HOW MUCH LIQUIDITY WOULD A LIQUIDITY-SAVING MECHANISM SAVE IF A LIQUIDITY-SAVING MECHANISM COULD SAVE LIQUIDITY? A SIMULATION APPROACH FOR CANADA’S LARGE-VALUE PAYMENT SYSTEM
How much liquidity would a liquidity-saving mechanism save if a liquidity-saving mechanism could save liquidity? A simulation approach for Canada’s large-value payment system

Shaun Byck\textsuperscript{a} and Ronald Heijmans\textsuperscript{a,b}

\textsuperscript{a}Payments Canada
\textsuperscript{b}De Nederlandsche Bank

Friday 14\textsuperscript{th} February, 2020

Abstract

Canada’s Large Value Transfer System (LVTS) is in the process of being replaced by a real-time gross settlement (RTGS) system. A pure RTGS system typically requires participants to hold large amounts of intraday liquidity in order to settle their payment obligations. Implementing one or more liquidity-saving mechanisms (LSMs) can reduce the amount of liquidity participants need to hold. This paper investigates how much liquidity requirements can be reduced with the implementation of different LSMs using LVTS transaction data from 2018. These LSMs include: 1) Bilateral offsetting, 2) FIFO-Bypass, 3) Multilateral offsetting, and 4) a combination of all LSMs. We simulate two different scenarios. In the first scenario, all payments from Tranche 1, which are considered time-critical, are settled in a pure RTGS payment stream, while less time-critical Tranche 2 payments are settled in a payment stream with LSMs. In the second scenario, we settle all payments (Tranche 1 and 2) in the LSM stream. Our results show that when there is ample liquidity available in the system, there is minimal benefit from LSMs, as payments are settled without much delay—the effectiveness of LSMs increases as the amount of intraday liquidity decreases. A combination of LSMs shows a reduction in liquidity requirements that is larger than any one individual LSM.

Keywords: Liquidity Saving Mechanism, Simulation, LVTS, RTGS, Financial Market Infrastructure, Intraday Liquidity, Collateral.


\textsuperscript{*}Byck (corresponding author) and Heijmans can be reached at sbyck@payments.ca and ronald.heijmans@dnb.nl, respectively. The views expressed in the paper are solely those of the authors and do not necessarily represent the views of Payments Canada, the Eurosystem, or De Nederlandsche Bank.

\textsuperscript{†}We are grateful to colleagues at Payments Canada, Segun Bewaji, Philippe-Olivier Blanchet, Viktoria Galociova, Fuchun Li, and David Millard, for helpful comments and discussions. The authors also gratefully acknowledge comments from Samantha Cook at Financial Network Analytics and Neville Arjani at Canada Deposit Insurance Corporation.
1 Introduction

As one of the most important financial market infrastructures in developed economies, large-value payment systems sit at the center of the financial system and provide the means for commercial banks to clear and settle financial obligations between one another. The Large Value Transfer System (LVTS), owned and operated by Payments Canada, provides the main clearing and settlement infrastructure for large-value payments between commercial banks in Canada, and is one of a handful of financial market infrastructures designated as systemically important by the Bank of Canada. However, the LVTS began operation more than 20 years ago. As the application and infrastructure age, they present technological and operational challenges. To ensure Payments Canada can continue to serve the financial system effectively, the LVTS will be replaced as part of Payments Canada’s multi-year project to modernize its core payment systems.

Historically, most countries’ large-value payment systems have been deferred net settlement (DNS) systems. In a DNS arrangement, final settlement of payments is deferred for some period of time, usually until the end of the day, and done on a multilateral net basis. However, as the volume and value of interbank payments has increased dramatically over the past several decades, most central banks around the world have opted to implement real-time gross settlement (RTGS) systems in place of their DNS counterparts (Bech and Hobijn, 2007). RTGS systems are attractive for central banks because they limit the credit exposures that can build up over the day in a DNS system by settling payments continuously throughout the day on an individual basis. However, limiting this credit exposure comes at a cost. The liquidity needed to settle payments gross in real-time is generally a considerable amount more than that needed to settle the net difference at the end of the day. Moreover, because access to this liquidity is costly for settlement banks, there is the potential for the underprovision of intraday liquidity and, as a result, delays in the settlement of transactions (Ball et al., 2011). In short, RTGS systems trade credit risk for liquidity risk.

Consequently, replacing the LVTS with an RTGS system has the potential to increase LVTS participants’ liquidity requirements significantly. As the operator of this future RTGS system, Payments Canada is responsible for ensuring its safety and soundness, but also its efficiency. Excessive liquidity requirements on the part of its participants can put significant limitations on the system’s efficiency while increasing liquidity risk. Today, liquidity-saving design features, commonly referred to as "liquidity-saving mechanisms" (LSMs), are commonplace around the world and have been adopted under the Principles for Financial Market Infrastructures as a way for RTGS systems to manage liquidity risk effectively (CPSS, 2012).  

This paper aims to identify the optimal combination of LSMs for Payments Canada’s future RTGS system. Using the payment system simulation engine provided by Financial Network Analytics (FNA)\textsuperscript{2}, we perform counterfactual simulations with LVTS data to examine the intraday liquidity requirements of moving to an RTGS system, and attempt to identify an optimal set of LSMs. We extend the work of Embree and Taylor (2015), which examines the impact of full collateralization of the LVTS, by introducing LSMs. We also revisit the work of Arjani (2006), which examines the trade-off between settlement delay and intraday liquidity in the LVTS, and examine the trade-off under an RTGS settlement model.

The relationship between settlement banks’ behaviour with respect to their use of intraday liquidity and incentive to delay the settlement of payments has been examined in the literature from several perspectives. Among others, Angelini (1998), Bech and Garratt (2003), and Galbiati and Soramäki (2011) have employed a range

\textsuperscript{1} See Table 1 in Fugal et al. (2018) for an overview of the LSMs in use in major countries, including Australia, the Eurozone, Korea, Japan, Mexico, Singapore, Sweden, Switzerland, and the UK.

