Research article

Multiskilled personnel assignment problem under uncertain demand: A benchmarking analysis

César Augusto Henao[1],*, Ana Batista[2], Andrés Felipe Porto[3] and Virginia I. González[1]

[1] Department of Industrial Engineering, Universidad del Norte, Km. 5 Vía Puerto Colombia, Barranquilla, Colombia

[2] Industrial and Systems Engineering Department, Pontificia Universidad Católica de Chile, Santiago, Chile

[3] Department of Industrial Engineering, Corporación Universitaria Americana, Barranquilla, Colombia

* Correspondence: Email: cahenao@uninorte.edu.co.

Abstract: The personnel assignment problem in different service industries aims to minimize the staff surplus/shortage costs. However, uncertainty in the staff demand challenges the accomplishment of that objective. This research studies the personnel assignment problem considering uncertain demand and multiskilled workforce configured through a 2-chaining strategy. We develop a two-stage stochastic optimization (TSSO) approach to calculate the multiskilling requirements that minimize the training costs and the expected costs of staff surplus/shortage. Later, we evaluate and compare the performance of the TSSO approach solutions with the solutions of two alternative optimization approaches under uncertainty - robust optimization (RO) and closed-form equation (CF). These two alternative approaches were published in Henao et al. [1] and Henao et al. [2], respectively. In addition, we compare the performance of the TSSO approach solutions with the solution of the deterministic (DT) approach and the solutions of myopic multiskilling approaches. To make performance comparisons between the different approaches, we used both real and simulated data derived from a retail store operating in Chile. The results show that, for different demand variability levels, TSSO, RO, and CF always belong to the set of approaches with the lowest average total cost. That is, in this group, there are no statistical differences from one approach to another, so these approaches are the most cost-effective. We also provide insights to retail decision-makers for addressing two key aspects. First, the methodology allows to address two fundamental multiskilling issues: how much multiskilling to add and how it should be added. Second, it is provided understanding on how to select the most

suitable approach according to the balance between the conservatism and the reliability associated with the solutions delivered by each approach. Finally, we identify some methodological challenges for future research, such as the evaluation of k-chaining strategies with $k \geq 2$.

Keywords: workforce flexibility; multiskilling; cross-training; 2-chaining; stochastic optimization; robust optimization; retail

1. Introduction

Retail companies must face constant changes in staffing demand caused by the seasonality and uncertainty of demand [1–4]. In general, retailers seek to reduce their labor costs and simultaneously improve the service levels offered to their customers [5,6]. However, even retail companies with effective personnel scheduling tools can have difficulty in managing staff surpluses or shortages due to the seasonality and uncertainty of demand [7–9]. Various authors addressed the problems of staff surpluses and shortages through the use of an optimal multiskilled staff (e.g., [10–13]).

In a retail store, multiskilled staff can work in multiple departments, allowing store supervisors to relocate multiskilled workers from departments with staff surpluses to departments with staff shortages [14]. Several articles have reported that the most profitable staff is one where multiskilled workers are skilled to perform up to a total of two store departments (e.g., [15–17]). This is known in the multiskilling literature as a 2-flexibility policy, and based on this policy, the 2-chaining strategy has been the most recommended to design training plans (e.g., [18–21]).

In a retail store context, 2-chaining requires that a set of workers be skilled up to a maximum of two departments and training schemes must be designed through closed long chains and/or closed short chains. In one closed long chain (CLC), each store department has one multiskilled worker, such that these workers connect all the departments through a single closed chain. In turn, in one closed short chain (CSC), only a subset of departments is connected through a closed chain, since only some of them have multiskilled workers. Particularly, Henao et al. [1,2] and Porto et al. [6,21] indicated that staff training plans built through of CLCs and CSCs provide maximum efficiency to assign and reassign workers in scenarios with unscheduled staff absenteeism and/or uncertainty in the staff demand.

In Figure 1 three examples of 2-chaining structures are shown. In each of the three structures, demand nodes (departments) are in the middle, whereas supply nodes (workers) are positioned on the right (single-skilled workers) and the left (multiskilled workers). Figure 1(a) shows that in one CLC, there is only one multiskilled worker at each store department. In contrast, Figure 1(b) shows that in two similar CLCs, each department has two multiskilled workers forming two CLCs that are indistinguishable. Finally, Figure 1(c) shows three multiskilled workers connecting departments 1, 2, and 3 by using of one CLC, and also shows two additional multiskilled workers connecting departments 2 and 3 by using of one CSC.
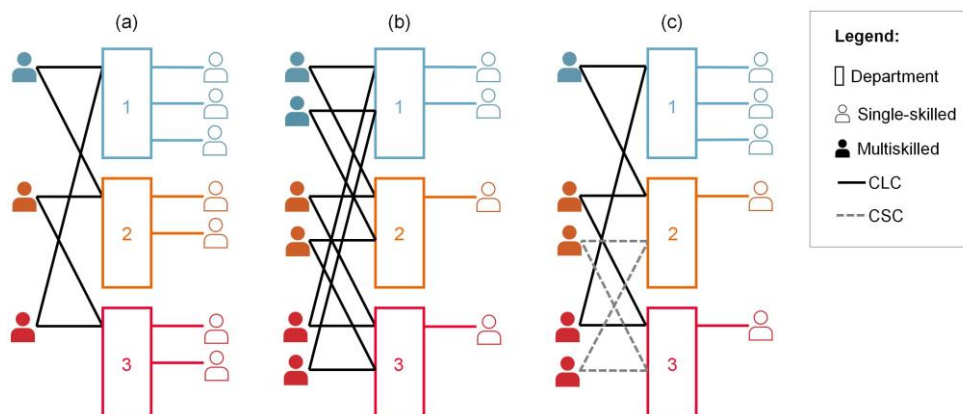
Figure 1. Some examples of 2-chaining structures: (a) One closed long chain; (b) Two similar closed long chains; (c) One closed long chain and one closed short chain.

For the same retail case study and assuming uncertainty in the staff demand (at the store level), Henao et al. [1] and Henao et al. [2] solved a weekly multiskilled personnel assignment problem (MPAP). They calculated the number of single-skilled workers who should become multiskilled and in which additional department they should be trained. In addition, both studies designed training plans through a 2-chaining strategy to minimize the total cost of under/overstaffing and training. In the first study, Henao et al. [1] developed a single-stage robust optimization approach to calculate the optimal number of multiskilled workers per store's department, whereas in the second study, Henao et al. [2] derived a closed-form equation to deliver an estimation of the optimal quantity of multiskilled workers per store's department. Even with different solution approaches, both studies showed that the most profitable staff is one where there is a mix of single-skilled and multiskilled workers. Their results also showed that the best 2-chaining structures combine CLCs and CSCs of different lengths.

The literature on personnel scheduling problems shows several approaches for optimizing under uncertainty that have received substantial theoretical development, among these are: (1) Closed-form equation (CF); under this approach, it is possible to develop an analytic expression to calculate the optimal solution (or an approximation of the optimal solution) for the addressed problem. This approach requires knowing the probability distribution function associated with the parameter under uncertainty[1]. (2) Two-stage stochastic optimization (TSSO); this approach assumes complete knowledge of the distribution of probability associated with the parameter under uncertainty and includes (in an explicit way) such uncertainty into the mathematical model [22]. (3) Robust optimization (RO); unlike the previous approaches, complete knowledge of the stochastic parameters is not required [23,24]. Stochastic parameters are represented by the so-called uncertainty set, i.e., a set where parameters can take any value. RO limits the variability in the parameter under uncertainty to an acceptable and predefined range to avoid unreasonable scenarios. (4) Distributionally robust optimization (DRO); this approach is known as a generalization of the RO and TSSO approaches. DRO tackles the uncertainty by including additional information into the uncertainty set (e.g., mean and covariance), which can be estimated from expert knowledge or historical data [25]. In short, CF and TSSO assume complete distributional information, DRO assumes partial distributional information, whereas RO assumes minimal distributional knowledge about the uncertain parameters.

---

[1] In the personnel scheduling literature, we have not found a closed-form equation with partial knowledge of the stochastic parameter. On the contrary, all the studies assumed complete knowledge. Thus, without loss of generality, the closed-form approach is considered to have access to full distributional knowledge of stochastic parameters in this paper.

Although all of these approaches guarantee robust solutions, they cannot ensure the same level of reliability and conservatism in the solutions. Note that, in a retail store context, a robust solution is not necessarily the best in every possible scenario of demand variability, but it remains good over a wide range of changes in staffing demand.

The current work builds on top of our previous findings from Henao et al. [1] and Henao et al. [2]. We also propose to solve the MPAP with 2-chaining in this paper. However, unlike other references in the state-of-the-art and our previous works, we formulate a TSSO model to calculate the multiskilling requirements that minimize the training costs and the expected costs of staff surplus/shortage. Particularly, this work presents two main contributions. First, we solve the MPAP using a TSSO approach. Second, we evaluate and compare the performance of three different optimization approaches under uncertainty (i.e., TSSO vs. RO vs. CF). Note that, the benchmark RO and CF approaches were already published in Henao et al. [1] and Henao et al. [2], respectively. Besides, these three optimization approaches under uncertainty will be compared with the deterministic (DT) approach and with two multiskilling approaches considered as myopic: zero multiskilling (ZM) and total multiskilling (TM). Such that, under a ZM approach, each worker is skilled to perform in a single store department (i.e., single-skilled), whereas under a TM approach, all workers are skilled to perform in all departments. The methodology is tested for a case study associated with a retail store operating in Chile, which has already been presented and solved by Henao et al. [1,2]. The performance of all approaches is assessed by means of out-of-sample simulations. Each approach is tested against different levels of variability in staff demand per store's department. We also provide insights on how to choose the most suitable approach according to the balance between conservatism and reliability associated with each of them.

2. Literature review

Personnel scheduling is a research area extensively studied. In this topic, literature review articles such as Ernst et al. [26], Brucker et al. [27], Van den Bergh et al. [28], and Özder et al. [29] addressed various problems and solution methodologies associated with multiple industries. Table 1 presents previous articles that have used optimization approaches under uncertainty to solve personnel scheduling problems in the service industries. The principal features of these articles are described below. We remark that each of the articles classified in Table 1 considered demand uncertainty, but none of them included supply uncertainty (i.e., unscheduled staff absenteeism).

1. *Type of personnel scheduling problem* (PSP) *studied*: May be classified as: (a) *Staffing* (S), that is, deciding the number of workers needed per task type and per labor contract type. (b) *Shift scheduling* (SS), that is, deciding what shift each worker should perform on their working days. (c) *Days-off scheduling* (DOS), that is, deciding the sequence of working and resting days that should be assigned to each worker (e.g., weekly, biweekly). (d) *Tour scheduling* (TS), where shifts and days-off are scheduled simultaneously (e.g., weekly, biweekly). (e) *Assignment* (A), where workers are assigned to particular task types, but it is not required to schedule shifts or rest days (e.g., daily, weekly, monthly). Note that, the last type of problem is known as the personnel assignment problem (PAP), and when it involves the use of multiskilled staff, it is known as the MPAP.

2. *Multiskilling* (MS): Specifies whether or not the PSP incorporates multiskilling as a source of labor flexibility. If the study includes multiskilling, we identify two alternatives. That is, the skill(s) of a worker can be incorporated into the math model as a parameter (par) or as a variable (var).

3. *2-chaining* (2-CH): Specifies when a study incorporated the multiskilling using the 2-chaining strategy.

4. *Solution method* (SM) *used*: May be: (a) *linear programming* (LP); (b) *integer programming* (IP); (c) *mixed integer programming* (MIP); (d) *heuristics* (H); (e) *metaheuristics* (M); (f) *benders decomposition* (BD); (g) *branch-and-cut* (B&C); (h) *column generation* (CG); or (i) *analytic* (AN).

5. *Approach to model uncertainty* (AMU): Specifies the optimization approach under uncertainty employed to handle the uncertain demand in the study. It can be classified as: (a) *two-stage stochastic optimization* (TSSO); (b) *distributionally robust optimization* (DRO); (c) *single-stage robust optimization* (1RO); (d) *two-stage robust optimization* (2RO); or (e) *closed-form equation* (CF).

