidity in the system and the value of the bilateral limits set by the participants, the impact of the technical default of a participant on the functioning of the system and the consequences for the non-defaulting participants are very significant. Indeed, a technical default leads, in all cases, to a considerable increase in the delay indicator and a substantial number of rejected payments between the non-defaulting participants.

¹¹ For technical reasons, due to the simulator's excessively long response time, the value $\beta = 0$, which models the functioning of PNS with the bilateral limits set at their lower bound (LBBL), was not able to be tested. Simulations were carried out with values of β sufficiently close ($\beta = 0.05$) to be able to extrapolate the functioning of the system at this level of limits.

Figure 11.4

$$\delta = f(\alpha, \beta)$$

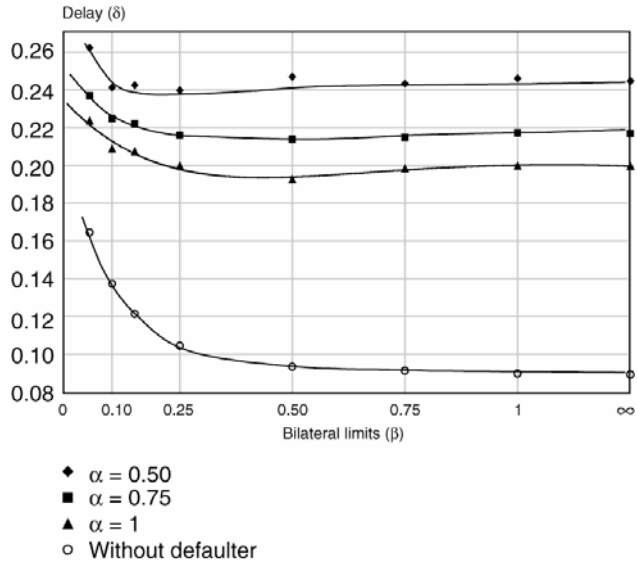


Figure 11.5

Rejected payments to the defaulter
(as a % of payments to the defaulter)

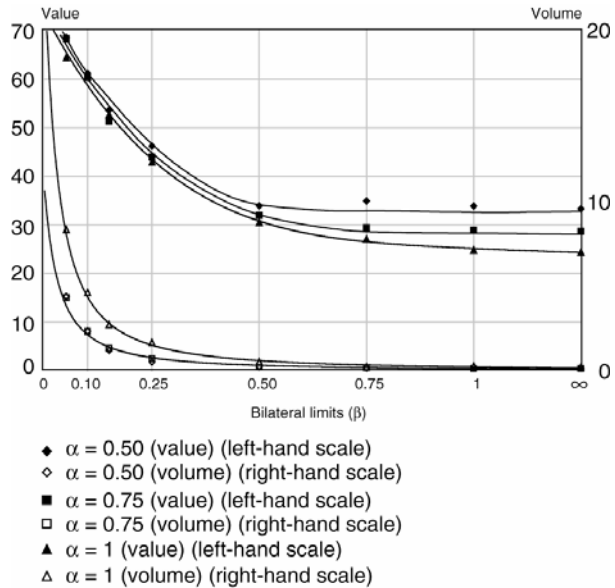
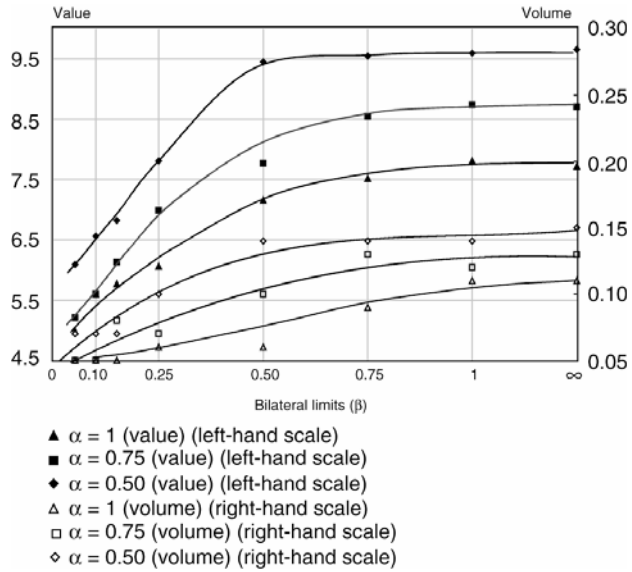


Figure 11.6

Rejected payments between non-defaulting participants (as a % of payments)



Almost 10% (in value terms) of payments between non-defaulting participants could be rejected in the event of a default. This figure is reduced by half at low values of β.

The delay indicator provides information on how smoothly the system is operating as a whole. It rises sharply in the event of a participant defaulting, which means that a greater number of payments are placed in the queue and remain there for a longer time. It also increases when the amount of liquidity in the system (α) declines. This result is intuitive, because a greater number of payments are queued due to insufficient balances. Moreover, the tighter the liquidity conditions in PNS (low values of α), the less sensitive the delay indicator will be to the value of the bilateral limits. This can mainly be ascribed to the fact that when there are liquidity strains in the system, payments are queued because of insufficient balances before the bilateral limits are reached. Lowering these limits therefore has little impact on the functioning of the system.

Table 11.4 **Queued payment data**

Simulations with $\alpha = \beta = 1$	Without defaulter	With defaulter
Queued payments (as a % of total payments)	42.90	63.15
Average time spent in the queue (in seconds)	1,850	2,677

Table 11.5 **Delay indicator as a function of β**

Delay indicator	Without defaulter	With defaulter
$\beta = 1$	$\delta = 0.090$	$\delta = 0.200$
$\beta = 0.05$	$\delta = 0.160$	$\delta = 0.225$

From figure 11.5, it can also be seen that the number and the value of rejected payments to the defaulter are highly dependent on β (the lower the bilateral limits are, the faster those set with the defaulter are reached as the latter does not submit any payments), but not very dependent on the level of liquidity α . If β falls from 1 to 0.05 the proportion (in value terms) of rejected payments rises from 25% to 65%.

The number and the value of rejected payments between non-defaulting participants are very sensitive to the level of liquidity and value of the bilateral limits β . They increase when α declines, and decrease when β declines. Furthermore, sensitivity to β is much greater for low values of this parameter. The following mechanism can be observed: the lower the bilateral limits are, the higher the number of rejected payments to the defaulter. This results in non-defaulting participants losing less liquidity and reduces the “liquidity sink” effect, as these counterparties would then have more liquidity to settle payments between themselves. Figure 11.6 illustrates this mechanism: the number of rejected payments between non-defaulting participants is reduced by almost half when β declines from 1 to 0.05.

Moreover, payments rejected between non-defaulting participants are mainly those of large value. The average value of a rejected payment (EUR 76.1 million, where $\alpha = 1$) is more than 50 times greater than that of the average payment in PNS (EUR 1.5 million).

In the event of a technical default, the value of the bilateral limits has a twofold influence. Firstly, the lower the bilateral limits are, the greater the number of queued payments. This increases the value of

the delay indicator and the number of rejected payments. This mechanical effect was highlighted in the first section of the study. However, low bilateral limits contribute to reducing liquidity flows (liquidity sink effect) to the defaulting participant, which increases the amount of liquidity available for the settlement of transactions between non-defaulting participants. This effect results in a reduction in the delay indicator and in the number of rejected payments among non-defaulting participants.

The second effect, whereby the number and the value decline when β decreases, is a determining factor for explaining rejected payments between non-defaulting participants. Moreover, the combination of the two effects gives an optimum value of β (close to 0.50 for $\alpha = 1$ or 0.75 and close to 0.25 for $\alpha = 0.50$), above and below which the delay indicator shows a higher value.

Table 11.6 **Payments received by the defaulter as a function of β (in EUR billions)**

β	∞	1	0.50	0.25	0.10	0.05
Payments received	6.005	5.972	5.583	4.484	3.113	2.522

11.5.2 Varying bilateral limits

In order to determine the extent to which changes in the behaviour of a defaulter's counterparties are likely to reduce the negative impact on the functioning of PNS, scenarios were carried out in which the bilateral limits of these counterparties vis-à-vis the defaulter were changed.

In the scenario tested, the bilateral limits of the non-defaulting participants vis-à-vis the defaulter were changed immediately (as of the opening of the system), and set at their theoretical lower bound¹². This value enables non-defaulting participants to ensure that all payments to the defaulter could be settled if the latter were also able to submit payments (below this level it would not be possible to meet this condition). However, this value also limits liquidity flows to the defaulter that is unable to submit payments.

¹² It should nevertheless be noted that this exercise remains theoretical because this value is not a priori known by the participants, as they are not necessarily aware of all the payments they expect from their counterparties.

Several simulations were carried out with a liquidity level $\alpha = 1$ and varying the bilateral limits among non-defaulting participants ($\beta = 0.25; 0.50; 0.75$ and 1). No simulations were conducted with both a selective change in limits vis-à-vis the defaulter and values of β below 0.25 . Indeed, these conditions are sufficiently close to those already simulated and presented in Section 11.4.1 of this article ($\beta = 0.05$; no selective change vis-à-vis the defaulter) for the results from the latter to be extrapolated. Furthermore, given the probability that a technical default may also affect the amount of liquidity transferred between TBF and PNS, the study was supplemented by a series of simulations in which the amount of liquidity α was lower than its actual value, $\alpha = 0.75$. This series was tested with a value of $\beta = 0.50$, which, on the basis of the preliminary results obtained for $\alpha = 1$, appeared to be the most appropriate choice.

