On the use of multiple criteria distance indexes to find robust cash management policies

Francisco Salas-Molina[a], Juan A. Rodriguez-Aguilar[b], David Pla-Santamaria[c]

[a]Hilaturas Ferre, S.A., Les Molines, 2, 03450 Banyeres de Mariola, Alicante, Spain;
[b]IIIA-CSIC, Campus UAB, 08913 Cerdanyola, Catalonia, Spain;
[c]Universidad Politcnica de Valencia, Campus Alcoy, 03801 Alcoy, Alicante, Spain.

ABSTRACT
Cash management decision making can be handled from a multiobjective perspective by optimizing not only cost but also risk. Nevertheless, choosing the best policies under a changing context is by no means straightforward. To this end, we rely on compromise programming to incorporate robustness as an additional goal to cost and risk within a multiobjective framework. As a result, we propose to calculate robustness as a multiple criteria distance index that is able to identify the best compromise policies in terms of cost and risk. Such policies are also robust to cash flow regime changes. We show its utility by transforming the Miller and Orr's cash management model into its robust counterpart using real data from an industrial company.

1. Introduction

Financial and economic methods are constantly driven by expectations. Decision makers usually have to deal with a set of relevant variables subject to regime changes because of the economic cycle or other sources of instability. From a macro-economic approach, international economic crises, credit restrictions, or monetary policies are notably influential on decision makers. Meanwhile, from the micro-economic approach, market changes are usually the rule rather than the exception. Cash flow management is particularly affected by regime changes. A cash manager must make daily decisions about her firm's optimal cash level for operational and precautionary purposes. The focus is placed on finding the balance between cash holdings and short-term investments. A number of cash management models have been proposed to control cash balances based on a set of levels or bounds. A comprehensive review of cash management models, from the first proposals to the most recent contributions can be found in da Costa Moraes et al. (2015); Gregory (1976); Srinivasan and Kim (1986).

The particular characteristics of cash flows (or cash balances) are an important dimension of the CMP. In this sense, cash flows used in the literature range from deterministic (Baumol 1952) to completely uncertain (Miller and Orr 1966). The most

---

usual cash flow probability distribution in the literature is the Gaussian either in the form of a Wiener process (Constantinides and Richard 1978; Premachandra 2004; Baccarin 2009), or as a number of samples drawn from normal distributions with known mean and standard deviation (da Costa Moraes and Nagano 2012, 2014). However, empirical distributions obtained from real data sets are hardly used with the exception of Gormley and Meade (2007). Beyond the discussion about the most appropriate distribution, cash management models are usually designed from a stationary perspective without considering possible cash flow regime changes. However, economic cycles, market evolution and customer behavior make cash flow a random variable. Thus, optimal solutions to mathematical programming models can be importantly affected by slight perturbations in the input data (Ben-Tal and Nemirovski 1999). Hence, we say that a system is robust when it is relatively insensitive to changes in environmental factors that can negatively affect its performance (Montgomery 2013). However, whether cash management models are robust to regime changes remains a question unanswered by previous cash management contributions in the literature.

In order to provide robust solutions to optimization problems, stochastic programming and robust optimization represent two alternative approaches. Stochastic programming assumes that randomness in problem parameters is restricted to a particular probability distribution (Sahinidis 2004; Abdelaziz et al. 2007; Aouni and La Torre 2010). As an alternative approach to stochastic programming, Soyster (1973) introduced the concept of interval uncertainty within the framework of a linear program and its robust counterpart. Later on, Ben-Tal and Nemirovski (1999); Ben-Tal et al. (2009), established a sound framework for robust optimization. Briefly, robust optimization aims to solve optimization problems in which input data is uncertain. Both approaches, stochastic programming and robust optimization, aim to transform any optimization problem under uncertainty to a collection of deterministic problems that can be solved using state-of-the-art mathematical programming solvers. However, the underlying assumption on the stationarity of probability distributions and uncertainty sets represents an important limitation.

A suitable technique to validate the utility of solutions in a varying (non-stationary) context is the widely used practice in machine learning known as hold-out validation (Mitchell 1997; Hastie et al. 2005; Provost and Fawcett 2013). Hold-out validation estimates the accuracy of predictive models on a data set different from the original data sample. On the other hand, compromise programming (Ballestero and Romero 1998; Yu 2013), allows a two-step decision making process in which alternative solutions are presented before selection according to their particular preferences.

As a result, we rely on hold-out validation and compromise programming to ensure the robustness of cash management models. More precisely, we follow a data-driven approach to incorporate robustness as an additional goal to cost and risk within a multiobjective decision making framework. To this end, we assume that regime changes (if any) are contained in a test data set of past cash flows. We later compute a robustness index for cash management policies that is finally used to select the best one. Our procedure departs from previous approaches to deal with uncertainty in the sense that:

(1) we learn uncertainty from experience through a data set of past observations;
(2) we propose a data-driven approach to test robustness in cash management when possible regime changes are contained in a test data set;
(3) we introduce robustness as an additional goal to cost and risk within a multiobjective optimization framework.

For illustrative purposes, we use a real cash flow data set to select the best compromise policies in terms of cost, risk and robustness using an extended version of the one by Miller and Orr (1966). By following our procedure, we show how to transform the simple Miller-Orr's model into its robust counterpart. Summarizing, we support cash managers aiming to analyze cost and risk of cash policies in a changing environment by providing a procedure to ensure robustness in real world applications. More precisely, our contribution:

(1) follows a novel data-driven approach to deal with uncertainty in cash management without assuming any theoretical probability distribution;
(2) tests robustness of cash management models within a changing environment;
(3) transforms a cash management model into its robust counterpart.

