



ISTITUTO DI ANALISI DEI SISTEMI ED INFORMATICA
“Antonio Ruberti”
CONSIGLIO NAZIONALE DELLE RICERCHE

S. Mattia, F. Rossi, M. Servilio, S. Smriglio

**STAFFING AND SCHEDULING FLEXIBLE CALL
CENTERS BY TWO-STAGE ROBUST
OPTIMIZATION**

R. 2, 2015

Sara Mattia – Istituto di Analisi dei Sistemi ed Informatica, Consiglio Nazionale delle Ricerche
— via dei Taurini 19, 00185 Roma, Italy (sara.mattia@iasi.cnr.it).

Fabrizio Rossi – Dipartimento di Ingegneria e Scienze dell'Informazione e Matematica, Uni-
versità di L'Aquila — via Vetoio I-67010 Coppito (AQ) (fabrizio.rossi@univaq.it).

Mara Servilio – Dipartimento di Elettronica, Informazione e Bioingegneria, Politecnico di
Milano — piazza Leonardo Da Vinci, 32 20133 Milano, Italy (mara.servilio@polimi.it).

Stefano Smriglio – Dipartimento di Ingegneria e Scienze dell'Informazione e Matematica, Uni-
versità di L'Aquila — via Vetoio I-67010 Coppito (AQ) (stefano.smriglio@univaq.it).

ISSN: 1128–3378

Collana dei Rapporti dell'Istituto di Analisi dei Sistemi ed Informatica "Antonio Ruberti",
CNR

via dei Taurini 19, 00185 ROMA, Italy

tel. ++39-06-49937101/02

fax ++39-06-49937106

email: iasi@iasi.cnr.it

URL: <http://www.iasi.cnr.it>

Abstract

We study the shift scheduling problem in a multi-shift, flexible call center where moving agents between front end and back office is allowed to follow the actual demand. Differently from previous approaches, the staffing levels, required to provide a service with desired quality, are considered uncertain. This perspective naturally leads to a two-stage robust integer program with right-hand-side uncertainty (R-IP-RSHU). Three different uncertainty sets are investigated, also including correlations between consecutive time slots. In contrast to the general R-LP-RSHU problem, we show that the associated LP relaxation is polynomially solvable. A branch-and-cut algorithm based on a Benders type reformulation is also devised to solve the integer problem and tested on real-world data from a large Italian call center. The algorithm turns out to be effective and suitable to support relevant managers decisions. In fact, a high level of protection is achieved with a limited additional cost.

1. Introduction

Workforce management (WFM) is a complex process and represents a prominent issue in call centers optimization. Its general goal is to find a satisfactory trade-off between the Level of Service (LoS) provided to customers and personnel costs. LoS is mostly based on waiting times and quality of the response, while personnel costs represent one major expense for call center Companies. An insightful description of both these aspects can be found in [15]. WFM is traditionally split into a sequence of almost separate steps [3]: forecasting call volumes; determining the *staffing levels*, defined as the number of agents required at each time period to guarantee the desired LoS; translating them into agents work shifts (*shift scheduling*); assigning agents to such shifts (*rostering*) and, finally, monitoring out-of-adherence situations at operational level and reacting accordingly. In practice, the arrival rate of calls is quite difficult to estimate, resulting in frequent *overstaffing* and *understaffing* situations [26, 29]. Overstaffing may lead to higher costs [1] and even affect agents satisfaction, while understaffing typically lowers the LoS and impacts on the revenues. To overcome these challenges, call centers throughout the industry are exploring different flexibility models, ranging from recruiting temporary operators to delaying less urgent calls [10]. We focus on a specific flexible setting, investigated in [17], that we also encountered in several applied projects. In this setting managers react to front-end understaffing by moving agents from back to front office, eventually accepting some delays in the back office work (e.g., paperworks, e-mails and calls with callback). From now on, a call center implementing such a practice is referred to as *flexible* call center.

WFM poses theoretical and algorithmic challenges in operations management as well. In fact, several topics in this context are object of huge research, and we refer the reader to the insightful surveys [10, 3] for a comprehensive view. According to the practical decomposition, analytic queueing models or simulation models have been developed for computing staffing levels able to guarantee the desired LoS, while integer programming algorithms have been used to determine optimal shift schedules able to cover such levels. However, most of the assumptions of standard queueing models are often not valid in practice, especially in multi-queue/multi-skill environments. For instance, the *Stationary Independent Period by Period* (SIPP) model assumes the independence of the arrival processes at consecutive periods, whereas significant correlations across time periods are well-known to occur in practice. These affect the quality of the computed staffing levels, as documented in [11, 5]. A new stream of research has been recently started with the purpose of overcoming the SIPP model and looking at more complex and realistic arrival processes. This naturally leads to integrate staffing and shift scheduling decisions into suitable mathematical programming formulations. For instance, in [5, 4], formulations are presented to minimize staffing costs while preserving the LoS and sophisticated simulation-based cutting plane methods are developed. Here, simulation is required to compute the LoS, which is not analytically accessible. In [16, 9, 28] similar methodologies are applied to a staffing-scheduling problem with a single global service level constraint. Unlike previous papers, in [28] a stochastic programming formulation is introduced. The migration towards this kind of models is further developed in recent papers. In [18], stochastic programming is combined with distributionally robust optimization for multi-period, multi-shift staffing problems under service level constraints, with uncertain calls arrival rates and an intra-day seasonality. In [17], a stochastic programming approach is compared to a robust optimization based method applied to the solution of a staffing problem in a multi-period, single-shift, flexible call center. Finally, in [8] the authors present a stochastic programming model in which the LoS is approximated by a linear program instead of

using simulation. This improves computational tractability at the price of reducing the accuracy.

In this paper we study a *two-stage* robust optimization model for shift scheduling in a multi-period, multi-shift, flexible call center. The distinguishing feature of this model is that the staffing levels required to guarantee the desired LoS are uncertain, being affected by errors in the call volume forecast and by point estimate approximations. This model introduces a new perspective, as it disdains the description of the arrival process as well as the detailed modeling of the LoS. In fact, it is complementary to any staffing method, being designed to react to unpredictable demand patterns at the shift scheduling stage. A preliminary version of this model has been presented in the conference paper [20], where its adherence with managers' practice in exploiting flexibility is discussed. Our model belongs to the class of two-stage problems with right-hand-side uncertainty (RHSU), typically tackled by a Benders-like reformulation. The continuous version of such problems has been investigated in [23], where the separation problem associated to the Benders constraints is proved to be strongly NP-hard in general. We identify three different form of the uncertainty set which are of practical relevance and investigate the complexity issue accordingly. In particular, we show that the separation problem is solvable in strongly polynomial time when the SIPP model is considered and in pseudo-polynomial time when correlation between consecutive time periods is taken into account. A branch-and-cut algorithm is then devised to manage the integer case and experimented on real-world data from a large Italian call center. The favorable theoretical complexity has also born out in practice, as the convergence of the algorithm is remarkably faster than what is typically observed from formulations of this kind. This allows to exploit the algorithm in addressing relevant managerial issues, such as the trade-off between personnel costs and protection against uncertainty. Interestingly, high level of protections turns out to be achievable with shift plans whose cost is comparable to the one computed from deterministic staffing levels. On the contrary, exploiting flexibility to re-adjust deterministic plans can be remarkably more expensive.

