Analysing the impact of operational incidents in large-value payment systems: A simulation approach

Paul Bedford

Stephen Millard

and

Jing Yang

Bank of England, Threadneedle Street, London, EC2R 8AH.
E-mail: paul.bedford@bankofengland.co.uk
stephen.millard@bankofengland.co.uk
jing.yang@bankofengland.co.uk

Copies of working papers may be obtained from Publications Group, Bank of England, Threadneedle Street, London, EC2R 8AH; telephone 020 7601 4030, fax 020 7601 3298, e-mail mapublications@bankofengland.co.uk

Working papers are also available at www.bankofengland.co.uk/wp/index.html

The Bank of England's working paper series is externally refereed.

ISSN 1368-5562

Contents

Abstract

In recent years, central banks have focussed considerable attention on improving the operational resilience of large-value payment systems (LVPS). These systems play a vital role in supporting financial market activity, often standing at the centre of a complex web of infrastructural arrangements for discharging payment obligations and clearing/settling securities trades. Their reliable operation is thus crucial to the continued stability of the financial system and, more generally, to establishing an environment that allows economic agents to exploit valuable opportunities for exchange of goods or assets. It is recognised, however, that the possibility for operational incidents to impair the ability of an LVPS to settle payments cannot be eliminated entirely; an element of residual operational risk will always remain.

In this paper, we propose a simulation-based approach to examining the *consequences* of different types of operational incident affecting an LVPS. Our method exhibits many similarities with the type of stress-testing exercises often used to evaluate the robustness of the banking system to financial shocks. In particular, we outline a three procedure: the first stage involves identification of a set of worst-case scenarios; the second, assessing the impact of operational incidents; the third step uncovers the empirical distribution of impacts of operational risk for various liquidity levels.

By way of illustration, we present results obtained from applying the proposed methodology to data extracted from CHAPS Sterling (the UK's main large-value payments system). We are able to conclude that the likelihood of an operational failure causing a significant disruption to CHAPS Sterling payments activity is low. This finding reflects the availability of a range of robust contingency arrangements and the abundance of available liquidity in the system.

Summary [To be written]

1 Background and motivation

During the 1990s, central banks devoted considerable resources towards reducing the financial risks banks and other financial institutions can become exposed to through their participation in large-value payment systems (LVPS). Most significantly, many countries implemented new systems based on models of real-time gross settlement (RTGS) in order to eliminate the credit exposures associated with systems that defer settlement in order to allow a netting process to be carried out.[1] For example, the UK's main LVPS – CHAPS Sterling – converted to RTGS in April 1996.

More recently, attention has shifted towards ensuring that key market infrastructures – and LVPS in particular – exhibit sufficiently robust levels of operational resilience. The value transferred by these systems each day often amounts to 20% or more of the annual GDP of the country concerned, which implies that a disruption to their operation has the potential to impact significantly on the users of these systems and, in extreme cases, undermine the stability of the financial system. To the extent that the disruption distorts economic agents' optimal trading decisions, there may also be an impact on the wider economy.

More specifically, a shock to the operation of an LVPS, for example because of a failure of the central payment processing infrastructure, has the potential to compromise one or more of the participants' ability to make payments discharging (possibly very large) settlement obligations. In turn, this increases the likelihood of financial distress at one institution having a knock-on impact on other institutions – that is, operational disruption can be a source of systemic risk. The immediate implication is that the payment system itself is of systemic importance.

The importance of operational resilience is recognised in the Core Principles for Systemically Important Payment Systems developed by the G10 central banks (BIS, 2001). In particular, Core Principle VII (CP VII) states that a system should ensure a high degree of security and operational reliability and should have contingency arrangements for timely completion of daily processing.

Compliance with CP VII is the minimum central bank overseers typically expect from an LVPS. But many such systems, motivated by heightened awareness of the possibility for deliberate, large-scale acts of terrorism to disrupt the financial institutions and markets, now aim to achieve even higher standards. New arrangements and procedures have been introduced to ensure continued operation in all but the most extreme of circumstances. It is recognised, however, that efforts to improve resilience cannot eliminate entirely the possibility for operational incidents to occur. Indeed, this is not what CP VII seeks to achieve; rather, the focus is on how operational risk in systemically important payment systems should be mitigated and controlled. This involves analysis of both the likelihood and impact dimensions of the risk.

[1] More detailed discussions of the features of the properties of RTGS and deferred net settlement systems can be found in BIS (1997) or McAndrews and Trundle (2001).

The focus of this paper is on the *impact* of various types of operational incidents. We propose a methodology for determining a set of 'worst-case' scenarios (in terms of scale of disruption), but do not consider the probability that a particular scenario will occur. This implies that the overall level of operational risk remains undetermined. Nevertheless, impact assessments (or stress-tests) based on simulations of remote probability events represent an integral part of the overall risk management process, particularly in the current, post-9/11 climate. It is our contention, therefore, that the approach proposed here can be a valuable addition to the range of methods available to the operators (and overseers) of LVPS for evaluating the extent to which a system is exposed to operational risk.