Figure 11.7

$$\delta = f(\beta)$$

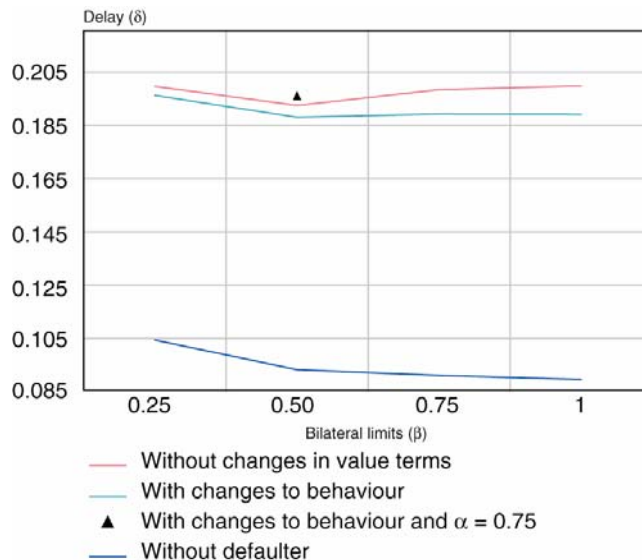


Figure 11.8

Rejected payments to the defaulter
(as % of payments to the defaults)

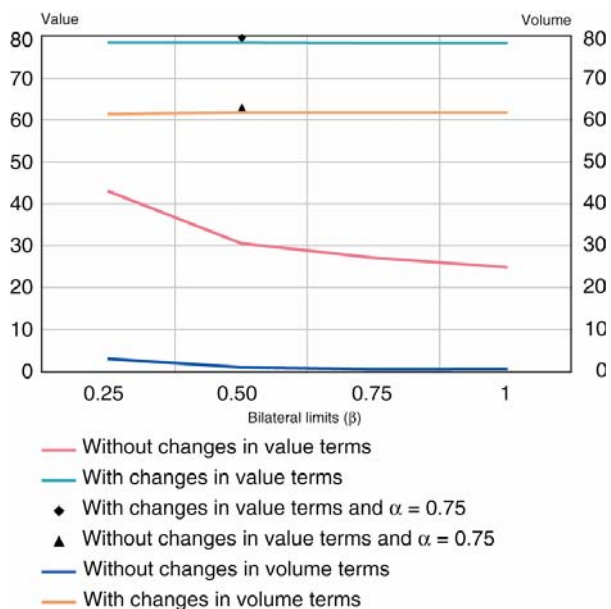


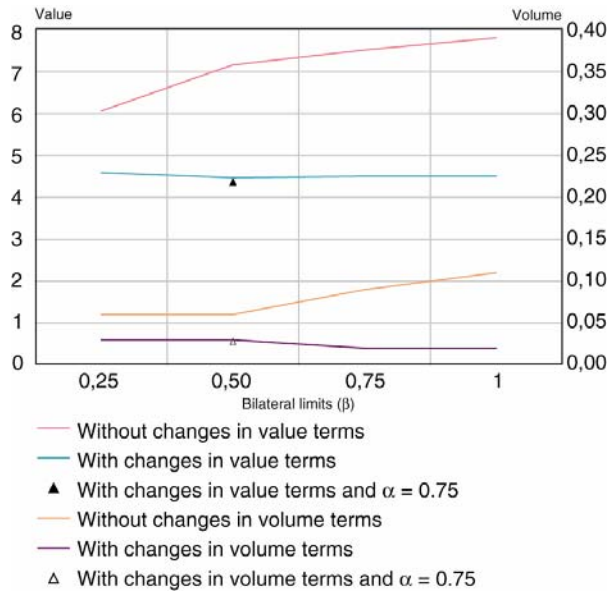
Table 11.7

Reason for rejected payments to the defaulter and payments received by the latter when $\beta = 0.5$
(Rejected payments as a %, payments received in euro)

	Rejected payments due to exceeding limits		Rejected payments due to insufficient balance		Payments received by the defaulter
	In value terms	In volume terms	In value terms	In volume terms	
Without changing behaviour	15.40	33.80	84.60	66.20	5,583,232,111
Changing behaviour	82.30	98.05	17.70	1.95	1,675,232,702

Figure 11.9

Rejected payments between non-defaulting participants (as % of payments between non-defaulting participants)



If non-defaulting participants rapidly lower their bilateral limit vis-à-vis the defaulter, the rejected payments of among these participants (in value terms) are reduced by over 40% when $\alpha = 1$ but remain significant (around 4.5%).

A preliminary analysis of the results obtained (figures 11.7, 11.8 and 11.9) shows that when non-defaulting participants react rapidly and set their bilateral limit vis-à-vis the defaulter at its theoretical lower bound, the consequences of the default are diminished, but remain significant. In fact, the delay indicator is only marginally reduced and the number of rejected payments between non-defaulting participants, although considerably lessened, remains substantial (around 4.5% in value terms).

The decrease in the delay indicator is indeed very slight irrespective of the value of β (around -2% for $\beta = 0.50$), which suggests that setting bilateral limits vis-à-vis the defaulter at their theoretical lower bound only slightly improves the fluidity of the system as a whole.

Furthermore, by a similar mechanism, as when non-defaulting participants do not react selectively, there is an optimum value of β

that minimises the delay indicator. This value is close to $\beta = 0.50$ for a value of $\alpha = 1$.

The impact of setting limits at their theoretical lower bound is more pronounced on the number and the value of rejected payments between non-defaulting participants, which decline from 7.2% to 4.5% in value terms and from 2% to 1% in volume terms for a value of $\beta = 0.5$. They are also relatively insensitive to variations of β .

Similarly to the observation in Section 11.4.1 of this article (without varying bilateral limits), the payments rejected between non-defaulting participants are mainly those of large-value. Indeed, the average value of a rejected payment (EUR 166.2 million, where $\alpha = 1$) is over 110 times greater than that of the average payment in PNS (EUR 1.5 million).

Moreover, the existence of such limits significantly increases the number and the value of rejected payments to the defaulter, as the latter rise from 30% to 80% in value terms and from 2% to 60% in volume terms. These limits therefore result in a substantial reduction of liquidity flows to the defaulter.

The indicators observed are relatively insensitive to small variations in the amount of liquidity α in the system. On the basis of the results obtained with $\alpha = 0.75$, the delay indicator and rejected payments between non-defaulting participants only increase marginally compared with results obtained with $\alpha = 1$.

Table 11.8 **Impact of the threshold below which payments can bypass the FIFO rule and optimisations** (Rejected payments as a % of the total payments)

Simulations with $\alpha = \beta = 1$	δ	Rejected payments to the defaulter		Rejected payments between non-defaulting participants	
		In value terms	In volume terms	In value terms	In volume terms
Normal case	0.1998	24.80	0.44	7.81	0.11
With 2 additional multilateral optimisations at 9.15 am and 2pm	0.2005	25.29	0.45	7.83	0.11
Simulations with a threshold of 500.000	0.1995	24.78	0.45	7.93	0.13
Simulations with a threshold of 100,000,000	0.2100	26.05	0.40	8.64	0.11

Furthermore, for all the indicators concerned, the lower the value of β is, the lower the impact of a selective setting of bilateral limits vis-à-vis the defaulter. This result is intuitive because the lower the value of

β , the closer the case observed will be to that observed in Section 11.4.1 of this article (uniformly low limits).

A number of simulations were also conducted by varying the value of the threshold below which payments can bypass the FIFO rule and by increasing the number of multilateral optimisations. These simulations showed that these two parameters have a relatively low impact on the indicators observed (see Table 11.8).

11.6 Conclusions

The results of the study show that the impact of the technical default of a participant on the fluidity of transactions in the large-value payment system PNS is significant and may disrupt the system's smooth functioning. Such a default results in an almost doubling of the delay indicator and causes payment rejections among non-defaulting participants of up to 10% of total transactions (in value terms). However, an appropriate use of risk management tools enables participants to greatly reduce the impact of a technical default. Preventive actions and a rapid response on the part of the other participants are required. Indeed, as a preventive measure, setting bilateral limits at a lower level than that actually observed in the system makes it possible to reduce liquidity loss in the event of a default, while having an insignificant impact if the system is functioning normally. Moreover, by responding rapidly to information indicating a technical default and adapting their bilateral limits vis-à-vis the defaulter, non-defaulting participants can significantly reduce the impact of such a default on their own transactions. Simulations also showed that there are optimum values for these limits in order to minimise the impact on payments between participants, which are lower than those actually set by participants in the system.

This study confirms the usefulness of the introduction, already planned by the system operator, of a function that makes it possible to submit payments on behalf of the defaulter. This would reduce the risk from such a default by partially redistributing the liquidity "trapped" in the defaulter's account.

These simulations supplement the assessments of PNS carried out by the Banque de France as part of its payment systems oversight duties, in particular by shedding more light on the capacity of the system to function smoothly in the even of "shocks" and highlighting the mechanisms to damp such shocks.

The results obtained for the PNS system make it possible to better understand the consequences that might arise from the technical default of a participant in a system with comparable characteristics, and above all illustrate the importance of the appropriate use of bilateral sender limits in a system with a similar risk management tool. This is the case for TARGET 2, which the Eurosystem has scheduled to go live in 2007, to replace its current network of large-value payment systems linked to TARGET.