The paper is organized as follows. In §2 we discuss the flexible version of the classic shift scheduling model. In §3 we illustrate the two-stage robust optimization approach, describe the uncertainty set and derive the Benders-like reformulation. The complexity of the separation problem is addressed in §4 and the solution algorithm described in §5.1. The computational experience is described in §5. Finally, in §6, some conclusions are drawn.

2. Flexible shift scheduling

We consider a discrete planning horizon $T = \{1, \dots, m\}$ and denote by b_t the staffing level at period t , $t \in T$, i.e., the (integer) number of agents required on duty in period t . Agents are assigned with work shifts, each characterized by a starting time and a duration. Work shifts do not include breaks which are assumed to be managed at real-time level. Let $J = \{1, \dots, n\}$ be the set of all possible shifts and c_j the cost associated to shift j . The classical shift scheduling problem consists in determining the number of agents to be assigned to each shift in order to satisfy the levels at minimum cost. A basic integer programming model for this problem was introduced in [25]. Let us define the shift matrix $\mathbf{A} \in \{0, 1\}^{m \times n}$, with $a_{tj} = 1$ if shift j covers period t and 0 otherwise, and denote by x_j the number of agents assigned to shift j , $j \in J$. The problem writes:

$$\min \sum_{j \in J} c_j x_j \tag{1}$$

$$\begin{aligned} \sum_{j \in J} a_{tj} x_j &\geq b_t \quad t \in T \\ \mathbf{x} &\in \mathbb{Z}_+^n \end{aligned}$$

Since shifts do not include breaks, each column has consecutive ones (C1P property), which allows to reformulate the problem as a minimum cost flow problem [25]. Here we show that the model can be adapted to include flexibility while preserving the network structure. Recall that flexibility consists in the possibility to react to (front-end) understaffing and overstaffing situations by moving personnel between back and front office. Of course, a natural property holds:

Property 2.1. *At any time slot, it is not possible to have overstaffing and understaffing simultaneously.*

Let us introduce variables $o_t, u_t \in \mathbb{Z}_+$ to represent the number of agents in excess (overstaffing) resp. defect (understaffing) at period t . These situations yield additional costs, induced by organizational and even behavioral issues, which may become quite relevant. Let w_t^o and w_t^u be the overstaffing and understaffing cost at period t and we suppose that for any $t \in T$, $w_t^o + w_t^u > 0$. The *flexible model* reads as:

$$\begin{aligned} \min \sum_{j \in J} c_j x_j + \sum_{t \in T} (w_t^o o_t + w_t^u u_t) & \quad (2) \\ \sum_{j \in J} a_{tj} x_j + u_t - o_t = b_t & \quad t \in T \\ \mathbf{x} \in \mathbb{Z}_+^n, \mathbf{o}, \mathbf{u} \in \mathbb{Z}_+^m & \end{aligned}$$

Variables \mathbf{o}, \mathbf{u} act as surplus and slack variables, modeling personnel reallocation between front and back office implemented to (exactly) balance the demand. It is well-known [2], that if a constraint matrix is in the form $[A \mid -I]$ with A having C1P by column, the problem can be reformulated as min-cost flow. The same holds for $[A \mid I \mid -I]$.

Theorem 2.1. *Consider problem*

$$\begin{aligned} \min c^T x + d^T y + g^T z \\ Ax + Iy - Iz = b \\ x \in \mathbb{Z}_+^n, y, z \in \mathbb{Z}_+^m \end{aligned}$$

If A has the C1P by column then the problem can be reformulated as min-cost flow.

PROOF. Using a standard technique, first add redundant row $\mathbf{0} = \mathbf{0}$ to the constraints matrix. Then, replace the i -th constraint by the difference of the $(i+1)$ -th and the i -th, for $i = m+1, m, \dots, 1$. The new constraints matrix corresponds to the node-arc incidence matrix of a directed graph $G(V, E_z \cup E_y \cup E_x)$ with $|V| = m+1$. E_y (E_z) includes arcs $y_i = (i, i+1)$ ($z_i = (i+1, i)$) whose cost is g_i (d_i) for $1 \leq i \leq m$, while E_x has an arc (u_j, l_j+1) of cost c_j for each $j \in J$, where u_j (l_j) are first and last row where the coefficient of x_j is 1 in A . Each node $i \neq m+1$ of G has a supply/demand equal to $b'_i = b_{i+1} - b_i$, while $b'_{m+1} = -b_m$. See Figure 1 for a graphical representation of G . By construction, $\sum_{i \in V} b'_i = 0$. Solving problem (2) amounts to compute a minimum cost flow on G . \square

Applying Theorem 2.1 to our problem, we obtain the following result.

Corollary 2.2. *Model (2) can be solved in $O((m+n)^2 \log m + m(n+m) \log^2 m)$.*

PROOF. Constraint matrix of problem (2) is in the form $[A \mid I \mid -I]$ with A having the $C1P$ property, therefore Lemma 2.1 holds. Since \mathbf{b} is integer, \mathbf{b}' is integer and the optimal flow will be integral. The graph has $O(m)$ nodes and $O(n+m)$ edges. As stated in [24], the complexity of solving a min-cost flow problem in graph $G(V, E)$ is $O(|E|^2 \log |V|^2 + |E||V| \log^2 |V|)$, hence problem (2) can be solved in $O((m+n)^2 \log m + m(n+m) \log^2 m)$. \square

An optimal solution to model (2) naturally satisfies Property 2.1, that is, for any time period $t \in T$, the optimal value of at most one between u_t and o_t will be positive. Moreover, having such a property is a necessary condition for a solution to be optimal.