This paper is organized as follows. We first provide in Section 2 basic background about the Miller and Orr's cash management model that will be later applied to our case study. We describe in Section 3 our multiobjective approach to the cash management problem. Section 4 introduces distance indexes as a measure of robustness and justifies its utility. Section 5 presents an illustrative example of our data-driven approach to find robust policies using real data from a Spanish industrial company. Furthermore, we discuss the advantages of our procedure with respect to alternative approaches in Section 6. Finally, we conclude and discuss further research in Section 7.

2.   The Miller-Orr's cash management model

The first approach to the cash management problem considering stochastic cash flows was formulated by Miller and Orr (1966). They developed a simple model in which cash flow is characterized as a sequence of independent and symmetric Bernoulli trials. They assumed that cash balance either increases or decreases by m dollars with probability $p = 1/2$. Main assumptions of this approach are independence, stationarity, zero-drift, and the absence of regular swings in the daily cash flow. Moreover they ignored shortage and variable transaction costs. The rationale for selecting the Miller-Orr model in our study is double. First, its relevance. It constitutes a widely known and referenced stochastic approach that has been used as a foundation for the construction of many modern cash management models. Second, its simplicity. It allows us to illustrate that our proposal is able to transform a simple cash management model into a robust one.

The Miller-Orr model sets its policy based on equation (1) below. The transfer (x_t) occurring at day t is computed by comparing the cash balance at the beginning of the day (b_{t-1}) to a low and a high bound, denoted as l and h, respectively. Although Miller and Orr initially obtained optimal values setting the low bound to zero, a version of the model with a low bound distinct from zero can be found in Ross et al. (2002) as follows:

$$x_t = \begin{cases} z - b_{t-1}, & \text{if} \quad b_{t-1} > h \\ 0, & \text{if} \quad l \le b_{t-1} \le h \\ z - b_{t-1}, & \text{if} \quad b_{t-1} < l \end{cases} \qquad (1)$$

where z is a target cash balance. As shown in Figure 1, cash balance is allowed to wander around between bounds l and h. When h is reached a withdrawal transfer is made to restore the balance to target level z. Analogously, when cash balance reaches

3

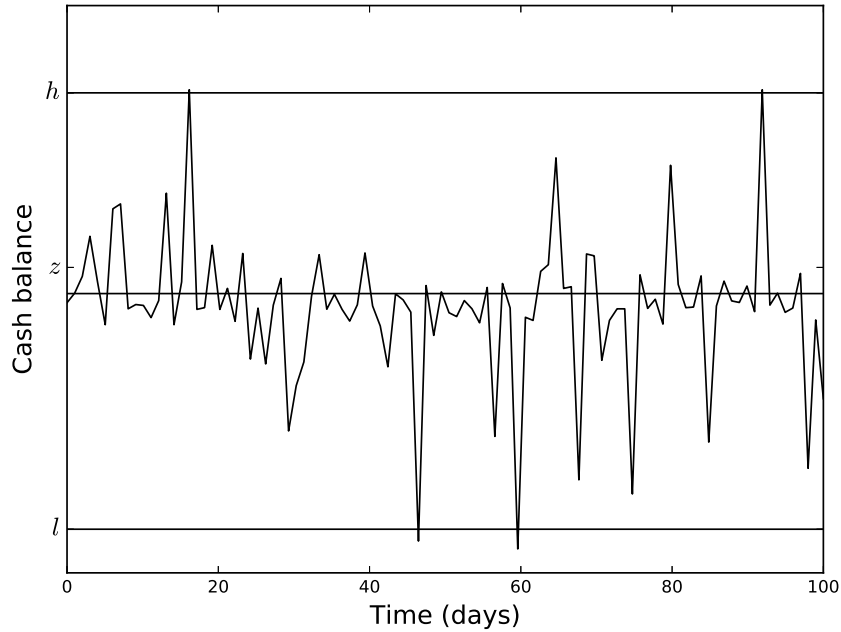l, a positive transfer is made to restore again the balance to z.

Figure 1 Example of cash balance evolution over time.

By relaxing most of Miller and Orr's assumptions, such as the use of shortage and variable transaction costs, we will later use their model in order to select robust policies determined by some control limits h, z, and l.

3. The multiobjective cash management problem definition

In this section, we formulate the multiobjective cash management problem (MOCMP) as proposed by Salas-Molina et al. (2016). Within a single objective framework, the cash management problem (CMP) is defined as an optimization problem whose goal is to find the best sequence of transactions $X = \langle x_1, x_2, \ldots, x_T \rangle \in \mathbb{R}$, what is called a policy, that minimizes a cost function $C(X, T)$ over a time horizon of T days. The CMP is characterized by its particular cost structure (Penttinen 1991; Gormley and Meade 2007). Any ordering transaction into a cash account may have a cost, which may include a fixed part (γ_0^+) and a variable part (γ_1^+). On the other hand, a return transaction from a cash account may also have a cost with a fixed part (γ_0^-) and a variable part (γ_1^-). Furthermore, at the end of the day, a holding cost v per money unit is charged if a positive cash balance occurs, or a penalty cost u per money unit is charged if a negative cash balance occurs. According to this cost structure, a general daily cost function can be defined as:

$$C_t(x_t, b_t) = \Gamma(x_t) + L(b_t) \tag{2}$$

where x_t is the transaction made at day t, b_t is the cash balance at the end of day t, $\Gamma(x_t)$ is a transfer cost function, and $L(b_t)$ a holding/shortage cost function. The

