DISCUSSION PAPER SERIES

No 1667

THE COMPARATIVE ANALYSIS OF SETTLEMENT SYSTEMS

Shuji Kobayakawa

FINANCIAL ECONOMICS

Centre for Economic Policy Research

Centre for Economic Policy Research
25–28 Old Burlington Street, London W1X 1LB

ISSN 0265-8003

THE COMPARATIVE ANALYSIS OF SETTLEMENT SYSTEMS

Shuji Kobayakawa

Discussion Paper No. 1667
July 1997

Centre for Economic Policy Research
25–28 Old Burlington Street
London W1X 1LB
Tel: (44 171) 878 2900
Fax: (44 171) 878 2999
Email: cepr@cepr.org

This Discussion Paper is issued under the auspices of the Centre's research programme in **Financial Economics**. Any opinions expressed here are those of the author and not those of the Centre for Economic Policy Research. Research disseminated by CEPR may include views on policy, but the Centre itself takes no institutional policy positions.

CEPR Discussion Paper No. 1667

July 1997

ABSTRACT

The Comparative Analysis of Settlement Systems*

After a rapid expansion of financial transactions, both the authorities and financial institutions became aware of the risk involved in interbank settlement systems. To cope with the risk the systems in most economies have been designed so that large-sized payments are settled in the real time gross settlement (RTGS) mode, whereas small ones are dealt with in the designated time net settlement system. Yet the introduction of the RTGS system imposes substantial costs of maintaining intraday liquidity on participants, which creates a fear among European as well as Japanese banking communities that they might lose competitiveness against their US counterparts whose main settlement system, Fedwire, does not require collateral. We first compare these different settlement systems from the participant's cost-benefit perspective and see if this concern always holds. Second, we develop the same framework to analyse the net settlement system, and then suggest if there is a rationale behind the coexistence of both the RTGS and net settlement systems.

JEL Classification: G29
Keywords: settlement risk, RTGS, intraday liquidity, risk management

Shuji Kobayakawa
Institute for Monetary and Economic Studies
The Bank of Japan
CPO Box 203
Tokyo 100–91
JAPAN
Tel: (81 3) 3279 1111 x.6544
Fax: (81 3) 3510 1265
Email: shuji@imes.boj.go.jp

*This paper is produced as part of CEPR's Network on the Japanese Economy, supported by a grant from the Daiwa Anglo-Japanese Foundation and the ESRC. Support from the Japan Foundation Endowment Committee for production costs is gratefully acknowledged. The author would like to thank

Alan Beggs, Richard Dale, Marco Rossi and Dirk Schoenmaker for helpful discussions and comments, and is also grateful to the participants of the CEPR/CEP/STICERD Workshop on the Japanese Economy, 13/14 December 1996. The views expressed here are the author's own and not necessarily those of the Bank of Japan. Any remaining deficiencies are the author's sole responsibility.

Submitted 4 April 1997

NON-TECHNICAL SUMMARY

This paper examines the structure of settlement systems. Most payments have long been dealt with in the designated time net settlement system, where there is a time lag between the payment order and the actual transfer of money. Traditionally, the receiver of the payment order has made funds available to its client before interbank settlement (i.e. actual receipt of money) takes place. Yet the receiver faces a credit risk if the sender of the order goes insolvent before it fulfills its obligation. The recent expansion of financial transactions forced the authorities, as well as financial institutions, to realise how much risk is involved in this scheme. One way to cope with this risk is to design the payment system in such a way that large-sized payments are dealt with in the real time gross settlement (RTGS) system, in which the payment order and the transfer of money are carried out simultaneously, while small payments are still dealt with in the net settlement, but incorporate some measures against risk rather than being left open-ended. This combination of two systems seems to have become a world standard; in fact, most OECD countries except Japan have already adopted this framework although structural details vary. Yet the RTGS system requires banks to hold more intraday liquidity in order to finalize each payment separately. It certainly imposes additional costs of liquidity maintenance on participants, which creates a consensus in the banking community that its introduction is costly, but inevitable in order to reduce the risk they are exposed to under the 'old-style' net settlement system, i.e. the system without risk measures.

This paper examines the cost concept in the settlement system more carefully and assesses if there is an economic rationale for the coexistence of the RTGS and net settlement systems from the participating bank's cost-benefit perspective. We define three cost components, namely the cost of settlement failure, the cost of settlement delay and the cost of liquidity maintenance. First, the payment is cancelled under the RTGS system, unless the bank either holds funds in the form of reserves or has access to an overdraft facility provided by the central bank. Similarly, its cancellation occurs in the net settlement system which adopts the credit limit. In other words, any payment leading to a net debit position greater than the limit will automatically be rejected. The cancellation is costly to the bank in the sense that it brings about the dissatisfaction of clients because the payment cannot be carried out as smoothly as under the 'old style' net settlement system. Second, the payment delay will be observed in the RTGS system, as the bank cannot undertake the payment unless it has enough reserves or an overdraft facility. The delay may

Giannini, Curzio and Carlo Monticelli. (1995) 'Which TARGET for Monetary Policy in Stage Three? Issues in the Shaping of the European Payment System.' Temi di discussione No. 257 Banca d'Italia.

Greenspan, Alan. (1996) 'Remarks on Evolving Payment System Issues.' *Journal of Money, Credit, and Banking* 28(4): 689-95.

Horii, Akinari and Bruce Summers. (1994) 'Large-Value Transfer Systems.' In Bruce Summers, ed., *The Payment System: Design, Management, and Supervision.* Washington D.C.: International Monetary Fund.

Prisman, Elizer, Myron Slovin and Marie Sushka. (1986) 'A General Model of the Banking Firm under Conditions of Monopoly, Uncertainty and Recourse.' *Journal of Monetary Economics* 17(2): 293-304.

Richards, Heidi. (1995) 'Daylight Overdraft Fees and the Federal Reserve's Payment System Risk Policy.' *Federal Reserve Bulletin*: 1065-77.

Rochet, Jean-Charles and Jean Tirole. (1995) 'Controlling Risk in Payment Systems.' mimeo.

Rossi, Marco. (1995) 'Pricing Intra-day Credit in Real Time Gross Settlement Systems.' Discussion Paper No. 211 LSE Financial Markets Group Discussion Paper Series.

Schoenmaker, Dirk. (1994) 'A Comparison of Net and Gross Settlement.' Special Paper No. 60 LSE Financial Markets Group Special Paper Series.

______ . (1995) 'A Comparison of Alternative Interbank Settlement Systems.' Discussion Paper No. 204 LSE Financial Markets Group Discussion Paper Series.

Summers, Bruce. (1994) *The Payment System: Design, Management, and Supervision.* Washington D.C.: International Monetary Fund.

VanHoose, David. (1991) 'Bank behaviour, interest rate determination, and monetary policy in a financial system with an intraday federal funds market.' *Journal of Banking and Finance* 15: 343-365.

require an additional overdraft, and if the participant fails to recover its overdraft position, the preposted collateral may be distrained. The delay is irrelevant in the net settlement system, as the transfer of money does not take place until some designated time. Finally, there is the cost of liquidity maintenance: i) the loss of revenue from interest-bearing assets (we assume the reserve does not bear interest); and ii) the opportunity cost of holding collateral. This is also applicable to the net settlement system that requires participants to post collateral in order to cope with the possible loss caused by a participant's insolvency.

With these concepts we first define the profit the bank retains in the settlement systems as the revenue obtained from interest-bearing assets in the portfolio minus the relevant cost specified above. Then the model incorporates two types of payments: one is known in advance while the other is not. The bank's strategy is to choose the settlement timing in the RTGS system and the level of credit limit in the net settlement system. We then derive the equilibrium and compare the level of profit between the two systems. We first compare the US and EU RTGS systems; the former charges overdraft fees while the latter requires collateral. Depending on the level of fees and the extent banks perceive collateral holding as costly, we claim that either system can provide higher profit to the bank. We then analyse the coexistence of the RTGS and net settlement systems, which is justified if the bank receives higher profit from the net settlement system when the size of payment is small, and the RTGS system provides higher profit when the payment is large. Because the profit in the net settlement system is less likely to be affected by payment size, as payments tend to be cancelled out by each other and only the net-out position is settled, we need to determine the conditions under which the profit in the RTGS system increases with payment size. We find that if the cost of customer dissatisfaction is relatively high while other costs are low, the coexistence is justified. This is because, as the size of the payment rises, the more the bank needs to post collateral in order to guarantee access to an overdraft, the higher the chances are that the bank cannot recover its overdraft position after having settled the large payment, but the lower the possibilities are that it has to cancel this payment. Thus, if the third cost saving factor outweighs the other two, the profit in the RTGS system increases with payment size. We then find this condition holds under reasonable values of costs. Our result seems to suggest that coexistence of the two systems is well explained from the participant's cost-benefit perspective, and it can be misleading to conclude that the RTGS system is always more costly than the other. Finally, the implication of the analysis is rediscussed. We are especially interested in how the settlement system in Japan, especially BOJ-NET, the system run by the Bank of Japan, should be restructured.

REFERENCE

Angelini, Paolo. (1994) 'About the Level of Daylight Credit, Speed of Settlement and Reserves in Electronic Payment Systems.' Temi di discussione No. 229 Banca d'Italia.

Bank of Japan. (1992) 'Functions of and Risks in Payment and Settlement Systems in Japan.' Special Paper No. 216.

Bank of Japan. (1995) *Wagakuni no Kin-yu Seido* (Financial Systems in Japan). Tokyo.

Benink, Harald. (1995) *Coping with Financial Fragility and Systemic Risk*. London: Kluwer Academic Publishers.

Berger, Allen, Diana Hancock and Jeffrey Marquardt. (1996) 'A Framework for Analyzing Efficiency, Risks, Costs, and Innovations in the Payments System.' *Journal of Money, Credit, and Banking* 28(4): 696-732.

Bhattacharya, Sudipto and Anjan Thakor. (1993) 'Contemporary Banking Theory.' *Journal of Financial Intermediation* 3: 2-50.

Borio, Claudio and Paul Van den Bergh. (1993) 'The Nature and Management of Payment System Risks: an International Perspective.' BIS Economic Papers No. 36.

Brealey, Richard and Stewart Myers. (1992) *Principles of Corporate Finance*. New York: McGraw-Hill.

Chakravorti, Sujit. (1995) 'Analysis of Systemic Risk in the Payments System.' mimeo. Barents Group, LLC of KPMG Peat Marwick, LLP.

Dewatripont, Mathias and Jean Tirole. (1994) *The Prudential Regulation of Banks*. Payot Lausanne: The MIT Press.

Folkerts-Landau, David, Peter Garber and Dirk Schoenmaker. (1996) 'The Reform of Wholesale Payment Systems and its Impact on Financial Markets.' Occasional Paper 51, Group of Thirty, Washington DC.

Fudenberg, Drew and Jean Tirole. (1993) *Game Theory*. Cambridge: The MIT Press.

(a) If $\Delta^{US}(0) > 0$, i.e.

$$s < \frac{\theta}{4L}(P + 2R_A),$$

then $\Delta^{US}(0)$ is always positive, implying that settling at period 1 is less costly than delaying. This will lead us to a unique equilibrium where two banks settle at period 1, because settling at period 2 is now strictly dominated by settling at period 1.

If $\Delta^{US}(0) < 0$, i.e.

$$s > \frac{\theta}{4L}(P + 2R_A),$$

then there are two possibilities. There exists a unique R_B^* which gives $\Delta^{US}(R_B^*) = 0$ ($R_B^* \in [0, P]$). This follows from $\Delta^{US}(0) < 0$, $\Delta^{US}(P) > 0$, and monotonicity of $\Delta^{US}(R_B)$ in $R_B \in [0, P]$.

(b-1) If $R_B \in [0, R_B^*]$, then $\Delta^{US}(R_B) < 0$, implying that the delay is less costly. This will derive two pure strategy equilibria where one bank settles at period 1 while the other settles later, and one equilibrium with mixed strategies.

(b-2) Finally if $R_B \in [R_B^*, P]$, then $\Delta^{US}(R_B) > 0$. This brings us back to the first case where there is a unique equilibrium in which two banks settle immediately. □

1. INTRODUCTION

Over the last decade, the volume of financial transactions in domestic as well as international markets has increased enormously, owing a great deal to the development of new financial commodities and the innovation in telecommunications. Borio and Van den Bergh (1993) report that the total value of fund transfers has become so large that in 1990 it took just 2 business days for Japan to turn over the equivalent of its annual GNP. Similarly it took 3 and 5.5 days respectively in case of the US and the UK. This tremendous expansion of financial transactions has raised doubt about the efficiency and safety of the settlement system which has long been regarded as an issue people in the back office should concern. The bankruptcy of Bankhaus Herstatt in 1974 is a classic example in which banks realised risks involved in the settlement lags in foreign exchange transactions. Nevertheless, the BCCI debacle in 1991 proved participants as well as the authorities in the system had yet to realise fully how much risk was immanent in the payment system. The banking community worldwide has only recently begun to adopt rigorous measures to ensure effective risk management.

There are essentially two major payment systems: one is the real time gross settlement (hereafter RTGS) system managed by the central bank, and the other is the designated time net settlement system run by a private institution. Currently, most countries seem to favour developing a framework where both systems can coexist. The US is leading this trend - having both Fedwire, the RTGS system run by the Federal Reserve, and CHIPS (Clearing House Interbank Payments System), the net settlement system governed by New York Clearing House. The EU member states currently plan to institute a system whereby both RTGS and net settlement systems are accommodated in each member state, and eventually all the RTGS systems will be linked as the TARGET system. Japan is being left behind. In its main payment system, BOJ-NET (Bank of Japan Financial Network System), while accommodating both the net and gross settlement modes, the bulk of payments goes through the net settlement mode, the restructuring of which has now become a matter of urgency. It has been discussed that BOJ-NET should relinquish its net settlement mode, and adopt the "genuine" RTGS system, but the question is, what type of the RTGS system fits BOJ-NET?

In order to find an appropriate structure for BOJ-NET, this paper first provides models which examine several features of two completely different RTGS systems, namely the EU-type and the US-type.[1] The US-type charges fees on the central bank's overdraft, but requires no collateral in principle, while the EU-type charges no fees, but full collateralisation is strictly required.[2] Folkerts-Landau, Garber and Schoenmaker (1996) claim one important issue arising from the evolution of two different RTGS systems is that the bulk of financial transactions will ultimately be carried out in US dollars, if, as they argue, the EU-type is more expensive than the US-type. This is based on the idea that the collateral holding is considered highly costly. Yet, if it were so costly, it is not clear why the European banking communities would have adopted this system, given the world of banking is under fierce competitive pressure. In fact, following the discussion of how CHAPS (Clearing House Automated Payment System) came to be transferred from the net settlement to the RTGS system with collateralised, interest-free overdrafts in April 1996, the clearing banks seemed to have been quite resistent to an introduction of fees, but were very understanding on the point that the Bank of England should not bear any credit risk; hence 100% collateralisation was inevitable. Indeed, Robin Leigh-Pemberton, the former Governor of the

[1] Rochet and Tirole (1995) compare RTGS systems which are categorised according to the availability of the central bank's intraday credit. Our interest of RTGS is, however, limited to the systems with the central bank credit, because the RTGS systems around the world seem to converge into one direction in which the central bank's intraday liquidity provision is already taken for granted.