6. *Application*: Service industry where the study has been applied.

Table 1. Principal features for the related literature on optimization methods under uncertainty for solving the PSP in the service sector; including this paper.

Reference	PSP	MS	2-CH	SM	AMU	Application
Bard et al. [30]	S + TS	No	–	IP + H	TSSO	Postal service
Zhu & Sherali [31]	S + SS	Var	No	MIP + BD	TSSO	Service centers
Robbins & Harrison [32]	S + TS	No	–	MIP + BD	TSSO	Call Centers
Campbell [33]	DOS	Par	No	IP + H	TSSO	Health
Liao et al. [34]	S + A	No	–	MIP	TSSO, 1RO	Call Centers
Liao et al. [35]	S + SS	No	–	MIP	TSSO, DRO	Call Centers
Gnanlet & Gilland [36]	S + A	Var	Yes	MIP + AN	TSSO, CF	Health
Paul & MacDonald [37]	S + A	Var	Yes	MIP + H + AN	CF	Health
Kim & Mehrotra [38]	S + TS	No	–	MIP + M	TSSO	Health
Parisio & Jones [39]	TS	No	–	MIP	TSSO	Retail
Henao et al. [1]	A	Var	Yes	MIP + H	1RO	Retail
Bodur & Luedtke [40]	S + SS	Par	Yes	MIP + BD + B&C	TSSO	Call Centers
Liu [18]	S + A	Par	Yes	IP + AN	TSSO, CF	–
Mattia et al. [41]	S + SS	No	–	IP + BD + B&C	2RO	Call Centers
Restrepo et al. [42]	TS	Par	No	MIP + M	TSSO	–
Altner et al. [43]	S + SS	No	–	IP + CG	TSSO	CSOC
Altner et al. [44]	DOS	Var	No	MIP + M	TSSO	CSOC
Henao et al. [2]	A	Var	Yes	AN + H + LP	CF	Retail
Fontalvo Echavez et al. [20]	DOS	Var	Yes	MIP	TSSO	Retail
This paper	*A*	*Var*	*Yes*	*MIP + H + AN*	*TSSO, RO, CF*	*Retail*

Below, we present a more detailed review of various articles listed in Table 1. We begin with those that did not incorporate multiskilling as a source of flexibility and then consider those that did. Then, we discuss those articles that solved the addressed problem using two or more optimization approaches under uncertainty.

2.1. Personnel scheduling problems without multiskilling

Table 1 shows that various articles did not use multiskilled staff to minimize overstaffing/understaffing levels (i.e., [30,32,34,35,38,39,41,43]). This is common in studies where

the personnel scheduling problem only considers a single type of task (or store department). In Bard et al. [30], a TSSO model is developed to solve a shift scheduling problem in a post office. Decisions on the number of required workers were modeled in the first stage, and weekly shift structures were modeled in the second stage. In Kim and Mehrotra [38], the authors proposed a TSSO model to determine the staff size and the shift structures in the first stage, whereas these shift structures were adjusted once new information about incoming demand was available in the second stage. Similarly, Parisio and Jones [39] proposed a TSSO model to solve a tour scheduling problem in retail. Optimal decisions on workers' schedules (i.e., rest days and shifts) were modeled in the first stage, while overtime decisions after demand realization were modeled in the second stage.

2.2. Personnel scheduling problems with multiskilling

Table 1 indicates that various articles included multiskilling as a source of flexibility. Thus, such articles addressed a personnel scheduling problem with at least two types of tasks (or store departments). The literature review shows that multiskilling can be incorporated into mathematical modeling in three ways.

First, by considering multiple multiskilling structures as an input parameter in the math model and then the model solution must identify the best performing structure (i.e., [18,33,40,42]). In Campbell [33], a TSSO model is developed to solve a days-off scheduling problem with multiskilled workers and a multi-department service environment. Decisions on rest days were modeled in the first stage, whereas reassigns available workers according to the observed demands were modeled in the second stage. In Restrepo et al. [42], the authors addressed a tour scheduling problem considering workers with identical skill sets. Decisions on rest days and shifts were represented in the first stage, while break times and task assignments were described in the second stage.

Second, by modeling the multiskilling as a decision variable into the math model (i.e., [31,44]). In Zhu and Sherali [31], the authors proposed a TSSO model to solve a staffing and shift scheduling problem. Decisions on staffing levels and shifts were modeled in the first stage, while in the second stage, worker reassignment decisions were taken. In Altner et al. [44], the authors studied a TSSO model to solve a days-off scheduling problem. The first stage modeled decisions on which workers should be multiskilled and which days each worker should rest considering a bi-weekly planning horizon. Then, for each day, the second stage considered how task types could be assigned to workers once the uncertain demand is revealed.

Third, by modeling the multiskilling as a decision variable into the math model and ensuring the formation of closed-chained multiskilling structures by including of a set of constraints into the model (i.e., [1,2,20,36,37]). Particularly in the retail industry, Henao et al. [1] and Henao et al. [2] solved a weekly MPAP with 2-chaining using different optimization approaches under uncertainty; RO and CF, respectively. Note that, to our knowledge, these are the only articles that solved the MPAP with 2-chaining (see Table 1). The MPAP allows adding and simplifying scheduling decisions while focusing on answering the multiskilling tactical questions. In both articles, the authors proposed two methodological stages to sequentially answer the following three tactical questions: (i) How many multiskilled workers are needed per department? (ii) In which second department should be trained each of the workers identified as multiskilled? (iii) How many CLCs and/or CSCs are needed, and which store departments should belong to each of the chains? In Henao et al. [1], the first stage of the methodology proposed a single-stage robust optimization model to answer the question (i), while the

second stage proposed a constructive heuristic to answer the questions (ii) and (iii). In Henao et al. [2], the first stage of the methodology derived a closed-form equation to answer the question (i), while the second stage used the constructive heuristic developed by Henao et al. [1] to answer the questions (ii) and (iii). However, these sequential methodologies could lead to suboptimal solutions; thus, integrated approaches are needed to guarantee optimality.

2.3. Personnel scheduling problems solved with various optimization approaches under uncertainty

Table 1 shows various articles that solved a personnel scheduling problem using two optimization approaches under uncertainty (i.e., [18,34,35,36]). On one side, Liao et al. [34] and Liao et al. [35] did not incorporate multiskilling in the mathematical modeling. On the other hand, Liu [18] incorporated multiskilling in the mathematical modeling as a parameter, while Gnanlet and Gilland [36] incorporated multiskilling as a variable, but they also used 2-chaining.

In Liao et al. [34], the authors addressed a call center staffing problem considering single-length shifts. They compared the performance of a stochastic programming approach and a robust programming formulation. Regarding the average cost, they explained that the TSSO approach was more efficient than the RO approach. This is because the RO approach is based on a worst-case-type analysis. Thus, the optimal RO solution generates higher staffing levels, which decreases the staff shortage costs but also increases the direct salary costs. In Liao et al. [35], for a case study in call centers, the authors proposed a methodology that combined the TSSO and DRO approaches to solve simultaneously a staffing and shift scheduling problem considering service level constraints. Liu [18] and Gnanlet and Gilland [36] proposed TSSO models to solve a staffing problem with only two task types for the health sector. In Liu [18], the authors developed an analytic expression to calculate the optimal number of workers. In Gnanlet and Gilland [36], the authors also developed analytic expressions to determine the optimal number of multiskilled nurses. Note that none of these four papers solved the MPAP, and they did not address the retail industry as a case study.

2.4. Main contributions

Our study seeks to fill two gaps evidenced in the literature review. First, for the MPAP with 2-chaining in the retail industry, we develop a TSSO model to simultaneously answer the tactical questions (i), (ii), and (iii) described previously, instead of solving them sequentially, as proposed by Henao et al. [1] and Henao et al. [2]. It is important to note that, Table 1 indicates that most of the authors used a TSSO approach to solve the PSP addressed, but none of them studied the MPAP with 2-chaining.

Second, we propose to evaluate and compare the performance of three different optimization approaches under uncertainty (i.e., TSSO vs. RO vs. CF). Thus, using an extensive computational study, we will quantify the reliability of the TSSO model by comparing it with the benchmark RO and CF approaches presented in Henao et al. [1] and Henao et al. [2], respectively. However, we will not make comparisons with additional benchmarks from the literature, because except for Henao et al. [1] and Henao et al. [2], the other articles in Table 1 studied different PSPs. In addition to these three optimization approaches, we propose to compare them with the deterministic (DT) approach and with the two myopic approaches ZM and TM.

3. Description and assumptions of the MPAP

Our problem considers a retail store with a known number of departments, such that each of them has contracted a group of workers initially single-skilled. Working in a store department requires specific skills, which differ among all departments. Although the working hours of workers are determined in their contracts, the store's departments face uncertainty in demand. Here, the skill training of workers plays a strategic role in handling the uncertainty in demands. That is, the costs associated with staff shortages and surpluses in the different store departments may be prevented by assigning some multiskilled workers from departments with staff surpluses to departments with staff shortages. Therefore, the goal of the problem is to decide how many workers should be multiskilled and in what additional department, such that they get ready to work in other departments in case needed.

Table 2. Notation.

Sets	
I	Workers, indexed by i
L	Departments of the retail store, indexed by l
S	Scenarios of demand, indexed by s
I_l	Set of single-skilled workers initially hired in department l, indexed by i
Parameters	
$r_l(s)$	Random parameter of staff demand in weekly hours per store's department l, $\forall l \in L, s \in S$
$\bar{r}_l$	Average value of staff demand in weekly hours per store's department l, $\forall l \in L$
c	Cost of workers training in any department; [US\$-week /Worker]
u	Staff shortage cost per store's department; [US\$/Hour]
b	Staff surplus cost per store's department; [US\$/Hour]
h	Weekly hours that workers have to work in accordance with their employment contract
m_i	Department in which the worker i is initially trained, $\forall i \in I$
Variables	
x_{il}	1 if worker i is trained in department l, 0 otherwise, $\forall i \in I, l \in L$
v_i	1 if worker i is multiskilled, 0 otherwise, $\forall i \in I$
λ_l	Number of multiskilled workers belonging to department l, $\forall l \in L$
Ω	Number of generated closed-chains
ω_{il}	Number of weekly working hours allocated to worker i in department l, $\forall i \in I, l \in L$
κ_l	Number of staff shortage hours per week in department l, $\forall l \in L$
δ_l	Number of staff surplus hours per week in department l, $\forall l \in L$
Solution approaches for the MPAP with 2-chaining	
DT	Deterministic optimization
RO	Robust optimization
CF	Closed-form equation
$TSSO$	Two-stage stochastic optimization

In short, in this paper, we will solve a weekly MPAP that at the same time determines how many multiskilled workers are needed, in what second department, and how to assign working hours to each worker over a one-week planning horizon. In addition, our formulation will provide a training plan that ensures the formation of CLCs and/or CSCs, which have shown to be especially useful in minimizing overstaffing and understaffing under uncertain demand. Thus, the solution of the MPAP

with 2-chaining, at the store level, must minimize the training costs and the expected costs of staff surplus/shortage. Table 2 presents the mathematical notation employed to formulate the MPAP with 2-chaining. It will also be possible to find additional notation at each subsection of the proposed methodology in Section 4.

Similar to Henao et al. [1] and Henao et al. [2], the formulation of the MPAP with 2-chaining considers various assumptions: (1) The demand in weekly hours per store's department, $r_l(s)$, is a random parameter. In this paper, it is assumed two different levels of knowledge about the uncertainty of the parameter $r_l(s)$: (i) for the TSSO and CF approaches, the probability distribution of the random parameter is completely known; (ii) for the RO approach, only the average value and the maximum value allowed of the random parameter are known. (2) There is no correlation between the demand of different pairs of departments. (3) Staffing demand per store's department may not be fully covered (i.e., unsatisfied demand). Thus, the objective function incorporates a staff shortage cost, and it is assumed that this cost does not differ between departments. (4) Staff surplus also has an associated cost, and it is also assumed that this cost does not vary between departments. Mattia et al. [41] and Henao et al. [1] indicated that this cost represents an opportunity cost, which is incurred when there are idle workers in the store. (5) The objective function also incorporates staff training costs, and they do not vary by department. (6) All store workers have the same type of contract and thus they must perform an equal number of weekly working hours. (7) It is assumed that, for each department, the available assignment weekly hours, $(|I_l|h)$, are equal to the average value of weekly hours demand $(\bar{r}_l)$. Notice that this assumption imposes a trivial solution to the deterministic formulation for the MPAP. That is, since there is no imbalance between staffing and staff demand in each store department, the solution of the deterministic MPAP formulation will not require any multiskilled worker. That is, the DT approach solution will be equivalent to the ZM approach solution (i.e., zero multiskilling). (8) All departments have contracted a group of workers initially single-skilled. (9) Workers can be trained in a single additional department and, thus, they could perform up to a total of two store departments. (10) Workers have equal productivity in those departments where they are skilled. Henao et al. [1,2,9] and Porto et al. [6,21], who also addressed multiskilled personnel scheduling problems for retail stores, explain that this last assumption is reasonable when multiskilled workers can only perform up to a maximum of two store departments, and the types of tasks performed have a low and uniform level of complexity.