\textsuperscript{2} For more information on the FNA platform, see www.fna.fi
of theoretical frameworks, including game-theory and agent-based modelling, to study participant behaviour in an RTGS system and the incentives to submit or delay payments. Abbink et al. (2017) have also investigated these incentives in an experimental study using a stylized version of the model of Bech and Garratt (2003). Several papers, including Martin and McAndrews (2008), Galbiati and Soramäki (2010), and Jurgilas and Martin (2010), introduce an LSM into a theoretical framework and demonstrate how liquidity-saving measures can increase welfare.

A related set of papers analyze the trade-off between efficiency, risks, and costs in an RTGS system based on the framework first introduced by Berger et al. (1996). Koponen and Soramäki (1998) was among the first papers to use this framework in a simulation approach, simulating a variety of system designs using data from the Finnish BoF-RTGS system. Enge and Øverli (2006) and Arjani (2006) each use a simulation approach to quantify the trade-off between liquidity and delay in the Norwegian Interbank Clearing System and Canada’s LVTS, respectively. A related set of papers quantify the benefit of introducing one or more LSMs to an existing RTGS system. Norman (2010) provides an overview of several such studies, reporting estimated liquidity savings of 20% in the Korean BOK-Wire+, and 15% in the Japanese BOJ-Net. Denbee and McLafferty (2012) employ a similar approach to estimate potential liquidity savings of 30% from introducing an LSM in the UK’s Clearing House Automated Payment System (CHAPS). This paper will contribute to the literature by simulating alternative system designs using data from the LVTS, and quantifying the potential benefits of implementing one or more LSMs.

The remainder of this paper is organized as follows. Section 2 provides a brief overview of the LVTS and how it differs from an RTGS system. Section 3 provides an overview of the methodology and assumptions as well as a description of the various LSMs that can be simulated in FNA. Section 4 details the results. Finally, section 5 concludes and discusses some of the policy implications.

## 2 Intraday liquidity in the LVTS and RTGS

### 2.1 A brief overview of the LVTS

The LVTS is an essential part of the Canadian financial system, processing an average of approximately 39,000 payments equivalent to $187 bn each day in 2018. As of January 2020, there are 15 financial institutions and the Bank of Canada participating directly in the LVTS. In addition, there are non-participant partners, to whom LVTS participants provide payment agent services through contractual arrangements.

The LVTS comprises two separate payment streams: Tranche 1 (T1) and Tranche 2 (T2). Participants are free to submit payments to either stream, subject to the risk controls and collateral requirements of each. T1 is very similar to an RTGS system because any net debit position incurred by an LVTS participant must be fully backed by eligible collateral. T2, on the other hand, is unique. Participants grant bilateral credit limits (BCLs) to one another, which establish the largest net debit position a grantee can incur against a grantor. Moreover, a participant cannot incur a multilateral net debit position in T2 greater than its "Tranche 2 Net Debit Cap" (T2NDC), which is equal to the sum of all BCLs granted to it, multiplied by the system-wide percentage

---

3 See Heijmans and Heuver (2012) for an overview of the simulation literature.

4 While each stream is characterized by its own risk controls and collateral requirements, settlement of the entire system is effected on a multilateral net basis at the end of the cycle.

5 For example, if Participant A grants a BCL of $100 to Participant B, Participant B can incur a net debit position with Participant A no greater than $100.
Participants are also required to apportion collateral to T2, but, unlike in T1, do not have to back the full value of their T2NDC with eligible collateral. Rather, the value of collateral each participant is required to apportion to T2 is equal to the largest BCL it chooses to grant any other LVTS participant, multiplied by the SWP.⁶

Throughout the day, payments that pass the applicable risk control are considered final and irrevocable; final cash settlement of the LVTS, however, does not take place until the end of the day and is effected on a multilateral net basis.⁷ In the event a single LVTS participant is unable to cover its end-of-day LVTS multilateral net position, there is sufficient collateral in the LVTS to settle the largest possible default of any participant. In the event of multiple participant defaults, however, the Bank of Canada will provide a guarantee of settlement by acting as an unsecured creditor for the residual amount, once the defaulters’ own collateral, as well as the shared collateral in T2, is exhausted. As part of the LVTS replacement, the Bank of Canada will no longer provide this guarantee. Moving to an RTGS settlement model will ensure all credit exposure is fully backed by system participants, thereby eliminating the need for the Bank of Canada’s guarantee.

Today, the majority of LVTS payments are sent through T2 because of its collateral savings for participants relative to T1. Figure 1 shows the amount of collateral in each settlement stream on a daily basis for the year 2018. The average amount of collateral in T1 is three times greater than the amount in T2. However, T1 payments account for only 1.1% and 26.6% of the total volume and value, respectively, of LVTS payments in 2018.

---

⁶ The SWP is currently set at 0.3. See LVTS Rule 2 (Section 2.3): https://www.payments.ca/sites/default/files/21-Aug-17/lvts_rule_2_eng.pdf

⁷ A participant’s T2 collateral is referred to as its Maximum Additional Settlement Obligation (Max ASO), and will be used to settle the LVTS in the event of a default. A participant’s Max ASO represents the largest financial loss it would incur in the event one or more LVTS participants default.

⁸ Intraday finality is achieved through novation netting. Once a payment has passed the applicable risk control, the bilateral net payment position is extinguished and replaced by a multilateral settlement obligation of the sending LVTS participant with all other participants. It is these multilateral net positions that LVTS participants settle at the end of the day.
2.2 Intraday liquidity and settlement delay in an RTGS

Moving to an RTGS system will present significant changes for LVTS participants’ intraday liquidity and collateral requirements. Participants in any large-value payment system rely on intraday liquidity to make payments, the need for which they can meet in different ways. One way participants in both the LVTS and RTGS systems can make payments is to re-use the liquidity generated by incoming payments from other participants.\textsuperscript{9} There is, however, a potential timing mismatch between settlement banks’ outgoing and incoming payments that can generate the need for additional intraday liquidity, the amount of which depends on a participant’s ability to rely on incoming payments as a source of liquidity (see e.g., Bech and Garratt, 2003; Kaliontzoglou and Müller, 2015; Heijmans and Heuver, 2014). The most significant change from the LVTS will be in how participants meet this need for additional intraday liquidity.