[2] Obviously, there are many practical issues which will not be incorporated in the model. Nonetheless, our aim is not to draw a grand design of the future payment system in Japan, but rather to provide a starting point of discussion which treats and assesses various RTGS systems.

earlier when the counterpart delays the payment.

$\Delta^{US}(R_B)$ compares two factors. The first integral in the equation shows the total fee charged by the central bank for use of the overdraft at period 1, which can be avoided by delaying the settlement. The second integral represents the additional loss from the payment cancellation. By delaying the payment, B will raise the probability of A's payment cancellation, i.e. A is forced to cancel ϵ if $\epsilon \in [R_A, L]$, which has a counter-effect on B by incurring further loss from customer dissatisfaction.

Given the assumption that ϵ is uniformly distributed in the region of $-L$ and L,

$$\Delta^{US}(R_B) = sR_B + [\frac{\theta}{4L}(P^2 + 2PR_A) - sP].$$

We know $\Delta^{US}(R_B)$ reaches the minimum value at $R_B = 0$, and monotonically increasing for $R_B \in [0, P]$. Furthermore,

$$\Delta^{US}(0) = \frac{\theta}{4L}(P^2 + 2PR_A) - sP$$
$$\Delta^{US}(P) = \frac{\theta}{4L}(P^2 + 2PR_A) > 0.$$

$\Delta^{US}(P)$ is always positive, suggesting if $R_B = P$, Bank B will prefer to settle without delay. The logic here is quite straightforward - if the bank holds reserve equal to P, there is no need for overdrafting in order to settle P, whereas to delay will be costly because the bank has to incur the cost of customer dissatisfaction. At this stage, $\Delta^{US}(0)$ could be either positive or negative; let us look at each possibility in turn.

$$\Delta^{EU}(R_A) = \int_{-L}^{-(R_B+\frac{P}{1+\rho})} \theta\epsilon\, f(\epsilon)d\epsilon - \int_{-L}^{-R_B} \theta\epsilon\, f(\epsilon)d\epsilon$$
$$= -\int_{-(R_B+\frac{P}{1+\rho})}^{-R_B} \theta\epsilon\, f(\epsilon)d\epsilon > 0$$

$\Delta^{EU}(R_A)$ is always positive, meaning that Bank A prefers to settle at period 1 to period 2. This again applies to Bank B by symmetry when A settles at period 2. Thus both banks' strategy to settle at period 2 is strictly dominated, and the only equilibrium is to settle at period 1. □

Proof of Proposition 2:

The proof follows the same procedure as in Proposition 1. First compare Bank B's profits when A settles at period 1, and second, follow the same for Bank A's profits when B settles at period 2. We define B's difference in the profit functions when A settles at period 1 by $\Delta^{US}(R_B)$ and A's when B settles at period 2 by $\Delta^{US}(R_A)$ in this model.

$$\Delta^{US}(R_B) = -\int_{-L}^{L} s(P-R_B)\, f(\epsilon)d\epsilon - \int_{R_A+\frac{P}{1+\rho}}^{L} \theta\epsilon\, f(\epsilon)d\epsilon + \int_{R_A}^{L} \theta\epsilon\, f(\epsilon)d\epsilon$$
$$= -\int_{-L}^{L} s(P-R_B) f(\epsilon)d\epsilon + \int_{R_A}^{R_A+\frac{P}{1+\rho}} \theta\epsilon\, f(\epsilon)d\epsilon$$

$$\Delta^{US}(R_A) = \int_{-L}^{-(R_B+\frac{P}{1+\rho})} \theta\epsilon\, f(\epsilon)d\epsilon - \int_{-L}^{-R_B} \theta\epsilon f(\epsilon)d\epsilon$$
$$= -\int_{-(R_B+\frac{P}{1+\rho})}^{-R_B} \theta\epsilon\, f(\epsilon)d\epsilon > 0$$

$\Delta^{US}(R_A)$ suggests that it is less costly for A to settle at period 1 when B settles later. This means that if every payment is delayed till the last moment, the probability of the counterpart rejecting the payment will be increased, which eventually turns out to be more costly for them as well due to the additional cost of customer dissatisfaction, so there is an incentive for the bank to settle

Bank of England, had announced as early as October 1992 that "the approach ... is for the central bank to provide settlement banks ... with an intraday credit facility, which does not incur interest charges but which is covered by the banks providing high quality assets as security to remove the risk to the central bank." This rightly suggests that there had already been a consensus about the future framework of RTGS, between the Bank of England and the clearing banks, more than 3 years earlier than the actual introduction of the system. It may furthermore suggest that they are primarily concerned with the fees, but not with the cost of collateral holding.

We analyse how two RTGS systems affect the profit of each participant who is due to settle payments each day. We then compare the profit obtained in each system, and assess which system is preferred by the participants. Secondly, we look into a different issue arising from the structure of the payment system: we try to find if there is any economic rationale behind the coexistence of the RTGS and the net settlement systems in the single economy. From our point of view, it is justified if the banks receive higher profits by settling larger payments in the RTGS system, while they prefer the net settlement system when the payment size is small. Our conclusion is that the two systems can coexist under certain conditions.

The paper is organised as follows. In Section 2, we assess the RTGS systems against two criteria: first, intraday liquidity provision by the central bank and second, risk management. In Section 3, we study the precise definition of costs, based on the recent contribution by Schoenmaker (1995). We specify three major cost components: the cost of settlement failure, the cost of settlement delay, and the cost of liquidity maintenance. The model building starts from Section 4, which deals with the EU-type RTGS system. We assume that the bank can choose whether or not to delay the payment (the timing of settlement). To the extent that the banks are guaranteed access to the central bank's overdraft facility without any charge, we find the equilibrium is to settle without delay. Section 5 looks into the US-type RTGS system. This time, the central bank's overdraft is charged fees, and there exists an incentive to free-ride on the incoming payments, thus being able to avoid overdrafting. Therefore, delaying the settlement is also possible as an equilibrium, depending on the relative effect of two variables: the fee charged by the central bank and the amount of reserve the banks hold initially. Section 6 deals with the comparative statics between these two types of the RTGS systems. Depending on (1) the level of fee set in the US-type, and (2) the extent banks perceive collateral holding as costly, we claim that the EU-type can provide higher profit to the bank. We then move on to the analysis of the net settlement system, in which the bank's strategy is to set the intraday credit it is willing to accept from others. Comparative study between the net settlement system and the RTGS system is carried out in Section 8. We conclude that the coexistence of two systems are justified if (1) the cost of customer dissatisfaction due to payment cancellation is comparatively large, and (2) the banks hold relatively little reserve. Finally, we discuss the implications of this analysis for reform of BOJ-NET.

2. POINT OF DISTINCTION IN REAL TIME GROSS SETTLEMENT SYSTEMS

In order to make the comparative analyses in later sections easier, let us focus on the following two factors, from which we can categorise the RTGS systems into several groups.

- Is the central bank's credit facility available?
- If so, how is its exposure to credit risk controlled?

2.1 Liquidity Provision

One way of distinguishing the settlement systems is to see how participants will fund the liquidity that is necessary to carry out the settlement obligation. The first and most common

The current trend in restructuring BOJ-NET favours the EU-type RTGS system, where daylight overdrafts are fully secured by collateral. It seems to be more consistent with Article 22 of the Bank of Japan Law that emphasises the soundness of the Bank's assets. However, our analysis shows that in certain conditions there can be advantages to the US-type system: if the overdraft fees are set appropriately, then participating banks can make more profit from this system. Ultimately, before deciding between the US-type and the EU-type RTGS systems, further empirical research on the actual costs of each system is required. This analysis provides a model for assessing such evidence.

APPENDIX

Proof of Proposition 1:

We first compare Bank B's profits. Let $\Delta^{EU}(R_B)$ be $\pi_B^{11} - \pi_B^{21}$. It represents Bank B's difference in profit if Bank A settles at period 1, i.e. (B's profit by settling at period 1) - (B's profit by settling at period 2).

$$\Delta^{EU}(R_B) = -\int_{R_A+\frac{p}{1+\rho}}^{L} \theta\epsilon f(\epsilon)d\epsilon + \int_{R_A}^{L} \theta\epsilon f(\epsilon)d\epsilon$$
$$= \int_{R_A}^{R_A+\frac{p}{1+\rho}} \theta\epsilon f(\epsilon)d\epsilon > 0$$

The equation is always positive. Hence we know Bank B's profit by settling at period 1 is always greater than settling at period 2, when A chooses to settle at period 1. By symmetry, this result applies to Bank A when B settles at period 1.

Next we compare Bank A's profit if Bank B settles at period 2. Let $\Delta^{EU}(R_A)$ be $\pi_A^{12} - \pi_A^{22}$, Bank A's difference in profit, i.e. (A's profit by settling at period 1) - (A's profit by settling at period 2).

The comparative analysis further identified conditions on which each equilibrium yields higher profit to the banks. Unless there is a delay in the US-type, we concluded that, depending on the level of fees and the cost of collateral holding, the EU-type could turn out to be more profitable than the US-type and vice versa. The balance of these two cost factors determines whether the bank would prefer to be a participant of the EU or US-type RTGS system. One cannot determine whether the present fee is higher than the cut-off level specified in the model, without conducting extensive analysis of actual data, which is beyond the scope of the current analysis. However, our model seems to suggest that further examination would be necessary before one can determine whether the bulk of financial transactions would in fact be redenominated into US dollars, as Folkerts-Landau *et al.* (1996) maintain.

Secondly, we modeled the net settlement system by taking the credit limit as a choice variable. There we analysed extensively what happens if banks set the limit below their overnight cash positions. We then look at whether there is an economic rationale for the coexistence of the net and gross settlement systems in a single economy. The profit in the NET system is not a function of payment size, as they cancel each other out; nonetheless, the profit in the RTGS system is affected by various factors. As the size of payments increases, both types of the RTGS systems benefit from the decline in the possibility of incurring the cost of customer dissatisfaction, but they are negatively affected by the rise of the cost of overnight credit, and the rise in fees in the US-type or the rise in the cost of collateral posting in the EU-type. The effect on profit in the RTGS systems will be determined by the relative magnitude of these factors. We then identified certain conditions, under which the coexistence of the RTGS and NET systems is justified, specifically that: (1) the cost of customer dissatisfaction is relatively large, and (2) the bank holds little reserve.

method is <u>reserve holding</u>. It is usually regarded as costly in a sense that it incurs opportunity costs of foregoing the revenue that could have been earned otherwise, had it been invested in the interest bearing assets. On the other hand, the reserve is an essential source of "free money," especially when the central bank charges fees for overdraft facility provided. The second source of liquidity is the <u>central bank credit facility</u>. It can take the form of the central bank providing an overdraft facility or offering a repurchase agreement. Overdraft facilities can either be fully collateralised or can be provided without collateral, in which case the central bank bears a credit risk. In the following analysis, we treat the full collateralised overdraft and the repurchase agreement as the same system, in that in both cases the central bank is free from credit risk. The third source is the <u>interbank money market</u> where sharing of spare liquidity among participating banks takes place. This source is arguably limited because it is extremely difficult to meet the vast amount of demand and supply which arise in the RTGS system but differ in time and condition. Although intraday liquidity provision through the interbank market is practised in Japan, it does not seem correct to argue that the market plays a major role of liquidity provision in its main settlement system, BOJ-NET, for the following reason: BOJ-NET is very peculiar - it possesses both gross and net settlement modes. The real time gross settlement mode in BOJ-NET handles less than 2% of its designated time net settlement mode in terms of transfer volumes (Horii and Summers (1994)). Thus the market transaction is derived by the demand to settle net balances between the designated times, and has little to do with the RTGS services. Let us now categorise the RTGS systems, according to which facility is used as the main source of liquidity:

Figure 2-1 Several Models of RTGS Systems

		Model 1	Model 2	Model 3
Main Source of Liquidity	Reserve	O	O	O
	C/B credit	O	-	-
	Interbank	-	(O)	-
Example		CHAPS Fedwire	(BOJ-NET)	SIC

As described, BOJ-NET does not correspond exactly to the RTGS model. SIC (Switzerland) is another unique system as neither the central bank credit nor the interbank market is available. Instead, SIC has developed a queuing system in which all payment orders are processed on the FIFO (first-in-first-out) base, once the funds become available.[3] All the other systems in Europe, including CHAPS (UK), EIL-ZV (Germany), RIX (Sweden) and TBF (France), have been transferred to the RTGS system with a central bank credit facility. We denote these and Fedwire of the US as examples of Model 1.

2.2 Credit Risk Management

Next point of distinction is to see how the central bank's credit exposure to the participants is managed. Here we divide Model 1 into two categories:

Figure 2-2 Risk Management

	Model 1-A (EU-type)	Model 1-B (US-type)
Collateral	O	-
Fee	-	O

[3] We will not deal with SIC type model as it is not a main stream of the world RTGS systems.

US-type RTGS system than the EU-type. This choice will depend on what level of overdraft fees is set by the central bank. What this result seems to suggest is that the possibility of adopting the US-type RTGS system should not be ruled out, just because it suits the central bank better.

It looks as though we need to give further consideration before deciding which type of the RTGS system is more appropriate. In practice, the Bank of Japan is capable of extending its loans without collateral, according to Article 25, which allows the Bank, with the Finance Minister's authority, to act beyond its ordinary scope of operation. The Federal Reserve is also acting in a flexible way: although its discount window credit is obliged to be backed by collateral according to Federal Reserve Act, the daylight overdraft is interpreted as "something else incident to payment system services," which enables the Federal Reserve to adopt the present Fedwire system.

10. CONCLUSION

We have studied how different settlement systems affect the bank's profit. In the models of gross settlement systems, we assumed that the bank can choose whether or not to delay the settlement, and found settling without delay was an equilibrium. However, in the US-type system, it was also an equilibrium to delay the settlement, depending on the relative effect of two variables: the fee charged by the central bank and the amount of reserve the banks hold initially. The higher the charged fees are, the more costly for the banks, therefore, the higher the incentive for them to avoid using overdrafts by waiting and free-riding on the incoming payments from the counterpart. Similarly, the less reserve the banks hold, the more likely it is that they will be forced to rely on overdrafts; thus, the higher the incentive is for them to delay the payments. These equilibria seem quite intuitive.

It would be incorrect to assume that each city bank needs to raise 10 times as much liquidity as in the above figure (2,815.2 billion Yen), because the call money is not solely used for payment purposes. Nevertheless, a huge demand for call money by city banks, the biggest borrowing group, could hardly be met by the call loans. Even the biggest lender, trust bank, is not capable of accommodating all the extra liquidity needs which would result from the transition to the RTGS system. Therefore it is right to suggest that this liquidity has to be supplied by the central bank, and thus the RTGS system has to be provided by the public sector. The first option seems inappropriate. However, there is little to choose between the second and third options. One difference is as follows:

- **Integration of the whole payment system** - the central bank provides both settlement modes but with various measures and the intraday credit facility. This means the private systems can be abolished.
- **Coexistence of public and private systems** - the central bank provides the RTGS system which is supported by the intraday credit and the private sector supplies the net settlement system.

Whether the net settlement mode should be provided by the public or private is beyond the scope of our analysis, but it may be an area of further research.