4. Methodology

This section proposes a framework to solve the MPAP with 2-chaining and uncertain demand, as described above. First, in Subsection 4.1, we develop a TSSO model to calculate the multiskilling requirements that minimize the training costs and the expected costs of staff surplus/shortage. Then, in Subsection 4.2, we present the two alternative optimization approaches under uncertainty - RO and CF - to solve the addressed problem. These two alternative approaches were already published in Henao et al. [1] and Henao et al. [2], respectively. Recall that the main contributions of this work are: (1) to solve the MPAP with 2-chaining using a TSSO approach and (2) to evaluate and compare the performance of three different optimization approaches under uncertainty (i.e., TSSO vs. RO vs. CF).

4.1. Stochastic MPAP formulation

In this subsection, we present a brief background on the TSSO approach, and then we present the TSSO version of the deterministic MPAP model (DT) developed by Henao et al. [1]. In Subsection 4.1.1, we make abuse of notation for the sake of consistency with the standard nomenclature used on the Stochastic Optimization literature. All symbols are described within the context.

4.1.1. Background on the TSSO framework

In a TSSO approach, decisions are modeled in two stages. The first-stage decisions are often denoted by the set of variables, x, and are called the "*here-and-now*" decisions. These decisions are made before the uncertain parameters become known. The second-stage decisions or recourse decisions are often described by the vector of variables, $y(x, \xi)$, and are defined as the "*wait and see*" decisions. These decisions are made once the uncertainty is exposed. Uncertain parameters are represented by ξ, and are assumed to be known. Thus, the second-stage decisions depend not only on first-stage decisions but also on the knowledge of the parameters under uncertainty, that is, $y(x, \xi)$. In addition, in the TSSO framework, non-anticipativity constraints are implicitly modeled to ensure that the first-stage decisions are not affected by the knowledge of the uncertain parameters. For details about the theory of TSSO, the reader is referred to Birge and Louveaux [22].

A general form for the classical TSSO model is expressed in Equation (1), distinguishing between first-stage and second-stage decisions.

$$\underset{x,y(x,\xi)}{Min} \quad c^T x + \mathrm{E}_p\left[f\left(y(x,\xi)\right)\right]$$

$$s.t. \tag{1}$$

$$x \in X$$

$$y(x,\xi) \in Y(x,\xi),$$

Note that the first term in the objective function, $c^T x$, corresponds cost incurred by the first-stage decisions. The second-stage term is the recourse expected cost function, $E_p[f(y(x,\xi))]$, which is represented as the expectation of the second-stage objective taken over the stochastic parameter ξ within a given probability measure p. Such that $f(y(x, \xi))$ is the operational value function.

The uncertain parameters, ξ, are usually continuous; thus, the TSSO framework is in most cases intractable. Sample Average Approximation (SAA) approach is commonly used to overcome this issue. The SAA is based on the generation of a set of random samples from the true stochastic parameter ξ that would represent them. Birge and Louveaux [22] explain that, for each possible scenario, the SAA clones the variables set of second-stage. Thus, by taking independent and identically distributed (iid) random samples, ξ, from the assumed probability distribution, it is possible to approximate the problem (1). By doing so, we represent each generated sample by $\xi(s) \in \varXi_s$, where $\varXi_s$ is the set of all iid random samples generated and $|\varXi_s|$ represents the cardinality of $\varXi_s$, that is, number of scenarios. Such that, when $|\varXi_s|$ is large enough (but a finite number), the SAA approach has a high quality. However, for a large set of scenarios in the second-stage, the SAA can be too expensive to solve. Below, for a finite set of scenarios, the SAA formulation of (1) is defined.

$$\underset{x,y(x,\xi(s))}{Min} \quad c^T x + \frac{1}{|\Xi_s|} \sum_{\xi(s)\in\Xi_s} f\left(y(x,\xi(s))\right)$$

$s.t.$ (2)

$$x \in X$$

$$y(x,\xi(s)) \in Y(x,\xi(s)), \quad \xi(s) \in \Xi_s$$

Then, (2) is a finite-dimensional problem equivalent to (1).

4.1.2. TSSO model for the MPAP

For the MPAP problem, we assume complete knowledge of the distribution of probability associated with the uncertain demand, which is accessible from historical data. The deterministic linear MIP model presented by Henao et al. [1] can be extended to add such uncertainty. For full details on the DT model, the reader is referred to Subsection 3.1 in Henao et al. [1].

In this subsection, we proposed a TSSO model for the MPAP with 2-chaining. To facilitate comparisons of the TSSO version with the DT version developed by Henao et al. [1], we use the same mathematical notation. The first-stage variables, x_{il}, v_i, and λ_l are tactical decisions associated with the workers training to work in one additional store department. These decisions are taken before the uncertainty in weekly demands has been revealed. The second-stage continuous recourse variables, $\omega_{il}(s)$, $\kappa_l(s)$, $\delta_l(s)$, are operational decisions associated with the workers allocation to minimize staff surplus/shortage levels. These decisions are taken after the uncertainty in weekly demands has been revealed. To solve the TSSO version for the MPAP, we implemented an SAA approach.

Below, the mathematical formulation of our TSSO model is presented. Also recall that, in Table 2 was defined the math notation for each of the sets, parameters, and variables used in the TSSO model.

$$Min \quad \underbrace{\sum_{i\in I}\sum_{l\in L:l\neq m_i} cx_{il}}_{first-stage} + \frac{1}{|S|}\underbrace{\left[\sum_{s\in S}\sum_{l\in L} u\kappa_l(s) + \sum_{s\in S}\sum_{l\in L} b\delta_l(s)\right]}_{second-stage} \tag{3}$$

$s.t.$

$$\sum_{i\in I} \omega_{il}(s) + \kappa_l(s) - \delta_l(s) = r_l(s) \quad \forall l \in L, s \in S \tag{4}$$

$$\sum_{l\in L} \omega_{il}(s) = h \quad \forall i \in I, s \in S \tag{5}$$

$$\omega_{il}(s) \leq hx_{il} \quad \forall i \in I, l \in L, s \in S \tag{6}$$

$$x_{il} = 1 \quad \forall i \in I, l \in L: l = m_i \tag{7}$$

$$v_i = \sum_{l\in L:l\neq m_i} x_{il} \quad \forall i \in I \tag{8}$$

$$\lambda_l = \sum_{l\in I_l} v_i \quad \forall l \in L \tag{9}$$

$$v_i \leq 1 \quad \forall i \in I \tag{10}$$

$$\sum_{i\in I_l} v_i = \sum_{i\in\{I-I_l\}} x_{il} \quad \forall l \in L \tag{11}$$

$$x_{il} \in \{0,1\} \quad \forall i \in I, l \in L \tag{12}$$

$$v_i \geq 0 \qquad\qquad \forall i \in I \qquad\qquad (13)$$

$$\lambda_l \geq 0 \qquad\qquad \forall l \in L \qquad\qquad (14)$$

$$\omega_{il}(s) \geq 0 \qquad\qquad \forall i \in I, l \in L, s \in S \qquad\qquad (15)$$

$$\kappa_l(s) \geq 0 \qquad\qquad \forall l \in L, s \in S \qquad\qquad (16)$$

$$\delta_l(s) \geq 0 \qquad\qquad \forall l \in L, s \in S \qquad\qquad (17)$$

The objective function (3) minimizes the total weekly cost. On the one hand, and associated with the first-stage decisions, the first term in the objective function minimizes the total training cost. On the other hand, and associated with the second-stage decisions, the next two terms in the objective function minimize the total expected cost of staff shortages and surpluses. In the TSSO version for the MPAP, the decisions about how many workers should be multiskilled in the first-stage will impact the staff surplus/shortage levels in the second-stage. Besides, the second-stage decisions are recourse variables that seek to take corrective actions in the workers' assignment after the uncertainty is revealed, but such corrective actions have a cost. Thus, a TSSO approach seeks to select the variables of first-stage that minimizes both first-stage costs and expected second-stage random costs.

Constraints (4) account for demand fulfillment for each department and demand scenario. They also calculate the staff surplus/shortages levels produced by the mismatch between the staffing and staff demand. Constraints (5) guarantee that, for each demand scenario, all workers are assigned to work the working hours stipulated in their contracts. Constraints (6) ensure that, for each demand scenario, workers can only be assigned to work in those departments where they received training. Constraints (7) indicate the department where each worker was originally trained. Constraints (8) indicate whether a worker was trained to be multiskilled, or he/she continues as a single-skilled. Constraints (9) calculate the optimal number of multiskilled workers belonging to each department. Constraints (10) ensure that workers can be trained in a single additional department. Constraints (11) ensure that the sum of multiskilled workers originally trained in the department l and who were also trained to perform in another second department ($\sum_{i \in I_l} v_i$), is equal to the sum of multiskilled workers originally trained in a department different to l but that were also trained to perform in l as additional department ($\sum_{i \in \{I - I_l\}} x_{il}$). Constraints (10) and (11) guarantee a 2-chaining policy, and thus, the generation of CLCs or CSCs that allow the transfer of multiskilled workers among store departments. Note that, Henao et al. [1] explained that for both CLCs and CSCs formation, it is required that the outdegree of every one demand nodes belonging to a chain must be equal to the indegree of its corresponding supply node. The constraints (11) guarantee the satisfaction of such requirement. Constraints (12)-(17) indicate the possible values that each decision variable can take. Note that, given the auxiliary variables v_i must take values less than or equal to one (see constraints (10)) and they are calculated through the sum of the binary variables x_{il} (see constraints (8)), such variables do not require to be declared as binary. Finally, note that constraints (4)-(6) and (15)-(17) were accordingly extended from Henao et al. [1] to model that the second-stage variables depend on the random demand scenarios.

The optimal solution of the TSSO simultaneously answer the following three tactical questions: (i) How many multiskilled workers are needed per department, λ_l? (ii) In which second department should be trained each of the workers identified as multiskilled, x_{il}?, and (iii) How many CLCs and/or CSCs are needed, and which store departments should belong to each of the chains, Ω?

4.2. Alternative optimization approaches for solving the MPAP under uncertainty

In this subsection, we present the alternative optimization approaches, RO and CF, presented in Henao et al. [1] and Henao et al. [2] respectively, for solving the MPAP with 2-chaining and uncertain demand. Later, in Section 6, the solution of the proposed TSSO model will be compared with the benchmark approaches RO and CF.

4.2.1. Robust MPAP formulation

Robust optimization is widely studied in the literature. In Beyer and Sendhoff [45], Kouvelis and Yu [46], and Gabrel et al. [47], the authors exhaustively reviewed and discussed articles that present methodological advances in RO as well as a range of problems that have used RO. In this subsection, we present the single-stage RO approach developed by Henao et al. [1] to solve the MPAP with 2-chaining and uncertain demand. The classic first-stage RO is an alternative approach to TSSO, where the probability distribution associated with the random parameter is not known; instead, the uncertainty of the parameter is modeled using the so-called uncertainty sets that usually are obtained by using historical data or estimates. In the RO approach from Henao et al. [1], it is assumed that the uncertain demand of each store department has a known average value, $\bar{r}_l$, and the maximum value allowed for the uncertain demand is equal to $\hat{r}_l = \alpha \bar{r}_l$. Note that, this maximum allowed variation is defined as a fraction of the average value, where α is equivalent to the coefficient of variation (CV) for the demand in a department, $0 \leq \alpha \leq 1$. Thus, the weekly hours demand $\tilde{r}_l$ is a random variable that can only take values within the interval $[\bar{r}_l - \hat{r}_l, \bar{r}_l + \hat{r}_l]$.

