Simulating retaliation in payment systems: Can banks control their exposure to a failing participant?
Simulating retaliation in payment systems
Can banks control their exposure to a failing participant?

Elisabeth Ledrut*

* Views are those of the individual author and do not necessarily reflect official positions of De Nederlandsche Bank or the Bank for International Settlements
Simulating retaliation in payment systems
Can banks control their exposure to a failing participant?

Elisabeth Ledrut

Abstract

This paper assesses the impact of an operational failure at one of the biggest participants in the Dutch interbank payment system, varying the time at which the disruption takes place. Liquidity levels equal historical levels. The impact of such a disruption is quantified in terms of the additional liquidity needed in order to settle all payments than can settle given the banks’ intraday reserves and collateral facilities. Assuming the disruption lasts for the remainder of the day, banks are faced with costs, as they need to borrow this additional liquidity overnight from the market or from the central bank. As could be expected, the second-round effect (the number of unsettled payments among healthy banks) of an operational disruption is highest when it occurs early during the day and lasts for the remainder of the day. So are the additional overnight liquidity needed and the costs of overnight credit.

Furthermore, the paper introduces different possible reaction patterns from the stricken bank’s counterparties. These counterparties can react according to two basic rules: they stop sending payments to the stricken bank either after some pre-determined time or after their exposure to the stricken bank reaches a certain level. From a cost perspective, reacting is more effective when determined by the individual exposure of the stricken banks’ counterparties. However, even an immediate reaction does not prevent banks from running losses following the failure of a major participant.

This leads to a reflection about the bilateral relations between the stricken bank, considered to be a node in a partial star network, and the other banks. How much each payment system participant can control its exposure to the stricken bank depends on the degree of reciprocity in the value of bilateral payments.

Key words: payment system, operational disruption, liquidity
JEL Classifications: C88, G21

1 Elisabeth Ledrut is on secondment from De Nederlandsche Bank at the Secretariat of the Committee on Payment and Settlement Systems (Bank for International Settlements). Contact information: tel. +41 61 280 81 71 or email elisabeth.ledrut@bis.org.

I would like to thank Judith Looman, Odette Min, Gwan Tjia and Richard Heuver for their very helpful statistical support; Jaap Mauritz for his insights in Dutch banks’ behaviour; Matti Helqvist for his always speedy and efficient advice with the simulator; Harry Leinonen and participants to the Helsinki seminar, as well as Ronald Heijmans and Ronald Bosman (who refereed this paper) for their valuable comments on an earlier version of this work. All remaining errors are my own. The views expressed in this paper are those of the author and do not necessarily reflect those of De Nederlandsche Bank or the Bank for International Settlements.
Introduction

The Payment and Settlement Simulator (“the simulator”), developed by the Bank of Finland, has proven to be a useful tool to gain insight in risks and efficiency in payment systems. It has been used to estimate the trade-off between the liquidity needs of different types of systems and their level of risk (Koponen & Soramäki, 1998, or Leinonen & Soramäki, in Leinonen eds., 2005, among others). Delayed net settlement (DNS) systems, which settle all payment orders at the end of a pre-determined period of time by transfers of net liabilities, are least demanding in terms of liquidity. However, they exhibit the highest level of settlement risk as, in case of a participant failure, all processed transactions may be unwound. Real-time gross settlement (RTGS) systems, which settle payments immediately, require the biggest share of funding from participants, but they offer certainty with regard to the receipt of funds. In-between RTGS and DNS lies a whole range of system designs offering various liquidity-saving mechanisms.

The simulator has also been extensively used to test the potential effect of an operational disruption, either at a major market participant (where the “stricken bank” fails to submit payments during a day or part of the day) or at the system itself (ceasing to function as from a certain time). These studies rely on historical data, adapted in order to create various scenarios for the disruption (McVanel, 2005, among others). In general, the objective of these studies is to assess the systemic impact of a disruption. However, using historical data as such presupposes that, were a disruption to occur, (other) payment system participants would not adapt their behaviour to the new environment – an exercise that could be subject to a form of Lucas critique.

Behavioural aspects have received less attention. This can be explained by a lack of knowledge about participants’ reaction in the face of a disruption, as well as about the factors affecting participants’ reaction patterns. There is some historical evidence with regard to wide-scale disasters like September 11, but one should be cautious when generalising findings from such exceptional events. There is also some evidence, mainly about participant’ reactions to failures to pay by other participants, as observed by payment system operators. Work by Bedford, Millard & Yang (2005, further Bedford et al.) is based on the assumption that participants in the United Kingdom’s payment system would react to such a failure within ten minutes. More generally, one study in particular (Conover, 2005) researched occurrences of disruptions and estimated the strength of other participants’ reactions. ²

In theory, the reaction of payment system participants to (the absence) of incoming payments can be expressed in the form of a reaction function (McAndrews, 2002). Such a reaction function

---

² It appears that, depending on the gravity of the disruption, participants in the U.S.’ large-value payment system Fedwire tend to retain up to 27% of the value of their payments during disruptions (up to 15% when September 11-14 are excluded), while the volume of payments is not affected.
is based on the assumption that part of participants’ outgoing payments is triggered by payment receipts and that payment orders exhibit a certain degree of synchronicity.  

What would be the consequences of a failure to pay by one of the major participants in the Dutch interbank payment system? And how would other participants’ reactions determine the impact of a disruption on the payment system? Insight in the influence of payment system participants’ behaviour on the system would allow central banks to identify and encourage stabilising behaviours, thereby lowering the potential systemic impact of operational disruptions.

This paper assesses the impact of an operational failure, varying the time at which the disruption takes place, with the help of the simulator. Although the relation between the timing of the disruption and its gravity has not been analysed before – most papers concentrate on the “worst-case scenario”, derived from the banks’ intraday payment patterns - the main added value of this exercise comes from the explicit quantification of the impact at each hour of the day. On the contrary to other work, the liquidity available to banks is considered as fixed. When a participant fails to pay, other participants bear the burden of that failure by having to rely on intraday credit from the central bank, which is provided against collateral. If banks have not received the expected payments at the end of the day, they will need to borrow overnight either from the central bank (this is called the marginal lending facility in the Euro area) or from the market. The additional liquidity needed at the end of the day and the cost of borrowing this liquidity from the market or from the central bank have been quantified.

Furthermore, the paper introduces different possible reaction patterns from the stricken banks’ counterparties in the worst-case scenario of the first round of simulations. These counterparties react according to two basic rules: they stop sending payments to the stricken bank either after some pre-determined time or after their exposure to that bank reaches a certain level. Here as well, the impact of the disruption is quantified in terms of the additional liquidity needs and in terms of the costs of overnight lending from the market and from the central bank.

However, even if participants’ reaction is as rapid as it can possibly be, it seems that participants cannot fully control their exposure to the failing bank. This conclusion leads to some thoughts about how the ability to control exposures could be quantified. With that objective in mind, the set of bilateral relationships between the stricken bank and all other banks is analysed as a ‘star’ network (a network with one central node). The degree of reciprocity, mainly in terms of the value of expected payments between the failing bank and each of its counterparties, is considered as one of the factors driving the ability of the stricken banks’ counterparties to control their exposures. The paper finally offers a measure of this reciprocity in the partial network created by the stricken bank and its counterparties in TOP.

3 McAndrews empirically estimated the slope of that function, both under “normal circumstances” and for the days following September 11, and found it to be generally quite steep (indicating a high level of coordination) and significantly less so in the days following the attacks (coordination broke down).
1. TOP and the Bank of Finland Payment and Settlement Simulator

1.1. The Dutch large-value payment system TOP

Analysing the effects of a disruption on a payment system with the help of the simulator necessitates, first, to replicate the functioning of the given payment system under normal conditions. This is done by calibrating the algorithms and parameters defining payment system design in the simulator, with a view to obtaining the best possible fit between the pattern of payment flows in the resulting benchmark simulation and in the historical settlement data.

TOP, the Dutch payment system, is relatively easy to approximate with the simulator: it is a simple RTGS system, with queues and priorities, but no liquidity-saving mechanisms. TOP is part of the pan-European payment system TARGET. In TARGET, it is the fourth largest TARGET component in terms of value (see table 1 for the main features of both TOP and TARGET). The system has 102 participants. Most banks located in the Netherlands participate directly in the system and formalised indirect participation does not exist. Some financial institutions may use correspondents for their TOP transactions and de facto participate indirectly in the system, but the importance of such indirect participation is considered to be relatively small. The large bank failing to send payments in the simulation accounts for about one fifth of TOP’s volume and value.

