LIQUIDITY RISK IN SECURITIES SETTLEMENT SYSTEMS

Johan Devriese

National Bank of Belgium*

First draft: January 2005

Very preliminary, please do not quote

Abstract

This paper studies the potential impact on securities settlement systems (SSSs) of a major market disruption, caused by the default of the largest player. A multi-period, multi-security model with intraday credit is used to simulate direct and second-round settlement failures triggered by the default, as well as the dynamics of settlement failures, arising from a lag in settlement relative to the date of trades. The effects of the defaulter's net trade position, the numbers of securities and participants in the market, and participants' trading behavior (risky versus conservative) are also analyzed.

We show that in SSSs – contrary to payment systems – large and enduring settlement failures are possible even when ample liquidity is provided. Central bank liquidity support to SSSs thus cannot eliminate settlement failures due to major market disruptions. This is due to the fact that securities transactions involve a cash leg and a securities leg, while liquidity can only affect one side of a transaction. Whereas a broad program of securities borrowing and lending might help, it is precisely during periods of market disruption that participants will be least willing to lend securities.

Interestingly, settlement failures continue to occur beyond the period corresponding to the lag in settlement. This is due to the fact that, upon observation of a default, market participants must form expectations about settlement of their previous trades not yet settled, and these expectations affect current trading behavior. If, ex post, fewer of the previous trades settle than expected, new settlement failures will occur. This result has interesting implications for financial stability. On the one hand, conservative reactions by market participants to a default – for example by severely limiting their trades – will result in a more rapid return of the settlement system to a normal level of efficiency. On the other hand, severe limitation of trading by market participants can sharply reduce market liquidity, which may have a significant, negative impact on financial stability.

Keywords: Security clearing and settlement, liquidity risk, systemic risk,
JEL-classification numbers: C6, D4, G20, O33

* E-mail: johan.devriese@nbb.be. Many thanks to Janet Mitchell for the feedback, comments and support and to Konstantijn Maes and Hans Degryse for their valuable comments. The views and findings of this paper are entirely those of the author and do not necessarily represent the views of the National Bank of Belgium.

1 INTRODUCTION

A prerequisite for the development of a viable capital market is a well functioning transactions infrastructure. The settlement of securities transactions is an important component of this infrastructure, as it determines the legal transfer of the securities that are traded. This infrastructure must operate in a seamless and integrated manner, in order to minimize the costs and risks for the end users in the market and to facilitate the allocation of capital. Hence, securities settlement systems (SSSs) are crucial to the financial system and are often supported by the central bank as lender of last resort.

Disruptions in the settlement infrastructure can lead to increased transaction costs and to a possible erosion of market liquidity which, if serious enough, may undermine financial stability. An extreme example of the potential severity of settlement failures was provided by the September 11 attacks. Settlement failures in the U.S. Treasury market jumped from $1.7 billion per day in the week ending September 5 to $190 billion per day in the week ending September 19 (see Flemming and Garbade, 2002). Failures rose initially because of the destruction of communication facilities, but remained high because the lending and borrowing program was ill-suited to absorb the massive shock.

This paper explores the potential consequences of a market disruption that is less severe than the Sept. 11 attack, but that is nevertheless serious; namely, the default of the largest participant in the market. Although the impact of the default on other market participants will be gauged in terms of the resulting liquidity problems, in order for the default to have serious consequences, it must be assumed to derive from a fundamental problem such as insolvency of the defaulting institution. A simulation is used to study questions such as the direct and second-round settlement failures triggered by the default, the dynamics of settlement failures, and effects of the defaulter's net trade position.

Despite the recognized importance of securities settlement for the functioning of financial markets, risk (other than operational risk) in SSSs is often perceived to be relatively low. This may be explained by the fact that the existence of a cash leg and a securities leg has led to widespread adoption of delivery-versus-payment (DVP) systems, in which the security and cash are transferred at the same time. By ensuring that securities are transferred from the seller to the buyer if and only if funds are transferred from the buyer to the seller, DVP systems eliminate principal risk.[1] Hence, solvency risk of institutions arising from participation in securities markets is virtually eliminated. DVP systems, however, do not eliminate liquidity risk arising from unsettled trades.

Indeed, the presence of a cash and a securities leg in securities transactions suggests that the impact on the SSS of a default by a major player may depend, in addition to the player's size, upon

---

[1] Principal risk is the risk that the seller of a security delivers the security but does not receive payment, or that the buyer of a security makes the payment but does not receive the security. Principal risk in securities

the player's trade position: net buyer versus net seller. Since cash is used in all transactions, default by a large player with a net buy position will result in a withdrawal of cash from the system, which can impact trades of all securities. This implies that liquidity provision by the SSS or the central bank may help to eliminate settlement failures; however, it may not be sufficient to avoid significant disruptions when, for example, the defaulting participant is a net seller. In this case, some securities (which cannot be provided by the SSS if there is no securities borrowing and lending program) are no longer available for the system.

Another important distinction between SSSs and payment systems is one of deferred versus real-time settlement. Transactions can either settle in real time (t+0), immediately after trade, or at some later time (e.g., t+2 or t+3). Contrary to most payment systems, securities transactions typically settle at a date later than the date of trade.[2] Although this lag gives participants extra time to find the necessary funds to finance the trades, it also increases replacement cost risk.[3] In the delay between trade and settlement, asset prices may have changed, making it possibly more expensive to trade the securities elsewhere if the initial trade does not settle. Moreover, a disruption in the settlement process on a given day *t* (settling trades from day *t-3*, for example) will also have consequences for the trades from days *t-2* and *t-1*, which are already committed but not yet settled. Due to the initial settlement failures (of trades on day *t-3*), counterparties of the defaulting institution may end up with less cash or securities in their accounts than was anticipated when they committed to the trades on days *t-2* and *t-1*. It may, therefore, become difficult for these participants to meet the settlement obligations for those days; thus further settlement failures may be triggered. Hence, the possibility of contagion in securities settlement systems is not limited to a single day, but may last for a longer period.

The differences in payment and settlement systems give rise to a number of important questions for SSSs which have not been addressed by existing literature. What are the dynamic effects on settlement – both direct and contagion effects – of a major disruption in the market? Is the first-day impact greater or smaller than the impact in subsequent days? How many days does it take for settlement efficiency (the percentage of trades settled) to return to its normal level? How does the existence of a cash leg and a securities leg influence the degree of contagion, relative to a payment system where only cash is involved? Can central bank support of the SSS through credit provision prevent contagion? If so, how much credit would be needed? How does the trade position of the defaulter (e.g., size, net buyer versus net seller) affect the magnitude of the disruption?

---

markets is analogous to credit risk in the interbank market. Liquidity risk is the risk that a counterparty will not settle an obligation for full value when due, but on some unspecified date thereafter (BIS, 1992)

[2] Settlement usually takes place only two or three days following trade. Under t+0 settlement, operational risk can be high, as there is little time to solve operational problems. In addition, t+0 settlement is often technically infeasible, due to the time required to match the cash and securities legs in transactions involving multiple currencies, different time zones, and cross-border settlement.

[3] Replacement cost risk is the risk that a counterparty may default prior to settlement, denying the non defaulting party the gain on the transaction (BIS, 1992).

This paper uses a simulation to address these questions. Settlement is assumed to occur in a DVP system with gross (trade-by-trade) settlement and a two-day lag (i.e., settlement of trades on day *t* occurs at the end of day *t+2*). The SSS may provide liquidity in the form of credit and results are compared across scenarios with differing assumptions regarding the amount of liquidity provided. Default by the largest player triggers the initial settlement failures. The direct and second-round effects of the default are measured over a period of five days following the default. The impact of the defaulter's net trade position, as well as the numbers of securities and participants in the market, and participants' trading behavior (risky versus conservative) are analyzed.

Several results emerge from the analysis. First, the two-day lag in settlement implies that settlement failures will last for at least two days following default. Settlement efficiency (measured as the percentage of trades that actually settle) is in fact lower on the day following default than on the day that default occurs, due to continuing contagion.[4] Thus, the crisis situation initially worsens before improving. Interestingly, settlement efficiency may not return to normal after two days (and indeed does not return to normal in the simulations reported here), despite the two-day settlement lag. The reason is that upon observation of a default on day *t*, market participants must form expectations about the impact that the default will have on settlement of the trades they committed to on the two days preceding the default. These expectations concern the amounts of securities and cash that the participants will ultimately have in their accounts. The expectations thus affect the budget constraints that the participants use to determine their trades on days *t+1* and *t+2*. If, ex post, fewer trades actually settle than anticipated (i.e., if, ex post, traders' expectations about their actual budget constraints are not conservative enough), then participants may commit to trades in the two days following the default that later turn out to be infeasible. Thus, settlement inefficiency can continue for a longer period than the length of the settlement lag.

This result has interesting implications for financial stability. On the one hand, conservative reactions by market participants to a default – for example by severely limiting their trades – will result in a more rapid return of the SSS to a normal level of efficiency, and an end to the crisis. On the other hand, severe limitation of trading by market participants will sharply reduce market liquidity, which may have a significant, negative impact on financial stability. In addition, severe limitation of trading can have negative welfare effects on participants, due to lost benefits from trading.