This additional intraday liquidity is not provided free of charge by the central bank, but in the form of intraday credit backed by eligible collateral.\textsuperscript{10} Today in the LVTS, a participant’s T2NDC determines the amount of

---

\textsuperscript{9} In the LVTS, receiving a payment will decrease a participant’s multilateral net debit position (or increase its multilateral net credit position), relative to its limit.

\textsuperscript{10} Almost unanimously, central banks have agreed that free intraday credit is not a viable policy because it incentivizes the overuse of credit for which the central bank (i.e., taxpayers) is the ultimate guarantor (Bech and Garratt, 2003). Some central banks also provide intraday credit to settlement banks for a fee, but neither the LVTS nor its future replacement will rely on priced intraday credit.
intraday credit available to that participant in T2. This intraday credit, unlike in T1, is not backed in full by eligible collateral. Rather, it is typically much greater than the amount of collateral an LVTS participant is required to apportion to T2. On average, there is approximately $29.1 bn of intraday credit available across LVTS participants in T2 on each day, but just $5.2 bn in collateral required.\footnote{This is by design. Recall that a participant’s T2NDC, or intraday liquidity, is determined by the sum of all BCLs granted to it by other LVTS participants. A participant’s collateral requirement, however, is determined by just the largest BCL it chooses to grant to any other participant.}

In an RTGS system, on the other hand, individual payments are settled through an immediate transfer of funds, meaning they must be fully covered beforehand. As a result, many central banks require that all intraday credit be backed in full by eligible collateral (Bech et al., 2008). As an illustration, the $29.1 bn of intraday credit available to LVTS participants in T2 would need to be backed by $29.1 bn of collateral in an RTGS system. Due to the opportunity cost of posting this collateral, intraday liquidity is potentially much more costly in an RTGS system.\footnote{As Arjani and McVanel (2006) notes, the LVTS is a more efficient design than an RTGS system from a cost-minimization perspective. LVTS participants, though, are exposed to the risk that a counterparty defaults on their end of day obligations and their T2 collateral is used by the Bank of Canada to settle the defaulter’s final net debit position. This risk is eliminated in an RTGS system. However, Bewaji (2018) argues that LVTS participants factor this counterparty credit exposure into the value of their BCLs, which drives participants to maintain as low an exposure as possible.}

This increased cost of collateral can incentivize banks to rely on incoming payments to a far greater extent, as they are a costless source of liquidity (Angelini, 1998). In an effort to do so, settlement banks may choose to purposely delay payments to other settlement banks. Delayed payments can, however, impose their own costs on settlement banks. Delays in settling time-critical payments can result in financial or reputational penalties.\footnote{Certain payments, such as those to ancillary clearing and settlement systems, may be subject to penalties if not submitted by a certain deadline. Consistent delays in settling customer payments may also damage a bank’s reputation, leading to a loss of goodwill and future business (Bech and Garratt, 2003).} Moreover, delaying payments can impose an externality on the rest of the system, as each participant tends to rely on incoming liquidity to some degree, leading to further delays, and in extreme cases, gridlock. As a result, each settlement bank in an RTGS system faces a trade-off between the costs of intraday liquidity and the costs of settlement delay.

This trade-off can impede the smooth functioning of an RTGS system, which can in turn pose significant risks to the financial system (CPSS, 2012). Settlement banks in the Fedwire Funds Transfer service, for example, have been observed to concentrate their payments in the late afternoon, in an attempt to coordinate their payment activity and avoid more costly liquidity sources. These long delays can increase operational and settlement risk (McAndrews and Rajan, 2000). The settlement of a large number of payments during a short window can place a greater burden on the system’s processing capabilities than continuous settlement of payments throughout the day. Moreover, the uncertainty with regards to the settlement of a large proportion of payments close to the end of the day can be deleterious to the federal funds market. Too much liquidity usage can also pose problems. Failing to manage its available intraday liquidity prudently can leave a settlement bank short of the incoming liquidity it needs to continue making payments in the event the bank or one of its counterparties are stressed, or market conditions change (Ball et al., 2011).
2.3 Liquidity-saving mechanisms

There is a range of tools available to help settlement banks manage their intraday liquidity effectively. Many banks, including some LVTS participants, use internal schedulers and other internal risk management tools. Policy-makers and system operators can encourage prudent intraday liquidity management through rules and regulations. RTGS systems are also increasingly being designed with one or more specific features or tools available for participants that are intended to improve liquidity efficiency. These tools, referred to collectively as LSMs, are intended to make coordination between incoming and outgoing payments easier, encouraging the smooth flow of payments throughout the day and minimizing settlement banks’ liquidity needs.

In most cases, there are three key elements of an effective set of LSMs: a central queue, offsetting algorithms, and liquidity-reservation functionality (Ball et al., 2011). Introducing a central queue to the settlement process is perhaps the most important LSM, as it directly supports the functioning of other commonly-used LSMs.14 Allowing payments to queue automatically when there isn’t enough liquidity available for settlement encourages settlement banks to manage their payments through a central queue, rather than individual queueing arrangements. The liquidity-saving benefits that are derived from a central queue generally happen in one of two ways. The first is through offsetting queued payments between two or more participants; the second involves encouraging greater liquidity recycling. As a coordination device for settlement banks, a central queue can identify liquidity recycling and offsetting opportunities that may be difficult to achieve through individual coordination alone (CPSS, 2005).