Chapter 12

Stress testing securities clearing and settlement systems using simulations

Matti Hellqvist – Jenni Koskinen

12	Stress testing securities clearing and settlement systems using simulations.....	324
	Abstract	324
12.1	Introduction.....	324
12.2	Stress testing of settlement systems.....	326
12.2.1	Creating stress tests	328
12.2.2	Analysing stress test results.....	330
12.2.2.1	Participants' intraday liquidity.....	330
12.2.2.2	Comparing and combining scenario results	331
12.2.2.3	Empirical distribution of scenario impacts	332
12.2.3	Modelling with the BoF-PSS2 simulator	334
12.2.3.1	Executing stress tests.....	335
12.2.3.2	Limitations in modelling with BoF-PSS	336
12.2.3.3	Other applications of settlement system models	338
12.3	Stress testing the RM system.....	338
12.3.1	System-level results.....	342
12.3.2	Participant-level results	346
12.4	Conclusion	348
	References	349

12 Stress testing securities clearing and settlement systems using simulations

Abstract

This chapter presents a methodology for stress testing securities clearing and settlement systems with simulations based on actual system data. The method can be used for assessing the consequences of operational failures of clearing parties within the systems. It can also be used to quantify the possibility of spillover effects – ie contagion between linked systems – resulting from operational failures in securities clearing and settlement systems. As an example, the method is applied to the Finnish bond clearing and settlement system. The study makes use of the Bank of Finland’s payment and settlement system simulator BoF-PSS2.

12.1 Introduction

One of the tasks of the authorities responsible for financial market supervision and oversight is to foster the reliability and stability of the financial system.¹ In line with this, the importance of proper preparation for operational risks and the quantification of said risks has been emphasised in recent years. For individual banks this has been highlighted recently by the recommendations of the Basel Committee on Banking Supervision, which is coordinating the harmonisation of capital adequacy requirements². These recommendations include quantitative indicators for the minimum amount of capital reserved in case of operational risk.

In connection with securities clearing and settlement systems, operational risk is defined as ‘the risk that deficiencies in information systems or internal controls, human errors or management failures will result in unexpected losses.’³ The impact of an operational failure can

¹ In Finland, this is regulated by Section 12.3 of the Act on the Bank of Finland and Section 12.2 of the Act on the Financial Supervision Authority.

² BIS (2003) and (2004).

³ BIS (2001).

be viewed as final losses or as liquidity risk, ie temporary disturbances affecting the liquidity position of market participants during the failure. In a problem situation, payment and settlement systems can be direct channels of contagion through which problems can spread, as the number of open intraday positions between market participants is typically large in these systems. In a worst case scenario, the failure of one participant can cause a chain reaction and a crisis in the entire financial system. This is known as systemic risk.

This chapter presents how the impacts of operative disturbances in securities clearing and settlement systems can be measured using a simulation model and stress tests that reproduce problem situations. Simulations can be used to define how widely and quickly a disturbance spreads in the studied system. Thus the analysis enables the assessment of the systemic risk of a settlement system and of the adequacy of contingency measures in terms of liquidity. The simulation makes use of the Bank of Finland's payment and settlement system simulator BoF-PSS2.

As an example of the method the article reviews the Finnish bond clearing and settlement system. The focus is on assessing the extent of disturbances within the system and the spillover of the impact of disturbances from the settlement system to participants' liquidity on their RTGS accounts. As far as the authors are aware, this is one of the first attempts to stress test securities clearing and settlement system.

Section 12.2 presents background information on simulation modelling and on the stress testing of settlement systems. Subsection 12.2.1 discusses the possibilities of modifying the stress tests for settlement systems, and Subsection 12.2.2 discusses the analysis of stress test results and the background factors that affect this. Section 12.3 discusses the Finnish bond clearing and settlement system and the results of stress testing on the levels of the system and the participants.⁴

⁴ The study described in this chapter was conducted in summer 2004 at the Financial Markets Department of the Bank of Finland, in cooperation with the Finnish Central Securities Depository (APK), as part of an extensive project on the stress testing of financial markets infrastructure. Some of the results were published in November 2004 in the Bank of Finland's Financial Stability Report. The views expressed in this study are those of the author and do not reflect the views of the Bank of Finland.

12.2 Stress testing of settlement systems

The known advantage of simulation modelling is the opportunity it provides to examine systems and experiments, the studying of which would otherwise be too expensive or even impossible. Both these conditions are fulfilled in the context of severe disturbances in securities clearing and settlement systems, because real systems cannot, naturally, be destabilised and historical data on similar events is only rarely available.

By creating a model that simulates the structure and operational logic of the system to be studied, and by adding to it the transaction history of a real-life system, a laboratory setting is created in which the impact of exceptional events can be studied. The analysis is often performed in the form of stress testing, ie by modifying the situation according to a possible disturbance scenario and studying the severity of the impact caused. In settlement systems, typical stress situations involve the removal of one or more clearing parties from the system for varying lengths of time. Creating and justifying different stress scenarios is discussed in Section 12.2.1.

In addition to data based on real transaction histories, simulations can also be run on the basis of artificial transaction data. This requires that the stochastic process which describes the generation of transactions can be identified from the original data. Regarding payment system modelling it has been suggested that a comprehensive model should always include a component concerning the demand for payments, ie the actual exchange of commodities.⁵ In settlement systems, this means modelling the transfer of the traded book entries, but also that the closing of a deal should be described on the level of the final operators – the actual investors – and the individual asset type. In the processed transactions, the clearing party is often merely the trading intermediary, and thus the number of system participants is much higher than the number of clearing parties. The number of asset types, ie various kinds of shares and debt securities, is also high. Therefore even long-term historical data does not give many transactions per final participant and asset type. This makes it harder to create well-founded artificial simulation data and undermines the credibility of simulation results from such data. Hence payment and settlement system simulations are mainly based on real historical data.

One of the key challenges of stress testing is to define how changes in participants' behaviour – due to changes in conditions –

⁵ Zhou (2000).

affect the credibility of results, and whether participants' reactions can be taken into account in the modelling. In economics, this problem is more generally known as the Lucas Critique.⁶ In this case, estimating reaction functions is particularly difficult since the situations assessed are by their very nature rare and unexpected. The easiest solution would be to examine reactions in detail based on real historical data and thereby define the extent of damage a failure would cause if participants did not change their behaviour. This would show the scale of losses participants would be facing and also how quickly they would have to react to prevent these losses.

In examining operative disturbances we must always bear in mind the fact that, by definition, they are unexpected. Thus, the Lucas Critique is directed only towards the phase following the disturbance, as participants would be unable to prepare themselves for the problem in advance by adjusting their behaviour before the disturbance emerged.

Performing an analysis precisely as described in this chapter requires that at least the following facts are known about the settlement system being examined:

- the precise participant and account structure;
- the account balances of all book entries from the beginning of the reference period (incl. fund accounts);
- all events affecting the account balances, specifically:
 - time of transaction,
 - value of transaction,
 - flow through the accounts,
 - information on transaction type;
- the precise logic of clearing and settlement.

Information concerning transaction type reflects the need to separate emissions where new assets are introduced into the system and other exceptional transactions from ordinary trades. In creating stress tests it is also essential to know the agreed practices and contingency measures to be applied in a problem situation.

⁶ Lucas (1976).

12.2.1 Creating stress tests

A variety of scenarios can be created to describe the range of possible of operative disturbances. This section describes a classification of variables that can be used for modifying the scenarios of settlement systems' operative disturbances. The impact of the choices on the modelling process has been explained on a general level. Thus various types of scenario can be created for analysing different systems, depending on the possible weak points of each system and the complexity of the modelling that is going to be performed.

In all the scenarios, results are assessed by comparing the simulation situation of an operative disturbance to a separate reference simulation of normal operating conditions. The most distinct quantity to be assessed is the amount of liquidity held by each participant, ie the liquidity effect caused by the disturbances. Other possible factors that could be examined include changes in the amount of assets held by participants and the impact of this via a reduction in collateral.

Focus of disturbance

One or more of the system participants or the system operator can be subject to an operative disturbance. The latter case can be considered to be more severe, and as such it deserves analysis but requires no simulations. The impact can be obtained by simply calculating chronologically the sum of transactions to be settled per participant and asset type.

Number of affected participants

If the disturbance focuses on the system participants, one or several of these at a time can be made unable to continue normal daily operations in the scenario. The obligations that these participants are unable to meet cause the cancellation of some, or in a worst-case scenario, all of the transaction intended to be settled during the day. A case in which one or a few participants are affected can be justified in that the scenario could be the result of a disturbance which has affected all the participants, for example a computer virus, but only the selected participants have been unable to recover.

Duration of disturbance

Duration of disturbance is the variable that can be modified most in creating a stress situation. In the simplest case, the participant affected by the problem is out of the system for an entire settlement day or at least from the time of the disturbance until the end of the business day. The simplicity of this scenario can be justified by the fact that once the problem is over the unsettled transactions that are still queuing and awaiting settlement would naturally be settled and the system would resume standard operations. If the settlement day were to end before the problem is resolved, the transactions would remain unsettled, which means that the problem would persist for a longer period. As payment and settlement systems operate mainly on a fixed schedule, the end of the business day can also be found in all the other systems that are linked with the settlement system under examination. This enables calculation of participants' liquidity position or end-of-day balance; the latter is a more suitable indicator, as it shows the impact of all the payments settled in the various systems during the day and remains valid until the beginning of the following business day.⁷ The end of the day also has special significance in terms of participants' intraday liquidity management. This aspect is discussed in section 12.2.2.1 below.

