University of Groningen

Replenishment and denomination mix of automated teller machines with dynamic forecast demands

Heide, van der, Lieke; Callegari Coelho, Leandro; Vis, Iris; van Anholt, Roel G.

Published in:
Computers & Operations Research

DOI:
[10.1016/j.cor.2019.104828](https://doi.org/10.1016/j.cor.2019.104828)

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2020

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

Heide, van der, L., Callegari Coelho, L., Vis, I., & van Anholt, R. G. (2020). Replenishment and denomination mix of automated teller machines with dynamic forecast demands. *Computers & Operations Research*, 114, [104828]. <https://doi.org/10.1016/j.cor.2019.104828>

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

Replenishment and denomination mix of automated teller machines with dynamic forecast demands

Lieke M. van der Heide^a, Leandro C. Coelho^{a,c,*}, Iris F.A. Vis^a, Roel G. van Anholt^b

^a Faculty of Economics and Business, University of Groningen, the Netherlands

^b Geldmaat, the Netherlands

^c Canada Research Chair in Integrated Logistics and CIRRELT, Université Laval, Québec, Canada

ARTICLE INFO

Article history:

Received 24 July 2017

Revised 6 June 2019

Accepted 14 October 2019

Available online 14 October 2019

Keywords:

Inventory management
Automated teller machine
Denomination mix
Optimization
Replenishment
Cash logistics
Prescriptive analytics

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. We investigate whether allowing the denomination mix to vary over time based on the forecast withdrawals at an ATM reduces actual operating costs of an ATM. To verify this, 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 or a one-smallest strategy. In all strategies we simultaneously optimize denomination mixes and replenishment decisions. We define the problem and solution strategies as mixed integer programming formulations and solve them via a rolling horizon algorithm using different frequencies of denomination mix updates, rolling horizon lengths, numbers of ATMs, cost parameters, and forecast qualities. By implementing the time-varying denomination mix, we show that the operational costs of managing an ATM can be reduced by 21% or €153.77 per ATM per month on average, which can represent over €10 million per year in the Netherlands only.

© 2019 Elsevier Ltd. All rights reserved.

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 (Larrain et al., 2017; van Anholt et al., 2016; Westland, 2002). 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. This situation is depicted in Fig. 1. 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 there is no denomination mix in stock to completely satisfy this demand. 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, reducing the cash cycle and having a more cost-effective operation (EPC, 2018; Eurosystem, 2016). Recent opportunities in cash management allow the implementation of solutions that exploit this opportunity to reduce costs by varying the denomination mix.

* Corresponding author at: Canada Research Chair in Integrated Logistics and CIRRELT, Université Laval, Québec, Canada.

E-mail addresses: L.m.van.der.heide@rug.nl (L.M. van der Heide), l.c.coelho@rug.nl (L.C. Coelho), i.f.a.vis@rug.nl (I.F.A. Vis), Roel.van.Anholt@geldmaat.nl (R.G. van Anholt).

Fig. 1. Selecting a denomination mix, ATM icon courtesy of www.flaticon.com.

Such an opportunity arises for our industry partner Geldmaat (formerly Geldservice Nederland NOS (2018)) in the Netherlands to implement such solutions. Geldmaat is a joint venture owned by the three largest banks and is responsible for cash distribution and ATM maintenance in the Netherlands. The company stands before a major operation of merging three ATM networks into one. Not only is the objective to optimize and standardize back-office processes, also the number and distribution of ATMs in the country is optimized, leading to fewer ATMs with higher coverage. However, having fewer ATMs also implies that ATMs will be used more frequently, requiring more replenishments. Therefore, it is even more important that all denominations remain available just before running low on inventory. According to the company implementing a smart denomination mix is a prerequisite to the optimization of the merged ATM network.

This major operation is an example of coping with the trend where increasingly less cash is used in the Netherlands. This creates a dependence on technology, which can be difficult especially for the elderly. In the Dutch national politics a debate will be held on the importance of access to cash from ATMs (Hoekstra, 2018). Reducing the costs of ATM inventory management will simplify this access. This is in line with the policy of the Dutch National

Bank which states that because of the reduced amount of cash payments, maximum efficiency is needed to ensure cash can still play its role in current society (De Nederlandsche Bank, 2018). In countries with high use of cash as a payment instrument, such as Austria, the pressure on cash logistics is even higher. In Austria, transactions of less than €10 are paid in cash 92.4% of the times; for transactions between €10 and €20 this figure is 87.6%, decreasing to 77.6% for transactions between €20 and €50, 59.5% for payments of €50–€100, and still high at 47.2% for transactions of value higher than €100 (Rusu and Stix, 2016). In that country, 81.8% of the transactions were paid by cash, for a value share of 64.9% in 2016 (Rusu and Stix, 2016). Fig. 2 shows the number of cashless transactions and number of cards per capita in the Euro zone.