The remainder of this paper is structured as follows: Section 2 describes the main types of operational incident that can occur in a large-value payment system; Section 3 introduces our proposed methodology for identifying worst-case scenarios and evaluating their impact using simulation techniques; Section 4 applies the methodology to transactions data captured from the CHAPS Sterling system; and Section 5 presents conclusions.

2 Types of operational incident

2.1 Sources of operational disruption

It is a stylised fact that no two (LVPS) are identical. Indeed, there are many dimensions along which the design and structure of these systems can vary. The external environment within which a particular LVPS operates may also have an (idiosyncratic) impact on the risk profile of that system. But, at an abstract level, it is possible to characterise all LVPS in terms of four constituent parts:

- participants (usually banks);
- the provider/operator of the communications network;
- the provider(s)/operator(s) of the payment processing infrastructure; and
- the settlement agent.

For simplicity, in this paper we assume that a single entity – 'the system operator' – acts as both infrastructure provider/operator and settlement agent. This conforms relatively closely with real-world practices. In particular, most LVPS currently operating in the G10 settle their participants' payment obligations across accounts held with the local central bank. Moreover, it is common (albeit far from universal) for the central bank also to provide the payment processing infrastructure supporting the settlement process. Examples of LVPS covered by this characterisation include CHAPS Sterling in the UK and the American Fedwire system.[2]

With this assumption in hand, we are left with three possible sources of operational disruption in an LVPS. More precisely, any event that compromises the normal functioning of one or more participants; the provider/operator of the communications network; or the system operator would constitute an operational incident.

[2] In the case of Fedwire, the Federal Reserve also provides the communications network.

Operational disruption involving the *system operator* would typically constitute the unavailability of the core payment processing platform. Possible causes range from an IT (software or hardware) failure to the absence of sufficiently well-trained staff. External events such as natural disasters, power failures or terrorist action have the potential to have similar effects. Unavailability of the core platform would render the LVPS unable to handle payments in the normal way until the problem is resolved. Contingency arrangements that allow a limited number of payments to be processed and settled via alternative means may be available, but it is unlikely that these will be able to replicate fully the service offered by the LVPS under normal operating conditions.

A failure of the *communications network* supporting an LVPS would have the effect of preventing new payment instructions from reaching the system operator (except by contingency means, if these are available). While this clearly precludes the settlement of payments submitted after the failure occurs, it does not prevent those payments that have already been received by the system operator from being processed and settled in the normal way.

Finally, an operational incident could entail the inability of one or more of the *participants* to submit payment instructions to the LVPS. Such a situation would typically arise from a failure of the internal (back-office) systems of the participant(s) concerned, the possible sources of which are similar to those that may cause the core payment processing platform to be unavailable. Unaffected participants would be able to continue to operate as usual, and any payment instructions submitted to the system operator (including in favour of the 'stricken' bank or banks) would be processed and settled in the normal way.

2.2 Role of LVPS design

The likely impact of the operational incidents described above, expressed in terms of the volume and value of payments instructions affected, will typically depend on the design of the LVPS. The settlement model – RTGS or deferred net settlement (DNS) – is of particular significance.

In an RTGS system, settlement takes place on a continuous basis throughout the day. Conditional on the sending bank having sufficient liquidity available, a payment instruction is settled (with complete finality) immediately it is received by the system operator. In the absence of sufficient liquidity, payment instructions are queued to await the arrival of additional funds.[3] Crucially, this implies that all payments settled before an operational incident (of any type) occurred would not be affected.

Where settlement is deferred, on the other hand, all payment instructions submitted since the last net settlement event (and therefore not yet settled with finality) are 'at risk' from a disruption to normal operations. This is not to say, however, that the consequences of operational incidents are unambiguously greater in a DNS system than in an RTGS system handling an equivalent set of payment instructions. The

[3] There are two sources of additional funds: incoming payments (which may represent inter-bank loans) and intraday borrowing from the settlement agent.

temporary unavailability of any constituent part of a DNS system is unlikely to have a significant impact if the problem can be resolved in sufficient time to allow all payments to be processed and (net) settlement to take place as planned. By contrast, disruption to an RTGS system will almost certainly involve settlement of some payments taking place later in the day than would otherwise have been the case.

Beyond the chosen settlement model, a further important aspect of LVPS design concerns the development of contingency arrangements that aim to reduce the impact and, where possible, the likelihood of episodes of operational disruption. A wide range of alternative measures have been implemented, and no two systems operate an identical set of arrangements.[4] In the next Section, we abstract from the impact of these design features to develop a framework for analysing the impact of operational incidents affecting 'basic' RTGS and DNS systems.

3 A simulation-based framework for analysing the impact of operational incidents in LVPS

In this section, we focus upon developing a framework for assessing the impact of three distinct types of operational disruption that could compromise the ability of an LVPS to settle payments in the normal way:

i. the inability of one participant to send and receive payments;
ii. similar problems involving multiple participants (simultaneously); and
iii. the unavailability of the central payment processing infrastructure.