Lemma 2.3. *A necessary condition for a solution $(\mathbf{x}, \mathbf{o}, \mathbf{u})$ to be optimal for (2), is that Property 2.1 holds for (\mathbf{o}, \mathbf{u}) .*

PROOF. Observe that for each pair of nodes $(i, i+1)$ for $i = 1, \dots, m$, arcs u_i, o_i induce a positive cost cycle. Suppose that $(\mathbf{x}, \mathbf{o}, \mathbf{u})$ is an optimal solution such that there exists $\bar{t} \in T$ where $u_{\bar{t}} > 0$ and $o_{\bar{t}} > 0$. Suppose without loss of generality that $u_{\bar{t}} > o_{\bar{t}}$ and consider solution $(\mathbf{x}, \bar{\mathbf{u}}, \bar{\mathbf{o}})$, where: $\bar{u}_t = u_t$, $\bar{o}_t = o_t$ for any $t \in T \setminus \{\bar{t}\}$; $\bar{u}_{\bar{t}} = u_{\bar{t}} - o_{\bar{t}}$, $\bar{o}_{\bar{t}} = 0$. Solution $(\mathbf{x}, \bar{\mathbf{o}}, \bar{\mathbf{u}})$ is feasible and has a better objective value than $(\mathbf{x}, \mathbf{o}, \mathbf{u})$, as it is obtained by reducing flow along a positive cost cycle, then $(\mathbf{x}, \mathbf{o}, \mathbf{u})$ cannot be optimal. This means that, for any time period $t \in T$, the optimal value of at most one between u_t and o_t will be positive. \square

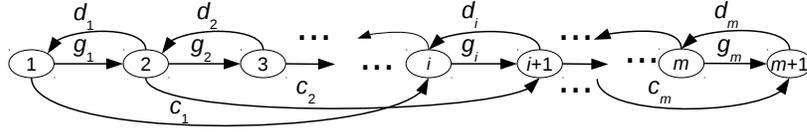


Figure 1: Graph G of Theorem 2.1

Model 2 represents our starting point and will be referred to as *nominal problem*. We now start investigating the case when uncertainty comes into play.

3. The robust optimization perspective

In the standard WFM process [15], the staffing levels \mathbf{b} , are computed in such a way that a desired Level of Service (LoS) is guaranteed. These represent the input to model (2) which returns a minimum cost shift schedule. Therefore, errors in estimating the demand (call volumes), as well as approximations in computing the staffing levels, affect the resulting LoS. The key idea of our modeling approach is to react to out-of-adherence at the scheduling stage by looking at the staffing levels as uncertain, that is, bounding into them multiple sources of randomness and approximations. One major advantage of this approach is that it does not require the explicit description of the LoS, which is typically complex and hard to manage computationally [4, 17, 18]. Conversely, we will show that this choice yields computationally tractable models. This perspective naturally leads to the application of robust optimization to shift scheduling. Notice that staffing levels correspond to right-hand-sides in model (2). Therefore, considering them uncertain gives rise to RHSU models.

In §3.1 and §3.2 we present a two-stage model for flexible shift-scheduling under uncertain staffing levels; in §3.3 we describe different uncertainty sets which are of interest for our application; finally in §3.4 we introduce a Benders-like reformulation.

3.1. The two-stage approach

Robust optimization methods can be classified into single-stage and multi-stage methods. In the former the solution is computed entirely before the realization of the uncertainty and the same solution is applied to any realization. In multi-stage optimization, the solution is computed in stages (generally two), that is: a part of the solution is computed before the realization of the uncertainty and a part is computed after. Both methods have advantages and disadvantages: single stage solutions are in general easier to compute, but they are often too conservative, whereas multi-stage solutions are less conservative (the solution can be adapted to the actual values), but the corresponding problem is harder to solve. RHSU single stage optimization reduces to solve a nominal problem with suitable right-hand-side values [22]. In our case, a single stage reformulation cannot even be applied, as the nominal problem is defined by equality constraints, which cause the robust single stage problem be infeasible citeBGGN04.

Interestingly, not only a two-stage formulation is the only option from the mathematical point of view, but it is also best-suited for the application. Indeed, it corresponds to a widespread call center practice, where the decision process is naturally characterized by a two-stage decomposition. In a first stage, typically at the beginning of the week, a daily shift schedule is computed. In a second stage (at the operational level), personnel reallocation between front and back office is implemented, according to the actual period by period needs. In this framework \mathbf{x} are first stage variables, whose value is computed in advance and kept fixed for any realization of the uncertainty, whereas \mathbf{o} and \mathbf{u} , which correspond to personnel reallocation, are second stage variables, whose value depends on the actual values of staffing levels \mathbf{b} . According to such process, we introduce the following two-stage robust version of model (2):

$$\min_{\mathbf{x} \in X} \left\{ \sum_{j \in J} c_j x_j + W_{\mathbf{x}}(U) \right\} \quad (3)$$

where X is the set of the feasible shift schedules, including the integrality of the \mathbf{x} variables; $W_{\mathbf{x}}(U)$ is the minimum reallocation cost for a fixed schedule \mathbf{x} (see §4) over the set of all possible realizations of the staffing levels, i.e., the uncertainty set U (see §3.3).

The solution of model (3) provides a double information: the value of first stage variables \mathbf{x} and an upper bound on the amount of personnel reallocation cost $W_{\mathbf{x}}(U)$. However, it neither provides the real personnel reallocation cost $R_{\mathbf{x}\mathbf{b}}$ nor the value of \mathbf{o} and \mathbf{u} variables, which must be recomputed ex-post, as they depend on both first stage variables \mathbf{x} and the actual realization of \mathbf{b} . We call this problem *personnel reallocation* problem.

3.2. The second stage decision: personnel reallocation

The personnel reallocation problem corresponds to model (2) with fixed \mathbf{x} , that is, the problem:

$$R_{\mathbf{x}\mathbf{b}} = \min \sum_{t \in T} (w_t^o o_t + w_t^u u_t) \quad (4)$$

$$u_t - o_t = b_t - \sum_{j \in J} a_{tj} x_j \quad t \in T$$

$$\mathbf{o}, \mathbf{u} \in \mathbb{Z}_+^m$$

Model (4) is still a min-cost flow problem. Moreover, the problem can be decomposed by time period, leading to the following result.

Lemma 3.1. *Personnel reallocation problem can be solved in $O(m)$ and the optimal solution is:*

$$o_t = \begin{cases} \sum_{j \in J} a_{tj} x_j - b_t & \text{if } b_t - \sum_{j \in J} a_{tj} x_j < 0 \\ 0 & \text{otherwise} \end{cases}$$

$$u_t = \begin{cases} b_t - \sum_{j \in J} a_{tj} x_j & \text{if } b_t - \sum_{j \in J} a_{tj} x_j \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

PROOF. Since \mathbf{x} are fixed, each time slot can be treated separately. The optimal solution must satisfy the necessary condition given by Lemma 2.3, therefore, for any $t \in T$, at most one between u_t and o_t can be positive. Since we have equality constraints, the only feasible solution having such property is the one above. \square

If \mathbf{x} is the optimal solution of (3), then the optimal value of (4) cannot exceed $W_{\mathbf{x}}(U)$. Since (4) is always feasible and bounded (for bounded \mathbf{b}), it can also be solved in its dual version.

$$R_{\mathbf{x}\mathbf{b}} = \max \sum_{t \in T} (b_t - \sum_{j \in J} a_{tj} x_j) y_t \quad (5)$$

$$-w_o^t \leq y_t \leq w_u^t \quad t \in T$$

where \mathbf{y} is the dual vector.