9.4 How to Provide Intraday Liquidity

Finally we briefly comment on alternative methods of intraday liquidity provision by the central bank. At the moment, BOJ-NET seems to be heading for the EU-type RTGS system in which the overdrafts are provided with no fees but with collateral. This view is based on Article 22 of the Bank of Japan Law which emphasises the soundness of assets. Our analysis, nevertheless, has shown that in some cases the participants in the system prefer to settle their payments in the

In Model 1-A, the central bank credit is fully covered with collateral. Some systems in Europe including EIL-ZV (Germany) and RIX (Sweden) provide overdraft facilities with full collateralisation, but others, such as CHAPS (UK) and TBF (France), adopt the repurchase agreement. The latter systems can nevertheless be thought as a variant of fully collateralised overdraft, in the sense that the central bank does not face any credit risk - the credit is always 100% covered with a pool of collateral. Model 1-B, the US-type, is quite different. In Fedwire (US), the central bank's credit is not fully collateralised, but fees are charged instead.[4] Having categorised the RTGS systems, our next task is to set up a framework which incorporates these contrasting features of the two dominant RTGS systems. In order to do so, it is necessary to (1) identify what costs there are in each system, and (2) see how the bank's profit is affected under different systems.

3. Costs in Settlement Systems

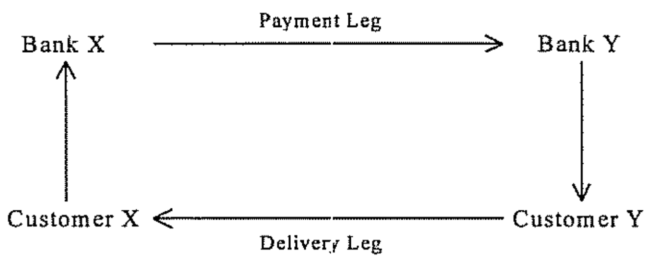
Schoenmaker (1995) has introduced three different concepts of cost in the settlement system: the cost of settlement failure, the cost of settlement delay, and the cost of collateral holding. We shall add different interpretation of these elements in turn.

While Schoenmaker's cost associated with the failure of settlement is negligible in the RTGS system, because the system does not adopt the "survivors pay" rule in which non-defaulting participants jointly share the defaulter's loss as an additional settlement obligation, it is hard to argue that the insolvency of the transacting partner in the system does not affect non-insolvent participants at all. We believe it should provoke certain effects on others even in the RTGS system. Our interpretation of settlement failure is as follows. There are two main causes of

[4] Charging fees started in April 1994. Before then, the overdraft was completely free. See later section about the fee policy in Fedwire.

settlement failure. First, the rejection of a payment order, and second, the failure of an actual transfer of good funds after the payment order has been made. Consider the following situation where there are two banks and two customers:

Figure 3-1 Circle of Transactions



Customer Y has sold goods to Customer X (delivery leg in the figure), and in return Customer X is due to pay the price for this good through his account held at Bank X. Let us start from the "old-style" situation where payments are undertaken in the net settlement system wherein participants are not concerned about the credit risk in the following two senses: (1) The sender (Bank X) provides a generous credit line to the customer (Customer X), so that Bank X is generally ready to transmit the payment order even if Customer X does not have good funds in his account. This has been customary in order to satisfy the customers' needs; (2) The receiver (Bank Y), upon obtaining the payment order (but not the actual funds) from the sender, advises the customer (Customer Y) right away that the funds have been received in his favour. This is again a long-standing custom in the business of banking. Obviously, these two customs could create a serious sequence of credit risk. For example, if Customer X goes bankrupt, Bank X will face the risk of not being able to collect the funds by a designated time for settlement with Bank

- A system with the net settlement mode, but with extensive measures to guarantee "finality"
- A system with both the net settlement mode with the finality rule, and the gross settlement mode with the central bank intraday credit
- A system with the gross settlement mode with the central bank intraday credit

The first option is excluded, because it will end up with all the systems adopting the net settlement mode. From our comparative analysis, we know that the banks can receive higher profits by settling large payments in the RTGS system under certain conditions, so we need to provide the gross settlement mode in at least one of the systems. Can it be provided by the private sector? This seems very unlikely. The enormous amount of intraday liquidity can hardly be financed either by the reserve holdings or in the interbank markets. For example, it is estimated that a city bank may need almost 10 times as much liquidity in the RTGS mode as in the present net settlement mode of BOJ-NET. At the moment, city banks raise most of the necessary liquidity in the call money market.

Figure 9-3 Average Outstanding Amounts by Borrowers in the Call Money Market in 1995

(100 million Yen)

	Total Amount	Average Amount
Total	421,289	-
City Banks	309,670	28,152
Regional Banks	30,311	474
Trust Banks	15,775	563
Long-term Credit Banks	20,981	6,994
Foreign Banks	25,560	272

Source: Bank of Japan (1996)

As the Federal Reserve increased the fees, they constantly had to take account of the possibility that substantial part of payments could be shifted to other settlement systems, if the participants perceived the fees to be too high.[34] In fact, this concern was one of the major factors that forced the Federal Reserve to scrap the plan to increase the fee to 0.48% in April 1995 and to 0.6% in April 1996. Instead it decided to set the fees at 0.36%, effective from April 1995. This evidence suggests that the Federal Reserve has always paid careful attention to substitutionary roles of Fedwire and CHIPS. We can see this more clearly by following how the Federal Reserve implemented a "twin-track risk reduction programme."

Figure 9-2 Risk Reduction Programme in US

	Fedwire	CHIPS
1984		bilateral credit limit
85	net debit cap	
86		net debit cap
90		loss sharing rule
94	fee charging (0.24%)	
95	fee charging (0.36%)	
97		loss sharing rule [34]
		net debit cap [35]

9.3 What Options Are There for BOJ-NET?

Given the issues denoted in Section 9.1, what options are left in terms of restructuring BOJ-NET? There seem to be three main choices:

[34] See Richards (1995) for detailed background of how the Federal Reserve faced the dilemma, namely to reduce the credit risk while avoiding the shifts of payments to private networks.

[35] These programmes are based on the announcement made by the New York Clearing House in July 1995. Loss sharing rule incorporates measures against multibank default case, and net debit cap is reduced further 20%.

Y. This may have a knock-on effect for Bank Y, namely that if Bank X fails to carry out the payment. Bank Y will automatically make a loss as long as the equivalent amount of funds has already been credited on its customer's account. Bearing these types of risk in mind, we now consider the following two "risk-conscious" situations.

Suppose, first, the payments are now carried out in the RTGS system, instead of the net settlement system. The virtue of the RTGS system is that the payment order and the actual transfer of funds take place at the same time, while in the net settlement system, there is always a time lag between these two. Bank X can still provide the credit line to Customer X, but this time the payment order to Bank Y cannot be made unless Bank X has the actual funds in its account and/or the central bank's overdraft facility is sufficient to cover the payment.[5] Bank Y, on the other hand, will not credit Customer Y's account, unless it receives the payment order from Bank X. Comparing this situation with the old-style one, we see that payments can be delayed or cancelled, which will certainly lead to inconvenience of the customers. Thus, adopting the RTGS mode is one way of dealing with the risk that banks face with respect to their customers, and we regard it is one significant source of the cost in the sense that the customer's convenience is sacrificed. In the following analysis, when a customer-related payment is delayed/cancelled, both sending and receiving banks will incur the cost due to customer dissatisfaction. We assume it is proportional to the size of payment by the factor θ ($0 < \theta < 1$). The customers are dissatisfied because the payment can no longer be carried out as smoothly as it used to be.

[5] In an extreme case, Bank X could reduce or abolish the credit line, if liquidity is tight, in which case Bank X would not accept the payment order of Customer X unless Customer X came up with the actual funds. The order would be delayed until Bank X could raise enough liquidity, or it could be cancelled if, for example, Customer X's order came at the very end of the day and Bank X failed to carry out this unexpected order.

Secondly, we consider the situation where the payments still take place in the net settlement system, but this time the system incorporates several risk measures. Along with the survivors pay rule, we introduce a bilateral credit limit which is the maximum amount of the net debit position the receiver is willing to accept vis-a-vis the sender. As described before, in the net settlement system, the receiver bears the risk against the sender until good funds are credited on its own account, since it has been customary for the receiver to make funds available to the customer before the interbank settlement takes place. In order to limit its exposure to this risk, the receiver sets the maximum limit bilaterally, beyond which the payment order is cancelled automatically. The sender cannot transfer the payment order unless its net-out bilateral debit position is within the region set by the limit. Again, payment cancellation can occur and in our interpretation both banks incur the cost of the customer dissatisfaction. Unlike the RTGS system, nonetheless, this bilateral limit cannot fully eliminate the credit risk. As long as the debit position is within the limit, it is possible for the payment order to be made, without the sender covering its net debit position at the end of the day. This could impose another cost on the banks. We assume if the sender of the payment order fails to raise enough funds by a designated time, it is regarded as insolvent and loses its assets. At the same time, the receiver of the order incurs the loss whose value is equivalent to the payment order, as it has already credited in favour of its customer; nevertheless, some of the loss can be recovered by the sender's asset. To sum up, settlement failure in the RTGS system results in banks incurring costs due to customer dissatisfaction. In the net settlement system, failure can also be caused by the introduction of a bilateral limit, in which case the bank also incurs the cost of customer dissatisfaction; otherwise, failure is due to lack of sufficient liquidity, in which case there is a loss of asset for the sender and the additional cost for the receiver.

Secondly, we adopt a different interpretation of the cost of delay. In the net settlement system, it takes the form of sending the payment order later than anticipated, which accompanies the

as privately-owned systems do. Nevertheless, this guarantee by the central bank creates moral hazard among the participants; therefore, if the net settlement mode continues to be accommodated in BOJ-NET, it is important that there is risk-sharing with the private sector. The participants in BOJ-NET must commit to a certain degree of risk, otherwise the Bank of Japan has to manage even the smallest risk.

To sum up, two issues seem to be currently under consideration: (1) Should BOJ-NET be restructured so that the Bank of Japan retains ownership only of the gross settlement mode, as is rapidly becoming the norm across the world? (2) A privately owned system (esp. FEYCS) does not offer protection against risk. What further steps should be taken at the same time as BOJ-NET is restructured?

9.2 Why Should the Private System Be Endurable Against the Risks?

- Development of Risk Management and Fee Policy in Fedwire (US)

Perhaps, it is helpful at this point to know how the systems in the US have been constructed. In Fedwire, nearly unlimited overdraft was available to the participants at no cost until the beginning of the 1980s. Then the rapid growth of intraday credit raised concerns about the credit risk the Federal Reserve had faced. The fee policy was finally implemented on 14 April 1994, but the Federal Reserve had taken several steps to reduce the amount of overdrafts. First of all, it introduced the net debit cap in 1985. Originally, the cap was set at three times as much as each participant's capital, but it was changed several times subsequently. By the late 1980s, the Federal Reserve concluded that further measures had to be taken to reduce its exposure to credit risk. A fee of 0.24% p.a. was introduced. The introduction of the fee had an immense impact on the use of overdrafts. According to Folkerts-Landau *et al.* (1996), peak overdrafts immediately fell on average by 40% (from nearly \$125 billion to \$70 billion after the introduction of fees in 1994), but they only underwent a marginal impact in 1995.

Of these systems, the Bill and Check Clearing System is different from the others, because of its nature as a regional clearing system rather than a nationwide system.[30] Among other three systems, the Zengin System, which is run by Domestic Funds Transfer Management Organisation, and deals with most of the interbank transactions, incorporates the most effective measures to ensure proper risk management. The amount of collateral reaches almost 50% of the total net debit cap of all the participants. This means, roughly speaking, that the system will survive even if half the participants default; this is a much higher standard than the one required by the Lamfalussy Report (November 1990). Moreover, there is an unwritten agreement that the Bank of Japan will always support the system, if necessary.[31] On the other hand, FEYCS (Foreign Exchange Yen Clearing System) run by the Tokyo Bankers Association, is left open-ended. Although the system supports the setting of bilateral net credit limit, this facility is hardly used among participants in practice. FEYCS has not introduced the net debit cap, so the system as a whole is still subject to unlimited risk. The loss-sharing rule just specifies that the participants with net credit position against the defaulter will share the loss; however, the participants do not offer any collateral in support of the scheme.[32] FEYCS can thus be exposed to a severe risk. Finally, the net settlement mode in BOJ-NET is also open-ended.[33] Of course, the system is guaranteed by the Bank of Japan, so that it may not require as exacting standards

[30] As part of risk management, if a bill or check for collection is dishonoured, there is a severe penalty: an issuer's name and address are publicly announced by the clearing house. Were this to happen twice within 6 months, the issuer is denied any transactions with the member banks of the house for two years.

[31] This is an arguable point as it may encourage moral hazard among participants.

[32] The Bankers Association recently (December 1996) announced that it will require the collateral to be preposted from the first half of 1998.

[33] As we discussed in the beginning, the gross settlement mode in BOJ-NET is virtually unused, mainly due to a lack of proper intraday liquidity provision. Horii and Summers (1994) report that the designated time net settlement mode is estimated to handle 50 times the transfer volume than the gross settlement mode.

customer dissatisfaction. In the RTGS system, it can take two forms. Consider the situation where two banks are expected to make certain payments simultaneously. If one of them fails to carry out its obligation on time, the counterpart will face a tighter liquidity position, which can lead to the following problems. Until the counterpart receives the payment, all the payments in subsequent periods may face higher chance of being cancelled. Should any payment be cancelled, the banks will again incur the cost of customer dissatisfaction. Even though cancellation of the payment can be prevented, the counterpart may have to rely on an additional overdraft to carry out the payment. In the US-type RTGS system, the overdraft is explicitly charged fees. In the EU-type, if it fails to overcome its overdraft position, the collateral can be distrained. In this case, the bank may have to incur an additional cost solely as a result of the settlement delay.

Thirdly, we look into the cost of liquidity maintenance. The reserve bears an opportunity cost as it does not earn any interest at all. The more reserve the bank holds, the higher the opportunity cost is, so there is always an incentive to hold as little reserve as possible. Nevertheless, the reserve plays the most important role along with the allowed overdraft facility, when it comes to settling the payments. The bank may be forced to reject more payments if it does not hold sufficient reserve. In our model, this cost of liquidity maintenance is reflected on the revenue side in the bank's profit function. The reserve is regarded as a non-interest bearing asset, and keeping too much reserve will reduce the amount of interest bearing assets in the bank's portfolio, and hence reduce potential earnings. In addition, we assume there is an opportunity cost of holding collateral, which is proportional to the amount of collateral by the factor c ($0 < c < 1$). This cost is borne due to the fact that the collateral cannot be traded freely in the market.[6]

[6] Some argue, however, the foregone interest by holding liquid assets should also be added as a justification of the opportunity cost, but we think it secondary. As in the case of the UK, the clearing banks hold quite substantial amount of liquid assets to meet a liquid asset ratio for prudential reasons (see Folkers-Landau *et al.* (1996) Footnote 78), there is little cost element which is solely incremented by the

The following model incorporates these concepts in the participants' profit functions. Each bank possesses certain amount of liquidity which can be invested either in reserves or short-term government bond (TB). Then banks are given the choice of settlement timings, i.e. whether or not to delay the payment, and through processing the payments, their portfolios change. At the end of the day, they will receive the interest revenue from the remaining TB, and the profit is derived by subtracting all the relevant costs. Our model identifies the equilibrium which turns out to be that the banks settle their payments without delay.