Table 1: Main features of TOP and TARGET

<table>
<thead>
<tr>
<th></th>
<th>TOP</th>
<th>TARGET</th>
</tr>
</thead>
<tbody>
<tr>
<td>Since</td>
<td>1997</td>
<td>1999</td>
</tr>
<tr>
<td>Number of direct participants</td>
<td>102</td>
<td>989</td>
</tr>
<tr>
<td>Number of transactions</td>
<td>4.98 million</td>
<td>69.2 million</td>
</tr>
<tr>
<td>Value of transactions</td>
<td>EUR 29 681 billion</td>
<td>EUR 443 993 billion</td>
</tr>
<tr>
<td>Average value per transaction</td>
<td>EUR 6.0 million</td>
<td>EUR 6.4 million</td>
</tr>
<tr>
<td>Open</td>
<td>7:00-18:00</td>
<td>7:00-18:00</td>
</tr>
<tr>
<td>Liquidity saving mechanisms</td>
<td>No</td>
<td>Yes, some national components</td>
</tr>
<tr>
<td>Intraday credit</td>
<td>Yes, based on pledged collateral</td>
<td>Yes, based on collateral (repo or pledge, depending on country)</td>
</tr>
</tbody>
</table>

Sources: ECB, “Payment and securities settlement systems in the EU and in the acceding countries”, March 2006 and BIS, “Payment and securities settlement systems in selected countries”, March 2006.

Simulations with historical payments data serve as a benchmark against which to assess the impact of disruptions. However, at least two features complicate the exercise. First, as TOP is part of TARGET; it is not possible, in TOP alone, to know the level of reserves and collateral facilities available to banks located in another country; I had to assume that these reserves were...

---

4 In other types of work, benchmark simulations have also been used to compare different system designs and to assess the efficiency of a system’s design compared to other possible designs.
unlimited. Therefore, part of the effect of a default (on other EU banks) cannot be estimated. Furthermore, at the time at which data was gathered and simulations carried out, it was not possible to know the identity of the European counterparties. The only information available concerned the country from which or to which payments were made via TARGET. It was not, therefore, possible to distinguish payments coming from branches of the failing bank. This is mainly of importance for simulations implying a reaction from the stricken bank’s counterparties, as counterparties belonging to the same group are likely to act differently from others. Second, some accounts in TOP, used for the settlement of ancillary systems, do not have any liquidity facilities, as their only purpose is to “pass on” payments (it is an administrative function). In practice, if they are not “fed”, they will not pay out. In the simulations, the effects of a default on ancillary systems have been ignored. In practice, they are likely to be limited: payments to these accounts would not stop following an operational disruption at one of the systems’ participants, as the failing participant’s central bank account would be automatically debited, allowing the settlement of ancillary systems to continue.

1.2. Data used for the simulations

The data used for the simulations concerns the month of June 2004 (22 days). Although this may have been motivated by practical reasons – by the time I started working with the simulator, these were the most recent figures, except for the summer months, which are atypical, with a lower number and value of transactions – the month of June happened to display a volume and value of transactions that were close to the daily average for that year, with a slightly higher number of transactions (24,4 thousand a day, compared to an average of 23,5 thousand) and a slightly lower value (EUR 160 billion a day compared to 164 billion). June offered the additional advantage, in that year, not to have any special days or holidays. Although big banks tend to operate round the clock, many smaller European banks are closed on bank holidays, which leads to a reduced activity in TOP. June 2004 did not have any American bank holidays either. U.S. bank holidays tend to have a significant impact on payments in the Netherlands, as most Dutch banks manage their liquidity by Euro-Dollar swaps. This impact is mainly felt on the day of the U.S. bank holiday itself, with a decrease of 35% and 23% in the value and number of transactions respectively. The day after the U.S. bank holidays, both the value and number of transactions increase by 23% and 17% (van Oord & Lin, 2005).

---

5 The use of the terms “default”, “operational default”, “failure”, “disruption” etc. in this paper refer to the bank’s inability to send payment messages. They do not, in any case, refer to the institution’s possible insolvency. An insolvency of a major bank would indeed not come as a surprise, and its counterparties would have taken measures in advance of the default. The aim of this exercise is to assess the impact of an unexpected failure to send payments, and to see whether
1.3. Benchmark simulations and three scenarios

A total of 484 simulations were carried out: 44 for the benchmark, as well as the higher and lower bounds of liquidity (this will be explained later in this paragraph); 264 for scenario 1, with default occurring at each hour of the day; 110 for scenario 2, with the non-failing participants reacting within 10, 30 minutes, 1, 2 or 4 hours; and finally 66 for scenario 3, with participants reacting when their individual exposure reached three different thresholds.

Scenario 1 explores the failure to pay of a major TOP participant, at different hours of the day. Payments from that participant to other banks are removed from the data. The stricken bank does not recover within the day, which means that the disruption lasts for the remainder of the day. Although it is possible to guess that, absent any reaction from other participants, the effect of a disruption will be stronger the earlier it occurs, this exercise makes it possible to see how the impact of a disruption varies within the day. Scenarios 2 and 3 introduce a reaction from other payment system participants. In scenario 2, once banks are aware of the disruption, they cease to pay to the stricken bank, in order to prevent its accounts to act as a liquidity sink. That way, they are able to reduce the impact of the operational disruption on the system. A reaction time of 10, 30 minutes, 1, 2 or 4 hours is simulated. In scenario 3, counterparties’ reaction is individually-tailored, as they stop sending payments once their exposure to the failing bank (the net amounts paid) reach a certain percentage of their regulatory capital, namely 25%, 10% or 5%. See table 2 for a summary of these different scenarios.

Table 2: Summary of the different scenarios

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Failure major participant</th>
<th>Reaction counterparties</th>
<th>Variations on reaction</th>
<th>Number of simulations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Benchmark</td>
<td>No</td>
<td></td>
<td></td>
<td>22 days</td>
</tr>
<tr>
<td>Scenario 1</td>
<td>Yes, at different times of day</td>
<td>No</td>
<td>-</td>
<td>22 days * 12 times = 264</td>
</tr>
<tr>
<td>Scenario 2</td>
<td>Yes, at 12:00 (worst-case from scenario 1)</td>
<td>Yes, timed stop sending</td>
<td>10 minutes 30 minutes 1 hour 2 hours 4 hours</td>
<td>22 days * 5 variations = 110</td>
</tr>
<tr>
<td>Scenario 3</td>
<td>Yes, at 12:00 (worst-case from scenario 1)</td>
<td>Yes, stop sending after exposure reaches % of regulatory capital</td>
<td>25% 10% 5%</td>
<td>22 days * 3 variations = 66</td>
</tr>
</tbody>
</table>

The benchmark simulation, carried out with historical data, can also be run with a view to assess the upper and lower bounds of liquidity that banks will need in order to settle all their payments (see Leinonen & Soramäki, in Leinonen ed., 2005, for a more thorough explanation of the

---

6 The 10 minutes interval corresponds to the assumption made by Bedford et al. (2005) for the UK payment system CHAPS. I moved away from the 10 minutes in order to see whether a very short reaction time really mattered. As from 30 minutes the intervals represent a doubling of the reaction time.

7 These percentages are discussed in chapter 3.3.
concept of liquidity bounds). These bounds of liquidity can be computed by allowing banks, starting the day without any balances on their accounts, to draw on unlimited amounts of intraday credit. All payments will thus settle immediately and queues will be empty. The sum of all the lowest (positive and negative) balances on banks’ accounts during the day is the upper bound of liquidity. The sum of all negative net balances is the lower bound of liquidity; it would be sufficient to allow participants to settle all their obligations in a DNS system.

In June 2004 in TOP, the upper bound of liquidity, which would have been needed to allow all payments to settle immediately, amounted to EUR 50.18 billions; and the lower bound of liquidity, which would have been enough for all payments to settle at end-of-day, to EUR 10.78 billion. Note that this includes payments that are not originated or received by banks; these payments may be sent from accounts that are not bound by a maximum overdraft facility, eg the accounts of the central bank itself. The liquidity available de facto in the system to banks for the same month was EUR 6.7 billion in start-of-day balances and EUR 53.85 billion in deposited collateral (minus haircuts), ie a total of more than EUR 60 billion. This is more than the upper bound of liquidity, which suggests that banks’ endowment in liquid funds and collateral is very comfortable for the payments that are scheduled to be settled on the days of that month.