Lemma 3.2. *The optimal solution of (5) is*

$$y_t = \begin{cases} w_u^t & \text{if } b_t - \sum_{j \in J} a_{tj} x_j \geq 0 \\ -w_o^t & \text{otherwise} \end{cases}$$

PROOF. Let (\mathbf{o}, \mathbf{u}) be the optimal solution to (4) given in Lemma 3.1 and let \mathbf{y} the dual solution defined above. It is easy to see that $\mathbf{y}, \mathbf{o}, \mathbf{u}$ satisfy the complementarity conditions. In fact, when $u_t > 0$, that is, $b_t - \sum_{j \in J} a_{tj} x_j > 0$ then the corresponding dual constraint is satisfied with equality, i.e. $y_t = w_u^t$. In the same way, when $o_t > 0$, that is $b_t - \sum_{j \in J} a_{tj} x_j < 0$, dual constraint $y_t \geq -w_o^t$ is satisfied with equality. \square

Both Lemmas 3.1 and 3.2 make use of the condition given in Lemma 2.3. Indeed, it is possible to prove that such property, although only necessary for a solution to be optimal for (2), provides a necessary and sufficient condition of optimality for the personnel reallocation problem.

Theorem 3.3. *A feasible solution (\mathbf{o}, \mathbf{u}) is optimal for (4) iff, for any $t \in T$, o_t and u_t are not simultaneously positive.*

PROOF. The necessity follows from Lemma 2.3. Suppose now that (\mathbf{u}, \mathbf{o}) is a feasible solution where for any $t \in T$, either $o_t = 0$ or $u_t = 0$. The only solution having such property is the one computed in Lemma 3.1. By Lemma 3.2, given such (\mathbf{o}, \mathbf{u}) there exists a dual solution \mathbf{y} with the property that $\mathbf{o}, \mathbf{u}, \mathbf{y}$ satisfy the complementarity conditions and then (\mathbf{o}, \mathbf{u}) is optimal. \square

Being \mathbf{b} given, the personnel reallocation problem is easy, independently of the uncertainty set. However, the uncertainty set plays a fundamental role in solving robust problem (3), therefore in the next section we present the descriptions of the set U relevant to our application.

3.3. The uncertainty set

The uncertainty set U consists of all the realizations of the uncertain parameters, and, traditionally, is expressed by a polyhedron or by a convex set [6, 7]. Here, it includes all staffing levels that might reasonably occur. Our starting point is that the nominal staffing levels \tilde{b}_t are supposed to be affected by some deviations δ_t ranging in some interval $[-D_t, D_t]$. Advanced staffing procedures, besides nominal levels \tilde{b}_t , provide estimates for the gap between these levels and their worst case realization D_t , i.e. [13]. Two further observations contribute to the definition of uncertainty sets, summarized in the following property:

Property 3.1. *Properties of the uncertainty set:*

- (i) *deviations typically occur only in a limited number Γ of time periods, while in the others can be considered negligible;*
- (ii) *deviations at consecutive time periods are often not independent (see e.g. [5]).*

Condition (i) naturally leads to look at the well-known cardinality constrained approach [7]. Formally, let z_t be the percentage deviation in period t and let ζ_t indicate whether a deviation occurs in period t or not. The set U is:

$$U_{\Gamma} = \{ \mathbf{b} \in \mathbb{Z}^m : b_t = \tilde{b}_t + D_t z_t; \sum_{t \in T} \zeta_t \leq \Gamma, |z_t| \leq \zeta_t, \zeta_t \in \{0, 1\}; z_t \in \mathbb{R}, t \in T \}$$

In order to model condition (ii), we assume that the difference between the deviations of two consecutive periods is limited by a parameter $\Delta(t)$. In this case, we have the following form:

$$U_{\Delta} = \{ \mathbf{b} \in \mathbb{Z}^m : b_t = \tilde{b}_t + D_t z_t; |D_t z_t - D_{t-1} z_{t-1}| \leq \Delta(t); z_t \in [-1, 1], t \in T \}$$

When both conditions are enforced we obtain the set $U_{\Gamma\Delta}$, which is of great practical relevance:

$$U_{\Gamma\Delta} = \{ \mathbf{b} \in \mathbb{Z}_+^m : b_t = \tilde{b}_t + D_t z_t; \sum_{t \in T} \zeta_t \leq \Gamma, |z_t| \leq \zeta_t, \zeta_t \in \{0, 1\}; |D_t z_t - D_{t-1} z_{t-1}| \leq \Delta(t); z_t \in \mathbb{R}, t \in T \}$$

Notice that b_t is integer, as it represents the number of agents on duty at period t . The integrality of b_t , $t \in T$, has some consequences. First, \tilde{b}_t is integer in order to guarantee the feasibility of the nominal scenario, that is, the one with $\mathbf{b} = \tilde{\mathbf{b}}$. If so, $D_t z_t$ has to be integer as well. Therefore, w.l.o.g., we assume D_t to be integer. This also implies $\Delta(t)$ integer. Clearly, $U_{\Gamma\Delta}$ is bounded and non-empty.

3.4. Benders reformulation

We now illustrate a Benders like reformulation of problem (3) and discuss the related algorithmic issues. Problem (3) can be rewritten as:

$$\min \sum_{j \in J} c_j x_j + \lambda \tag{6}$$

$$\begin{aligned} \lambda &\geq W_{\mathbf{x}}(U) \\ \mathbf{x} &\in X \end{aligned}$$

where $\sum_{j \in J} c_j x_j$ represents the total cost of work shifts and $W_{\mathbf{x}}(U)$ the worst-case personnel reallocation cost:

$$W_{\mathbf{x}}(U) = \max_{\mathbf{b} \in U} \{R_{\mathbf{x}\mathbf{b}}\} \quad (7)$$

Using the expression (5) for $R_{\mathbf{x}\mathbf{b}}$, $W_{\mathbf{x}}(U)$ is computed as:

$$\begin{aligned} W_{\mathbf{x}}(U) &= \max \sum_{t \in T} (b_t - \sum_{j \in J} a_{tj} x_j) y_t \\ &\quad - w_o^t \leq y_t \leq w_u^t \quad t \in T \\ &\quad b \in U \end{aligned} \quad (8)$$

Therefore, problem (6) becomes:

$$\begin{aligned} &\min \sum_{j \in J} c_j x_j + \lambda \\ &\lambda \geq \sum_{t \in T} (b_t - \sum_{j \in J} a_{tj} x_j) y_t \quad \mathbf{b} \in U, \mathbf{y} \in Y \\ &\mathbf{x} \in X \end{aligned} \quad (9)$$

where $Y = \{\mathbf{y} : -w_o^t \leq y_t \leq w_u^t, t \in T\}$. This formulation is non-compact, as it may have an exponential number of constraints. A standard algorithmic framework for such formulations is the Kelley cutting plane method, originally introduced in [14]. This method starts with a relaxed formulation including a suitable subset of constraints (9). Then, additional constraints are dynamically generated by a *separation oracle*. Let $(\bar{\mathbf{x}}, \bar{\lambda})$ be a solution to the current problem. The *Separation Problem* (SEP) consists of finding a realization $\bar{\mathbf{b}}$ and a vector $\bar{\mathbf{y}} \in Y$ such that $\bar{\lambda} < (\bar{b}_t - \sum_{j \in J} a_{tj} \bar{x}_j) \bar{y}_t$ or prove that none exist. In the former case, the corresponding (violated) inequality (9) is added to the formulation. In general, separation problems arising from two-stage RHSU models are typically nonconvex and strongly NP-hard, as shown in [23]. In §4 we investigate the theoretical complexity of SEP.

