CIRRELT

Centre interuniversitaire de recherche sur les réseaux d'entreprise, la logistique et le transport

Interuniversity Research Centre on Enterprise Networks, Logistics and Transportation

Replenishment and Denomination Mix of Automated Teller Machines with Dynamic and Stochastic Demands

Lieke M. van der Heide
Leandro C. Coelho
Iris F.A. Vis
Roel G. van Anholt

August 2017

CIRRELT-2017-47

Document de travail également publié par la Faculté des sciences de l'administration de l'Université Laval, sous le numéro FSA-2017-009.

Bureaux de Montréal :
Université de Montréal
Pavillon André-Aisenstadt
C.P. 6128, succursale Centre-ville
Montréal (Québec)
Canada H3C 3J7
Téléphone : 514 343-7575
Télécopie : 514 343-7121

Bureaux de Québec :
Université Laval
Pavillon Palasis-Prince
2325, de la Terrasse, bureau 2642
Québec (Québec)
Canada G1V 0A6
Téléphone : 418 656-2073
Télécopie : 418 656-2624

www.cirrelt.ca

Replenishment and Denomination Mix of Automated Teller Machines with Dynamic and Stochastic Demands

Lieke M. van der Heide[1,*], Leandro C. Coelho[2], Iris F.A. Vis[1], Roel G. van Anholt[3]

[1] Faculty of Economics and Business, University of Groningen, The Netherlands

[2] Interuniversity Research Centre on Enterprise Networks, Logistics and Transportation (CIRRELT) and Department of Operations and Decision Systems, 2325 de la Terrasse, Université Laval, Québec, Canada G1V 0A6

[3] Geld Service Nederland, The Netherlands

Abstract. In the inventory management of automated teller machines (ATMs) many activities affect the total costs such as forecasting, replenishments and the denomination mix used. The denomination mix is the combination of bills used to fulfill a customer's demand. Given a good demand forecast for the demands at an ATM, the challenge is to determine replenishments and denomination mix strategies. In this paper, we propose a time-varying denomination mix strategy, which is validated by benchmarking it against the case of a bank's denomination mix strategy. The bank's predetermined strategy typically consists of a least note strategy. Our alternative, the time-varying denomination mix, allows adjustments to the denomination mix over time. In both strategies we simultaneously optimize denomination mixes and replenishment decisions. We define the problem and solution policies as mixed integer formulations and solve them via a rolling horizon algorithm using different frequencies of denomination mix updates, rolling horizon lengths, numbers of ATMs, costs, and forecast qualities. By implementing the time-varying denomination mix the operational costs of managing an ATM can be reduced by 7.6% or Euro 46 per ATM per month on average, which can represent over Euro 3.5 million per year in the Netherlands only.

Keywords. Sustainability, multi-tier supply chain management, integrated optimization.

Acknowledgements. We thank the contacts at our industry partner for providing the data and validating our parameters. This project was partly funded by the Natural Sciences and Engineering Research Council of Canada (NSERC) under grant 2014-05764. This support is greatly acknowledged.

---

* Corresponding author: l.m.van.der.heide@rug.nl

Dépôt légal – Bibliothèque et Archives nationales du Québec
Bibliothèque et Archives Canada, 2017

1 Introduction

This paper considers the inventory management problem of determining the denomination mix that arises in an automated teller machine (ATM) network. Common decisions faced in ATM research are the replenishment frequency and quantities of each of the denominations. These can be varied in order to minimize replenishment, holding and lost-sales costs. Replenishment costs are those incurred because ATMs need to be replenished, such as fuel costs and driver wages. The holding costs can represent missed interest that could not be made from the money stored in the ATM. Lastly, we refer to lost-sales as the costs occurring when an ATM cannot satisfy the requested amount due to depletion of one or more denominations. This is not necessarily a complete stock-out, since other demanded withdrawals could still be available.

In order to fulfill a withdrawal, a denomination mix is necessary. This is the combination of bills used to fulfill a customer's demand. For instance, when a customer requests €50, there are several possible denomination mixes. One bill of €50 could be used, two bills of €20 and one of €10, 5 bills of €10, or any other combination totaling €50. Which of these is chosen influences the inventory of the ATM and, subsequently, the replenishment quantities and frequency as well. Usually this denomination mix is assumed to be predetermined by the bank for each value a customer could demand. However, we expect that varying the denomination mix can have a significant impact on all three types of costs. If, for instance, an ATM always runs out of €50 notes, it could be beneficial to satisfy some demands with notes of €10 instead. This would then also reduce lost-sales costs, on top of decreasing replenishment costs. These are lost-sales since a demand is refused if the necessary denomination mix is not in stock, even though another combination of bills could have been used. The decrease in the number of replenishments could also reduce the overall inventory by allowing an ATM to operate for a longer period without replenishments.

In this paper, we propose a model taking into account each of these aspects and analyze the effect of denomination mixes on the different types of costs and inventory. Varying denomination mixes have not yet been extensively covered, even though Paul and Mukherjee [9] mention that determining the optimal mix of denominations to satisfy demands is one of the biggest problems banks are facing. By having several options, one can skew the demand for a particular denomination leading to short supply, and postpone or even reduce the frequency of visits to ATMs, significantly reducing costs.

Inventory and distribution management for ATMs have been jointly studied as an inventory-routing problem (IRP) [3]. IRPs have been explicitly applied to ATM management [6, 8, 11]. In addition to the IRP, demand forecasting for ATMs has been covered in, for instance, Van Anholt [11] and Vanketesh et al. [12]. Moreover, the combination of these two aspects has been covered in Ekinci et al. [5], where time intervals between two replenishments are determined using a forecast dependent on clusters of nearby ATMs. Furthermore, inventory management in the case of lost-sales has been covered extensively [2].