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 (2010) 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) (Coelho et al., 2014b). IRPs have been explicitly applied to ATM management (Kurdal and Sebestyénová, 2013; Larrain et al., 2017; Van Anholt, 2014). In addition to the IRP, demand forecasting for ATMs has been covered in, for instance, Van Anholt (2014) and Vanketesh et al. (2014). Moreover, the combination of these two aspects has been covered in Ekinici et al. (2015), 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 (Bijvank and Vis, 2011).

Although the literature on optimization of denomination mixes for ATMs is scarce, industry is very active in this field, with

Fig. 2. Number of cashless transactions and number of card per capita in the Euro area, 2016. Source: European Payments Council document #EPC115-18, with data from the European Central Bank Statistical Warehouse.

numerous patents on ATM technology being registered in the past few years, e.g., Colvin et al. (2014), Huang et al. (2014), Vesely (2011). Policies similar to varying the denomination mix have been used in different fields with promising results. For instance, in Bassok et al. (1999) 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 (2010). In the case of ATMs many different withdrawal sizes (products in lot-sizing terminology) are considered where small denominations are preferred by customers over large denominations (Cleveland, 2011). 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. Shy (2019) studied how the flexibility in obtaining different denominations at ATMs affect a person's behavior in paying for his/her transactions in cash or card. This flexibility is highly related to the use of the ATM, and its cost effect is investigated in this paper.

We have partnered with Geldmaat, the operator responsible for replenishing ATMs for the three major banks in the Netherlands, covering approximately 90% of the Dutch ATM volume. Approximately 35% of their ATMs use a least note strategy to fulfil demands, i.e., the smallest possible number of bills is used to satisfy a demand. According to Geldmaat, the one-smallest, where at least one of the smallest denominations is included in the denomination mix, is implemented in approximately 55% of the ATMs. 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 or one-smallest strategy. These strategies are both examples of a 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. This allows for better demand satisfaction and lower operational costs. Using the withdrawal data Geldmaat 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, or on the one-smallest strategy which includes at least one of the smallest denomination in the mix. For a predetermined denomination mix, inventory management policies need to determine only the timing and quantities of the replenishments for each denomination. 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. The calculation of these costs is done in two phases. Firstly, we optimize an inventory management policy by minimizing the costs based on the forecast data in the planning phase. Then in the implementation phase, we use this policy and calculate the actual costs based on actual demands. In the experiments in Section 5, we use a rolling horizon method where for each time period we first optimize the inventory management policy in the planning phase and then use it in the implementation phase to calculate the actual costs and find the actual inventories, which is then used at the start of the next time period. We now focus on the planning phase and in Section 4 the implementation phase and the method for alternation used in the experiments are explained in more detail.

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 cost c_i corresponding to a replenishment at ATM i .

A lost-sale occurs when it is impossible to satisfy the complete demand using the current inventory. We assume in the planning phase that if the denomination mix necessary to fulfill the demand is not available in the ATM a lost-sale occurs, since the denomination mix cannot be varied within a day. This assumption is relaxed when the planned replenishment strategy is implemented and the actual costs are calculated, as is explained in Section 4. We need this assumption in the planning phase to ensure that the optimized replenishment strategy and denomination mix are robust and to reduce the computational complexity. 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 scheduling a lost-sale instead of a replenishment, a lost-sales cost M is considered in our model. We then calculate the actual lost-sales costs separately using a lost-sales cost of d , as is explained in Section 4.

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 $\mathcal{I} = \{1, \dots, m\}$ 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 $\mathcal{J} = \{1, \dots, J\}$ 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} . We define $\mathcal{O} = \{0, \dots, O\}$ as the set of number of the same bills used in a denomination mix, meaning that O represents the maximum number of the same bills used in a denomination mix. Moreover, the number of bills in box j of ATM i at the beginning of the planning

Table 1
Overview of parameters.

Parameter	Definition
p	Pack size in which bills are delivered
h	Interest rate
c_i	Fixed cost for a single replenishment at ATM i
M	Lost sales cost
D	Current period
T	Length of the planning horizon
b_{ij}	Denomination contained in box j of ATM i
Q_{ij}	Maximum capacity of box j of ATM i
e_{ij}	Number of bills in box j of ATM i at the beginning of the planning horizon
v_{ij}	Historical average number of bills in box j of ATM i
V_i	The number of values for which a withdrawal can occur at ATM i
l_{in}	Monetary values of the withdrawal n at ATM i
a_{in}^t	The forecast number of withdrawals at ATM i for value n at time t
o_{ijn}	The predetermined number of bills from box j of ATM i included in the denomination mix for value n
u_{ij}	Binary parameter, equals 1 when box j of ATM i is unused

Table 2
Overview of variables.