However, the distribution of this liquidity among banks is also important for the system to function well. Indeed, some banks may be ‘richer’ in terms of deposited funds and collateral, while others, for various reasons, may have chosen to economise on their reserves at the central bank. Given the fact that these reserves are costly and that collateral has an opportunity cost, banks would, in theory, be tempted to reduce their liquid holdings at the central bank down to the level at which they will be able to settle all their payments within an acceptable time. As we will see later, however, other factors play a role in the distribution of payment system participants’ liquidity. In the simulations, this distribution is important, as it influences the queuing – and eventually the rejection - of payments. It is implicitly assumed that there is no interbank market where banks with excess reserves could trade with banks lacking some necessary funds.

A number of papers vary the level of available intra day credit between the upper and lower bounds, and estimate how much this available liquidity influences the impact of a disruption (Bedford et al., 2005, and others). For this paper, however, all simulations are based on the real amounts of reserves and collateral at the central bank, in order to assess the potential effect of a disruption under the conditions that held in June 2004. As said previously, these conditions were relatively comfortable, which suggests that a disruption occurring under more stressed liquidity conditions would undoubtedly have a stronger impact on the system, assuming the distribution of reserves and collateral is kept constant.

---

8 These are theoretical lower and higher bounds, as the system is a “pure” RTGS system. No consideration is taken of time-critical payments which absolutely need to be settled at certain times during the day.

9 By banks I mean banks located in the Netherlands, which need to rely on either balances on their accounts at De Nederlandsche Bank or on collateral deposited at the central bank. This does not include banks located in other European countries which send payments to TOP through TARGET.
2. The timing of an operational failure

2.1. Scenario 1: varying the timing of an operational failure
The number and value of unsettled payments depend on the time at which the stricken bank ceases to send payments. The first-round effects of such an operational disruption stem directly from the payments not sent by that bank, while the secondary-round effects can be measured as the number and value of payments that other banks will not be able to send, due to a lack of funds caused by the fact that none of the payments from the stricken bank were received. Obviously, in the absence of any reaction from the failing bank’s counterparties, the secondary-round effects will be stronger the earlier the default.

Scenario 1 assumes that one of the major payment system participants in the Dutch interbank payment system TOP is unable to send payment orders as from some time during the day. This time is varied between 07:00 and 18:00 (respectively the opening and the closing time of the TARGET system). Default lasts for the remaining part of the day. If the operational inability to send payments were to be solved within a short period of time, the payment system would only be affected during the disruption; queues would arise for the duration of the disruption and disappear rapidly after the failing participant resumes its operations. At the end of the day no unsettled payments would remain. Unsettled payments, which provide information about the secondary-round effects of a disruption and about banks’ positions on their central bank accounts will be used further in the paper in order to assess the strength of the disruption and participants’ ability to weather a counterparty’s failure to pay.

How likely is an important payment system participant to experience an operational disruption lasting for the whole day? Not very likely, but the impact of such a disruption would certainly be significant. In 2003 one of the major European banks, accounting for slightly less than 10% of the overall TARGET volume, failed to send payments between 11.30 and 19.30 on a single day. Its counterparties were expecting to receive EUR 40 billion on that day. In the simulations, the stricken bank accounted for slightly more than 20% of all TOP transactions in volume, and slightly less than 20% in value. If a similar event were to occur, payment system participants would most likely reroute payments from TOP to another system (eg EURO 1, the system of the Euro Banking Association, which settles payments on a net basis, and therefore needs less liquidity to operate) or review the organisation of their payment flows for the day concerned, in order to avoid likely losses resulting from a lack of incoming funds. These more

---

10 TOP itself is, as many other systems, also open during the closing time of TARGET for the night-settlement of securities transactions. As the night-settlements only concern certain types of transactions, with a relatively low value, the possibility of a disruption occurring at these times of the day has not been simulated in this exercise.

11 It might, however, be interesting to vary the duration of such a disruption in order to gain insight in the patterns of queues at different times of the day and for different disruption durations.
complex behavioural reactions have not been rendered in the simulations – the only reaction simulated concerned retaining payments to the stricken bank, as we will see in chapter 4.

2.2. Second-round effects

2.2.1. Unsettled payments

The aim of this first exercise is to see how the timing of an operational disruption impacts the effect of such a disruption on the payment system. As expected, the secondary-round effects of a payment failure are strongest the earlier the failure. On average, over the month of June 2004, were a major Dutch bank to retain all its payments as from the opening time of the system, about 16 payments from other participants would remain unsettled as a consequence of liquidity shortages. This would represent more than EUR 1 billion (see figure 1, which shows all unsettled payments in the secondary round, excluding payments from one of the accounts of the failing bank). The number and value of unsettled payments first decreases slowly until a failure time of around 09:00, after which the decline steepens. The impact of the failure is close to nil as from one hour before the closing time of the system. Between 09:00 and 17:00, the volume and the value of stranded payments diverge slightly, as the number of unsettled payments declines faster than the value, which indicates that the stricken bank’s counterparties continue to await important funding for payments that they intend to send relatively laterly during the day.

Figure 1

![Time-dependency of impact of disruption at a major participant (excluding failing participant’s payments)](image)

The strength of the secondary-round effects can substantially differ from one day to another. With default occurring at the opening time of the TARGET system, it can vary between 3 and 45 unsettled payments, with a value comprised between about EUR 40 million and slightly
less than EUR 2.2 billion (see table 3. The last column of the table, showing the delay indicator, will be discussed later). On the worst day, the effect remains significant until 16:00, with 7 unsettled payments and a value of EUR 1.1 billion – and the only payment unsettled with a disruption at 17:00 still amounts for more than EUR 1 billion.

Table 3: Secondary-round effects of an operational disruption, depending on the time of the disruption (excluding failing participant’s payments)

<table>
<thead>
<tr>
<th>Time of disruption</th>
<th>Number unsettled</th>
<th>Value unsettled (EUR millions)</th>
<th>Delay ind.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Av</td>
<td>Min</td>
<td>Max</td>
</tr>
<tr>
<td>7</td>
<td>16.62</td>
<td>3.00</td>
<td>45.00</td>
</tr>
<tr>
<td>8</td>
<td>16.19</td>
<td>3.00</td>
<td>45.00</td>
</tr>
<tr>
<td>9</td>
<td>15.43</td>
<td>3.00</td>
<td>45.00</td>
</tr>
<tr>
<td>10</td>
<td>12.05</td>
<td>3.00</td>
<td>37.00</td>
</tr>
<tr>
<td>11</td>
<td>10.24</td>
<td>3.00</td>
<td>37.00</td>
</tr>
<tr>
<td>12</td>
<td>5.86</td>
<td>1.00</td>
<td>20.00</td>
</tr>
<tr>
<td>13</td>
<td>4.52</td>
<td>1.00</td>
<td>16.00</td>
</tr>
<tr>
<td>14</td>
<td>3.57</td>
<td>0.00</td>
<td>14.00</td>
</tr>
<tr>
<td>15</td>
<td>1.33</td>
<td>0.00</td>
<td>7.00</td>
</tr>
<tr>
<td>16</td>
<td>1.21</td>
<td>0.00</td>
<td>7.00</td>
</tr>
<tr>
<td>17</td>
<td>0.11</td>
<td>0.00</td>
<td>1.00</td>
</tr>
<tr>
<td>18</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
</tbody>
</table>

2.1.2 Settlement delays

When payments cannot be settled due to a lack of funds, they remain queued. The longer they remain in the queue, the longer settlement is delayed. From a system perspective, the higher the number (and value) of payments that cannot be settled immediately, and the longer their stay in the queue, the higher the settlement delay. RTGS systems with ample available liquidity and no queuing will show no settlement delay at all. Introducing the possibility of queuing, and reducing the funds available increases settlement delay, with DNS systems that settle all payments at the end of the day exhibiting the highest settlement delay. Such an aggregate settlement delay can be expressed by way of a settlement delay indicator:

\[ \rho = \frac{\sum_{t=1}^{T} Q_t}{\sum_{t=1}^{T} \sum_{i=1}^{V} V_i} \]

where, in the numerator, the value of payments in the queue is summed over each minute of the day and, in the denominator, the cumulative value of all sent payments is summed over each
minute of the day. With this metric, it is possible to calculate the settlement delay relative to a
certain portion of time (up to a whole day), for some or all payment system participants. It is
comprised between 0 and 1, with 0 indicating that all payments are settled immediately (as in a
pure RTGS system with no queues and ample liquidity), and 1 that all remain delayed until the
end of the period under consideration (this will typically be exhibited by a DNS system).