4

transfer cost function $\Gamma(x_t)$ is defined as:

$$\Gamma(x_t) = \begin{cases} \gamma_0^- - \gamma_1^- \cdot x_t & \text{if} \quad x_t < 0, \\ 0 & \text{if} \quad x_t = 0, \\ \gamma_0^+ + \gamma_1^+ \cdot x_t & \text{if} \quad x_t > 0. \end{cases} \tag{3}$$

Additionally, the holding/shortage cost function is expressed as:

$$L(b_t) = \begin{cases} -u \cdot b_t & \text{if} \quad b_t < 0; u > 0, \\ v \cdot b_t & \text{if} \quad b_t > 0; v > 0 \end{cases} \tag{4}$$

Under this cost structure, the ultimate goal of the MOCMP is to find the policy X that minimizes the expected cost and risk over the time horizon T. To this end, cost $C(X,T)$ is measured by the average daily cost:

$$C(X,T) = \frac{1}{T} \sum_{t=1}^{T} C_t(x_t, b_t) = \frac{1}{T} \sum_{t=1}^{T} [\Gamma(x_t) + L(b_t)] \tag{5}$$

and risk $R(X,T)$ is measured by the standard deviation of the daily cost:

$$R(X,T) = \left(\frac{1}{T} \sum_{t=1}^{T} (C(X,T) - C_t(x_t, b_t))^2 \right)^{1/2}. \tag{6}$$

Once the objective functions for cost and risk are defined, Salas-Molina et al. (2016) rely on compromise programming (Zeleny 1982; Ballestero and Romero 1998; Yu 2013), to solve the MOCMP by deriving an efficient frontier with the set of not-dominated policies evaluated in terms of cost and risk. Indeed, all policies are dominated by the policy with the lowest risk from the cost perspective. Likewise, from the risk perspective, all policies are dominated by the policy with the lowest cost. Then, cash managers have to choose a policy from the efficient frontier according to their risk/cost preferences. This is attained through the use of two normalized indexes as suggested by Ballestero and Romero (1998). First, a cost index θ_1, defined as:

$$\theta_1(X,T) = \frac{C(X,T) - C_{min}}{C_{max} - C_{min}} \tag{7}$$

where $C(X,T)$ is the expected daily cost over T days computed for a particular policy X using equation (5), and C_{max} and C_{min} are, respectively, the maximum and minimum daily costs. Second, a risk index, θ_2, defined as:

$$\theta_2(X,T) = \frac{R(X,T) - R_{min}}{R_{max} - R_{min}} \tag{8}$$

where $R(X,T)$ is a risk measure over T days computed for a particular policy X using equation (6), and R_{max}, R_{min} are, respectively, the maximum and minimum values for R. Note that θ_1 and θ_2 range in the interval [0,1], and that the ideal point with minimum cost and risk, $(\theta_1, \theta_2) = (0,0)$, is usually unfeasible. However, the closer to this ideal point, the better the solution. A family of normalized distance functions to

the ideal point that includes weights determining the decision maker's risk preferences can be expressed as:

$$\mathcal{L}_h = \left[w_1^h \cdot \theta_1^h + w_2^h \cdot \theta_2^h \right]^{1/h}. \tag{9}$$

Note that $\mathcal{L}_1$ is the Manhattan distance, $\mathcal{L}_2$ is the Euclidean distance, and $\mathcal{L}_\infty$ is the Chebyshev distance. In addition, weights $w_1, w_2 \in [0, 1]$ reflect the particular preferences of cash managers when $w_1 + w_2 = 1$. Then, under the framework of compromise programming for $\mathcal{L}_1$, the goal is to find the policy X that minimizes the weighted Manhattan distance to the ideal point $(0, 0)$:

$$\min \ \ [w_1 \cdot \theta_1(X, T) + w_2 \cdot \theta_2(X, T)] \tag{10}$$

subject to the following cash balance state equation:

$$\hat{b}_{t-1} + \hat{f}_t + x_t = \hat{b}_t \tag{11}$$

where $\hat{f}_t$ and $\hat{b}_t$ denote predicted cash flows and balances, respectively. Since cash managers make decisions in advance, predicted instead of actual values are used for random cash flows and balances resulting in an uncertain optimization problem. However, since neither θ_1 nor θ_2 are able to assess robustness of alternative cash policies, we require a new measure to do so. Next, we consider such a measure.

4. On the use of distances as a measure of robustness

A system or a process is robust when it is relatively insensitive to changes in environmental factors, operating conditions and components that can negatively affect its performance (Montgomery 2013). Therefore, cash management models considering cost and risk objectives must also be evaluated in terms of robustness to deal with changing conditions. Let us illustrate this concept with an example. Consider that an empirical probability distribution $p(b)$ of a cash balance b can be derived from an initial cash balance and a set F of observed past values of cash flows. A usual assumption is that the stochastic cash balance is stationary, i.e., the main attributes of $p(b)$, such as its mean and its standard deviation, do not vary with time. However, chances are that cash flow regime changes lead to variations in both cost and risk of a particular cash policy. For example, assume that some $p(b)$ for a particular company is characterized by a mean value of 80 and a standard deviation of 20. Also assume that, under the framework of the Miller and Orr's cash management model, the low bound of the best policy is set to 30. This policy clearly becomes suboptimal if the standard deviation rises to 40 due to, for example, a negative market evolution in the last 6 months. In what follows, we first describe our data-driven procedure to ensure robustness of cash policies. Second, we propose the use of a robustness index as a key input to the previous procedure.