Policies similar to varying the denomination mix have been used in different fields with promising results. For instance, in Bassok et al. [1] a similar form of firm-driven substitution of different products is used. In that case, the benefits of allowing substitution are greater when the demand variability is high and substitution costs are low.

Determining the replenishment quantities while simultaneously determining the denomination mixes is comparable to the lot-sizing problem with a flexible bill-of-materials, which is considered in Lang [7]. However, in the case of ATMs many different withdrawal sizes (products in lot-sizing terminology) are considered with perfect substitution, meaning that there is no preferred combination of bills, since the monetary value is the same. Moreover, the bill-of-materials is not completely flexible, but instead an optimal denomination mix is determined and used for a certain amount of time, after which it can be changed. Therefore, firm-driven substitution at ATMs has not been covered sufficiently in current research, even though it appears beneficial to do so.

We have partnered with the company Geldservice Nederland (GSN), the operator responsible for replenishing ATMs for the three major banks in the Netherlands. Typically, banks use a least note strategy to fulfill demands, i.e., the smallest possible number of bills is used to satisfy a demand. According to our industry partner, more inefficient strategies were deployed until recently such as including at least one of the smallest denominations in the denomination mix. Moreover, optimal replenishment plans are often not determined. In order to estimate the costs of the bank's activities, we consequently underestimate their operational costs by computing an optimal replenishment plan under the least notes strategy, which is the best strategy that is deployed in practice. This strategy is called the predetermined denomination mix, which is used as a benchmark for our proposed strategy: the time-varying denomination mix. This strategy consists of an optimal denomination mix based on the forecast data for the next periods, which is allowed to vary periodically. Using the withdrawal data GSN has collected at over 400 ATMs for a period of three months, we compare the two different denomination mix strategies.

A rolling horizon algorithm is designed and implemented to benchmark a bank's predetermined denomination mix strategy against the time-varying one. An optimal replenishment strategy is derived by simultaneously determining for each period both the denomination mix and the replenishments that result in the lowest total cost. After computing an optimal replenishment strategy for the next periods, the denomination mix is updated. The time-varying denomination mix can be adjusted by the frequency of changes allowed to the denomination mix. The performance of these two policies will be extensively compared, while an optimal replenishment strategy is also implemented.

The remainder of this paper is organized as follows. In Section 2 we formally describe the problem at hand, and in Section 3 the mathematical formulations are provided. In Section 4 we introduce in detail our solution algorithm, which is based on a rolling horizon framework. Section 5 is devoted to the computational experiments to assess the performance of our methods on real data from an industry partner in the Netherlands, including robustness and sensitivity analyses. Finally, Section 6 is devoted to our conclusions.

2 Problem description

A bank's predetermined denomination mix is typically based on the least notes strategy, which uses the smallest possible number of notes to fulfill a withdrawal request. For this strategy, inventory management policies need to determine only the timing and quantities of the replenishments for each denomination. We propose a time-varying denomination mix policy that allows the denomination mix for a certain demand to vary over time. The aim is to design an inventory management policy that is determined by simultaneously computing the denomination mix, replenishment timings and replenishment quantities. The objective is to minimize the total costs, consisting of replenishment costs, holding costs and lost-sales costs.

We assume, as is realistic in the Netherlands, that all bills are delivered in packs of size p. Moreover, it is assumed that all customers arrive after the replenishments of that day. Therefore, the calculation of the holding costs is the value of the bills that remain in the ATM after these demands, multiplied by the interest rate h. There is a fixed costs c corresponding to a single replenishment and a lost sale occurs if the denomination mix necessary to fulfill the demand is not available in the ATM. We assume that the denomination mix cannot be varied within a day. Moreover, a lost sale cannot occur to prevent later lost-sales or to reduce the number of replenishments. For instance, refusing to satisfy demands would reduce the replenishment costs. To prevent this, lost-sales costs M are implemented.

In current ATM technology, all denominations are available in separate boxes. Finally, in order to find a sustainable solution, each planned schedule must start and end with at least the historical average number of bills in inventory for each of the used denominations.

The problem is formally defined over a set $\mathcal{I} = \{1, \ldots, m\}$ of nodes, each representing an ATM location. Even though we assume that ATMs do not influence each other, we consider multiple ATMs simultaneously, as in practice the replenishment strategies for many ATMs need to be determined. Each ATM contains a set $\mathcal{J} = \{1, \ldots, J\}$ of boxes storing notes of a single denomination b_{ij}, $i \in \mathcal{I}, j \in \mathcal{J}$. Each box has a maximum capacity Q_{ij}. Moreover, the number of bills in box j of ATM i at the beginning of the planning horizon is defined as e_{ij}. The historical average number of bills in box j of ATM i is v_{ij}. At each ATM i there are withdrawals for a number of values V_i, $i \in \mathcal{I}$. These withdrawals have monetary values l_{in}, $i \in \mathcal{I}, n \in \mathcal{N} = \{1, \ldots, V_i\}$. The forecast number of withdrawals at ATM i, for value n at time t is a^t_{in}, $i \in \mathcal{I}, n \in \mathcal{N}, t \in \mathcal{T} = \{D, \ldots, D+T\}$, where $D \geq 1$ is the current period and T is the length of the planning horizon. Important input for all models are the forecast data instead of the actual data and, therefore, they try to minimize the forecast costs. We explain in Section 4 how we have calculated the actual costs using withdrawal data after the mathematical model was solved.

3 Mathematical models

In this section we present the mathematical models used to formulate the bank's predetermined denomination mix and the time-varying denomination mix that will be evaluated in Section 5. First, we describe the part of the model that is the same for both denomination mix strategies. Subsequently, in Section 3.1 we present the additional constraints specific for the bank's predetermined denomination mix, and in Section 3.2 the additional constraints for our proposed the time-varying denomination mix.