Settlement delay means settlement risk. As long as queued payments have not been
settled with finality, the recipient remains exposed to credit risk (which would materialise if the
sender proves unable to fund its account), to liquidity risk (if the payment arrives later than the
recipient would have expected given its own funding needs) and to operational risk (a disruption
at either the sending institution or the system operator can prevent the payment from being
settled in due time). In our exercise, the settlement delay indicator drops between 07:00 and
12:00: the later the disruption, the faster payments are settled, the less settlement risk remains in
the system. The indicator then remains fairly constant for all afternoon failure times, only dropping
at the end of the day to a level similar to that in the benchmark simulation (see table 3 above).

2.1.3 The liquidity sink effect
When a bank stops to send payments, its account has the potential to develop into a liquidity
sink, as it continues to receive incoming payments from its counterparties. The liquidity that
remains on the failing participant’s account cannot be recycled and fulfill its purpose of smoothing
the functioning of the system. If this process continues for a certain time, the failing participant’s
counterparties can be short of funds on their accounts, and appear unable to settle their own
outgoing payments, as we have seen here above. In June 2004, counterparties to the failing bank
were expected to receive, on average, about 3.4 thousand payments and EUR 30 billion on each
day. Figure 2 shows how strong the liquidity sink effect would be, depending on the time of
failure. Obviously, the amplitude of the effect declines when failure occurs later: the earlier it is,
the more liquidity gets trapped on the stricken bank’s accounts. When operational failure occurs
at 07:00, the value trapped on these accounts reaches about 5.5 times its normal end-of-day
value. The decline is smooth and regular, which suggests the stricken bank’s accounts receive
liquidity fairly regularly during the day; this continues until the liquidity on the failing bank’s
account at the end of the day approximates its normal end-of-day-value (which would be with
failure occurring at about 17:00).

The net credit position of the failing bank (figure 3) shows, at each hour of the day, by how
much the bank expects its position to change during the day, as it gives the difference between
due payments and expected payments. At the opening time of the system, due and expected
payments are reasonably aligned. However, as that bank tends to be an early payer, for most
part of the day after 08:00 expected payments are much higher than payments due. Only after
15:00 does the gap decrease and disappear. This suggests that counterparties should be able to
control their exposure by delaying payments relatively well for failure times from 08:00 onwards. With failure occurring at 07:00, this may still seem true for the aggregate positions, as the stricken bank still expects more than it intends to send, but the distribution of positions among banks would certainly matter. Indeed, one bank might be expecting payments from the stricken bank, while it barely has payments due; another bank might find itself in the opposite situation. In that example, the first bank’s ability to control its exposure would be limited, but not that of the second bank. We will explore this further in chapter 5.

**Figure 2**

Value on the failing bank’s accounts depending on time of disruption

![Graph showing value on the failing bank’s accounts](chart2)

**Figure 3**

Net credit position of the failing bank

(payments expected - payments due as from indicated time, in % of start-of-day value)

![Graph showing net credit position](chart3)
2.2 The costs of an operational failure

Central banks are interested in ensuring the smooth functioning of interbank payment systems, and therefore seek to prevent systemic crises; such crises would bear the highest cost for the financial system as a whole. For payment system participants, however, even a disruption with a fairly minor impact can lead to real costs. And, provided the disruption lasts for the whole day, these costs would be highly visible, as participants would be forced to fund their additional intraday liquidity needs by borrowing overnight.

The trade-offs between settlement delay and intraday liquidity costs have often been discussed; they have been used to explain the choice participants face when deciding to send their payments relatively early or relatively late during the day (Bech & Garratt, 2003). When choosing to delay their outgoing payments, system participants risk their reputation, as their counterparties need to wait for their payments. When they do pay early during the day (ie before having received any payments themselves), they can be forced to incur liquidity costs, related to the funding of these payments. At the end of the day, this problem would be posed with even more stringency. Indeed, continuing to delay payments to the next day would possibly lead, not only to even higher reputation costs, but also, for part of these payments related to the repayment of money market deals, to costs caused by the de facto roll-over of these deals. Given the high costs of delaying payments to the next day, I assumed for this exercise that payment system participants would fund their outgoing payments up to the maximum during the day with central bank intraday credit. At the end of the day, as the disruption would not be resolved, remaining negative positions would need to be funded overnight, either on the overnight interbank market, or at the central bank, making use of what is called the marginal lending facility.

It is possible to see how much additional liquidity banks incorporated in the Netherlands would need at the end of the day, compared to the benchmark simulation. This is equal to the difference between the sum of all negative values on central bank accounts at the end of the day and the use of the marginal lending facility in the month of June 2004, which averaged EUR 159.76 thousand. For the same month, the interest rate of the marginal lending facility was 3% p.a., while the EONIA rate, the market rate, amounted to 2.03% p.a. Although I have calculated the costs of additional liquidity in accordance with these rates, a major disruption could influence the market rate and push it upward, as demand for overnight liquidity would significantly rise compared to the available liquidity in the market. However, such an effect would be dampened by the width of the Euro market, as the liquidity needs of Dutch banks might be fulfilled by other Euro area institutions. This effect can be expected to be stronger in a small closed market, when all participants need to weather the same disruption in their main payment system.

Table 4 and figure 4 show the average value of the additional liquidity required by non-defaulting participants, as well as the costs of overnight funding, according to the time at which the disruption occurs. As expected, the later the disruption, the lower the additional liquidity.
needed by the stricken bank’s counterparties, and the lower the associated costs of overnight borrowing. On average, the maximum level of intraday borrowing that would remain outstanding at the end of the day would amount to more than EUR 1.5 billion. The maximum cost of overnight funding would be about EUR 200 thousand for the marginal lending facility, and nearly EUR 140 thousand for interbank lending. At 18:00 these values would be down to a few thousand euros, which are minor costs if borne by the whole banking sector.

Table 4: Additional liquidity and related costs, depending on the time of a disruption

<table>
<thead>
<tr>
<th>Time of disruption</th>
<th>Value (EUR million)</th>
<th>Costs of overnight overdraft (EUR thousands)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>EONIA</td>
<td>marginal lending</td>
</tr>
<tr>
<td>07:00</td>
<td>1,687.98</td>
<td>137.06</td>
</tr>
<tr>
<td>08:00</td>
<td>1,617.98</td>
<td>131.38</td>
</tr>
<tr>
<td>09:00</td>
<td>1,339.25</td>
<td>108.75</td>
</tr>
<tr>
<td>10:00</td>
<td>1,374.76</td>
<td>111.63</td>
</tr>
<tr>
<td>11:00</td>
<td>1,195.81</td>
<td>97.10</td>
</tr>
<tr>
<td>12:00</td>
<td>851.86</td>
<td>69.17</td>
</tr>
<tr>
<td>13:00</td>
<td>747.98</td>
<td>60.74</td>
</tr>
<tr>
<td>14:00</td>
<td>499.30</td>
<td>40.54</td>
</tr>
<tr>
<td>15:00</td>
<td>125.98</td>
<td>10.23</td>
</tr>
<tr>
<td>16:00</td>
<td>44.11</td>
<td>3.58</td>
</tr>
<tr>
<td>17:00</td>
<td>26.62</td>
<td>2.16</td>
</tr>
<tr>
<td>18:00</td>
<td>20.65</td>
<td>1.68</td>
</tr>
</tbody>
</table>

Interestingly, the value of additional borrowing does not decline in a consequent way as the disruption is shifted to a later time. Between 09:00 and 10:00 it increases slightly, and with it the associated costs. Why does this happen? It would not happen if participants would be allowed to incur unlimited overdrafts on their central bank accounts. The possible overdrafts are, however, limited, which leads to the queuing – and ultimately to the rejection – of payments for which there is insufficient liquidity. Queued payments do not draw upon the liquidity available to banks. When the stricken bank defaults at 09:00, one of its counterparties lacks the funds to successfully settle one of its big outgoing payments. This payment remains in the queue until the end of the day; it is one of the unsettled payments that we report under the secondary-round effects. However, when
the stricken bank defaults one hour later, the payments that it could make in this additional hour to the counterparty concerned is sufficient to fund part of this big outgoing payment. The payment is thus made and settled with finality in TOP. But the settlement of that payment forces the sending participant deeply into the red. At the end of the day, this negative position shows up in the additional value of liquidity that needs to be borrowed overnight. Although this is a fairly exceptional situation, it nevertheless leads to the conclusion that, although in general the timing of a disruption seems to explain the relative level of its impact and costs, certain particular payment patterns may engender a counterintuitive relationship.