4. THE MODEL OF GROSS SETTLEMENT SYSTEM: CASE OF MODEL 1-A[7]

The analysis is based on the demand for reserves under uncertainty. This framework is first developed by Prisman, Slovin and Sushka (1986) and VanHoose (1991). Angelini (1994) develops VanHoose's model further by choosing the bank's borrowing and lending in the interbank market as a control variable. In this model, however, we would like to focus on the bank's strategy to coordinate the amount of reserve holdings and the use of the overdraft, which has not been rigorously analysed so far.

4.1 Notation and Timeline

There are three periods in a day. Because we will deal with an overnight credit, our model needs four periods - the last period represents the beginning of the following day.

intraday settlement needs. In other words, banks have to hold liquid assets anyway, no matter whether they settle payments every day.

[7] In this section, we regard the system which provides overdraft with full preposted collateralisation (eg. EIL-ZV in Germany) as a prototype of the EU-type RTGS system.

system. These results imply that the authorities, although they still need to monitor the level of the cost of customer dissatisfaction and the level of the fees charged on the overdraft in the case of the US-type, they could introduce both systems and allow participating banks to adopt their own strategies with respect to using the two systems.

9. SOME REMARKS ON THE DESIGN OF SETTLEMENT SYSTEMS IN JAPAN

We shall now examine what policy implications this analysis may have on the current restructuring of the settlement systems in Japan.

9.1 Issues of Concern

Broadly speaking, there are four systems currently under operation: three of them are private, and one is run by the Bank of Japan (see Figure 9-1).

Figure 9-1 Average Daily Volume in 1993

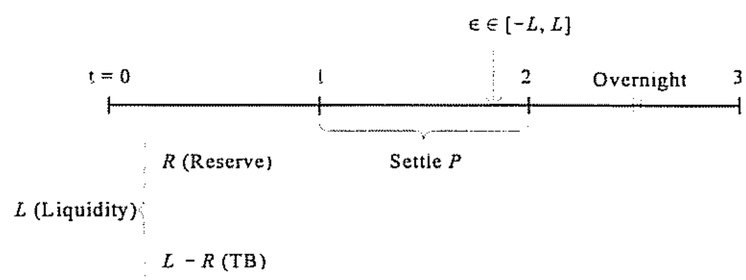
	Number of Transactions	Value (trillion Yen)	Value per Transaction (million Yen)
Bill & Cheque	1,332,000	13.2	10.0
Zengin	3,359,000	7.3	2.2
FEYCS	27,000	24.2	900.0
BOJ-NET	23,000	283.6	12,200.0

Source: Bank of Japan (1995)

benefit in the NET system. Taking account of the fact that there is an incentive for the bank to hold less reserve, Case 3 represents best what the bank faces, but the question is, which region is more likely, Region I or II? Let us first look at the numerical example to compare the area of two regions for feasible values of ρ and θ. Suppose $r = 0.04$. In the NET vs EU-type, the ratio of Region I and II reaches 1 : 99. In the NET vs US-type, it is almost 1 : 73, due to the fact that the slope constraint is shifted upwards by the overdraft fee equivalent to 0.36%. This example suggests that Region I is undoubtedly smaller than Region II. Nevertheless, the likelihood depends on the value of the cost of customer dissatisfaction. We know if this cost is larger than ρ/2 in the NET vs EU-type, Region II always prevails, no matter what value ρ takes. In the NET vs US-type, if θ is larger than $(s+\rho)/2$, the same result, i.e. Region II always prevails, applies, but if θ is smaller than $s/2$, the converse is true, i.e. Region I always comes into existence. It should be admitted that the cost of customer dissatisfaction is difficult to quantify. As we described in Section 3, however, the bank-customer relationship has been build on the long-established practice that the bank provides a credit line to the customer, and makes funds available as soon as the payment order is received. A breakdown in this established practice would damage the relationship severely. This may suggest that the cost of customer dissatisfaction is quite high.

To sum up, it seems that the coexistence of two systems can be understood clearly from the cost-benefit perspective of the participating banks. When the amount of reserve and P are both small, the bank has to cancel the stochastic payment in a wider range in the RTGS system than the NET system, which turns out to be more costly for the RTGS system, given the unit cost of customer dissatisfaction is high. In this case, the NET system is preferred. However, as P increases, the bank in the RTGS system builds up the collateral and as it starts undertaking the payment, the cost of customer dissatisfaction declines much faster in the RTGS system, which is replaced by the cost of overnight credit from the central bank. Because the central bank credit is cheaper than raising funds in the market, the bank will eventually reach the point where it prefers the RTGS

Figure 4-1 Timeline of the Model



There exist two banks, namely Bank A and B, whose total liquidity positions are L. L can be thought as a sum of reserves and TB (Treasury Bill), some of which is due to be preposted as collateral. At period 0, they decide an amount of reserve R_i ($i = A, B$) and the rest of the liquidity position is invested in TB which yields a return equivalent with r.[8] The banks' initial balance sheets are thus as follows:

Bank A		Bank B	
(Reserve)	(Liquidity)	(Reserve)	(Liquidity)
R_A	L	R_B	L
(TB)		(TB)	
$L - R_A$		$L - R_B$	

We assume each bank owes a payment P to the other. P represents a bank's normal transaction which occurs regularly as part of their daily interbank operation. Both of them initially expect P to be settled at period 1, but it is not time critical - as long as they settle P on the designated

[8] In the net settlement system, we assume banks can raise overnight funds in the market at the rate of r.

day, the payments are regarded as being finalised. On our timeline, we assume they are free to choose either period 1 or 2 to settle this transaction.

Banks have an access to the central bank's daylight overdraft at these two periods by posting collateral. In the RTGS system, the amount of collateral to be posted is usually determined by the peak daylight overdraft. We assume that they post collateral equivalent to P. Posting P would enable them to overdraft and settle their part of payment P without having any reserves or waiting for the counterpart to transfer P in advance of its own payment. So, the lower boundary for $L - R_i$ is $\min(L - R_i) = P$ $(i = A, B)$; hence, $R_i \leq L - P$. Their intraday cash position, defined by the sum of the reserve and the available overdraft, is thus $R_i + P$. Any payment greater than the cash position cannot be processed. If this happens at the end of the day, i.e. period 2 in our model, the overdraft can be converted into an overnight credit, where the collateral is required and the official discount rate ρ, which is less than the market rate r, is charged. So, with TB amounting to P, the bank receives the overnight credit equivalent to $P/(1+\rho)$. We thus define the bank's overnight cash position as the sum of the reserve and the available overnight credit, i.e. $R_i + P/(1+\rho)$. This overnight credit has to be returned to the central bank at period 3.

At the beginning of period 2, a stochastic payment obligation arises, which is due to be settled at period 2. This payment, denoted by ϵ, is best represented as customer-related, which can happen any time of the day when the customer walks in the bank and asks for the transfer of money to his client's account in another bank. For simplicity, the expected value of ϵ is zero ($E(\epsilon) = 0$), and ϵ is uniformly distributed in $[-L, L]$, where positive ϵ is interpreted as the payment from Bank A to B, and negative ϵ as the payment from Bank B to A. For convenience in the later analysis, we further assume that P has a priority over ϵ if there is a situation in which two payments have to be settled at the same time. So it could happen that the bank does not have

8.3 Discussion

We first summarise the result obtained in the previous sections. We only show the cases where cL/R and sL/R are greater than 1, as they seem to represent the most likely outcomes. We also omit the non-generic cases ($R = R^*$ <Section 8.1.2> and $R = R^{**}$) from the summary.

Figure 8-5 Summary of Results for NET vs EU-type RTGS Systems

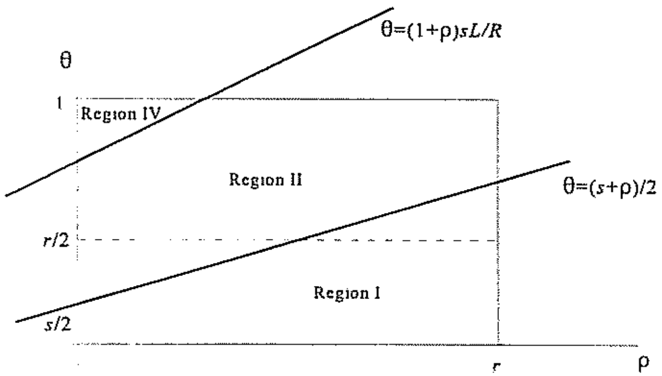
Case 1 ($R > R^*$)	Region I	$\pi^{EU} > \pi^{NET}$ ($P < P^*$) $\pi^{EU} < \pi^{NET}$ ($P^* < P$)
	Region II	$\pi^{EU} > \pi^{NET}$
Case 3 ($R < R^*$)	Region I	$\pi^{EU} < \pi^{NET}$
	Region II	$\pi^{EU} < \pi^{NET}$ ($P < P^*$) $\pi^{EU} > \pi^{NET}$ ($P^* < P$)

Figure 8-6 Summary of Results for NET vs US-type RTGS Systems

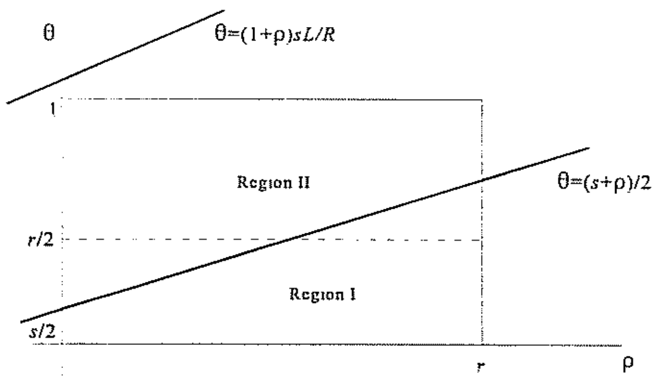
Case 1 ($R > R^{**}$)	Region I	$\pi^{US} > \pi^{NET}$ ($P < P^*$) $\pi^{US} < \pi^{NET}$ ($P^* < P$)
	Region II	$\pi^{US} > \pi^{NET}$
Case 3 ($R < R^{**}$)	Region I	$\pi^{US} < \pi^{NET}$
	Region II	$\pi^{US} < \pi^{NET}$ ($P < P^*$) $\pi^{US} > \pi^{NET}$ ($P^* < P$)

Both comparative analyses lead to the same conclusion. When the bank holds a relatively large amount of reserve, it prefers to settle small payments in the RTGS system and large ones in the NET system. When the bank does not hold a large reserve, it generally profits more from the RTGS system. As the bank retains less reserve, the picture changes. We find the region where our proposition holds both in the EU-type and the US-type. Otherwise, they will retain higher

(2) $r/2 < sL/R < 1$



(3) $1 < sL/R$



We omit Case 2: $R = R^{**}$ and Case 3: $R < R^{**}$, as the interpretation is the same as Figure 8-2 in 8.1.2 and Figure 8-3 in 8.1.3 respectively. We have a region which supports our proposition (Region II), and it becomes prevalent as sL/R shifts upward.

enough cash position to settle both P and ϵ; nevertheless if settling either one of two payments is feasible, the bank will go ahead and settle P but cancel ϵ.

We focus our analysis in case of $P \geq R_i$ $(i = A, B)$. There is no point holding reserve more than P since its payment obligation P can always be settled with the help of an overdraft. Under these circumstances, following situations may arise. If they both want to settle P at period 1, they need to overdraft $P - R_i$, but the overdraft position will disappear as the counterpart fulfills its obligation P at the same time. At period 2, further overdraft may come into existence, depending on the size and the flow of ϵ. Alternatively, the following situation might arise. Suppose, for example, Bank A decides to settle at period 1 while B decides not to. In the RTGS system, the payment order and the settlement are carried out simultaneously, so that P is transferred from Bank A to B at period 1. Now Bank B has an extra reserve P which can be used for settling its own part of payment P back to Bank A, or settling the stochastic payment whose size is yet known. We now formally analyse these cases. For the time being, we will neglect the cost of collateral holding, which we denote by cP $(0 < c < 1)$, until we come to the comparative study of different settlement regimes. No matter which period banks decide to settle, they will always incur cP with the preposted collateral amounting P. After having derived all the profit functions, we will move to the equilibrium analysis in Section 4.3, where the unique equilibrium turns out to be two banks settling without delay.

4.2 Choice of Settlement Timing

4.2.1 Both Bank A and B settle at period 1

The first case in the game examines when both banks settle their payments at period 1. In order to execute the payment, they have to overdraft by the amount of $P - R_i$. Then they receive P which becomes available as payments are finalised right away. At the end of period 1, two banks' balance sheets become identical to their initial ones. Then the stochastic payment ϵ, which lies between $-L$ and L, is realised. Note Bank A's intraday cash position is $R_A + P$ and B's is $R_B + P$. We derive the following cases, depending on the value of ϵ, which yield revenues and costs for Bank A and B.

Figure 4-2 Derivation of Profits for Bank A and B

ϵ	Situation	Revenue	Cost
$[R_A + P/(1+\rho), L]$	Bank A cancels the payment.	$r(L-R_A)$	$-\theta\epsilon$
		$r(L-R_B)$	$-\theta\epsilon$
$[R_A, R_A + P/(1+\rho)]$	Bank A withdraws R_A & overdrafts $\epsilon - R_A$.	$r(L-R_A)$	$-\rho(\epsilon-R_A)$
		$r(L-R_B)$	0
$[0, R_A]$	Bank A withdraws ϵ from reserve.	$r(L-R_A)$	0
		$r(L-R_B)$	0
$[-R_B, 0]$	Bank B withdraws $-\epsilon$ from reserve.	$r(L-R_A)$	0
		$r(L-R_B)$	0
$[-(R_B + P/(1+\rho)), -R_B]$	Bank B withdraws R_B & overdrafts $-\epsilon - R_B$.	$r(L-R_A)$	0
		$r(L-R_B)$	$-\rho(-\epsilon-R_B)$
$[-L, -(R_B + P/(1+\rho))]$	Bank B cancels the payment.	$r(L-R_A)$	$\theta\epsilon$
		$r(L-R_B)$	$\theta\epsilon$

Bank A's revenue & cost

Bank B's revenue & cost

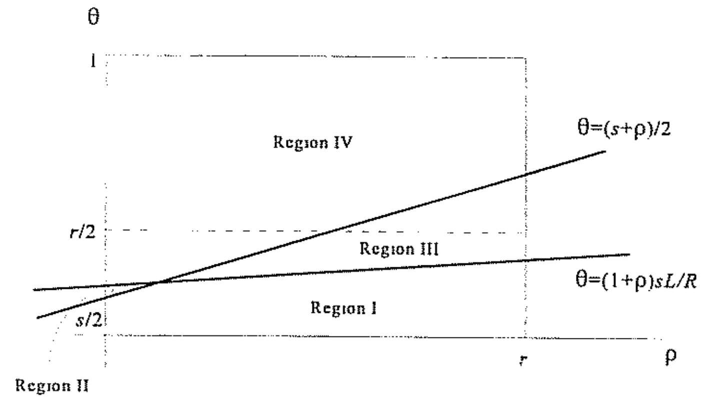
When ϵ exceeds Bank A's overnight cash position, i.e. $R_A + P/(1+\rho) < \epsilon < L$, the payment is cancelled, and both incur customer dissatisfaction costs. The revenue is straightforward - they receive the interest revenue from their holdings of TB. It will be credited as long as banks do not go insolvent, in which case the asset is taken away. As we will not observe insolvency caused

Figure 8-4 Comparison of Profits in RTGS and NET Systems

	$\Gamma(P)$ - concave $\theta < (\rho+s)/2$	$\Gamma(P)$ - convex $\theta > (\rho+s)/2$
Negative slope at $P = 0$ $\theta < (1+\rho)sL/R$	**Region I** $\pi^{US} > \pi^{NET}$ $(P < Q^*)$ $\pi^{US} < \pi^{NET}$ $(Q^* < P)$	**Region II** $\pi^{US} > \pi^{NET}$ $\pi^{US} < \pi^{NET}$ $(Q^{**} < P < Q^*)$
Positive slope at $P = 0$ $\theta > (1+\rho)sL/R$	**Region III** $\pi^{US} > \pi^{NET}$ $(P < Q^*)$ $\pi^{US} < \pi^{NET}$ $(Q^* < P)$	**Region IV** $\pi^{US} > \pi^{NET}$

All the regions represent the same as in Figure 8-1, and depending on where sL/R lies, we have three possible outcomes. The slight difference with Section 8.1.1 is, unlike Equation 8-3, the curvature constraint, Equation 8-5 does not go through the origin in the diagram, due to the existence of s.