All models require the following decision variables: let y^t_i be a binary variable indicating whether ATM i is visited at time t and w^t_{ij} be a binary variable indicating whether box j of ATM i was replenished at time t. The number of bills delivered to box j of ATM i in period t is q^t_{ij}. Let also I^t_{ij} be the number of bills that remains in box j of ATM i from day t to day $t+1$. Moreover, let x^t_{ij} be an integer representing the number of bills lacking in box j of ATM i at time t due to lost-sales. Integer variables f^t_{ij} ensure that all deliveries will be made in packs of size p. Lastly, g_{ij} are the variables representing the minimum number of bills in box j of ATM i at time $D+T$.

All the subsequent models are based on the following:

$$\min \sum_{t\in\mathcal{T}} \sum_{i\in\mathcal{I}} \left(y_i^t c + \sum_{j\in\mathcal{J}} \left(hI_{ij}^t b_{ij} + Mx_{ij}^t \right) \right) \tag{1}$$

subject to

$$w_{ij}^t \leq y_i^t \qquad \forall i \in \mathcal{I}, j \in \mathcal{J}, t \in \mathcal{T} \tag{2}$$

$$q_{ij}^t \leq w_{ij}^t Q_{ij} \qquad \forall i \in \mathcal{I}, j \in \mathcal{J}, t \in \mathcal{T} \tag{3}$$

$$q_{ij}^t \leq Q_{ij} - I_{ij}^{t-1} \qquad \forall i \in \mathcal{I}, j \in \mathcal{J}, t \in \mathcal{T} \tag{4}$$

$$q_{ij}^t = pf_{ij}^t \qquad \forall i \in \mathcal{I}, j \in \mathcal{J}, t \in \mathcal{T} \tag{5}$$

$$I_{ij}^T \geq g_{ij} \qquad \forall i \in \mathcal{I}, j \in \mathcal{J} \tag{6}$$

$$w_{ij}^t, y_i^t \in \{0, 1\} \qquad \forall i \in \mathcal{I}, j \in \mathcal{J}, t \in \mathcal{T} \tag{7}$$

$$q_{ij}^t, f_{ij}^t, g_{ij} \in \mathbb{N} \qquad \forall i \in \mathcal{I}, j \in \mathcal{J}, t \in \mathcal{T} \tag{8}$$

$$I_{ij}^t, x_{ij}^t \in \mathbb{R}^+ \qquad \forall i \in \mathcal{I}, j \in \mathcal{J}, t \in \mathcal{T}. \tag{9}$$

The objective function (1) minimizes the total costs composed of replenishment, holding and lost-sales costs. Coefficient M should be sufficiently large to prevent lost-sales unless there is no other option. Constraints (2) ensure that boxes can only be replenished if their ATM is visited. Constraints (3) and (4) impose the number of bills used to replenish box j. Constraints (3) ensure that the number of bills that is used to replenish a box is less than its maximum capacity, and is zero if the box was not filled. Moreover, constraints (4) state that the number of bills cannot be greater than the maximum capacity minus the current inventory. This allows active replenishment and prevents the number of bills in a box from exceeding its capacity. Constraints (5) ensure all deliveries are made in packs of p bills. Constraints (6) ensure a sustainable solution by imposing a minimum inventory, the historical average, at the end of the planning horizon. Constraints (7)–(9) define the domain and nature of the variables.

3.1 Bank's predetermined denomination mix

We now present the model that defines the bank's predetermined denomination mix which will be used as a baseline for our comparison. Implementing this denomination mix means that a demand for €50 is fulfilled by one bill of €50. Therefore, the denomination mixes for each value that can be withdrawn are known and fixed and consequently parameters in the model. Let o_{ijn} determine how many bills from box j of ATM i are included in the denomination mix for value $n \in \mathcal{N}$. Note that o_{ijn} depends on i because the number of demanded values can differ per ATM, since different types of bills are available. Note also that the set $\mathcal{N}$ contains all possible demanded values that can be withdrawn. We then model the bank's predetermined denomination mix by (1)–(9) and by:

$$I_{ij}^{t} = I_{ij}^{t-1} + q_{ij}^{t} - \sum_{n \in \mathcal{N}} (o_{ijn} a_{in}^{t}) + x_{ij}^{t} \qquad \forall i \in \mathcal{I}, j \in \mathcal{J}, t \in \mathcal{T} \qquad (10)$$

$$g_{ij} = \min(v_{ij}, Q_{ij} - \sum_{n \in \mathcal{N}} (o_{ijn} a_{in}^{T}) - p + 1) \qquad \forall i \in \mathcal{I}, j \in \mathcal{J} \qquad (11)$$

$$I_{ij}^{D} = e_{ij} \qquad \forall i \in \mathcal{I}, j \in \mathcal{J}. \qquad (12)$$

Constraints (10) define the inventory conservation. They state that the inventory in period t equals the overnight inventory in period $t-1$, minus the required bills to fulfill demands from period t, plus the bills delivered in period t, plus the lost-sales quantity. The number of bills necessary to fulfill demands is defined as the number of demands for a certain value times the predetermined denomination combination. Constraints (11) ensure that at the end of the planning horizon the inventory of each box is the minimum of the average and the maximum possible number of bills in box j. Moreover, constraints (12) are added to ensure that the initial inventory is correctly considered.