Intuitively, it seems obvious that the impact of incident (iii) will be greater than that of incident (ii), which will itself have more significant consequences that incident (i). But the actual extent of the disruption likely to be caused by any one of the three scenarios is much less easily determined. We therefore propose a generic framework (intended to be applicable to any type of LVPS) for the quantitative assessment of the maximum impact of particular types of operational disruption.

For each of the three types of incident listed above, we develop a basic two-step procedure. The first step involves using data on payment flows through the LVPS concerned to identify, on the basis of objective criteria, the 'worst-case' point in time for an operational incident to occur. The second stage involves simulating the pattern of payment activity in the LVPS conditional on the assumption that an incident actually occurs at exactly the worst-case time, and comparing this to activity under normal operating conditions.[5]

Our framework assumes that there are no circumstances under which the operational disruption creates doubts regarding the financial soundness of LVPS participants. Moreover, we assume in all cases that the operational disruption is of sufficient

[4] Section 4 includes a description of the arrangements implemented in CHAPS Sterling.

[5] It is also necessary to make some simple, intuitive assumptions regarding the behaviour of the (unaffected) participants following the operational incident.

severity to preclude a resumption of normal payment processing activities during the same business day.

3.1 Identifying worst-case operational disruptions

The 'worst-case' operational disruption is defined as a point of time (within a given time interval) when the occurrence of an operational incident entails the largest potential impact to the LVPS concerned.

3.1.1 Operational failure affecting one or more settlement bank(s)

We consider, first, an operational failure affecting one settlement bank. In this scenario, a single settlement bank – described as the 'stricken bank' – is unable to submit payments to the LVPS concerned owing to a failure of its internal back-office systems. For the stricken bank itself, the operational failure immediately creates a risk that it will not be possible to complete its payment activities by the end of the business day. The impact on the remainder of the LVPS participants will depend, however, on the design of the system.

In a DNS system, the operational failure is unlikely to have a large-scale impact. Even if, as assumed, the problem persists to the point of settlement, it is probable that the system operator will be able to invoke contingency arrangements in order to complete successfully the settlement process.

By contrast, in an RTGS system there is a much more significant possibility for the operational failure to disrupt the payment activity of the unaffected participants. In particular, there is a risk that a 'liquidity sink' effect will develop where available liquidity becomes concentrated on the settlement account of the stricken bank which is unable to 'recycle' the funds by initiating outgoing payments. In this case, those banks who are unaffected by the operational disruption may still be unable to settle their outgoing payments at the desired time (or even, in extreme cases, not at all) owing to shortages of available liquidity. The extent to which additional risks are created by system-wide liquidity shortages would depend on the whether the affected payments are 'time-critical' and thus require settlement at (or before) a certain time. For a large proportion of the payments settled through a typical LVPS, the precise time of settlement is unlikely to be of major significance. But a certain sub-set of payments are appropriately described as time-critical, and delays to any of these payments would thus represent a crystallisation of liquidity risk.

The likelihood of liquidity shortages arising can be reduced, however, by the behaviour of the unaffected participants. Indeed, on learning of the operational problem, one response would be immediately to stop sending payments to the stricken bank (thereby preserving liquidity). In practice, anecdotal evidence from CHAPS Sterling suggests that the time-lag between a settlement bank experiencing an operational failure and the flow of payments to that bank slowing significantly is likely to be of the order of ten minutes.

Nevertheless, the possibility for a liquidity sink to develop remains real, particularly if the stricken bank holds a large positive balance on its account with the settlement agent (and is therefore controlling a significant amount of liquidity) at the time of the

operational failure. The risk is also greater where the stricken bank is due to receive a large gross value of payments in the few minutes immediately following the operational failure. One of the potentially most problematic scenarios therefore involves the operational incident occurring at a point in time when the stricken bank has the potential to act as a liquidity sink and there remain a large volume (and value) of payments still to settle.

On this basis, the potential for an LVPS participant settlement bank to become a liquidity sink can be assessed using actual payment flow data. In particular, a 'virtual' credit balance – defined as the actual balance on account with the settlement agent plus the gross inflow of funds over the next ten minutes – can be calculated for each participant at every point in time. The worst-case date and time for the operational failure to occur may then be determined by identifying the peak virtual credit position (for any participant) observed during the sample period, subject to this occurring before a particular time of day.[6]

The risk implications of an operational disruption affecting the ability of multiple settlement banks to submit payments to an LVPS are similar to those arising from the operational failure of a single settlement bank. In particular, the scale of the risk will be related to the aggregate volume and value of payments that remain unsettled at the time of the initial shock, along with the potential of the stricken banks to act as liquidity sinks. On both measures, the risks are intuitively likely to be greater than in the case of a single stricken bank.

3.1.2 Operational failure of the core payment processing platform

In this scenario, the participants in an LVPS are unable to settle their payment obligations after a particular point in time (and for the remainder of the day concerned) owing to a complete failure of the core payment processing platform. If we assume that the system is at least able to open as normal, the worst-case scenario is clearly one in which the incident occurs very early on high-volume day (i.e. when there remains a large number of payments still be settled). This applies equally to RTGS and DNS systems.