Variable	Definition
y_i^t	Binary variable, equals 1 if ATM i is visited at time t
l_{ij}^t	Inventory of box j of ATM i at the end of time t
x_{ij}^t	Number of bills lacking in box j of ATM i due to lost-sales at time t
f_{ij}^t	Integer variable that ensures all replenishments are made in packs of size p
m_{ijn}	Number of bills from box j of ATM i included in the denomination mix for value n
k_{ijnq}	Number of bills from box j of ATM i included in denomination mix q for value n
h_{inq}^t	Number of times denomination mix q is used for value n at ATM i at time t

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 can have monetary values l_{in} , $i \in \mathcal{I}$, $n \in \mathcal{N} = \{1, \dots, V_i\}$, but may not be demanded. The forecast number of withdrawals at ATM i , for value n at time t is a_{in}^t , $i \in \mathcal{I}$, $n \in \mathcal{N}$, $t \in \mathcal{T} = \{D, \dots, D+T\}$, 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. An overview of the parameters is presented in Table 1.

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 model, the time-varying denomination mix.

All models require the following decision variables: let y_i^t be a binary variable indicating whether ATM i is visited at time t . The number of bills that remains in box j of ATM i from day t to day $t+1$ is l_{ij}^t . Moreover, x_{ij}^t represents the number of bills that were necessary to fulfill customers' demands but were not in inventory in box j of ATM i at time t , meaning bills corresponding to lost-sales. Since we will consider forecast demands, these variables do not have to be integer. Integer variables f_{ij}^t ensure that all deliveries will be made in packs of size p . A summary of all the required variables is presented in Table 2.

All subsequent models are based on the following:

$$\min \sum_{t \in \mathcal{T}} \sum_{i \in \mathcal{I}} \left(y_i^t c_i + \sum_{j \in \mathcal{J}} (h l_{ij}^t b_{ij} + M x_{ij}^t) \right) \quad (1)$$

subject to

$$p f_{ij}^t \leq y_i^t Q_{ij} \quad \forall i \in \mathcal{I}, j \in \mathcal{J}, t \in \mathcal{T} \quad (2)$$

$$p f_{ij}^t \leq Q_{ij} - l_{ij}^{t-1} \quad \forall i \in \mathcal{I}, j \in \mathcal{J}, t \in \mathcal{T} \quad (3)$$

$$y_i^t \in \{0, 1\} \quad \forall i \in \mathcal{I}, j \in \mathcal{J}, t \in \mathcal{T} \quad (4)$$

$$f_{ij}^t \in \mathbb{N} \quad \forall i \in \mathcal{I}, j \in \mathcal{J}, t \in \mathcal{T} \quad (5)$$

$$l_{ij}^t, x_{ij}^t \in \mathbb{R}^+ \quad \forall i \in \mathcal{I}, j \in \mathcal{J}, t \in \mathcal{T}. \quad (6)$$

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 such as a replenishment. Constraints (2) ensure that boxes can only be replenished if their ATM is visited. Constraints (3) impose the number of bills used to replenish box j . Constraints (2) 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 (3) state that the number of bills cannot be greater than the maximum capacity minus the current inventory. This allows active replenishment before the box is empty and prevents the number of bills in a box from exceeding its capacity. Constraints (4)–(6) 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)–(6) and by:

$$I_{ij}^t = I_{ij}^{t-1} + pf_{ij}^t - \sum_{n \in \mathcal{N}} (o_{ijn} a_{in}^t) + x_{ij}^t \quad \forall i \in \mathcal{I}, j \in \mathcal{J}, t \in \mathcal{T} \quad (7)$$

$$I_{ij}^T \geq \min \left(v_{ij}, Q_{ij} - \sum_{n \in \mathcal{N}} (o_{ijn} a_{in}^T) - p + 1 \right) \quad \forall i \in \mathcal{I}, j \in \mathcal{J} \quad (8)$$

$$I_{ij}^D = e_{ij} \quad \forall i \in \mathcal{I}, j \in \mathcal{J}. \quad (9)$$

Constraints (7) 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. However, some of these demands are not fulfilled due to lost-sales and need to be added to the inventory. Constraints (8) 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 . The maximum possible number of bills in a box is attained when at the start of the period it contains Q_{ij} , then all demand is fulfilled using the denomination mix. Finally, we need to subtract $(p - 1)$, since bills are delivered in packs of size p . This can make it impossible to replenish the ATM to capacity and hence reduces the maximum number of bills in a box. Moreover, constraints (9) are added to ensure that the initial inventory is correctly considered.