A second result is that the net trade position of the defaulting institution can have a significant impact on the severity of the crisis. When the SSS provides little or no liquidity, a large net buy position of the defaulter will cause a significantly higher fall in settlement efficiency than will a large net sell position. As one might expect, generous liquidity provision by the SSS can eliminate the differential effects of the defaulter's trade position on settlement efficiency. Importantly, however, liquidity provision cannot completely eliminate the crisis: settlement efficiency still falls significantly following the default of the largest player even when plenty of liquidity is available.[5] As noted above, this is due to the fact that securities transactions involve a cash and a securities leg. Liquidity provision by a central bank or a central security depository (CSD) can eliminate problems on the cash side of transactions but not on the securities side. Thus, default by a major player can still have an impact on the system. In order for a crisis to be completely avoided, a broad, well functioning securities borrowing and lending program would have to be in place, in addition to adequate liquidity provision. Yet, a securities borrowing and lending program will only work if enough participants are willing to lend a wide enough range of securities, and it is precisely at the time of a crisis that uncertainty about repayment is greatest and holders of securities will be the least willing to lend.

Additional results concern the implications of differing numbers of participants and securities, and differences in trading behavior. Not surprisingly, the severity of the crisis (in terms of settlement inefficiency) decreases as the number of participants increases. This is due in part to the fact that the size of the largest participant relative to the market (or the degree of concentration) is directly linked to the number of participants. A larger number of participants translates into a smaller direct impact of default by the largest player. Less intuitive, however, is the result that the aggregate amount of credit that the SSS will have to provide (for a given credit-granting rule) will increase with the number of participants. This is because a higher number of participants leads to a longer chain of trades, and therefore to a greater number of trades for which credit may need to be extended in order to avoid settlement failures. So, whereas a larger number of participants increases settlement efficiency, it can also place a greater burden on the liquidity provider when market disruptions occur. Finally, the severity of the crisis increases with the number of securities traded. The crisis is also more severe when the participants typically trade closer to the boundaries of their trading budget constraints (more "extreme" trades) than not.

Section 2 gives an overview of the existing literature. Section 3 presents the model. Section 4 discusses the simulation results. Section 5 concludes.

2 LITERATURE REVIEW

Contagion has become a topic much investigated in the finance literature during the last decade. Starting with bank runs as a channel of financial contagion when agents do not have complete information (Diamond and Dybvig (1983), authors have shown that even under perfect information financial contagion is possible (see, e.g., Rochet and Tirole (1996), Allen and Gale (2000) and Diamond and Rajan (2003)). These papers concentrate on contagion in the interbank market, where it is assumed that banks have uncollateralized exposures to each other, and the default of

---

[4] Default during day *t*, prior to settlement on that day, implies that the trades from day *t-2* are the first to be affected, as these trades are settled at the end of day *t*. Settlement on the day following default concerns trades that occurred on day *t-1*.

[5] Indeed, the impact of additional liquidity provision above some threshold level appears to be limited.

one bank can cause other banks to become insolvent and default as well. Hence, credit risk and solvency risk are at the fore.

This is also the idea behind several empirical studies investigating financial contagion. Humphrey (1986), Angilini, Maresca and Russo (1996) and Norhtcott (2002)) all use payments data from a single day in payments systems in which net settlement occurs. Humphrey uses data from a randomly selected business day in CHIPS (US) and simulates the impact of a major participant's failure by unwinding all of the day's transactions to and from that participant, calculating the balances of the remaining participants, comparing this with their capital buffers and iterating the unwind. Humphrey finds that on average, 37% of the institutions fail after the initial participant's failure. Angilini, Maresca and Russo (1996) and Norhtcott (2002) use a similar method for the Italian and Canadian netting systems, respectively, and conclude that systemic risk in those systems is very low or nonexistent.

Other empirical studies use data on large interbank exposures as reported in banks' balance sheets (see Upper and Worms (2002), Furfine (2003) and Degryse and Nguyen (2004)). On average, these papers also find low degrees of potential contagion. In reaction to these findings, however, Cifuentes, Ferruci and Shin (2004) argue that in reality systemic risk is greater than that identified by interbank contagion simulations because the risk that actually materialises is not credit risk but market risk, due to a fall in asset prices when banks liquidate collateral in response to defaults by their interbank borrowers in order to meet their own obligations. In order to illustrate the potential importance of market risk, Cifuentes et al use a simulation where banks' asset sales in response to interbank defaults cause a fall in market prices of securities.

The small degree of interbank contagion found empirically, together with the virtual absence of principal risk (or credit risk) in SSSs may explain why there are very few studies on systemic risk in security settlement systems. Indeed, De Bandt and Hartmann (2000) observe in their extensive literature review on systemic risk that: "Empirical studies of systemic risk in securities settlement systems appear to be non-existent". However, much like the argument made by Cifuentes et al (2004) with respect to interbank markets, one can argue that credit risk (or principal risk) is not the only possible risk in SSSs. Liquidity risk is also important and can even, lead to market risk when participants must liquidate collateral in response to market disruptions such as settlement failures.

The nature of systemic risk in SSSs, however, seems to depend upon the type of settlement system: net or gross settlement. In net settlement systems, transactions are settled on a net basis, which economizes on the amount of liquidity needed by participants. However, default by a participant in a net settlement system causes trade unwinds, whereby some or all of the transfers involving that participant are deleted and the settlement obligations of the other participants are recalculated, which may lead to possible further unwinds. This increases replacement cost risk, as settlement is only final at the end of the entire settlement process.

Gross settlement systems, on the other hand, transfer instructions for both securities and funds on a trade-by-trade basis during the settlement process. Failure of a participant to meet a delivery or payment obligation on a given transaction will not lead to costly unwinds of multiple transactions. Yet, DVP systems with gross settlement require substantial intraday liquidity. If participants are unable to adjust their cash balances during the processing cycle, they will have to hold enough cash to cover at least the largest debit position during processing. Hence, liquidity risk becomes more important than with net settlement systems. If sufficient money balances are not available, high “fail” rates may result, implying substantial liquidity risk and replacement cost risk to counterparties.

Most of the literature on payments settlement has focused on the differential effects of gross versus net settlement. Angilini (1998) uses a real-time gross settlement model (RTGS) to show that if daylight liquidity is costly, banks may be induced to postpone payment, hence increasing liquidity risk in the system. Kahn and Roberds (1998) note that although net settlement is less costly due to the lower need for liquidity, net settlement increases moral hazard, as banks have an option to revoke their trades, which distorts incentives. Kahn, McAndrews and Roberds (2003) analyze more fully the prospect of strategic default in settlement systems and again conclude that net settlement causes less payment gridlock. Leinonen and Soramaki (1999) use Finnish data in an attempt to quantify the relationship between liquidity usage and settlement delay in net settlement systems and RTGS systems with queuing. When the central bank provides low-cost intraday credit, liquidity costs are low relative to delay costs and RTGS systems with queuing are more efficient.

The main conclusion from this line of research appears to be that there is no liquidity risk in payment systems using gross settlement as long as there is sufficient and cheap intraday credit. However, for the reasons noted in the Introduction (e.g., securities and cash leg, settlement lags), this argument will not necessarily hold for SSSs. The only paper to our knowledge that investigates liquidity risk in SSSs is that of Iori (2004), which analyzes the importance of operational risk with respect to differing lag times between trade and settlement in both gross and net settlement systems. In this model, only one security is traded, and no cash or budget constraints exist. Trades occur at periodic intervals, and operational delays result in settlement failures whenever the operational delay is longer than the lag between trade and settlement. While shortening the lag between trade and settlement has the advantage of reducing replacement cost following the failure of a participant to settle, it also increases the likelihood of settlement failures caused by an operational problem. Thus, even under real-time settlement (t+0), significant settlement contagion is still possible.

Much of the empirical literature on contagion in payments systems and interbank markets makes use of simulations with strong underlying assumptions, which are necessary, for example, due to the inability to obtain data on participants' bilateral positions. This will be all the more true for simulations of SSSs, which generally will not be able to make use of any real data. Not only are data relating to individual trades in SSSs highly confidential, but also would the amount of data needed for an empirical study be massive, due to the need to have data on participants' cash and securities holdings as well as their trades. Only the SSSs themselves are able to use real data in simulations or stress tests. Unfortunately, such exercises are for internal use only and often suffer from a number of shortcomings when viewed from a financial stability perspective. First, SSSs are mainly concerned about their own exposure in case the largest participant fails. Second, stress tests often take into account only the direct effects of a participant's failure, which underestimates systemic risk. Moreover, when second-round effects are incorporated, they only cover a single day of trade data, while disruptions in the settlement system may last for several days. Finally, trading behavior in times of stress likely differs signficiantly from behavior on "normal" days, raising the question as to whether the use of trading data from a "normal" day is valid for simulating a stress event. The only apparent way around this problem is to conduct empirical studies based upon real stress events, as in Flemming and Garbade (2002). (Un)fortunately, these events are rare.

3 MODEL

3.1 Description

We model a SSS with DVP and gross settlement, where settlement occurs with a two-day lag. All securities prices are assumed to be fixed and normalized to one.[6] Participants are randomly allocated initial quantities of cash and securities. The distribution of initial endowments can be varied to result in greater or lesser size concentration of the largest participant.

Trades are assumed to occur randomly, and trades are computed between all possible combinations of counterparties and securities. Once a given security and participant pair are randomly selected, feasible trades are determined by the two participants' "expected" budget constraints (i.e., their expected holdings of the security and cash). A trade is then randomly chosen from the set of feasible trades, and the "expected" budget constraints of the participants involved in the transaction are updated to reflect the trade. Selection of the trade is determined via a Beta distribution, which has the advantage that different parameter values lead to more or less "extreme" trades (i.e., how close the trade is to the boundaries of participants' budget constraints). Simulations with conservative and "extreme" trading behavior can thus be compared.