4. The separation problem

Let us discuss the complexity of SEP, starting from U_{Γ} . Similarly to the personnel reallocation problem, in SEP understaffing and overstaffing situations can be addressed one period at a time. This has a remarkable consequence when set U_{Γ} is used to model uncertainty. In fact, the worst-case realization for a given \mathbf{x} can be easily identified:

Theorem 4.1. [27] *Given \mathbf{x} , the corresponding worst case realization \mathbf{b} is:*

$$b_t = \begin{cases} \tilde{b}_t + D_t & \text{if } t \in I \text{ and } (\tilde{b}_t + D_t - \sum_{j \in J} a_{tj} x_j) w_t^u \geq (\sum_{j \in J} a_{tj} x_j - \tilde{b}_t) w_t^o \\ \tilde{b}_t - D_t & \text{if } t \in I \text{ and } (\tilde{b}_t + D_t - \sum_{j \in J} a_{tj} x_j) w_t^u < (\sum_{j \in J} a_{tj} x_j - \tilde{b}_t) w_t^o \\ \tilde{b}_t & \text{if } t \notin I \end{cases}$$

where set $I \subseteq T$ includes the first Γ time periods according to non decreasing values of τ defined below.

$$\tau_t = \max \left\{ \left(\tilde{b}_t + D_t - \sum_{j \in J} a_{tj} x_j \right) w_t^u, \left(\sum_{j \in J} a_{tj} x_j - \tilde{b}_t - D_t \right) w_t^o \right\} \\ - \max \left\{ \left(\tilde{b}_t - \sum_{j \in J} a_{tj} x_j \right) w_t^u, \left(\sum_{j \in J} a_{tj} x_j - \tilde{b}_t \right) w_t^o \right\} \quad t \in T$$

Based on the result in [27], the complexity of the second stage problem for U_Γ is:

Corollary 4.2. *SEP for U_Γ can be solved in $O(m \log m)$.*

PROOF. Given \mathbf{x} , compute worst-case realization \mathbf{b} as in Theorem 4.1. Vectors \mathbf{o} , \mathbf{u} and \mathbf{y} can be obtained solving the personnel reallocation problem for the given \mathbf{x} and \mathbf{b} , which can be done in $O(m)$ according to Lemmas 3.1 and 3.2. Therefore, the complexity reduces to the one of ordering the τ values, that is, $O(m \log m)$. \square

When U_Δ or $U_{\Gamma\Delta}$ are considered, it is no longer possible to use such a method, as those sets include correlation between the deviations of consecutive time periods. However, in this case, SEP can be solved in pseudo-polynomial time by reducing it to computing paths on a suitable Δ -uncertainty graph $G(V, E)$. Let K_t be the set of integers in $[-D_t, D_t]$, $t \in T$, the graph can be formally defined as follows (see Figure 2).

Definition 4.3. Δ -uncertainty graph $G(V, E)$ is defined as:

$$V = \{\sigma, \tau\} \cup \{v_{tk}, t \in T, k \in K_t\};$$

$$E = E_\sigma \cup E_\tau \cup E_\Delta, \text{ where:}$$

$$E_\sigma = \{(\sigma, v_{1k}) \text{ for each } k \in K_1\};$$

$$E_\tau = \{(v_{mk}, \tau) \text{ for each } k \in K_m\};$$

$$E_\Delta = \{(v_{tp}, v_{(t+1)q}) \text{, for each } p \in K_t, q \in K_{t+1}, t \in T \setminus \{m\} \text{ and } |p - q| \leq \Delta(q)\}.$$

Note that $G(V, E)$ is acyclic, $|V| = \sum_{t \in T} |K_t| + 2$ and $|E| = \sum_{t=1}^{m-1} |K_t| |K_{t+1}| + |K_m| + |K_1|$. Furthermore, $|K_t| = 2D_t$. If we let $\bar{D} = \max_{t \in T} D_t$, then $|V| = O(\bar{D}m)$ and $|E| = O(\bar{D}^2m)$. We show that solving the problem with U_Δ reduces to computing a longest-path in $G(V, E)$, which can be done in pseudo-polynomial time. For any $t \in T$ and $k \in K_t$, let c_{tk} be the minimum reallocation cost for time slot t when $b_t = \tilde{b}_t + k$ (see §3.2).

Theorem 4.4. *SEP for U_Δ can be solved in $O(\bar{D}^2m)$ time.*

PROOF. Let us define arc weights as: $c_e = 0$, $e \in E_\sigma$, $c_e = c_{mk}$ for $e \in E_\tau$, $c_e = c_{tp}$ for $e \in E_\Delta$. By construction, SEP amounts to computing a longest $\sigma - \tau$ path on $G(V, E)$. Since G is acyclic, the problem can be solved in $O(|E|) = O(\bar{D}^2m)$ time. \square

The Δ -uncertainty graph also allows us to prove that the separation problem for $U_{\Gamma\Delta}$ can still be solved in pseudo-polynomial time, as it amounts to solve a resource constrained shortest-path problem on G . If we let $\bar{w} = \max_{t \in T} \max\{w_t^o, w_t^u\}$ and $\Omega = \max_{t \in T} \max\{o_t, u_t\} = \max_{t \in T} (\tilde{b}_t + D_t)$, the problem complexity is given in the theorem below.