Recognising the potential for a failure of the core payments processing platform to prevent the settlement of a large volume of payments, most LVPS have, however, developed contingency arrangements designed to allow the most urgent payments to be settled. The presence of these back-up arrangements, which typically differ from system to system, can potentially change the nature of the worst-case scenario.

In the main, contingency arrangements are not able to cater for the full volume of payments processed under normal operating conditions; the possibility of some transactions remaining unsettled is therefore real. An exception to this is the 'RTGS by-pass mode' facility implemented in CHAPS Sterling. In the event that the RTGS infrastructure at the centre of the system is unavailable, CHAPS Sterling is able to revert to a model of deferred net settlement that should allow the vast majority of payments to settle on a same-day basis. Section 4 below

[6] A time-of-day constraint serves to ensure that there will be a significant volume and value of payments still to settle following the initial shock to the system.

provides further details on by-pass mode and its consequences for the (operational) risk profile of CHAPS Sterling. In particular, we consider how the switch from RTGS to DNS can re-introduce an element of settlement risk and the impact this has on the worst-case scenario in respect of a failure of the core payment processing platform.

3.2 Assessing the impact of operational disruption

The second step of the overall process involves simulating activity in the LVPS under the assumption that the worst-case scenario does actually occur. This allows the maximum likely impact of a particular type of operational incident to be identified, by comparing the outcome under stress conditions with the situation where the system is able to operate normally. The impact of an operational incident is measured as the relative delay compared to a benchmark case.

To assess the impact of exceptional events, it is first necessary to establish a 'benchmark' against which the results of simulations of operational events may be compared. This involved conducting a simple simulation of a LVPS under normal operating conditions using actual transaction data. For the purposes of establishing benchmark liquidity levels, the simulation permitted all settlement banks to draw on unlimited amounts of intraday credit.

Output from the benchmark simulation allowed the calculation of two *hypothetical* liquidity levels useful to simulation-based analysis of RTGS systems. In particular, it was possible to derive, for each settlement bank on each day within the sample period, the 'upper bound' and the 'lower bound' of liquidity.

The upper bound of liquidity measures the amount of intraday credit a settlement bank would need to obtain in order for all its outgoing payments to settle immediately upon their submission to the central processor (that is, without being placed in the central queue to await the arrival of additional liquidity). The lower bound of liquidity, on the other hand, refers to the amount required for the settlement bank just to cover its net outflow of funds across the day as a whole.[7] Although, the total amount of intraday credit obtained in an LVPS may exceed the upper bound and will always exceed the lower bound in an RTGS system, the upper and lower bounds nonetheless define a suitable range across which settlement banks' ability to draw on intraday credit can be varied in order to examine the impact of changes to the amount of liquidity available in the system.[8]

To facilitate investigation of the extent to which the amount of liquidity available in a LVPS influences the ability of the system to withstand different types of operational disruption, additional benchmark simulations were conducted (each using the same transaction data and with the LVPS operating

[7] For further discussion of upper and lower bound of liquidity concept, see, for example, Bech and Sorämaki (2001).
[8] For example, the total amount of intraday credit obtained by the CHAPS Sterling settlement banks typically significantly exceeds (by about half) the upper bound; this largely reflects the low overall cost of the liquidity. In addition, settlement banks face further demands on their liquidity positions that are unrelated to CHAPS Sterling (for example, from the CREST securities settlement system).

under normal conditions). In these experiments, the amount of intraday credit available to the settlement banks was constrained to levels between the upper bound (UB) and the lower bound (LB), defined on the basis of the following expression

$$\text{UB} - \alpha\,(\text{UB} - \text{LB}) \qquad \textbf{(1)}$$

with α set equal to 1, 0.75, 0.5, 0.25 and 0.

The second stage of the analysis then involved comparing the benchmark cases to the results from simulations of the LVPS operating under stress conditions. We use four delay measures to capture the impact of a disruption. The first two are straight forward: total value and volume of unsettled payments. The third measure – average queue value – refers to the average (across the entire day) value of payments held in the central queue. The forth measure is the 'delay indicator' introduced by Bech and Soramäki (2001). This statistic is based on the amount of time each individual payment spends in the central queue relative to its maximum possible queuing time; a value-weighted average of the (relative) delay to all payments may be viewed as a measure of the aggregated level of delay in a payment system.

Algebraically, the delay indicator (which must lie between zero and one) is defined as:

$$\frac{\sum_i \left(t_{2,i} - t_{1,i}\right) a_i}{\sum_i \left(t_{end} - t_{1,i}\right) a_i} \qquad \textbf{(2)}$$

where a_i is the value of payment i, $t_{1,i}$ and $t_{2,i}$ are, respectively, the submission and settlement times for payment i and t_{end} is the time for the end of the business day. A value of one shows that every payment has been held in the queue for the maximum possible time (that is, from the point it was first submitted to the system until the end of the business day).

For each given level of liquidity, the impact of an operational incident is gauged by the difference between delay in the benchmark case and the stress case.