A numerical example is provided to help understand this formulation, specifically the use of constraints (8). Consider a case with initial inventory equal to zero notes, maximum capacity equal to 50, total demand equal to 31 notes, pack size equal to 10, and (historical) average inventory equal to 30. The second term of the right-hand side of constraints (8) then evaluates to $50 - 31 - 9 = 10$. If the term -1 was not present in these constraints, they would be evaluated to 9. With a final inventory of 10 and demand of 31, one should order at least 41 units; because the pack size is 10, one should order 50 units, which is the maximum capacity. Without the term -1 , the final inventory would have been only 9, which with the demand of 31 adds up to 40, which would have been the order size. Note that ordering 50 instead of 40 leads to a final inventory closer to the historical average.

To linearize constraints (8), we will use two auxiliary variables s_{ij}^+ and s_{ij}^- as follows:

$$I_{ij}^T \geq v_{ij} - s_{ij}^+ \quad \forall i \in \mathcal{I}, j \in \mathcal{J} \quad (10)$$

$$I_{ij}^T \geq Q_{ij} - \sum_{n \in \mathcal{N}} o_{ijn} a_{in}^T - p + 1 - s_{ij}^- \quad \forall i \in \mathcal{I}, j \in \mathcal{J} \quad (11)$$

$$s_{ij}^+ \geq v_{ij} - \left(Q_{ij} - \sum_{n \in \mathcal{N}} o_{ijn} a_{in}^T - p + 1 \right) \quad \forall i \in \mathcal{I}, j \in \mathcal{J} \quad (12)$$

$$s_{ij}^- \geq Q_{ij} - \sum_{n \in \mathcal{N}} o_{ijn} a_{in}^T - p + 1 - v_{ij} \quad \forall i \in \mathcal{I}, j \in \mathcal{J} \quad (13)$$

$$s_{ij}^+, s_{ij}^- \in \mathbb{N} \quad \forall i \in \mathcal{I}, j \in \mathcal{J}, \quad (14)$$

and penalize s_{ij}^- and s_{ij}^+ in the objective function. s_{ij}^+ then equals the amount v_{ij} exceeds $Q_{ij} - \sum_{n \in \mathcal{N}} o_{ijn} a_{in}^T - p + 1$ and s_{ij}^- the other way around. Constraints (10) ensure that the inventory is at least equal to v_{ij} if it is the smallest. If it is greater, we subtract the difference and the inventory is at least equal to the other term. Constraints (11) are used similarly.

We also derive the following valid inequalities:

$$\sum_{t \in \mathcal{T}} pf_{ij}^t \geq I_{ij}^T - e_{ij} + \sum_{t \in \mathcal{T}} \sum_{n \in \mathcal{N}} o_{ijn} a_{in}^t - \sum_{t \in \mathcal{T}} x_{ij}^t \quad \forall i \in \mathcal{I}, j \in \mathcal{J}. \quad (15)$$

Inequalities (15) 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 m_{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. This means that if the bills from that box are not allowed in the denomination mixes u_{ij} equals 1, and it is 0 otherwise. The model for the time-varying denomination mix is then defined by (1)–(6) and by:

$$\sum_{j \in \mathcal{J}} m_{ijn} b_{ij} = l_{in} \quad \forall i \in \mathcal{I}, n \in \mathcal{N} \quad (16)$$

$$I_{ij}^t = I_{ij}^{t-1} + pf_{ij}^t - \sum_{n \in \mathcal{N}} (m_{ijn} a_{in}^t) + x_{ij}^t \quad \forall i \in \mathcal{I}, j \in \mathcal{J}, t \in \mathcal{T} \quad (17)$$

$$I_{ij}^T \geq v_{ij} \quad \forall i \in \mathcal{I}, j \in \mathcal{J} \quad (18)$$

$$I_{ij}^D = e_{ij} \quad \forall i \in \mathcal{I}, j \in \mathcal{J} \quad (19)$$

$$pf_{ij}^{D+1} > -e_{ij} - u_{ij} \quad \forall i \in \mathcal{I}, j \in \mathcal{J} \quad (20)$$

$$m_{ijn} \in \mathcal{O} \quad \forall i \in \mathcal{I}, j \in \mathcal{J}, n \in \mathcal{N}. \quad (21)$$

Constraints (16) ensure that the denomination mix multiplied by the values of the bills that are used in the mix equals the requested value. Constraints (17) ensure inventory conservation. Constraints (18) ensure that at the end of the planning horizon the inventory equals the average number of bills again to ensure a sustainable solution. Constraints (19) guarantee that the initial inventory is taken into consideration here. Constraints (20) 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 (21) ensure that optimal denomination combinations do not use too many of the same bills, by restricting the number of bills used.