In a scenario which ends before the end of the business day, the system can recover by settling the cancelled transactions during the same day, and the impact of the disturbance has to be assessed based on the situation at the time when the disturbance ended, or alternatively as a time series for the duration of the disturbance. A scenario in which the disturbance is short is better because it is more realistic: the duration of the disturbance can be set by assessing how long it would take for participants to resume operations with the help of contingency measures. In such a scenario, the shock can affect each of the participants for different lengths of time, and participants' recovery times can be modelled as random values.

Scope of simulation

The focus of simulation results can be changed by modifying the scope of the stress testing. In the above cases, the settlement system

⁷ The payment systems of various countries differ considerably in this respect. For example, in Finland, the transfer of covering funds in the interbank retail payment system (PMJ) is also executed during the night, in the form of batch runs.

alone is being simulated. This enables assessment of the deviation in the distribution of assets and money among the clearing parties compared to a standard situation. These values can then be weighed against the participants' liquid assets. However, the result does not show whether the parties would have acutely needed the missing assets or money for some other purpose. It would thus be appropriate to extend the modelling of settlement systems to also cover the system supplying the liquidity for settlement and possibly even further, to other linked systems.

On a practical level, extending the scope of simulations would normally proceed to clearing parties' accounts at the central bank (RTGS accounts) and their payment transactions. At the same time the whole RTGS system would have to be included in the simulations. In such a scenario, it would be possible to assess – minute by minute – the impact of an operative disturbance in the settlement system on participants' liquidity position. It would be interesting to study eg the probability of time-critical payments from an RTGS account being delayed or cancelled due to participants' not having received their money from the securities settlement system. Extending the model would require a higher volume of input data and a more complex model structure.

12.2.2 Analysing stress test results

The impact of operative disturbances can be analysed from two viewpoints – on either the system or the participant level. On the level of systems, it is reasonable to examine aggregated data. A good example of this is the total number or value of unsettled transactions, which gives an overview of the number of transactions that remain unsettled or are cancelled, thus causing a deviation in the participants' position compared to a normal situation. Another perspective is to study in detail the impact on participants. In this scenario, the aim is to find out the adequacy of participants' liquidity reserves in a disturbance. Below, the concept of liquid reserves is discussed, as well as the question of how the impacts of disturbances can be converted into relative values which are comparable between participants.

12.2.2.1 Participants' intraday liquidity

Assessing liquidity effects requires knowledge of banks' intraday liquidity management. In Finland, the counterparties for monetary

policy manage their intraday liquidity centrally through their RTGS accounts. Domestic and cross-border interbank payments are made via these accounts, and liquidity is transferred to various settlement systems, eg the Continuous Linked Settlement (CLS) system (a settlement system for foreign exchange transactions) and domestic securities clearing and settlement systems.

The liquidity a central bank provides a commercial bank is typically intraday liquidity and meant for covering the liquidity need caused by the settlement of payments in an RTGS system. Intraday liquidity is provided in the various RTGS systems in different ways. Central banks usually provide account holding institutions with liquidity in the form of interest-free credit, which must be collateralised. In Finland and some other countries in the Eurosystem, this means using pledged securities as collateral. In the United Kingdom, for example, repurchase (REPO) agreements are used. The US Fedwire system uses uncollateralised credit with a low interest rate up to a certain level.

Overnight credit, in contrast, is usually acquired on the interbank market. Overnight credit is used in stabilising the uneven distribution of liquidity among banks. Banks' need for end-of-day liquidity arises, on one hand, due to overnight transfers of covering funds and, on the other hand, due to an obligation to pay back the intraday credit provided by the central bank at the end of the business day. Banks' liquidity position is also affected by the minimum reserve requirement the central bank has defined as part of its monetary policy. Compliance of the requirement is determined typically on the basis of the average overnight balances of RTGS accounts over a fixed period.

If a bank's RTGS account balance remains negative at the end of the day, the marginal lending facility – which corresponds to intraday credit – provided by the central bank can be used to cover the liquidity need. The interest rate on the marginal lending facility is always higher than that of market-based financing. However, in unexpected situations banks may sometimes have to resort to this type of financing.

12.2.2.2 Comparing and combining scenario results

Before comparing and combining the results from different simulated scenarios it may be necessary to scale the results into proportional measures. On the level of participants, the objective is to take into account differences in size and liquidity reserves. On the system level, the objective may be to compare the total volume of liquidity during a

particular day. Converting the liquidity effect into a relative value must be done only after all the possibilities of spillover and contagion between participants and systems have been taken into account in order to get a true picture of participants' exposure to disturbances.

On the participant level, in scenarios that endure to the end of the day, it is natural to compare the liquidity effects on participants' available liquidity reserves. The measure for liquidity reserves used in this study was the RTGS limit, ie the maximum amount of intraday credit or marginal lending facility available to a participant. This can be augmented by the possible end-of-day balance on the participant's RTGS account, which will give the amount of liquidity that is available to the participant quickly and definitely. If there is data available on the participant's other available reserves, such as eligible securities, they could also be used as a benchmark.

The problem in this selection of liquidity measure is that only those participants that are monetary policy counterparties, and therefore have an RTGS limit, can be included in the comparison. Modelling the possibilities of market-based financing in a reliable and comparable way would in this case be a more realistic option, but at the same time also more difficult.

If the simulation has also included transactions that have taken place outside the settlement system, for example on an RTGS account, the liquidity effect can be assessed based on true RTGS account balances at a given time and payments that are delayed due to a shortage of liquidity.

12.2.2.3 Empirical distribution of scenario impacts

The results of stress tests give a better picture than those of individual random tests if a method corresponding to Monte Carlo simulation is applied, ie a method in which the same test is repeated several times and the results are combined to assess the empirical distribution of impacts. In the case of settlement systems, one way of doing this is to aim the selected stress situation on every day of the reference period at each of the participants, one at a time, thereby creating the most extensive sample of different stress situations to be included in the impact distribution. In this scenario, the severity of the impact of each stress situation must be described by an individual value. On the system level this can be eg settled transactions as a proportion of the total volume, and on participant level it can be eg the highest relative or absolute liquidity deficit that any participant has suffered as a result of a disturbance.

The problem in the method presented above is that the distribution of shock impact will include several individual elementary events from the same day if all the disturbances aimed at the various participants are included. In consequence, the elementary events in the distribution are not fully independent. If the data used is extensive enough, this problem can be bypassed by randomly selecting the time of the operative disturbance and the participants at which the disturbance is aimed.

If comprehensive statistics of disturbances in the system are available, they can be used to estimate participants' exposure to disturbances and thereby to define participants' a priori likelihood of experiencing problems. These estimates can be applied for scaling the disturbances caused by the various participants, or directly in randomly selecting the participant at which the disturbance will be aimed. This will give a more realistic impact distribution. If reliable a priori likelihoods are not available, the shock exposure of all the participants can be assumed to be equal. This assumption must be made eg in forming an impact distribution that includes the maximum number of stress tests, as described above.

Knowing the impact values from a large number of scenarios allows estimation of the probabilities for different impact levels in the system under study. Here the different simulated scenarios are assumed to be independent trials with the same distribution for potential impacts. The data set also has to be assumed to cover all significant features of the system under study, eg days with peak values in settlement activity.

Let p_Y denote the real probability of having an impact level less than the fixed value Y in the simulated scenario, and X_Y the number of times when an impact less than Y is actually observed. Now, the value of X_Y follows binomial distribution:

$$X_Y \sim \text{Bin}(n, p_Y) \quad (12.1)$$

where n is the total number of simulated scenarios. An unbiased estimator \hat{p}_Y for the unknown probability p_Y is calculated as frequency

$$\hat{p}_Y = \frac{X_Y}{n} \quad (12.2)$$

A cumulative distribution function can be obtained for the empirical impact distribution by calculating the frequency estimates for all

different observed impact levels Y . Also, confidence intervals can be determined for these estimates with the pivot variable

$$V = \frac{\hat{p}_Y - p}{\sqrt{\frac{\hat{p}_Y(1 - \hat{p}_Y)}{n}}} \quad (12.3)$$

which is known to be asymptotically normally distributed, ie $V \sim N(0,1)$ when equation (12.1) holds.⁸ By solving equation (12.3) and selecting a confidence level of eg 95%, the confidence interval for the estimates can be stated as

$$P\left(p_Y \in \left[\hat{p}_Y \pm 1.96 \sqrt{\frac{\hat{p}_Y(1 - \hat{p}_Y)}{n}}\right]\right) = 0.95 \quad (12.4)$$

Two examples of impact distributions and corresponding confidence intervals are presented in sections 12.3.1 and 12.3.2 below.