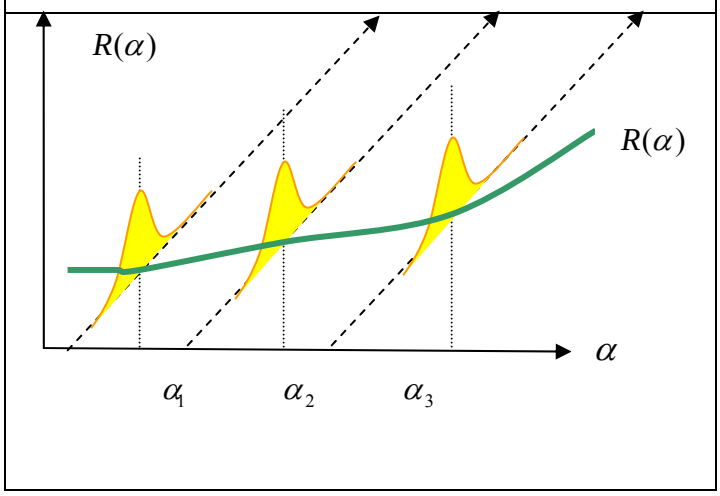
3.3 Uncovering an empirical distribution for the impact of operational disruption

Although in the application below we look at the impact of different operational disruptions for a particular LVPS in a particular data period, it is, perhaps, more useful to develop a feel for where and when the impact of various operational disruptions is more or less worrying and whether or not particular systems are becoming more or less resilient. In what follows, we discuss how one might build up a distribution for the effects of operational disturbances that can be used to answer these questions.

Consider a payment data set with T time periods that can be divided into n equal time intervals, denoted as $\{\tau_1, \tau_2 ... \tau_n\}$. (Relating this to the application we discuss in section 4, we can think of the time intervals as being months.) If the type of operational incident is denoted $j \in J$, then the impact of an incident of type j for a

given level of liquidity α over the time interval τ_i can be written as $R_{t_i}(j,\alpha)$, where t_i represents the worst-case point in interval τ_i (obtained using the methodology discussed in section 3.1). We think of the impact functions $R_{t_i}(j,\alpha)$ as the delay measures discussed in the previous section but any variable capturing the impact of an operational incident will do. Given a particular data set, we can calculate $R_{t_i}(j,\alpha)$ for all *i*, *j* and α. Then, for a large enough *n* we can uncover the distribution of *R*, for each pair of $\{j,\alpha\}$. Chart 1 illustrates the idea.

Chart 1: Impact of operational incident *j*



4 Application: Assessing the resilience of CHAPS Sterling

4.1 Operational risk controls in CHAPS Sterling

The Bank of England aims to ensure that the CHAPS Sterling system's central payment processing infrastructure is available at all times during every business day. A range of controls and procedures are in place to reduce the likelihood of disruption from internal sources of operational risk. In addition, the Bank maintains remote back-up facilities capable of assuming fully the payment processing role should the primary system experience problems.

In the unlikely event of the primary and back-up payment processing infrastructures being unavailable simultaneously, an additional layer of contingency exists in the form of 'RTGS by-pass mode'. In by-pass mode, which has been extensively tested but never required for actual operations, CHAPS Sterling reverts to a model of deferred net settlement. Payment information continues to flow between settlement banks, but central bank money is no longer transferred in real-time. Rather, obligations are settled (across accounts held with the Bank of England) on a multilateral net basis at the end of the business day.

The aim of by-pass mode is to ensure all critical CHAPS Sterling payments can be settled same-day. If the disruption occurs sufficiently late in the day, by-pass

mode may not be required to achieve this objective; the Bank of England could process a small volume of unsettled payments manually. Where it is necessary to invoke by-pass mode, however, the switch to deferred net settlement re-introduces some of the financial risks eliminated by RTGS. In particular, the default of a settlement bank holding a net debit position could cause the other settlement banks (and their customers) to incur financial losses.[9]

To control the scale of the risks to which settlement banks could be exposed in the event of by-pass mode ever being needed, an arrangement based upon multilateral net sender (or net debit) caps was put in place in July 2003. It is intended that these (self-imposed) caps would be set equal to the amount of collateralised intraday liquidity each settlement bank is able to draw from the Bank of England at the point by-pass mode is invoked. As a result, the credit risk associated with by-pass mode would be tightly controlled – all net debit positions would be fully backed by collateral.

Although this model of full collateralisation is the preferred approach to controlling credit risk in by-pass mode, its feasibility depends on the ability of the Bank of England to determine settlement banks' account positions at the time by-pass mode is invoked. In circumstances where this is not possible, each settlement bank would set its net sender cap at a maximum level of £1 billion.

In this second variant of by-pass mode, it is possible that the net sender caps would prevent the settlement of some payments. Such a situation could arise when the net value of payments an individual settlement bank had to make during a period of by-pass mode operation exceeded £1 billion. The constraint imposed by the sender caps could, however, in practice be relaxed by means of inter-bank loans agreed between the settlement banks, which, when processed by CHAPS Sterling (operating in by-pass mode), would reduce the net debit position of the borrower.

Arrangements for mitigating operational risk affecting the central payment processing infrastructure are necessary, but not sufficient, to ensure the overall robustness of CHAPS Sterling: it is also important for the individual settlement banks to implement suitable risk controls. As a condition of membership of the system, therefore, settlement banks are required to comply with certain minimum standards in respect of their internal back-office arrangements for handling payments. An auditable system of self-certification is in place to ensure these standards are satisfied.