To linearize constraints (11), we will use two auxiliary variables s_{ij}^{+} and s_{ij}^{-} as follows:

$$g_{ij} \geq v_{ij} - s_{ij}^{+} - s_{ij}^{-} \qquad \forall i \in \mathcal{I}, j \in \mathcal{J} \tag{13}$$

$$g_{ij} \geq Q_{ij} - \sum_{n \in \mathcal{N}} o_{ijn} a_{in}^{T} - p + 1 - s_{ij}^{+} - s_{ij}^{-} \qquad \forall i \in \mathcal{I}, j \in \mathcal{J} \tag{14}$$

$$s_{ij}^{+} \geq v_{ij} - (Q_{ij} - \sum_{n \in \mathcal{N}} o_{ijn} a_{in}^{T} - p + 1) \qquad \forall i \in \mathcal{I}, j \in \mathcal{J} \tag{15}$$

$$s_{ij}^{-} \geq Q_{ij} - \sum_{n \in \mathcal{N}} o_{ijn} a_{in}^{T} - p + 1 - v_{ij} \qquad \forall i \in \mathcal{I}, j \in \mathcal{J} \tag{16}$$

$$s_{ij}^{+}, s_{ij}^{-} \in \mathbb{N} \qquad \forall i \in \mathcal{I}, j \in \mathcal{J}, \tag{17}$$

and penalize s_{ij}^{-} and s_{ij}^{+} in the objective function. We also derive the following valid inequalities:

$$\sum_{t \in \mathcal{T}} q_{ijt} \geq g_{ij} - e_{ij} + \sum_{t \in \mathcal{T}} \sum_{n \in \mathcal{N}} o_{ijn} a_{in}^{t} - \sum_{t \in \mathcal{T}} x_{ijt} \qquad \forall i \in \mathcal{I}, j \in \mathcal{J}. \tag{18}$$

Inequalities (18) state that the number of bills delivered in the whole planning horizon is the number of bills that should be in inventory at the end of the planning horizon, minus how many were in the box at the beginning of the period, plus the number of bills needed to fulfill demands, minus the lost-sales variables.

3.2 Time-varying denomination mix

We now present the model for our proposed time-varying denomination mix. This model takes into consideration a number T of periods ahead which are used to determine the best denomination mix to minimize the total costs. This model for the optimal denomination mix uses the same variables, parameters and basic model previously defined in Sections 2 and 3. Moreover, a new set of variables o_{ijn} is defined, which specifies how many bills of box j of ATM i are included in the denomination mix for value n. New binary parameters u_{ij} are defined and indicate whether box j of ATM i is unused, meaning whether the bills from that box are not allowed in the denomination mixes. The model for the time-varying denomination mix is then defined by (1)–(9) and by:

$$\sum_{j\in\mathcal{J}} o_{ijn} b_{ij} = l_{in} \qquad \forall i \in \mathcal{I}, n \in \mathcal{N} \tag{19}$$

$$I_{ij}^{t} = I_{ij}^{t-1} + q_{ij}^{t} - \sum_{n\in\mathcal{N}} (o_{ijn} a_{in}^{t}) + x_{ij}^{t} \qquad \forall i \in \mathcal{I}, j \in \mathcal{J}, t \in \mathcal{T} \tag{20}$$

$$g_{ij} = v_{ij} \qquad \forall i \in \mathcal{I}, j \in \mathcal{J} \tag{21}$$

$$I_{ij}^{D} = e_{ij} \qquad \forall i \in \mathcal{I}, j \in \mathcal{J} \tag{22}$$

$$q_{ij}^{D+1} > -e_{ij} - u_{ij} \qquad \forall i \in \mathcal{I}, j \in \mathcal{J} \tag{23}$$

$$o_{ijn} \in \mathcal{O} = \{0, \ldots, O\} \qquad \forall i \in \mathcal{I}, j \in \mathcal{J}, n \in \mathcal{N}. \tag{24}$$

Constraints (19) ensure that the denomination mix multiplied by the values of the bills that are used in the mix equals the requested value. Constraints (20) ensure inventory conservation. Constraints (21) ensure that at the end of the planning horizon the inventory equals the average number of bills again to ensure a sustainable solution. Constraints (22) guarantee that the initial inventory is taken into consideration. Constraints (23) ensure that if a box is empty at the beginning of the planning horizon, it must be replenished the first period if it is to be used. We include it to ensure that the optimal denomination mix is not affected because a box is empty. Lastly, constraints (24) ensure that optimal denomination combinations do not use too many of the same bills.

We also derive the following valid inequalities:

$$\sum_{t\in\mathcal{T}} q_{ijt} \geq g_{ij} - e_{ij} + \sum_{t\in\mathcal{T}} \sum_{n\in\mathcal{N}} o_{ijn} a_{in}^{t} - \sum_{t\in\mathcal{T}} x_{ijt} \qquad \forall i \in \mathcal{I}, j \in \mathcal{J}. \tag{25}$$

Constraints (25) state that the number of bills delivered in the whole planning horizon is the number of bills that should be in inventory at the end of the planning horizon, minus how many were in the box at the beginning of the period, plus the number of bills necessary to fulfill demand, minus the lost-sales variables.

4 Solution algorithm

We have implemented the models described in Sections 3.1 and 3.2 using a rolling horizon method, meaning that in each period we plan replenishments for the next T periods, we implement the decisions made for the first period, add one period to the end of the planning horizon, update demands and inventories, and solve the problem again for the next T periods. In some of the periods the denomination mixes are also determined for the next L periods and implemented. We use the rolling horizon paradigm for three main reasons: (i) to plan using information for more than one period ahead, which entails a better overall solution; (ii) forecast data may not be available or be reliable for many periods in advance; (iii) the models are too time-consuming to be solved for planning horizons that are too long. Each of these reasons is later validated in our experiments.