Settlement banks, though, may prefer to ensure their intraday liquidity is reserved for settling time-critical payments, and to allow only those that are not time-sensitive to queue.15 The absence of a mechanism to manage their liquidity accordingly can discourage the use of the central queue. Implementing a design feature that allows settlement banks to reserve some of their liquidity for certain payments can help mitigate this. For example, participants can choose to reserve a certain portion of their liquidity for time-sensitive payments, and allow another, likely smaller, portion of their liquidity to be used for settling less urgent payments, either gross or through offsetting.

Figure 2 illustrates the trade-off between liquidity and delay in an RTGS system with and without an LSM. In both cases, settlement banks can reduce the amount of intraday liquidity they hold, but only at the cost of delays in settlement. In the absence of an LSM, reductions in intraday liquidity usage can lead to long delays in settlement. This relationship is represented by the solid line in Figure 2, along which each combination of liquidity and delay presents the same total cost for a given system design and payments that must be settled. Introducing an LSM will shift the entire trade-off curve, represented by the dotted line in Figure 2. Implementing an LSM, such as a central queue with an offsetting algorithm, with which banks can manage their liquidity should incentivize timely submission of payments, leading to greater coordination and reducing the amount of settlement delay for a given level of liquidity. Participants are better off at any point in this case because their total costs, i.e., the costs of settlement delay and intraday liquidity taken together, are lower. Some banks may even translate this reduction in delay to increased liquidity savings. The optimal set of LSMs will shift this trade-off curve as close to the origin as possible, and represent the most efficient system design.

14 In RTGS systems with no central queue, payments submitted without sufficient liquidity for settlement are typically rejected.
15 An example of payments that are considered very time-critical are pay-ins to Continuous Linked Settlement (CLS). CLS is a financial institution that provides settlement services to its members in the foreign exchange market and connects large-value payment systems worldwide in various time zones.
3 Methodology

3.1 Simulation methodology

The trade-off between the costs of intraday liquidity and settlement delay for settlement banks is formalized below. As profit-maximizing agents, banks prefer to minimize their total costs involved in settling payments. Each settlement banks’ cost minimization problem is:

\[
\begin{align*}
\min_L \quad & C = f(L, D, \theta, Y) \\
\text{s.t.} \quad & Y = g(\phi) \\
& D = h(L)
\end{align*}
\]  

where settlement banks’ total costs, \( C \), depend on the amount of intraday liquidity borrowed against eligible collateral, \( L \), and the resulting level of settlement delay, \( D \). Settlement banks’ total costs also depend on exogenous factors, such as the design of the RTGS system and whether one or more LSMs are available. These features are captured by \( \theta \). The value and volume of payments that must be settled in a given day, \( Y \), either on behalf of the settlement bank’s customers or proprietary operations, and the timing of these payments, is also assumed to be exogenous. In the present case, \( Y \) is captured by LVTS transaction data, where \( g(\phi) \) represents the data generating process assumed to capture the economic activity of the bank and its customers. This economic activity determines the value and volume of payments that must be settled in a given day, as well as the urgency with which a payment must be settled. We discuss this further in section 3.4.
We account for the trade-off between intraday liquidity and settlement delay by the constraint $D = h(L)$.\footnote{The exact functional form of these costs are beyond the scope of this paper, however, Bewaji (2018) provides a more detailed model of participant behaviour under the current LVTS settlement model in T2, including detailed functional forms for LVTS participants’ costs.} Assigning an amount of intraday liquidity, $L$, available to each LVTS participant in the simulations will produce a resulting settlement delay, $D$. Simulating a range of intraday liquidity levels will allow us to observe whether the expected relationship between liquidity and delay does indeed hold empirically, and examine its characteristics. Further, simulating system designs with different LSMs will allow us to investigate the impact of those LSMs on the trade-off between intraday liquidity and settlement delay.

We restrict our analysis to the range of intraday liquidity levels between zero and what is commonly referred to as the “upper bound” of intraday liquidity. If each bank had no intraday liquidity available upon which to draw, the system would be deadlocked, and every payment would be delayed the longest amount of time possible. At the other extreme is the upper bound of liquidity, which is the amount of intraday liquidity each bank would need to hold in order to settle every payment immediately upon submission. Any intraday liquidity above the upper bound would, in theory, be unnecessary as it would result in no reduction in settlement delay.\footnote{In practice, having this exact amount of intraday liquidity available at the start of the day will likely prove difficult for a number of reasons. Not all payments that will need to be made over the course of a given day are known at the start of the day. Partly because of this, settlement banks may also prefer to have a buffer of intraday liquidity available for precautionary purposes.} Accordingly, an LSM will only have an impact within this range of liquidity.

A settlement bank’s upper bound of liquidity is equal to their largest cumulative net debit position over the course of the day:

\[
L_{i,d}^u = \min\left[0, \sum_{t=0}^{T} (P_{j,i}(t) - P_{i,j}(t))\right] \quad (2)
\]

where $L_{i,d}^u$ represents the largest cumulative net debit position of each bank, $i$, on each day, $d$, and $P_{i,j}(t)$ represents a payment of value $P$ from bank $i$ to bank $j$, submitted, and in this case settled, at time $t$. Not every settlement bank will incur a net debit position at some point intraday, and may be able to rely entirely on payments received from other settlement banks. In this case, it is possible for a settlement bank’s upper bound intraday liquidity to be zero. Using LVTS transaction data, we can calculate each participant’s upper bound of liquidity for every day in our sample of data.

To establish a trade-off between intraday liquidity and settlement delay, the same sample of LVTS data, in other words holding $Y$ constant, is simulated with less than the upper bound of intraday liquidity available to each LVTS participant, thereby causing the settlement of some payments to be delayed. We let $L_{i,d}(t)$ denote the intraday liquidity available to bank $i$ at time $t$. The amount of intraday liquidity available to each LVTS participant at the beginning of the day, $t = 0$, is calculated as:

\[
L_{i,d}(0) = \alpha L_{i,d}^u, \quad 0 < \alpha < 1 \quad (3)
\]

The entire sample of LVTS data is simulated for each level of $\alpha$. For example, when $\alpha = 0.95$, each day is simulated with every LVTS participant having 95% of the amount of intraday liquidity that would be required to settle all payments immediately on that day. We vary $\alpha$ uniformly across LVTS participants in 5% increments for each participant on each day in the sample.