Henao et al. [1] also included into the RO model the set J, which refers to the set of departments whose staff demands can vary simultaneously. They assumed that $|J| = |L|$, so that, the staff demand in each department is subject to uncertainty. In order to address issues of overconservative solutions in the RO framework, it was possible to define an uncertainty budget parameter or level of conservatism, Γ, which was stated as a fraction, $0 \leq \Gamma \leq 1$, of all departments with uncertain demand, $|J|$. If $\Gamma = 0$, RO model is equivalent to DT model, since uncertainty is not considered and demand values are the nominal values of the uncertainty set. If $\Gamma = 1$, solutions are optimized accordingly for the worst-case scenario, which considers the maximum allowed uncertainty.

Considering the previous assumptions and the approaches developed by Bertsimas and Sim [24], Bertsimas and Thiele [48] and Bohle et al. [49]; Henao et al. [1] presented a RO version of their DT formulation. The RO model seeks to minimize the worst-case operational costs of staff shortages and surpluses for the training decisions, x_{il}. For full details on the RO version, the reader is referred to Subsection 3.2.2 in Henao et al. [1].

Finally, we found two key methodological differences between the TSSO approach and the RO approach. First, while the TSSO approach solutions are driven by the best-expected cost, solutions of the RO approach are driven by the best performance in the worst-case scenario. Second, the TSSO approach simultaneously answers the tactical questions (i), (ii), and (iii) established at the end of Subsection 4.1.2, which ensures consistency in the decision approach. On the contrary, the single-stage RO approach answers these questions sequentially through two methodological stages. The first stage of the methodology proposed a RO model to answer the question (i). The second stage proposed a constructive heuristic to answer the questions (ii) and (iii). For full details on the constructive heuristic, the reader is referred to Subsection 3.3 in Henao et al. [1]. In short, the first stage determines the

optimal number of multiskilled workers needed per department (λ_l). Then, the second stage considering the solution provided by the first stage decides in which second department should be trained each of the workers identified as multiskilled (x_{il}), whereas the minimum number of closed chains is built (Ω).

4.2.2. Closed-form MPAP formulation

In this subsection, we present the approach developed by Henao et al. [2] for solving the MPAP with 2-chaining and uncertain demand. Similar to Henao et al. [1], Henao et al. [2] answered the tactical questions (i), (ii), and (iii) sequentially through two methodological stages. The first stage of the methodology proposed a CF equation to estimate: (i) the number of multiskilled workers needed per department (λ_l). The second stage used the constructive heuristic developed by Henao et al. [1] to determine: (ii) in which second department should be trained each of the multiskilled workers (x_{il}); and (iii) the number of generated closed-chains (Ω). Note that, the second stage considers the solution provided by the first stage.

The CF expression developed by Henao et al. [2] was motivated by various problems in the literature where the authors seek to find an approximate solution using local optimality criteria. In Robuste et al. [50], Del Castillo [51], and Daganzo [52], the authors present a range of complex logistics problems (e.g., distribution centers location and vehicle routing) that have used a closed-form equation. In particular, in Henao et al. [2], the authors derived a novel CF equation that examines every single department in separate to deliver an estimation of the optimal quantity of multiskilled workers. Note that, under a CF approach, it is assumed that the local conditions are considered to be explicable for the complete region under evaluation.

CF is an alternative framework to TSSO and RO, where is assumed complete knowledge of the distribution of probability associated with the uncertain demand. The CF approach considered the following two key assumptions: (1) the probability density function for the staff demand per department is independently and normally distributed with known average value $\bar{r}_l$ and standard deviation σ_l; (2) the staff hours supply per store's department are equal to the average value of staff hours demand ($|I_l|h = \bar{r}_l$). Note that the second assumption was also considered in the TSSO and RO approaches.

Under these assumptions, Henao et al. [2] developed a closed-form equation to estimate the number of multiskilled workers needed per department (λ_l). The CF equation seeks to minimize the incurred training cost (but per-hour instead of per-week) and the expected staff shortage cost. For further details about the development of the CF equation, the reader is referred to Subsection 4.2.1 in Henao et al. [2].

We also found two key methodological differences between the TSSO approach and the CF approach. First, the TSSO approach solutions are driven by the best-expected cost. In contrast, the solutions of the CF approach are driven by the best performance in the average-case scenario. Second, the TSSO approach simultaneously answers the tactical questions (i), (ii), and (iii), whereas the CF approach answers these questions sequentially. Finally, unlike TSSO and RO approaches that propose sophisticated optimization models for assessing the benefits of the 2-chaining structures, the CF approach offers decision-makers with a valuable and practical formula to estimate cost-effectiveness levels of multiskilling.

5. Case study

We propose a case study to validate the methodology previously presented. Such case study is associated with a home improvement store in Santiago de Chile, only for its sales personnel. In Subsection 5.1, we present the performance metrics used in Section 6 to assess and compare the 2-chaining structures constructed by the optimization approaches under uncertainty (i.e., TSSO, RO, and CF). Besides, in Section 6, these three optimization approaches under uncertainty will be compared with the two myopic approaches ZM and TM. Recall that the ZM approach represents the solution of the DT formulation. Also, note that the TM approach will achieve the lower bound of the total weekly staff surplus/shortage cost, whereas the ZM approach will obtain the upper bound. Later, we explain the data requirements of the MPAP model in Subsection 5.2.

5.1. Metrics of performance

To evaluate the performance of the solution approaches presented (i.e., DT, TSSO, RO, CF, ZM, and TM), we use metrics to measure: the required multiskilling levels ($\%ME, \%TM$), the incurred costs (φ, ψ), and the reliability ($\%SSS$) associated with the solutions delivered by each approach. Below, we describe each of these metrics in more detail.

1) *Multiskilled workers in percentage.* It is calculated as the ratio between the total number of multiskilled workers and the total number of hired workers.

$$\%ME = \frac{\sum_{i \in L} \lambda_l}{|I|} . 100$$

2) *Total multiskilling in percentage.* It is calculated as the ratio between the number of additional departments in which workers were trained and the maximum number feasible.

$$\%TM = \frac{\sum_{i \in I} \sum_{l \in L : l \neq m_i} x_{il}}{|I|(|L| - 1)} . 100$$

3) *Total weekly cost.* It is equivalent to the sum of incurred costs of training and staff surplus/shortage.

$$\varphi = \sum_{i \in I} \sum_{l \in L : l \neq m_i} c x_{il} + \sum_{l \in L} u k_l + \sum_{l \in L} b \delta_l$$

4) *Total weekly staff* surplus/shortage *cost.* It is defined as the incurred cost of staff surplus/shortage associated with any approach.

$$\psi = \sum_{l \in L} u k_l + \sum_{l \in L} b \delta_l$$

5) *Savings on the total weekly cost of staff surplus/shortage for an approach.* Consider that, ψ_{ZM} indicates the incurred cost of staff surplus/shortage when all workers are single-skilled, whereas ψ_{TM} indicates the incurred cost when each worker is trained to work in all departments. Then, the difference between ψ_{ZM} and ψ_{TM} represents the maximum possible savings in this cost when there is total multiskilling.

$$\%SSS = \frac{\psi_{ZM} - \psi}{\psi_{ZM} - \psi_{TM}} . 100$$

5.2. Data requirements of the MPAP

This subsection describes the data used in this study. To fairly compare the three optimization approaches under uncertainty (i.e., TSSO, RO, CF), we used the same case study that was already presented and solved by Henao et al. [1] and Henao et al. [2]. The database includes real and simulated data that were derived from a retail store operating in Chile. The real dataset indicates the store size, the number of single-skilled workers initially hired in each store department, working hours established in the employment contract, average weekly staff hours demand per department, and the cost parameters values. The simulated datasets describe the random weekly staff hours demand per department. Below we present a description of both dataset types. For more details of the database, the reader is referred to the data article presented by Henao et al. [53].

5.2.1.   Real dataset

The case study evaluates a store with $|I| = 30$ and $|L| = 6$. Such that each worker has a full-time contract according to Chilean labor legislation, that is, $h = 45$. Let $\bar{r}_l$ be the average value of demand in weekly hours per store's department l, $\forall l \in L$. Such that, $\bar{r}_1 = 315$ (i.e., 7 workers), $\bar{r}_2 = 225$, $\bar{r}_3 = 135$, $\bar{r}_4 = 135$, $\bar{r}_5 = 180$, and $\bar{r}_6 = 360$. Regarding the cost parameters, the training cost per worker is assumed to be the same minimum value for each store department ($c = US\$1 - week/worker$). Henao et al. [1,2,9] and Porto et al. [6,21,54] indicate that the results obtained under this assumption can be interpreted as an upper bound of the requirements multiskilling. The staff shortage cost is equal to $u = US\$60/h$, whereas the staff surplus cost is equal to $b = US\$15/h$. Note that the TSSO, RO, and CF approaches use the same real dataset.

5.2.2.   Simulated datasets

The case study considers in-sample data along with out-of-sample data, both associated with the uncertain demand. Regarding the in-sample data, these consider the data utilized to find the in-sample solutions with the TSSO approach. Regarding the out-of-sample data, these consider the data utilized to make fair comparisons between the performance of the solutions obtained by the optimization approaches under uncertainty (i.e., TSSO, RO, and CF) and the approaches considered as myopic (i.e., ZM and TM). In both sample data types, to evaluate how the 2-chaining structures work in the face of different levels of variability demand, we test six coefficients of variation: $CV = 5, 10, 20, 30, 40, 50\%$. In addition, in both sample data types, we consider that the random realizations of the weekly staff hours demand per store's department are independent and normally distributed with known average value $\bar{r}_l$ and standard deviation $\sigma_l = \bar{r}_l CV$. Note that, in Henao et al. [1,2], Abello et al. [3], Fontalvo Echavez et al. [20], Porto et al. [6,21], and Mercado et al. [14], articles with case studies in the retail industry, authors also used the same probability density function (pdf) to test the benefits of using multiskilled workers.

For the in-sample data, because an SAA method is used for solving the TSSO model (3)-(17), we employ a Monte Carlo simulation to randomly create 2000 scenarios for the staff demand in weekly hours per store's department $r_l(s)$, $\forall l \in L, s \in S$, such that $|S| = 2000$. This simulation is performed for six different CVs of demand. To determine the appropriate sample size of $|S|$, we performed a data sample evaluation and identified that the solutions stabilize from 1000 scenarios. We choose the value

of $|S| = 2000$, considered as a representative sample size for this case study. Note that 2000 weekly demand scenarios are equivalent to a sample of 38.5 years. Later, we created two different datasets to evaluate and compare the conservatism level in the TSSO approach solutions. Each dataset has a different truncation type in the pdf's of the departments' demand. The first dataset considers that the pdfs are truncated between the percentiles 5th and 95th. That is, 10% of the data distribution is excluded, which avoids the generation of extreme demand values. In addition, the pdfs are truncated at zero. The second dataset considers that the pdfs are zero-truncated, which avoids the generation of negative demand values. However, in this second dataset, extreme demand values could be generated more frequently (in both pdf's tails) as the %CV of demand increases. Thus, the second dataset offers a more conservative point of view than the first dataset. That is, TSSO approach solutions obtained under a more conservative truncation may require higher multiskilling levels than those obtained under less conservative truncation. Note that the CF approach also assumes that the pdf for each department's demand follows a normal distribution truncated at zero.

For the out-of-sample data, to simplify the analyses, we employ a Monte Carlo simulation to create a single dataset for the realizations of the random staff demand in weekly hours in all departments. In this case, following a zero-truncated normal distribution and, for each store department and CV, we generated 10000 scenarios for the demand parameters.