(1) $0 < sL/R < r/2$



charged by the central bank, s, appears in the coefficients of P^2 and P. Taking the first and second derivatives, we get both the curvature constraint (Equation 8-5) and the slope constraint (Equation 8-6).

$$\frac{\partial^2 \Gamma(P)}{\partial P^2} = \frac{1}{2L}\left[\frac{2\theta-\rho-s}{(1+\rho)^2}\right] \quad (8\text{-}5)$$

$$\frac{1}{L}\left[\frac{\theta R}{1+\rho}-sL\right]. \quad (8\text{-}6)$$

We are not sure about the value of $\Gamma(0)$. From Equation 8-4, it is trivial $G(0) < 0$ and $G(L) > 0$, so there exists a unique $R = R^{**}$ where $G(R^{**}) = 0$. Depending on the value of $G(R)$, we again divide into following three cases. Let us denote Q^* and Q^{**} ($< Q^*$) as two values, if any, which derive $\Gamma(Q^*) = \Gamma(Q^{**}) = 0$.

8.2.1 Case 1: $R > R^{**}$

Qualitatively, the result does not alter from Section 8.1.1, and the following interpretation applies:

by the settlement obligation in the RTGS system,[9] banks always get revenues from TB, $r(L - R_A)$ and $r(L - R_B)$ for Bank A and B respectively. Once ϵ stays within the range of 0 and $R_A + P/(1+\rho)$, Bank A can deal with the payment without cancellation. When ϵ lies between R_A and $R_A + P/(1+\rho)$, it is carried out by withdrawing R_A from the reserve and overdrafting $(\epsilon - R_A)$ for the time being. The balance sheets for two banks, just after settling ϵ, look as follows:

Bank A		Bank B	
(Reserve)	(Liquidity)	(Reserve)	(Liquidity)
$R_A - \epsilon$	$L - \epsilon$	$R_B + \epsilon$	$L + \epsilon$
(TB)		(TB)	
$L - R_A$		$L - R_B$	

A's total liquidity position is $L - \epsilon$ after having settled ϵ. There is no change in the amount of TB, but the reserve is less than zero ($R_A - \epsilon < 0$), indicating it is in an overdraft position. In the meantime, B's reserve as well as the liquidity position increase by the amount of ϵ. After ϵ is settled at period 2, however, there is no further incoming payment for A. Bank A has no chance of recovering the overdraft position. At this point, the central bank could distrain A's collateral by an amount equivalent to the overdraft, $\epsilon - R_A$, to set off against the overdraft. But it is customary for this intraday credit to be converted into an overnight credit with the collateral and the official discount rate ρ ($< r$).[10] Bank A's balance sheet after the central bank extended the

[9] The bank would just cancel the payment, rather than carry it out and go bankrupt.

[10] The reason why it is converted overnight is that in practice the participating banks may be expecting incoming payments in the following day, so that the collateral distraint can be avoided. Our model, nonetheless, does not incorporate next day transactions, so we simply assume that the bank pays the amount of credit plus the interest rate. Whether it is paid back from the incoming transaction in the following day or not (in which case, collateral is taken), the cost incurred by having relied on the central bank's overnight credit does not alter. Another concern arising from the conversion of the intraday credit into overnight is that it may overestimate the cost factor. If the collateral is simply distrained on the same day, banks need not pay the interest rate on top of it. As we will see later in the comparative analysis, however, we are not interested in an absolute level of profits but the relative magnitude, obtained in different settlement regimes, and this possible overestimation bias of the cost factor will be

overnight credit is rewritten as follows:

Bank A		Bank B	
(Reserve)	(Liquidity)	(Reserve)	(Liquidity)
0	$L - R_A$	$R_B + \epsilon$	$L + \epsilon$
(TB)		(TB)	
$L - R_A$		$L - R_B$	

Next day (period 3), $(1 + \rho)(\epsilon - R_A)$ is charged, and Bank A incurs the net-out cost $\rho(\epsilon - R_A)$. Once ϵ lies between 0 and R_A, the payment is small enough for Bank A to handle with its reserves. No overdraft is necessary, with no further costs incurred.

Next we deal with cases if ϵ is negative, meaning Bank B is now responsible for paying ϵ to Bank A. If ϵ lies between 0 and $-R_B$, B manages to pay out from its own reserve, and no cost is incurred. Once ϵ exceeds $-R_B$, Bank B has to overdraft by the amount of $-R_B - \epsilon$. The balance sheets after B's overdraft are shown below:

Bank A		Bank B	
(Reserve)	(Liquidity)	(Reserve)	(Liquidity)
$R_A - \epsilon$	$L - \epsilon$	$R_B + \epsilon$	$L + \epsilon$
(TB)		(TB)	
$L - R_A$		$L - R_B$	

ϵ itself is negative and its absolute value is greater than R_B, so $R_B + \epsilon$ is negative. This time, Bank B's reserve position is less than zero. Similar to the previous case, where Bank A had to rely on the overnight credit, B's overdraft position cannot be recovered, and it is converted into the central bank's overnight credit. The balance sheets after the overnight credit are as follows:

cancelled out each other.

dissatisfaction while the NET system imposes the cost of the overnight credit. Therefore, high θ is likely to benefit the NET system, i.e. the bank is likely to make a higher profit in the NET system, while low θ may result in a higher profit in the RTGS system. As P goes up, however, the profit in the RTGS system catches up and eventually exceeds the profit in the NET system, so it is interpreted that the bank makes a higher profit in the NET system if the payment size is small but earns more in the RTGS system if payments are large. The three diagrams are again the same as Case I; nevertheless, whether Region II or IV prevails, our result does not alter - we have found the case in which the proposition predominantly holds true.

8.2 Net Settlement vs Gross Settlement (US-type)

Our second analysis deals with the comparison between the NET and the US-type RTGS systems. We find that all the interpretations derived in 8.1 hold. We will briefly show the results and the diagrams in this sub-section; analysis of the results is set out in Section 8.3. We denote the profit from this type of the RTGS system by $\pi^{US}(P)$. Because we only deal with the non-delay equilibrium in the US-type, $\pi^{US}(P)$ is equivalent with the profit in Equation 5-1.[29] Let $\Gamma(P)$ be $\pi^{US}(P) - \pi^{NET}(\sigma)$. Again, we are interested in the curvature of this function and the sign as well as the slope of $\Gamma(0)$. First of all, we calculate $\Gamma(P)$.

$$\Gamma(P) = \frac{1}{4L}\left[\frac{2\theta-\rho-s}{(1+\rho)^2}\right]P^2 + \frac{1}{L}\left[\frac{\theta R}{1+\rho} - sL\right]P + \frac{1}{L}\left[\frac{\theta(\theta-r)}{2\theta-r}\right]R^2 + G(R)$$

where (8-4)

$$G(R) = K(R) + sR.$$

In comparison with Equation 8-1, this equation is again a quadratic function of P, and the fee

[29] Again, we drop the subscripts, so π_A^{II} and π_B^{II} become identical. This is what we call $\pi^{US}(P)$.

interpretation in Section 8.1.1); hence the NET system continues to be preferred. In Region II, on the other hand, any combination of ρ and θ, except high ρ and low θ, will have an overall positive effect on the profit in the RTGS system, as the payment size increases. Thus the RTGS system is eventually preferred.

8.1.3 Case 3: $R < R^*$

Finally we have the case where $\Omega(0)$ is negative. It again contains the region which supports our proposition.

Figure 8-3 Comparison of Profits in RTGS and NET

	$\Omega(P)$ - concave $\theta < \rho/2$	$\Omega(P)$ - convex $\theta > \rho/2$
Negative slope at $P = 0$ $\theta < (1+\rho)cL/R$	**Region I** $\pi^{EU} < \pi^{NET}$	**Region II** $\pi^{EU} < \pi^{NET}$ $(P < P^*)$ $\pi^{EU} > \pi^{NET}$ $(P^* < P)$
Positive slope at $P = 0$ $\theta > (1+\rho)cL/R$	**Region III** $\pi^{EU} < \pi^{NET}$ $\pi^{EU} > \pi^{NET}$ $(P^{**} < P < P^*)$	**Region IV** $\pi^{EU} < \pi^{NET}$ $(P < P^*)$ $\pi^{EU} > \pi^{NET}$ $(P^* < P)$

What is different from Figure 8-2 is Region IV. As in Region II, this region also suggests that the bank receives higher profit in the NET system, if the payment size is small and in the RTGS system if the payment is large. Consider, for example, $R = P = 0$. The cash position in the RTGS system is 0, so any payment will be cancelled, incurring the cost of customer dissatisfaction for $\epsilon \in [0, L]$. In the case of the NET system, the cash position is still $L/(1+r)^2$, so the bank incurs the cost of customer dissatisfaction if $\epsilon \in [L/(1+r)^2, L]$ but if $\epsilon \in [0, L/(1+r)^2]$, the payment is undertaken, and the cost of the overnight credit is levied. The difference between the costs in RTGS and NET systems is, for $\epsilon \in [0, L/(1+r)^2]$, the RTGS system imposes the cost of customer

Bank A (Reserve)	Bank A (Liquidity)	Bank B (Reserve)	Bank B (Liquidity)
$R_A - \epsilon$	$L - \epsilon$	0	$L - R_B$
(TB)		(TB)	
$L - R_A$		$L - R_B$	

Next day, $(1 + \rho)(-\epsilon - R_B)$ is charged, so Bank B's net-out loss is $\rho(-\epsilon - R_B)$. If ϵ exceeds $-(R_B + P)$, the payment is rejected, and the cost of customer dissatisfaction is incurred. We can now show the overall profit function for Bank A as π_A^{ij}, meaning Bank A's profit if it settles P at period i while Bank B settles at period j, and similarly π_B^{ij} for Bank B.

$$\begin{aligned}
\pi_A^{11} &= \int_{-L}^{L} r(L - R_A)\, f(\epsilon)d\epsilon - \int_{R_A + \frac{P}{1+\rho}}^{L} \theta\epsilon\, f(\epsilon)d\epsilon \\
&\quad - \int_{R_A}^{R_A + \frac{P}{1+\rho}} \rho(\epsilon - R_A)\, f(\epsilon)d\epsilon + \int_{-L}^{-(R_B + \frac{P}{1+\rho})} \theta\epsilon\, f(\epsilon)d\epsilon \\
\pi_B^{11} &= \int_{-L}^{L} r(L - R_B)\, f(\epsilon)d\epsilon - \int_{R_A + \frac{P}{1+\rho}}^{L} \theta\epsilon\, f(\epsilon)d\epsilon \\
&\quad - \int_{-(R_B + \frac{P}{1+\rho})}^{-R_B} \rho(-\epsilon - R_B)\, f(\epsilon)d\epsilon + \int_{-L}^{-(R_B + \frac{P}{1+\rho})} \theta\epsilon\, f(\epsilon)d\epsilon
\end{aligned} \tag{4-1}$$

$f(\epsilon)$ is the density function of ϵ. Bank A's profit consists of four elements: the revenue from TB holding ($\epsilon \in [-L, L]$); the cost of customer dissatisfaction due to A's cancellation of payment ($\epsilon \in [R_A + P/(1+\rho), L]$); the cost of overnight credit ($\epsilon \in [R_A, R_A + P/(1+\rho)]$); and the cost of customer dissatisfaction due to B's cancellation of payment ($\epsilon \in [-L, -(R_B + P/(1+\rho))]$). Bank B's profit also consists of the same four elements.[11]

[11] In the following, we will not show the profit function in each section, as they are summarised in the profit matrix later in the equilibrium analysis.

4.2.2 Bank A settles at period 1, but B settles at period 2

The next case involves a mismatching of settlement timings between two banks. It is only Bank A which executes the payment at period 1. Already overdrafting by the amount of $P - R_A$, the remaining overdraft available for A is R_A (= $P - (P - R_A)$), which represents A's overall intraday cash position. On the other hand, B's intraday cash position is $R_B + 2P$ as B holds $R_B + P$ as its reserve and the full overdraft by the amount of P is still accessible. We have the following situations depending on the value of ϵ:

Figure 4-3 Derivation of Profits for Bank A and B

ϵ	Situation	Revenue	Cost
$[R_A, L]$	Bank A cancels the payment.	$r(L-R_A)$	$-\theta\epsilon$
		$r(L-R_B)$	$-\theta\epsilon$
$[0, R_A]$	Bank A overdrafts ϵ, but the position is recovered.	$r(L-R_A)$	0
		$r(L-R_B)$	0
$[-R_B, 0]$	Bank B withdraws $-\epsilon$ from reserve.	$r(L-R_A)$	0
		$r(L-R_B)$	0
$[-(R_B + P/(1+\rho)), -R_B]$	Bank B withdraws R_B & overdrafts $-\epsilon-R_B$.	$r(L-R_A)$	0
		$r(L-R_B)$	$-\rho(-\epsilon-R_B)$
$[-L, -(R_B + P/(1+\rho))]$	Bank B cancels the payment.	$r(L-R_A)$	$\theta\epsilon$
		$r(L-R_B)$	$\theta\epsilon$

Reflecting A's limited cash position, whenever ϵ exceeds R_A, the payment is rejected. This higher probability of payment ϵ being rejected is apparently caused by the shortage of liquidity on A's part. If ϵ lies between 0 and R_A, i.e. ϵ is within the range of A's intraday cash position, Bank A manages to carry out its obligation by overdrafting ϵ. If Bank B is obliged to pay ϵ, i.e. ϵ is negative, the situation is the same as in 4.2.1. This is due to the fact that Bank B's balance

raised if $\epsilon \in [R, R + P/(1+\rho)]$, so lowering R does not have any effect on the size of this region. However, in the NET system, the overnight credit has to be raised if $\epsilon \in [R, \sigma]$. Given the optimal value of $\sigma = rR/(1+r) + L/(1+r)^2$, lowering R does actually widen the size of this region, hence the bank has to incur the greater cost of the overnight credit.