12.2.3 Modelling with the BoF-PSS2 simulator

Analysing payment and settlement systems with conventional analytical methods is difficult. This is due to the often complex system structure, which includes a considerable amount of feedback and interaction that affect the position of system participants. An alternative approach would be to use simulation and proceed from the grassroots level of individual events to statistical indicators of the system as a whole.⁹

The BoF-PSS2 simulator (hereinafter referred to as ‘the simulator’) can be characterised as a laboratory, in which risks in and impact on payment and settlement systems caused by various system structures or exceptional events can be studied. The basic structure of modelling is discrete event simulation. The participants, their position, ie assets and funds, and the events that affect their position are included in the input data and the system logic is described piece by piece by algorithms that process various tasks. For example, several alternative algorithms are available for describing the operational

⁸ See eg Råde and Westergren (1998).

⁹ The possibilities of simulation modelling in this area of research are presented in the recent publications Leinonen (2004) and Leinonen and Soramäki (2003).

logic of queues or netting rules.¹⁰ If functionality cannot be described with the existing algorithms, the user of the simulator can add new or modified algorithm modules.

The modelling of settlement systems introduces a few new features into modelling that do not exist in payment system models, the most important of which is the Delivery Versus Payment (DVP) principle. DVP requires simultaneous settlement of the funds and securities linked together. This reduces settlement risk, ie the possibility that one of the parties to a transaction does not receive the payment or asset despite fulfilling its own obligations. Due to the DVP requirement, an individual transaction always involves four accounts, instead of the two involved in payment systems. Therefore, not all the algorithms and functionalities designed for modelling payment systems are suitable as such for modelling settlement systems. In the simulator, the DVP requirement has been met by keeping the funds and asset transactions separate, but adding to partial events a link code that connects them.

The second typical feature of settlement systems is the increased complexity of the account structure to be modelled. This is due to the fact that each asset type, ie the various types of security, have to be described as if they were each a separate currency. This means that a separate account has to be created for each type of security held by every simulated participant, which will increase the number of required accounts significantly.

The third typical feature of settlement systems is the issuance of new assets. These have to be separated from other transactions, because in the simulator an asset cannot be created out of nowhere. Thus, the source of issuances has to be defined as eg a bottomless account separate from the original account structure. A similar situation also arises, but in reverse, with the maturing of debt instruments or when an asset is otherwise removed from the register.

12.2.3.1 Executing stress tests

Stress tests conducted with the simulator are based on a modelling that reproduces the structure of the examined system with adequate accuracy. The constructed model must be validated by comparing its output with the performance of the real system. When adequate correspondence is achieved, a reference situation can be simulated for

¹⁰ Detailed information on the simulator is available at www.bof.fi/sc/bof-pss.

the entire reference period, thus creating a benchmark of a disturbance-free normal situation.

After this, transaction data for an individual stress test is created by taking into account the information structure of the system to be studied and the practices for dealing with a disturbance. The data is formulated by filtering out those transactions which cannot take place during a disturbance. Alternatively, the events that are affected by the disturbance can be transferred to the selected end point of the disturbance. Useful sources in forming a disturbance situation include the rules of the settlement system concerned, the Disclosure Framework¹¹, and possibly the Market Practice Regulations published by the market participants.

If the settlement system's transaction data has initially been presented in a form in which both the purchase price and the transferred assets are specified in the same entity, the entity must be split for entry in the simulator as DVP partial events. Splitting can be easily carried out by adding a link code to the original data and copying the entire data into two parallel versions. One of the versions is then modified into asset transactions and the other into funds transactions by making sure that the selected account structure is duplicated correctly based on the original transaction data.

An analysis which includes a large number of independent stress situations can be run more easily if one transaction data set includes events that describe several stress situations. This can be done with the simulator if the stress tests to be executed simultaneously are separated by at least one end-of-day interlude, as the situation can be reset to normal only by setting the initial balance for the day. Simulating exceptional events that last several days has been possible since the introduction of version 1.2.0 of the simulator.¹²

12.2.3.2 Limitations in modelling with BoF-PSS

Some clear restrictions or problems were detected in the simulator's modelling alternatives for settlement systems, due to the above-described special features of settlement systems. Some have already

¹¹ Disclosure Framework refers to the settlement system operator's responses to a questionnaire of the same name. It is used to improve the awareness of the participants operating in the settlement system of their rights and obligations in the system. See BIS (1997) for the questions.

¹² Version 1.2.0 introduces the possibility of simultaneously transferring balances to the next day and setting initial balances using input data.

been amended or removed during the development of the simulator, and some further improvements are also planned.

The key restriction is the lack of gridlock-resolution algorithms that are able to process DVP data. In payment system modelling, algorithms based on payment splitting or various netting algorithms are used for releasing transactions from queues. In partial netting, a specific subset of transactions waiting in the queue is selected, while in multilateral netting all transactions that are queuing have to be cleared and settled simultaneously. In the latter case, netting refers to computational netting, in which the events to be executed are run separately but simultaneously. In the simulator version 1.2.0, DVP material can be processed only with an algorithm that executes multilateral netting.

In modelling, another restriction concerning netting is caused by the defining of the algorithms' operating range. The simulator enables the modelling of groups of several interlinked but differently operating systems. A good example of this is the parallel simulation of a securities settlement system and RTGS accounts, the advantages of which are discussed in Section 12.2.1, under 'Scope of simulation'. In the simulation of several systems, the linking of transactions between various systems should be avoided even though it is technically possible with the simulator. This is due to the fact that the queue-solving mechanisms, if such are needed, can settle only intrasystem transactions. This means that transactions linked between systems can easily block the queues. This should, however, not be a problem, because cross-system DVP-linked transactions are rare in real systems. For example, Europe currently lacks the technical capacity to execute DVP links in central bank money between central securities depositories.

Problems were also detected in the simulator (at least up to version 1.10) concerning the processing of data including both DVP transactions and normal unlinked transactions. This can be solved, if necessary, by creating dummy transactions separate from the rest of the account and transaction structure as pairs for normal transactions. All processed transactions will then be DVP-linked. Dummy transactions have to be included in the same settlement structure due to the problems inherent in intersystem links (see previous paragraph), and this has to be taken into account in interpreting the simulation results.

The simulator's data structures may currently also be inefficient in processing DVP-linked data. This can be concluded from the fact that the only time the simulation run took longer than the corresponding processing in a real system was when extensive DVP data was

simulated.¹³ Improvements in this area are being planned for version 2.0.0 of the simulator.

12.2.3.3 Other applications of settlement system models

A simulation model of a settlement system can naturally be used for other types of analysis besides quantifying the impact of operative disturbances. For example, the impact of participants' insolvency or shortage of liquidity can be simulated, or the settlement system's functionalities such as the efficiency of liquidity usage can be analysed. These analyses were performed for the settlement system for Finnish money market instruments, which is discussed in section 12.3 concerning only the perspective of operative disturbances. Some of the other results are presented in another publication.¹⁴

The direct analysis of transaction data compiled for the simulation also enables several assessments. For central banks, it is interesting from the perspective of oversight to classify in more detail possible free-of-payment (FOP) transactions or transactions settled with commercial bank money. Another angle is to study participants' operating practices from the perspective of consumer protection, ie whether the assets of an end customer are used for settling the trades of other customers. It is also possible to study participants' practices and risk attitudes by examining at how early a stage trades are entered irrevocably in the settlement system, and whether there are differences in the time of entering depending on the counterparties to the transaction or other circumstances.

In addition to studying the results of stress testing and transaction data, constructing a modelling process that reproduces the structure and functionality of the settlement system can be considered valuable as such because it requires a detailed conceptualisation of the system and thus enables the identification of new risk perspectives and sets of issues.

12.3 Stress testing the RM system

Finnish securities are issued through the Finnish Central Securities Depository Ltd (APK) in book-entry form in a centralised register.

¹³ Bech et al (2004).

¹⁴ Bank of Finland (2004), 81 and 84–85.

The register is divided into two parts: the OM register for equities and the RM system for money market instruments. Both registers have their own settlement system: the OM register has the HEXClear system, and the RM system has the Ramses system. The registers can also be defined so that the OM register covers the retail market in securities and the RM register the wholesale market.

In summer 2004, the Bank of Finland conducted stress tests on the RM register and the Ramses settlement system (both hereinafter referred to as ‘the RM system’). In these tests, the risks inherent in the RM system were studied, as were the possibilities of contagion of problems from the RM system to the other parts of the Finnish financial system in a case where a failure in the RM system was assumed to have happened. The study focused on the liquidity effect on the clearing parties, ie the amount of money the parties do not receive in a disturbance and the impact of this on their RTGS accounts.

For the stress tests, a model was constructed that was as close as possible to the structure of the settlement system. Attention was paid particularly to the account structure and the correct flow of payments. A weak point of the modelling was the lack of an identical gridlock resolution mechanism for solving queues. In the RM system, the gridlock feature called chaining enables simultaneous settlement of independent transactions, which can be described as computational partial netting. In the modelling, chaining was replaced with multilateral netting, and the results may therefore not be as close to the benchmark value as they would have been in a corresponding situation in the real system.

The RM system currently has 15 clearing parties.¹⁵ In 2004, the average monthly nominal value of settled trades was EUR 13.4 billion.¹⁶ The studied data included 2,130 trades, the nominal value of which was EUR 15.9 billion, ie slightly higher than the annual average.

The analysis was performed based on real historical data from a single month (April 2004). The created stress scenarios described fictitious, severe operative disturbances affecting the participants in the settlement system. In the scenario, the participants that were affected by the disturbance were totally unable to operate in the system. The operative disturbance was aimed in turn at each of the clearing parties for one settlement day. As a result, a total of 260

¹⁵ See Ramses Clearing Parties and Account Operators in APK (2004d).