Summarizing, we use three simulated datasets. Two in-sample datasets, of which the first is percentile-truncated and the second is zero-truncated; whereas the out-of-sample dataset is zero-truncated only. There are six files in each dataset (one for each CV), and each of them contains random realizations of the weekly staff hours demand for six departments (2000 for in-sample and 10000 for out-of-sample). The simulated datasets are presented in 18 files classified according to: (i) sample data type (in-sample, out-of-sample); (ii) In which way the normal distribution was truncated (percentile-truncated, zero-truncated); and (iii) coefficient of variation for demand in a department (5, 10, 20, 30, 40, 50%). As mentioned before, these files are available in the data article presented in Henao et al. [53].

6. Computational tests, results, and discussion

This section presents the computational experiments, the results, and the discussions associated with the case study. In Subsection 6.1, the in-sample solutions in terms of the training decisions of the TSSO approach are compared with the solutions of RO, CF, and DT approaches. In Subsection 6.2, we carry out an out-of-sample analysis to evaluate cost and reliability benchmarking the TSSO approach to the alternative optimization approaches under uncertainty (i.e., RO and CF), the deterministic (DT) approach, and the two myopic approaches ZM and TM. The models were written in the AMPL language and solved through ILOG CPLEX 12.9.0, such that for all instances, we obtained optimality gaps of less than 0.1%. Additionally, all test instances were run on Intel Xeon Gold 6148, 2.4GHz - 2.39GHz (two processors) with 256 GB of RAM.

6.1. Benchmarking analysis: In-sample

We compared the proposed TSSO model with the benchmark approaches RO, CF, and DT. The DT model presented in Henao et al. [1] was solved using the nominal values of the uncertain demand. The RO and CF approaches were solved as we defined in Subsection 4.2.1 and 4.2.2, respectively. Regarding the RO approach, we evaluated two levels of conservatism or uncertainty budget; $\Gamma = 0.6$

and 1. On one side, $\Gamma = 1$ represents the worst-case scenario and, therefore, the problem's most conservative solution. On the other hand, Henao et al. [1] indicated in their results that $\Gamma = 0.6$ represents a cost-effective value of Γ, since such value achieves savings of at least a 95% on the average total weekly cost of staff surplus/shortage (i.e., $\%\overline{SSS} \geq 95\%$) in the face of different levels of demand uncertainty. The TSSO model (3)-(17) was solved using an SAA approach, as described in Subsection 5.2. The TSSO model solutions are presented considering two in-sample datasets: (i) zero-truncated and (ii) percentile-truncated.

Summarizing, the in-sample analysis compares the performance of the following approaches: (i) two-stage stochastic optimization with a pdf zero-truncated (TSSO-ZT); (ii) two-stage stochastic optimization with a pdf percentile-truncated (TSSO-PT); (iii) robust optimization with $\Gamma = 1$ (RO-1); (iv) robust optimization with $\Gamma = 0.6$ (RO-0.6); (v) closed-form equation (CF); and (vi) deterministic optimization (DT). To compare in-sample solutions in terms of the training decisions, and only for the approaches TSSO-ZT, TSSO-PT, RO-1, RO-0.6 and CF, we evaluate six different demand variability levels: $CV = 5, 10, 20, 30, 40, 50\%$.

Below, we present the results and discussions of the in-sample analysis in three subsections.

6.1.1. Characteristics of the models

Table 3 reports the computational times and size for each model and CV. We observe that the computational times of the TSSO model are much greater than the RO and CF models, which is because the size of the TSSO model is also larger than the others. TSSO-ZT and TSSO-PT models take an average of 39.2 hours and 28.1 hours, respectively, to solve to optimality. Despite the computational time difference compared to the other models, the TSSO approach has acceptable computational times for real-life applications since decisions related to workers training are considered as tactical-strategic. Thus, staff training decisions are taken over planning horizons of medium and long-term (e.g., months, seasons, semesters).

Table 3. Computational times and size of the models.

Approach	CPU time						Average CPU time	Constraints	Binary variables	Continuous variables
	5%	10%	20%	30%	40%	50%				
TSSO-ZT	41.9 h	109.6 h	23.3 h	41.1 h	13.6 h	5.4 h	39.2 h	384,054	150	396,048
TSSO-PT	11.3 h	18.0 h	37.7 h	32.8 h	15.3 h	53.4 h	28.1 h	384,054	150	396,048
RO-1	0.39 s	0.33 s	0.3 s	3.83 s	4.56 s	0.53 s	1.66 s	254	150	247
RO-0.6	0.23 s	0.23 s	0.7 s	0.06 s	0.3 s	5.66 s	1.20 s	254	150	247
CF	0.02 s	0.02 s	0.02 s	0.02 s	0.02 s	0.02 s	0.02 s	-	-	-
DT			0.01 s				0.01 s	240	150	240

In addition, as was expected, the DT model has the lowest computational time. However, the computational times of the RO and CF models are also quite short. Regarding RO models, these have very low computational times since their size is almost identical to the DT model. Finally, the computational times of the CF model are also very low because this approach is solved using conventional spreadsheets, which demonstrates its high practical value. Thus, the DT, CF, and RO approaches have the same order of magnitude for finding an optimal solution, while the TSSO

approach is several orders of magnitude larger, but reasonable for the tactical MPAP.

6.1.2. How much multiskilling should be added?

The objective in this subsection is to respond the following tactical question: (i) How many multiskilled workers are needed per department? Given this purpose, we calculate the optimal number of multiskilled workers per department (λ_l^*) considering six coefficients of variation (CV) in demand and five optimization approaches under uncertainty (TSSO-ZT, TSSO-PT, RO-1, RO-0.6, CF). Figure 2 presents the multiskilling requirements associated with each combination between optimization approach and CV. Recall that, our case study considers a store with six department sizes (DS), whose average values of weekly staff hours are respectively 315, 225, 135, 135, 180, and 360. Finally, the DT approach was not considered in this comparison, since its solution does not require any multiskilled worker (i.e., zero multiskilling). This is because, for each department, the available assignment weekly hours ($|I_l|h$) are equal to the average value of weekly hours demand ($\bar{r}_l$). Thus, there is no imbalance between staffing and staff demand in each store department.

TSSO-ZT

%CV	DS 315	225	135	135	180	360	%ME	%TM
5	1	1	1	1	1	1	20	4
10	2	2	1	1	2	2	33	7
20	4	3	2	2	3	5	63	13
30	5	4	3	3	4	7	87	17
40	7	5	3	3	4	8	100	20
50	7	5	3	3	4	8	100	20
	D1	D2	D3	D4	D5	D6		
	Deparment ID							

TSSO-PT

%CV	DS 315	225	135	135	180	360	%ME	%TM
5	1	1	1	1	1	1	20	4
10	2	1	1	1	1	2	27	5
20	3	2	1	1	2	3	40	8
30	4	3	2	2	2	5	60	12
40	5	4	2	2	3	6	73	15
50	6	5	3	3	4	7	93	19
	D1	D2	D3	D4	D5	D6		
	Deparment ID							

RO-1

%CV	DS 315	225	135	135	180	360	%ME	%TM
5	1	1	1	1	1	1	20	4
10	2	1	1	1	1	2	27	5
20	3	2	2	2	2	4	50	10
30	5	3	2	2	3	5	67	13
40	7	4	3	3	3	7	90	18
50	7	5	3	3	4	8	100	20
	D1	D2	D3	D4	D5	D6		
	Deparment ID							

RO-0.6

%CV	DS 315	225	135	135	180	360	%ME	%TM
5	1	1	1	1	1	1	20	4
10	1	1	1	1	1	2	23	5
20	3	2	1	1	1	3	37	7
30	4	2	2	2	2	4	53	11
40	4	3	3	2	3	5	67	13
50	6	4	3	3	3	7	87	17
	D1	D2	D3	D4	D5	D6		
	Deparment ID							

CF

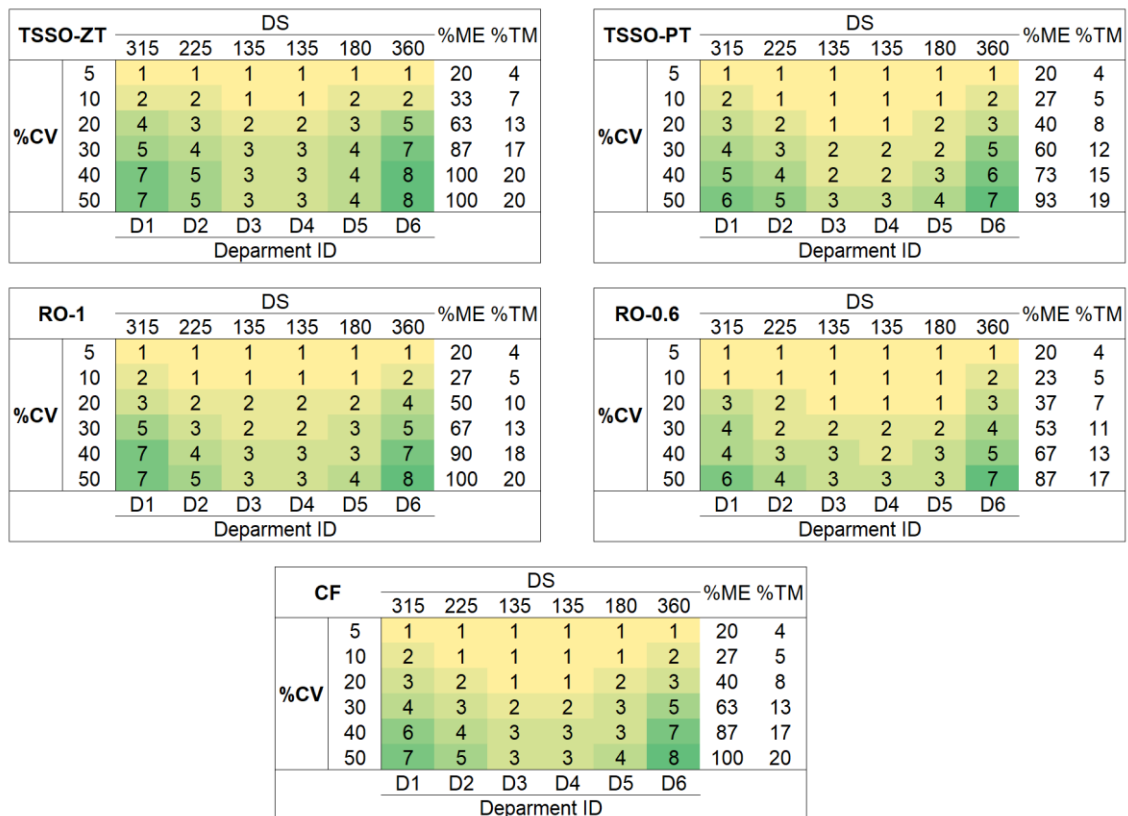| %CV	DS 315	225	135	135	180	360	%ME	%TM
5	1	1	1	1	1	1	20	4
10	2	1	1	1	1	2	27	5
20	3	2	1	1	2	3	40	8
30	4	3	2	2	3	5	63	13
40	6	4	3	3	3	7	87	17
50	7	5	3	3	4	8	100	20
	D1	D2	D3	D4	D5	D6		
	Deparment ID							

Figure 2. Metrics of multiskilling for each CV and optimization approach under uncertainty.

In Figure 2, the colored area shows the optimal number of multiskilled workers in each department size (λ_l^*). It is observed that as the %CV of demand and the DS increase, λ_l^* also increases. Similarly, for each approach, the sum of the multiskilling requirements for all departments ($\sum_{l \in L} \lambda_l^*$, %ME, %TM) increase progressively as %CV of demand rises. There are several aspects worth discussing. It is observed that, for each of the five approaches, when CV=5% (i.e., the minimum

value tested for the %CV) a single multiskilled worker is required per department even when the training cost was considered almost free ($c = 1$). Such a result can be seen as the required minimum multiskilling level (i.e., $ME = 20\%$ and $TM = 4\%$). In contrast, for the RO-1, CF, and TSSO-ZT approaches, when CV=50% (i.e., the maximum value tested for the %CV) all store workers are required to be multiskilled. Such result can be seen as the theoretical maximum (i.e., $\sum_{l \in L} \lambda_l^* = |I|$, $ME = 100\%$, $TM = 20\%$). In fact, once the CV is greater than or equal to 40%, the TSSO-ZT approach requires that 100% of the single-skilled workers be trained to be multiskilled.