Because settlement occurs with a lag, the budget constraints that are used for determining feasible trades are actually participants' expected holdings of securities and cash. For example, the expected holdings at the beginning of day *t* will be the amounts of securities and cash that participants believe will actually be in their accounts following settlement of the trades from days *t-2* and *t-1*. (The two-day settlement lag means that trades from day *t-1* will only be settled at the end of day *t+1*.) If participants want to trade today, they need to have an idea about the amount of cash and securities they have to back these trades. However, since this position will only become known with certainty in two days, they must form expectations. It is assumed that as long as no defaults have occurred (i.e., in "normal" times), participants expect that all their trades will settle (which will actually be the case). Thus, their expected budgets will be perfectly consistent with the

[6] This implies that in the version of the model presented here, market risk does not arise.

actual budgets (i.e., the actual deposits of securities and cash in their accounts) which will be used for settling the trades.

When a default has occurred, (i.e., in "crisis" times), participants' expected holdings may differ from the holdings that they will actually have in their accounts after settlement. This is due to the fact that although participants are assumed to know their counterparties and can, therefore, accurately predict the direct effect of a default by one of their own counterparties, participants do not know the counterparties of their counterparties and, hence, cannot predict which of their nondefaulting counterparties have traded with the defaulting counterparty and may thus be unable to fulfill their settlement obligations.

Second-round, or indirect, effects of defaults are thus unknown, and participants must form some expectations about them. If, ex post, the expectations turn out to be incorrect, then additional settlement failures can occur. The simulations presented below use an ad hoc rule for determining participants' expectations regarding second-round effects: participants' expected holdings of securities and cash are assumed to be 80% of what they would be if all trades settled. These expectations, while fairly conservative, still turn out to generate additional settlement failures. Although expectations formation in the model is quite mechanical, it is not the mechanical nature that leads to the ex post errors. Any more sophisticated model of expectations formation could also lead to ex post errors, as long as participants' information about the trades of their counterparties with other counterparties, and information about all counterparties' budget constraints, is not perfect.

Settlement of trades is assumed to occur in the same order as the order in which the trades were undertaken. This maximizes settlement efficiency (the percentage of trades that actually settle). A further aid to settlement is the assumption of a queue of unsettled trades, which also reflects practice in SSS's. That is, the settlement process – which occurs at the end of a given day and applies to the trades undertaken two days earlier – is assumed to consist of three batches, or iterations. Trades that are not settled in the first iteration are tried again in the second ,and the process continues through the third iteration. Trades that are still unsettled at the end of the three batches are then placed in the queue for settlement at the end of the following day. Because trades that are not settled in the first batch often settle in the second or third batch, allowing for a queue of unsettled trades reduces the negative impact on settlement of default by a participant.

Another feature of the model that can reduce settlement failures is the provision of intraday credit, which may be drawn upon during the settlement process. Note that at the point when trades for day t are settled (i.e., at the end of day $t+2$), the holdings of securities and cash that participants have in their accounts reflects the settlement of all trades that were undertaken up to day t (i.e., through day $t-1$). Therefore, whereas trades are undertaken on the basis of expected holdings of cash and securities, the settlement process uses actual (legal) holdings. When intraday credit is available, a participant who is short in cash for settlement of a trade can draw on the credit during the settlement process and avoid settlement failure. Simulations with differing amounts of credit availability are compared in Section 4.

The initial shock in settlement is assumed to stem from an exogenous default of the largest participant, as measured by the amount of initial securities holdings plus cash. This does not imply, however, that solvency risk is playing a role in the model. The simulation takes into account liquidity risk only, gauged in terms of the trades that fail to settle because of insufficient cash or securities holdings by the transaction participants. Unlike the interbank contagion literature, participants' losses due to failed trades are not compared with a solvency constraint.

3.1 Notation and sequencing of events

Three "events" occur during each day t in the following order: (1) participants' determination of their expected holdings of securities and cash; (2) trading; and (3) settlement of trades undertaken on day t-2 (which accomplishes the legal transfer of cash and securities).

N number of participants

K number of securities, with k=1,…,K. Each security is infinitely divisible into tradable shares. Hence, if a participant holds 0.01 of security k, it holds 1% of the total outstanding amount of this security.

t time index, representing one day.

T(t) Three-dimensional trade flow matrix (N,N,K) of trades occurring during day t, where the elements T_{ijk} are the number of securities of type k that participant i sells to participant j at day t. $T_{ijk}= -T_{jik}$ and i, j=1,…,N. $T_{ijk} = 0$ when no trade occurs.

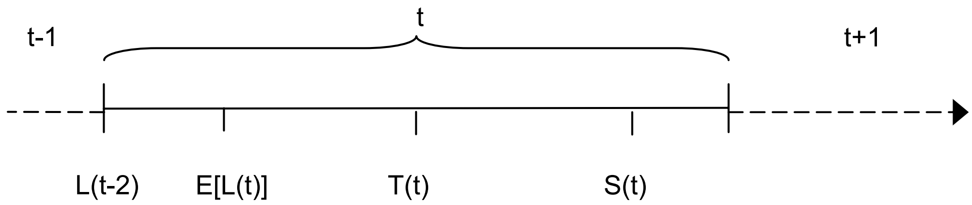
S(t) Three-dimensional settlement flow matrix (N,N,K), whose elements are defined analogously to those of T(t-2) **containing all trades from day t-2 to be settled at time t** (plus any unsettled trades from day t-3 that are still in the queue). Unsettled trades enter as zeros in the matrix.

L(t-2) (K+1,N) matrix containing the legal (or post-settlement) holdings of the K securities and cash by each participant held at the beginning of day t-2 (after settlement of trades from day t-3), but only known at the beginning of day t. L(t-2) will be used for settlement of trades from day t-2 and becomes only known at the beginning of day t (or equivalently, after the settlement process at the end of day t-1). The matrix reflects past settlement; i.e. L(t-2)=S(t-1)+S(t-2)+S(t-3)+…

E[L(t)│L(t-2), T(t-2), T(t-1)]

Expectations of L(t) formed by each participant prior to trading on day t and based upon information on all previous trades and on settlement for that participant. Because L(t-2) will already have been determined by the beginning of day t, E[L(t)] will ultimately differ from L(t) only if some of the participant's trades on days t-2 and t-1 do not settle (which will become known at the end of days t and t+1, respectively).

The figure below depicts the timing of "events" during a single day t.



At the beginning of the day participants **form expectations** about their holdings of cash and securities; i.e. E[L(t)] is formed. Expectations are based upon information available at the beginning of time t, which includes the history of trades and settlement, as reflected in L(t-2), T(t-2) and T(t-1). Because settlement of trades undertaken in the two previous days has not yet occurred, participants need to form expectations about the stocks of securities and cash that they have available for trading during day t. For instance, trades occurring on day t-1 have not been settled at time t; therefore, participants do not know if all of their t-1 trades will actually settle. In the simulations reported below, the assumption is made that in normal times (when no default has occurred), participants expect that all trades from T(t-2) and T(t-1) will settle.

Once E[L(t)] is determined, participants then **make trades** on the basis of these expected holdings of securities and cash. The trades then make up the elements of the matrix T(t). Trades occur on a one by one basis, so that after each trade T_{ijk}, the expected budget constraint as reflected by E[L(t)] is updated before a new trade is determined.

At the end of the day, trades that were committed on day t-2 (plus any unsettled trades from days prior to t-2 that are still in the queue) are presented for **settlement.** By comparing the individual trades from day t-2 with the true holdings reflected in L(t-2), it is determined if a particular trade is settled or not. Hence S(t) becomes known.

In normal times (i.e. no settlement failures), the legal holdings of securities and cash for each participant will simply equal the initial holdings modified by the accumulated trades. Hence, each participant's expected stocks of securities and cash will equal the true stocks; i.e., E[L(t-2)]=L(t-2), and participants' trades will have been based upon an accurate perception of their true budget constraints. However, if during the settlement process on day t some trades from day t-2 do not settle – due to a failure of a participant, for example – then S(t) will differ from T(t-2) and the real holdings of securities and cash L(t-1) will differ from what the expected holdings E[L(t-1)] were at the time of trading on day t-1. Therefore, the budget constraints of some participants that will be used for settling the t-1 trades will not equal the expected budget constraints that were used to determine these trades. Certain trades that were feasible given the expected budget constraint may not be feasible after all. Thus, further settlement failures may follow the initial failure.

3.2 Determination of initial securities and cash holdings, L(1)

In the simulation the starting position of the assets and cash held, L(1), is determined under the following restrictions:

- All securities are randomly distributed among the N participants following the positive plane of a multivariate normal distribution, with mean 0 and a KxK variance matrix A[7]:

  $$A = \sigma^2 \begin{bmatrix} 1 & \rho & \cdots & \rho \\ \rho & \ddots & & \vdots \\ \vdots & & \ddots & \rho \\ \rho & \cdots & \rho & 1 \end{bmatrix}$$ and σ^2=variance, ρ=correlation

  By varying the variance and covariance of this distribution, different types of portfolio structures can be simulated. E.g. a higher variance increases the concentration of certain securities in the hands of a few participants. A high ρ increases the likelihood that if a participant has a large position in one type of security, it has a large position in the other securities as well. So a higher correlation results in larger differences in the total size of participants portfolios, i.e. larger concentration of asset size.
- For each security k, the total amount of securities outstanding is normalized to 1. Hence, if a participant holds 0.01 of security k, it holds 1% of the total outstanding amount of this security.
- Cash positions are assumed to equal a percentage C of securities held. This reflects the idea that cash bears no return and is only held for trading purposes.