In order to understand how different LSMs can impact the trade-off between liquidity and delay, we change $\theta$ by...
introducing an LSM, and simulate the entire sample of LVTS data again for each $\alpha$. We first test each individual LSM, before introducing a combination of LSMs.

### 3.2 Data

The data used for the simulations consists of transaction data from the LVTS, which includes the date and time of each payment as well as its value, the counterparties involved, and whether the payment was settled in T1 or T2. The sample consists of all 252 days during 2018 on which the LVTS was operational, representing more than 9.5 million individual payments worth approximately $45.6 tn.

### 3.3 Liquidity-saving mechanisms

The baseline set of simulations is based on a "first-in first-out" (FIFO) queueing arrangement. The FIFO principle presents an additional condition for settling a payment—not only does the sending bank require sufficient liquidity on hand, but its queue must also be empty. FIFO also impacts the sequence in which queued payments are released from a participant’s queue when incoming liquidity becomes available. The payment at the top of the queue is released when there is sufficient liquidity, and only once this first queued payment has settled is the next queued payment considered for settlement. FIFO queueing arrangements can lead to long delays in settlement as large-value transactions at the top of the queue block the settlement of subsequent lower-value transactions. To address this problem, a number of LSMs have been developed and implemented in RTGS systems around the world, including alternative queueing arrangements and offsetting algorithms (CPSS, 2005). The LSMs we implement in the simulations are described in detail below.

#### 3.3.1 Queueing arrangement

A widely-used alternative to the strict FIFO principle is "FIFO-Bypass", which we will refer to simply as Bypass. This queueing arrangement allows the settlement of payments to bypass a strict FIFO ordering. Payments in the sending participant’s queue are not considered—the sending participant must only have sufficient liquidity at the time the payment is submitted for settlement. As a result, the condition for settlement of a payment is simplified to:

$$ P_{i,j}(t) \leq L_i(t) $$

(4)

Implementing a Bypass queueing arrangement will also have an impact on the conditions under which queued payments are released when incoming funds become available. Consider an incoming payment from bank $j$ at time $b$, which increases bank $i$’s liquidity, such that $L_i(b) > L_i(a)$, where $a < b$. The system will now check each of bank $i$’s queued payments against its new liquidity position to see if they can be released for settlement. This condition is given by (5), where $q \in Q_i$ indicates a payment’s position in the ordered set of all bank $i$’s queued payments, $Q_i$. As implied by the term Bypass, payments that do not meet this condition can be bypassed, and the next queued payment, $q + 1$, is tested.

---

18 FIFO-Bypass is used in RITS (Australia), TARGET2 (Eurozone), BOK-Wire+ (Korea), BOJ-Net (Japan), and RIX (Sweden). See Fugal et al. (2018).
\[ P_{i,j}^q(t) \leq L_i(b) \]  

(5)

### 3.3.2 Bilateral offsetting

One type of offsetting algorithm implemented in RTGS systems is a bilateral offsetting algorithm.\(^{19}\) When a bilateral offsetting algorithm is implemented, the system will search for an offsetting payment from the intended recipient before submitting a payment for gross settlement. Consider a payment from bank \(i\) to bank \(j\). The system will first search for a queued payment from bank \(j\) to bank \(i\) that can be offset.\(^{20}\) If the two payments do not offset perfectly, i.e., \(P_{i,j}(t) \neq P_{j,i}^q(t)\), the system will attempt to use liquidity from either bank to settle the two payments. The algorithm will search for the first payment to bank \(i\) in bank \(j\)’s queue, and attempt to offset the two payments subject to:

\[
\begin{align*}
L_i(t) &\geq P_{i,j}(t) - P_{j,i}^q(t), \quad \text{if } P_{i,j}(t) > P_{j,i}^q(t) \\
L_j(t) &\geq P_{j,i}^q(t) - P_{i,j}(t), \quad \text{if } P_{i,j}(t) < P_{j,i}^q(t)
\end{align*}
\]  

(6)

where \(\hat{q} \leq q \forall q \in Q_j\).

### 3.3.3 Multilateral offsetting

Another type of offsetting algorithm, designed to offset payments between as many banks as possible, is a multilateral offsetting algorithm. The particular algorithm varies widely by system, but we use the Bech-Soromäki algorithm, which is available in FNA. The algorithm works by calculating the multilateral net position of each bank, were all queued payments to be offset, taking into account each bank’s available liquidity. If these positions cannot be settled because one or more banks have insufficient liquidity, the algorithm will choose any such bank and remove the latest queued payment from the solution, and attempt to offset the remaining payments.\(^{21}\) The process repeats until a set of payments that can be offset is identified, or all queued payments have been removed.\(^{22}\)

While bilateral offsetting can only be run continuously in the FNA platform, the frequency with which the multilateral offsetting algorithm runs is not fixed. In RTGS systems that use multilateral offsetting, there is a wide range of frequencies at which these algorithms are run.\(^{23}\) Running the multilateral offsetting as frequently

---

\(^{19}\) Bilateral offsetting is widely used in RTGS systems around the world. RITS (Australia), TARGET2 (Eurozone), BOK-Wire+ (Korea), BOJ-Net (Japan), SPEI (Mexico), and MEPS+ (Singapore) run a bilateral offsetting algorithm continuously, while RIX (Sweden), SIC (Switzerland), and CHAPS (UK) run an algorithm at frequent intervals. See Fugal et al. (2018).

\(^{20}\) There are various ways in which the offsetting algorithm can search queued payments from bank \(j\) to find an applicable offset in FNA. We limit our analysis to a search method called "First", which tests the first payment to bank \(i\) in bank \(j\)’s queue. This search method can break FIFO ordering because it searches for only the first payment to bank \(i\), rather than the first payment to any bank, as in FIFO. As a result, this bilateral offsetting algorithm achieves some of the benefits of Bypass in addition to offsetting.