Sethi and Sorger [10] state that forecasting future events is costly and that these costs may depend on when the forecast is made and how far ahead they are in the future. Therefore, implementing a rolling horizon method would be less costly than scheduling the complete time horizon. In research concerning similar problems, such as that of Coelho et al. [4], the rolling horizon method was effectively implemented as well.

Since the rolling horizon is implemented using forecast data, the actual costs incurred and the forecast costs that are found in the optimization of the models have to be separated. The actual costs are calculated a posteriori after each iteration of the rolling horizon algorithm. This is done by implementing the replenishments that were scheduled for the first day and subtract all actual demand from the inventory using the appropriate denomination mix, yielding the actual inventory. The actual costs are then calculated by adding the replenishment costs for the scheduled replenishments to the holding costs times the actual inventory. Lastly, lost-sales costs due to all actual demands that were not satisfied are added.

In the case of the bank's predetermined denomination mix, the constraints described in Section 3.1 are updated at each iteration. Subsequently, the replenishments that were scheduled for the first day are implemented and the actual demands are satisfied using the least notes denomination mix. We then compute the actual costs for this first day based on the actual demands. Before starting the next iteration the necessary parameters are updated, namely e_{ij} and a_{in}^t. It is important to note that the parameter e_{ij} represents the number of bills in box j of ATM i after the first day of actual demands and scheduled replenishments are implemented.

In the case of the time-varying denomination mix, the model described in Section 3.2 is only solved once every L iterations in order to determine a denomination mix. After these iterations, the scheduled replenishments for the first period are implemented and the actual demands are satisfied using the denomination mix that was optimized for the subsequent T periods. An overview of the algorithm is provided in Figure 1. If the current L periods have finished, we solve the model presented in Section 3.2 and if not, we solve the model presented in Section 3.1 using the previously found denomination mix. Note that again these models are solved using forecast data. Once these models are solved, actual daily demands are revealed and costs are calculated based on the scheduled replenishments and actual lost-sales. Moreover, the current inventory of the ATMs is updated. This is repeated each period for L periods, after which both the denomination mixes and the replenishments are reoptimized. The process iterates until the end of the overall planning horizon T.

5 Computational experiments

In this section we report the results of extensive computational experiments to evaluate the benefits of varying the denomination mix compared to the bank's least note strategy.

We use two datasets of actual withdrawals obtained from GSN, the operator responsible for the replenishment of the ATMs of the three major banks in the Netherlands. The first dataset contains all the demands at 300 ATMs of the standard configuration for a three month period. In the standard configuration, ATMs hold notes of €10, €20 and €50.

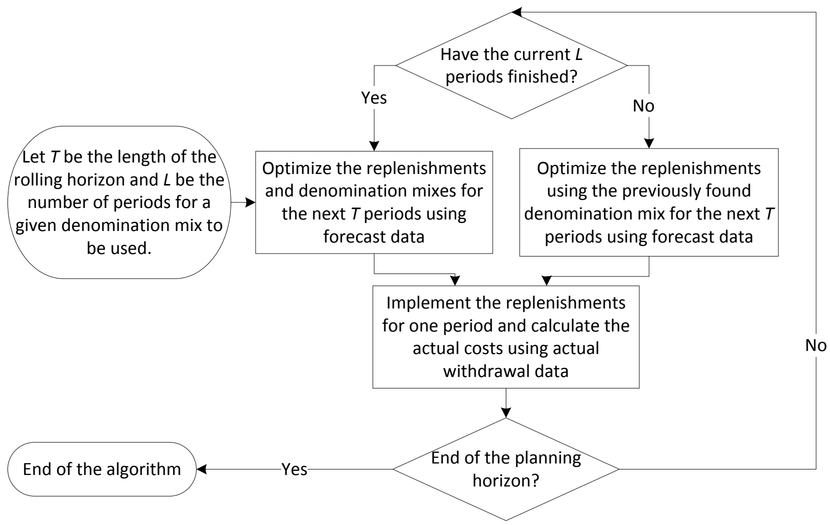

Figure 1: Overall scheme of the time-varying algorithm

The other dataset contains all demands for these months at 108 ATMs that also contain €5 notes.

The forecasts of demands for these days were obtained according to Van Anholt [11] and are denoted a_{in}^t. The corresponding monetary values l_{in} are also derived from these forecasts. This provides a well-founded base for the forecasts that we use to plan replenishment strategies. These forecasts proved to be robust, on average being only 0.49 demands off for the standard ATM, and 0.46 for ATMs containing notes of €5. For robustness purposes we also compare the bank's predetermined denomination mix against the time-varying denomination mix using different forecasts in Section 5.3.

These two datasets were used to create five instances containing 50 ATMs and five with 100 ATMS of both configurations, and five instances of 200 ATMs of the standard configuration. We consider these amounts of ATMs simultaneously in order to determine if the model is useful in practice, where many ATMs have to be considered. Each of these instances contains all the demands for the first 30 days. In order to gain meaningful results, valid values for the interest rate h, the replenishment costs c, the capacity of one box Q_{ij}, the pack size p and the average number of bills in the ATM if its type is used, v_{ij}, were provided by GSN and their corresponding values are shown in Table 1. Per ATM we model five boxes, the first one containing notes of €5, the second of €10, the third of €20, the fourth of €50 and the final box contains €100. When a type of bill is not used, the capacity of that box is set to 0. In reality each ATM contains four boxes, therefore at least one capacity is set to 0. Three of the boxes are filled with notes of €10, €20 and €50. The content of the last box is varied in the experiments and contains notes of either €5, €20, €50, or €100.