Note that when %ME and %TM reach their theoretical maximum, this may involve three possible conclusions. First, $ME = 100\%$ and $TM = 20\%$ represent sufficient multiskilling levels to minimize the staff surplus/shortage costs. Second, for very high demand variability levels, 2-chaining structures could not be enough to minimize the staff surplus/shortage costs, and thus, k-chaining structures with $k \geq 2$ may be explored. That is, chaining structures where a number of workers are trained to perform in two or more departments. Third, for very high demand variability levels, increases in labor flexibility are required but also increases in staffing levels. That is, more workers should be hired.

It is also noteworthy that the conservatism levels in the solutions vary according to the demand variability and the selected approach. Note that, the higher are the required multiskilling levels (%ME, %TM) associated with an approach, the higher is its conservatism level. Considering this criterion, from Figure 2, we order the five approaches according to its conservatism level: (1) TSSO-ZT (the most conservative), (2) RO-1, (3) CF, (4) TSSO-PT, and (5) RO-0.6 (the least conservative). The solutions of the RO-0.6 are the least conservative, mainly because it considers an intermediate uncertainty budget (i.e., $\Gamma = 0.6$). In contrast, the RO-1 approach presents solutions more conservative than the solutions of the CF and TSSO-PT approaches. This intuitive result is explained by the fact that the RO-1 approach aims to protect against the worst-case scenario. That is to say, the RO approach as presented in Subsection 4.2.1, implies that for each CV, the weekly hours demand $\tilde{r}_l$ is a bound random variable taking the maximum permitted variation, $\hat{r}_l = CV\bar{r}_l$, on the interval $[\bar{r}_l - \hat{r}_l, \bar{r}_l + \hat{r}_l]$.

Finally, a less intuitive result indicates that the solutions of the TSSO-ZT are more conservative than the solutions of the RO-1 approach, although not very different. However, this result can be explained because the TSSO-ZT approach considers a range of possible demand realizations wider than the RO-1 approach. Recall that, although under the RO-1 approach the demand realizations take the maximum or minimum value on the interval $[\bar{r}_l - \hat{r}_l, \bar{r}_l + \hat{r}_l]$, the TSSO-ZT uses a dataset where the pdf's of the departments' demand are zero-truncated. Such a dataset represents the maximum theoretical range for the demand realizations in our case of study.

6.1.3. How should multiskilling be added?

The objective in this subsection is to respond the following two tactical questions: (ii) In which second department should be trained each of the workers identified as multiskilled? and (iii) How many CLCs and/or CSCs are needed, and which store departments should belong to each of the chains? Given this purpose, we analyzed the 2-chaining structures generated according to the demand variability (CV) and the evaluated approach (TSSO-ZT, TSSO-PT, RO-1, RO-0.6, CF). Regarding the RO and CF approaches, once the question (i) has been answered in Subsection 6.1.2, it is required to apply the constructive heuristic developed by Henao et al. [1] to answer the questions (ii) and (iii). The constructive heuristic builds the minimum required number of closed chains, replicating the CLC concept. Unlike RO and CF approaches, the TSSO approach simultaneously answers the tactical

questions (i), (ii), and (iii).

As an example, Table 4 presents the set of 2-chaining structures obtained for each approach and a CV of 20%. For each approach, Table 4 shows the optimal total number of multiskilled workers ($\sum_{l \in L} \lambda_l^*$), the constructed closed-chains, and what types of closed chains were built.

Table 4. 2-chaining structures constructed for each approach and for CV=20%.

Approach	$\sum_{l \in L} \lambda_l^*$	Closed chains	Type of closed chain
TSSO-ZT	19	D1-D3-D5-D6-D4-D2-D1	CLC
		D5-D4-D1-D6-D2-D5	CSC
		D2-D3-D6-D2	CSC
		D1-D5-D6-D1	CSC
		D1-D6-D1	CSC
RO-1	15	D3-D1-D4-D5-D6-D2-D3	CLC
		D1-D2-D6-D3-D5-D1	CSC
		D4-D6-D4	CSC
		D1-D6-D1	CSC
CF	12	D3-D1-D4-D5-D6-D2-D3	CLC
		D1-D5-D2-D6-D1	CSC
		D1-D6-D1	CSC
TSSO-PT	12	D1-D3-D6-D5-D4-D1	CSC
		D1-D2-D5-D6-D1	CSC
		D1-D6-D2-D1	CSC
RO-0.6	11	D1-D2-D4-D3-D6-D5-D1	CLC
		D2-D1-D6-D2	CSC
		D1-D6-D1	CSC

Table 4 shows that under the TSSO-ZT approach, five closed chains were built, the first one is a CLC while the other four are CSCs. For example, the constructed chain "D1-D5-D6-D1" indicated that a worker from Department 1 was also trained in Department 5, a worker from Department 5 was also trained in Department 6, and a worker from Department 6 was also trained in Department 1. The RO-1, CF, and RO-0.6 approaches also combine the two closed-chain types (CSCs of not the same length and CLCs). In addition, the TSSO-PT approach constructs three closed short chains of not the same length but no closed long chain. This is a valuable result since it indicates that under a TSSO approach, where all tactical questions were answered simultaneously, various CSCs, each one connecting a significant number of departments, perform as well as a combination of CLCs and CSCs. This result extends the findings reported by Henao et al. [1] and Henao et al. [2].

In addition, for the RO-1, CF, and R0-0.6 approaches, the constructive heuristic seeks to maximize flexibility in the training plans. That is, if possible, it is avoided that the multiskilled workers belonging to a given department have been trained in the same additional departments. The solutions of the TSSO-ZT and TSSO-PT approaches show also that the flexibility is maximized. However, in the set of constructed closed-chains for the TSSO-ZT approach, four training arcs are repeated: "D1-D6", "D5-D6", "D6-D2", and "D6-D1". This is also a valuable result since it indicates that, under a TSSO approach, connecting pairs of departments with large demands more than once is a good strategy. Recall that D1, D2, D5, and D6 represent the departments with the largest sizes in our case study. This result also extends the findings reported by Henao et al. [1] and Henao et al. [2].

Finally, it is important to highlight that, given the methodology used by the RO and CF approaches, it would not have been possible to find these new findings with such approaches. That is, the second

stage of the RO and CF approaches consists of using a constructive heuristic to answer the questions (ii) and (iii), and thus, such a heuristic uses a set of intuitive decision rules (set by the modeler) to create an attractive set of CLCs and CSCs. However, being a heuristic procedure, it explores only a portion of the entire feasible region of solutions. This does not happen with the simultaneous TSSO approach, which explores the entire feasible set of solutions. This feature of the TSSO approach, although it comes with a high computational cost, allowed us to identify and provide new managerial insights to practitioners and retail decision-makers.

6.2. Benchmarking analysis: Out-of-sample

The out-of-sample evaluation presented in this subsection aims to fairly compare the solutions of the TSSO-ZT and TSSO-PT approaches with the solutions of the alternative optimization approaches, RO-1, RO-0.6, and CF. Besides, we evaluate the myopic approaches, i.e., ZM and TM. We remark that, in our case study, the solution of the DT approach is equivalent to the ZM approach. The benchmark evaluation is made by measuring and comparing the performance solutions of these seven approaches. We employ metrics that measure the incurred costs (φ, ψ) and the reliability (%SSS) associated with the solutions delivered by each approach. We use a Monte Carlo simulation to generate 10000 random realizations of the weekly demand per store's department and CV of demand using a normal probability distribution truncated at zero. Then, for each approach and CV, we compute the total weekly cost (φ) and the weekly staff surplus/shortage cost (ψ) related to each of the 10000 demand scenarios by using the multiskilling structures obtained in Subsection 6.1. After, the average costs, $\bar{\varphi}$ and $\bar{\psi}$, are calculated.

For each approach and CV, Table 5 presents in-sample results along with out-of-sample results. Columns 2 and 3 show the requirements of multiskilling (%ME, %TM) obtained in the in-sample results. Columns 4, 5, and 6 show the metrics related to the out-of-sample results. Column 4 indicates the percentage savings on the average total weekly cost of staff surplus/shortage (%$\overline{SSS}$); which represents a measure of reliability. Column 5 reports the average total weekly cost ($\bar{\varphi}$). Finally, Column 6 shows a metric that we call homogeneous group metric. This metric summarizes the results of an inferential analysis performed on the average total costs ($\bar{\varphi}$) of the seven approaches. Following the Fisher's least significant difference (LSD) procedure, we construct 21 confidence intervals for the difference of the average total costs, as a result of the possible combinations between each pair of approaches ($C_2^7 = 7!/2!\,(7-2)! = 21$). The LSD procedure allows to detect significant difference between each pair of approaches. Considering a confidence level equal to 95%, the approaches with X's in the same subcolumn do not present significant differences among them, whereas those that are not in the same subcolumn do present significant differences. Note that, approaches with X's (in bold) in subcolumn "A" have the best performance (i.e., the most cost-effective approaches) in the face of uncertain demand.

From Table 5, several aspects can be highlighted. First, the TSSO-ZT approach always delivers the solution with the lowest average total cost (i.e., $\bar{\varphi}$), whereas the ZM approach always delivers the solution with the highest average total cost. This result is expected given that, except for the TM approach, the TSSO-ZT approach has the highest conservatism level, and the ZM approach does not provide any protection for its solutions since %ME = %TM = 0%.

Table 5. Benchmarking performance between approaches for different CVs of demand.

Approaches	In-sample metrics		Out-of-sample metrics					
	%ME	%TM	%$\overline{SSS}$	$\bar{\varphi}$ (US$)	Homogeneous groups			
					A	B	C	D
CV = 5%								
TSSO-ZT	20	4	100	882	X			
RO-1	20	4	100	882	X			
CF	20	4	100	882	X			
TSSO-PT	20	4	100	882	X			
RO-0.6	20	4	100	882	X			
TM	100	100	100	1,026		X		
ZM (DT)	0	0	0	2,013			X	
CV = 10%								
TSSO-ZT	33	7	100	1,798	X			
TSSO-PT	27	5	100	1,806	X			
RO-1	27	5	99	1,813	X	X		
CF	27	5	99	1,813	X	X		
RO-0.6	23	5	97	1,859		X		
TM	100	100	100	1,937			X	
ZM (DT)	0	0	0	4,046				X
CV = 20%								
TSSO-ZT	63	13	100	3,561	X			
RO-1	50	10	100	3,573	X			
TSSO-PT	40	8	98	3,649	X	X		
CF	40	8	98	3,660	X	X		
TM	100	100	100	3,692		X		
RO-0.6	37	7	94	3,811			X	
ZM (DT)	0	0	0	8,124				X
CV = 30%								
TSSO-ZT	87	17	100	5,324	X			
RO-1	67	13	99	5,363	X			
CF	63	13	99	5,371	X			
TSSO-PT	60	12	99	5,413	X			
TM	100	100	100	5,445	X	X		
RO-0.6	53	11	96	5,594		X		
ZM (DT)	0	0	0	12,154			X	
CV = 40%								
TSSO-ZT	100	20	100	7,190	X			
RO-1	90	18	99	7,237	X			
CF	87	17	99	7,238	X			
TM	100	100	100	7,297	X	X		
TSSO-PT	73	15	98	7,344	X	X		
RO-0.6	67	13	97	7,464		X		
ZM (DT)	0	0	0	16,104			X	
CV = 50%								
TSSO-ZT	100	20	99	9,309	X			
RO-1	100	20	99	9,326	X			
CF	100	20	99	9,326	X			
TSSO-PT	93	19	99	9,354	X			
TM	100	100	100	9,365	X			
RO-0.6	87	17	98	9,508	X			
ZM (DT)	0	0	0	20,170		X		

Second, for each CV, the metric %$\overline{SSS}$ allows us to assess the reliability and robustness of the solutions. As expected, the TM approach exhibits the highest possible reliability and robustness. That is, the TM approach reaches the maximum feasible reduction on the average total weekly cost of staff surplus/shortage (%$\overline{SSS}$=100%), although it also needs the maximum investment in multiskilling

since %ME = %TM = 100%. Conversely, the ZM approach represents no investment in multiskilling, such that %ME = %TM = %$\overline{SSS}$= 0%. In addition, when the approaches TSSO-ZT, RO-1, CF, TSSO-PT, and RO-0.6 are compared, we found that there are similar values of %$\overline{SSS}$ among these five approaches. In turn, the values of this metric were also very close to those reported for the TM approach.