4.1. A data-driven procedure to ensure robustness

Within the MOCMP in Salas-Molina et al. (2016), policies are evaluated through cost and risk indexes θ_1 and θ_2, introduced in Section 3. However, these indexes do not

provide information about policy robustness. In order to ensure robustness in cash management, we here rely on a data-driven procedure. More precisely, we apply the widely used practice in the field of machine learning consisting of dividing the available data in a training set and a test set (Mitchell 1997; Hastie et al. 2005; Provost and Fawcett 2013). The utility of predictive models is then assessed by estimating the accuracy of forecasts on a data set (test set) different from the original data sample (training set). In the context of the MOCMP, we propose to use a training data set as a business-as-usual situation, and a test set that may contain changes in the context faced by cash managers. As long as the training set is statistically different to the test set (but possibly many others), we are in a position to measure policy robustness as the change in performance over two different data sets. Thus, we define policy performance as follows:

Definition 1. (**Policy performance**). The performance P_i of a cash policy i in a normalized cost-risk space is given by the pair of points $P_i = (\theta_{1i}, \theta_{2i})$, computed using (7) and (8), over a given cash flow data set.

Recall that we do not assume any cash flow probability distribution but we are provided with a data set of past observations. Since we are dealing with cash flow time-series, we may be reasonably interested in evaluating the effect of recent cash flow data on policy performance. Thus, we perform an empirical performance test with the most up-to-date data. If any regime change occurred in recent past, it will be reflected in the performance of policies derived from solving the MOCMP. A feasible solution to the MOCMP is a policy derived from a cash management model (e.g., a policy of the form in equation (1) for the Miller and Orr 1966 model), that satisfies the cash balance state equation (11). Hence, given a feasible policy derived from a particular cash management model, we assess its robustness by computing the difference in performance over a training set (e.g., with the first 80 % of the observations), and over a test set (with the remaining 20% of the observations). Summarizing, our data-driven procedure to ensure robustness in cash management comprises the following steps:

(1) Select a cash management model.
(2) Generate a set of alternative feasible policies.
(3) Divide the cash flow data set in a training set and a test set.
(4) Evaluate policy performance in the training set.
(5) Evaluate policy performance in the test set.
(6) Measure robustness through a new index that considers policy performance with respect to the training and test sets.
(7) Select the best policies using a robustness index.

Next, we propose a collection of distances to measure policy robustness in a multi-objective framework in which cost and risk are goals to minimize.

4.2. *Measuring robustness through distance indexes*

In order to derive a robustness index for cash policies, we rely on the concept of distance. A comprehensive work on distances can be found in Deza and Deza (2014). We here define distance as a function in a normalized cost-risk space:

Definition 2. (**Distance**). In a normalized cost-risk space $S = [0, 1]^2$, a distance is a function $D : S \times S \to \mathbb{R}^+$, with the following properties:

(1) $D(P, P') \geq 0$ (non-negativity);
(2) $D(P, P') = D(P', P)$ (symmetry);
(3) $D(P, P) = 0$ (reflexivity);
(4) $D(P, P'') \leq D(P, P') + D(P', P'')$ (triangle inequality);

for all points $P, P', P'' \in S$.

Hereafter, we propose to test the robustness of policies to cash flow regime changes by computing distances between performance points derived from policies in a training set denoted as $P_i = (\theta_1, \theta_2)$, and performance points derived from policies in a test set denoted as $P'_i = (\theta'_1, \theta'_2)$. Thus, we propose to use the following collection of distances to measure robustness of policies that we illustrate in Figure 2. In the business-as-usual context, policy performance can be synthetically evaluated in terms of cost and risk by means of the (training) distance to the ideal performance $(0, 0)$.

Definition 3. (**Training distance**). The training distance δ_i of policy i is the Euclidean distance between performance P_i and the ideal policy performance $(0, 0)$.

The shorter the training distance the better the policy within a business-as-usual context. However, regime perturbations may result in performance changes of alternative policies. Then, we introduce the concept of deviation as the difference in performance between policies over the training and test sets in a normalized cost-risk space.

Definition 4. (**Deviation**). The deviation d_i of policy i is the Euclidean distance between performance P_i and performance P'_i, namely:

$$d_i = \sqrt{(\theta_{1i} - \theta'_{1i})^2 + (\theta_{2i} - \theta'_{2i})^2}. \tag{12}$$

According only to the magnitude of deviation, we can initially state that policy 2 in Figure 2, with deviation d_2, is more robust than policy 3, with deviation d_3, because $d_2 < d_3$. However, although d_1 and d_2 have the same magnitude, a rational cash manager should select policy 2 instead of 1 because the former is closer to the ideal point $(0, 0)$ than the latter. This motivates the need for an additional distance measure.

Definition 5. (**Proximity or test distance**). The proximity δ'_i of policy i is the Euclidean distance between performance P'_i and the ideal policy performance $(0, 0)$.

In addition, a policy that improves its performance after a context change is preferred to a policy that worsens its performance. For example, policy 2 in Figure 2 is an even better policy than policy 1 because, after the evaluation over the test set, performance P'_2 moves closer to the ideal point while P'_1 moves away from the ideal point. Thus, we should also measure movement of policies.