**Figure 4**

**Additional liquidity needed at end-of day, depending on time of disruption**

<table>
<thead>
<tr>
<th>Time (EUR million)</th>
<th>07:00</th>
<th>08:00</th>
<th>09:00</th>
<th>10:00</th>
<th>11:00</th>
<th>12:00</th>
<th>13:00</th>
<th>14:00</th>
<th>15:00</th>
<th>16:00</th>
<th>17:00</th>
<th>18:00</th>
</tr>
</thead>
<tbody>
<tr>
<td>EUR million</td>
<td>0</td>
<td>200</td>
<td>400</td>
<td>600</td>
<td>800</td>
<td>1,000</td>
<td>1,200</td>
<td>1,400</td>
<td>1,600</td>
<td>1,800</td>
<td>1,800</td>
<td>1,800</td>
</tr>
<tr>
<td>07:00</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>08:00</td>
<td>200</td>
<td>200</td>
<td>200</td>
<td>200</td>
<td>200</td>
<td>200</td>
<td>200</td>
<td>200</td>
<td>200</td>
<td>200</td>
<td>200</td>
<td>200</td>
</tr>
<tr>
<td>09:00</td>
<td>400</td>
<td>400</td>
<td>400</td>
<td>400</td>
<td>400</td>
<td>400</td>
<td>400</td>
<td>400</td>
<td>400</td>
<td>400</td>
<td>400</td>
<td>400</td>
</tr>
<tr>
<td>10:00</td>
<td>600</td>
<td>600</td>
<td>600</td>
<td>600</td>
<td>600</td>
<td>600</td>
<td>600</td>
<td>600</td>
<td>600</td>
<td>600</td>
<td>600</td>
<td>600</td>
</tr>
<tr>
<td>11:00</td>
<td>800</td>
<td>800</td>
<td>800</td>
<td>800</td>
<td>800</td>
<td>800</td>
<td>800</td>
<td>800</td>
<td>800</td>
<td>800</td>
<td>800</td>
<td>800</td>
</tr>
<tr>
<td>12:00</td>
<td>1,000</td>
<td>1,000</td>
<td>1,000</td>
<td>1,000</td>
<td>1,000</td>
<td>1,000</td>
<td>1,000</td>
<td>1,000</td>
<td>1,000</td>
<td>1,000</td>
<td>1,000</td>
<td>1,000</td>
</tr>
<tr>
<td>13:00</td>
<td>1,200</td>
<td>1,200</td>
<td>1,200</td>
<td>1,200</td>
<td>1,200</td>
<td>1,200</td>
<td>1,200</td>
<td>1,200</td>
<td>1,200</td>
<td>1,200</td>
<td>1,200</td>
<td>1,200</td>
</tr>
<tr>
<td>14:00</td>
<td>1,400</td>
<td>1,400</td>
<td>1,400</td>
<td>1,400</td>
<td>1,400</td>
<td>1,400</td>
<td>1,400</td>
<td>1,400</td>
<td>1,400</td>
<td>1,400</td>
<td>1,400</td>
<td>1,400</td>
</tr>
<tr>
<td>15:00</td>
<td>1,600</td>
<td>1,600</td>
<td>1,600</td>
<td>1,600</td>
<td>1,600</td>
<td>1,600</td>
<td>1,600</td>
<td>1,600</td>
<td>1,600</td>
<td>1,600</td>
<td>1,600</td>
<td>1,600</td>
</tr>
<tr>
<td>16:00</td>
<td>1,800</td>
<td>1,800</td>
<td>1,800</td>
<td>1,800</td>
<td>1,800</td>
<td>1,800</td>
<td>1,800</td>
<td>1,800</td>
<td>1,800</td>
<td>1,800</td>
<td>1,800</td>
<td>1,800</td>
</tr>
<tr>
<td>17:00</td>
<td>1,800</td>
<td>1,800</td>
<td>1,800</td>
<td>1,800</td>
<td>1,800</td>
<td>1,800</td>
<td>1,800</td>
<td>1,800</td>
<td>1,800</td>
<td>1,800</td>
<td>1,800</td>
<td>1,800</td>
</tr>
<tr>
<td>18:00</td>
<td>1,800</td>
<td>1,800</td>
<td>1,800</td>
<td>1,800</td>
<td>1,800</td>
<td>1,800</td>
<td>1,800</td>
<td>1,800</td>
<td>1,800</td>
<td>1,800</td>
<td>1,800</td>
<td>1,800</td>
</tr>
</tbody>
</table>
3 The influence of bank behaviour on the impact of a disruption

3.1 Participants’ reaction to a failure

Little is known about how payment system participants react to an operational disruption at one of their counterparties. Part of this stems from the fact that banks’ reaction is related to the intraday counterparty credit limits that they have built into their internal systems. These limits are not communicated to central banks. Payment system operators’ interactions with market participants suggest that the limits are hard for front office dealers’ trading – but less so with regard to payment initiated by clients, and also that banks in some countries are more stringent than in other countries. But none of these affirmations is based on empirical research, although some central banks keep an “operational events database” (eg the Bank of Canada), including information on participants’ reactions. It is also difficult to generalise banks’ behaviour, because disruptions have different causes and effects, and also because behaviour can change, and in the presence of thin knowledge, information may be contradictory.

One of the most interesting systematic studies of payment system participants’ behaviour following a disruption, by D. Conover & D. Amanuel (2005), has been presented at the bank of Finland’s simulator seminar in 2005. In that work, the authors have analysed real failures in the U.S.’ interbank payment system Fedwire, defined as abnormally long periods of time between payments from any participant. It seems that, during such an event, the value of payments received by the outage bank is down by 15.3%; as to the number of payments, it increases by 1.9%, which is probably within the bounds of normality. This is conform anecdotal evidence, which suggests that bigger payments tend to be delayed first, as banks see their exposure to the “stricken bank” increase. Payments also seem to stay longer in the queue, as the average delay of payments to the outage bank increases by 3 minutes. What this work does not say, however, is how long participants tend to wait before delaying their payments to the outage bank, nor how participants’ reaction changes over time, as the disruption lengthens. Furthermore, it would be extremely interesting to have some additional insights in the distribution of the length of disruptions, although the presentation does give some indications about the relation between the time gap between two payments from an outage bank (which defines a disruption) and the length of the outage, and identifies the number of outages corresponding to four pre-defined gap lengths.

With regard to TOP, payment system operators observe that, in the face of a disruption, foreign banks are expected to be the first to stop sending payments. Transactions to the failing bank can be cancelled, or, in some situations, re-routed via EURO1, in which they would be

---

12 This excludes the period between September 11-14, 2001, ie an exogenous event, clearly not related to operational disruptions at specific participants.
settled net later during the day, and therefore impede less on the sender’s liquidity. In our example, given the reputation of the failing bank and its habit of sending payments early in the day, and therefore of supplying other banks with liquidity, domestic banks would probably not rapidly cancel or re-route any transactions. In general, in case of a morning failure, most banks might be expected to delay any reaction until early in the afternoon. We will see later in this paper that differentiating between domestic and foreign banks can also be justified by the fact that all banks wish to limit their relative exposure to the failing institution.

The following exercises assume two possible triggers for counterparties’ reaction: either the length of the disruption or their individual exposure. In scenario 2, payment system participants stop sending payments to the failing institution after 10, 30 minutes, 1, 2 or 4 hours. In scenario 3, each payment system participant stops to send payments when its exposure to the stricken bank reaches 5%, 10% or 25% of its regulatory capital. These simulations are based on the worst-case situation under scenario 1, where the operational disruption occurs at 7:00 and lasts for the remainder of the day.