¹⁶ Based on APK statistics. The monthly average for 2004 does not include the data for December 2004.

stress tests could be created.¹⁷ In the stress tests, only the intrasystem transactions of the settlement system were simulated as well as funds transfers between the RTGS account and the settlement system.

The transaction data for April shows that operating as a clearing party is fairly concentrated: in terms of nominal values, four of the largest parties were involved in 73% of the trades as a clearing party of the buyer or the seller. In terms of the number of transactions the corresponding figure is 66%.

Another description of the structure and concentration of the studied market is given below by three Herfindahl indices for daily settlement transactions (figure 12.1). The index is calculated by

$$HI = \sum_{i=1}^n \left(\frac{\text{Transactions of bank } i}{\text{All transactions}} \right)^2 \quad (12.5)$$

where n is the number of parties. The higher the index, the higher the concentration of the settlement transactions. In this case the index can be as high as 1 if all the settlement transactions on one specific day are internal transactions of one clearing party ie between the customers of the one clearing party in question. At its lowest, the index can be as low as $1/n$ if market shares are divided equally among all the parties. In the case examined, the index can thus have values within the range of [0,0667 ... 1].

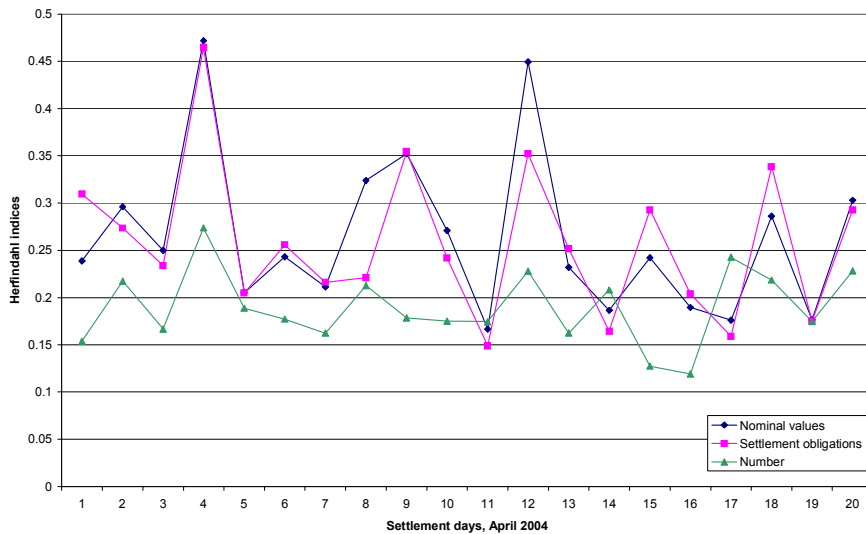
Figure 12.1 shows the Herfindahl index calculated in three different ways, based on:

1. **Nominal values**, ie by including the nominal value of each of the settled asset items in the buyer's and seller's transactions.
2. **Settlement obligations**, ie by registering under the seller's transactions the nominal value of the asset item to be traded, and the purchase price under the buyer's transactions.
3. **Number of trades**, ie by including each settlement in the buyer's and the seller's transactions.

¹⁷ The data included 20 settlement days and the disturbance was aimed at 13 parties. Two clearing parties (Bank of Finland and APK) were excluded from the disturbance scenarios due to their exceptional position.

Figure 12.1

Three Herfindahl indices of the concentration of operating as a clearing party



Source: Bank of Finland.

The highest average of the index, 0.264, is achieved when calculating based on nominal values. The average based on settlement obligations, 0.258, is only slightly lower. In both cases, the value is above 0.25, which would correspond to a situation in which the transactions are divided evenly among four parties. On an individual day, these indices are almost as high as 0.5, ie close to a situation in which the number of parties operating in the settlement system is only two if the parties have an equal share of the market. The index is lower when calculated based on the number of trades. The average value is then 0.190, which corresponds to a situation in which the transactions are divided equally between 5 to 6 parties.

Both the indices, market share of the four largest participants and Herfindahl, show that operating as a clearing party is fairly concentrated. The impact is emphasised further by the fact that the processing of transactions of a high nominal value is even more concentrated on individual parties than the number of transactions. This makes the operative disturbances of small clearing parties less significant by expectation value, but on the other hand, it correspondingly increases the large clearing parties' potential for causing disturbances.

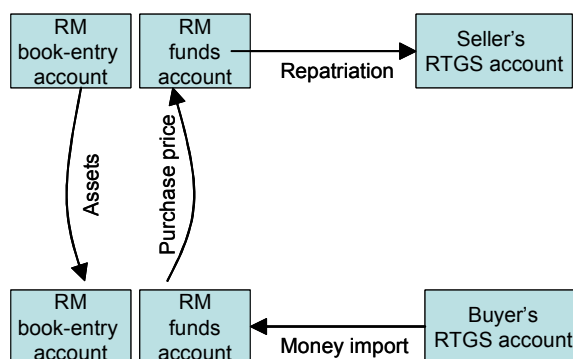
12.3.1 System-level results

The impact of operative disturbances on the system level show how extensively the problems of an individual participant can affect the entire settlement system. To assess the impact it is meaningful to relate the sum of cancelled transactions to eg the total value of transactions of an entire day, or to the size of the settlement obligations of the participant that is affected by the disturbance, which can be considered as the size of the source of the disturbance. If the results are related to the total volume of the day, the impact of even a small participant may be emphasised if the settlement day has been exceptionally quiet. This problem does not arise if the results are related to settlement obligations.

In examining the results, it should be kept in mind that the values describe all the simulated transactions. These include the asset and funds transfers of trades separately, as well as funds transfers from RTGS accounts into the RM system. Therefore the cancellation of transactions worth EUR 1 million (both in nominal value and purchase price) can show in the statistics as EUR 4 million. In such a case, the DVP parts of trades are cancelled: ie asset and purchase price transfers as well as the inflow of the purchase price from the RTGS accounts and the repatriation of the purchase price to the RTGS accounts. This scenario is described below in Figure 12.2. The arrows represent simulated events.

Figure 12.2

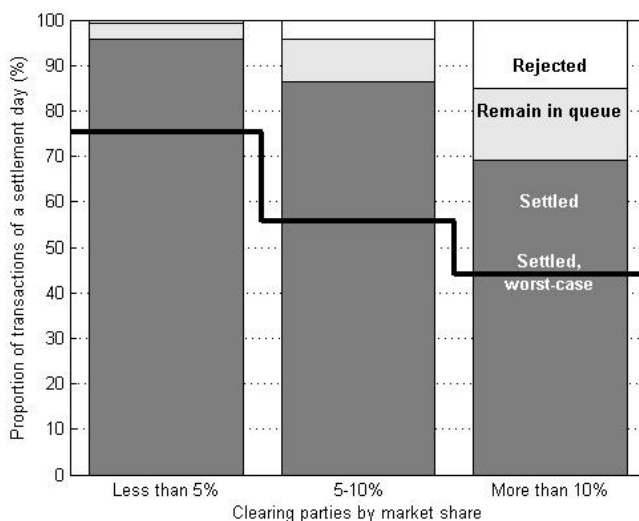
Possible partial events of a settlement transaction in the simulator



On the system level, the results can be examined participant by participant or by market share. Figure 12.3 shows the impact of a disturbance affecting clearing parties in the RM system, broken down into three categories. Market shares have been calculated based on the nominal value of transactions, as in the calculation of Herfindahl indices described above. Each category includes at least three clearing parties. The presented values are averages of corresponding numbers from all individual parties within a category.

Figure 12.3

**Shock impact on the system level:
the proportion of transactions of an entire
settlement day that is settled, remains in the
queue or is rejected after a disturbance
caused by clearing parties of various sizes**



Source: Bank of Finland.

Figure 12.3 shows that in disturbances involving the largest clearing parties, on average 69% of transactions are settled normally, while 16% enter the settlement system but still remain unsettled, ie remain in the queue. The remaining 15% are rejected totally. In one case – the most severe one simulated (it is not shown in the figure as such) – involving a shock situation of one large party, only 20% of the transactions could be settled.

The category of rejected transactions described above includes those settlement transactions the processing of which would have required that the participant affected by the disturbance take measures

during the disturbance. Thus this category also includes trades that were signed only during the simulated disturbance¹⁸ and would therefore not have been executed on the day in question in a real problem situation. The breakdown or exclusion of these types of transaction from the benchmark simulation is, however, not justified because this would inevitably result in an inadequate attempt at modelling the parties' reactions to a disturbance. It is clearer to include all the transactions settled in a real-life situation in the benchmark simulation and to treat all the deviations from this historical data as unexpected disturbances affecting the position of the parties.

