European Journal of Operational Research 000 (2016) 1–9

Contents lists available at ScienceDirect

European Journal of Operational Research

journal homepage: www.elsevier.com/locate/ejor

Interfaces with Other Disciplines

Pareto improvement and joint cash management optimisation for banks and cash-in-transit firms

Kolos Cs. Ágoston[a,*], Gábor Benedek[b], Zsolt Gilányi[c]

[a] Department of Operational Research and Actuarial Sciences, Corvinus University of Budapest, 1092 Budapest, Fővám tér 8., Hungary
[b] Thesys SEA Pte Ltd. 16 Ayer Rajah Crescent, #03-03, Singapore 139965 / Corvinus University of Budapest, 1092 Budapest, Fővám tér 8, Hungary
[c] Institute for Economic Sciences and Methodology, Faculty of Economics, University of West Hungary, 9400 Sopron, Erzsébet út 9., Hungary

ARTICLE INFO

Article history:
Received 25 September 2015
Accepted 23 April 2016
Available online xxx

Keywords:
OR in banking
Cash-management
Group decisions and negotiations
Supply chain management

ABSTRACT

Improving the ATM cash management techniques of banks has already received significant attention in the literature as a separate optimisation problem for banks and the independent firms that supply cash to automated teller machines. This article concentrates instead on a further possibility of cost reduction: optimising the cash management problem as one single problem. Doing so, contractual prices between banks and the cash in transit firms can be in general modified allowing for further cost reduction relative to individual optimisations. In order to show the pertinence of this procedure, we have determined possible Pareto-improvement re-contracting schemes based on a Baumol-type cash demand forecast for a Hungarian commercial bank resulting in substantial cost reduction.

1. Introduction

Due to dramatic changes in interest rates and in the availability of funds, especially during crisis, cash management of banks received much attention. This focus on cash management was partly motivated by the cost effective functioning—though cash management costs are relatively small compared to the overall banking costs—and partly by better assuring liquidity in periods of shortage in liquidity.

The proper cash management assures the optimal quantity of cash at the right place and time. Accordingly, the focus in the literature on cash management of banks is to improve the stochastic cash demand forecasts. Cash management policies can be built upon these forecasts, for example by the help of linear (stochastic multistage) programming. In brief, the basic idea of cost reduction is optimisation under given constraints in various environments. In concrete terms, the cash management of ATMs is equivalent to the problem of finding the optimal amount of cash and the optimal number of transports at given interest rate (opportunity cost of holding cash) and given transportation fee. The transportation fee comprises in general two components: a fix fee per transport and a variable fee on the amount handled (replenished or emptied). This logic allows thus for cost reduction by the help of better prediction and optimisation techniques.

As far as we know, none of the contributions to the cash management literature on ATMs has investigated another potential cost gain which is independent of the prediction and optimisation techniques. Namely, in general it is possible to reduce the overall cash management costs by holding the overall transportation costs and the interest rate constant. For this eventuality, the composition of the fix and variable parts of the transportation fee should induce an appropriate variation of the amount and the number of transported cash; the gain comes from the reduction of the total opportunity cost of holding cash (interest). The problem is not just finding the optimal cash management scheme for given prices but offer a new Pareto improvement pricing scheme which makes the bank strictly better off.

In this paper we use the simple Baumol technic to predict the cash demand though better prediction methods are available. This is because our purpose is to hold the model as simple as possible in order to insure that the bank management understands and accepts it. The model is applied for a Hungarian commercial bank's cash logistics of withdrawal ATMs replenished by an independent firm from a central hub by way of replacing unopened cash cassettes. While the Baumol prediction oversimplifies some real world facts, the policy built upon this prediction method can still significantly reduces cash management cost in our empirical case.

Clearly, the re-contracting policy of the bank with the cash in transit firm following the Pareto improvement logic applies in any case when the cash management scheme is optimised without

* Corresponding author. Tel.: +36 302107338.
E-mail addresses: kolos.agoston@uni-corvinus.hu (K. Cs. Ágoston), gabor.benedek@thesys.com (G. Benedek), gilanyi@ktk.nyme.hu (Zs. Gilányi).

http://dx.doi.org/10.1016/j.ejor.2016.04.045
0377-2217/

taking into consideration the concrete cost structure of the transportation; especially, if the cash transportation is assured by an independent firm. The point is that independent optimisation (CIT optimises independently of the bank) cannot produce better result than joint optimisation. Also, the re- contracting policy suggested in this paper is not dependent on the optimisation and prediction technique. In this regard, any limitations coming from the Baumol technic are irrelevant.

This article determines just the Pareto improvement recontracting payment schemes between banks and the cash supplying firms, a precondition of strategic behaviour for the implementation of enhanced payment schemes. The implementation of the possible ensuing contracts between the bank and the independent firm is outside the scope of this article; a straightforward framework for handling the implementation problem is to define an appropriate bargaining game.

Following this introduction, Section 2 presents the relevant literature. Section 3 constructs a Baumol-type model for predicting the cash demand and examines four scenarios based on the amount of the unloaded cash y_{un} and the independent firm's (henceforth the Firm) cost CT. Section 4 presents the Hungarian commercial bank (henceforth the Bank). Afterwards, this theoretical model is applied on the Bank's representative ATM and an augmented simulation model on its overall ATM network—the positive test outcomes suggest three pricing schemes. Finally, Section 5 concludes the article by summarising the research findings and suggesting possible avenues for policy implications and further methodological amendment.

2. Relevant literature on cash management

The cash management literature starts with the pioneering contribution on the transactions demand for money by Baumol (1952) and Tobin (1956). The basic model investigates the optimal size of cash holdings in a deterministic setting—where the choice is between cash and an interest-bearing asset, with strictly positive trade-off costs between the two assets. Integrating other (transaction) cost elements and stochastic money flows is a straightforward extension of this basic model (Miller and Orr (1966), Heyman (1973)). To this end, Eppen and Fama (1968) used linear programming and assumed stationary net cash flows to satisfy the Markov process. Elton and Gruber (1974) used dynamic programming. Following these early works a great number of authors contributed to the cash management literature (the list is non-exhaustive): Hinderer and Waldmann (2001) generalised the so called (u,U,D,d) role for a wider environment, Cardo (2009) conceived the cash management problem as a stochastic problem and solved it with (mixed) integer linear programming. Bar-Ilan, Perry, and Stadje (2004) used a generalised impulse control model. Yao, Chen, and Lu (2006) and Arora and Saini (2014) used the fuzzy integral method. Simutis, Dilijonas, Bastina, Friman, and Drobinov (2007) constructed an artificial neural network to forecast the daily cash demands for ATMs and estimate the optimal ATM cash loads—Venkatesh, Ravi, Prinzie, and Van Den Poel (2014) used neural networks too, but clustered ATMs preliminarily. Armenise, Birtolo, Sangianantoni, and Troiano (2010) used a specific genetic algorithm. Teddy and Ng (2011) forecasted the cashdemands with the help of cerebellar associative memory network, Ekinci, Lu, and Duman (2015) used group-demand forecasts. Baker, Jayaraman, and Ashley (2013) dropped the assumption of normally distributed errors to improve the cash demand forecast. Dilijonas, Sakalauskas, Kriksciuniene, and Simutis (2009) combined optimisation with ATM management.