We can draw exactly the same three diagrams as in Case 1, but the interpretation is absolutely different. We still have Region IV where the RTGS system is strictly preferred. As cL/R rises, Region II starts appearing and eventually it will dominate most of the region. Our understanding of Region II is identical to our proposition - the bank prefers to settle the small size payment in the NET system and large ones in the RTGS system. Here, we have shown that the coexistence of the RTGS and NET systems is feasible in Region II and the region can become dominant under certain conditions, namely that the bank holds a certain amount of reserve ($R = R^*$),[28] and cL/R is large. We are not sure whether the cost of collateral holding has to be large, from the second condition above. In fact, it is very likely that even if c is small, cL/R as a whole is still larger than unity, due to the fact that the total liquidity the bank holds tend to be much larger than the amount of reserve. So, out of the three diagrams, we think (3) $1 < cL/R$ is the most likely scenario. In this case, we have two variants. In Region I, we have the combination of high ρ and low θ. As we saw, holding a lower level of reserve will result in increased costs of customer dissatisfaction in the RTGS system, while it will result in increased costs of overnight credit in the NET system. If the rate of interest on the central bank credit is high, then the cost of overnight funds is increased and there are additional costs of customer dissatisfaction. In this case, the RTGS system becomes a less attractive proposition, and the bank tends to prefer to settle in the NET system. Even though the payment size increases, the combination of high ρ and low θ is likely to have an overall negative effect on the profit in the RTGS system (see the

[28] Admittedly, this is non-generic.

Figure 8-2 Comparison of Profits in RTGS and NET Systems

	$\Omega(P)$ - concave $\theta < \rho/2$	$\Omega(P)$ - convex $\theta > \rho/2$
Negative slope at $P = 0$ $\theta < (1+\rho)cL/R$	**Region I** $\pi^{EU} < \pi^{NET}$	**Region II** $\pi^{EU} < \pi^{NET}$ $(P < P^*)$ $\pi^{EU} > \pi^{NET}$ $(P^* < P)$
Positive slope at $P = 0$ $\theta > (1+\rho)cL/R$	**Region III** $\pi^{EU} > \pi^{NET}$ $(P < P^*)$ $\pi^{EU} < \pi^{NET}$ $(P^* < P)$	**Region IV** $\pi^{EU} > \pi^{NET}$

In Region I, the bank prefers the NET system all the time, while in Region IV, it prefers the RTGS system. Region II yields higher profit in the NET system if the payment is small, and it yields higher profit in the RTGS system if the payment is large. Region III is the opposite of Region II. Compared with Figure 8-1, Regions I and II show different situations.

What makes Figure 8-2 different from Figure 8-1 is the amount of reserve.[27] We first look at the effect on the cost of customer dissatisfaction. With respect to the profit in the RTGS system, any payment greater than the overnight cash position, $R + P/(1+\rho)$, has to be cancelled, so lowering R will fully shift this cash position downward, increasing the possibility of payment cancellation, hence increasing the cost of customer dissatisfaction. For the profit in the NET system, the optimal value of the credit limit is a function of R, so holding less reserve will lower this value, but only by the factor of $r/(1+r)$ (see Equation 7-2). This will also increase the possibility of payment cancellation, thus increasing the cost of customer dissatisfaction, but with relatively smaller magnitude, compared with the RTGS system. Secondly, holding less reserve will oblige the bank to rely more on overnight credit. In the case of the RTGS system, the credit has to be

[27] The following interpretation also applies to Case 3.

sheet becomes identical to 4.2.1, once it carries out its obligation P at period 2.

4.2.3 Bank A settles at period 2, but B settles at period 1

We have another mismatching case where it is Bank A this time which delays its payment. Because this case and the previous one are symmetric, all the revenue and the cost identified in the previous sub-section carry over to this case by swapping the role of Bank A and B.

4.2.4 Both Bank A and B settle at period 2

The final case is if both banks decide to delay their payments. The balance sheet has not changed from the initial state, but none of the transactions has taken place yet. The intraday cash positions of two banks are $R_A + P$ and $R_B + P$ respectively. Depending on the value of ϵ, we have the following situations:

Figure 4-4 Derivation of Profits for Bank A and B

ϵ	Situation	Revenue	Cost
$[R_A, L]$	Bank A cancels the payment.	$r(L-R_A)$	$-\theta\epsilon$
		$r(L-R_B)$	$-\theta\epsilon$
$[0, R_A]$	Bank A overdrafts ϵ, but no overnight credit.	$r(L-R_A)$	0
		$r(L-R_B)$	0
$[-R_B, 0]$	Bank B overdrafts $-\epsilon$, but no overnight credit.	$r(L-R_A)$	0
		$r(L-R_B)$	0
$[-L, -R_B]$	Bank B cancels the payment.	$r(L-R_A)$	$\theta\epsilon$
		$r(L-R_B)$	$\theta\epsilon$

If ϵ is larger than R_A, P is carried out but ϵ cannot be made due to lack of sufficient cash position,[12] and Bank A cancels the payment. If ϵ is smaller than R_A, A has enough cash position to settle two transactions. At the same time as A's two payments are made, B's payment of P will eventually deletes A's overdraft. So, the overnight credit need not be relied on. Similarly, if ϵ turns out to be B's payment obligation, it can be dealt with only if ϵ is in the range of 0 and $-R_B$. Again A's payment of P will delete B's overdraft position; hence no further cost is incurred. Finally we have a situation where Bank B cannot execute two payments due to a lack of cash. B executes P under priority rule, but ϵ is rejected.

4.3 Equilibrium Analysis

The four cases above are summarised in the profit matrix (see Figure 4-5), and we derive the following proposition.

Proposition 1:

There exists a unique equilibrium where two banks choose to settle at period 1.

Proof:

See Appendix.

The equilibrium seems quite intuitive. Because the overdraft is provided without any charge, there is no incentive for the participants to delay the payments. In addition, the delay will increase the probability of its counterpart being forced to reject the subsequent payment due to a lack of cash, and this has a counter-effect on the delayed agent which has to incur the cost of customer dissatisfaction due to the counterpart's payment cancellation.

[12] P has priority over ϵ, by assumption, as it has been realised earlier than the stochastic payment.

payment size increases.[26] The reduction of the customer dissatisfaction cost (high θ) surpasses the increases in the cost of overnight credit (ρ) and the cost of collateral holding (low c). Hence, the RTGS system is preferred, as in Region IV. On the other hand, the combination of high ρ, low θ and low c, is likely to lead to an overall negative effect - as the payment size goes up, the negative factors outweigh the benefit of saving the cost of customer dissatisfaction, and eventually the NET system yields a higher profit for the bank, as in Region I and III. As cL/R increases, the picture is not so much altered. The region below the curvature constraint always represents the one in which the bank prefers to settle large payments in the NET system, and small ones in the RTGS system (Region I). We have another region, above the curvature constraint, where the RTGS system is predominantly preferred (Region II).

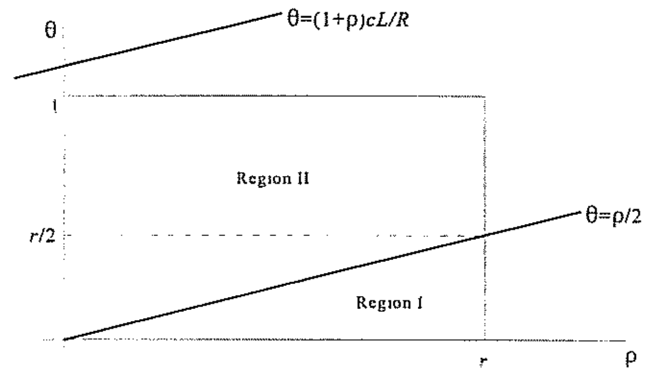
The result does not support our proposition that the large payments should be dealt with in the RTGS system and small ones in the NET system. In fact, none of the regions coincides with what we wanted to show. Instead, what we derived here is a completely different picture. The bank would be better off settling large payments in the NET system and small ones in the RTGS system (Region I and III), as the higher cost of overnight credit and the lower cost of customer dissatisfaction are likely to make the profit in the RTGS system a decreasing function of the payment size. Otherwise, the bank predominantly prefers the RTGS system (Region II), since the lower cost of overnight credit and the higher cost of customer dissatisfaction make the profit in the RTGS system to an increasing function of the payment size.

8.1.2 Case 2: $R = R^*$

Next we have the case where $\Omega(0) = 0$. This is completely different from the previous case. The interpretation of four regions is as follows:

[26] c may have to be extremely low to make cL/R less than $r/2$, given the total liquidity (L) tends to be much larger than the reserve (R).

(3) ι < cL/R



These diagrams show that there are three factors which affect the bank's non-delay equilibrium profit in the RTGS system. As the payment size increases, it has a positive effect on the profit because the cost of customer dissatisfaction declines - the bank can deal with larger-sized payments, which will reduce the possibility of the subsequent payment being cancelled as a result of preposting more collateral, enabling the bank to have a greater overdraft. Nevertheless, there are two channels which negatively affect the profit. One is an increase in the total cost of overnight credit - having more collateral, the bank can manage even larger-sized payments, hence there is a greater chance of the daylight overdraft being converted into the overnight credit, and the total cost of paying additional charges on this credit increases. The other is an increase in the cost of collateral holding, which is proportional to the amount of collateral. Let us examine these positive and negative effects on the profit more closely.

First, if cL/R is low, we have two dominating situations. The combination of low ρ, high θ and low c, is likely to provide an overall positive effect on the profit in the RTGS system as the

Figure 4-5 Profit Matrix of Model 1-A

		BANK A: Period 1	BANK A: Period 2
BANK B	Period 1	$\int_{-L}^{L} r(L-R_A)f(\epsilon)d\epsilon - \int_{R_A+\frac{\rho}{1+\rho}}^{L}\theta\epsilon f(\epsilon)d\epsilon$ $- \int_{R_A}^{R_A+\frac{\rho}{1+\rho}}\rho(\epsilon-R_A)f(\epsilon)d\epsilon + \int_{-L}^{-(R_B+\frac{\rho}{1+\rho})}\theta\epsilon f(\epsilon)d\epsilon$ $\int_{-L}^{L} r(L-R_B)f(\epsilon)d\epsilon - \int_{R_A+\frac{\rho}{1+\rho}}^{L}\theta\epsilon f(\epsilon)d\epsilon$ $- \int_{-(R_B+\frac{\rho}{1+\rho})}^{-R_B}\rho(-\epsilon-R_B)f(\epsilon)d\epsilon + \int_{-L}^{-(R_B+\frac{\rho}{1+\rho})}\theta\epsilon f(\epsilon)d\epsilon$	$\int_{-L}^{L} r(L-R_A)f(\epsilon)d\epsilon - \int_{R_A+\frac{\rho}{1+\rho}}^{L}\theta\epsilon f(\epsilon)d\epsilon$ $- \int_{R_A}^{R_A+\frac{\rho}{1+\rho}}\rho(\epsilon-R_A)f(\epsilon)d\epsilon + \int_{-L}^{-R_B}\theta\epsilon f(\epsilon)d\epsilon$ $\int_{-L}^{L} r(L-R_B)f(\epsilon)d\epsilon$ $- \int_{R_A+\frac{\rho}{1+\rho}}^{L}\theta\epsilon f(\epsilon)d\epsilon + \int_{-L}^{-R_B}\theta\epsilon f(\epsilon)d\epsilon$
	Period 2	$\int_{-L}^{L} r(L-R_A)f(\epsilon)d\epsilon$ $- \int_{R_A}^{L}\theta\epsilon f(\epsilon)d\epsilon + \int_{-L}^{-(R_B+\frac{\rho}{1+\rho})}\theta\epsilon f(\epsilon)d\epsilon$ $\int_{-L}^{L} r(L-R_B)f(\epsilon)d\epsilon - \int_{R_A}^{L}\theta\epsilon f(\epsilon)d\epsilon$ $- \int_{-(R_B+\frac{\rho}{1+\rho})}^{-R_B}\rho(-\epsilon-R_B)f(\epsilon)d\epsilon + \int_{-L}^{-(R_B+\frac{\rho}{1+\rho})}\theta\epsilon f(\epsilon)d\epsilon$	$\int_{-L}^{L} r(L-R_A)f(\epsilon)d\epsilon$ $- \int_{R_A}^{L}\theta\epsilon f(\epsilon)d\epsilon + \int_{-L}^{-R_B}\theta\epsilon f(\epsilon)d\epsilon$ $\int_{-L}^{L} r(L-R_B)f(\epsilon)d\epsilon$ $- \int_{R_A}^{L}\theta\epsilon f(\epsilon)d\epsilon + \int_{-L}^{-R_B}\theta\epsilon f(\epsilon)d\epsilon$

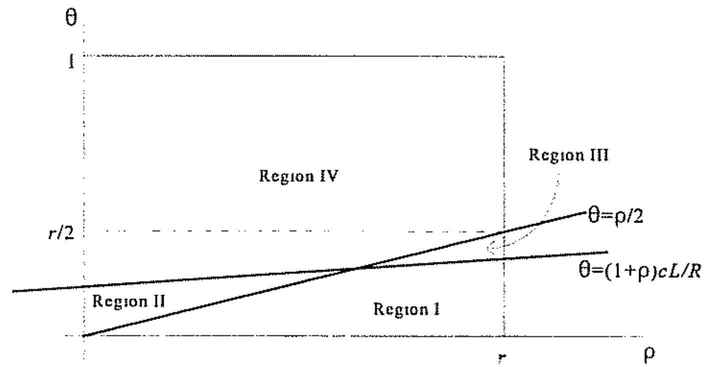
5. THE MODEL OF GROSS SETTLEMENT SYSTEM: CASE OF MODEL 1-B

We now move to the second model of the RTGS system, which is the crude example of Fedwire in the US. The main distinction with Model 1-A is as follows: (1) We again assume the central bank's intraday credit is available. Unlike Model 1-A, however, it requires no collateral but charges s if overdraft is realised at either period; (2) Although the overdraft does not require collateral, which normally sets the maximum overdraft available to the collateral poster, banks in this model do not have unlimited access to overdraft in this model, either. The cap is normally set based on each bank's risk-based capital. Moreover, the Federal Reserve holds the right to refuse this credit provision if they have any doubt about a particular bank's creditworthiness or operational capabilities; in an extreme case, they can require collateral to secure the credit. In this model, we simply assume the debit cap is set at the level P, in order to make the comparison with Model 1-A easier, and we call the sum of the reserve and the debit cap, "intraday cash position" in this model; (3) Because the overdraft is not tied up with collateral posting, if the participant cannot recover its overdraft position after having settled all its obligations, the overdraft is switched to the overnight credit which requires the collateral and charges the official discount rate. Note this facility is usually provided only to help the "solvent but illiquid" bank.[13] [14]

[13] The famous example of supporting the solvent but illiquid bank is Bank of New York in 1985. The deficit in the bank's reserve reached more than 22 billion US dollars by the end of the day, due to the computer mishandling of the security transaction. This liquidity risk was resolved by the intervention of the Federal Reserve, which provided the overnight loan from the discount window.

[14] Admittedly, it is still arguable whether the Federal Reserve is capable of distinguishing between the solvent and insolvent banks through monitoring the daily operation of participants in real time.