Participants can go short in cash, by using credit granted by the CSD. The credit limit is set as a percentage λ ($0<\lambda<1$) of the initial total allocation of assets. The credit in the model can be thought of as representing either the collateralised credit provided by a CSD or liquidity that may otherwise be available through the interbank market (which is not formally modeled here). However, since credit is costly, we assume that participants try not to use it in normal days. That is, participants do not include credit in the budget constraints that determine their trades. Participants are assumed to make use of their credit line within the SSS only as a backup facility, that is, when there are some unanticipated settlement failures[8]. Note that for the initial period (day 1), E[L(1)] = L(1) by definition.

3.3 Determining trades, T(t)

Trades are determined by randomly choosing two counterparties and a security and generating a trade lying within the expected budget constraints of the two participants with respect to that security. The budget constraints are derived from the expected stocks of the particular security and cash, (i.e. from the relevant cells of E[L(t)]), and will determine the maximum amount of the security

---

[7] Drawing from the positive plane is done by randomly drawing from a multivariate normal distribution with mean 0 and then taking the absolute values. To avoid extreme portfolios, values bigger than one are set equal to one. The result is a K by N matrix. To make sure that the total quantity of each security k equals one, each entry in the kth row of the matrix is divided by the sum of all entries in the kth row.

[8] Allowing participants to draw on their credit lines for trading would not change any of the qualitative results in the model. It would simply widen budget constraints used to determine feasible trades.

that each player could sell or buy. For example, trade T_{ijk} will represent a trade between players i and j, from player i's point of view. Negative numbers represent purchases of the security by player i from j and positive numbers represent sales of the security by player i to player j. Note that T_{ijk}=-T_{jik}. All feasible trades can be represented by the interval [-A,B], where

$$A = \min\left(E\left\lfloor L(t)_{j,k} \right\rfloor, E\left\lfloor L(t)_{i,K+1} \right\rfloor\right) \qquad \text{eq. 1}$$
$$B = \min\left(E\left[L(t)_{i,k}\right], E\left[L(t)_{j,K+1}\right]\right)$$

and

$E\left\lfloor L(t)_{j,k} \right\rfloor$ is the expected amount of security k held by participant j (idem for i).

$E\left\lfloor L(t)_{j,K+1} \right\rfloor$ is the expected amount of cash held participant j (idem for i).

A is the expected maximum amount of security k that player i can buy from player j

B is the expected maximum amount of security k that player i can sell to player j.

The intuition behind equation 1 is that the maximum amount of a security that a participant can sell is determined by both the amount of security held by that participant and the amount of cash held by its counterparty. Conversely, the maximum amount of a security a participant can buy is determined by the amount of cash held by the participant and the amount of the security held by its counterparty.

The entry T_{ijk} is randomly chosen from the interval [-A,B] following a symmetrical beta distribution with parameter β_i (β_i>0 and i=1,2). The Beta distribution makes it possible to vary participants' simulated trade behavior by varying the value of β. Thus, it is possible to asses the impact of settlement failures of relatively extreme vs. moderate traders. For the simulation results presented in Section 4, different values of β have been used for trades involving the largest (hence, the defaulting) participant versus trades not involving this participant. A value β_2 is used for trades involving the largest participant and β_1 is used for all other trades.[9]

After T_{ijk} is determined, the expected cash and security positions of the participants involved in that trade are updated. Then another security and participant combination is selected. The trading process ends when all possible combinations of counterparties and securities have been sampled.

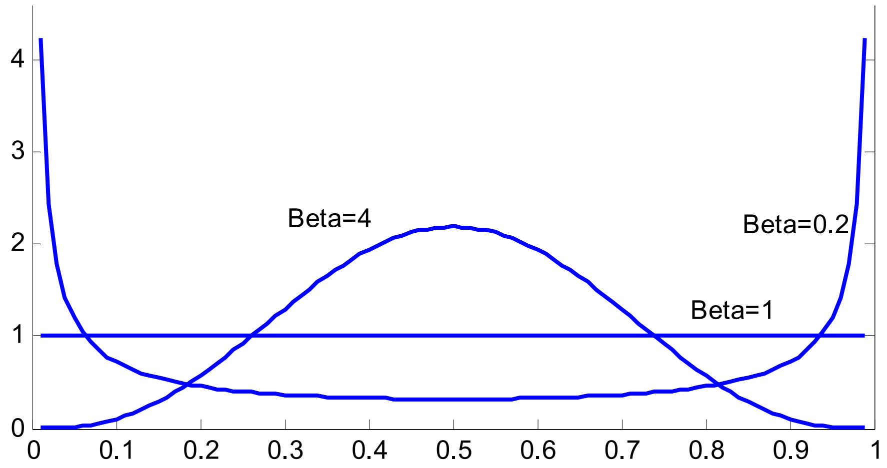
[9] Varying the values of β in this way can allow for risky or conservative behavior of the largest participant relative to others in the market. More analysis of the impact of this parameter are being undertaken in ongoing work with the model.

The beta distribution

The beta distribution describes a family of curves that are unique in that they are nonzero only on the interval [0,1]. The shape of the beta distribution is quite variable depending on the values of the parameter β, as illustrated by the plot below. The constant pdf (the flat line) shows that the standard uniform distribution is a special case of the beta distribution. 0< β <1 represents participants with extreme trade behaviour as they frequently use the limits of the budget constraint. β >1 represents participants with less extreme trade behaviour.

In the simulation, β <1, resembling broker/dealers which only hold securities and cash on their account for trading purpose. In order to minimise the costs for their clients, they try to use all their margin.

Figure 1: PDF of the beta distribution



3.4 Determine S(t)

S(t) is the matrix containing all settled trades from day t-2 –as reflected in T(t-2) – plus unsettled trades from the queue. Trades do not settle if settlement would imply any one of the following:

- A counterparty's cash position becomes overdrawn and exceeds the credit limit granted by the CSD.
- A counterparty becomes short in securities.
- One of the counterparties defaulted prior to settlement

Trades are settled one-by-one (gross settlement) in the same order as they were traded, thereby allowing participants to manage their cash positions.[10] Each trade is compared with the holdings of cash and securities represented by the matrix L(t-2), plus available credit. For example, trade T_{ijk}(t-2) will be settled on day t if it lies in the interval [-$\overline{A}$,$\overline{B}$], where

[10] This ordering of settlement reflects actual practice and also maximizes the proportion of trades that will be settled.

$$\overline{A} = \min\left(L(t-2)_{j,k}, L(t-2)_{i,K+1} + credit_i\right)$$
$$\overline{B} = \min\left(L(t-2)_{i,k}, L(t-2)_{j,K+1} + credit_j\right) \quad \text{eq. 2}$$

and

$L(t-2)_{j,k}$ is the quantity of security k in participant j's account at time t (idem for i)

$L(t-2)_{j,K+1}$ is the amount of cash in j's account at time t (idem for i)

$Credit_j$, is the credit of j available from the SSS

$\overline{A}$ is the maximum amount of security k that i can buy from j

$\overline{B}$ is the maximum amount of security k that player i can sell to player j.

Trades that are not settled become zero entries in the S(t) matrix. Trades that do settle are filled in and L(t-2) is updated. Unsettled trades are put in a queue Q(t)[11]. As explained above, these trades will be examined again, in a second "batch", once all the trades ought to be settled have been examined once. In the simulation three batches are used. Trades that do not settle after three batches stay in the queue and are added at the end of the first batch in the next settlement period. At the end of the settlement process on day t, L(t-1) becomes known.

3.5 Determining the stress scenario

We assume that due to external factors, the participant with the largest total asset value is not able to fulfill its obligations on day D and the days thereafter. We also assume that during day D rumors of the imminent default begin circulating in the market, and participants react by avoiding all trades on day D with the troubled participant. This allows investigation of the impact of a default that is anticipated and which will result in less of a shock than an unanticipated default. Actual failure of the participant is assumed to happen at the close of the trading period on day D, but before the settlement period begins. Before settlement begins, all of the trades of the failing participant from day D-2 and D-1are deleted, and settlement of other trades from D-2 proceeds.

3.6 Reactions of counterparties

From the moment a counterparty fails, settlement failures may be expected. First of all, there is the direct effect, i.e. the failure of all unsettled transactions with the failing participant. This implies that all trades committed to by the defaulter but not yet settled will be deleted. In our case with a settlement lag of two days and with anticipation on day D of the default (and, therefore, no trades undertaken on day D with the defaulting institution), the direct effect will include settlement failures of trades undertaken on days D-2 and D-1, to be settled on days D and D+1, respectively.

[11] Note that as in reality, partial settlement , i.e.; splitting a big trade into many small ones that are than settled separately, is not allowed. This is because participants may be unhappy with such a procedure as some prefer full or no settlement rather than partial settlement.

In addition to the direct effects of the default, there will be indirect effects; i.e. second-round effects as counterparties of the defaulter become short in cash and/or securities needed to fulfill their settlement obligations, in turn causing further settlement failures and further contagion.

In addition, knowing that some previously committed trades are likely not to settle as a result of second-round effects, participants know that expected holdings under the assumption of full settlement, $E[L(D)]$ and $E[L(D+1)]$, will differ from the true holdings L(D) and L(D+1) that will be determined through settlement. Hence, participants need to update their expectations about the budget constraints upon which new trades are based. We assume that participants know their counterparties; therefore, they can calculate the direct effect of the default by deleting the unsettled trades they undertook with the defaulter. It is impossible to know how their other counterparties will be affected by the default, and therefore, to calculate what the indirect effects of the default will be. The only thing they can do is to narrow their trading within the expected budget constraint that they would have had if all previous trades had settled. Thus, we assume that from day D onwards, budget constraints will be narrowed by an adjustment margin ε, with 0<ε<1. A value of ε=1 would imply that there is no longer any trading.