In summary, much of the literature on cash management is concerned with cash supply optimisation within a given cash allocation scheme. According to Cabello (2013, p. 334), most analyses 'in the financial literature assume the standard allocation models of cash from central hubs to branches as given'. However, Pokutta and Schmaltz (2011) showed that choosing the appropriate liquidity hub scheme results in significant economies.

Batlin and Hinko (1982) furthered the literature by replacing the standard allocation framework with a game theoretical framework and by analysing the interaction between cash flow acceleration and liquidity. This novel approach also shed light on an important lacuna of the standard allocation framework—cash management costs may indeed be reduced by enhanced cash flow predictions, but also by enhanced contractual pricing between banks and the firms that supply cash to bank branches and ATMs. Consequently, this article builds on Baumol (1952) and Tobin (1956) by suggesting a Baumol-type model and advocating further economies within the standard framework from enhanced contractual pricing.

Naturally, cash optimisation may be analysed from perspectives different from that adopted in this article. Blackman, Holland, and Westcott (2013) and Popa (2013) did indeed do so, from the perspective of global financial supply chain strategies, as did de Haan and van den (2013), from the perspective of regulatory implications (see also Basel Committee on Banking Supervision (2013) and Burcea, Bălău, Bâldan, Avrămescu, and Ungureanu (2013) on liquidity management). Also, in a reversal of emphasis, Snellmana and Virenb (2009) examined how markets and ATM network structures affect payment choices; Bátiz-Lazo (2009) examined ATM network structures in the UK; and Opasanon and Lertsanti (2013) examined the cost efficiencies of companies that provide logistics services for banks. However, such perspectives are outside the scope of this article.

Finally some articles are aiming at modelling the cost structure related to the cash management problem. Ou, Hung, Yen, and Liu (2009) investigate the effect of ATM deployment on cost efficiency. Stavins (2000) analyses the use of non-customer ATM fees, while Bjø rndal, Hamers, and Koster (2004) investigate the ATM fees with the help of cooperative game theory.

3. Methodology and discussion

The ATM cash management literature is mainly concerned with developing better cash demand predictions and better replenishment policy by the help of optimisation programs. This article concentrates instead on the further possibility of cost reduction—optimising the ATM cash management problem as one single problem. Therefore, we use a simple Baumol-type model to forecast the cash demand. The limitations of this simple model are secondary for now, because the possible economies are not based on better predictions or better optimisation but on joint optimisation of the ATM cash management problem instead of the optimisation of two separate programs of banks and cash in transit firms.

The Baumol type model advanced here below is applied for withdrawal ATMs; for deposit ATMs see Appendix A. The model divides time into periods—days—and uses a number of symbols:

- a 1 + value added tax (VAT),
- C_T an independent firm's total daily cost for a specific ATM,
- e an independent firm's initial total daily earning for a specific ATM,
- E an independent firm's total daily earning for a specific ATM,
- E^* an independent firm's optimal total daily earning for a specific ATM,
- f the daily turnover of a specific ATM (defined as the total successful withdrawal from that ATM),
- k an independent firm's journey fee for a specific ATM (defined as the price the independent firm charges for a specific number of transports to that ATM),

- l_{un} an independent firm's unloaded quantity fee for a specific ATM (defined as the price the independent firm charges for the cash unloaded from that ATM),
- l_{up} an independent firm's uploaded quantity fee for a specific ATM (defined as the price the independent firm charges for the cash uploaded into that ATM),
- p an independent firm's initial total daily profit for a specific ATM,
- P an independent firm's total daily profit for a specific ATM,
- r a bank's daily interest rate,
- tc a bank's initial total daily cost for a specific ATM,
- TC a bank's total daily cost for a specific ATM,
- TC^* a bank's optimal total daily cost for a specific ATM,
- x the number of days between two consecutive transports to a specific ATM,
- x^* the optimal number of days between two consecutive transports to a specific ATM,
- y_{un} the cash unloaded from a specific ATM,
- y_{up} the cash uploaded into a specific ATM,

and

- y^*_{up} the optimal cash uploaded into a specific ATM.

Evidently, $x = \frac{y_{up} - y_{un}}{f}$.

The model also makes a number of assumptions:

- banks own their respective ATM networks,
- the ATMs in a bank's network are replenished by an independent firm,
- ATMs can be replenished at any moment in time (during the day),
- f is constant in time,
- k is constant in time and identical for all the ATMs in a bank's network,

and

- r is constant in time—there is no compound interest.

We note that the assumption of identical journey fee k and quantity fee l_{up} will be raised in the next point.

For a specific transport to a specific ATM, independent firms charge a cash transportation fee $\frac{l_{up}y_{up}+l_{un}y_{un}}{x} + \frac{ka}{x}$ involving the two quantity fees l_{up} and l_{un} and the journey fee k. In Hungary, k incorporates VAT, because financial institutions are not allowed reimbursement, but l_{up} and l_{un} do not, because they are exempt from VAT—in different taxation circumstances, the quantity fees would incorporate VAT too.

In addition to the cash transportation fee, banks incur the interest loss on ATM cash stocks—withdrawals by customers from ATMs translate into identical concomitant withdrawals by banks from customers' bank accounts. Following Baumol and the assumption that ATM cash stocks decrease to y_{un} constantly, the average ATM cash stock is $\frac{y_{up}+y_{un}}{2} = \frac{fx}{2} + y_{un}$ and the daily interest loss is $r\frac{fx}{2} + ry_{un}$, where ry_{un} is a fixed cost with no bearing on optimisation. The daily interest loss is proportional with the average ATM cash stock and tax distortions are ignored—if taken into account, r would need to be modified accordingly.