3.2 Timed stop-sending

Although such information is not available for TOP, there is some observed evidence for other payment systems that participants would react within a certain timeframe following an operation disruption at one of their counterparties. A paper by Bedford et al. (2005) is based on such evidence for CHAPS, the UK payment system. Because their work incorporates the effect of counterparties’ timed reaction, the worst-case moment for a disruption to occur is not at the beginning of the day, but at a moment at which a so-called “virtual credit balance” (for each minute, the sum of the stricken bank’s reserves and its expected payments in the next ten minutes) is at its highest, and a lot of payments still need to be made. Such a virtual credit balance gives an indication of the potential for the stricken bank to develop into a liquidity sink. At the moment at which the operational disruption is meant to occur, the stricken bank is due to send or receive 46,000 payments, with a value of GBP 45.7 billion. As a share of the total value processed by CHAPS, this is in the same order of magnitude as the bank considered in this paper, although the UK bank accounts for a much higher relative volume of payments. Ten minutes after the disruption, all payments to the failing participant cease. Varying the quantity of liquidity available, Bedford et al. find that the secondary-round effects of a disruption are only significant at levels of liquidity which are much below the liquidity available in the system. However, the impact of a disruption at actual levels of liquidity is not tested. Could it be that, although the total amount of liquidity exceeds the upper bound, the distribution of that liquidity may still lead to some local liquidity shortages? Nor is the effect of participant’s behaviour on the impact of a disruption estimated.
Scenario 2 is based partly on the work by Bedford et al., as I simulate that participants stop sending payments to the stricken bank some time after the disruption occurs. However, it focuses on the unanswered questions above: by how much can participants reduce the impact of a disruption, and is this a function of the reaction time, given actual liquidity levels? The stricken banks’ counter-parties stop to send payments 10, 30 minutes, 1, 2 or 4 hours after the disruption. Note that such a reaction pattern assumes perfect information in the market about the occurrence and timing of a disruption. Even if, de facto, not all banks react at the same time, as they have not all planned to send payments to the stricken bank within the reaction period, they would stop if they had planned to send such payments within that period.

When participants do not react following a disruption, and continue to send payments to the stricken bank, about 16 additional payments get stalled due to liquidity shortages. When participants do react, the secondary-round effect decrease to an average of about 7 rejected payments at the end of the day, for a reaction time of 4 hours (a reaction time in concordance with payment system operators’ observations for TOP). It thus seems that even a relatively long reaction time would be effective in limiting the secondary-round effects of a failure. Counter-intuitively, reducing the reaction time from 4 to 2 hours and less has a much more limited effect. Indeed, the secondary-round effects are even quasi-similar for a reaction time of between 10 minutes and 2 hours, with about 5 stalled payments. It seems that even an immediate reaction would not totally eliminate the secondary-round effects of an operational disruption.

Results

Although a fast reaction does not eliminate secondary-round effects entirely, it appears to be quite efficient in reducing the costs of additional liquidity for other payment system participants. Where such costs reach about EUR 202 thousand and slightly less than EUR 140 thousand for respectively central bank and money market lending (for an additional EUR 1.7 billion) in the worst-case simulation, they can be reduced to less than half these amount if banks react within ten minutes (see table 5). But no matter how fast they react, banks appear unable to fully eliminate the need to rely on additional liquidity following a disruption; even when they stop sending payments to the failing institution within 10 minutes after a disruption, they still incur costs of about EUR 100 thousand and EUR 67 thousand for respectively central bank and money market lending. For a reaction time of 10 to 30 minutes, the costs of additional liquidity increase, as expected, while the unsettled payments in the secondary-round effects remain constant: for these levels, banks can apparently continue to incur higher overdrafts; their payments are only rejected for a longer reaction time. But between 30 minutes and 1 hour, the additional liquidity needed is reduced, as more payments remain unsettled. The difference between days is relatively important (see table in annex). On the day with the lowest second-round effects, all

---

13  “Anecdotal evidence from CHAPS Sterling suggests that the time-lag between an individual bank experiencing an
payments settle as from a reaction of 2 hours, while on the day with the highest effects, even a reaction time of ten minutes does not reduce the number of unsettled payments below 12.

Table 5: The impact of timed stop-sending on secondary round effects and liquidity costs

<table>
<thead>
<tr>
<th>Time before participants react</th>
<th>Unsettled payments</th>
<th>Additional liquidity</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Number</td>
<td>Value (EUR million)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10 min</td>
<td>4.86</td>
<td>303.37</td>
</tr>
<tr>
<td>30 min</td>
<td>4.86</td>
<td>303.37</td>
</tr>
<tr>
<td>1 hour</td>
<td>4.90</td>
<td>319.23</td>
</tr>
<tr>
<td>2 hours</td>
<td>5.00</td>
<td>329.65</td>
</tr>
<tr>
<td>4 hours</td>
<td>7.43</td>
<td>510.46</td>
</tr>
<tr>
<td>Unlimited (&gt; 11 hours)</td>
<td>16.62</td>
<td>1075.18</td>
</tr>
</tbody>
</table>

On a certain day, an interesting situation can be witnessed, as participants’ reaction lead to an increased number of unsettled payments. On that day, the payments sent by the defaulting participant before disruption, but waiting in the queue for incoming payments, can remain unsettled as these expected incoming payments do not arrive.

Compared to the simulations carried out under the first scenario, settlement delays increase when the failing bank’s counterparties stop sending payments to that bank. The indicator has a value of 0.14 for all reaction times, except for the longest (4 hours), for which it takes the value 0.12. Apparently, a longer reaction time reduces settlement delay. The delay indicators for the third scenario confirm these findings. For all reaction options, it takes the value 0.16. As counterparties’ reactions are expected to limit the liquidity strains in the system, these results are somewhat surprising. But they may be explained by the way the indicator is calculated. Indeed, the number (and value) of payments decreases very steeply when participants react, while the values in the queue are much less affected. For example, the value of queued payments declines by 21% in the 5% exposure control scenario, while the overall value of all sent payments is reduced by 32% as a consequence of participants’ reaction.

In view of the anecdotal evidence with regard to the behaviour of TOP participants described above, it would be interesting to differentiate between participants (for example, letting only a fraction of participants react, or different types of participants react at different time intervals, with smaller or foreign banks reacting immediately, and bigger or domestic banks being less rapid in their reaction). It would also be possible to simulate a timed stop-sending where the operational failure and the flow of payments to that bank slowing significantly is typically of the order of ten minutes*.
time lag is calculated not as from the occurrence of the operational disruption – which assumes perfect information in the market about the operational health of any given bank – but as from the moment at which each bank expects to receive payments from the failing institution – which assumes that banks are only informed about an operational failure when they miss payments due to them. However, instead of differentiating in terms of participants’ reaction time, I have worked on another scenario, where differentiation was a natural consequence of the assumption: participants react when their exposure reaches a certain threshold.

3.3 Exposure control
The hypothesis underpinning scenario 3 is that once the stricken bank’s counterparties’ liquidity is getting scarcer, they block outgoing payments to that bank. Doing that, they rely on their internal limits. Their real limits are not known. For this exercise, they have been approximated by setting bilateral credit limits vis-à-vis the failing bank at 25%, 10% and 5% of counterparties’ regulatory capital. 25% of regulatory capital corresponds to the maximum allowed exposure toward any individual counterparty. Under what is called the “large positions rule” (grote postenregeling), banks are not allowed to exceed this threshold. Although this rule applies for overnight and longer exposures, I have used it as a maximum benchmark for intraday exposures in this exercise.

Setting a higher threshold does not seem to make sense in TOP.

In its 1986 paper concerning bank failures in CHIPS, Humphrey finds that banks' exposures intraday, stemming from the failure of one payment system participant, are sufficient to endanger their continuity under an unprotected netting arrangement. For banks to have an overall exposure to such an event exceeding their own capital, their individual exposure to the failing participant would need to be relatively high – almost certainly exceeding the 25% threshold chosen for this exercise – or the secondary-round effects would need to have a substantial impact. We have seen here above that, in our example, the secondary-round effects are relatively limited. As to individual intraday exposures to the failing institution exceeding the 25% threshold, they are quite rare, with some 8 banks sending payments to the stricken bank in excess of 25% of their regulatory capital. Banks are furthermore obliged to report any (overnight) exposure exceeding 3% of their regulatory capital (see Lelyveld & Liedorp, 2004). For intraday liquidity, I have set the lowest limit at 5%; when carrying out this exercise, it did not seem that lowering that threshold to 3% would add meaningful information, as most payments remaining under the 5% threshold are already sent by the banks with the highest regulatory capital.

A practical difficulty related to this exercise concerned bank branches operating under a European passport; these branches do not need to maintain separate regulatory capital for their Dutch activities. There were 7 such branches participating in TOP in June 2004. They accounted

---

14 When setting their intraday limits, it does not seem that banks take into account any additional overnight lending to the institutions concerned. Such overnight lending does not need to be subtracted from the maximum allowed exposure in order to calculate a threshold for intraday lending.
for about 2.6% of all payments sent by banks in volume, and 3.8% in value. I have assumed that their payment activity in the Netherlands was indicative of their strength in terms of capital in that country, and have created a peer group for each bank, made up of 4 to 6 Dutch banks with comparable payment patterns in TOP. I have then taken the average regulatory capital of the Dutch banks in the peer group as a proxy for the regulatory capital of the foreign banks. Although this may not be totally correct in theory, as the solvability of these banks is related to the strength of their mother institution, from a payment system perspective it seems reasonable to assume that a bank’s ability to manage its payment activities in a foreign country is related to its payment flows. Discussions with banking supervisors and payment system operators confirm that view.