Third, for each CV, the homogeneous group metric indicates that TSSO-ZT, RO-1, CF, and TSSO-PT always belong to the set of approaches with the lowest average total cost (see subcolumn A). In this group, there are not statistically significant differences from one approach to another, so these four approaches are the most cost-effective. Also, the RO-0.6 approach is shown as cost-effective only when CV=5% or CV=50%, i.e., the minimum and maximum values tested for the %CV. In turn, the TM approach is cost-effective only when the CV is high (≥30%). This result is interesting because it can reinforce the hypothesis described in the in-sample analysis, in which it was expressed that for very high demand variability levels, 2-chaining structures could not be enough to minimize the staff surplus/shortage costs. Therefore, it would be valuable to study two additional policies: (i) evaluate k-chaining structures with $k \geq 2$; and (ii) consider multiskilling and staff size decisions simultaneously.

Fourth, we carry out an analysis to provide understanding on how to select the most suitable approach(es) according to the balance between the conservatism (i.e., %ME reported in the in-sample analysis) and the reliability (i.e., %$\overline{SSS}$ reported in the out-of-sample analysis) associated with the solutions delivered by each approach. Figure 3 shows that as the CV of demand rises, the percentage of multiskilled workers (%ME) also increases for each approach. In addition, for each combination of %CV and approach, Figure 3 also shows the value of the %$\overline{SSS}$ metric. Note that, the TM and ZM (DT) approaches represent the upper and lower bound of the level of conservatism, respectively. In addition, it is interesting to note that, as indicated in Table 5 and Figure 3, even when a minimal training cost was assumed, the approaches TSSO-ZT, RO-1, CF, and TSSO-PT achieve between 98% and 100% of the possible benefits of total multiskilling but requiring much lower investments. Besides, the RO-0.6 approach achieves savings of at least a 94% on the average total weekly cost of staff surplus/shortage (i.e., %$\overline{SSS}$ ≥ 94%,∀ CV), even with a lower multiskilling investment than the one required by the other four optimization approaches under uncertainty.

Finally, from Figure 3, we order the approaches according to the level of conservatism as reported in the solutions: (1) TM (the most conservative), (2) TSSO-ZT, (3) RO-1, (4) CF, (5) TSSO-PT, (6) RO-0.6, and (7) ZM (DT) (the least conservative). We remark that the TSSO-ZT, TSSO-PT, and CF approaches assume complete knowledge of the distributional information, whereas the RO-1 and RO-0.6 approaches do not require this full knowledge of the stochastic parameters. In addition, the TSSO approach can simultaneously answer the tactical questions (i), (ii), and (iii), and not sequentially as RO and CF approaches do. Thus, from the results, we observe that the suitable approach under demand uncertainty will depend on the preferred solution method (simultaneous or sequential), the decision-maker's knowledge on the probability distribution, and his/her degree of risk-aversion over the decision.
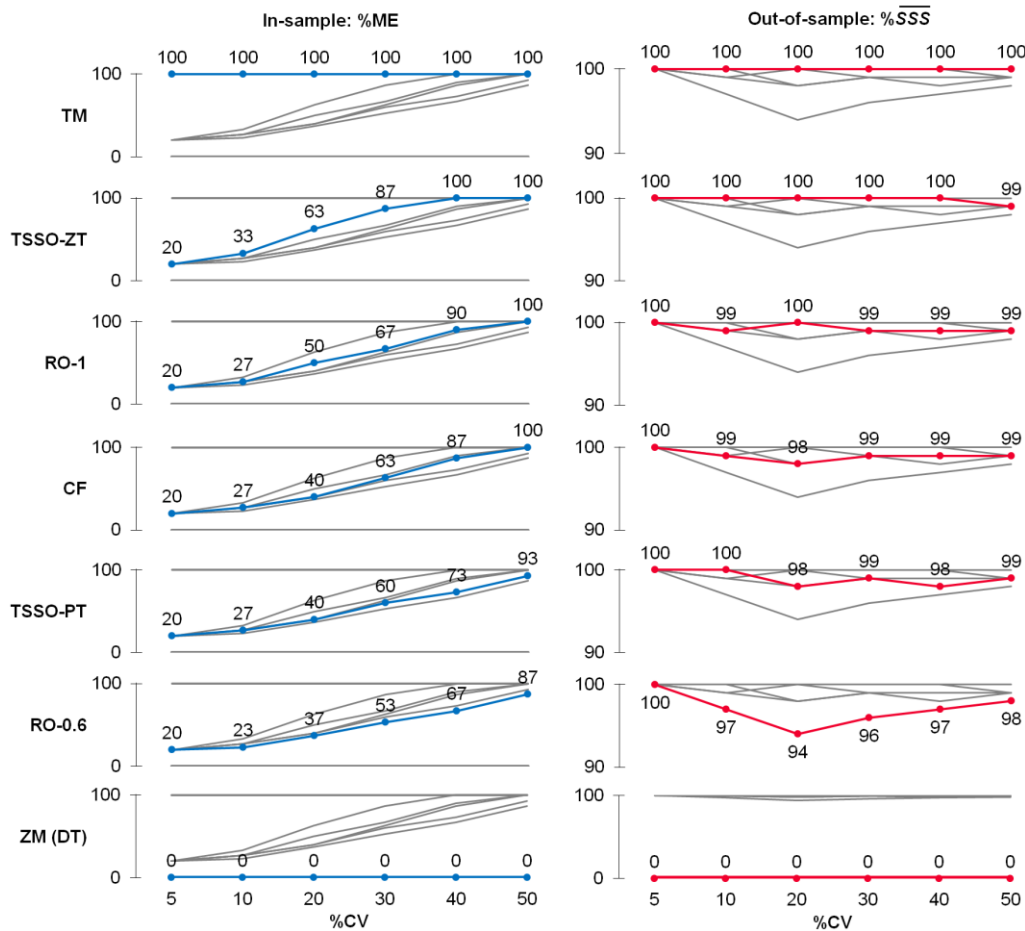
Figure 3. Multiskilling levels and reliability by CV and approach.

7. Concluding remarks

This article studies the weekly multiskilled personnel assignment problem (MPAP) considering 2-chaining and uncertain demand. The solution of the MPAP generates staff training plans structured through closed chains (CLCs, and CSCs) and where the workers could be trained for only one additional task type. The contribution of this paper is twofold. Firstly, we developed a TSSO approach to calculate the multiskilling requirements that minimize the training costs and the expected costs of staff surplus/shortage. Secondly, we evaluated and compared the performance of the solutions generated by the TSSO approach with the solutions of alternative optimization approaches under uncertainty (i.e., RO and CF), the solution of deterministic (DT) approach, and the solutions of myopic approaches (i.e., ZM and TM).

The methodology was tested for a case study associated with a retail store operating in Chile. Through in-sample and out-of-sample analyses employing real and simulated data, we evaluated and compared the different approaches by testing six representative levels of staff demand per department. On the in-sample analysis, we evaluated the performance of the optimization approaches under uncertainty TSSO, RO, and CF, such that the findings can be grouped under the following two aspects. First, given the reported computational times, the RO and CF approaches can be useful for larger problems; more store departments and workers. Second, the obtained results under the TSSO approach

and its comparison with the other approaches allow us to incorporate new guidelines to design practical and effective policies that answer the following three tactical questions: (i) How many multiskilled workers are needed per department? (ii) In which second department should be trained each of the workers identified as multiskilled? (iii) How many CLCs and/or CSCs are needed, and which store departments should belong to each of the chains? We remark that, unlike RO and CF approaches, the TSSO approach simultaneously answers the tactical questions (i), (ii), and (iii).

Regarding the first question, the results of the TSSO, RO, and CF approaches showed that, even when the cost of training was considered almost free, total multiskilling is impractical and is not required to achieve the full possible benefits associated with multiskilling. In addition, our results (intuitively) showed that the required number of multiskilled workers rises as the department size and the level of demand variability grow. Finally, we found a less intuitive result when the solutions to the optimization approaches under uncertainty were compared. The solutions of the TSSO approach where the demand parameters follow a zero-truncated normal distribution (i.e., TSSO-ZT) were more conservative than the solutions of the RO approach with the maximum uncertainty budget (i.e., worst-case scenario). Note that, the higher are the required multiskilling levels associated with an approach, the higher is its level of conservatism.

Regarding the second and third questions, we found that the most profitable staff is one where there is a mix of single-skilled and multiskilled workers. This result is observed with all optimization approaches under uncertainty. In turn, and in a novel way, we also found that under a TSSO approach, various CSCs, each one connecting a significant number of departments, perform as well as a combination of CLCs and CSCs. Finally, under the RO and CF approaches, it seeks to maximize flexibility in the training plans, that is, avoiding that the multiskilled workers belonging to a given department were trained in the same additional departments. However, under the TSSO approach, we discovered that closed chains connecting pairs of departments with large demands more than once also exhibited excellent performance.

On the out-of-sample analysis, we evaluated the performance of the optimization approaches under uncertainty TSSO, RO, and CF, the deterministic approach DT, and the two myopic approaches TM and ZM. The findings provided understanding on how to select the most suitable approach(es) according to the balance between the conservatism and the reliability associated with the solutions delivered by each approach. Particularly, it was shown that the TSSO-ZT approach always delivers the solution with the lowest average total cost. However, the main finding was that TSSO-ZT, RO-1, CF, and TSSO-PT always belong to the set of approaches with the lowest average total cost. That is, in this group, there are not statistical differences from one approach to another, so these approaches are the most cost-effective. This finding showed that there is not only one approach to effectively solve the MPAP with 2-chaining under uncertain demand. Finally, considering the overall results, we concluded that the choice of the most suitable approach would depend on the preferred solution method (i.e., simultaneous or sequential methodology), the available information of the distribution of probability associated with the uncertain demand, and the level of risk tolerance of the store manager.

As for future research, the methodology could incorporate new additional features, or some assumptions could be relaxed. (1) Assess k-chaining structures with $k \geq 2$ where a number of workers are trained to perform in two or more departments. (2) Solve simultaneously a staffing and personnel assignment problem considering closed-chained multiskilling structures. (3) Solve the MPAP under uncertain demand by a distributionally robust optimization, which is a suitable approach for ambiguity-averse decision-makers.

Acknowledgments

Conflict of interest

All authors declare no conflicts of interest in this paper.

References

1. C. A. Henao, J. C. Ferrer, J. C. Muñoz, J. A. Vera, Multiskilling with closed chains in the service sector: a robust optimization approach, *Int. J. Prod. Econ.*, **179** (2016), 166–178. https://doi.org/10.1016/j.ijpe.2016.06.013

2. C. A. Henao, J. C. Muñoz, J. C. Ferrer, Multiskilled workforce management by utilizing closed chains under uncertain demand: a retail industry case, *Comput. Ind. Eng.*, **127** (2019), 74–88. https://doi.org/10.1016/j.cie.2018.11.061

3. M. A. Abello, N. M. Ospina, J. M. De la Ossa, C. A. Henao, V. I. González, Using the k-chaining approach to solve a stochastic days-off-scheduling problem in a retail store, *in Production Research: ICPR-Americas 2020, Communications in Computer and Information Science* (eds. D. A. Rossit, F. Tohmé and G. Mejía Delgadillo), Springer, Cham, **1407** (2021), 156–170. https://doi.org/10.1007/978-3-030-76307-7_12

4. R. Muñoz, J. C. Muñoz, J. C. Ferrer, V. I. González, C. A. Henao, When should shelf stocking be done at night? A workforce management optimization approach for retailers, *Work in progress*.