In addition, a range of contingency arrangements are in place to limit the impact on CHAPS Sterling of instances where a settlement bank is unable to submit payments to the system. For example, the settlement bank concerned may send authenticated faxes or use the RTGS Enquiry Link facility to instruct the Bank of England to process manually a small number of high-priority payments.[10] This mechanism is particularly important for ensuring that 'time-critical' payments,

[9] For a description of the risk properties of deferred net settlement (relative to the RTGS alternative), see Bank for International Settlements (1997) or McAndrews and Trundle (2001).
[10] The Enquiry Link facility is an interactive information and communication service available to all CHAPS Sterling settlement banks.

for example those associated with other market infrastructures (such as multilateral net settlement in BACS or pay-ins to the Continuous Linked Settlement (CLS) system, which settles foreign exchange transactions on a payment-versus-payment basis) are processed at the required time.

The inability of a CHAPS Sterling settlement bank to send payments also raises the possibility of a 'liquidity sink' developing as available liquidity becomes concentrated on the settlement account of the bank concerned. Intraday, this could cause liquidity shortages elsewhere in the system, which may in turn lead to significant delays to the settlement of payments between the unaffected settlement banks. It is therefore important for prompt action to be taken to prevent large flows of liquidity to a settlement bank experiencing operational problems – this would typically be achieved by issuing a general request (a 'stop-sending') for no payments to be sent to the settlement bank concerned.

A liquidity sink that occurred late in the business day would also be likely to leave the unaffected settlement banks short of funds and unable to re-pay their intraday borrowing from the Bank of England.[11] The CHAPS Sterling 'stricken bank scheme' (which has never been used) addresses this problem by requiring a settlement bank suffering operational problems after 4pm to extend uncollateralised overnight loans (settled manually by the Bank of England) to any unaffected settlement bank requiring additional liquidity.

4.2 Data

To illustrate the method we discussed in section 3.2, we selected one time interval, February 2004, with a total of 2.1 million payments. For this interval, we first identify the worst-case scenario for each type of operational incidents, and then simulate possible delay impact across five different liquidity levels. As only having tried on one interval, we can't get the distribution of impact for a given liquidity level, but rather one point for each liquidity level. Ideally, the similar exercise should be repeated for a few years data with enough intervals to generate the whole distribution.

The findings presented below are based on a series of simulation experiments carried out using a payment system simulator developed by the Bank of Finland.[12] Each experiment was conducted using a simulation set-up intended to replicate CHAPS Sterling as closely as possible, including in respect of central queuing arrangements and procedures for gridlock resolution.[13]

---

[11] The Bank of England extends interest-free intraday credit to all CHAPS Sterling settlement banks, in return for suitable collateral. Intraday borrowing not re-paid by the end of the business day is converted into collateralised overnight borrowing and charged a penal rate of interest.

[12] For further details regarding the simulator and its capabilities, see www.bof.fi/eng/3_rahoitusmarkkinat/3.4_Maksujarjestelmat/3.4.3_Kehittaminen/3.4.3.3_Bof-pss2/index.stm. The authors would like to thank Harry Leinonen and colleagues at the Bank of Finland for their work in developing the simulator.

[13] Under current operational procedures, the CHAPS Sterling central queuing mechanism is not used; this implies that settlement banks must queue payments for which insufficient liquidity is available within their own back-office systems. The simulations carried out for this article relax this constraint and allow payments to queue centrally. In addition, the simulations employed gridlock resolution procedures at the end of the day only.

4.3 Results and discussion

4.3.1 Operational failure affecting one settlement bank

One of the potentially most problematic scenarios therefore involves the operational incident occurring at a point in time when the stricken bank has the potential to act as a liquidity sink and there remain a large volume (and value) of CHAPS Sterling payments still to settle.

Applying this reasoning, a set of simulation experiments (using five different initial levels of liquidity ranging from the lower bound to the upper bound) were carried out under the assumption that an operational problem affected one particular CHAPS Sterling settlement bank at a point when it held a credit position of £4.2 billion on its account with the Bank of England. Over the remainder of the day concerned, this settlement bank was due to be either sender or receiver for some 46,000 payments with an overall value of £45.7 billion. This illustrates a point made by James (2003) – an operational incident of the kind considered here has the potential to have a large impact on the total volume and value of payments actually settled in CHAPS Sterling.

More significant from a systemic perspective, however, is the extent to which the operational failure creates liquidity shortages in the system as a whole. In particular, it is interesting to determine the extent to which payments between unaffected settlement banks are either delayed or prevented from being settled.

Under normal operating conditions, any level of initial liquidity at least equal to the lower bound would be sufficient to allow all CHAPS Sterling payments to settle same-day. This is not necessarily the case, however, following an operational incident – the failure of the unaffected settlement banks to receive payments from the stricken bank may leave them short of liquidity. In fact, the simulation results (reported in Table A) reveal that, at most liquidity levels, the disruption did not prevent settlement of a substantial volume and value of payments between unaffected settlement banks. Indeed, a significant impact (in terms of unsettled payments) was observed at the lower bound of liquidity only.