\(^{21}\) The choice of bank from which to remove queued payments does not influence the final solution.

\(^{22}\) The full details are available in Bech and Soromäki (2001).

\(^{23}\) Fugal et al. (2018) provides an overview of existing centralized central queueing mechanisms in major countries with an RTGS system. BOJ-NET (Japan) runs a multilateral offsetting algorithm just 4 times a day. TARGET2 (Eurozone), SPEI (Mexico), and MEPS+ (Singapore) run a multilateral offsetting algorithm continuously throughout the day. CHAPS (UK) and BOK-Wire+ (Korea) run less frequently, at 2 and 30 minute intervals, respectively.
as possible produced the best results when tested, and is further supported by the fact that TARGET2 (Eurozone), SPEI (Mexico), and MEPS+ (Singapore) all run a multilateral offsetting algorithm continuously (Fugal et al., 2018).

### 3.3.4 Liquidity reservation

There are different ways in which the ability for participants to reserve liquidity for time-critical payments can be achieved in practice. One such way is for the system to operate multiple separate settlement streams, each with its own settlement account. However, the presence of multiple settlement streams can increase the role of participant behaviour in determining the levels of risk, cost, and efficiency of the system (CPSS, 2005). Even with multiple settlement streams, for example, some participants may still prefer to actively manage their intraday liquidity in a single settlement account. Many RTGS systems with a central queue offer control features that make it easier for participants to manage liquidity this way, such as the possibility to change the location of a payment in the queue, prioritize the release of certain payments, or set bilateral and multilateral limits to control the outflow of liquidity (CPSS, 2005).

To obtain more robust estimates of intraday liquidity requirements and the impact of implementing one or more LSMs, we simulate multiple scenarios that account for different participant behaviour with respect to settling time-critical payments. In the first scenario, we model an RTGS system with multiple settlement streams. One settlement stream, which we refer to as the LSM stream for simplicity, will offer liquidity savings at the cost of potential delays in settlement. A second settlement stream, the RTGS stream, will also be available to participants to reserve certain liquidity for higher priority payments that must settle quickly. In this scenario, participants send all time-critical payments to the RTGS stream, and less urgent payments to the LSM stream. We vary only the amount of intraday liquidity that participants assign to settling less urgent payments, holding constant the amount of liquidity participants need to settle time-critical payments with no delay.

In the second scenario, we model an RTGS system with a single settlement stream, extending the work of Embree and Taylor (2015). To account for time-critical payments in this situation, we employ the "force settlement" functionality of the FNA platform. Payments can be marked for forced settlement, meaning they will be settled immediately and in full, ignoring any liquidity constraints. The participant’s liquidity constraint remains unchanged, however, meaning non time-critical payments will still be subject to the constraint and, therefore, potential delays. The liquidity used to force settle payments will be accounted for in that participant’s intraday liquidity usage, and will also serve as incoming liquidity for the recipient. The FNA platform allows payments to be marked for forced settlement by any field in the transaction data, such as time, value, intended recipient, etc.

### 3.4 Assumptions and limitations

In order to carry out these simulations, we must make certain assumptions regarding LVTS participants’ future behaviour. First, we assume that T1 payments are time-critical. As discussed in section 2.1, the design of the 24 For processing purposes it was not feasible to run the algorithm continuously, but at 30 second intervals instead.

25 The BOJ-NET system in Japan and the RIX system in Sweden operate with multiple settlement accounts. See Fugal et al. (2018) for an overview.
LVTS can help reveal participants’ preferences with respect to the costs involved in settling payments. The cost of sending payments in T1 is much higher than in T2 because of the collateral requirements. As such, only a small number of high-value payments are made in T1 each day. We assume that, in order to incur this cost, the settlement of T1 payments must be considered time-critical, either by the participant itself or its client.

Second, the intraday timing of payments is assumed to be exogenous. As discussed in section 2.2, participants in a RTGS system may prefer to submit their payments in such a way that ensures other banks will also be sending payments during that time. It is possible that participant behaviour changes drastically in either of our scenarios. However, the opportunity for liquidity recycling and payment offsetting in the LSM stream should, in principle, incentivize participants to continue to submit their payments in a timely manner. Moreover, certain LVTS payments are particularly time-sensitive and will likely not change.

Lastly, participants’ intraday liquidity in the simulations is assigned as a proportion of their upper bound liquidity requirements. Participants, however, generally do not know the amount of intraday liquidity they will require over the course of the day, as not all payments that must be made are known to a bank at the start of the day. In practice, intraday liquidity would not be apportioned across participants in as optimal a way as it is in the case of the upper bound of liquidity, and the trade-off curves that could realistically be achieved would present slightly worse-off combinations of intraday liquidity and settlement delay.

Together, these assumptions present limitations to the analysis, and the results should serve only to provide insight on the potential implications for intraday liquidity and settlement delay.

### 3.5 Measures of intraday liquidity and settlement delay

Each simulation will produce a record of each individual payment—when it was submitted, when it was settled, and how it was settled, along with a range of other output statistics. Using these outputs, we calculate the following set of measures.

#### 3.5.1 Intraday liquidity usage

The amount of intraday liquidity used by each participant is calculated as the largest cumulative net debit position incurred by that participant on a given day in the sample. Intraday liquidity usage, $U_d$, is calculated for the entire system by summing each individual LVTS participant’s largest cumulative net debit position:

\[
U_d = \sum_{i=1, i\neq j}^{N} \left[ \min[0, \sum_{t=0}^{T} (P_{j,i}(t) - P_{i,j}(t))] \right]
\]  

#### 3.5.2 Settlement delay

Settlement delay is measured by calculating the difference between the time a payment was submitted for settlement and the time it actually settled. To capture settlement delay, we introduce $\tau$ as a payment’s settlement

---

26 Payments are rarely queued in the LVTS, making it more difficult to discern participants’ settlement delay preferences. In fact, LVTS participants are encouraged to submit only those payments that will not be queued. The majority of payments that do queue are T2 payments, indicating a greater cost of settlement delay in T1. Moreover, LVTS participants have also indicated anecdotally that T1 payments are time-critical, particularly compared to T2 payments.
time. A payment between banks $i$ and $j$ that was submitted at time $t$ and settled at time $\tau$ is denoted by $P_{i,j}(t, \tau)$. For $t = \tau$ there is no settlement delay, while for $t < \tau < T$ settlement of the payment was delayed by $\tau - t$.