Consequently, a bank's total daily cost of securing a specific ATM's cash supply is

$$TC = \frac{rfx}{2} + ry_{un} + \frac{l_{up}y_{up} + l_{un}y_{un}}{x} + \frac{ka}{x} = \\ = \frac{rfx}{2} + l_{up}f + \frac{(l_{up}+l_{un})y_{un}}{x} + \frac{ka}{x} + ry_{un} \quad (1)$$

If l_{un} is considered exogenously given for the bank and the above expression is rearranged, we get:

$$TC = \frac{rfx}{2} + l_{up}f + \frac{(l_{up}+l_{un})y_{un} + ka}{x} + ry_{un} \quad (2)$$

Given that a, k, f, l_{up}, and r are known parameters, the first order condition for the optimal total daily cost TC^* with respect to x is

$$\frac{rf}{2} - \frac{(l_{up}+l_{un})y_{un} + ka}{x^2} = 0 \quad (3)$$

Consequently, the optimal number of days between two consecutive transports to a specific ATM is

$$x^* = \sqrt{\frac{2[(l_{up}+l_{un})y_{un} + ka]}{rf}} \quad (4)$$

and the optimal amount of money uploaded into that ATM is

$$y^*_{up} = \sqrt{\frac{2f[(l_{up}+l_{un})y_{un} + ka]}{r}} + y_{un} \quad (5)$$

These results are common sense—x decreases with increases in f and/or r as well as with decreases in k, l_{up}, and/or y_{un}; x increases with increases in k and/or l_{up}. Hence, the optimal total daily cost for a bank is

$$TC^* = \sqrt{\frac{rf[(l_{up}+l_{un})y_{un} + ka]}{2}} + fl_{up} \\ + \sqrt{\frac{rf[(l_{up}+l_{un})y_{un} + ka]}{2}} + ry_{un} \\ = f\left(\sqrt{\frac{2r[(l_{up}+l_{un})y_{un} + ka]}{f}} + l_{up}\right) + ry_{un} \quad (6)$$

However, the likelihood of mutually beneficial arrangements can only be evaluated by considering the independent firm's profit P—for convenience, the signs of the variations in earnings and profits are assumed to coincide. An independent firm's optimal earning is the sum of the second and third terms of the bank's total daily cost less VAT:

$$E^* = fl_{up} + \frac{\sqrt{rf}[(l_{up}+l_{un})y_{un} + k]}{\sqrt{2[(l_{up}+l_{un})y_{un} + ka]}} \\ = f\left(\frac{\sqrt{r}[(l_{up}+l_{un})y_{un} + k]}{\sqrt{2f[(l_{up}+l_{un})y_{un} + ka]}} + l_{up}\right) \quad (7)$$

By the help of these general expressions, the Pareto optimal points can be determined for different parameter scenarios:

3.1. The $y_{un} = 0$ and $C_T = 0$ scenario

The assumption $y_{un} = 0$ is unrealistic, but reduces the relevant mathematical formulae to $TC^* = f(\sqrt{(2rka)/f} + l_{up})$ and $E^* = f(\sqrt{(rk)/(2fa)} + l_{up})$, thus revealing the point of mutually beneficial arrangements—l_{up} is both a cost for the bank and an income for the independent firm, but the same is not true for k. A marginal increase in k results in a marginal increase in TC^*—$\partial TC^*/\partial k = f\sqrt{(ra)/(2fk)}$—greater than the marginal increase in E^*— $\partial E^*/\partial k = f\sqrt{r/(8afk)}$. This result too is common sense—the bank pays VAT, not the independent firm. Moreover, the result is valid even if VAT = 0 and, implicitly, $a = 1$—to compensate for effects due to variations in k, the bank can always adjust the number of transports.

For Pareto improvement, judicious adjustments to k and the number of transports transform the interest loss from a mere cost for the bank into an income for the independent firm. Intuitively, reductions in k lead to increases in the number of transports, increases in the independent firm's earning, and decreases in the bank's interest loss. Moreover, increases in the independent firm's earning have to be accompanied by increases in the independent firm's profit, while eventual increases in TC have to be balanced

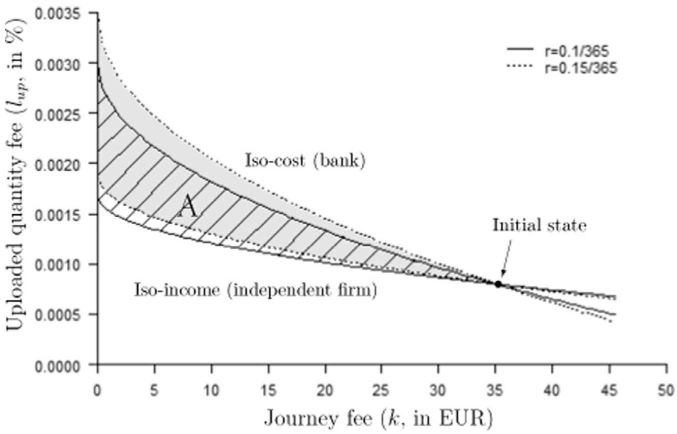

Fig. 1. Iso-cost (bank) and iso-income (independent firm) curves for zero cash unloaded from an ATM and zero independent firm costs ($f =$ EUR 5,000, $a = 1.25$).

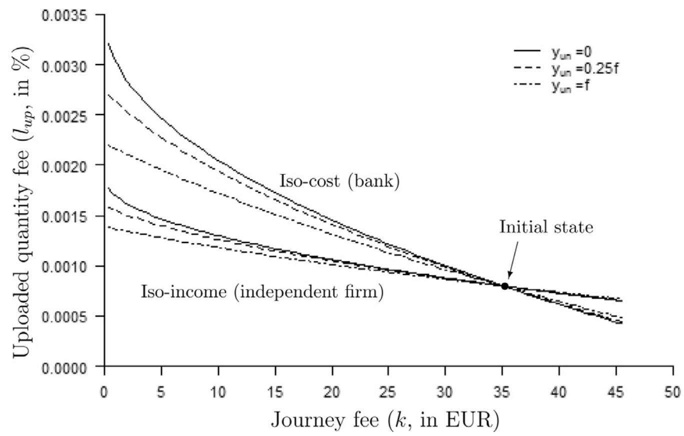

Fig. 2. Iso-cost (bank) and iso-income (independent firm) curves for positive cash unloaded from an ATM and zero independent firm costs ($f =$ EUR 5,000, $a = 1.25$, $r = 0:15 = 365$, $l_{un} = 0:000247$).

by reductions in l_{up}. That intuition is correct can be proved algebraically, by determining the Pareto optimal points—if tc and e are the initial total daily cost for the bank and respectively earning for the independent firm, then the iso-cost curve for the bank is

$$l_{up} = \frac{tc}{f} - \sqrt{\frac{2ra}{f}}\sqrt{k} \tag{8}$$

with the slope $-\sqrt{(ra)/(2fk)}$ and the iso-income curve for the independent firm is

$$l_{up} = \frac{e}{f} - \sqrt{\frac{r}{2fa}}\sqrt{k} \tag{9}$$

with the slope $-\sqrt{r/(8afk)}$.

Fig. 1 shows the iso-cost and iso-income curves for $r = 0.1/365$—the initial state is characterised by k = EUR 35 and $l_{up} = 0.0008$. (Contracts between banks and independent firms are highly confidential—as well as subject to significant variations—consequently, the seed values here and elsewhere in this article were chosen close to industrial averages and by omitting contractual aspects such as penalties, liabilities, etc.) The hashed area A represents the negotiation space—the Pareto improvement points—where the bank and the independent firm would both be better off than in the initial state. For comparison, Fig. 1 also shows the iso-cost and iso-income curves for $r = 0.15/365$. The slope of the iso-cost curve is always greater than the slope of the iso-income curve—in absolute terms, independently of tc and e. Pareto improvement is therefore always possible, by decreasing k—if $k > 0$—and increasing l_{up}. Also, quite logically, increases in r lead to increases in the negotiation space—the higher the r, the higher the interest loss for cash stocks held in ATMs and the larger the number of available options.