Definition 6. (**Movement**). The movement of policy i is the difference in distance to the ideal performance between policy performance P'_i and P_i, computed as $\delta'_i - \delta_i$.

We next summarize how each of the aforementioned distances to measure robustness rank policies when the rest of factors (distances) are equal:

- Deviation: policy i is better than policy j if $d_i < d_j$.
- Proximity: policy i is better than policy j if $\delta'_i < \delta'_j$.

- Movement: policy i is better than policy j if $\delta'_i - \delta_i < \delta'_j - \delta_j$.
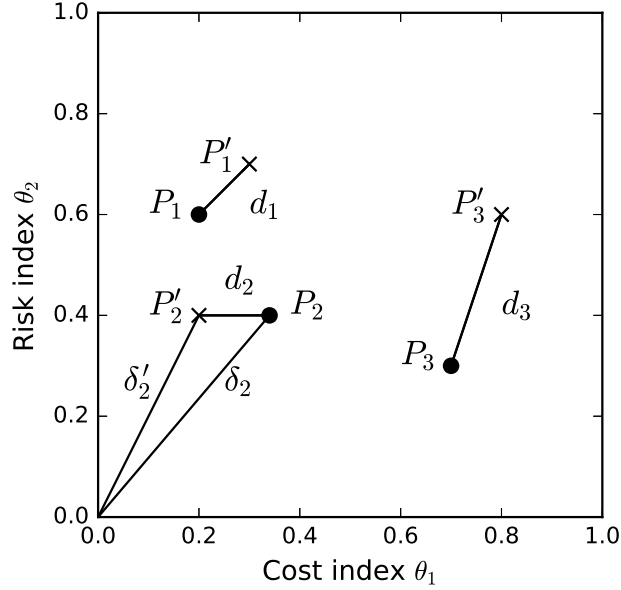

Figure 2 Distances between cash policies evaluated over two data sets.

The collection of distances mentioned above allows us to introduce different measures of robustness according to the particular preferences of cash managers. Next, we introduce two alternative distance functions to measure robustness that we later compare in a numerical example:

$$D_1(P_i, P'_i) = \begin{cases} 2\alpha(\delta_i - \delta'_i) & \text{if} \quad \delta_i > \delta'_i, \\ 2(1-\alpha)(\delta'_i - \delta_i) & \text{if} \quad \delta_i \le \delta'_i \end{cases} \tag{13}$$

as a linear asymmetric loss function, with $\alpha \in [0, 1]$, and:

$$D_2(P_i, P'_i) = \alpha d_i + \beta \delta'_i + (1 - \alpha - \beta)(\sqrt{2} + \delta'_i - \delta_i) \tag{14}$$

as a synthetic loss function where $\alpha, \beta \in [0, 1]$ are weights to set preferences objectives, and $\sqrt{2}$ is only added to ensure non-negativity. Notice that (13) considers only movement but (14) takes into account deviation, proximity and movement. Cash managers can also consider their particular preferences by tuning parameters α and β. In (13), parameter α can be tuned to weigh movement. Alternatively, in (14), parameter α is linked to deviation in performance between regimes so that the higher the value of α the higher the importance of deviation. Parameter β is linked to the proximity of policies to the ideal point so that the higher the value of β the higher the importance of proximity. Finally, the values given to α and β determine the importance of the movement, since it is weighed by $1 - \alpha - \beta$.

Next, we apply compromise programming (Ballestero and Romero 1998) to find robust policies. Recall that this approach is based on the concept of ideal point where both the minimum cost and the minimum risk occur simultaneously, i.e., $(0,0)$. This point is usually unfeasible and it is necessary to look for compromise solutions that

can be evaluated by computing the distance to the ideal point. Thus, once an entire cash flow data set is divided in a training set and a test set, we first evaluate a given set of feasible policies over the training set to obtain performance points $P_i = (\theta_{1i}, \theta_{2i})$. However, when selecting a policy from the whole cost-risk space, cash managers may be interested in testing the robustness to cash flow regime changes. To this end, we propose the use of a third index θ_3, namely, a distance index as a measure of robustness when applied to data not considered in the first selection of policies. As a result, given a policy X_i, with expected performance $P_i = (\theta_{1i}, \theta_{2i})$, obtained from the training set, and with empirically tested performance $P'_i = (\theta'_{1i}, \theta'_{2i})$, obtained from the test set, we use a distance function (e.g., (14)), to compute the following distance index:

$$\theta_3(X_i) = \frac{D(P_i, P'_i) - D_{min}}{D_{max} - D_{min}} \tag{15}$$

where D_{max} and D_{min} are, respectively, the maximum and minimum values of performance distances $D(P_i, P'_i)$ between the training and the test set for each policy i. At this point, we are in a position to derive an efficient frontier with the set of policies that are not dominated by any other policy in terms of cost and robustness. Graphically, the efficient frontier in the normalized cost-robustness space, is built from a set of two-dimensional points (θ_1, θ_3), as in Figure 3. Note that since we compute robustness using both cost and risk indexes, we are implicitly considering risk in the selection of policies. As a result, by following our distance-based procedure, cash managers aiming to analyze cost and risk of cash management policies in a changing environment are able to ensure robustness in real world applications. Next, we illustrate our procedure and its benefits through a numerical example.

5. An illustrative example using the Miller and Orr's model

In this section, we follow the procedure described in Section 4.1 to determine the best compromise cash policies in terms of cost and risk that are also robust to cash flow regime changes.