We also derive the following valid inequalities:

$$p \sum_{t \in \mathcal{T}} f_{ij}^t \geq I_{ij}^T - e_{ij} + \sum_{t \in \mathcal{T}} \sum_{n \in \mathcal{N}} m_{ijn} a_{in}^t - \sum_{t \in \mathcal{T}} x_{ijt} \quad \forall i \in \mathcal{I}, j \in \mathcal{J}. \quad (22)$$

Constraints (22) 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.

3.3. Different denomination mixes

We now present a model that allows the use of different denomination mixes within a time period for our proposed time-varying denomination mix. Again we use the variables, parameters and the basic model presented in Sections 2 and 3. In Section 3.2 we introduced the variables m_{ijn} , which we now replace by the more general variables k_{ijnq} such that multiple denomination mixes for one withdrawal value can be used. Variables k_{ijnq} indicate the number of bills from box j of ATM i included in denomination mix q for value n . Additionally, we use a new set of variables h_{inq}^t , indicating the number of times denomination mix q is used to fulfill a demand for value n at time t for ATM i . The nonlinear model for the time-varying denomination mix with multiple denomination mixes is then defined by (1)–(6), (18)–(20) and by:

$$\sum_{j \in \mathcal{J}} k_{ijnq} b_{ij} = l_{in} \quad \forall i \in \mathcal{I}, n \in \mathcal{N}, q \in \mathcal{Q} \quad (23)$$

$$l_{ij}^t = l_{ij}^{t-1} + p f_{ij}^t - \sum_{n \in \mathcal{N}} \sum_{q \in \mathcal{Q}} (k_{ijnq} h_{inq}^t) + x_{ij}^t \quad \forall i \in \mathcal{I}, j \in \mathcal{J}, t \in \mathcal{T} \quad (24)$$

$$\sum_{q \in \mathcal{Q}} h_{inq}^t = a_{in}^t \quad \forall i \in \mathcal{I}, n \in \mathcal{N}, t \in \mathcal{T} \quad (25)$$

$$k_{ijnq} \in \mathcal{O} \quad \forall i \in \mathcal{I}, j \in \mathcal{J}, n \in \mathcal{N}, q \in \mathcal{Q} \quad (26)$$

$$h_{inq}^t \in \mathbb{R}^+ \quad \forall i \in \mathcal{I}, n \in \mathcal{N}, q \in \mathcal{Q}, t \in \mathcal{T} \quad (27)$$

Constraints (23) state that the denomination mix multiplied by the value of the bills that are used in the mix equals the requested value, for each of the denomination mixes. Constraints (24) enforce inventory conservation. Moreover, the number of times a denomination mix is used equals the number of forecast demands due to constraints (25). The denomination mixes cannot consist of too many of the same bills due to constraints (26), and the domain of the variable h_{inq}^t is defined in constraints (27).

Similar to the previous section, we derive the following valid inequalities:

$$p \sum_{t \in \mathcal{T}} f_{ij}^t \geq l_{ij}^T - e_{ij} + \sum_{t \in \mathcal{T}} \sum_{n \in \mathcal{N}} \sum_{q \in \mathcal{Q}} k_{ijnq} h_{inq}^t - \sum_{t \in \mathcal{T}} x_{ijt} \quad \forall i \in \mathcal{I}, j \in \mathcal{J}. \quad (28)$$

These valid inequalities state that the number of bills delivered in the whole planning horizon should equal the number of bills in inventory at the end of the planning horizon, minus how many were in inventory at the beginning, plus the number of bills necessary to fulfill demand, minus the lost sales.

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. Since the model presented in

Section 3.3 is nonlinear and it is a more general form of the model from Section 3.2, we do not include it in our experiments.

Sethi and Sorger (1991) 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. (2014a), 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. Then for the demands that cannot be fulfilled using this denomination mix, we calculate a new least-notes mix using the bills in inventory. In that sense, our proposed algorithm is more flexible than the mathematical models used to optimize the denomination combination because we actually allow a least note strategy to satisfy a demand in case the optimized denomination combination is no longer available. Note also that in case a given denomination runs out of inventory, the least note strategy is automatically updated for that period using the remaining bank notes. If a demand cannot be satisfied using any denomination mix, it is denoted a lost-sale and the corresponding costs are calculated. This yields the actual inventory. For example: an ATM contains only notes of €20 and €10. For demands of €50 the (time-varying or predetermined) denomination mix one note of €50 should be used. However, when a customer arrives demanding €50, they will receive two notes of €20 and one note of €10. When the ATM contains only €10, five notes of €10 will be used. A lost-sale will only occur when the notes in the ATM cannot fulfill the demand. For instance, when there are no €20 notes and less than five €10 notes, or if there are only €20 notes.