3.7 Calculation of settlement efficiency

As an aggregate measure of liquidity risk, settlement efficiency is used. Settlement efficiency is determined by dividing the aggregate value of settled trades by the aggregate value of trades needing to be settled. We distinguish between two measures: total settlement efficiency (θ(t)) and indirect settlement efficiency ($\theta^*(t)$). In the first measure the denominator includes all trades committed two days earlier, plus all the trades that did not settle previously; i.e., the ones in the queue Q(t). For $\theta^*(t)$ on the other hand, the trades that are to be settled do not include the trades of the defaulting participant. Hence, this is a measure of contagion in the settlement system. In symbols,

$$\theta(t) = \frac{\sum_{i>j}^{N}\sum_{j=1}^{N}\sum_{k=1}^{K}\left| S(t)_{ijk} \right|}{\sum_{i>j}^{N}\sum_{j=1}^{N}\sum_{k=1}^{K}\left| T(t-2)_{ijk} \right| + \sum_{t}\sum_{i>j}^{N}\sum_{j=1}^{N}\sum_{k=1}^{K}\left| Q(t)_{ijk} \right|} * 100 \text{ and} \qquad \text{eq. 3}$$

$$\theta^*(t) = \frac{\sum_{i>j}^{N}\sum_{j=2}^{N}\sum_{k=1}^{K}\left| S(t)_{ijk} \right|}{\sum_{i>j}^{N}\sum_{j=2}^{N}\sum_{k=1}^{K}\left| T(t-2)_{ijk} \right| + \sum_{t}\sum_{i>j}^{N}\sum_{j=2}^{N}\sum_{k=1}^{K}\left| Q(t)_{ijk} \right|} * 100 \qquad \text{eq. 4}$$

where participant j=1 is the defaultor.

4 SIMULATION RESULTS

4.1 Parameters

Simulations have been undertaken for the following combinations of parameter values: N=5 or 15; K=20 or 30; σ^2=1; ρ=0.9; β_1=0.2; β_2=0.2 or 1; ε=0.2, λ=0, 0.15 or 0.3; C=0.05. For every possible combination of parameters (24 combinations in all) we run the model 100 times, giving a total of 2400 runs[12].

Although N and K, which represent the numbers of participants and securities, are set at relatively low levels (partly for computational reasons), it may be argued that, compared with observed numbers of participants and securities in some real SSSs, these parameter values are not necessarily low. In reality, although many securities may exist, a relatively small subset of securities may be traded among a few active participants such as large custodians, central counterparties or specialized traders. Even in systems with many active participants, the bulk of trades is often initiated by a relatively small number, say a dozen, players. There are also many SSSs with only a dozen participants.

Varying the parameters σ^2 and ρ influences the distribution of initial allocations, making it possible to create scenarios where participants have very different initial portfolios and asset sizes. The parameter β_1 of the trade distribution, which is used for all trades not involving the defaulting participant, is set at a relatively low value of 0.2 in order to simulate a market with extreme trading behavior; i.e., where participants tend to submit trades close to their budget constraints. This will of course increase the number of unsettled trades when a settlement disruption occurs. On the other hand, β_2 is allowed to vary, in order to asses the impact on settlement efficiency of different trading profiles for the defaulting participant.

The parameter ε is the adjustment margin that surviving participants use to narrow their budget constraints for trading in response to a crisis. A value of ε=0.2 implies that participants assume that indirect settlement failures will account for around 20% of previously committed trades. The credit availability parameter λ varies from 0 to 0.15 up to 0.3, representing the proportion of the value of the initial portfolio which can be borrowed during the settlement process.[13] A low value of λ represents an SSS with only a limited liquidity program. Depending on the level of λ, the liquidity program may or may not be able to significantly increase settlement efficiency. The parameter C determines the amount of cash held by participants as a percentage of total assets in the initial portfolio.

---

[12] Future versions of the paper report results for 10,000 simulations.

[13] λ is deliberately set lower than 1, since even with full collateralisation, not all assets may be eligible, and haircuts need to be applied. It can be shown that even for λ=1 settlement failures will remain

4.2 Results

4.2.1 Turnover and trade statistics

For illustrative purposes, Table 1 displays some turnover statistics of the SSS for a normal day (i.e., where no defaults have occurred) for different values of N and K. The reported figures are those for β_2=0.2 (extreme trades for transactions involving the largest player). The order of magnitude is more or less the same for β_2=1 (not reported). The number of possible trades on a single day is N*N*K (i.e. each participant trades with every counterparty in every security). The table shows that turnover increases with N and K. Hence, in this model a rise in the total number of outstanding securities and counterparties increases the turnover. Note that total turnover can be higher than the amount of securities outstanding. This may seem counterintuitive, as in reality turnover as a percentage of securities outstanding is often less than 10%. However, this should not be cause for concern if one interprets K as the number of outstanding securities which are actively traded and posted on the account of the participants in the SSS itself and not on a sub-account with a client of the participant.

Table 1: Basic turnover statistics	**K=20**		**K=30**	
	N=5	**N=15**	**N=5**	**N=15**
Average	13,98	44,02	22,37	77,75
Std. dev.	1,51	1,98	2,52	3,09
Max	18,04	49,18	29,55	86,99
Min	9,73	34,90	14,91	68,90
Number obs	300	300	300	300

Note: only β_2=0.2 is reported. Results are more or less similar for β_2=1

Table 2 reports statistics, for differing values of N and K, relating to the market share on a normal day of the largest participant (i.e., largest initial portfolio). This is also the participant that is assumed to default in our simulations. Obviously, for low values of N the market share of the biggest participant is higher (ranging from 37% to 75%), than for high values of N (ranging from 12% to 25%). The reported values for N=15 are in line with what is observed in many SSSs in Europe, where often a few participants generate the largest proportion of the business.

Table 2: Market share of largest participant (defaultor)	**K=20**		**K=30**	
	N=5	**N=15**	**N=5**	**N=15**
Average	0,56	0,17	0,61	0,20
Std. Dev.	0,06	0,02	0,05	0,02
Max	0,76	0,23	0,74	0,26
Min	0,37	0,12	0,46	0,16

Note: only β_2=0.2. Similar results for β_2=1

4.2.2 First day impact (D-Day)

The largest participant is assumed to default on all of its outstanding obligations on day D. Tables 3 (a and b) and 4 (a and b) illustrate the impact on average values of settlement efficiency θ and θ^* for differing values of N, K and λ. Several observations can be made. First, θ and θ^* vary considerably across different parameter combinations. The value of average θ (total settlement efficiency) ranges from 18% to 77%, while θ^* ranges from 47% to 93%. Settlement efficiency appears to be higher for higher β_2 (less extreme trades with the defaulting participant). Also, settlement efficiency varies considerably across the simulations for a given parameter combination, as can be seen from the relatively high standard deviations. θ appears to be positively related to λ and N and negatively related to K. A higher λ increases participants' credit limits, thereby softening their budget constraints during settlement. A higher N results in a lower market share for the failing participant. In turn, the number of deleted trades upon default of this participant is relatively lower and, hence, the initial shock is smaller, leading to higher values of θ and θ^*.

Table 3a: total settlement efficiency θ

(β_2=0.2)

		K=20		**K=30**	
λ		**5**	**15**	**5**	**15**
0	average	22,04	44,14	18,45	42,68
	stdev	9,08	4,89	9,34	5,08
0,15	average	35,22	69,70	30,40	64,57
	stdev	6,45	3,95	4,93	3,24
0,3	average	35,82	72,85	30,82	66,34
	stdev	6,56	2,71	4,96	2,55

Table 3b:indirect settlement efficiency θ^*

(β_2=0.2)

		K=20		**K=30**	
λ		**5**	**15**	**5**	**15**
0	average	49,41	53,42	47,12	53,62
	stdev	18,57	5,83	22,60	6,26
0,15	average	79,97	84,49	76,45	80,64
	stdev	9,14	3,52	8,85	3,23
0,3	average	81,14	88,13	77,70	83,22
	stdev	7,82	2,34	7,31	2,35

Table 4a: total settlement efficiency θ

(β_2=1)

		K=20		K=30	
λ		**5**	**15**	**5**	**15**
0	average	32,80	64,25	29,83	60,80
	stdev	8,13	4,20	6,72	3,95
0,15	average	42,26	77,55	37,97	73,08
	stdev	4,36	2,09	3,50	1,641
0,3	average	41,99	77,84	38,44	73,55
	stdev	4,48	1,95	3,48	1,71

Table 4b: Indirect settlement efficiency θ^*

(β_2=1)

		K=20		K=30	
λ		**5**	**15**	**5**	**15**
0	average	68,61	77,30	67,33	75,55
	stdev	15,66	4,70	14,84	4,88
0,15	average	88,28	93,18	85,66	90,80
	stdev	4,82	1,56	5,33	1,59
0,3	average	87,92	93,78	86,36	91,30
	stdev	4,98	1,31	4,70	1,43

Increasing λ from 0.15 to 0.3 does not appear to have a large impact on settlement efficiency (contrary to the increase from 0 to 0.15). This is possibly because liquidity cannot solve settlement problems for participants who find themselves short in securities.