Determining mutually advantageous arrangements and Pareto optimal points amounts to solving the algebraic problem

$$\min_{k,l_{up}} f\left(\sqrt{\frac{2rka}{f}} + l_{up}\right)$$
$$s.t.$$
$$f\left[\sqrt{\frac{rk}{2fa}} + l_{up}\right] \geq \bar{e}$$
$$k, l_{up} \geq 0 \tag{10}$$

If the objective value is not higher than tc, then the optimal solution to (10) is a Pareto optimal point—changes in $\bar{e}$ result in a set of Pareto optimal points. The optimal solution to (10) is the same corner solution seen in Fig. 1, with $k = 0$—and $x = 0$—and $l_{up} = \frac{\bar{e}}{f}$. In other words, an infinite number of transports with infinitely small amounts of money uploaded into the ATM will minimise the cash stock—and the interest loss—to zero.

3.2. The $y_{un} > 0$ and $C_T = 0$ scenario

However, if $y_{un} > 0$, TC^* and E^* become cumbersome equations and an infinite number of transports is not the optimal solution—banks would have to pay l_{up} and l_{un} proportional with y_{un} an infinite number of times. By decreasing k and increasing l_{up} as in Section 3.1, banks would upload ATMs more frequently, to avoid increases in the interest loss. Moreover, increases in l_{up} would have effects similar to increases in k—in (6) and (7), the term $(l_{up} + l_{un})y_{un}$ has the same bearing as k, but the term fl_{up} counterbalances its effect. For Pareto optimal, $k = 0$ again, but with a finite number of transports.

Fig. 2 shows the effect of positive cash unloaded from ATMs—the iso-cost and iso-income curves both shift downwards with increases in y_{un}—a reasonable expectation, since the term $(l_{up} + l_{un})y_{un}$ does not diminish as k tends to zero. The independent firm's earning remains unchanged with moderate increases in l_{up}, which compensates the bank for cost increases.

3.3. The $y_{un} = 0$ and $C_T > 0$ scenario

Generally, independent firms have significant fixed costs—car fleet amortisation, wages, and many others. However, for Pareto improvement, fixed costs may be ignored and, similarly with their earnings, the total daily costs of independent firms may be split in two—one part proportional to l_{up}, the other proportional to k:

$$C_T = \frac{y_{up}l' + k'}{x} = f\left(\frac{k'}{\sqrt{k}}\sqrt{\frac{r}{2fa}} + l'\right) \tag{11}$$

where l' and k' are the cost parameters. Consequently, the independent firm's profit is

$$P = E - C_T = f\left(\sqrt{\frac{r}{2fa}}\frac{k - k'}{\sqrt{k}} + l_{up} - l'\right) \tag{12}$$

If p is an independent firm's initial total daily profit, then–similarly to (8)–the iso-income curve is

$$l_{up} = \frac{p}{f} - \sqrt{\frac{r}{2fa}}\frac{k - k'}{\sqrt{k}} + l' \tag{13}$$

with the slope $-\sqrt{r/(2af)}(k + k')/(2k\sqrt{k})$.

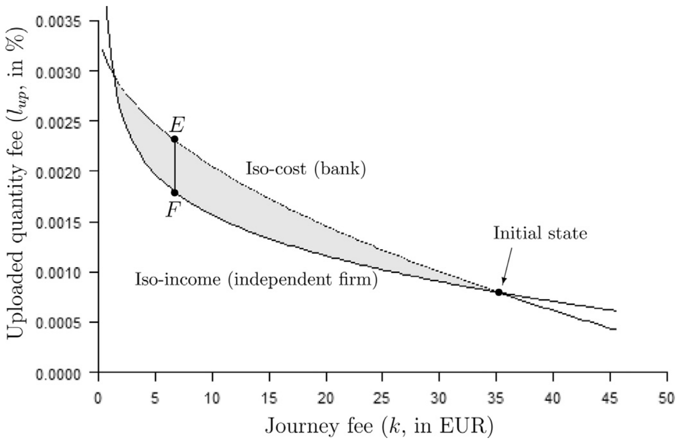

Fig. 3. Iso-cost (bank) and iso-income (independent firm) curves for zero cash unloaded from an ATM and positive independent firm costs ($f =$ EUR $5,000$, $r = 0.15/365$, $a = 1.25$, $k' =$ EUR 10, $l' = 0.0247$ percent and $y_{un} = 0$).

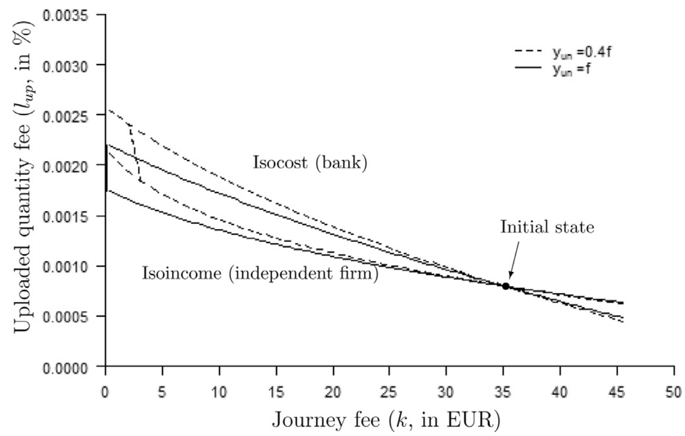

Fig. 4. Iso-cost (bank) and iso-income (independent firm) curves and Pareto optimal sets for positive cash unloaded from an ATM and positive independent firm costs ($f =$ EUR $5,000$, $a = 1.25$, $k' =$ EUR 10, $l' = 0.0247$ percent).

If $k = 0$, $l_{up} = \frac{tc}{f}$ in the iso-cost curve; if k tends to zero, l_{up} tends to infinity in the iso-profit curve and there is no corner solution; and if $k = k'/(2a-1)$, the slopes of the two curves are equal. However, if k is not 'too small' and k' is not 'too high', then the slope of the iso-income curve is less steep than the slope of the iso-cost curve in absolute terms—decreases in k result in win-win situations. The slopes of the iso-cost and iso-income curves are independent of tc and p—consequently, the Pareto optimal points are aligned vertically (see the straight line between points E and F in Fig. 3).

Determining mutually advantageous arrangements and Pareto optimal points amounts again to solving an algebraic problem:

$$\min_{k,l} f\left(\sqrt{\frac{2rka}{f}} + l\right)$$

$$s.t.$$

$$f\left(\sqrt{\frac{r}{2fa}}\frac{k-k'}{\sqrt{k}} + l - l'\right) \geq \overline{p}$$

$$k, l \geq 0 \tag{14}$$

If the objective value is not higher than tc, then the optimal solution is a Pareto optimal point—changes in $\overline{p}$ result in a set of Pareto optimal points.