Some experiments to investigate the difference between using the bank's predetermined denomination mix and the time-varying denomination mix are defined. Five different values of T, the length of the rolling horizon, are considered namely: 1, 3, 7, 15 and 30 days. Moreover, we change the denomination mix of the time-varying denomination mix once every $L = 1$, 3, 7, 15 or 30 periods. Combining these configurations with the different lengths of the rolling horizon T and the implementation times L, we have run a total of 1650 experiments.

Table 1: Parameter setting values obtained from our industrial partner GSN

Parameter	h	c	Q_{ij}	p	v_{i1}	v_{i2}	v_{i3}	v_{i4}	v_{i5}
Values	0.05	100	2000	100	933	858	975	1.395	551

These experiments were run on a server using 92 GB of RAM implemented in C++, using CPLEX 12.6 from IBM Concert Technology. We imposed limits on the maximum time allowed for each part of the solution algorithm. We ensured that the iterations in which a new denomination mix had to be computed took no more than 10 hours altogether, since this is the most time-consuming part of the algorithm. Moreover, the iterations in between were allowed to take up to 4 hours altogether. These time restrictions were implemented to enable practical use.

5.1 Comparison of bank's predetermined and time-varying denomination mixes

We now report the results of extensive experiments run in order to evaluate the difference between the bank's predetermined denomination mix and that of our time-varying one. We start by assessing the impact of different rolling horizons T and implementation times L, for all numbers of ATMs considered. For now we only consider boxes containing notes of €10, €20 and €50, which is the standard configuration in the Netherlands. The costs of different configurations with other bills will be evaluated later.

Table 2: Costs per ATM (standard deviation) in € of implementing the different replenishment and denomination mix strategies using standard configuration

	Bank's	Time-varying denomination mix (L)				
T	mix	1	3	7	15	30
1	3140.56	3256.36	3262.99	3302.14	3303.20	3305.78
	(53.51)	(29.52)	(27.43)	(31.41)	(29.26)	(30.80)
3	637.63	556.99*	763.62	947.55	963.26	1046.95
	(18.96)	(14.03)	(22.85)	(44.10)	(65.68)	(97.12)
7	650.20	669.56	614.02*	633.24*	657.18	655.34
	(17.92)	(131.07)	(26.62)	(14.20)	(13.20)	(18.21)
15	602.99	1048.1	662.03	615.05	586.78*	579.78*
	(16.81)	(217.36)	(60.05)	(41.72)	(23.11)	(19.12)
30	608.96	1113.53	1084.51	981.55	827.93	754.52
	(20.68)	(215.14)	(224.44)	(280.54)	(208.19)	(46.95)

* indicates significant improvement over the bank's mix at 1% level.

In Table 2 we show the average monthly costs per ATM of implementing the bank's strategy and the time-varying denomination mix for several values of T and L. We used paired t-tests to identify when the costs of the time-varying denomination mix were significantly lower than the bank's mix using the same rolling horizon T. From this table it is clear that a short rolling horizon T, i.e., considering only a few periods ahead for the replenishment plan, is extremely costly; these costs decrease as more information from the future is taken into consideration. However, when a rolling horizon of 30 days is implemented, the quality of the solutions decreases. This can be partly explained by the fact that a time limit was imposed on the solution algorithms, and those were truncated while computing the solutions. Since the algorithm is more time-consuming if more data is used, the 30-day forecast runs were cut short more often. This was especially true when 200 ATMs were considered. Similarly, when L is smaller, the denomination mix has to be determined more frequently, which is also more time-consuming. Therefore, small values of L are only beneficial if a small rolling horizon is used. Reasoning from Table 2, we find that even when the denomination mix is only updated every month, it can still outperform the bank's predetermined denomination mix.

It is clear from Table 2 that the time-varying denomination mix outperforms the bank's predetermined denomination mix for some combinations of rolling horizon T and the update frequency of the denomination mix L. Moreover, when comparing these solutions, we find that using a 3-day forecast and changing the denomination mix daily costs only €556.99 per ATM per month. This is the least costly solution of the time-varying denomination mix and is found using forecast information for only three days ahead. This solution is 7.6% or €46 per ATM less costly than the most efficient solution of the bank's predetermined denomination mix, which is based on a 15-day rolling horizon and costs €602.99 per ATM per month. This reduction in costs can be decomposed into a reduction of 3.1% in replenishment costs and of 11.6% in holding costs. While lost-sales costs increase, they amount to less than 0.23% of the total operational cost. This combination of a rolling horizon of 3 days and changing the denomination mix daily remains an optimal choice even for the different numbers of ATMs considered simultaneously.

There are a few downsides to the implementation of the time-varying denomination mix. First of all, it is more time-consuming than simply using the bank's predetermined denomination mix. Specifically, the best solution of the bank's predetermined denomination mix takes on average 827 seconds, whereas the time-varying denomination mix takes 6295 seconds. However, it still allows practical use. Moreover, there is a significant increase in the amount of lost-sales when the time-varying denomination mix is implemented. The least costly solution of the bank's predetermined denomination mix includes only €495.10 lost-sales per month per ATM on average, while €969.50 lost-sales per month per ATM occurred when the optimal setting of the time-varying denomination mix is implemented. However, the costs corresponding to these lost-sales remain less than 1% of the total production costs. We also note that the increased running time remains fully operational and applicable in practice.

5.2 Configurations of ATMs

We now consider the different configurations of the boxes within an ATM. As mentioned before, an ATM consists of four different boxes that can each contain one denomination. In the Netherlands the standard configuration contains only €10, €20 and two boxes with €50. We compare configurations with €10, €20, €50 and vary the contents of the fourth box, so that it either contains €5, €20, or €100 bills.