Given that the actual amount of liquidity available in CHAPS Sterling typically exceeds the upper bound, the findings reported in Table A imply that CHAPS Sterling is well-placed to withstand an operational incident affecting one settlement bank. This conclusion is strengthened further by the presence of the stricken bank scheme, the effects of which were not captured in the simulation experiments. If the scheme was invoked, the resulting re-distribution of liquidity should allow some (or all) of the outstanding payments to be settled.

Table A: Effect of an operational failure affecting one CHAPS Sterling settlement bank on payments between other settlement banks – unsettled payments

Liquidity level	Value of unsettled payments (£ billions)	Volume of unsettled payments
α = 0 (UB)	0.00	0
α = 0.25	0.00	0
α = 0.5	0.01	7
α = 0.75	0.35	66
α = 1 (LB)	4.03	4,086

Table A captures the extreme outcome of the operational failure causing liquidity shortages that are of sufficient scale to prevent same-day settlement of all CHAPS Sterling payments. Also significant, however, is the extent to which settlement of individual payments (particularly those that are time-critical) is delayed beyond the preferred time – that is, the amount of extra time payments spend in the central queue awaiting the arrival of additional liquidity.

Table B draws on the simulation results to present two alternative measures of queuing and delay, in each case expressed relative to results obtained from the benchmark simulations of CHAPS Sterling operating under normal conditions (with the appropriate amount of available liquidity). The first measure is average queue value and the second measure is the 'delay indicator' defined in section 3.2.

Table B: Effect of an operational failure affecting one CHAPS Sterling settlement bank on payments between other settlement banks – delayed payments

Liquidity level	Average queue value (£ billions)	Delay
α = 0 (UB)	0.01	0.00
α = 0.25	0.03	0.01
α = 0.5	0.07	0.01
α = 0.75	0.15	0.04
α = 1 (LB)	0.23	0.06

The results reported in Table B point to a similar conclusion to that drawn from Table A. In particular, an operational failure involving a single settlement bank is unlikely to have a significant impact on the system as a whole (or, more precisely, on the ability of the unaffected settlement banks to make payments to each other), except in situations where the initial level of liquidity is low.

Moreover, current levels of liquidity in CHAPS Sterling imply that such an outcome is improbable.

4.3.2 Operational failures affecting multiple settlement banks

To examine this further, output from the benchmark simulation was used to calculate the aggregate credit position held on the Bank of England accounts of three randomly-selected settlement banks at every point in time during the sample period. The peak value of this metric (again subject to the before 12 noon constraint) was then used to identify the worst-case date and time for an operational incident simultaneously to affect the three settlement banks in question.[14]

A further set of simulations were then carried out under the assumption that the three settlement banks would all encounter operational problems at a point when they (collectively) controlled £4.8 billion of liquidity. The immediate consequence of this was that nearly 51,000 payments with an overall value of £143.4 billion could not be settled same-day because they involved one of the stricken banks as either payer or payee.

In order to assess the impact on the system as a whole, Table C describes the outcome of the simulations in terms of the volume and value of payments between settlement banks unaffected by the operational incident that were left unsettled at the end of the day.

A comparison of Tables A and C reveals that the impact of operational disruption involving multiple settlement banks is, unsurprisingly, significantly greater than in circumstances where a single settlement bank is affected. Nevertheless, it remained the case that no payments between unaffected settlement banks were left unsettled at the upper bound of liquidity (by implication, this result would also hold at actual levels of liquidity in CHAPS Sterling).

Table C: Effect of an operational failure affecting three CHAPS Sterling settlement banks on payments between other settlement banks – unsettled payments

Liquidity level	Value of unsettled payments (£ billion)	Volume of unsettled payments
$\alpha = 0$ (UB)	0.00	0
$\alpha = 0.25$	0.68	24
$\alpha = 0.5$	2.84	1,078
$\alpha = 0.75$	7.58	3,225
$\alpha = 1$ (LB)	13.08	6,299

[14] For simplicity, it is assumed that any operational incident affecting multiple settlement banks (for example, a localised general power failure) will be of sufficient severity that it is immediately visible to all settlement banks; the flow of payments to the stricken banks would therefore cease almost immediately.

Table D describes the outcome of the simulations on the basis of settlement delay, using the two measures described above. As previously, the results are (in each case) reported relative to the outcome of the benchmark simulations of CHAPS Sterling operating under normal conditions.

Table D: Effect of an operational failure affecting three CHAPS Sterling settlement bank on payments between other settlement banks – delayed payments

Liquidity level	Average queue value (£ billions)	Delay
α = 0 (UB)	0.02	0.03
α = 0.25	0.05	0.05
α = 0.5	0.07	0.06
α = 0.75	0.13	0.20
α = 1 (LB)	0.02	0.25

It is clear from Table D that an operational event affecting three settlement banks is likely to lead to a significant increase in settlement delays, even at upper bound levels of liquidity. A caveat to this finding, however, is that the removal of all payments involving the stricken banks has a significant impact on the set of payments upon which the measures of queuing and delay are based (relative to the benchmark simulations). This explains why the relationship between liquidity level and the average queue value is not monotonic.