3.4. The $y_{un} > 0$ and $C_T > 0$ scenario

Fig. 4 illustrates the case of positive cash unloaded from an ATM and positive independent firm costs. When y_{un} is 'relatively' small ($y_{un} = 0.4f$), the situation is similar to that of zero cash unloaded from the ATM, but not identical. No longer a vertical line, the Pareto optimal set in this case is situated to the left of the Pareto optimal set in that case (a vertical line at $k = 5$)—a reasonable difference, given that with the earning from the unloaded cash, the independent firm can compensate for the income loss from the number of transports. When y_{un} is high ($y_{un} = f$), there is again a corner solution.

3.5. Cost heterogeneity

In the previous sections the journey fee (k) and the quantity lee (l_{up} and l_{un}) were fixed and identical. In the banking practice these prices are not identical in general; most often the cash transportation fee for onsite and offsite ATMs differ. This difference can easily be handled with the same model by clustering ATMs preliminarily: the Bank and the Firm negotiate on prices separately for onsite and offsite ATMs. The same method applies if ATMs are clustered on the basis of any other properties if the transportation fees are fixed.

In fact, the difficulty comes rather from the constraint if—for some reason—the same price is to be set for all types of ATMs. In that case, the Firm has to determine a single price for the whole network or cluster though his cost differ significantly from ATM to ATM inside the network or the cluster. The representative ATM solution still works in this case where the cost parameter is a weighted average of the ATMs' cost parameters in the cluster. To illustrate this, let us take the simple case with two different ATMs (1 and 2) in a cluster with different cost parameters (k_1', k_2', l_1', l_2').

If y_{un}=0 and f is identical for both ATMs, the total profit is:

$$P = f\left(\sqrt{\frac{r}{2fa}}\frac{k-k_1'}{\sqrt{k}} + l_{up} - l_1'\right) + f\left(\sqrt{\frac{r}{2fa}}\frac{k-k_2'}{\sqrt{k}} + l_{up} - l_2'\right)$$
$$= 2f\left(\sqrt{\frac{r}{2fa}}\frac{k-\frac{(k_1'+k_2')}{2}}{\sqrt{k}} + l_{up} - \frac{(l_1'+l_2')}{2}\right)$$

That is, we obtain the same result as if there were two identical ATMs with the average cost parameters. Therefore, if the turnover f does not differ significantly, we can work with average (representative) ATMs.

However, if the turnover f differs from ATM to ATM —let us say the daily turnover of ATM 1 equals β times the daily turnover of ATM 2— we get a rather complicated expression:

$$P = f\left(\sqrt{\frac{r}{2fa}}\frac{k-k_1'}{\sqrt{k}} + l_{up} - l_1'\right) + \beta f\left(\sqrt{\frac{r}{2\beta fa}}\frac{k-k_2'}{\sqrt{k}} + l_{up} - l_2'\right)$$
$$= \sqrt{\frac{fr}{2a}}\left(\frac{k-k_1'}{\sqrt{k}} + \frac{\sqrt{\beta}(k-k_2')}{\sqrt{k}}\right) + (1+\beta)f\left(l_{up} - \frac{l_1'+\beta l_2'}{1+\beta}\right)$$
$$= (1+\beta)\sqrt{\frac{fr}{2a}}\left(\frac{\frac{(1+\sqrt{\beta})}{1+\beta}k - \frac{(k_1'+\sqrt{\beta}k_2')}{1+\beta}}{\sqrt{k}}\right)$$
$$+ (1+\beta)f\left(l_{up} - \frac{l_1'+\beta l_2'}{1+\beta}\right)$$
$$= (1+\beta)f\left(\sqrt{\frac{r\left(\frac{(1+\sqrt{\beta})}{1+\beta}\right)^2}{2fa}}\left(\frac{k - \frac{(k_1'+\sqrt{\beta}k_2')}{1+\sqrt{\beta}}}{\sqrt{k}}\right) l_{up} - \frac{l_1'+\beta l_2'}{1+\beta}\right)$$

Please cite this article as: K. Cs. Ágoston et al., Pareto improvement and joint cash management optimisation for banks and cash-in-transit firms, European Journal of Operational Research (2016), http://dx.doi.org/10.1016/j.ejor.2016.04.045

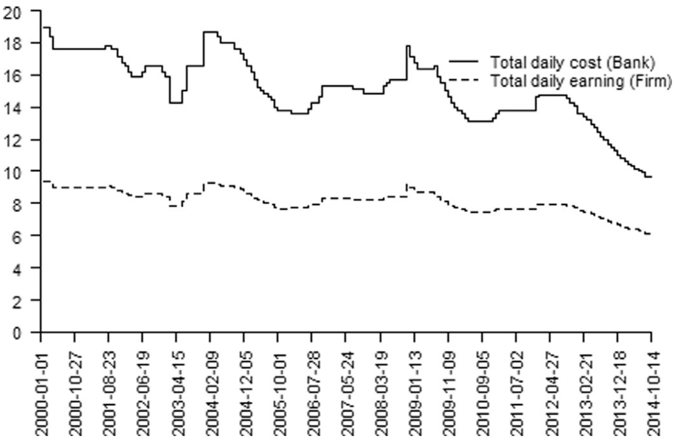

Fig. 5. Total daily costs (Bank) and total daily earnings (Firm) for a random ATM, 01.01.2000–14.10.2014, (f = EUR 5, 000; k = EUR 35; $l_{un} = 0.000247$; $l_{up} = 0.0008$; and $y_{un} = f$).

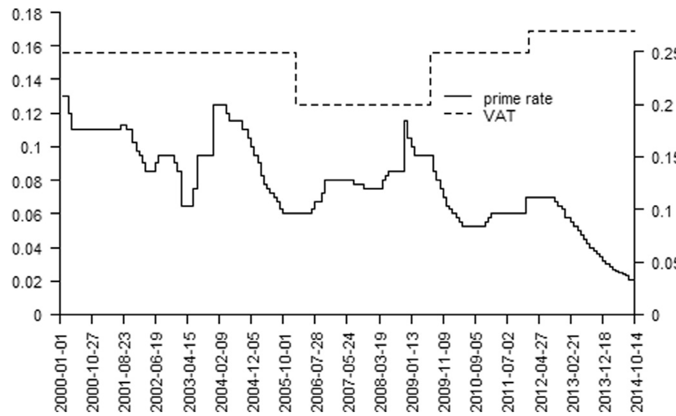

Fig. 6. The Hungarian National Bank prime interest/lending rates (left axis) and VATs (right axis), 01.01.2000–14.10.2014.

This result is basically identical to the previous one: we can still use representative ATMs, but now the cost parameter is not a simple average of the different cost parameters but a weighted average (with a modified r parameter). In clear this procedure can be generalised for any number of ATMs.