5. C. A. Henao, *Diseño de una fuerza laboral polifuncional para el sector servicios: caso aplicado a la industria del retail*, (Tesis Doctoral, Pontificia Universidad Católica de Chile, Santiago, Chile), 2015. [Online]. Available from: https://repositorio.uc.cl/handle/11534/11764

6. A. F. Porto, C. A. Henao, H. López-Ospina, E. R. González, Hybrid flexibility strategic on personnel scheduling: retail case study, *Comput. Ind. Eng.*, **133** (2019), 220–230. https://doi.org/10.1016/j.cie.2019.04.049

7. E. Álvarez, J. C. Ferrer, J. C. Muñoz, C. A. Henao, Efficient shift scheduling with multiple breaks for full-time employees: A retail industry case, *Comput. Ind. Eng.*, **150** (2020), 106884. https://doi.org/10.1016/j.cie.2020.106884

8. M. Mac-Vicar, J. C. Ferrer, J. C. Muñoz, C. A. Henao, Real-time recovering strategies on personnel scheduling in the retail industry, *Comput. Ind. Eng.*, **113** (2017), 589–601. https://doi.org/10.1016/j.cie.2017.09.045

9. C. A. Henao, J. C. Muñoz, J. C. Ferrer, The impact of multi–skilling on personnel scheduling in the service sector: a retail industry case, *J. Oper. Res. Soc.*, **66** (2015), 1949–1959. https://doi.org/10.1057/jors.2015.9

10. Y. Wang, J. Tang, Optimized skill configuration for the seru production system under an uncertain demand, *Ann. Oper. Res.*, **2020** (2020), 1-21. https://doi.org/10.1007/s10479-020-03805-3

11. Y. A. Mercado, C. A. Henao, Benefits of multiskilling in the retail industry: K-chaining approach

with uncertain demand, in *Production Research: ICPR-Americas 2020, Communications in Computer and Information Science* (eds. D. A. Rossit, F. Tohmé and G. Mejía Delgadillo), Springer, Cham, **1407** (2021), 126–141. https://doi.org/10.1007/978-3-030-76307-7_10

12. S. Vergara, J. Del Villar, J. Masson, N. Pérez, C. A. Henao, V. I. González, Impact of labor productivity and multiskilling on staff management: A retail industry case, *in Production Research: ICPR-Americas 2020, Communications in Computer and Information Science* (eds. D. A. Rossit, F. Tohmé and G. Mejía Delgadillo), Springer, Cham, **1408** (2021), 223–237. https://doi.org/10.1007/978-3-030-76310-7_18

13. C. Liu, Z. Li, J. Tang, X. Wang, M. J. Yao, How SERU production system improves manufacturing flexibility and firm performance: an empirical study in China, *Ann. Oper. Res.*, **2021** (2021), 1–26. https://doi.org/10.1007/s10479-020-03850-y

14. Y. A. Mercado, C. A. Henao, V. I. González, A two-stage stochastic optimization model for the retail multiskilled personnel scheduling problem: A k-chaining policy with k≥2, *Math. Biosci. Eng.*, **19** (2022), 892–917. https://doi.org/10.3934/mbe.2022041

15. W. C. Jordan, S. C. Graves, Principles on the benefits of manufacturing process flexibility, *Manag. Sci.*, **41** (1995), 577–594. https://doi.org/10.1287/mnsc.41.4.577

16. M. J. Brusco, T. R. Johns, Staffing a multiskilled workforce with varying levels of productivity: An analysis of cross-training policies, *Decision Sci.,* **29** (1998), 499–515. https://doi.org/10.1111/j.1540-5915.1998.tb01586.x

17. G. K. Taskiran, X. Zhang, Mathematical models and solution approach for cross-training staff scheduling at call centers, *Comput. Oper. Res.*, **87** (2017), 258-269. https://doi.org/10.1016/j.cor.2016.07.001

18. H. Liu, The optimization of worker's quantity based on cross-utilization in two departments, *Intell. Decis. Technol.*, **11** (2017), 3–13. https://doi.org/10.3233/IDT-160273

19. D. Simchi-Levi, Y. Wei, Understanding the performance of the long chain and sparse designs in process flexibility, *Oper. Res.*, **60** (2012), 1125–1141. https://doi.org/10.1287/opre.1120.1081

20. O. Fontalvo Echavez, L. Fuentes Quintero, C. A. Henao, V. I. González, Two-stage stochastic optimization model for personnel days–off scheduling using closed-chained multiskilling structures, *in Production Research: ICPR-Americas 2020, Communications in Computer and Information Science* (eds. D. A. Rossit, F. Tohmé and G. Mejía Delgadillo), Springer, Cham, **1407** (2021), 19–32. https://doi.org/10.1007/978-3-030-76307-7_2

21. A. F. Porto, C. A. Henao, A. Lusa, O. Polo Mejía, R. Porto Solano, Solving a staffing problem with annualized hours, multiskilling with 2-chaining, and overtime: a retail industry case, *Comput. Ind. Eng.*, **167** (2022), 107999. https://doi.org/10.1016/j.cie.2022.107999

22. J. Birge, F. Louveaux F, *Introduction to stochastic programming*, (2011), Springer, New York. https://doi.org/10.1007/978-1-4614-0237-4

23. D. Bertsimas, M. Sim, Robust discrete optimization and network flows, *Math. Program.*, **98** (2003), 49–71. https://doi.org/10.1007/s10107-003-0396-4

24. D. Bertsimas, M. Sim, The price of robustness, *Oper. Res.*, **52** (2004), 35–53. https://doi.org/10.1287/opre.1030.0065

25. W. Wiesemann, D. Kuhn, M. Sim, Distributionally robust convex optimization, *Oper. Res.*, **62** (2014), 1358–1376. https://doi.org/10.1287/opre.2014.1314

26. A. Ernst, H. Jiang, M. Krishnamoorthy, B. Owens, D. Sier, An annotated bibliography of personnel scheduling and rostering, *Ann. Oper. Res.*, **127** (2004), 21–144.

https://doi.org/10.1023/B:ANOR.0000019087.46656.e2

27. P. Brucker, R. Qu, E. Burke, Personnel scheduling: Models and complexity, *Eur. J. Oper. Res.*, **210** (2011), 467–473. https://doi.org/10.1016/j.ejor.2010.11.017

28. J. Van den Bergh, J. Belien, P. De Bruecker, E. Demeulemeester, L. De Boeck, Personnel scheduling: A literature review, *Eur. J. Oper. Res.*, **226** (2013), 367–385. https://doi.org/10.1016/j.ejor.2012.11.029

29. E. H. Özder, E. Özcan, T. Eren, A systematic literature review for personnel scheduling problems, *Int. J. Inform. Technol. Decision Making*, **19** (2020), 1695–1735. https://doi.org/10.1142/S0219622020300050

30. J. F. Bard, D. P. Morton, Y. M. Wang, Workforce planning at USPS mail processing and distribution centers using stochastic optimization, *Ann. Oper. Res.*, **155** (2007), 51–78. https://doi.org/10.1007/s10479-007-0213-1

31. X. Zhu, H. D. Sherali, Two-stage workforce planning under demand fluctuations and uncertainty, *J. Oper. Res. Soc.*, **60** (2009), 94–103. https://doi.org/10.1057/palgrave.jors.2602522

32. T. R. Robbins, T. P. Harrison, A stochastic programming model for scheduling call centers with global service level agreements, *Eur. J. Oper. Res.*, **207** (2010), 1608–1619. https://doi.org/10.1016/j.ejor.2010.06.013

33. G.M. Campbell, A two-stage stochastic program for scheduling and allocating cross-trained workers, *J. Oper. Res. Soc.*, **62** (2011), 1038–1047. https://doi.org/10.1057/jors.2010.16

34. S. Liao, G. Koole, C. Van Delft, O. Jouini, Staffing a call center with uncertain non-stationary arrival rate and flexibility, *OR Spectrum*, **34** (2012), 691–721. https://doi.org/10.1007/s00291-011-0257-0

35. S. Liao, C. Van Delft, J. P. Vial, Distributionally robust workforce scheduling in call centres with uncertain arrival rates, *Optim. Methods Software*, **28** (2013), 501–522. https://doi.org/10.1080/10556788.2012.694166

36. A. Gnanlet, W. Gilland, Impact of productivity on cross-training configurations and optimal staffing decisions in hospitals, *Eur. J. Oper. Res.*, **238** (2014), 254–269. https://doi.org/10.1016/j.ejor.2014.03.033

37. J. Paul, L. MacDonald, Modeling the benefits of cross-training to address the nursing shortage, *Int. J. Prod. Econ.*, **150** (2014), 83–95. https://doi.org/10.1016/j.ijpe.2013.11.025

38. K. Kim, S. Mehrotra, A two-stage stochastic integer programming approach to integrated staffing and scheduling with application to nurse management, *Oper. Res.*, **63** (2015), 1431–1451. https://doi.org/10.1287/opre.2015.1421

39. A. Parisio, C. N. Jones, A two-stage stochastic programming approach to employee scheduling in retail outlets with uncertain demand, *Omega*, **53** (2015), 97–103. https://doi.org/10.1016/j.omega.2015.01.003

40. M. Bodur, J. R. Luedtke, Mixed-integer rounding enhanced benders decomposition for multiclass service-system staffing and scheduling with arrival rate uncertainty, *Manag. Sci.*, **63** (2017), 2073–2091. https://doi.org/10.1287/mnsc.2016.2455

41. S. Mattia, F. Rossi, M. Servilio, S. Smriglio, Staffing and scheduling flexible call centers by two-stage robust optimization, *Omega*, **72** (2017), 25–37. https://doi.org/10.1016/j.omega.2016.11.001

42. M. I. Restrepo, B. Gendron, L. M. Rousseau, A two-stage stochastic programming approach for multi-activity tour scheduling, *Eur. J. Oper. Res.*, **262** (2017), 620–635. https://doi.org/10.1016/j.ejor.2017.04.055

43. D. S. Altner, A. C. Rojas, L. D. Servi, A two-stage stochastic program for multi-shift, multi-analyst, workforce optimization with multiple on-call options, *J. Schedul.*, **21** (2018), 517–531. https://doi.org/ 10.1007/s10951-017-0554-9

44. D. S. Altner, E. K. Mason, L. D. Servi, Two–stage stochastic days–off scheduling of multi–skilled analysts with training options, *J. Comb. Optim.*, **38** (2019), 111–129. https://doi.org/10.1007/s10878-018-0368-5

45. H.G. Beyer, B. Sendhoff, Robust optimization - A comprehensive survey, *Computer Methods Appl. Mech. Eng.*, **196** (2007), 3190–3218. https://doi.org/10.1016/j.cma.2007.03.003

46. P. Kouvelis, G. Yu, *Robust discrete optimization and its applications*, Springer, Boston, 2013. https://doi.org/10.1007/978-1-4757-2620-6

47. V. Gabrel, C. Murat, A. Thiele, Recent advances in robust optimization: An overview, *Eur. J. Oper. Res.*, **235** (2014), 471–483. https://doi.org/10.1016/j.ejor.2013.09.036

48. D. Bertsimas, A. Thiele, A robust optimization approach to inventory theory, *Oper. Res.*, **54** (2006), 150–168. https://doi.org/10.1287/opre.1050.0238

49. C. Bohle, S. Maturana, J. Vera, A robust optimization approach to wine grape harvesting scheduling, *Eur. J. Oper. Res.*, **200** (2010), 245–252. https://doi.org/10.1016/j.ejor.2008.12.003

50. F. Robuste, C. F. Daganzo, R. R. Souleyrette II, Implementing vehicle routing models, *Transport. Res. Part B Methodol.*, **24** (1990), 263–286. https://doi.org/10.1016/0191-2615(90)90002-G

51. J. M. Del Castillo, A heuristic for the traveling salesman problem based on a continuous approximation, *Transport. Res. Part B Methodol.*, **33** (1990), 123–152. https://doi.org/10.1016/S0191-2615(98)00034-4

52. C. F. Daganzo, *Logistics Systems Analysis* (4th edition), Springer, Berlin, 2005.

53. C. A. Henao, A. Batista, A. F. Porto, V. I. González, A benchmark dataset for the retail multiskilled personnel planning under uncertain demand, *Data Brief*, Forthcoming 2022.

54. A. F. Porto, C. A. Henao, H. López-Ospina, E. R. González, V. I. González, Dataset for solving a hybrid flexibility strategy on personnel scheduling problem in the retail industry, *Data Brief*, **32** (2020), 106066. https://doi.org/10.1016/j.dib.2020.106066