Theorem 4.5. *SEP for $U_{\Gamma\Delta}$ can be solved in $O(\bar{w}\Omega\bar{D}^3\Gamma m^2)$.*

PROOF. Let us associate to the arcs edge weights c_e defined as: $c_e = 0$, $e \in E_\sigma$, $c_e = c_{mk}$ for $e \in E_\tau$, $c_e = c_{tp}$ for $e \in E_\Delta$. Moreover, let us define for each arc a further binary weight r_e as: $r_e = 1$ for $e = (v_{tk}, j) \in E_\Delta \cup E_\tau$ such that node v_{tk} corresponds to a deviation from the nominal value, $r_e = 0$ otherwise. Hence, $r_e = 0$ if $e = (\sigma, j)$ or $e = (v_{tk}, j)$, with $v_{tk} = 0$ and $r_e = 1$ otherwise. Since we have a bound Γ on number of possible deviations, any feasible path must also satisfy the additional requirement $\sum_{e \in E} r_e \leq \Gamma$. Therefore, SEP reduces to a (single-)resource constrained longest path problem. In [21] it is shown that can the resource constrained shortest path problem with a single resource can be solved by $O(\log(|V|RC))$ shortest path computations, where C is the maximum arc cost and R the maximum weight. In our case, $|V| = O(\bar{D}m)$, $R = O(\Gamma)$, $C = O(\bar{w}\Omega)$ and the longest path problem can be solved in $O(\bar{D}^2m)$. The overall complexity is then $O(\bar{w}\Omega\bar{D}^3\Gamma m^2)$. \square

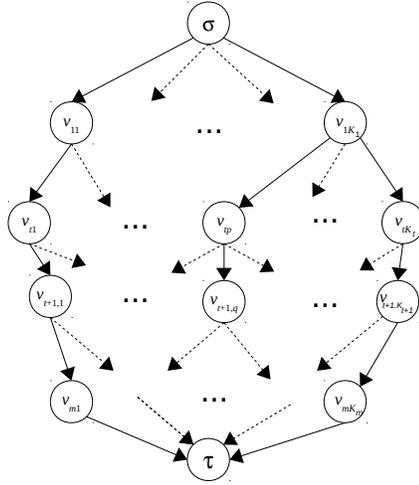


Figure 2: The Δ -uncertainty graph

These results have implications on the complexity of the continuous version of problem (3), where the integrality requirement on variables is relaxed. In detail, the following corollary holds:

Corollary 4.6. *The LP relaxation of (3) is solvable in:*

1. *strongly polynomial time for U_Γ ;*
2. *pseudo-polynomial time for U_Δ and $U_{\Gamma\Delta}$.*

PROOF. The equivalence between optimization and separation [12] implies that the LP robust problem has the same worst case complexity as SEP. Therefore, it is solvable in polynomial time for U_Γ thanks to Lemma 4.2 and pseudo-polynomial for U_Δ and $U_{\Gamma\Delta}$ thanks to Lemmas 4.4 and 4.5. \square

Interestingly, similar results do not hold for several well-structured LP-RHSU problems [23], which turn often out to be NP-hard in the strong sense. The relevance of this result is also confirmed by the practical evidence that rather standard rounding techniques often allow to compute good quality feasible solutions [21].

5. Computational experience

In this section we investigate the application of the robust optimization approach in practical settings. We first give the details of our implementation and describe the test-bed. Then, the algorithm performance is discussed. Finally, we analyze from managers' perspective the advantage of the robust methodology. The experiments are run on 2 Intel Xeon 5150 processors clocked at 2.6 GHz with 8 GB of RAM in 4-thread mode. The commercial framework IBM Cplex 12.6. is used to implement a branch-and-cut algorithm in which we integrated our primal heuristic and separation routine. Computations are stopped either by 1 hour time limit or 0.05% optimality tolerance. A preliminary experience showed that the best performance is obtained by Cplex default settings with cutting plane generation (for all families of cuts) and MIP heuristics turned off. Separation is performed on all integer solutions, while fractional ones are tested only at the root node.

5.1. Branch-and-cut details

The primal heuristic consists of rounding the current LP solution and computing the associated worst case uncertainty cost $W_{\mathbf{x}}(U)$, again by solving SEP. A time limit of 50 seconds is imposed to the heuristic: if an optimal solution has been obtained, then solution $(\chi, W_{\chi}(U))$ is returned, otherwise the heuristic fails.

We tested different algorithms for SEP. We first implemented an algorithm based on *label-setting* shortest path computations (see theorems 4.4, 4.2, 4.5). We experienced that the performance of these algorithms suffer, in some cases, from the size of the graph which depends on the deviations and the number of time periods. However, a more robust behavior has been observed by solving SEP through the following MIP reformulation.

$$\begin{aligned}
 \max \quad & \sum_{t \in T} (w_t^o o_t + w_t^u u_t) & (10) \\
 u_t - o_t = & b_t - \sum_{j \in J} a_{tj} x_j & t \in T \\
 o_t \leq & M \alpha_t & t \in T \\
 u_t \leq & M(1 - \alpha_t) & t \in T \\
 \mathbf{b} \in U, \mathbf{o}, \mathbf{u} \geq & 0, \alpha \in \{0, 1\}^{|T|}
 \end{aligned}$$

This problem is equivalent to computing $W_{\mathbf{x}}(U)$ using expression (4). In fact, variables α impose that, for any time slot $t \in T$, either $u_t = 0$ or $o_t = 0$, which is, by Theorem 3.3, a necessary and sufficient condition for (\mathbf{o}, \mathbf{u}) to be optimal for the personnel reallocation problem. Then, the results that follow are those obtained using the MIP-based separation procedure.

5.2. Test-bed

The instances are based on real data, gathered during year 2008 from a large, distributed call center of an Italian Public Agency receiving more than 1800 daily calls. The call center is on duty on working days from 7:45 a.m. to 8:00 p.m. corresponding to 49 time slots (15 minutes). Three labor contracts are used, with 4, 6 or 8 hour shifts. The costs ascribed to these shifts are 72, 96 and 112 € respectively. Agents are skilled for both front and back office and job flexibility

is implemented by call center managers, who dynamically allocate agents to different activities according to operational needs. However, personnel reallocation requires some organizational set-up and should be limited. Inbound calls are prioritized over back office. Therefore, (front office) understaffing is considered more critical than overstaffing and $c_j/s_j < w_t^o < w_t^u$, $t \in T$, where c_j is the cost of the most expensive shift, s_j the number of its time periods. Here, we considered as fair values $w_t^u = 10$ and $w_t^o = 5$ € for every $t \in T$.

Staffing levels of a reference day were provided us by the managers. Starting from this data, we generated 60 instances (36 for $U_{\Gamma\Delta}$, 18 for U_{Γ} and 6 for U_{Δ}) by systematically varying the percentage deviation $dev\%$ of the actual staffing level with respect to the nominal value, the number Γ of time slots affected by uncertainty and the allowed deviation difference Δ between two consecutive time periods, supposing it the same for every pair. We consider the values $dev\% \in \{5, 10, 20\}$, $\Gamma \in \{5, 10, 15, 20, 25, 30\}$ and $\Delta \in \{0.5\theta, \theta\}$, where θ is the average difference between two consecutive nominal staffing levels. We measured $\theta = 45$, yielding $\Delta \in \{22, 45\}$. The variation of the $dev\%$ parameter models the confidence of the managers in the nominal values, ranging from reliable estimates (e.g. days with standard demand patterns) to weak confidence caused by critical situations, such as strikes. Parameter Γ controls the trade-off between the robustness and the corresponding cost. Again, we sample a wide range of values. Parameter Δ captures the correlation between the variations of consecutive time periods, intrinsic in the call center dynamic.