4. Data and results

4.1. Data

In Hungary as elsewhere in Central and Eastern Europe (CEE), most onsite ATMs used to be replenished by bank employees—only offsite ATMs used to be replenished by independent firms. However, from around 2005, the development of reasonably accurate cash forecasting and cash logistics optimisation tools by Hewlett-Packard, NCR, SAS, and others led to independent firms replacing bank employees in the replenishment of all ATMs—forecast by bank employees is much less accurate. The Bank has a very small number of deposit ATMs—however, their 'replenishment' is covered by a separate pricing contract with the Firm. The Bank also has multifunctional ATMs which allow both depositation and withdrawal—however, they were just being tested and not yet installed during the observation period 01.11.2009–31.10.2010. ATMs could still be 'topped up', as they usually were when bank employees replenished onsite ATMs. Currently, however, the cassette-in-cassette-out method–whereby locked cassettes containing cash can only be opened at designated locations—prevails in CEE too. The Bank had adopted this method prior to the observation period. The model advanced in this article can be applied to withdrawal ATMs as well as to deposit ATMs and multifunctional ATMs—however, this article is concerned only with withdrawal ATMs, the type of ATM vastly predominant at the Bank, including during the observation period.

The contractual pricing between Bank and Firm used to depend on ATM location—for example, onsite or offsite and/or urban or rural. The model advanced in this article can easily accommodate this arrangement predating 2009—as long as each journey fee k is constant in time, irrespective of actual ATM location. However, during the observation period, the onsite and offsite journey fees k were identical as well as constant in time, irrespective of actual ATM location, as were the onsite and offsite quantity fees l.

During the observation period, the Bank's ATM transactions were primarily intra-bank, involving the Bank's own customers—even so, correct fee allocation to specific ATMs was practically impossible. Consequently, the majority of these transactions were free of charge—more precisely, the fee for the first two ATM transactions in any month was integrated either into the monthly account fee or into the bank card fee.

The Bank operates in Hungary as well as other CEE countries, but this article concentrates solely on its operations in Hungary. To this end, the following data were made available to the authors for all the Bank's ATMs in Hungary during the observation period 01.11.2009–31.10.2010:

- descriptive data: distinct identifier code (ID), location (address), type (onsite/offsite), and currency (EUR/HUF) (for convenience, EUR ATMs were excluded from this research);
- turnover data: distinct identifier code (ID), date (yyyy-mm-dd), and turnover (f);
- replenishment data: distinct identifier code (ID), date (yyyy-mm-dd), unloaded amount (y_{un}), and uploaded amount (y_{up});

and

- closing balance data: distinct identifier code (ID), date (yyyy-mm-dd), and closing balance ($y_{up} - y_{un}$).

However, to avoid commercially harmful identification, the detailed data is omitted from this article.

The theoretical model advanced in Section 3 determined the pricing mutually advantageous to banks and their respective independent firms. Using the comprehensive data provided by the Bank, this section tests the model's descriptive power—first on the representative ATM and then on the Bank's overall ATM network, allowing for parameter differences among ATMs.

For a random ATM, Fig. 5 shows significant variations in the Bank's total daily costs and the Firm's total daily earnings, due to significant variations in the Hungarian National Bank prime interest/lending rates and VATs (see Fig. 6).

Consequently, x^* and y^*_{up} also vary significantly. The interest loss for cash stocks decreases with decreases in prime interest/lending rates— consequently, banks order higher amounts of money and fewer transports. At constant prices, this leads to decreases in E, as it did in Hungary in 2004 and 2009.

4.2. The representative ATM model

To apply the theoretical model to the representative ATM, the average daily turnover between 01.11.2009 and 31.10.2010 is considered representative daily turnover as well as unit of measure, with $a = 1.25$, $l_{un} = 0.0247$ percent, $l_{up} = 0.077$ percent and $r =$ 5.5 percent/365. and $(k) = 0.033$ percent f (in percentage of f).

Table 1 compares the actual costs in the second column with the modelled costs for $y_{un} = 0$ and $y_{un} > 0$ in the third and respec-

Table 1
Actual and modelled daily averages and other quantities for the Bank's ATM network, 01.11.2009–31.10.2010 ($r = 0.055/365$, $a = 1.250$, $l_{up} = 0.770$ permille $l_{un} = 0.247$ permille, $k = 0.033$).

Daily averages 10^{-6}	Actual	Modelled		
		Representative ATM		Simulation
		$y_{un} = 0$	$y_{un} = f$	
(i) Journey fee	386	447	401	393
(ii) VAT	97	112	100	98
(iii) Quantity fee for uploads	877	770	863	902
(iv) Quantity fee for unloads	34	0	30	42
(v) Material cost (i)+(ii)+(iii)+(iv)	1,394	1,329	1,394	1,435
(vi) Interest loss	712	559	775	607
(vii) Total daily cost (v)+(vi)	2,106	1,888	2,169	2,042
(viii) CIT's income (vii)−(ii)−(vi)	1,297	1,217	1,294	1,337
(ix) Uploaded amount of cash	1,139,569	1,000,000	1,120,726	1,170,834
(x) Unloaded amount of cash	139,569	0	120,726	170,834
(xi) Cash stock	4,722,007	3,711,905	5,141,598	4,028,788
Other quantities				
(xii) Number of transports	0.116	0.135	0.121	0.120
(xiii) Number of days between two consecutive transports (1/xii)	8.604	7.424	8.283	8.334
(xiv) Uploaded cash (ix)∗(xiii)	9,804,366	7,423,811	9,283,195	9,757,400

tively fourth columns. In the $y_{un} = 0$ scenario, the Firm uploads ATMs when cash stocks are exactly zero. In the $y_{un} > 0$ scenario, modelled y_{un} equals f was preferred to modelled y_{un} equals actual y_{un} for coherence with the assumption of no runouts with minimal unloads. Firm costs aside, the theoretical model fits the representative ATM parameters well—even when $y_{un} = 0$. However, the modelled daily cash stock and interest loss are by 9 percent higher than the actual daily cash stock and interest loss—even when $y_{un} > 0$. Such distortions—due mainly to the simplifications embedded in the representative ATM model—suggest the need for a more refined, simulation model. In reality, ATMs differ in turnover and loading frequency, and their turnover fluctuates with time and location—it is higher during the working week and lower at weekends, if located in a business quarter; lower during the working week and higher at weekends, if located in a shopping centre; while ATMs in proximity of factories experience high turnovers on paydays and in their immediate aftermath.