4.2.3 Net buy vs. net sell position of the defaulter

We expect there to be a relation between the net trade position of the defaulter and settlement efficiency. The default of a net buyer extracts cash from the system. Hence, some participants end up with an unanticipated short position on the cash side. As cash is used in every transaction, this may lead to significant contagion and hence low settlement efficiency. In contrast, when the defaulter had a net sell position, counterparties become constrained on the securities side. As each security is only used in transactions of that particular security, contagion may be weaker and settlement efficiency is higher. One may argue that as long as there is enough liquidity (cash), all participants can fulfill their buy obligations, making it impossible to become short in securities for the next trade that needs to be settled. This reasoning, however, does not take into account that due to the initial default, some securities remain with the defaulting participant. As a result, the aggregate pool of securities available for the remaining participants is smaller. Even unlimited credit cannot replace these missing securities.

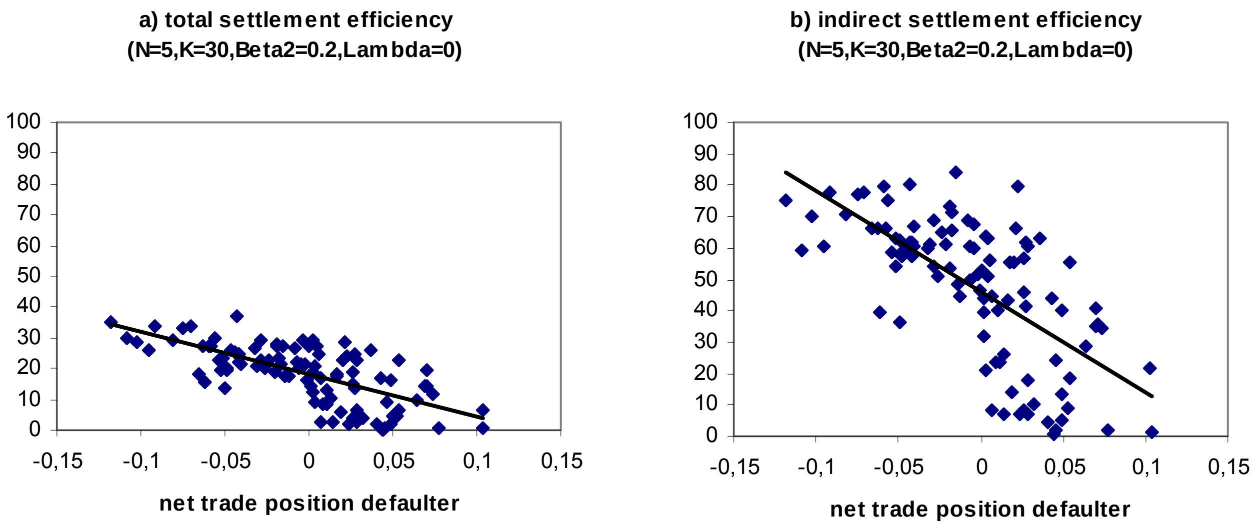
Figure 2 (a, b, c, d) plots direct and indirect settlement efficiency against the net trade position of the defaulter on trades that should have settled on day D (i.e., on trades from D-2). The measure for the net trade position of participant j=1 (the defaultor) controls for the volume traded and is defined by:

$$\frac{\sum_{i>j}^{N}\sum_{k=1}^{K} T_{ijk}}{\sum_{i>j}^{N}\sum_{k=1}^{K} \left|T_{ijk}\right|}. \qquad \text{eq. 5}$$

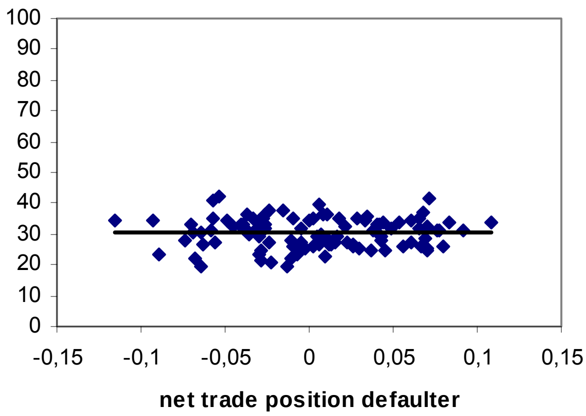
Note that a negative value of T_{ijk} represents a purchase of security k by participant i from j, or a equivalently sale of the security by j to i. Similarly, a positive value denotes a purchase of security k by participant j from i. Thus, a negative value for the trade position measure signifies that participant j is a net seller and a positive value signifies that j is a net buyer.

The graphs in Figure 2 show that for low values of λ, the net trade position of the defaulter seems to have an important impact on settlement efficiency. A net buy position of the defaulter has a negative impact on settlement efficiency and in some cases leads to a near complete breakdown of settlement. A defaulter with a net sell position, on the other hand, leads to relative higher settlement efficiency. The net buy position also seems to play a more important role in determining indirect settlement (θ^*) than total settlement (θ). In addition, once participants are able to draw on credit lines during settlement (higher λ), the impact of net trade position on settlement efficiency seems to disappear.

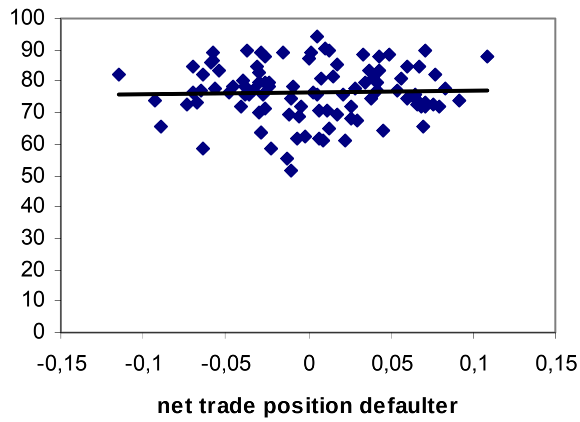
Figure 2 (a, b, c, d): net trade position and settlement efficiency



c) total settlement efficiency
(N=5,K=30,Beta2=0.2,Lambda=0.15)



d) indirect settlement efficiency
(N=5,K=30,Beta2=0.2,Lambda=0.15)



In order to get some idea of the relative importance of the different parameters in the model, we regress the values of settlement efficiency obtained from the simulations on a constant, N, K, λ and β_2 and the variable net trade position of the defaulter.[14] The results are presented in Table 5. Columns 1 and 2 of this table present the OLS estimates based on results from all simulations, with regressions using θ as the dependent variable in column 1 and $\theta*$ in column 2. The remaining columns present separate regression results depending on the value of λ. All variables in the regressions are significantly different from zero, with the exception of the coefficient of net trade position when $\lambda = 0.15$.

One potential remark that may be made with respect to Table 5, however, is that the net trade position may not be independent of some of the other independent variables (e.g. N and K). The largest participant's net trade position might be determined by the market structure, which itself depends on N and K. However, regressions of net trade position on N and K result in no coefficient being significantly different from zero[15].

---

[14] For λ only the value 0 and 0.15 are included here.

[15] However, there still seems to be a problem of heteroskedasticity for the regressions included in table 5. While this does not bias the coefficients, it may partly explain the very high values for the t-statistics. Future versions of the paper will investigate this issue further.

Table 5: OLS estimates

R^2	0,90	0,69	0,89	0,64	0,95	0,49
R^2-adj	0,90	0,69	0,89	0,63	0,95	0,48
Nobs	1600	1600	800	800	800	800
λ	0-0,15	0-0,15	0,00	0,00	0,15	0,15
dependent variable	θ	$\theta*$	θ	$\theta*$	θ	$\theta*$
Variable	Coefficient	Coefficient	Coefficient	Coefficient	Coefficient	Coefficient
constant	9,24 (10,68)	49,79 (34,89)	8,17 (7,11)	42,19 (19,83)	24,97 (30,62)	81,16 (73,25)
N	3,10 (103,31)	0,58 (11,67)	2,73 (67,51)	0,78 (9,45)	3,48 (121,08)	0,47 (11,9)
K	-0,38 (-12,55)	-0,22 (-4,39)	-0,30 (-7,39)	-0,15 (-2,04)	-0,47 (-16,29)	-0,31 (-7,92)
λ	96,63 (48,33)	156,26 (47,35)				
β_2	14,29 (38,11)	19,03 (30,75)	18,97 (37,54)	26,85 (28,67)	9,68 (26,96)	11,36 (23,3)
net trade	-58,49 (-13,97)	-117,61 (-17,02)	-115,26 (-20,53)	-231,26 (-22,23)	0,46 (0,12)	-2,80 (-0,51)

Note: t-statistic between brackets, N: number of participants, K: number of security issues, λ : credit margin, β_2 extremeness of trades of defaulter, net trade: net buy position of defaulter as a percentage of its total trades.

Table 5 indicates that settlement efficiency is positively related to N, the number of participants. This can be explained by the fact that the higher is N, the more the initial shock is distributed among multiple participants, making it easier to absorb. The impact of N, however, appears much stronger for θ than for $\theta*$, reflecting the importance of the number of participants to the initial shock but less so for the second-round effects reflected in $\theta*$. The coefficient of K, the number of securities, is statistically significant and negatively related to settlement efficiency. The greater the number of securities traded in the model, the longer the chain of contagion can be, resulting in lower settlement efficiency. These results suggest that small SSSs, with a low number of active participants but relatively high number of securities, may be most vulnerable to contagion.