5.1. *Empirical settings*

In what follows, we use the Miller and Orr's cash management model introduced in Section 2 to obtain control limits h, z and l, that determine policies according to equation (1). In our case study, we use a real data set from a Spanish industrial company composed by 1000 observations of cash flows at non-bank-holidays for a period of about 4 years. We divide the entire data set in a training set with the first 80% of the observations, as the expected conditions, and a test set with the remaining 20%, as representative of changes in current conditions.

For illustrative purposes, we make computations using a cost structure adjusted to current bank practices in Spain, summarized as follows: daily holding cost, $v = 0.1\%$; daily shortage cost, $u = 30\%$, to be used in equation (4); fixed transfer cost into account, $\gamma_0^+ = 5$ €; fixed transfer from account $\gamma_0^- = 5$ €; variable transfer cost into account, $\gamma_1^+ = 0$ €, and variable transfer cost from account, $\gamma_1^- = 0$ €, to be used in equation (3). It is important to note that $\gamma_1^- < v$ and $\gamma_1^+ < u$. Otherwise, no transaction would be made since the unitary costs of transferring money are higher than those of holding the same amount of money.

5.2. *Selecting robust policies*

In order to derive the efficient frontier in the cost-robustness space, shown in Figure 3, we develop a complete grid search over different policies derived from the Miller and Orr's model. These policies are determined by the control limits h, z, and l. Since our cost structure presents a high shortage cost, we expect better policies to be given by high control limits. Consequently, we iterate over feasible combinations of h, z, and l subject to $h \geq z \geq l$, ranging in $[h - \sigma, h + 3\sigma], [z - \sigma, z + 3\sigma], [l - \sigma, l + 3\sigma]$, in steps of 100000 €, where σ is the standard deviation of cash flows in the training set.
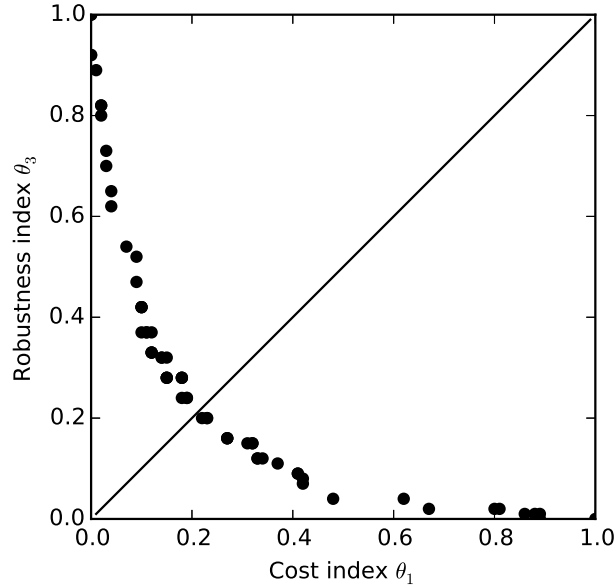

Figure 3 Efficient frontier in the cost-robustness space (θ_1, θ_3).

In this paper, we present alternative policies before selection. Therefore, we evaluate all possible policies in terms of cost and risk. Under the compromise programming framework, the closer to the ideal point $(0, 0)$, the better the policy. A sample of the best policies of the efficient frontier, in terms of Euclidean distance to the ideal is shown in Table 1. To compute θ_3 we use distance function (14) with parameters $\alpha = 0.33$ and $\beta = 0.33$, to equally weigh deviation, proximity and movement. Then, a cash manager would select either policy 35, 38, 39 or 40 to minimize the distance to the ideal $(0, 0)$. However, the particular preferences of cash managers should be considered to select other policies. A risky cash manager may select one of the policies with larger distances (e.g., policy number 32), in order to reduce daily cost. On the other hand, a conservative cash manager may consider accepting a higher expected cost (e.g., by selecting policy 45), in order to maximize robustness.

5.3. *Analysis*

In this example, we use a robustness index to select the best compromise policy in terms of cost and robustness to regime changes. However, an additional evaluation in terms of cost and risk may be useful for comparative purposes. Such an evaluation over the test set in our illustrative example is shown in Table 2. As a baseline, we use the

Table 1 Central sample of the efficient frontier (Id=policy id; Bounds h, z and l in thousands of euros; $D_2(P_i, P'_i)$ = distance function; θ_1 = cost index; θ_3 = robustness index; Dist=Euclidean distance to the ideal point $(\theta_1, \theta_3) = (0,0)$).

Id	High(h)	Target(z)	Low(l)	$D_2(P_i, P'_i)$	θ_1	θ_3	Dist
32	6310	4760	1780	0,05	0,18	0,28	0,33
33	6410	4760	1780	0,05	0,18	0,28	0,33
34	6510	4760	1780	0,05	0,18	0,28	0,33
35	6210	4660	1580	0,05	0,18	0,24	0,30
36	6310	4660	1580	0,04	0,19	0,24	0,31
37	6410	4660	1580	0,04	0,19	0,24	0,31
38	6110	4560	1480	0,04	0,22	0,20	0,30
39	6210	4560	1480	0,04	0,23	0,20	0,30
40	6310	4560	1480	0,04	0,23	0,20	0,30
41	6010	4460	1380	0,03	0,27	0,16	0,31
42	6110	4460	1380	0,03	0,27	0,16	0,31
43	6210	4460	1380	0,03	0,27	0,16	0,31
44	6010	4460	1480	0,03	0,31	0,15	0,34
45	6110	4460	1480	0,03	0,32	0,15	0,35

best compromise policy obtained using the entire data set but without any robustness index as in Salas-Molina et al. 2016. In addition, we evaluate three different distance measures: (i) the asymmetric loss function in (13) with $\alpha = 0.4$ to slightly underweigh the loss of policies that move closer to the ideal point and yield an improvement in cost; (ii) the synthetic loss function in (14) considering both deviation and proximity leading to a smaller improvement than the asymmetric function and an increase in risk; (iii) the synthetic loss function in (14) but considering deviation, proximity and movement, resulting in a reduction in cost but an increase in risk. Summarizing, the use of a robustness index in our example results in a cost reduction of policies. In addition, considering deviation, proximity and movement produces the best results in terms of cost. However, this cost reduction is only possible by accepting a higher level of risk.