5.3. Branch-and-cut performance

Traditionally, Benders reformulations may suffer from the weakness of the LP relaxation along with numerical difficulties due to nasty coefficients. In fact, specialized techniques have been recently proposed to overcome this drawback [8]. Interestingly, our formulation turns out to be not significantly affected by such problems. In tables 1, 2 and 3 the branch-and-cut statistics are reported for U_{Γ} , U_{Δ} and $U_{\Gamma\Delta}$ respectively. Besides the instance parameters, the tables contain:

- the objective value,
- the value found by the primal heuristic at the root node,
- the value of the LP relaxation and the percentage gap before enumeration,
- the number of B&B nodes,
- the total CPU time,
- the CPU time required by the separation algorithm,
- the number of generated cuts and
- the CPU time spent by the rounding heuristic (all times are expressed in seconds).

From the tables, one can observe that all the instances are solved in reasonable CPU time and limited number of branch-and-bound nodes (only in one case the latter gets over 1,000 and no significant differences have been observed among the different uncertainty sets from this point of view).

Two major evidences explain such a nice behavior. The first deals with the quality of the LP relaxation. In fact, the gap between the optimal LP value and the value of the solution returned

by the primal heuristic at the root node turns out to be very often quite small, never exceeding 1%. This is particularly valuable, as the results of §4 show that solving the LP relaxation is computationally accessible.

The second evidence concerns the cuts. In our case:

- integrality of b_t, w_t^u, w_t^o implies that all coefficients involved in inequalities (9) turn out to be integer (see Lemma 3.2),
- figures arising from real-world instances (m, b_t, w_t^u, w_t^o) always produce cuts with limited *coefficient dynamism* [19].

This mitigates known numerical difficulties of Benders reformulation and the number of generated cuts remains small, guaranteeing a good convergence of the algorithm. Looking at the effect of instance parameters, we observe that CPU times tend to increase as *dev%* gets larger, while it is not significantly affected by Γ . Conversely, times increase as Δ decreases. Thanks to this nice computational behavior, the economical impact of robustness can be evaluated in a real-world setting. This is illustrated in the next subsection.

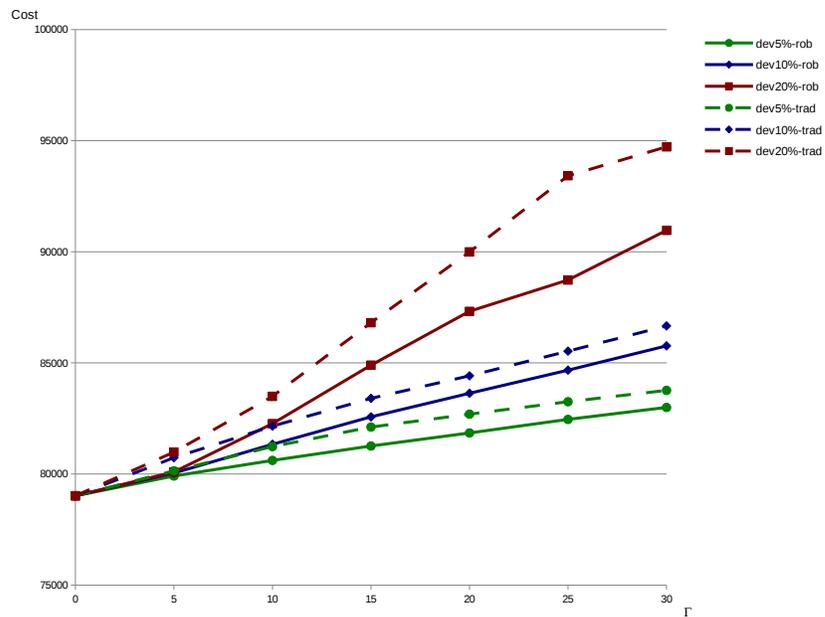
	Γ	Obj. value	Root primal	Root LP rel.	Root % gap	B&B nodes	Total time	Sep. time	# cuts	Heur. time
<i>dev%</i> = 5	5	80021	80068	80013.33	0.07	73	6	5	111	1
	10	80742	80742	80732.33	0.01	0	5	5	126	< 1
	15	81393	81575	81389.00	0.23	49	6	6	147	< 1
	20	82017	82122	82014.00	0.13	25	7	7	158	< 1
	25	82628	82690	82613.00	0.09	73	7	6	142	< 1
	30	83174	83251	83173.33	0.09	30	8	5	151	3
<i>dev%</i> = 10	5	80984	80984	80976.00	0.01	0	6	6	126	< 1
	10	82419	82511	82406.00	0.13	2	10	9	149	1
	15	83709	83771	83702.00	0.08	153	15	14	165	1
	20	84962	84994	84950.33	0.05	71	9	9	131	< 1
	25	86151	86222	86137.67	0.10	106	19	18	166	1
	30	87286	87286	87243.67	0.05	0	9	9	153	< 1
<i>dev%</i> = 20	5	82953	82953	82930.00	0.03	0	10	9	101	1
	10	85786	85876	85778.00	0.11	38	18	17	120	1
	15	88350	88422	88347.67	0.08	42	21	20	145	1
	20	90835	90895	90834.33	0.07	324	22	20	128	2
	25	93218	93298	93204.00	0.10	401	23	22	141	1

Table 1: Branch-and-cut statistics for U_Γ

5.4. Economical analysis

The trade-off between level of protection and personnel cost is investigated. We consider both the robust model presented in §3 and the traditional approach. The latter consists in assigning agents to shifts according to the nominal levels and then adjusting the solution depending on the actual realization. The personnel cost has two components: the cost of the shifts, associated with the \mathbf{x} variables, and the cost of flexibility, associated with \mathbf{o} and \mathbf{u} variables. When the staffing levels are subject to uncertainty, the cost of flexibility depends on the realization and

		Obj. value	Root primal	Root LP rel.	Root % gap	B&B nodes	Total time	Sep. time	# cuts	Heur. time
$dev\% = 5$	$\Delta = 22$	83961	84008	83946.00	0.07	1031	10	10	146	< 1
	$\Delta = 45$	84438	84487	84427.67	0.07	40	7	7	143	< 1
$dev\% = 10$	$\Delta = 22$	88237	88294	88219.93	0.08	163	60	60	121	< 1
	$\Delta = 45$	88924	88924	88901.33	0.03	0	15	15	126	< 1
$dev\% = 15$	$\Delta = 22$	97340	97450	97311.75	0.14	248	29	27	66	2
	$\Delta = 45$	98128	98595	98113.30	0.49	97	22	22	73	< 1