(1) $0 < cL/R < r/2$



(2) $r/2 < cL/R < 1$

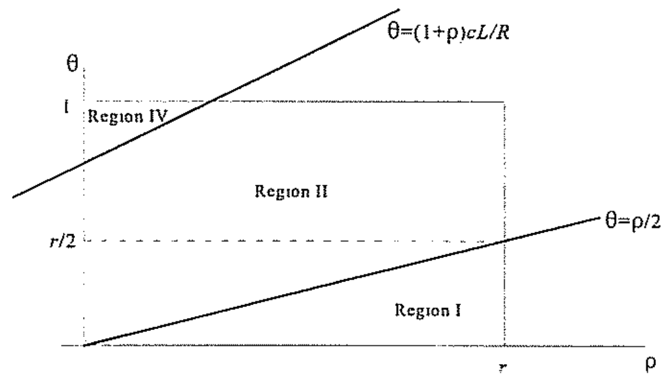


Figure 8-1 Comparison of Profits in RTGS and NET Systems

	$\Omega(P)$ - concave $\theta < \rho/2$	$\Omega(P)$ - convex $\theta > \rho/2$
Negative slope at $P = 0$ $\theta < (1+\rho)cL/R$	**Region I** $\pi^{EU} > \pi^{NET}$ $(P < P^*)$ $\pi^{EU} < \pi^{NET}$ $(P^* < P)$	**Region II** $\pi^{EU} > \pi^{NET}$ $\pi^{EU} < \pi^{NET}$ $(P^{**} < P < P^*)$
Positive slope at $P = 0$ $\theta > (1+\rho)cL/R$	**Region III** $\pi^{EU} > \pi^{NET}$ $(P < P^*)$ $\pi^{EU} < \pi^{NET}$ $(P^* < P)$	**Region IV** $\pi^{EU} > \pi^{NET}$

Two regions, Region I and III, show that the bank receives higher profit by settling P in the RTGS system if the payment is smaller than P^*, and it does so in the NET system if the payment is larger than P^* In Region II, it is generally better for the bank to settle in the RTGS system; however, there may be cases the bank benefits more from the NET system, depending on whether there exist P^* and P^{**}. In Region IV, the bank is always better off in the RTGS system, irrespective of the size of payments. Which region is more likely observed? We can summarise the answer in the following diagrams. In the ρ-θ dimension, the curvature constraint, $\theta = \rho/2$, is fixed, whereas the position of the slope constraint, $\theta = (1+\rho)cL/R$, depends on the value of cL/R. We will see how the area of each region changes as we shift up the slope constraint. Note that the value of ρ is restricted to $\rho \in [0, r]$, as the official discount rate normally sets the lower boundary for the interest rate in the market.

We now move to the derivation of the profit, followed by the equilibrium analysis, where we find the strategy to delay by one of the two banks also consists an equilibrium along with the non-delay equilibrium.

5.1 Choice of Settlement Timing

5.1.1 Both Bank A and B settle at period 1

All the situations are exactly identical to the synchronous case in the previous model (see Figure 4-2 of Section 4.2.1), except one difference in the cost side - a fee is now charged for the use of an overdraft, introducing another major cost component. If we again denote Bank A's profit by π_A^{11} if both banks settle P at period 1, and similarly π_B^{11} for Bank B, we have the following:

$$\begin{aligned}
\pi_A^{11} &= \int_{-L}^{L} r(L-R_A)\, f(\epsilon)d\epsilon \;-\; \int_{-L}^{L} s(P-R_A)\, f(\epsilon)d\epsilon \\
&- \int_{R_A}^{R_A+\frac{P}{1+\rho}} s(\epsilon-R_A)\, f(\epsilon)d\epsilon \;-\; \int_{R_A+\frac{P}{1+\rho}}^{L} \theta\epsilon\, f(\epsilon)d\epsilon \\
&- \int_{R_A}^{R_A+\frac{P}{1+\rho}} \rho(\epsilon-R_A)\, f(\epsilon)d\epsilon \;+\; \int_{-L}^{-(R_B+\frac{P}{1+\rho})} \theta\epsilon\, f(\epsilon)d\epsilon \\
\pi_B^{11} &= \int_{-L}^{L} r(L-R_B)\, f(\epsilon)d\epsilon \;-\; \int_{-L}^{L} s(P-R_B)\, f(\epsilon)d\epsilon \\
&- \int_{-(R_B+\frac{P}{1+\rho})}^{-R_B} s(-\epsilon-R_B)\, f(\epsilon)d\epsilon \;-\; \int_{R_A+\frac{P}{1+\rho}}^{L} \theta\epsilon\, f(\epsilon)d\epsilon \\
&- \int_{-(R_B+\frac{P}{1+\rho})}^{-R_B} \rho(-\epsilon-R_B)\, f(\epsilon)d\epsilon \;+\; \int_{-L}^{-(R_B+\frac{P}{1+\rho})} \theta\epsilon\, f(\epsilon)d\epsilon
\end{aligned} \tag{5-1}$$

The profit function for Bank A consists of six elements: the revenue from TB holdings ($\epsilon \in [-L, L]$); the fees charged for the overdraft ($\epsilon \in [-L, L]$ and $\epsilon \in [R_A, R_A + P/(1+\rho)]$)[15]; the cost of

[15] If $R_A < \epsilon < R_A + P/(1+\rho)$: the payment is within Bank A's cash position, Bank A has to overdraft by the amount of $\epsilon - R_A$ at period 2, which further generates the fee $s(\epsilon - R_A)$ on top of the fee charged for the overdraft at period 1.

customer dissatisfaction due to A's cancellation of payment ($\epsilon \in [R_A + P/(1+\rho), L]$); the same cost due to B's cancellation of payment ($\epsilon \in [-L, -(R_B + P/(1+\rho))]$); and finally the cost of overnight credit ($\epsilon \in [R_A, R_A + P/(1+\rho)]$). Bank B's function qualitatively consists of the same elements.[16]

5.1.2 Bank A settles at period 1, but B settles at period 2

Again the situations are identical to Section 4.2.2, except the fee charged on the overdraft. This time, there is an advantage in delaying the payment. While A is still charged $s(P - R_A)$ for its overdraft at period 1, Bank B has avoided paying the fee, since it has decided to delay the payment, so it will not incur $s(P - R_B)$. If ϵ exceeds R_A, ϵ will be cancelled. If ϵ is less than R_A, A will be further forced to overdraft ϵ, which will add another fee charged on A. However, as Bank B pays P at period 2, A's overdraft position is overcome, so Bank A need not rely on the overnight credit. If Bank B pays ϵ to A (i.e. ϵ is negative), all the relevant revenues and costs are identical to 5.1.1, except that B does not bear the overdraft fee at period 1.

5.1.3 Bank A settles at period 2, but B settles at period 1

We have again the symmetric version of the previous case. This time it is Bank A which has delayed the payment, so it has avoided being charged the fee. The revenue and cost are exactly the same if one exchanges the role of A and B in 5.1.2.

5.1.4 Both Bank A and B settle at period 2

Finally we have a case where two banks delay the payments (see Section 4.2.4). If ϵ exceeds A's cash position, A cancels the payment. The same for $-L < \epsilon < -R_B$, where B rejects the payment. If $0 < \epsilon < R_A$, Bank A overdrafts $R_A - P$ to settle P at period 2 and further ϵ to settle the payment

[16] In the following, profit functions are not shown until we deal with the equilibrium analysis.

$$\frac{\partial^2 \Omega(P)}{\partial P^2} = \frac{1}{2L}\left[\frac{2\theta - \rho}{(1+\rho)^2}\right] \qquad (8\text{-}3)$$

shows the curvature constraint. $\Omega(0)$ can be either positive or negative, as it depends on the constant term, $K(R)$. Given,

$$K(0) = -\left[\frac{2\theta - r}{4(1+r)^4}\right]L < 0$$

$$K(L) = \left[\frac{4\theta r + 6\theta r^2 + 4\theta r^3 + r^3}{4L(1+r)^4}\right]L^2 > 0,$$

and $K(R)$ is convex, there exists a unique $R = R^*$ where $R^* \in [0, L]$ and $K(R^*) = 0$. With these algebraic results, let us examine the following three cases in turn.

8.1.1 Case 1: $R > R^*$

First of all, if R is greater than R^*, $K(R)$ will be positive. This means $\Omega(P)$ will be positive at $P = 0$. We know $\Omega(P)$ is a quadratic function of P. Depending on the curvature and the slope at $P = 0$, we have the following possibilities summarised in Figure 8-1. Let us denote P^* and P^{**} ($< P^*$), if any, as $\Omega(P^*) = \Omega(P^{**}) = 0$.

manner, because they are in a perfectly symmetric situation. We denote the profit from the RTGS system by $\pi^{EU}(P)$, which is given by Equation 4-1.[25] The profit from the NET system is given in Equation 7-2. Let $\Omega(P)$ be $\pi^{EU}(P) - \pi^{NET}(\sigma)$. In the analysis, what we need to find is the following two elements:

- curvature of $\Omega(P)$
- sign of $\Omega(0)$ and its slope

These two conditions will cover all the possibilities of how $\Omega(P)$ behaves. For convenience, we call the equation which disciplines the curvature of $\Omega(P)$, "curvature constraint," and similarly "slope constraint" for the equation which determines the slope of the function at $P = 0$. Substituting Equation 7-2 and 4-1 into $\Omega(P)$, we get the following result.

$$\Omega(P) = \frac{1}{4L}[\frac{2\theta-\rho}{(1+\rho)^2}]P^2 + \frac{1}{L}[\frac{\theta R}{1+\rho}-cL]P + K(R)$$

where (8-1)

$$K(R) = [\frac{2\theta+4\theta r+r}{4L(1+r)^2}]R^2 - [\frac{(2\theta+1)r}{2(1+r)^3}]R - [\frac{2\theta-r}{4(1+r)^4}]L.$$

Taking the first derivative, the slope constraint is

$$\frac{1}{L}[\frac{\theta R}{1+\rho}-cL]. \qquad (8\text{-}2)$$

And the second derivative,

[25] Dropping the subscripts of π_A^{11} and π_B^{11} in Equation 4-1, they become identical. Adding the cost of collateral holding, cP, it comes to be what we define as $\pi^{EU}(P)$.

ϵ, whereas B overdrafts by the amount of $R_B - P$ at period 2, in order to settle its obligation. Finally, if $-R_B < \epsilon < 0$, B overdrafts $R_B - P$ and ϵ to settle two payments while A overdrafts $R_A - P$

5.2 Equilibrium Analysis

The profit matrix, Figure 5-1, summarises the analysis in Section 5.1, and we have the following proposition.

Proposition 2a (non-delay case):

$$If \quad s < \frac{\theta}{4L}(P+2R_A),$$

then there exists a unique equilibrium where two banks settle at period 1.

Proposition 2b-1 (delay case):

Define $\Delta^{US}(R_i)$ as Bank j's difference in profit functions when Bank i settles at period 1.

$$If \;\; s > \frac{\theta}{4L}(P+2R_A), \; and \; R_i \in [0, R_i^*] \;\; where \;\; \Delta^{US}(R_i^*)=0 \; and \; j \neq i,$$

then there exist two pure strategy equilibria where one bank settles at period 1 while the other settles at period 2, and one equilibrium with mixed strategies.

Proposition 2b-2 (non-delay case):

$$If \quad s > \frac{\theta}{4L}(P+2R_A), \quad and \quad R_i \in [R_i^*, P],$$

then there exists a unique equilibrium where both banks settle at period 1.

Proof:

See Appendix.

Figure 5-1 Profit Matrix of Model 1-B

		BANK A: Period 1	BANK A: Period 2
BANK B	Period 1	$\int_{-L}^{L} r(L-R_A)f(\epsilon)d\epsilon - \int_{-L}^{L} s(P-R_A)f(\epsilon)d\epsilon - \int_{R_A}^{R_A+\frac{P}{1+\rho}} s(\epsilon-R_A)f(\epsilon)d\epsilon$ $- \int_{R_A+\frac{P}{1+\rho}}^{L} \theta\epsilon f(\epsilon)d\epsilon - \int_{R_A}^{R_A+\frac{P}{1+\rho}} \rho(\epsilon-R_A)f(\epsilon)d\epsilon + \int_{-L}^{-(R_B+\frac{P}{1+\rho})} \theta\epsilon f(\epsilon)d\epsilon$	$\int_{-L}^{L} r(L-R_A)f(\epsilon)d\epsilon - \int_{R_A}^{R_A+\frac{P}{1+\rho}} s(\epsilon-R_A)f(\epsilon)d\epsilon$ $- \int_{R_A+\frac{P}{1+\rho}}^{L} \theta\epsilon f(\epsilon)d\epsilon - \int_{R_A}^{R_A+\frac{P}{1+\rho}} \rho(\epsilon-R_A)f(\epsilon)d\epsilon + \int_{-L}^{-(R_B+\frac{P}{1+\rho})} \theta\epsilon f(\epsilon)d\epsilon$
		$\int_{-L}^{L} r(L-R_B)f(\epsilon)d\epsilon - \int_{-L}^{L} s(P-R_B)f(\epsilon)d\epsilon - \int_{-(R_B+\frac{P}{1+\rho})}^{-R_B} s(-\epsilon-R_B)f(\epsilon)d\epsilon$ $- \int_{R_A+\frac{P}{1+\rho}}^{L} \theta\epsilon f(\epsilon)d\epsilon - \int_{-(R_B+\frac{P}{1+\rho})}^{-R_B} \rho(-\epsilon-R_B)f(\epsilon)d\epsilon + \int_{-L}^{-(R_B+\frac{P}{1+\rho})} \theta\epsilon f(\epsilon)d\epsilon$	$\int_{-L}^{L} r(L-R_B)f(\epsilon)d\epsilon - \int_{-L}^{L} s(P-R_B)f(\epsilon)d\epsilon$ $+ \int_{-R_B}^{0} s\epsilon f(\epsilon)d\epsilon - \int_{R_A+\frac{P}{1+\rho}}^{L} \theta\epsilon f(\epsilon)d\epsilon + \int_{-L}^{-R_B} \theta\epsilon f(\epsilon)d\epsilon$
	Period 2	$\int_{-L}^{L} r(L-R_A)f(\epsilon)d\epsilon - \int_{-L}^{L} s(P-R_A)f(\epsilon)d\epsilon$ $- \int_{0}^{R_A} s\epsilon f(\epsilon)d\epsilon - \int_{R_A}^{L} \theta\epsilon f(\epsilon)d\epsilon + \int_{-L}^{-(R_B+\frac{P}{1+\rho})} \theta\epsilon f(\epsilon)d\epsilon$	$\int_{-L}^{L} r(L-R_A)f(\epsilon)d\epsilon - \int_{-L}^{L} s(P-R_A)f(\epsilon)d\epsilon$ $- \int_{0}^{R_A} s\epsilon f(\epsilon)d\epsilon - \int_{R_A}^{L} \theta\epsilon f(\epsilon)d\epsilon + \int_{-L}^{-R_B} \theta\epsilon f(\epsilon)d\epsilon$
		$\int_{-L}^{L} r(L-R_B)f(\epsilon)d\epsilon - \int_{-(R_B+\frac{P}{1+\rho})}^{-R_B} s(-\epsilon-R_B)f(\epsilon)d\epsilon$ $- \int_{R_A+\frac{P}{1+\rho}}^{L} \theta\epsilon f(\epsilon)d\epsilon - \int_{-(R_B+\frac{P}{1+\rho})}^{-R_B} \rho(-\epsilon-R_B)f(\epsilon)d\epsilon + \int_{-L}^{-(R_B+\frac{P}{1+\rho})} \theta\epsilon f(\epsilon)d\epsilon$	$\int_{-L}^{L} r(L-R_B)f(\epsilon)d\epsilon - \int_{-L}^{L} s(P-R_B)f(\epsilon)d\epsilon$ $+ \int_{-R_B}^{0} s\epsilon f(\epsilon)d\epsilon - \int_{R_A}^{L} \theta\epsilon f(\epsilon)d\epsilon + \int_{-L}^{-R_B} \theta\epsilon f(\epsilon)d\epsilon$

Note again that σ_B is a strategic variable for Bank A and σ_A is for Bank B. Because Bank A and B are completely in a symmetric situation, we assume $R_A = R_B$, hence $\sigma_A = \sigma_B$. This enables us to simplify Equation 7-1, and the maximum profit is written as follows.

$$\pi^{NET}(\sigma = \frac{r}{1+r}R + \frac{1}{(1+r)^2}L) = r(L-R) + L[\frac{2\theta - r - 2\theta(1+r)^4}{4(1+r)^4}] + R[\frac{(2\theta+1)r}{2(1+r)^3}] + R^2[\frac{(2\theta r-1)r}{4L(1+r)^2}]. \qquad (7\text{-}2)$$

Here, we derived the optimal value of the bilateral credit limit as well. This leads us to the comparative analysis between the NET and the two RTGS systems.