Table 2 Evaluation of policies using different distance measures. Bounds h, z and l in thousands of euros.

Distance	High(h)	Target(z)	Low(l)	α	β	Cost(%)	Risk(%)
Salas-Molina et al. (2016)	6570	4910	3930	-	-	100	100
Asymmetric loss D_1 in (13)	6010	4460	3580	0,4	-	91	99
Synthetic loss D_2 in (14)	6610	4960	3380	0,5	0,5	94	112
Synthetic loss D_2 in (14)	6110	4560	1480	0,33	0,33	75	132

6. Discussion

In this section, we discuss our distance-based procedure in comparison to alternative approaches to obtain robust solutions to optimization problems such as stochastic programming and robust optimization. In this paper, we handle uncertainty from an empirical or data-driven approach which considerably departs from previous approaches. More precisely, the use of a distance index as a measure of robustness of cash policies allows to include an additional goal to cost and risk within the framework of multiobjective decision making. This additional goal can be designed according to the particular preferences of decision makers. Furthermore, since the selection of the best

compromise policies is affected by the uncertainty in cash flows, our approach is specially designed to deal with input data that can be learned from experience through a data set of past observations such as cash flows. At the same time, we follow a hold-out validation procedure based on the comparative analysis of a training data set and a test set that may contain a context change. Finally, we allow practitioners to transform any cash management model into its robust counterpart by following our distance-based approach to find robust cash management policies.

For comparative purposes, let us consider two different approaches to deal with uncertain optimization problems, namely, stochastic programming and robust optimization. Under a general stochastic programming approach (Prékopa 2013), a deterministic counterpart of the program (10)-(11) can be considered as follows:

$$\min \quad [w_1 \cdot E\left(\theta_1(X,T)\right) + w_2 \cdot E\left(\theta_2(X,T)\right)] \tag{16}$$

subject to:

$$P(\hat{b}_{t-1} + \hat{f}_t + x_t \geq \hat{b}_t) \geq \zeta \tag{17}$$

where E is the common expectation operator, P denotes probability and ζ is a safety threshold determined by a cash manager. A typical value for ζ is 95%, meaning that cash balances are above the expected values at least 95% of the times, reducing then the possibility of an overdraft. In other words, by solving the previous program we are forcing cash balances to be above a certain level determined by the safety threshold ζ. However, the introduction of probabilities implies the assumption of a particular distribution for cash flows resulting in two serious limitations due to: (i) the lack of empirical evidence on the common assumption of normally distributed cash flows (Emery 1981; Pindado and Vico 1996); (ii) the additional assumption on the stationarity of cash flow distributions (Constantinides and Richard 1978; Premachandra 2004; Baccarin 2009).

An alternative approach was proposed by Soyster (1973) and, later on, by Ben-Tal and Nemirovski (1999); Ben-Tal et al. (2009), who tackled uncertainty by proposing a robust counterpart to an optimization problem when uncertainty is determined either by a system of linear inequality constraints or by a system of conic quadratic inequalities. For instance, a robust counterpart of the program (10)-(11), when uncertainty in cash flows is determined by $\hat{f} - \varepsilon \leq f \leq \hat{f} + \varepsilon$, being ε the maximum forecasting error, is given by:

$$\min \quad [w_1 \cdot \theta_1(X,T) + w_2 \cdot \theta_2(X,T)] \tag{18}$$

subject to:

$$\hat{b}_{t-1} + \hat{f}_t + x_t \geq \hat{b}_t + \varepsilon. \tag{19}$$

The assumption of a maximum forecasting error ε may lead either to unrealistic or ultraconservative problem formulations. For instance, consider a data set of empirical errors ranging in $[-100, 100]$, in which the maximum negative error -100 occurred only once out of 1 million times. Setting ε to 100 results in an ultraconservative strategy. In addition, two important issues remain unsolved: (i) the stationarity assumption, meaning that estimation errors keep unaltered; and (ii) the possibility of adapting cash

policies to different scenarios according to the particular preferences of decision makers. These limitations are solved by our data-driven distance procedure as described above.

7. Conclusions

In the cash management problem, decision making can be enhanced by focusing on optimizing cost, risk and robustness of the available policies rather than just cost. This is particularly true under the realistic assumption of time-varying circumstances. Usual approaches to deal with uncertainty within the framework of optimization problems such as stochastic programming or robust optimization present limitations such as imposing a theoretical probability distribution or the stationarity assumption. To overcome these limitations, we propose a data-driven procedure to test the robustness of cash policies in a multicriteria decision making process. Particularly, we introduce the use of a distance-based robustness index to select the best compromise set of cash policies when cash flow regime changes are learned from experience through the use of a data set of past observations. As a result, we argue that our procedure is able to transform a cash management model into its robust counterpart when a data set of past cash flow observations is available through:

- the validation estimated cost-risk results from cash policies when possible regime changes are contained in a test set;
- the use of distance indexes to measure policy robustness;
- the selection of the best compromise policies in terms of cost and risk that are also robust to regime changes through compromise programming.