Firstly, let us focus on using notes of €5, for which the results are presented in Table 3. It is clear that this configuration is the most costly option, with the bank's predetermined denomination mix costing at least €1098.02 and the smallest value for the time-varying denomination mix being €1016.48. These costs were obtained using a different dataset that included demands for €5 and is, for this reason, not included in the comparison in the previous section. This dataset contains higher demands than the previous one, which also partly explains the higher costs.

When the holding, replenishment and lost-sales costs are considered separately, it becomes clear that the replenishment costs are the main reason for the increase in costs compared to other configurations. The smallest replenishment costs when the bank's predetermined denomination mix is implemented are €870, whereas they are only €255 when the standard configuration is used. The same relation occurs when the time-varying denomination

Table 3: Costs per ATM (standard deviation) in € of implementing the different replenishment and denomination mix strategies, considering ATMs containing €5

	Bank's	Time-varying denomination mix (L)				
T	mix	1	3	7	15	30
1	3317.65	3313.63	3346.39	3397.55	3398.02	3393.77
	(41.29)	(37.96)	(35.82)	(26.47)	(29.21)	(28.67)
3	1397.98	1550.44	1524.94	1611.36	1573.69	1566.98
	(56.11)	(209.09)	(109.98)	(80.59)	(63.516)	(43.15)
7	1150.57	1846.51	1536.58	1278.83	1047.74**	1016.48**
	(59.86)	(245.95)	(277.74)	(205.04)	(47.81)	(44.05)
15	1098.02	2604.13	2084.34	1496.26	1587.85	1055.28*
	(54.20)	(617.05)	(419.89)	(327.58)	(400.61)	(89.02)
30	1162.71	3546.10	2846.08	2752.43	2120.90	2071.86
	(80.09)	(348.23)	(585.37)	(606.80)	(793.94)	(806.09)

* indicates significant improvement over the bank's mix at 5% level, ** at 1% level.

mix is implemented: €778 when €5 was used and €266 when the standard configuration was used. Conversely, the holding costs are lower at only €193 compared to €319 when the bank's predetermined denomination mix was used and €200 compared to €263 when the time-varying denomination mix was used. Therefore, it becomes clear that the most influential costs when notes of €5 are considered are the replenishment ones. In any case, our proposed method can significantly reduce all costs.

We now focus on the configuration where the fourth box contains notes of €20. Here the smallest costs when the bank's predetermined denomination mix is used are €679.79, while when the time-varying denomination mix was used the total costs were at least €607.10, meaning a 10.7% reduction in costs. These minimum costs of the time-varying denomination mix were found using a 3-day forecast and allowing the denomination mix to change daily, always within limits on running times.

When notes of €100 were available, the lowest costs for the bank's predetermined denomination mix were €693.81, whereas the lowest costs using the time-varying denomination mix were €623.38, meaning a 10.15% reduction. The lowest costs using the time-varying denomination mix were obtained when a 15-day forecast was used and the denomination mix changed only once every 30 days.

Overall, we find that both the time-varying denomination mix and the bank's predetermined mix never perform as well as when the standard configuration is used. Note that including notes of €5 almost doubles the total costs. However, the time-varying denomination mix does decrease the total costs compared to the bank's mix for these different configurations up to 10.7%. Therefore, even when this method would be implemented on an ATM using a different configuration, our time-varying denomination mix reduces costs significantly.

5.3 Robustness analysis

In this section some robustness checks are considered, to ensure that the reduction in costs when the time-varying denomination mix is used remains when different parameters are used. The parameters that will be varied are the number of ATMs considered, the different costs, the forecast quality and the time limits imposed on the algorithms. We now assess the difference in total costs per ATM between finding a solution for 50, 100 or 200 ATMs.

When all configurations except the one containing notes of €5 are taken into account, we find that there are quite a few combinations of forecast and implementation time, for which there is a significant difference between the costs per ATM when 50, 100 or 200 ATMs are considered. This is partly explained by the fact that in all cases the same time limit was implemented.

When we consider the standard configuration in particular, we find that the costs are significantly different as well. It seems that determining the optimal strategy for only 50 ATMs produces the lowest costs per ATM, compared to using 100 or 200 ATMs. Due to this difference, the same combinations of implementation and forecast time do not always produce costs that are significantly smaller than the bank's predetermined denomination mix. For instance, using a 15-day forecast and allowing the denomination mix to change every 7 days produces costs that are significantly smaller than the bank's predetermined denomination mix when only 50 ATMs were considered. However, when 100 ATMs were considered the costs of this combination are no longer smaller than the bank's predetermined denomination mix, since its costs are 5.2% higher than when 50 ATMs were considered. These costs were increased by 14.5% when 200 ATMs were considered. This is due to the fact that the algorithm could not finish within the time limit when more ATMs are considered and, therefore, higher costs were produced. When 200 ATMs are considered, the complete algorithm takes on average 36461 seconds to finish, while it only takes 3739 seconds when 50 ATMs were considered.

Even though there are significant differences between using different numbers of ATMs, the optimal strategy for the time-varying denomination mix, using a 3-day forecast and changing the denomination mix daily remains optimal across the different numbers of ATMs.

The interest rate, replenishment costs and lost-sales costs have each been separately doubled and halved for the combinations of forecast length and rate of changing the denomination mix that were a significant improvement over the bank's predetermined denomination mix using different numbers of ATMs. We find that all these combinations still yield total costs lower than when the bank's predetermined denomination mix was implemented, even when the number of ATMs was varied. This proves that our method of selecting denomination combinations and replenishment schedules is robust with respect to the costs parameters.