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 could not be satisfied are added. For the calculation of actual lost-sales cost we use unit cost d instead of M .

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 or one-smallest 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 Fig. 3. 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,

Fig. 3. Overall scheme of the time-varying algorithm.

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 or one-smallest strategy.

We use two datasets of actual withdrawals obtained from Geldmaat, 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. 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 (2014) and are denoted a_{in}^t . The demand forecasts are generated using the statistical forecasting model for the multi-step ahead prediction of causal hourly and daily time series. The forecasting model adequately deals with seasonality, trends, holidays, events and outliers. 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 different from the actual demands for a certain value for the standard ATM, and 0.46 different 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 robustness. Since we consider forecasts with three decimals, we set the coefficient

M equal to $1000c_i$, making a lost-sale of 0.0001 more expensive than an additional replenishment.

These two datasets were used to create five instances containing 50 ATMs and five with 100 ATMs of the standard configuration, both starting from ATMs 0, 50, 100, 150 and 200 in the dataset. Moreover, five instances of 200 ATMs of the standard configuration were created from ATMs 0, 25, 50, 75 and 100. Additionally, 5 instances containing 50 ATMs of the other dataset were created from ATMs 0, 15, 30, 35 and 60. 5 instances of 100 ATMs were created from ATMs 0, 2, 4, 6, 8. 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 , actual lost-sales cost d , 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 Geldmaat and their corresponding values are shown in Table 3. Moreover, for the replenishment costs c_i we used the average value provided. 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. This means that $v_{i1}, v_{i2}, v_{i3}, v_{i4}, v_{i5}$ are the average number of bills in the ATM if €5, €10, €20, €50, or €100 were used, respectively. 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 one of the bank's predetermined denomination mixes and the time-varying denomination mix are defined. Five different values of T , the length of the rolling horizon, are considered namely: 1, 3,

Table 3
Parameter setting values obtained from our industrial partner Geldmaat.

Parameter	h	d	c_i	Q_{ij}	p	v_{i1}	v_{i2}	v_{i3}	v_{i4}	v_{i5}
Values	0.05	0.10	100	2000	100	933	858	975	1.395	551

Table 4

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

T	One smallest	Least notes	Time-varying denomination mix (L)				
			1	3	7	15	30
1	3206.40 (41.61)	3141.23 (45.31)	3255.75 (30.58)	3224.09 (36.98)	3269.76 (30.98)	3284.89 (25.89)	3287.88 (23.48)
3	827.88 (75.80)	788.01 (96.59)	565.38* (11.17)	860.46 (33.43)	1076.48 (44.82)	1116.21 (64.64)	1209.07 (63.71)
7	777.00 (77.25)	803.41 (96.51)	599.34* (31.46)	659.48* (39.72)	705.82* (48.29)	747.63* (53.96)	762.41* (67.58)
15	730.23 (76.66)	730.95 (82.54)	701.25 (104.46)	656.42* (29.86)	661.86* (32.98)	686.41* (66.37)	708.56* (83.50)
30	730.21 (69.69)	719.15 (76.81)	1184.93 (227.09)	976.01 (249.26)	876.22 (367.07)	709.595 (61.34)	703.00* (69.99)

* Indicates significant improvement over the minimum of the least notes and one-smallest mixes at 1% level.

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 1925 experiments.

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.

In Table 4 we show the average monthly costs per ATM of implementing the least notes and one-smallest predetermined strategies 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 best 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 4, 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 4 that the time-varying denomination mix outperforms both the bank's predetermined denomination mixes 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 €565.38 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 21% or €153.77 per ATM less costly than the most efficient solution of the bank's predetermined least notes denomination mix, which is based on a 30-day rolling horizon and costs €719.15 per ATM per month. This reduction in costs can be decomposed into a reduction of 34% in replenishment costs and of 3% in holding costs compared to the bank's predetermined denomination replenishment and holding costs, respectively. While lost-sales costs increase, they amount to less than 0.1% 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 is a downside to the implementation of the time-varying denomination mix. It is more time-consuming than simply using the bank's predetermined denomination mix. Specifically, the solution of the bank's predetermined denomination mix takes on average 945 s, whereas the best time-varying denomination mix takes 2628.6 s. However, it still allows practical use. However, there is a significant decrease in the amount of lost-sales when the time-varying denomination mix is implemented. The least costly solution of the bank's least notes denomination mix includes €1666.87 lost-sales per month per ATM on average, while €455.05 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 in both settings. Moreover, when the time-varying denomination mix is implemented, the demands are met using smaller or equally big denomination. Since customers prefer notes from smaller denominations, its implementation would increase customer satisfaction (Cleveland, 2011). 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