We take a measure of the average settlement delay across all payments on a given day in the sample, weighted by their relative value. The value-weighted average settlement delay is given by:

$$V_d = \frac{\sum_{i=1, i \neq j}^{N} \sum_{t=0}^{T} (\tau - t) P_{i,j}(t, \tau)}{\sum_{i=1, i \neq j}^{N} \sum_{t=0}^{T} P_{i,j}(t, \tau)}$$  \hspace{1cm} (8)

### 3.5.3 Unsettled transactions

Not all payments will necessarily be settled in the simulations. Each bank’s available intraday liquidity is set at the beginning of the day, and there is no change other than as a result of incoming or outgoing payments, or the additional liquidity required for time-critical payments. Were a bank approaching the end of the day, $T$, with one or more unsettled payments, it could in reality acquire additional liquidity. The absence of active intraday liquidity management in these simulations does present a limitation to the analysis, as it is very unlikely LVTS participants would tolerate unsettled transactions. Measuring the unsettled transactions in the simulations can, however, provide a measure of efficiency. If introducing an LSM reduces the amount of unsettled transactions in the simulations, the system is more efficient, as it can settle more payments with a given amount of available intraday liquidity.

We measure the value of transactions that are submitted at time $t$ but remain unsettled at the end of the day, $T$, as a proportion of the total value of payments submitted during the course of the day:

$$R_d = \frac{\sum_{i=1, i \neq j}^{N} P_{i,j}(t, T)}{\sum_{i=1, i \neq j}^{N} P_{i,j}(t)}$$  \hspace{1cm} (9)

### 4 Results

The results of the simulations are presented in sections 4.1 and 4.2, where two trade-off curves are presented under both the multiple and single settlement stream assumptions. Each figure presents a trade-off curve under five distinct combinations of the LSMs presented in section 3.3. We compare the results under our base RTGS system with no LSMs, labelled "FIFO", to the addition of each individual LSM. These individual LSM trade-off curves are labelled "Bypass", "Bilateral offsetting", and "Multilateral offsetting". Finally, we compare these trade-off curves to a system design with all previous LSMs combined, labelled "All LSM".

We present the results based on simulation averages for our measures of intraday liquidity usage, settlement delay, and unsettled transactions:
\[
U_{\alpha} = \frac{\sum_{d=1}^{D} U_d}{D}
\]
(10)

\[
V_{\alpha} = \frac{\sum_{d=1}^{D} V_d}{D}
\]
(11)

\[
R_{\alpha} = \frac{\sum_{d=1}^{D} R_d}{D}
\]
(12)

where \( D = 252 \), the total number of days in our sample of LVTS data.

4.1 Multiple settlement streams

Figure 3 shows the trade-off between intraday liquidity and value-weighted settlement delay for scenario 1, in which participants submit T1 payments to an RTGS settlement stream and T2 payments are submitted to an LSM settlement stream. The intraday liquidity required to settle all payments with no settlement delay is $26.2 bn on average, which consists of $8.7 bn to settle T1 payments and $17.5 bn to settle T2 payments. Intraday liquidity for LVTS participants is only reduced in the LSM stream, varying from $16.6 bn down to $0.9 bn, while it is held constant at $8.7 bn in the RTGS stream. The results are shown as the total of both settlement streams.

The results show that with multiple settlement streams, participants face a significant trade-off between intraday liquidity usage and settlement delay. To settle all payments quickly throughout the day, intraday liquidity of more than $20 bn is needed on average. As participants reduce their intraday liquidity usage, settlement delay increases dramatically, up to an average of more than five hours when no LSMs are implemented. The results also show that implementing LSMs does indeed shift the trade-off curve. Moreover, while each individual LSM shifts the trade-off curve a similar amount, combining all LSMs is by far the most effective. By implementing multiple LSMs, participants are presented with a more efficient trade-off between intraday liquidity and settlement delay, allowing them to reduce their total costs.

Figure 4 shows the trade-off between intraday liquidity and unsettled transaction value for scenario 1. The results are very similar to those shown in Figure 3; as participants reduce their intraday liquidity, more and more payments are left unsettled at the end of the day. LSMs are also effective at shifting this trade-off curve, but combining all LSMs again presents the most efficient design. While Bypass and Bilateral offsetting each have a similar impact on unsettled transactions, Multilateral offsetting has a unique impact, resulting in relatively little change from FIFO at low levels of liquidity and more significant benefits at higher levels of liquidity, approaching a similar value of unsettled transactions as the combination of all LSMs.

The results for scenario 1 show that LSMs have relatively little impact at higher levels of intraday liquidity on both value-weighted settlement delay and unsettled transactions. As intraday liquidity is reduced further, past approximately $17.5 bn, the benefits of an LSM gradually increase, having the greatest impact at low levels of
intraday liquidity. At the lowest level of intraday liquidity, the combination of LSMs can reduce the average settlement delay from more than five hours to less than three, a reduction of 45%. The LSM stream approaches deadlock at the lowest liquidity level, with more than 80% of payments remaining queued at the end of the day, but the combination of LSMs also has a significant impact here, reducing this number to less than 50%. These benefits are, of course, relative, as an average settlement delay of three hours and 50% of payments queued at the end of the day are not desirable outcomes. What these results do serve to illustrate is that the benefit of implementing an LSM in this scenario is increasing in the amount of intraday liquidity savings.