Moreover, the quality of the forecast on which the replenishments are based has also been varied. We subtract 0.5 demands for each value on each day from the number of demands that is expected due to the forecast that was used. When the replenishments were scheduled using these adjusted forecasts, we find that the time-varying denomination mix yields lower costs than the bank's predetermined denomination mix, unless 15 days of forecast were used and the denomination mix changes every 15 days. However, when 0.5 demands were added to the number of demands for each value on each day the time-varying denomination mix is never an improvement over the bank's predetermined denomination mix for the combinations we consider. This is partially due to the fact that the forecasts that were used already slightly overestimate the number of demands. On average, the forecast is 0.028 demands overestimated when compared to the actual demands. The impact of this finding is that underestimating the demand does not have a negative impact on the final solution, however, overestimating it forces the algorithm to plan more replenishments and higher inventory levels, which significantly affects the total costs.

Finally, we also evaluate how the performance and quality of the solutions are affected by the allotted running time. Some combination of forecast and denomination changes that did not yield a significant reduction in costs when compared to the bank's predetermined denomination mix are now allowed to run for twice as long as before. This means that all the time limitations that were mentioned before are now doubled to 20 hours when a new denomination mix is determined and 8 hours for all other iterations. These new limits are not completely without reach of practical use. These new settings were implemented using the standard configuration of an ATM and for the following combinations of forecast time and number iterations the denomination remains the same. For 50 ATMs we considered (f, L) equal to (15,1), (30,3), (30,7), (30,15), (30,30); for 100 ATMs we used (15,1), (15,3), (15,7), (30,15), (30,30); and for 200 ATMs we examined (7,1), (15,3), (15,7), (15,15), (30,30). Even though the time limit is doubled in most cases, the solutions still do not achieve lower costs than that of the bank's predetermined denomination mix. However, the number of times an iteration of the algorithm is truncated is not significantly reduced when compared to the previous time limit. Therefore, it appears that the combinations that did not yield a significant improvement over the bank's predetermined denomination mix need more time to finish calculations than the 30 hours that are now available.

The cases where it does find a better solution are (15,7) for 100 ATMs and (15,15) for 200 ATMs. These are the two combinations that were a significant improvement over the bank's predetermined denomination mix when a smaller number of ATMs were considered.

6 Conclusion

In this paper we have combined the scheduling of replenishments for inventory management with different denomination mix strategies for ATMs and tested our methods using real data from the Netherlands. The denomination mix is the combination of bills used to fulfill a customer's demand at an ATM. This is the first paper that simultaneously determines an optimal denomination mix and replenishment strategy.

We have proposed a time-varying denomination mix strategy and compared it to a typical bank's predetermined denomination mix. The predetermined denomination mix implements a least notes strategy, while the time-varying denomination mix strategy optimizes the denomination mix and implements it for some periods, after which it is allowed to change. The smallest total costs generated by the time-varying denomination mix were found using a 3-day forecast and changing the denomination mix daily. This yielded total costs that were 7.6% or €46 lower per ATM per month than the lowest total costs that were found using the bank's predetermined denomination mix. This could represent over €3.5 million per year in the Netherlands. When notes of €5 are considered, the total costs almost double, but our strategy clearly remains better than the bank's. To determine the robustness of this solution, the number of ATMs considered simultaneously is varied, the interest rate, replenishment costs and lost-sales costs are separately doubled and halved and the forecast quality is varied. Even in these cases, the time-varying denomination mix remains less costly than the bank's predetermined denomination mix.

Future research includes the incorporation of the possibility for customers to either choose a specific denomination mix or let the ATM determine the mix based on its inventory. Moreover, a less time-consuming method such as a heuristic could be explored. In practice, the time available to find the schedule and denomination mixes might be restrictive. Furthermore, using a different denomination mix to prevent lost-sales, when the current denomination mix is not in stock, could also be considered. This inventory-driven substitution has been considered in other contexts, but not yet in the cash distribution.

References

[1] Y. Bassok, R. Anupindi, and R. Akella. Single-period multiproduct inventory models with substitution. *Operations Research*, 47(4):632–642, 1999.

[2] M. Bijvank and I.F.A. Vis. Lost-sales inventory theory: A review. *European Journal of Operational Research*, 215:1–13, 2011.

[3] L.C. Coelho, J.-F. Cordeau, and G. Laporte. Thirty years of inventory-routing. *Transportation Science*, 48(1):1–19, 2014.

[4] L.C. Coelho, J.-F. Cordeau, and G. Laporte. Heuristics for dynamic and stochastic invenotry-routing. *Computers and Operations Research*, 52:55–67, 2014.

[5] Y. Ekinci, J.-C. Lu, and E. Duman. Optimization of ATM cash replenishment with group-demand forecasts. *Expert Systems with Applications*, 42(7):3480–3490, 2015.

[6] P. Kurdel and J. Sebestyénová. Routing optimization for ATM cash replenishment. *International Journal of Computers*, 7(4):135–144, 2013.

[7] C. Lang. *Production and inventory management with substitutions.* PhD thesis, Technische Universität Darmstadt, Germany, 2010.

[8] H. Larrain, L.C. Coelho, and A. Cataldo. A variable MIP neighborhood descent algorithm for managing inventory and distribution of cash in automated teller machines. *Computers and Operations Research*, 85:22–31, 2017.

[9] J. Paul and A. Mukherjee. ATMs and cash demand forecasting: A study of two commercial banks. *Journal of Regional Development*, 2(2):653–671, 2010.

[10] S. Sethi and G. Sorger. A theory of rolling horizon decision making. *Annals of Operations Research*, 29:387–416, 1991.

[11] R.G. Van Anholt. *Optimizing logistics processes in cash supply chains*. PhD thesis, VU University Amsterdam, 2014.

[12] K. Vanketesh, R. Vadlamani, A. Prinzie, and D. Van den Poel. Cash demand forecasting in ATMs by clustering and neural networks. *European Journal of Operational Research*, 232:383–392, 2014.