Table 5

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

T	One	Least	Time-varying denomination mix (L)				
	smallest	notes	1	3	7	15	30
1	3341.25 (45.18)	3300.97 (47.83)	3309.06 (38.22)	3336.07 (39.79)	3386.70 (28.70)	3387.75 (29.79)	3383.03 (34.19)
3	1464.92 (56.23)	1479.91 (61.61)	1238.97* (95.92)	1474.42 (78.61)	1587.75 (22.16)	1554.26 (62.31)	1514.03 (51.59)
7	1167.24 (53.76)	1260.71 (64.90)	1440.08 (192.82)	1374.40 (226.11)	1144.54 (93.41)	1103.55* (63.01)	1103.15* (52.02)
15	1123.58 (58.03)	1199.92 (60.27)	2415.41 (262.07)	1943.95 (380.50)	1549.09 (358.11)	1655.85 (456.46)	1198.36 (510.89)
30	1351.73 (75.09)	1249.03 (66.42)	3171.55 (204.69)	3058.34 (244.83)	2773.62 (294.62)	2441.32 (281.39)	2702.44 (146.47)

* indicates significant improvement over the minimum of the one smallest and the least notes mixes at 1% level.

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 5. It is clear that this configuration is the most costly option, with the bank's predetermined one-smallest denomination mix costing at least €1123.58 and the smallest value for the time-varying denomination mix being €1103.15. These costs were obtained using a different dataset that included demands for €5 which is 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 previous dataset is used to evaluate the performance of this configuration, we find that the best (T, L) strategy for the time-varying denomination mix performs 13.2% worse than when the best (T, L) was implemented for the standard configuration.

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 a predetermined denomination mix is implemented are €893.30, whereas they are only €415.20 when the standard configuration is used. The same relation occurs when the time-varying denomination mix is implemented: €869.90 when €5 was used and €272.16 when the standard configuration was used. Conversely, the holding costs are lower at only €223.29 compared to €301.67 when a bank's predetermined denomination mix was used and €229.74 compared to €292.58 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 one smallest denomination mix is used are €738.10, while when the time-varying denomination mix was used the total costs were at least €601.59, meaning a 23% 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 least notes denomination mix were €809.74, whereas the lowest costs using the time-varying denomination mix were €672.86, meaning a 17% reduction. The lowest costs using the time-varying denomination mix were obtained when a 7-day forecast was used and the denomination mix changed only once every 3 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 23%. 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 100 ATMs produces the lowest costs per ATM, compared to using 50 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 3 days produces costs that are significantly smaller than both the bank's predetermined denomination mixes when only 100 ATMs were considered. However, when 200 ATMs were considered the costs of this combination are no longer smaller than both the bank's predetermined denomination mixes, since its costs are 4% higher than when 100 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 32,808 s to finish, while it only takes 9792 s when 50 ATMs were considered. When 50 ATMs are considered, the costs are also 4% higher than when 100 ATMs were considered, however, this was the case for both the bank's predetermined and the time-varying denomination mix. This can be explained by the calculation of lost-sales. When the replenishments are scheduled and in some cases the denomination mixes are determined, we assume lost-sales occur when the denomination mix

is not in stock. However, when the actual lost-sales are computed we allow different denomination mixes to be used.

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. Moreover, there is no significant difference in the costs per ATM for the different numbers of ATMs considered using this strategy.

The interest rate, replenishment costs and lost-sales costs have each been separately doubled and halved for the best 5 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 least notes 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 only yields lower costs than the bank's predetermined least notes denomination mix, when 3 days of forecast were used and the denomination mix changes every day. 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 best bank's predetermined least notes 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 the success of the time-varying denomination mix is dependent on the quality of the forecast.

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 h when a new denomination mix is determined and 8 h 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 (30,1), (30,3); for 100 ATMs we used (30,1), (30,3), (30,7); and for 200 ATMs we examined (15,1), (30,1), (30,3), (30,7). Even though the time limit is doubled in most cases, the solutions still do not achieve lower costs than that of the bank's least costly 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 h that are now available. The cases where it does find a better solution are (30,3) for 100 ATMs and (30,7) for 200 ATMs.

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 two typical bank's predetermined denomination mixes. The predetermined denomination mixes consist of a least notes or a one-smallest 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 21% or €153.77 lower per ATM per month than the lowest total costs that were found using one of the bank's predetermined denomination mixes. This could represent over €10 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. However, the performance of the time-varying denomination mix does depend on the quality of the forecast.