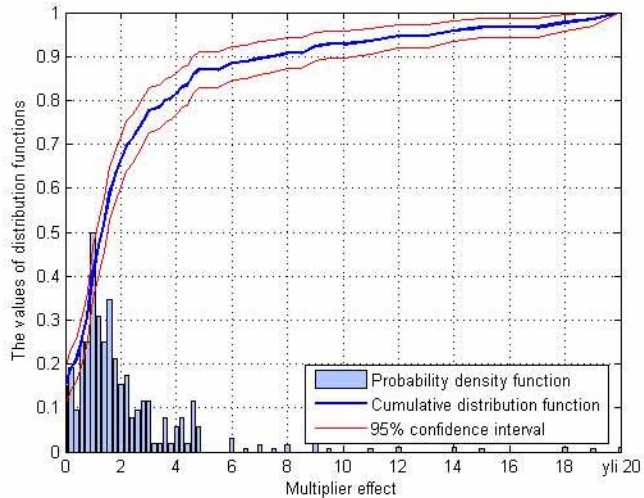
Figure 12.3 shows that small clearing parties, too, can in the worst case cause the cancellation of a large proportion of trades. A disturbance involving a clearing party with a market share of less than 5% could in the worst case result in the cancellation of 25% of transactions. In assessing the results, it should be taken into account that the breakdown by market shares is based on the data for the entire month and not on daily data.

The impact of a disturbance caused by one participant can also be measured by relating the total of all the transactions that remain unsettled in the system as a consequence of a disturbance to the total of the participant's cancelled obligations. This indicator could be called the multiplier effect of cancelled transactions. Figure 12.4 shows an empirical distribution of multiplier effects based on the 260 simulated stress situations.

¹⁸ In the RM system, participants can agree among themselves the settlement day for the trade. Settlement on the actual day of the trade (T+0) is possible, but trades are typically settled only on the second or third settlement day following the trade (T+2 or T+3).

Figure 12.4

Multiplier effects of a disturbance, ie the total of transactions that remain unsettled in the entire system in relation to the settlement obligations of the party causing the disturbance



Source: Bank of Finland.

Figure 12.4 shows that in 15% of cases (39 cases), the disturbance has no impact and the multiplier effect is thus zero. This is possible in a situation in which the party causing the disturbance has no payment obligations and the settlement queue does not include transactions on which the party causing the disturbance should take action. The days on which a party does not have any transactions have not been filtered from the data, because, according to the assumption, the parties' exposure to a disturbance is the same every day, irrespective of their settlement activity.

The highest value for the density function is achieved with a multiplier effect of 1. In practice, this is usually a situation in which the transactions of the party causing the disturbance are not connected in any way with the other trades settled during the day, and therefore they alone remain unsettled. A multiplier effect of 1 or less is achieved in 106 cases, ie in 40.8% of the findings. A multiplier effect of less than 1 is possible in a situation in which a clearing party is involved in transactions and some of the party's settlement obligations can be covered in the settlement with credits received as a result of other trades. The median of the estimated distribution is 1.48: ie half the transactions cause a higher multiplier effect and half a lower one.

Multiplier effects between 1 and 5 are achieved in 120 cases, and in 34 cases the effect is more than fivefold. The highest value was achieved in a case in which the total of cancelled transactions was 451-fold in relation to the payment obligations of the party causing the disturbance. This is possible if small missing sums cause higher-value trades to gridlock and remain in the queue.

12.3.2 Participant-level results

On the participant level, the results of stress tests can be examined from two perspectives.

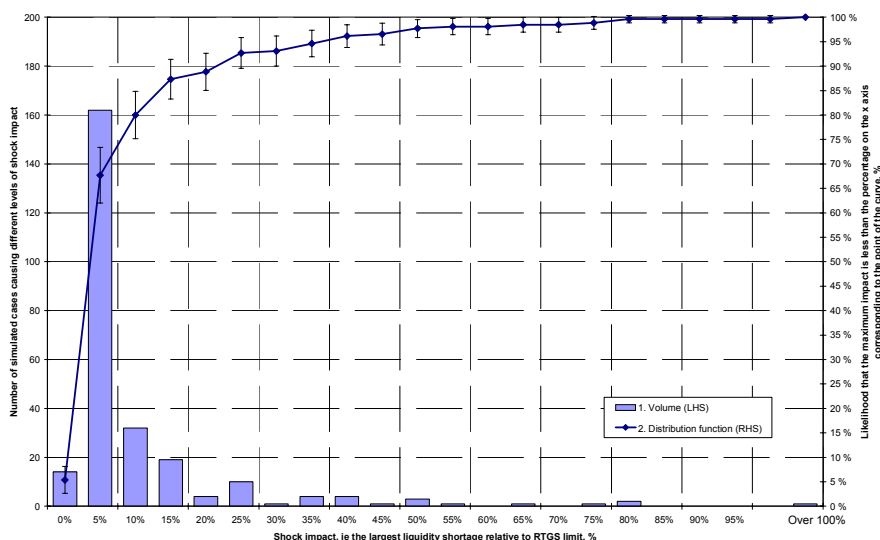
1. The extent of disturbance each party can in the worst case cause itself and other clearing parties, ie the worst effect among all the clearing parties from the scenario currently being tested. This is measured by an indicator, eg liquidity deficit.
2. The extent of disturbance each party can be affected by in a worst case scenario.

By taking both these perspectives into account, the need for and adequacy of the participants' contingency measures can be assessed.

Taking the latter perspective, the level of the liquidity effect caused by a single operative disturbance was measured based on the largest liquidity shortage relative to the RTGS limit caused to any counterparty for monetary policy participating in the settlement system. This indicator was defined as *shock impact*, and a value for it was established in all 260 stress tests (13 parties on 20 days). By relating the liquidity effect to the RTGS limit the shock impact was made comparable between parties of different size. As a side effect, those clearing parties that do not have an RTGS account at the Bank of Finland had to be excluded from the assessment. The chart below shows the distribution of shock impact in the form of a histogram and a cumulative distribution function.

Figure 12.5

Impact of an operative disturbance on the liquidity of the counterparties for monetary policy: empirical distribution



Source: Bank of Finland.

The diagram in figure 12.5 shows that the distribution of shock impacts is weighted heavily towards smaller impacts of less than 5%. The median of observed shock impacts is 1.616%, ie in half of the cases the impact is less than that, whereas in 75% of the cases studied the shock impact was less than 8.03%, and in 95% of cases less than 35.01%. The value of the estimated cumulative distribution function for shock impacts of less than 35% is $94.6 \pm 2.7\%$. The values of the distribution function and their confidence intervals are calculated using equations (12.2 and 12.4) described in section 12.2.2.3.

As an example, the probability of the shock impact being less than 15% can be calculated if we know that in 227 of the 260 cases the value of the relatively largest liquidity deficit is less than 15%. The estimated probability is $\hat{p}_{15\%} = \frac{227}{260} = 0.83077$, and the confidence interval for the estimate is $[0.831 \pm 0.040]$.

The key assumption in calculating the above empirical distribution of shock impact is that the likelihood of facing problems and causing an operative disturbance is the same for all the clearing parties on all

the days of the sample period.¹⁹ In reality, there are differences in the parties' contingency measures and ability to recover from disturbances. The contingency measures of larger parties can be assumed to be more extensive based merely on eg the higher number of people operating the system. Therefore the number of large shock impacts may be higher in the results than in a real-life situation. If reliable and comparable long-term statistics on the parties' level of contingency measures or shock exposure were available, the information could be used in assessing the a priori probability of parties being affected by a disturbance.

12.4 Conclusion

Examining securities clearing and settlement systems by using stress tests and simulation enables the assessment of risks associated with operative or other disturbances. Stress tests can be used to define the extent of the assumed disturbances within systems and the strength of their impact on the other financial market systems. The impact can be examined on the system level or by studying the position of individual parties.

The main challenge in using the method described in this chapter is to compile data that is comprehensive and describes the studied systems in sufficient detail. Key areas for improvement are to make the participant-level results more realistic by using improved modelling of the parties' overall liquidity position and to develop the modelling of participants' reactions in a disturbance, which is a bigger challenge.

The Bank of Finland's payment and settlement system simulator was also found to be a useful tool for studying settlement systems, despite the fact that some of its features clearly require further development. The most necessary enhancement would be gridlock resolution algorithms for DVP data with partial netting.

Stress testing the settlement system for Finnish money market instruments shows that the system does not seem to pose a significant systemic risk to the Finnish financial market. Even in a severe disturbance affecting one participant the liquidity effect on the other participants was for the most part minor when taking into consideration the participants' liquidity reserves.

¹⁹ All 260 cases were included. The data therefore includes days on which the party at which the operative disturbance is aimed had no transactions in the RM system.

References

- APK (2004a) **Book-entry System Historical Data as at 30.11.2004.** Finnish Central Securities Depository, <http://www.apk.fi/English/Statistics.htm>.
- APK (2004b) **Disclosure framework 2004.** Finnish Central Securities Depository, <http://www.apk.fi/English/SettlementServices/ShareTrade/DisclosureFramework.htm>.
- APK (2004c) **Rules of The Finnish Central Securities Depository Ltd.** Finnish Central Securities Depository, http://www.apk.fi/English/Regulation_Rules_of.htm.
- APK (2004d) **Ramses clearing parties and account operators.** Finnish Central Securities Depository, <http://www.apk.fi/English/SettlementServices/ClearingParties/Operator.htm>.
- Bank of Finland (2004) **Financial stability.** Bank of Finland Bulletin, special issue 2004.
- Bech, M, Johnson, K and Soramäki, K (2004, forthcoming) **Simulating the Fedwire Securities Service System.** Ongoing study presented at the 2nd payment and settlement simulation seminar and workshop, 26–27 August 2004 at the Bank of Finland.
- BIS (1997) **Disclosure framework for securities settlement systems.** Committee on Payment and Settlement Systems, Bank for International Settlements, Basel.
- BIS (2001) **Recommendations for securities settlement systems.** Committee on Payment and Settlement Systems and Technical Committee of the International Organisation of Securities Commissions, Bank for International Settlements, Basel.
- BIS (2003) **Sound Practices for the Management and Supervision of Operational Risk.** Basel Committee on Banking Supervision, Bank for International Settlements, Basel.