It comes as no surprise that the trading behaviour of participants is also important. The less extreme the trades (i.e. the higher is β_2) and the smaller is the net buy position, the higher are θ and $\theta*$. The credit limit λ is also an important variable in determining settlement efficiency. The estimated coefficients suggest that an increase in participants' liquidity margin from 0% to 15% raises θ by more than 14% and $\theta*$ by more than 23%.

The discussion above has suggested that there may be a structural break for different values of λ. This is tested by estimating the OLS coefficients for two separate subsamples, one for λ=0 and the other for λ=0.15. For λ=0, net trade is highly statistically significant. For λ=0.15 on the other hand, one cannot reject the null hypothesis that the coefficient of net trade equals zero. The importance of β_2 also appears to be lower (although still significant) when λ=0.15. These results provide some support for the idea that SSS liquidity may be an important tool in absorbing the contagion stemming from extreme trading behavior.

4.2.4 Length of the crisis

Not only is the impact of a disruption on the first day important, but it is also of interest to know what happens in subsequent days. This is especially true if settlement efficiency is used as a measure of liquidity risk in SSSs. For example, if a large number of unsettled trades on the first day of a stress event can be recycled into the next settlement period, and if settlement then succeeds, then replacement cost risk and liquidity risk will be judged to be limited, as settlement will have only been delayed by one day. However, if settlement failures persist, uncertainty will remain as well. Participants in need of the cash or securities from these failed transaction may prefer to revoke their trades and to conclude new transactions, even under less favorable terms.

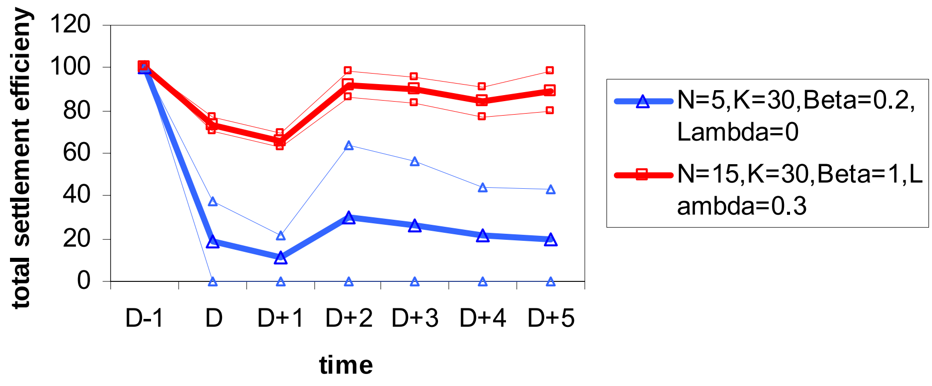
Given the results discussed above, the three graphs in Figure 3 focus on three parameter combinations: a worst-case; a best-case; and an intermediate case. In the worst-case scenario, N=5, K=30, β_2=0.2 and λ=0. Here, settlement efficiency suffers the most, as participants have no credit margin to rely on, while at the same time trades are extreme and the market share of the defaulting participant – hence the first-day impact – is high. In the best-case, N=15, K=30, β_2=1 and λ=0.3.[16] In this case, credit margins are high, trades average and first day impact is small. The third case is more intermediate. The only difference is N=5 instead of N=15 as in the best case.

---

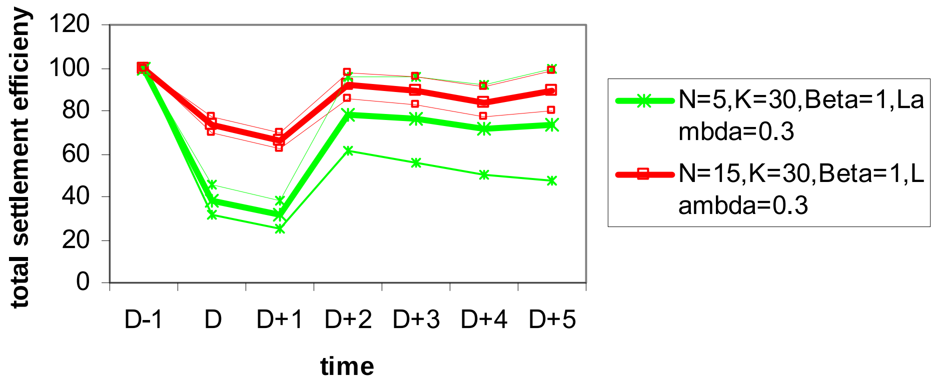
[16] Here, the term best-case is not an absolute term but refers only to the most favorable combination of parameters among those used in the simulations.

Figure 3 (a, b, c): Length of the crisis

a) Best-case vs. worst-case



b) Best case vs. intermediate case



c) Worst case vs. intermediate case

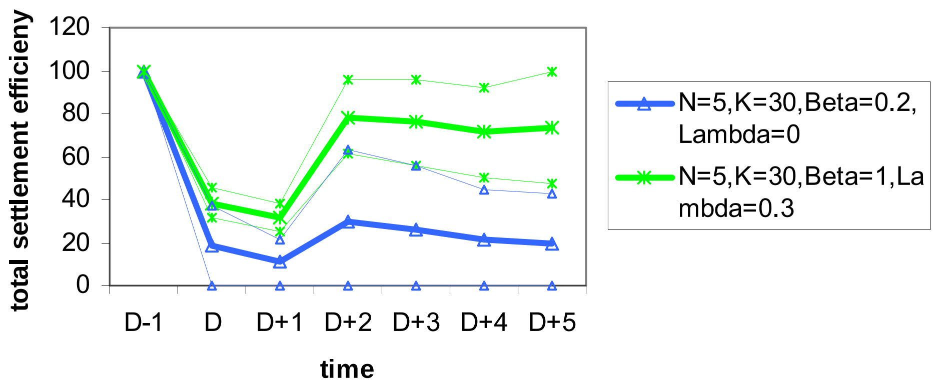


Figure 3 presents graphs of settlement efficiency over a period of several days (from D-1 up to D+5). In each figure two parameter combinations are compared.

Figure 3a compares the best-case with the worst-case. Figure 3b compares the best case with the intermediate case and Figure 3c compares the worst-case with the intermediate case. Each time, the thick lines represent the average value of total settlement efficiency across simulations, and the thin lines represent two standard-deviations from the average. Only total settlement efficiency is presented in the graphs, as from day D+2 onwards there are no more trades from the defaulted participant that need to be settled. Hence, direct and total settlement efficiency become the same.

Figure 3 clearly shows that even in a SSS with DVP and gross settlement there is still a possibility of a significant, multi-period disruption of settlement activity when a large participant fails. This is even the case when the default event was anticipated by the market (as is the case in our model). In the worst-case scenario, settlement efficiency falls drastically and does not improve rapidly. This is in sharp contrast with the best-case scenario, where the initial fall in efficiency is low and increases immediately thereafter. This rapid increase in settlement efficiency occurs as a result of ample liquidity provision by the SSS. This is also the case for the intermediate outcome, where the initial shock is relatively high but efficiency is restored to a large extent, due to the liquidity provided.

An important observation arising from all three figures is that even with generous liquidity provision, settlement efficiency does not return to its "pre-stress event" levels within 5 days of the crisis. This may appear counterintuitive, as the settlement lag is only two days, suggesting that settlement efficiency should be restored to its pre-crisis level after day D+2. Why is settlement efficiency not restored? It is clear that one impact of the default by a participant is to cause previously unanticipated settlement failures on day D and D+1 (of trades undertaken on days D-2 and D-1), because the trades that need to be settled on these dates were committed prior to knowledge of the default. Hence, participants could not take the upcoming default into account at the moment when they undertook these trades.

Settlement on day D+2, however, pertains to trades undertaken on day D, but at this point the default was anticipated and no participants traded with the defaulting institution. The explanation for the persisting settlement failures is linked to expectations formation by the surviving participants. As mentioned in section 3, although participants stop trading with the defaulter on day D and although they can calculate the direct effects of the default on the trades they undertook with the defaulting institution on days D-2 and D-1, they do not know how their counterparties will be affected, or equivalently, what the second-round effects will be. The only thing they can do is to assume that some proportion of their previously committed trades will not settle and to limit their current trading as a result. It is assumed that participants narrow their expected budget constraints by 20%. Thus it is still possible that actual holdings of cash and securities after settlement will be lower than the assumed 80%, and that additional settlement failures will occur. This is indeed what happens in the simulations. This does not mean, however, that settlement efficiency is never completely restored in any simulation. In almost 10% of the simulations settlement efficiency is back up to 95% on day D+2. In less than 1% of the simulations settlement efficiency reaches 100%

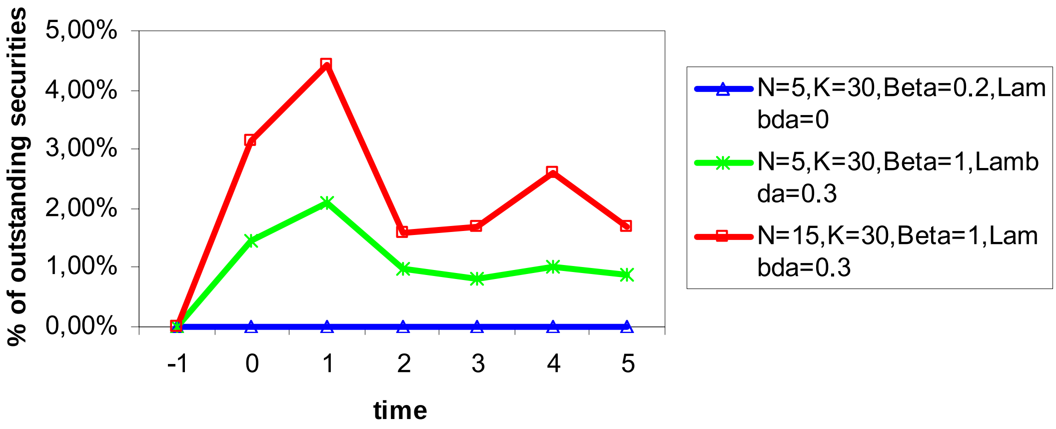
on day D+2[17] Also, a higher value of ε (representing a more conservative assumption about second-round effects) and, therefore, more restriction of trading on days D and D+1 would result in higher settlement efficiency after day D+2.