4.3. The simulation model

To best forecast turnovers, banks use sophisticated software in conjunction with numerous outer parameters, an approach useful in the given conditions, but useless in any other conditions. Consequently, the (deterministic) simulation model advanced in this article aims to combine the comprehensive advantage of the theoretical model with practical applicability. To this end, it allows for ATMs with different turnovers and loading frequencies and for customers who withdraw cash randomly, but with a systematic seasonality. This model is based on the following assumptions:

- the daily turnover of a specific ATM (f) is the average of the next seven days' forecast turnover;
- $a = 1.25$, $k = 0.0332$, $l_{un} = 0.000247$, $l_up = 0.00077$, $r = 0.055/365$ and $y_{un} = f$, as in Section 4.2;
- ATMs are replenished on the days when the cash stocks shrink to less than 1.5 times the turnovers forecast for the next day—to avoid cash runouts at weekends and on public holidays, when ATMs are not replenished, this assumption is modified accordingly, with the turnovers of non-replenishment days;

and

- y^*_{up} is calculated in accordance with (5), by omitting denomination imbalances (that is, cash spread unevenly over banknote denominations);

The fifth column of Table 1 shows that the modelled daily average cash stock and interest loss are lower than the actual daily average cash stock and interest loss.

Table 2
Modelled costs for ATM network. $r = 0.055/365$, $a = 1.250$ and $l_{un} = 0.247$ permille.

Daily averages 10^{-6}	Pricing scheme A		Pricing scheme B		Pricing scheme C	
	Representative ATM	Simulation	Representative ATM	Simulation	Representative ATM	Simulation
(i) Journey fee	85	88	235	209	303	299
(ii) VAT	21	22	59	52	76	74
(iii) Quantity fee for uploads	1,161	1,228	1,020	1,095	970	1,023
(iv) Quantity fee for unloads	52	69	39	60	35	50
(v) Material cost (i)+(ii)+(iii)+(iv)	1,319	1,407	1,353	1,416	1,384	1,446
(vi) Interest loss	510	308	624	389	684	481
(vii) Total daily cost (v)+(vi)	1,829	1,715	1,977	1,805	2,068	1,927
(viii) CIT's income (vii)−(ii)−(vi)	1,298	1,385	1,294	1,364	1,308	1,372
(ix) Uploaded amount of cash	1,209,885	1,279,578	1,159,206	1,244,238	1,141,155	1,203,826
(x) Unloaded amount of cash	209,885	279,578	159,206	244,238	141,155	203,826
(xi) Cash stock	3,382,254	2,044,628	4,140,582	2,579,369	4,542,201	3,189,723
Other quantities						
(xii) Number of transports	0.210	0.272	0.159	0.193	0.141	0.148
(xiii) Number of days between two consecutive transports x (1/xii)	4.765	3.674	6.281	5.174	7.084	6.735
(xiv) Uploaded cash (ix)∗(xiii)	5,764,508	4,701,489	7,281,163	6,437,186	8,084,403	8,107,545
(xv) Journey fee	0.0031	0.0031	0.0113	0.0113	0.0215	0.0215
(xvi) Quantity fee for uploads	0.096	0.096	0.088	0.088	0.085	0.085

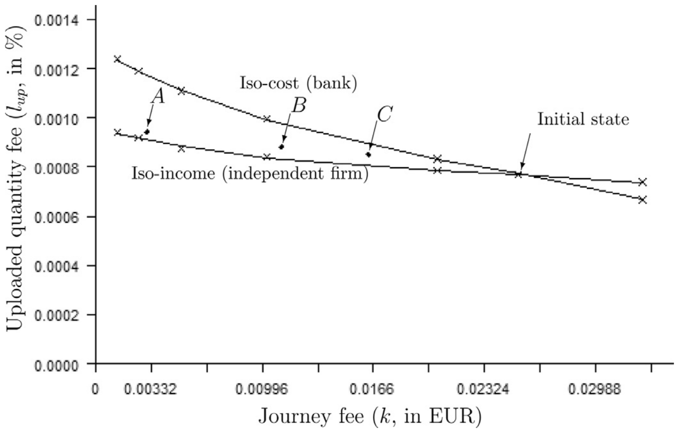

Fig. 7. Iso-cost (Bank) and iso-income (Firm) curves for the overall ATM network.

Fig. 7 shows the iso-cost and iso-income curves for the overall ATM network—they were determined by numerical trial (that is, by calculating concrete points on the iso-curves and fitting curves to these points)—and tree possible pricing schemes (A, B, and C). Table 2 shows the modelled costs for these pricing schemes

A shift from the pricing scheme A to C implies an increase of the cash stock and the journey fee and a parallel decreases in number of transports and the amount of uploaded cash. The outcome of the bargaining, i.e. the re-contracted pricing scheme makes certainly both parties better off. Nevertheless, without prior knowledge on the independent firm's costs it is impossible to determine the Pareto optimal states.

5. Conclusions

5.1. Main findings

This article examined the cash management optimisation for ATMs using a Hungarian commercial bank's data which owns a withdrawal ATM network replenished by an independent firm. At the time of this research, the Bank had already exploited—and exhausted—the usual cash management techniques for cost effectiveness stemming primarily from enhanced cash flow prediction. Nevertheless, even with the help of a simple Baumol-type model, this research succeeded in reducing further the Bank's cash management costs by determining mutually advantageous pricing schemes between Bank and Firm. For enhanced predictive power, the representative ATM model was dropped in favour of different ATMs, with different turnovers and loading frequencies—and with customers who withdraw cash randomly, but with a systematic seasonality. However, unlike the representative ATM model, it does not allow for finding Pareto optimal points through closed mathematical formulae—hence, the deterministic nature of the simulation model.

5.2. Policy implications

The usual cash management techniques for cost effectiveness stemming primarily from enhanced cash flow prediction and optimisation are not the only source for cost reduction. Indeed, a recontracting policy of the Bank with the Firm following a Pareto improvement logic applies in any case when the cash management scheme is optimised without taking into consideration the concrete cost structure of the cash transportation, because the independent optimisation cannot provide better result than joint optimisation. The simulation model revealed three mutually beneficial pricing schemes and, in practice, resulted in Bank and Firm re-contracting with each other successfully.

Naturally, the simulation model could be extended by removing assumption limitations such as those regarding ATM type and location, to include deposit/multifunctional ATMs and onsite/offsite cost differentiations. It could also be extended by analysing banks' profit sensitivity to changes in interest rates—increases in interest rates lead to increases not only in interest marges but also in cash management costs. Most excitingly, however, the simulation model could be extended by investigating the game theoretical implementations of Pareto improvement possibilities—a likely future research avenue for the authors of this article, facilitated by further data from the Bank.

Acknowledgements

The authors are very grateful to Lorenzo Peccati, for rejecting their earlier article firmly but kindly and for encouraging them to think afresh and start anew; to the two anonymous reviewers, for detailed and knowledgeable comments on the earlier article; and to László Halaj, Ferenc Forgó and to Miklós Pintér for their precious comments and suggestions.