The results also illustrate the effectiveness of LSMs at all levels of intraday liquidity. Though they are most effective at low levels of intraday liquidity, the trade-off curves under each system design with an LSM dominate the FIFO trade-off curve. On their own, each LSM improves the trade-off curve from FIFO, but results in a relatively similar reduction in settlement delay. Together, however, the combination of all LSMs presents a significant improvement over FIFO.
4.2 Single settlement stream

Figure 5 shows the trade-off between intraday liquidity and value-weighted settlement delay for scenario 2, in which all LVTS payments are submitted to a single LSM settlement stream. The intraday liquidity required to settle all payments with no settlement delay is, on average, $19.4 bn. Accordingly, the intraday liquidity available to the system is varied from $18.4 bn down to just less than $1 bn. Participants’ intraday liquidity usage is much greater than $1 bn, however, as time-critical payments must be settled immediately. In addition to the $1 bn available at the start of the day, participants require additional liquidity of between $10.7 bn and $11 bn, depending on the system design. Participants, on average, require additional intraday liquidity in all cases, except for the highest level of intraday liquidity assigned, $18.4 bn.

The settlement delay reduction benefits are clearly illustrated in this scenario, but are less drastic than with multiple settlement streams. In this scenario, the combination of multiple LSMs still presents the most efficient trade-off, but there is a clearer ranking of individual LSMs in terms of their impact. Bilateral offsetting appears to be the most effective at reducing settlement delay for a given amount of intraday liquidity, followed by Multilateral offsetting and Bypass. There is also less of an overall improvement from the worst trade-off (FIFO) to the best trade-off (All LSM) than in the first scenario. This suggests there is an implicit benefit to managing liquidity in a single settlement stream. Indeed, the upper bound of liquidity is approximately $6.8 bn lower when participants use a single settlement stream in place of multiple settlement streams. This result is rather intuitive, as opportunities for liquidity recycling are lost when payments are settled in separate settlement mechanisms. Comparing the two scenarios allows us to quantify the substantial impact of this liquidity recycling benefit.

Figure 6 demonstrates an interesting relationship between intraday liquidity and unsettled transactions in this
Figure 5: The effect of LSMs on value-weighted average settlement delay.

scenario. Notably, there are fewer unsettled transactions at any level of liquidity and under any system design, with the highest level being just over 15%. In terms of unsettled transaction value, Bypass performs relatively worse than the other LSMs, and even worse than FIFO for a range of intraday liquidity. At first, the fact that FIFO is more effective than Bypass at settling transactions seems counterintuitive given the previous results. Though this only represents a part of the curve, it is notable—there is relatively little else to take away from this figure given the relative efficiency of each system design. It can be explained by the fact that a Bypass queueing arrangement is effective at reducing long delays that are, under FIFO, caused by large-value transactions at the front of the queue. Whereas under FIFO, a participant must accumulate sufficient liquidity with which to settle this large-value transaction, under Bypass a participant may never accumulate this amount of liquidity, constantly using its intraday liquidity for lower-value transactions. This can lead to long delays in settlement for some large-value transactions that result in their being unsettled by the end of the day. This fact may also help explain why Bypass is the least effective individual LSM in Figure 5.

Another interesting feature of Figure 6 is the fact that Multilateral offsetting appears to be as effective alone as it is when combined with Bilateral offsetting and Bypass. Given the low value of unsettled transactions overall in this scenario, and the effectiveness of all LSMs at reducing settlement delay, this fact does not lead us to conclude that Multilateral offsetting alone is a viable alternative to the combination of LSMs.
Given the replacement of the LVTS, this paper investigates the potential for LSMs to reduce the amount of intraday liquidity that participants will need to hold in an RTGS system. We use the LSMs available in the FNA simulation engine, which include FIFO-Bypass queueing, Bilateral offsetting, and Multilateral offsetting. In addition to simulating each individual LSM, we test a combination of all available LSMs. We also simulate two different scenarios based on participant behaviour with respect to time-critical payments. First, we simulate T1 payments, which are assumed to be time-critical, and T2 payments in separate settlement streams, where LSMs are only available in the T2 stream. Second, we simulate all payments in a single settlement stream with LSMs.

Our results demonstrate that there is a clear benefit from implementing LSMs in both scenarios. The introduction of LSMs, whether individually or in combination, results in a clear improvement to the trade-off curves under each of our assumptions. For all levels of intraday liquidity that were simulated, the system design that incorporated all available LSMs reduced both settlement delay and unsettled transactions, demonstrating clear benefits for the management of liquidity and settlement risk. Though participants are ultimately responsible for managing these risks, LSMs can provide risk-reduction benefits that participants will likely not be able to achieve through individual coordination. Moreover, the effectiveness of these LSMs have been shown to be robust to different participant behaviour regarding time-critical payments. These results suggest that a combination of LSMs should strongly be considered when designing the LVTS replacement.

The results also show, however, that participant behaviour will play a significant role in determining the levels of risk, cost, and efficiency that are achievable under any RTGS system design. We expected the results for our two measures, settlement delay and unsettled transactions, to differ significantly between scenarios. It is clear...
from our results that they do, in fact, differ—both measures are lower across the entire trade-off curve in the single settlement stream scenario. It is not immediately clear, however, which scenario is preferable from the perspective of minimizing total costs, because the costs of settlement delay are not evenly distributed across payments. Some payments may face a very high cost of settlement delay. The ability to reserve liquidity allows participants to ensure the delay cost for these payments is as close to zero as possible, but at the cost of increased liquidity requirements and the costs associated with managing that liquidity more actively. The actual costs are ultimately only observable to the individual participants, however. As the system operator, Payments Canada can only attempt to present participants with the most cost-efficient system possible within certain parameters, such as an RTGS settlement model. The results suggest that, in order to achieve the efficiency of a single settlement stream, there may be a need for further tools, such as the ability for participants to set payment priorities or to manage their queued payments actively throughout the day.
References