- BIS (2004) **International Convergence of Capital Measurement and Capital Standards**. Basel Committee on Banking Supervision, Bank for International Settlements, Basel.
- Laki Suomen Pankista** (Act on the Bank of Finland) 27.3.1998/214, Finlex.
- Laki Rahoitustarkastuksesta** (Act on the Financial Supervision Authority) 27.6.2003/587, Finlex.
- Leinonen, H (2004) **Simulointimallit maksu- ja selvitysjärjestelmien uutena tutkimusvälineenä**. Kansantaloudellinen aikakauskirja. 2/2004, 154–161.
- Leinonen, H and Soramäki, K (2003) **Simulating interbank payment and securities settlement mechanisms with the BoF-PSS2 simulator**. Discussion Paper 23/2003, Bank of Finland.
- Lucas, R E (1976) **Econometric Policy Evaluation: A Critique**. Carnegie-Rochester Conference Series on Public Policy, 1, 19–46.
- Råde, L and Westergren, B (1998) **Mathematics handbook for science and engineering**. Studentlitteratur.
- Zhou, R (2000) **Understanding Intraday Credit in Large-Value Payment Systems**. Federal Reserve Bank of Chicago, Economic Perspectives XXIV, Third Quarter, 29–44.

Publications of the Bank of Finland

Series E (ISSN 1238-1691, print) (ISSN 1456-5951, online)

(Series E replaces the Bank of Finland's research publications series B, C and D.)

- E:1 Jukka Vesala **Testing for Competition in Banking: Behavioral Evidence from Finland.** 1995. 206 p. ISBN 951-686-447-3.
- E:2 Juha Tarkka **Approaches to Deposit Pricing: A Study in the Determination of Deposit Interest and Bank Service Charges.** 1995. 166 p. ISBN 951-686-457-0.
- E:3 Timo Tyrväinen **Wage Determination, Taxes, and Employment: Evidence from Finland.** 1995. 212 p. ISBN 951-686-459-7.
- E:4 Sinimaaria Ranki **Realignment Expectations in the ERM: Causes and Measurement.** 1996. 164 p. ISBN 951-686-507-0.
- E:5 Juhana Hukkinen **Kilpailukyky, ulkomaankaupan rakenne ja taloudellinen kasvu** (Competitiveness, structure of foreign trade and economic growth). 1996. 134 p. ISBN 951-686-512-7.
- E:6 Eelis Hein **Deposit Insurance: Pricing and Incentives.** 1996. 120 p. ISBN 951-686-517-8.
- E:7 Vesa Vihriälä **Banks and the Finnish Credit Cycle 1986–1995.** 1997. 200 p. ISBN 951-686-537-2.
- E:8 Anne Brunila **Fiscal Policy and Private Consumption-Saving Decisions: European Evidence.** 1997. 147 p. ISBN 951-686-558-5. (Published also as A-131, Helsinki School of Economics and Business Administration, ISBN 951-791-225-0, ISSN 1237-556X)
- E:9 Sinimaaria Ranki **Exchange Rates in European Monetary Integration.** 1998. 221 p. ISBN 951-686-564-X.
- E:10 Kimmo Virolainen **Tax Incentives and Corporate Borrowing: Evidence from Finnish Company Panel Data.** 1998. 151 p. ISBN 951-686-573-9. (Published also as A-137, Helsinki School of Economics and Business Administration, ISBN 951-791-290-0, ISSN 1237-556X)
- E:11 Monica Ahlstedt **Analysis of Financial Risks in a GARCH Framework.** 1998. 181 p. ISBN 951-686-575-5.
- E:12 Olli Castrén **Fiscal-Monetary Policy Coordination and Central Bank Independence.** 1998. 153 p. ISBN 951-686-580-1.
- E:13 Antti Ripatti **Demand for Money in Inflation-Targeting Monetary Policy.** 1998. 136 p. ISBN 951-686-581-X.
- E:14 Risto Koponen – Kimmo Soramäki **Intraday Liquidity Needs in a Modern Interbank Payment System. A Simulation Approach.** 1998. 135 p. ISBN 951-686-601-8.

- E:15 Liisa Halme **Pankkisääntely ja valvonta. Oikeuspoliittinen tutkimus säästöpankkien riskinotosta** (Banking regulation and supervision: A legal policy study of risk taking by savings banks). 1999. XLIV + 560 p. ISBN 951-686-606-9, print; ISBN 951-686-607-7, online.
- E:16 Juha Kasanen **Ilmoitusvelvollisten osakeomistus ja -kaupat Helsingin Pörssissä** (Corporate insiders shareholdings and trading on the HEX Helsinki Exchanges). 1999. 146 p. ISBN 951-686-630-1, print; ISBN 951-686-631-X, online.
- E:17 Mikko Spolander **Measuring Exchange Market Pressure and Central Bank Intervention**. 1999. 118 p. ISBN 951-686-645-X, print; ISBN 951-686-646-8, online.
- E:18 Karlo Kauko **The Microeconomics of Innovation: Oligopoly Theoretic Analyses with Applications to Banking and Patenting**. 2000. 193 p. ISBN 951-686-651-4, print; ISBN 951-686-652-2, online. (Published also as A-166, Helsinki School of Economics and Business Administration, ISBN 951-791-442-3, ISSN 1237-556X)
- E:19 Juha Kilponen **The Political Economy of Monetary Policy and Wage Bargaining. Theory and Econometric Evidence**. 2000. 180 p. ISBN 951-686-665-4, print; ISBN 951-686-666-2, online.
- E:20 Jukka Vesala **Technological Transformation and Retail Banking Competition: Implications and Measurement**. 2000. 211 p. ISBN 951-686-695-6, print; ISBN 951-686-696-4, online. (Published also as A-184, Helsinki School of Economics and Business Administration, ISBN 951-791-518-7, ISSN 1237-556X)
- E:21 Jian-Guang Shen **Models of Currency Crises with Banking Sector and Imperfectly Competitive Labor Markets**. 2001. 159 p. ISBN 951-686-711-1, print; ISBN 951-686-712-X, online.
- E:22 Kari Takala **Studies in Time Series Analysis of Consumption, Asset Prices and Forecasting**. 2001. 300 p. ISBN 951-686-759-6, print; ISBN 951-686-760-X, online.
- E:23 Mika Kortelainen **Edge: a model of the euro area with applications to monetary policy**. 2002. 166 p. ISBN 952-462-001-4, print; ISBN 952-462-002-2, online. (Published also as A-204, Helsinki School of Economics and Business Administration, ISBN 951-791-715-5, ISSN 1237-556X)
- E:24 Jukka Topi **Effects of moral hazard and monitoring on monetary policy transmission**. 2003. 148 p. ISBN 952-462-031-6, print; ISBN 952-462-032-4, online.
- E:25 Hanna Freystätter **Price setting behavior in an open economy and the determination of Finnish foreign trade prices**. 2003. 84 p. ISBN 952-462-045-6, print; ISBN 952-462-046-4, online.
- E:26 Tuomas Välimäki **Central bank tenders: three essays on money market liquidity auctions**. 2003. 232 p. ISBN 952-462-051-0, print; ISBN 952-462-052-9, online. (Published also as A-218, Helsinki School of Economics, Acta Universitatis Oeconomicae Helsingiensis, ISBN 951-791-762-7, ISSN 1237-556X)
- E:27 Heikki Hella **On robust ESACF identification of mixed ARIMA models**. 2003. 159 p. ISBN 952-462-112-6, print; ISBN 952-462-113-4, online.

- E:28 Heiko Schmiedel **Performance of international securities markets**. 2004. 275 p.
ISBN 952-462-132-0, print; ISBN 952-462-133-9, online.
- E:29 Tuomas Komulainen **Essays on financial crises in emerging markets**. 2004. 173 p.
ISBN 952-462-140-1, print; ISBN 952-462-141-X, online.
- E:30 Jukka Vauhkonen **Essays on financial contracting**. 2004. 134 p.
ISBN 952-462-172-X, print; ISBN 952-462-173-8, online.
- E:31 Harry Leinonen (ed.) **Liquidity, risks and speed in payment and settlement systems – a simulation approach**. 2005. Compilation. 350 p. ISBN 952-462-194-0, print;
ISBN 952-462-195-9, online.

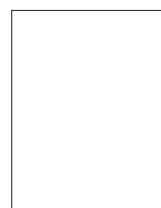
**Suomen Pankki
Bank of Finland**

PO BOX 160, FI – 00101 HELSINKI, FINLAND

Tel +358 10 8311, fax +358 9 174 872, email publications@bof.fi

Subscriptions to the E series of of the Bank of Finland's research publications and changes in address details

Old address details
Company
Name
Address
New address details/subscriptions
Company
Name
Address
New subscription <input type="radio"/> Cancellation <input type="radio"/> Number of copies



SUOMEN PANKKI
BANK OF FINLAND

PO BOX 160
FI – 00101 HELSINKI
FINLAND

ISBN 952-462-194-0
ISSN 1238-1691

Edita Prima Oy
Helsinki 2005