Appendix A

For deposit ATMs, the daily withdrawals f are partly balanced by cash deposits f_{in}. Accordingly, equation (1) varies because ATMs have $xf_{in}/2$ additional cash stocks, where x has the same meaning as before —the number of days elapsing between two uploads:

$$TC = \frac{r(f+f_{in})x}{2} + ry_{un} + \frac{l_{up}y_{up} + l_{un}y_{un}}{x} + \frac{ka}{x}$$
$$= \frac{r(f+f_{in})x}{2} + l_{up}f + \frac{(l_{up}+l_{un})y_{un}}{x} + \frac{ka}{x} + ry_{un} \quad \text{(A.1)}$$

The first order condition:

$$\frac{r(f+f_{in})}{2} - \frac{(l_{up}+l_{un})y_{un} + ka}{x^2} = 0$$

and the optimum is

$$x^* = \sqrt{\frac{2[(l_{up}+l_{un})y_{un}+ka]}{r(f+f_{in})}}$$

That is, we have a very similar result to the withdrawal ATM case with $f + f_{in}$ daily turnover. The only difference is that the bank's total cost is shifted with the fix cost $l_{un}f_{in}$, the unloading cost of the deposited cash.

References

Armenise, R., Birtolo, C., Sangianantoni, E., & Troiano, L. (2010). A generative solution for ATM cash management. *Soft Computing and Pattern Recognition (SoCPaR), 2010*, 349–356. doi:10.1109/SOCPAR.2010.5686730.

Arora, N., & Saini, J. K. R. (2014). Approximating methodology: Managing cash in automated teller machines using fuzzy ARTMAP network. *International Journal of Enhanced Research in Science Technology & Engineering, 3*(2), 318–326. ISSN: 2319-7463.

Baker, T., Jayaraman, V., & Ashley, N. (2013). A data-driven inventory control policy for cash logistics operations: An exploratory case study application at a financial institution. *Decision Sciences, 44*(1), 205–226.

Bar-Ilan, A., Perry, D., & Stadje, W. (2004). A generalized impulse control model of cash management. *Journal of Economic Dynamics and Control, 28*(6), 1013–1033.

Basel Committee on Banking Supervision (2013). Basel III: The liquidity coverage ratio and liquidity risk monitoring tools. Basel: Bank for International Settlements. http://www.bis.org/publ/bcbs238.pdf Accessed 26.06.15.

Bátiz-Lazo, B. (2009). Emergence and evolution of ATM networks in the UK, 1967–2000. *Business History, 51*(1), 1–27. doi:10.1080/00076790802602164.

Batlin, C. A., & Hinko, S. (1982). A game theoretic approach to cash management. *The Journal of Business, 55*(3), 367–381.

Baumol, J. (1952). The transactions demand for cash: An inventory theoretic approach. *Quarterly Journal of Economics, 66*, 545–556.

Bjø rndal, E., Hamers, H., & Koster, M. (2004). Cost allocation in a bank ATM network. *mathematical methods of operations research, 59*(3), 405–418.

Blackman, I. D., Holland, C. P., & Westcott, T. (2013). Motorola's global financial supply chain strategy. *Supply Chain Management, 18*(2), 132–147. doi:10.1108/13598541311318782.

Burcea, F.-C., Bălău, V., Bâldan, C., Avrămescu, T.-C., & Ungureanu, E. (2013). Central banks between classicism and modernity. In A. Karasavvoglou, & P. Polychronidou (Eds.), *Balkan and eastern european countries in the midst of the global economic crisis* (pp. 77–85). Heidelberg: Physica-Verlag.10.1007/978-3-7908-2873-3_6

Cabello, J. G. (2013). *Cash efficiency for bank branches*. SpringerPlus. 2(1).

Cardo, J. (2009). A stochastic programming approach to cash management in banking. *European Journal of Operational Research, 192*, 963–974.

Dilijonas, D., Sakalauskas, V., Kriksciuniene, D., & Simutis, R. (2009). Intelligent systems for retail banking optimization: optimization and management of ATM network system. In J. Cordeiro, & J. Filipe (Eds.), *Proceedings of the 11th international conference on enterprise information systems, volume AIDSS, ICEIS* (pp. 321–324). Milan, Italy. may 6–10.

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

Elton, E., & Gruber, M. (1974). On the cash balance problem. *Operational Research Quarterly, 25*(4), 553–572.

Eppen, G., & Fama, E. (1968). Solutions for cash-balance and simple dynamic-portfolio problems. *The Journal of Business, 41*(1), 94–112.

de Haan, L., & van den, E. J. W. (2013). Bank liquidity, the maturity ladder, and regulation. *Journal of Banking and Finance, 37*(10), 3930–3950. doi:10.1016/j.jbankfin.2013.07.008.

Heyman, D. (1973). A model for cash balance management. *Management Science, 19*(12), 1407–1413.

Hinderer, K., & Waldmann, K.-H. (2001). Cash management in a randomly varying environment. *European Journal of Operational Research, 130*, 468–485.

Miller, M., & Orr, D. (1966). A model of the demand for money by firms. *The Quarterly Journal of Economics, 80*(3), 413–435.

Opasanon, S., & Lertsanti, P. (2013). Impact analysis of logistics facility relocation using the analytic hierarchy process (AHP). *International Transactions in Operational Research, 20*(3), 325–339. doi:10.1111/itor.12002.

Ou, C.-S., Hung, S.-Y., Yen, D. C., & Liu, F.-C. (2009). Impact of ATM intensity on cost efficiency: An empirical evaluation in taiwan. *Information & Management, 46*, 442–447.

Pokutta, S., & Schmaltz, C. (2011). Managing liquidity: Optimal degree of centralization. *Journal of Banking & Finance, 35*(3), 627–638.

Popa, V. (2013). The financial supply chain management: a new solution for supply chain resilience. *Amfiteatru Economic [Economic Amphitheatre], 15*(33), 140–153. http://www.amfiteatrueconomic.ro/temp/Article_1181.pdf Accessed 26.06.15.

Simutis, R., Dilijonas, D., Bastina, L., Friman, J., & Drobinov, P. (2007). Optimization of cash management for ATM network. *Information Technology and Control, 36*(1A), 117–121.

Snellmana, H., & Virenb, M. (2009). ATM networks and cash usage. *Applied Financial Economics, 19*(10), 841–851.

Stavins, J. (2000). ATM fees: Does bank size matter? *New England Economic Review*, 13–24.

Teddy, S. D., & Ng, S. K. (2011). Forecasting ATM cash demands using a local learning model of cerebellar associative memory network. *International Journal of Forecasting, 27*, 760–776.

Tobin, J. (1956). The interest elasticity of the transactions demand for cash. *Review of Economics and Statistics, 38*(3), 241–247.

Venkatesh, K., Ravi, V., Prinzie, A., & Van Den Poel, D. (2014). Cash demand forecasting in ATMs by clustering and neural networks. *European Journal of Operational Research, 232*, 383–392.

Yao, J.-S., Chen, M.-S., & Lu, H.-F. (2006). A fuzzy stochastic single-period model for cash management. *European Journal of Operational Research, 170*, 72–90.