This exercise bears resemblance to the work of Mazars & Woelfel (2005), who simulate the technical failure of a major participant in the French PNS system at start-of-day. Up to 10% of payments get rejected in the second round. The authors experience with different levels of bilateral limits – these limits are managed centrally in the system - and conclude that “the consequences of a technical default could be greatly reduced if the participants set their bilateral sender limits at a lower level than that currently observed and if they reacted rapidly to information indicating a technical default by reducing their bilateral limits with the defaulting participant”. By reducing the level of exposure from 25% to 10% and 5%, I have tested if this conclusion also holds for TOP.

**Results**

In TOP, compared to no threshold at all, the existence of a threshold matters, as even a threshold of 25% reduces the secondary-round effects nearly by half, to slightly more than 9 unsettled payments, with a value of EUR 850 millions (see table 6). But there as well, reducing the level of that threshold to 10% and even 5% does not bring about the reduction of rejected payments that would be expected. Even when payment system participants do not allow for an exposure higher than 5% of their regulatory capital, the secondary-round effects appear to be stronger than when they only stop sending payments four hours after the start of the disruption. But for a similar level of secondary-round failures, the additional liquidity needed is lower, with EUR 1.1 billion for the 5% threshold and about 1.5 billion for the 4-hour stop-sending scenario. The difference between days is relatively important (see table in annex), with a minimum of 2 and a maximum of 20 unsettled payments at a 5% threshold.

These thresholds show that foreign banks (mainly subsidiaries of foreign banks that do need to hold capital in the Netherlands) and smaller banks will tend to delay payments earlier, as their regulatory capital is relatively low. With a threshold of 5%, most of the payments of such foreign participants to the stricken bank remain unset, as the value of these payments would amount to more than 5% of their regulatory capital.
Table 6: The impact of exposure control on secondary round effects and liquidity costs

<table>
<thead>
<tr>
<th>Exposure limited to (in % of regulatory capital)</th>
<th>Unsettled payments</th>
<th>Additional liquidity</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Number</td>
<td>Value (EUR million)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5%</td>
<td>8.71</td>
<td>756.96</td>
</tr>
<tr>
<td>10%</td>
<td>9.00</td>
<td>783.71</td>
</tr>
<tr>
<td>25%</td>
<td>9.38</td>
<td>850.42</td>
</tr>
<tr>
<td>Unlimited</td>
<td>16.62</td>
<td>1075.18</td>
</tr>
</tbody>
</table>

3.4 Limitations of these behavioural simulations

The simulation of payment systems participants’ behaviour makes it necessary to rely on a number of assumptions, which could be debated. Assuming all participants react in an identical matter is one of them. For example, the nationality or location of the counterparty would almost certainly play a role, as Dutch banks are more likely to continue to send payments to each other for a longer time after a disruption than would foreign banks. This is mainly due to their lower regulatory capital (as we have seen in 4.3), but it could also find its roots in a different relationship with the failing institution, less accommodative than would be Dutch banks, which have a long standing (and personal) relationship with the failing bank. Such differences can also be found among Dutch banks, with close cooperative relationships leading to less retained payments and vice versa. The size of the bank could also matter.

More specifically for the exercise with exposure reduction, the relation between regulatory capital and payments is not straightforward. Some banks have a high capital ratio and do few payments and vice versa. There are also some special banks, which are close to single purpose institutions and do many payments of very large values in spite of a relatively low capital ratio. The construction of the proxy could also be discussed, as foreign banks would be fully backed by the mother company.

Not only have I assumed identical behaviour among payment systems participants, but no differentiation has been made among payments sent to the stricken bank. However, the type of payment, its value, and its originator could all influence the sending bank’s likeliness to retain it in the face of a disruption. Some payments need to be made on a given day in order to discharge an obligation incurred for a certain amount of time, like the repayment of money market loans. These payments are not likely to be retained longer than the end of the day, in order to avoid overnight penalties. Other payments may be “coupled” to payments that the failing bank is about
to send; these payments would only be released after receipt of the incoming payment. The value of the payment also matters, as small value payments, which account for an important number of transactions settled by large-value payment systems\textsuperscript{15}, would be most likely to pass in spite of limits. Very large payments would be the first to be delayed. Note that this does happen under scenario 3, as limits happen to be often reached once a very large payment needs to be sent.

With regard to the originator, customer payments are more likely to be sent following a customer’s order, independently of the operational health of the receiving bank.

Finally, the order of payments is only of marginal importance in scenario 2, but can play a significant role in scenario 3. Due to the construction of the reaction function, all payments above the threshold remain unsent. This means that very large value payments that exceed the threshold automatically get blocked, independently of their timing. In practice, some banks send such payments very early in the day to the failing institution, before receiving any payments from that bank. In practice, these payments would unlikely be stopped at start-of-day. However, in the simulations, when very large payments precede smaller ones, both are blocked. This is not true when the smaller payments come first, so the rule can have different effects on different days (depending on the order of, among others, recurring payments). The influence of these variations on the results is however likely to be minor, as the large payments are the ones that matter.

\textsuperscript{15} As much as 25\% of all TARGET transactions have a value of less than EUR 1250.
4 Weighted networks and reciprocity: can participants control their exposure?

The results of the simulations suggest that the rapidity of participants’ reaction strongly reduces the secondary-round effects and the costs of a disruption. However, even with immediate reaction, payment system participants cannot fully insulate themselves from the impact of an operational disruption. In a bilateral relationship, this residual exposure can be expressed as the net position of the healthy institution vis-à-vis the failing bank, over the period starting from the moment at which the disruption takes place until the end of the day. If the healthy bank is in a net debit position, or if its position toward the failing institution is balanced, it will be able to protect itself from the effects of a disruption by stopping to send payments. If, on the contrary, it has a net credit position, its ability to protect itself will be limited, and it will retain a residual exposure equal to that net position.

From a system perspective, network analysis can give some insights in the aggregate sensitivity of the failing banks’ counterparties to an operational disruption at that bank. For that purpose, the constellation formed by the defaulting participant and all the banks to which it sends payments, or from which it receives payments can be considered as a partial network. Other relationships, where the failing institution is not involved, will not be considered here. Such a partial network resembles a “star” network, where any two nodes (here banks) are connected by exactly one path (see figure 5, where bank B is the stricken bank).

Figure 5: star network
In the example used for the simulations, there are 73 nodes, including the stricken bank (but excluding the central bank or ancillary systems), and 72 links, as all banks are linked to this bank. The connectivity of this network, which is the unconditional probability that two nodes share a link, is 1/72 or 1.4%. In a complete network, where all nodes would be linked, connectivity would be 100%. But the elimination from our analysis of links between healthy institutions substantially drives this figure down. Reciprocity gives an indication of the fraction of directional links for which there is a link in the opposite direction. In a payment system, this would indicate how likely a bank that is sending payments to another bank would be to also receive payments from that bank. In our example, the degree of reciprocity is very high, with 88.9% of two-way relationships. The stricken bank is clearly an important hub, being one of the big Dutch correspondent banks.

Figure 6

![Relation between In- and Out-strength of the node](image)

As we have seen in the very simple reasoning held above about banks’ ability to control their exposure in bilateral relationships, reciprocity between payment flows can enable banks to control their exposure to their counterparty. But this reciprocity ought not to be limited to the existence of bilateral payments, but also to be true for the associated values. The weight of links

---

Note that analysing only the partial network does not allow us to make any conclusions with regard to the way the impact of a disruption would spread among payment system participants.

According to Soramäki et al. (2006), in Fedwire, “less than a quarter of the relationships that exist between banks had payments going in both directions”.

---

16 Note that analysing only the partial network does not allow us to make any conclusions with regard to the way the impact of a disruption would spread among payment system participants.

17 According to Soramäki et al. (2006), in Fedwire, “less than a quarter of the relationships that exist between banks had payments going in both directions”.
originating in the node – ie the value of payments sent by the disrupted bank – is called “out-strength”, while the weight of links leading to the node – the value of payments received by the stricken bank – is called “in-strength”. Each participant which in a net debit or neutral position vis-à-vis the centre node will be able to withstand a failure from that node. Under these conditions the “in-strength” of node B is equal or superior to its “out-strength” vis-à-vis that participant. Therefore, all participants whose bilateral relationships with bank B place them in the upper quadrant of figure 6, or on the median line, will be able to protect themselves by delaying payments.