A final remark must be done in the sense that we claim that a new procedure is available to optimize cost and robustness by using two different data sets, a training set and a test set. Our procedure can be replicated as many times as needed by considering alternative (and possibly very different) test sets to incorporate a number of future scenarios. In addition, our procedure can be straightforwardly extended to the use of any other cross-validation method. However, further research would be needed to test the ability in reducing both cost and risk in a systematic way by reducing distances as a measure of robustness. In that case, the procedure presented in this paper could help.

References

Abdelaziz, F. B., Aouni, B., and El Fayedh, R. (2007). Multi-objective stochastic programming for portfolio selection. *European Journal of Operational Research*, 177(3):1811–1823.

Aouni, B. and La Torre, D. (2010). A generalized stochastic goal programming model. *Applied Mathematics and Computation*, 215(12):4347–4357.

Baccarin, S. (2009). Optimal impulse control for a multidimensional cash management system with generalized cost functions. *European Journal of Operational Research*, 196(1):198–206.

Ballestero, E. and Romero, C. (1998). *Multiple criteria decision making and its applications to economic problems*. Springer Science & Business Media.

Baumol, W. J. (1952). The transactions demand for cash: An inventory theoretic approach. *The Quarterly Journal of Economics*, 66(4):545–556.

Ben-Tal, A., El Ghaoui, L., and Nemirovski, A. (2009). *Robust optimization*. Princeton University Press.

Ben-Tal, A. and Nemirovski, A. (1999). Robust solutions of uncertain linear programs. *Operations research letters*, 25(1):1–13.

Constantinides, G. M. and Richard, S. F. (1978). Existence of optimal simple policies for discounted-cost inventory and cash management in continuous time. *Operations Research*, 26(4):620–636.

da Costa Moraes, M. B. and Nagano, M. S. (2012). Cash balance management: A comparison between genetic algorithms and particle swarm optimization. *Acta Scientiarum. Technology*, 34(4):373–379.

da Costa Moraes, M. B. and Nagano, M. S. (2014). Evolutionary models in cash management policies with multiple assets. *Economic Modelling*, 39:1–7.

da Costa Moraes, M. B., Nagano, M. S., and Sobreiro, V. A. (2015). Stochastic cash flow management models: A literature review since the 1980s. In *Decision Models in Engineering and Management*, pages 11–28. Springer International Publishing.

Deza, M. M. and Deza, E. (2014). *Encyclopedia of distances*. Springer, third edition.

Emery, G. W. (1981). Some empirical evidence on the properties of daily cash flow. *Financial management*, 10(1):21–28.

Gormley, F. M. and Meade, N. (2007). The utility of cash flow forecasts in the management of corporate cash balances. *European journal of operational research*, 182(2):923–935.

Gregory, G. (1976). Cash flow models: a review. *Omega*, 4(6):643–656.

Hastie, T., Tibshirani, R., Friedman, J., and Franklin, J. (2005). The elements of statistical learning: data mining, inference and prediction. *The Mathematical Intelligencer*, 27(2):83–85.

Miller, M. H. and Orr, D. (1966). A model of the demand for money by firms. *The Quarterly journal of economics*, 80(3):413–435.

Mitchell, T. M. (1997). *Machine learning*. McGraw-Hill.

Montgomery, D. C. (2013). *Design and analysis of experiments*. Wiley, eighth edition.

Penttinen, M. J. (1991). Myopic and stationary solutions for stochastic cash balance problems. *European journal of operational research*, 52(2):155–166.

Pindado, J. and Vico, J. (1996). Evidencia empírica sobre los flujos de caja. un nuevo enfoque en su tratamiento. *Revista Española de Financiación y Contabilidad*, 25(87):497–517.

Prékopa, A. (2013). *Stochastic programming*, volume 324. Springer Science & Business Media.

Premachandra, I. (2004). A diffusion approximation model for managing cash in firms: An alternative approach to the Miller-Orr model. *European Journal of Operational Research*, 157(1):218–226.

Provost, F. and Fawcett, T. (2013). Data science for business: What you need to know about data mining and data-analytic thinking.

Ross, S. A., Westerfield, R., and Jordan, B. D. (2002). *Fundamentals of corporate finance*. McGraw-Hill, sixth edition.

Sahinidis, N. V. (2004). Optimization under uncertainty: state-of-the-art and opportunities. *Computers & Chemical Engineering*, 28(6):971–983.

Salas-Molina, F., Pla-Santamaria, D., and Rodriguez-Aguilar, J. A. (2016). A multi-objective approach to the cash management problem. *Annals of Operations Research*, pages 1–15.

Soyster, A. L. (1973). Technical noteconvex programming with set-inclusive constraints and applications to inexact linear programming. *Operations research*, 21(5):1154–1157.

Srinivasan, V. and Kim, Y. H. (1986). Deterministic cash flow management: state of the art and research directions. *Omega*, 14(2):145–166.

Yu, P.-L. (2013). *Multiple-criteria decision making: concepts, techniques, and extensions*, volume 30. Springer Science & Business Media.

Zeleny, M. (1982). *Multiple criteria decision making.* McGraw-Hill.