Differences across Figures 3a-c suggest that λ and ε are partial substitutes. That is, either very conservative reactions by participants to the crisis (large ε) or ample liquidity (high λ) can lead to relatively high settlement efficiency. However, these two alternatives lead to a tradeoff from a financial stability perspective. Generous liquidity provision places a heavy burden on the liquidity provider but does not reduce trading activity, while conservative reactions by market participants avoid the burden on the liquidity provider but entail a fall in trading activity (resulting in less liquid markets). From a financial stability perspective a balance will need to be struck between the two.

4.2.5 Credit use

The above discussion has suggested that liquidity provision by the SSS (or a central bank supporting the SSS) is an important policy tool for improving settlement efficiency in periods of market disruption. Access to liquidity in times of stress weakens participants' cash constraints, resulting in higher settlement efficiency. Unfortunately, however, this solution may not always be possible, as generous liquidity provision may be judged by the SSS to be too costly or too risky. Even if the SSS is supported by a central bank, there may be limits on the amount of credit that the central bank is willing to provide. Figure 4 shows the amount of end-of-day credit provided by the SSS in the model of this paper as a percentage of the value of total securities in the system.

Figure 4: Average end of day credit provided by the SSS



In the best-case scenario for settlement efficiency, where the SSS provides participants an emergency credit line equal to 30% of their total assets, aggregate end of day credit reaches a

[17] Note that settlement efficiency is measured in value terms and not in volume (i.e. number of trades), as the former measure seems more appropriate for drawing conclusions relating to financial stability. If only a single trade does not settle (of which it is highly likely to be a large one), settlement efficiency easily drops to 95%, while settlement efficiency in volume terms would still be around 99.99%.

peak on D+1 of 4% of outstanding securities in the SSS[18]. It is questionable that an SSS would be willing or able to provide such a large amount of liquidity over several days, even if the SSS is supported by a central bank, as such an amount might have an impact on monetary policy objectives.

Figure 4 shows that end of day credit increases with N, as can be seen when comparing the best-case scenario with the intermediate case. This implies that whereas settlement efficiency increases with higher values of N, the cost of restoring efficiency is higher as well. The intuition for this result is straightforward. Whereas in this model a larger number of participants in the market makes it easier to absorb the initial shock, the larger number of participants also implies a longer chain of trades, which increases the likelihood that liquidity will be required somewhere down the chain. On aggregate, end of day credit use increases with N.

5 CONCLUSIONS, POLICY IMPLICATIONS AND ONGOING RESEARCH

This paper has demonstrated that liquidity risk may be important in SSSs with gross settlement, even in systems with delivery versus payment and adequate liquidity provision. Although DVP systems eliminate principal risk, these systems do not eliminate replacement cost and liquidity risk. Moreover, settlement disruptions may last over a period of days. The paper uses a multi-period model to analyze the extent and dynamics of settlement failures that may occur. From a financial stability point of view, it is important to understand the mechanics of breakdowns in settlement efficiency, the factors exacerbating disruptions, and the policy tools that may help to resolve crises.

The results suggest that settlement failures due to the default of a large participant may be higher in SSSs with only a limited number of participants but relatively many traded securities. The trading behavior of participants also appears to be important. A defaulting participant who tends to trade close to the borders of its budget constraint and who has a net buy position will cause larger settlement failures than a defaulting participant with less extreme trade behavior and a net sell position. The importance of the net trade position can be explained by the fact that cash is used in every transaction, while a security is only used for transactions involving that security.

When enough liquidity is provided by the SSS, the relation between the defaulting institution's net trade position and the degree of settlement efficiency appears to disappear. In other words, liquidity provision by the SSS can help participants to absorb the shock created by the default of a player. This suggests that liquidity provision can be an important policy tool for central banks in supporting the functioning of financial markets. However, injecting enough liquidity in the system to prevent severe contagion of settlement failures may prove to be quite costly or may interfere with other policy objectives. Moreover, because securities transactions involve both a security and a cash leg,

---

[18] The term credit line makes reference to the total credit available to settle transactions. The term overdraft makes reference to the end of day credit opening. Hence, if λ =0.3, the aggregate intraday peak credit usage is maximum 30% of outstanding assets. Generally, end of day overdraft will be much smaller as sell positions partly set off the credit needed to finance part of the buy positions.

liquidity provision cannot completely eliminate settlement failures due to major market disruptions, as cash affects only one side of transactions.

Along these lines, it might be desirable in future versions of the model to examine the effect of securities lending and borrowing. It would be necessary, however, to make sure that the size of the lending pool is endogenously determined, and to allow for changes in participants' willingness to lend securities during crises. Also, market risk might be included in the model, by allowing relative asset prices to change over time.

The result that settlement failures can be severe over a period of days is potentially important information for SSSs. This result suggests that assessments of liquidity risk that only focus on the day of the disruption may significantly underestimate the total amount of settlement failures and the ultimate amount of liquidity needed to guarantee timely settlement in case the largest participant fails. One potential way to shorten the potential length of crises is to try to limit the lag between trade and settlement. If technology would allow for real time t+0 settlement, for example, participants would not need to form expectations about their cash and security holdings. Although settlement failures in response to a major disruption would still arise, there would no longer be multi-day second-round effects.

Given what appears to be the technical infeasibility of real-time settlement, one might wonder whether there are significant benefits in moving to T+1 settlement versus T+2. What would be the magnitude of such benefits relative to those of moving from T+3 to T+2? This question is being addressed in ongoing work.

Other questions can also be addressed. For example, one of the features of the current model is that the market share of the largest participant (size) is linked to the number of participants in the market (structure). One might want to separate these two effects. The same is true for the volume of transactions and the length of transaction chains, which increase mechanically with the number of participants.

The welfare dimension might also be further developed – at the expense of adding more complexity to the model – by embedding participants' trading behavior in a utility framework. Such a framework might enable one to address questions such as what the optimal level of liquidity is in response to a market disruption. A utility framework might also allow for endogenous determination of participants' reactions to default; i.e., in setting the ε parameter. Is there a welfare trade-off between providing liquidity and allowing participants to limit their trading? Finally, a utility framework might allow for a richer response by participants to default, such as opting for less extreme trading behavior (an increase in β_1) in addition to using the parameter ε to limit trades.

6 REFERENCES

Allan F. and Gale D., 2000, " Financial contagion", *Journal of political economy,* vol 108, no 1, 1-33

Angelini P., Maresca G., Russo D., 1996, "Systemic risk in the netting system", *journal of banking and finance*, 20, 853-868

Angilini P., 1998, "An analysis of competitive externalities in gross settlement systems", *Journal of banking and finance,* 22, 1-18

BIS, 1992, "Delivery versus payment in security settlement systems", CPSS

Cifuentes R., Ferruci G. and Shin H., 2004, "Liquidity risk and contagion", *forthcoming in Journal of European economic association*

De Bandt O. and Hartmann P., 2000, "Systemic risk: a survey", *CEPR discussion paper 2634*

Diamond D. and Dybvig P., 1983, "Bank runs deposit insurance and liquidity", J*ournal of political economy,* 91, 401-419

Diamond D. and Rajan R. 2003, "Liquidity shortage and banking crises", *NBER Working Paper* No. w10071

Flemming M. and Garbade K., 2002, "When the back office moved to the front burner: Settlement fails in the treasury market after 9/11", *FRBNY policy review*, November

Furfine, Craig H., 2003, "Interbank Exposures: Quantifying the Risk of Contagion." *Journal of Money, Credit and Banking* 35, February, 111-28.

Humphrey D., 1986, "Payments finality and riks of settlement failure", in *Technology and the regulation of financial markets: Securities, futures and banking*, edited by A. Saunders and L. White, Lexington, MA: Lexington books, 97-120

Iori G., 2004, "An analysis of systemic risk in alternative securities settlement architectures", *ECB working paper*, no 404, November

Kahn C. and Roberds W., 1998, "Payment system settlement and bank incentives", *The review of financial studies*, vol 11, no 4, 845-870

Kahn C., MCandrews J. and Roberds W., 2003, "Settlement risk under gross and net settlement", J*ournal of money credit and banking*, vol 35, no 4

Leinonen H. and Soramaki K., 1999, "Optimising liquidity usage and settlement speed in payment systems", *Bank of Finland discussion papers*, 16/99, 1-39

Nguyen and Degryse, 2004, "Interbank exposures: An empirical examination of systemic risk in the Belgian banking system", *National bank of Belgium working papers*, no 43

Northcott, 2002, "Estimating settlement risk and the potential for contagion in Canada's automated clearing settlement system", *Bank of Canada working paper 2002-41*

Rochet J. and Tirole J., 1996, "Interbank lending and systemic risk", *Journal of money credit and banking*, 28, no 4, 733-762

Upper C. and Worms A., 2002, "Estimating bilateral exposures in the German interbank market: Is there danger of contagion?", Economic research centre of the Deutsche Bundesbank, discussion paper, no 09/02