Although there is no aggregate measure of such specific vulnerability to node B, the degree of weighted reciprocity can provide some insights in the level of bilateral exposures between node B and each of its counterparties. That degree of weighted reciprocity can be calculated as a specific correlation function (Valverde & Solé, 2005)\(^\text{18}\):

\[
\rho^{a} = \frac{\sum_{i \neq j} (e_{ij} - \bar{e})(e_{ji} - \bar{e})}{\sum_{i \neq j} (e_{ij} - \bar{e})^2}
\]

where \(\bar{e} = \frac{\sum_{i \neq j} e_{ij}}{N(N-1)}\) is the average link weight (here the average payment value).

When \(\rho^a > 0\), the network is said to be (weighted) reciprocal; it is antireciprocal for \(\rho^a < 0\) and neutral for \(\rho^a = 0\).

In a partial (star) network, this would give us:

\[
\rho_{B}^{a} = \frac{\sum_{i \neq B} (e_{iB} - \bar{e}) (e_{Bi} - \bar{e})}{\sum_{i \neq B} (e_{iB} - \bar{e})^2 + \sum_{i \neq B} (e_{Bi} - \bar{e})^2}
\]

In our example, \(\rho_{B}^{a} = 0.44\)

Although the partial network under consideration is reciprocal in terms of the number of links, once these links are weighted, link weight reciprocity amounts to about 0.44. When, instead of

\(^\text{18}\) This is the only reference that I have found to such a measure. Many thanks to Goetz von Peter at the BIS for helping me through this.
taking each link per day separately, we aggregate the weights of the links between identical nodes over the course of the month (in other words, we add all payment values during the month), weighted reciprocity is somewhat higher and the network is closer to reciprocal, with 0.49. Over such a period, indeed, some transactions, like money market loans, tend to be reversed, bringing the values of bi-directional payment flows closer. However, it is difficult to assess reciprocity by considering only one isolated result (see Garlaschelli & Loffredo, 2006). It would be interesting to further measure reciprocity in TOP for other partial networks and for the whole system, and to repeat that measure over time. An international comparison in that respect would also certainly be of value.

19 The authors do not base their work on weighted reciprocity, but only on reciprocity. They find that similar numerical results can reflect different conditions in different networks.
5. Conclusion

This paper presents simulations carried out with the payment and settlement system simulator developed by the Bank of Finland. The operational failure of a major Dutch payment system participant is engineered, and the time-dependency of this effect estimated with the help of the simulator. This work also shows how participants’ behaviour in the face of a disruption can mitigate the impact of such a disruption, both for themselves – in terms of the costs of fulfilling their own payment obligations – and for the system as a whole – in terms of the payments rejected following a liquidity shortage.

It seems that these secondary effects of a disruption at a major TOP participant would be relatively limited at current liquidity levels, even if this disruption occurred at the start of the day. Liquidity levels that are higher than the upper bounds of liquidity, which correspond to the liquidity needed to settle all payments immediately, are helpful to reduce the impact of a disruption – but do not guarantee that such a disruption will be weathered without further payment rejections, as the distribution of this liquidity in the system matters (abstracting from an interbank money market). The cost of such a disruption could however be large for the healthy banks, as these banks normally rely on incoming payments to fund part of their outgoing transactions. These would then need to be funded with intraday credit and with overnight credit if the stricken bank were to prove unable to resume operations at the end of the day.

Payment system participants can control their exposure to their counterparties by way of internal bilateral limits, or by way of monitoring the activity of other participants and reacting to outages. Simulations have shown that even a slow reaction (e.g., 4 hours) can be effective in reducing the secondary effects of a disruption. However, in some situations, participants’ reactions can aggravate the problem, as payments already sent by the failing participant before the occurrence of a disruption, and waiting in the queue, can remain unsettled by lack of funding. Although participants’ actions are effective, they cannot completely eliminate the secondary effects of a disruption, nor can they eliminate the costs they have to make in order to obtain additional overnight funding when their positive reserves are depleted.

These findings suggest that, in general, payment system participants need to keep stringent bilateral limits in order to mitigate the individual costs of operational outages. A timed stop-sending rule, though less effective from a cost perspective, is to be favoured in order to prevent systemic effects. Even a reaction delay of 1 or 2 hours would substantially reduce the second-round impact of disruptions, without being as burdensome to implement as a 10-minutes rule. Central banks could share information about participants’ outages so as to allow banks to respond timely – although such transparency might be incompatible with the level of confidentiality associated with individual institutions’ behaviour.
It would be interesting to see whether these conclusions hold for different periods of time, for other large players in TOP and for different systems. Furthermore, the cost of intraday credit, not taken into account in this exercise (collateral being deposited ex ante at the central bank) could be used to assess the impact of a disruption in a system which offers such credit against a direct fee (e.g., Fedwire).

Individual participants’ ability to control their own exposure is determined by their relation with the failing bank. They can only control exposures when they do send a significant amount of payments to that bank. The analysis of the partial network formed by the outage bank and its counterparties shows that daily relationships cannot be qualified as fully reciprocal. The degree of (weighted link) reciprocity gives an indication of the bilateral ability to control exposures following an operational disruption. However, the information it provides about the central nodes’ counterparties’ ability to weather a failure of that central node is limited, as it does not distinguish between situations where the counterparties have a net debit or a net credit position. The relation between network characteristics and the ability for participants to control their exposure would deserve to be explored further. In order to understand this relation, it would be interesting to vary the network characteristics (of which the level of reciprocity and the role or weight of the central node), to conduct this analysis for different star networks, around different banks, or for a complete network, and to do this for different periods of time. Such a measure of reciprocity might also be used to analyse the efficiency of queue management and offsetting mechanisms in payment systems.
Annex: daily variation of results for timed stop-sending and exposure control

Timed stop-sending

<table>
<thead>
<tr>
<th>Reaction time</th>
<th>Number unsettled</th>
<th>Value unsettled (EUR millions)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Av</td>
<td>Min</td>
</tr>
<tr>
<td>10 mn</td>
<td>4.86</td>
<td>0.00</td>
</tr>
<tr>
<td>30 mn</td>
<td>4.86</td>
<td>0.00</td>
</tr>
<tr>
<td>1 h</td>
<td>4.90</td>
<td>0.00</td>
</tr>
<tr>
<td>2h</td>
<td>5.00</td>
<td>0.00</td>
</tr>
<tr>
<td>4h</td>
<td>7.43</td>
<td>2.00</td>
</tr>
</tbody>
</table>

Exposure control

<table>
<thead>
<tr>
<th>Reaction threshold</th>
<th>Number unsettled</th>
<th>Value unsettled (EUR millions)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Av</td>
<td>Min</td>
</tr>
<tr>
<td>5%</td>
<td>8.71</td>
<td>2.00</td>
</tr>
<tr>
<td>10%</td>
<td>9.00</td>
<td>2.00</td>
</tr>
<tr>
<td>25%</td>
<td>9.38</td>
<td>3.00</td>
</tr>
</tbody>
</table>
Bibliography


Previous DNB Working Papers in 2007

No. 124  Andrew Hughes, Rasmus Kattai and John Lewis, Early Warning or Just Wise After the Event? The Problem of Using Cyclically Adjusted Budget Deficits for Fiscal Surveillance

No. 125  Kerstin Bernoth, Juergen von Hagen and Casper de Vries, The Forward Premium Puzzle: new Evidence from Futures Contracts

No. 126  Alexandra Lai, Nikil Chande and Sean O’Connor, Credit in a Tiered Payments System

No. 127  Jean-Charles Rochet and Jean Tirole, Must-Take Cards and the Tourist Test

No. 128  James McAndrews and Zhu Wang, Microfoundations of Two-sided Markets: The Payment Card Example

No. 129  John P. Jackson and Mark J. Manning, Central Bank intraday collateral policy and implications for tiering in RTGS payment system

No. 130  John Lewis, Hitting and Hoping? Meeting the Exchange Rate and Inflation Criteria During a Period of Nominal Convergence

No. 131  Paul T. de Beer and Robert H. Mosch, The waning and restoration of social norms: a formal model of the dynamics of norm compliance and norm violation

No. 132  Carin van der Cruijsen and Sylvester Eijffinger, The economic impact of central bank transparency: a survey