8. COMPARATIVE STUDY OF GROSS AND NET SETTLEMENT SYSTEMS

We now formally analyse the second question we asked in the introduction: can we give an economic rationale for a coexistence of the RTGS and NET systems? Why is it better for large-size payments to be dealt in the RTGS system while small ones in the NET system? We try to find if our model sustains this proposition, and if so, under what conditions. Intuitively, what we are looking for is a condition under which the bank's profit in the RTGS system is an increasing function of P. This is because the profit in the NET system is not a function of P as P is cancelled out each other; thus if we find the condition where (1) the bank receives higher profit in the NET system when P is small, and (2) the profit in the RTGS system increases with P, we can conclude that the bank prefers to settle small payments in the NET system and large ones in the RTGS system - the coexistence of two systems is justified.

8.1 Net Settlement vs Gross Settlement (EU-type)

Our first analysis deals with the comparison of the NET and EU-type RTGS system. As in Section 7.2, we drop the subscript $i = A, B$, assuming that Bank A and B behave in the same

Figure 7-1 Derivation of Profits for Bank A and B

ϵ	Situation	Revenue	Cost
$[\sigma_A, L]$	Bank A cancels the payment.	$r(L-R_A)$	$-\theta\epsilon$
		$r(L-R_B)$	$-\theta\epsilon$
$[R_A, \sigma_A]$	Bank A raises $\epsilon-R_A$.	$r(L-R_A)$	$-r(\epsilon-R_A)$
		$r(L-R_B)$	0
$[0, R_A]$	Bank A uses its reserve.	$r(L-R_A)$	0
		$r(L-R_B)$	0
$[-R_B, 0]$	Bank B uses its reserve.	$r(L-R_A)$	0
		$r(L-R_B)$	0
$[-\sigma_B, -R_B]$	Bank B raises $-\epsilon-R_B$.	$r(L-R_A)$	0
		$r(L-R_B)$	$-r(-\epsilon-R_B)$
$[-L, -\sigma_B]$	Bank B cancels the payment.	$r(L-R_A)$	$\theta\epsilon$
		$r(L-R_B)$	$\theta\epsilon$

If ϵ is beyond the bilateral limit, it is cancelled and both banks incur the cost of customer dissatisfaction. Otherwise, ϵ is processed, partly with the help of overnight credit. Thus we derive the following profit functions for Bank A and B respectively:

$$\pi^{NET}(\sigma_B) = \int_{-L}^{L} r(L-R_A)\, f(\epsilon)d\epsilon - \int_{\sigma_A}^{L} \theta\epsilon\, f(\epsilon)d\epsilon - \int_{R_A}^{\sigma_A} r(\epsilon-R_A)\, f(\epsilon)d\epsilon + \int_{-L}^{-\sigma_B} \theta\epsilon\, f(\epsilon)d\epsilon \tag{7-1}$$

$$\pi^{NET}(\sigma_A) = \int_{-L}^{L} r(L-R_B)\, f(\epsilon)d\epsilon + \int_{-L}^{-\sigma_B} \theta\epsilon\, f(\epsilon)d\epsilon - \int_{-\sigma_B}^{-R_B} r(-\epsilon-R_B)\, f(\epsilon)d\epsilon - \int_{\sigma_A}^{L} \theta\epsilon\, f(\epsilon)d\epsilon$$

The intuition of this proposition is as follows. What we are comparing is the fee charged for the use of overdraft in case the bank settles at period 1 and the additional cost by the counterpart's cancellation of payment in case the bank settles at period 2. Depending on the relative magnitude of these two cost factors, we observe both delay and non-delay as an equilibrium strategy. If the fee is set low, as specified in Proposition 2a, the bank will prefer to settle without delay. However, if it is set high, there is an incentive for the bank to delay the payment in order to avoid a fee being charged. Nevertheless, this result also depends on how much the bank holds in reserve. The more reserve it keeps, the lower the overdraft needed to settle the stochastic payment, hence lower the fee it will be charged. Ultimately as the cost of the counterpart's payment cancellation increases, the bank will tend to prefer not to delay again. Thus the payment delay is observed if the fee is set high and the banks do not hold enough reserve; hence they have to overdraft a large amount.

6. Comparative Study of Two Gross Settlement Systems

By now we have all the necessary tools for answering the question: which system is more costly for banks, the US-type or the EU-type? We compare the non-delay case in Proposition 2 with the unique equilibrium in Proposition 1.[17] Let $\Delta^{non\text{-}delay}(R_A)$ be the difference in the profit functions of Bank A, i.e. (profit by settling at period 1 in Model 1-A) − (profit by settling at period 1 in Model 1-B). Note in Section 4, we neglected the cost of collateral holding, which has to be introduced in this comparative analysis.

[17] We will not get into the delay case, because it can be avoided by setting the fee low as specified in Proposition 2. At least, it is what the Federal Reserve intends to do in Fedwire (see Section 9).

$$\Delta^{non\text{-}delay}(R_A) = -\int_{-L}^{L} cP\, f(\epsilon)d\epsilon + \int_{-L}^{L} s(P-R_A)\, f(\epsilon)d\epsilon + \int_{R_A}^{R_A+P} s(\epsilon-R_A)\, f(\epsilon)d\epsilon \qquad (6\text{-}1)$$

Then we derive the following proposition:

Proposition 3a:

$$If \quad s\,[(P-R_A) + \int_{R_A}^{R_A+P}(\epsilon-R_A)\, f(\epsilon)d\epsilon] > cP,$$

then the EU-type RTGS system yields higher profit for Bank A than the US-type.

Proposition 3b:

$$If \quad s\,[(P-R_A) + \int_{R_A}^{R_A+P}(\epsilon-R_A)\, f(\epsilon)d\epsilon] < cP,$$

then the US-type RTGS system yields higher profit than the EU-type.[18]

What $\Delta^{non\text{-}delay}(R_A)$ compares is the relative magnitude of the two cost components in the models: the fee charged for the overdraft in Model 1-B and the cost of collateral holding. It should be quite straightforward that if the cost of the fee is higher, the banks prefer the EU-type, and if the cost of collateral posting is higher, they prefer the US-type. Let us solve Equation 6-1. Because ϵ is uniformly distributed,

$$\Delta^{non\text{-}delay}(R_A) = -cP + s(P-R_A) + \frac{s}{4L}P^2$$

[18] Proof is omitted because it is trivial from Equation 6-1.

$$\min[\frac{\sigma_{ij}}{\sum_{i\neq j}\sigma_{ij}} d_i, \; C_i].$$

The survivor will not be asked to cover the loss which exceeds his collateral value; otherwise, this survivor could go insolvent as well, which might create a domino-effect of bankruptcy. In this sense, we cannot fully eliminate the possibility that the payment might not be finalised in the NET system. Nowadays, various measures have been taken in order to achieve as "close" finality as in the RTGS system.[23]

7.2 Case of Low Bilateral Credit Limit

We define the low credit limit as σ lying in the following region:

$$\sigma_i \in [R_i, \; \frac{r}{1+r}R_i + \frac{1}{(1+r)^2}L]$$

The upper boundary is the overnight cash position without collateral posted to the system operator,[24] and the lower boundary is equal to the amount of reserve. Because any payment greater than the limit (hence the overnight cash position) is rejected, the bankruptcy is prevented. Accordingly, we have the following situations:

[23] See Section 9.2 for recent development of risk management in CHIPS.

[24] In the RTGS system, the collateral is deposited with the central bank, and the overnight credit, if necessary, is provided by the central bank. In this centralised system, the preposted collateral, which was originally designed to guarantee the intraday credit, can be used for the overnight purpose. Nonetheless, in the NET system, the collateral is deposited with the operator of the system, while the overnight funds have to be raised in the market. This decentralised structure does not allow the preposted collateral to be used for raising the overnight credit; hence the upper boundary is the sum of R_i, the reserve, and $[L/(1+r)-R_i]/(1+r)$, the amount of collateral available for raising the overnight liquidity.

First mechanism is the adoption of bilateral credit limits and sender net debit caps. In CHIPS, the bilateral credit limit is set so as to prevent the participants from being exposed to an excessive settlement risk. Each participant determines the maximum amount he is willing to receive from other participant. As we will see next, this credit limit determines the amount of loss that each participant is imposed if any of the participants in the system is to go bankrupt. Likewise, the sender debit cap is implemented in order to limit the maximum amount each participant may owe to the entire system. The cap is set equal to 5% of the sum of the bilateral credit limits assigned by other participants. These two devices are linked closely each other, so that, for example, when other banks have some doubts about the soundness of one specific participant, they will reduce the credit limit (it can be zero in an extreme case) in order to prevent themselves from bearing further losses in case of bankruptcy, which automatically reduces the CHIPS' exposure to its credit risk as a whole.

The other option of interest in the NET system is to guarantee that the system can finalise the payments, even in the face of a participant's bailout. This is almost achieved by the loss sharing arrangement. Under this framework, each participant's additional settlement obligation (ASO) is derived as the ratio of his own credit limit extended to the defaulter over the summation of every single participant's bilateral limit to the defaulter. Namely, define σ_{ij} as Bank j's bilateral credit limit set by Bank i, and d_j as the possible loss to be covered by the non-defaulting banks upon Bank j's default. The additional obligation for Bank i is calculated as follows:

$$\frac{\sigma_{ij}}{\sum_{i \neq j} \sigma_{ij}} d_j$$

Accordingly, each participant is obliged to prepost collateral, C_i, which covers the maximum ASO. In reality, it could happen that the ASO exceeds the preposted value of collateral, due to an unexpectedly big loss. To be more precise, Bank i's additional obligation will be

So the critical value of s for Proposition 3 is

$$s' = \frac{4cLP}{4(P-R_A)L + P^2} \qquad (6\text{-}2)$$

Our conclusion is that if the fee is set below s^*, the US-type turns out to be less costly for the participants than the EU-type. The result also depends on how costly it is for the banks to post collateral.

Now we can derive the following policy implication.[19] We know, from Proposition 2b-1, the settlement delay occurs when (1) a relatively high fee is charged, and (2) the participants hold less reserve initially. The less reserve the participants keep, the more likely they have to rely on the overdraft, and the more costly it is for them when the fee is set high, which creates a stronger incentive to free-ride on the incoming payment from the counterpart; hence the delay is observed. In order to prevent this, it seems the Federal Reserve has to either (1) reduce the fee (transition to the non-delay case in Proposition 2a) or (2) encourage the participants to hold more reserve (transition to the non-delay case in Proposition 2b-2). Nonetheless, from Proposition 3a, charging high fees gives comparative disadvantage over the EU-type - it may produce an outflow of payments from the US-type to the EU-type, if the participants have the choice between the two. In order to avoid this, the Federal Reserve has to adopt the first choice and set the fee low; the second choice, encouraging the participants to hold high levels of reserve, may be excluded.

[19] The detailed analysis including the fee policy of the Federal Reserve is dealt with in Section 9.

7. THE MODEL OF NET SETTLEMENT SYSTEM

Our next analysis deals with the net settlement (hereafter NET) system, a polar type of settlement to the RTGS system. We first summarise the main differences with respect to the RTGS system.[20]

7.1 Point of Distinction

Operator of the System: The NET system is owned by a private institution, unlike the RTGS system where the central bank normally owns the system. In the case of CHIPS, it is governed by the New York Clearing House.

Type of Settlement: The crucial difference with respect to the RTGS system is that under the NET system, payments will not be finalised at the same time as the payment orders are transmitted from the payer to the payee. In other words, commitments to transfer money keep accumulating for certain intervals, usually until the end of day, and each participant transfers only the net-out position vis-a-vis all other participants in the case of the multilateral netting arrangement.[21] The NET system reduces the need for cash balances and moreover an incremental reduction of the number of transactions is achieved. This efficiency derived from the need for less liquidity is offset by a risk factor, namely that the participants are exposed to a possible loss on an outstanding promise to pay.

Credit Facility: Unlike in the RTGS system, where the central bank provides credit (mainly

[20] The following institutional features are based on CHIPS of the US. We will take this system as a prototype.

[21] This becomes the bilateral netting arrangement if the position is netted out between two participants.

intraday), the NET system's participants are not guaranteed an access to this credit facility. As long as the whole settlement takes place at the end of the day (as in CHIPS), there is no need for banks to actually raise intraday funds. The credit is *de facto* extended among participants. Eventually, when the designated time arrives, those participants with net debit position have to raise funds, and the interbank market plays the important role in providing overnight credits in this situation. With no central bank credit, the reserve and the funds raised in the market are the primary source of settling payments. We assume each bank can make some amount of TB as collateral for overnight credit. So with the interest rate equal to r, which is larger than ρ, the official discount rate of the central bank, the overnight cash position becomes

$$R_t + \frac{L-R_t}{1+r},$$

sum of the reserve and TB discounted by $(1 + r)$. Any net debit position greater than this position cannot be dealt with.[22]

Risk Management: The implicit intraday borrowing and lending caused by the settlement lag raise a further concern - how should one cope with the risk of a settlement failure? Broadly speaking, there are four mechanisms which have been implemented to achieve the "finality" of payments in the NET system: a contract design such as adding a closeout provision, in which an outstanding transaction is immediately reformulated into a payable amount; the membership criterion which restricts the participants, based on certain qualification; an introduction of bilateral credit limits and sender net debit caps; and the loss sharing arrangement. The first two mechanisms are self-explanatory; for our purposes, let us focus on the last two.

[22] As we will see in the model, some of TB has to be preposted to the NET system operator, which reduces the total amount of TB available for the overnight fund-raising purpose in this decentralised system.