4.4 Operational failure of the central payment processing infrastructure

In this scenario, it is assumed that an operational incident renders the CHAPS Sterling central payment processing infrastructure inoperable. As discussed in Section 3.1, the standard response to such a situation (which has never occurred) would be to invoke RTGS by-pass mode. If the operational incident occurred late in the day, however, it may not be necessary to take this step – the Bank of England would be able to process manually a small number of unsettled payments, which would avoid the introduction of additional risks. This highlights the importance, from an operational risk perspective, of rapid payments throughput in CHAPS Sterling.

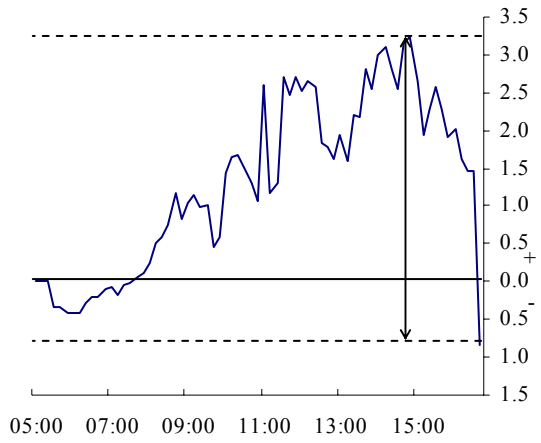
Invoking by-pass mode would allow payment processing to continue, but would also create additional settlement risks. In particular, the departure from RTGS would result in intraday credit exposures between the CHAPS Sterling settlement banks, the scale of which would be directly related to the absolute size of individual settlement banks' net debit positions.

Under the preferred variant of by-pass mode, net debit positions are fully backed by liquidity held with the Bank of England. The credit risks are thus tightly controlled. This may not be the case, however, in situations where it is necessary to implement the alternative version of by-pass mode and impose net sender caps equal to £1 billion for each settlement bank. In such circumstances, net debit

positions may be only partially collateralised – a default could therefore result in the remaining settlement banks (and their customers) incurring financial losses. Moreover, there are currently no formal arrangements in place to provide the additional liquidity that would be required to complete the net settlement.[15]

One of the most difficult scenarios therefore involves invoking by-pass mode at a point of time when an individual settlement bank would incur a large net debit position during the period CHAPS Sterling is operating in by-pass mode. On any given day, the maximum possible value of this position may be calculated (using the output from the benchmark simulation) by measuring the difference between each settlement bank's largest intraday net credit position (on its account with the Bank of England) and its end-of-day balance. This is illustrated in Chart 2, which depicts a randomly-generated settlement account balance over the course of one business day. The vertical distance shown by the arrow represents the maximum possible net debit position the settlement bank concerned could incur were by-pass mode to be invoked intraday.

Chart 2: Sample settlement account balance



Using data from February 2004, the worst-case outcome was identified by determining the time at which invoking by-pass mode would lead to the largest single net debit position. A simulation experiment was then performed under the assumption that by-pass mode was invoked (with net sender caps set at £1 billion for each settlement bank) at this exact time.

The simulation results indicate that under such circumstances 23 payments with a total value of £3.8 billion would have remained unsettled at the end of the day. This finding stems from the fact that one settlement bank needed to make net payments in excess of £1 billion between the time that bypass mode was invoked and the end of the day, but was unable to do so as a result of the sender cap.

[15] The CHAPS Company intends to consider the introduction of such arrangements once a similar financial risk management scheme involving the UK's main retail systems (BACS and the Cheque and Credit Clearings), both of which settle on a net basis, has been implemented.

An important aspect of this analysis is that it has assumed that settlement bank behaviour is unchanged. In practice, a settlement bank would be likely to raise additional liquidity by borrowing in the inter-bank market (or elsewhere) in order to ensure it can settle all of its outstanding transactions. This has the effect of transferring, but not eliminating, the credit risk associated with the net debit position; rather than being within CHAPS Sterling, the exposure would then be held outside the system by the lending bank.

5 Conclusions

This article has employed a simulation approach to explore the consequences of operational disruption affecting CHAPS Sterling, the UK's main large-value payment system. In particular, a series of experiments have been carried out to quantify the liquidity and credit risks that the settlement banks could incur in the event that normal payment processing activity was to be disrupted.

The simulation results illustrate that, while there remains scope for CHAPS Sterling contingency arrangements to be improved further (for example through the agreement of a loss-sharing mechanism for use in the event of a settlement bank being declared insolvent while the system is operating in by-pass mode), the system is already highly resilient. More specifically, the analysis has shown that CHAPS Sterling is well-placed to withstand a variety of plausible, though low-probability, types of operational disruption, and thus that the likelihood of operational risk in CHAPS Sterling acting as a source of financial instability appears to be reasonably small. Nevertheless, a degree of liquidity risk is still potentially present; this is especially the case in situations where an operational incident affects many settlement banks simultaneously (as illustrated by the simulation results presented in Tables C and D).

References