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 during the optimization. This inventory-driven substitution has been considered in other contexts, but not yet in the cash distribution. Future research could also include the design of a dynamic strategy for setting an optimal denomination mix for each arriving customer based on the current inventory level and testing the benefits.

Acknowledgments

We thank the contacts at our industry partner for providing the data and validating our parameters. This project was partly funded by the [Canadian Natural Sciences and Engineering Research Council](#) (NSERC) under grant 2019-00094. This support is greatly acknowledged. We thank the editor, an associate editor, and three anonymous referees for their valuable comments on an earlier version of this paper.

References

- Bassok, Y., Anupindi, R., Akella, R., 1999. Single-period multiproduct inventory models with substitution. *Oper. Res.* 47 (4), 632–642.
- Bijvank, M., Vis, I., 2011. Lost-sales inventory theory: a review. *Eur. J. Oper. Res.* 215, 1–13.
- Cleveland, V., 2011. Challenges in note circulation - speech by Victoria Cleveland. URL: <http://www.bankofengland.co.uk/archive/Documents/historicpubs/speeches/2011/speech482.pdf>.
- Coelho, L., Cordeau, J.-F., Laporte, G., 2014. Heuristics for dynamic and stochastic inventory-routing. *Comput. Oper. Res.* 52, 55–67.
- Coelho, L., Cordeau, J.-F., Laporte, G., 2014. Thirty years of inventory-routing. *Trans. Sci.* 48 (1), 1–19.
- Colvin, R., Norman, R., Siekman, J., 2014. Currency Recycler. US Patent 8,756,158.
- De Nederlandsche Bank, 2018. *Visie op Betalen 2018–2021*.
- Ekinci, Y., Lu, J.-C., Duman, E., 2015. Optimization of ATM cash replenishment with group-demand forecasts. *Expert Syst. Appl.* 42 (7), 3480–3490.

- EPC, 2018. Recirculation paper – improving the efficiency of the handling of cash. URL: <https://www.europeanpaymentscouncil.eu/sites/default/files/kb/file/2018-06/EPC115-18%20Recirculation%20Paper%20-%20Improving%20the%20efficiency%20of%20the%20handling%20of%20cash.pdf>.
- Eurosystem, 2016. Annual report 2016. URL: <https://www.ecb.europa.eu/pub/pdf/annrep/ar2016en.pdf>.
- Hoekstra, W., 2018. Reactie op het verzoek van het lid Van Rooijen, gedaan tijdens de Regeling van werkzaamheden van 7 juni 2018, over het gebruik van contant geld in de samenleving.
- Huang, J. C., Bullard, B. E., Vasilkos, G. J., 2014. Method for Determining a Bill Mix in an Electronic Cash Dispenser. US Patent App. 13/758,784.
- Kurdal, P., Sebestyénová, J., 2013. Routing optimization for ATM cash replenishment. *Int. J. Comput.* 7 (4), 135–144.
- Lang, C., 2010. Production and Inventory Management with Substitutions. Technische Universität Darmstadt, Germany.
- Larain, H., Coelho, L., Cataldo, A., 2017. A variable MIP neighborhood descent algorithm for managing inventory and distribution of cash in automated teller machines. *Comput. Oper. Res.* 85, 22–31.
- NOS, 2018. ING, Rabo en ABN werken samen met nieuwe gele geldautomaat.
- Paul, J., Mukherjee, A., 2010. ATMS and cash demand forecasting: a study of two commercial banks. *Journal of Regional Development* 2 (2), 653–671.
- Rusu, C., Stix, H., 2016. Technical Report. oesterreichische nationalbank.
- Sethi, S., Sorger, G., 1991. A theory of rolling horizon decision making. *Ann. Oper. Res.* 29, 387–416.
- Shy, O., 2019. How the ATM Affects the Way We Pay. Technical Report. Federal Reserve Bank of Atlanta. Working Paper 2019-2.
- Van Anholt, R., 2014. Optimizing Logistics Processes in Cash Supply Chains. VU University Amsterdam.
- van Anholt, R., Coelho, L., Laporte, G., Vis, I., 2016. Inventory-routing problem with pickups and deliveries. *Transp. Sci.* 50 (3), 1077–1091.
- Vanketesh, K., Vadrilamani, R., Prinzie, A., Van den Poel, D., 2014. Cash demand forecasting in ATMs by clustering and neural networks. *Eur. J. Oper. Res.* 232, 383–392.
- Vesely, J., 2011. State-Driven Self-Service Terminal. US Patent App. 12/749,840.
- Westland, C., 2002. Preference ordering cash, near cash and electronic cash. *J. Organ. Comput. Electron. Commerce* 12 (3), 223–242.