Table 2: Branch-and-cut statistics for U_Δ Figure 3: $dev\%$ values when $\Delta = 22$

	Γ	Obj. value	Root primal	Root LP rel.	Root % gap	B&B nodes	Total time	Sep. time	# cuts	Heur. time	
$dev\% = 5$	$\Delta = 22$	5	79909	79967	79902.67	0.08	485	64	63	135	< 1
		10	80611	80662	80586.00	0.09	2	27	25	149	2
		15	81258	81258	81240.33	0.02	0	31	31	151	< 1
		20	81847	81847	81832.33	0.02	0	20	20	126	< 1
		25	82456	82456	82416.11	0.05	0	17	15	126	2
		30	82995	82995	82960.33	0.04	0	19	18	151	1
	$\Delta = 45$	5	80043	80128	80010.67	0.15	770	12	12	139	< 1
		10	80751	80751	80714.17	0.05	0	6	5	127	< 1
		15	81374	81434	81360.00	0.09	113	16	16	132	< 1
		20	81992	82021	81974.67	0.06	137	16	15	144	1
		25	82577	82611	82548.00	0.08	390	16	16	116	< 1
		30	83129	83129	83087.67	0.05	0	9	9	139	< 1
$dev\% = 10$	$\Delta = 22$	5	80043	80043	80018.83	0.03	0	65	62	125	3
		10	81336	81354	81303.83	0.06	2	171	160	199	11
		15	82572	82572	82530.96	0.05	0	133	118	151	15
		20	83634	83697	83612.36	0.10	129	108	102	114	6
		25	84674	84941	84656.64	0.34	109	131	126	121	5
		30	85766	85891	85764.67	0.15	63	144	136	140	8
	$\Delta = 45$	5	80813	80813	80805.33	0.01	0	10	8	100	2
		10	82163	82261	82159.67	0.12	58	29	28	109	1
		15	83473	83473	83453.00	0.02	0	40	38	126	2
		20	84663	84663	84639.53	0.03	0	41	39	126	2
		25	85846	85884	85815.67	0.08	545	220	217	146	1
		30	86909	86909	86887.33	0.02	0	54	50	176	4
$dev\% = 20$	$\Delta = 22$	5	80091	80134	80081.33	0.07	231	576	564	119	12
		10	82267	82397	82264.67	0.16	39	804	763	150	41
		15	84899	85195	84897.64	0.35	39	301	287	105	14
		20	87319	87537	87284.82	0.29	813	1420	1409	127	11
		25	88730	88779	88692.67	0.10	121	1254	1222	99	32
		30	90962	91035	90944.59	0.10	60	271	249	89	22
	$\Delta = 45$	5	81137	81364	81127.57	0.29	94	111	108	123	3
		10	83780	83844	83754.67	0.11	120	255	248	122	7
		15	86399	86435	86387.54	0.05	2	136	125	124	11
		20	88584	88772	88567.67	0.23	75	106	101	93	5
		25	90765	91242	90727.75	0.57	2	100	94	99	6
		30	93052	93261	93010.78	0.27	359	171	165	125	6

Table 3: Branch-and-cut statistics for $U_{\Gamma\Delta}$

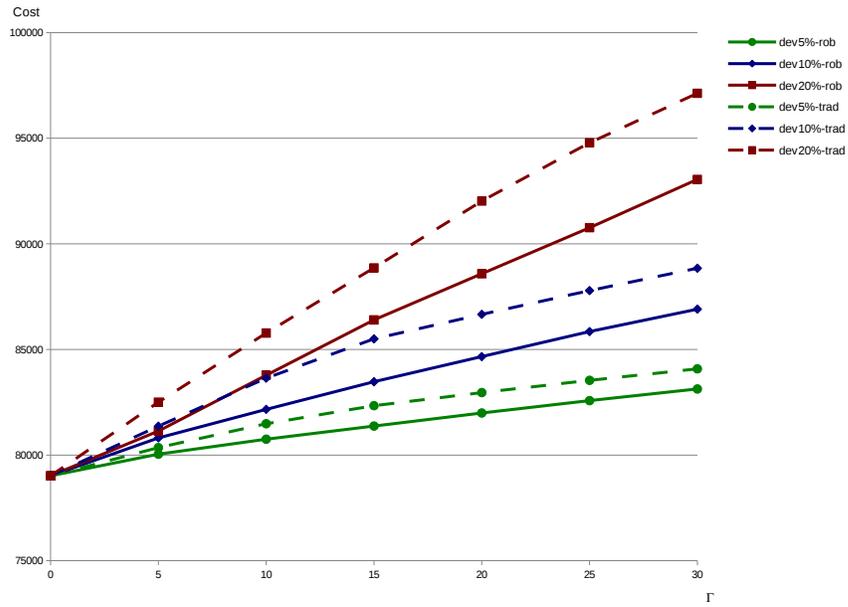


Figure 4: *dev%* values when $\Delta = 45$

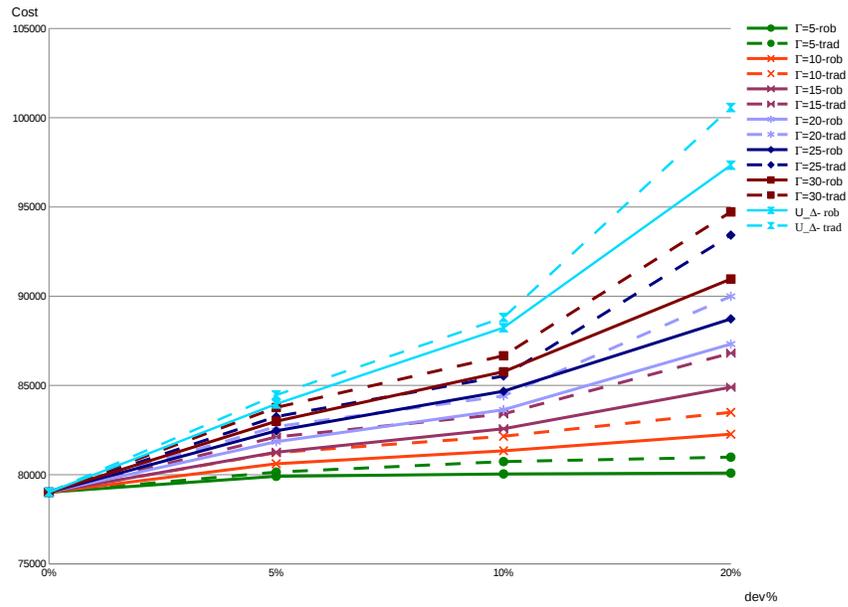
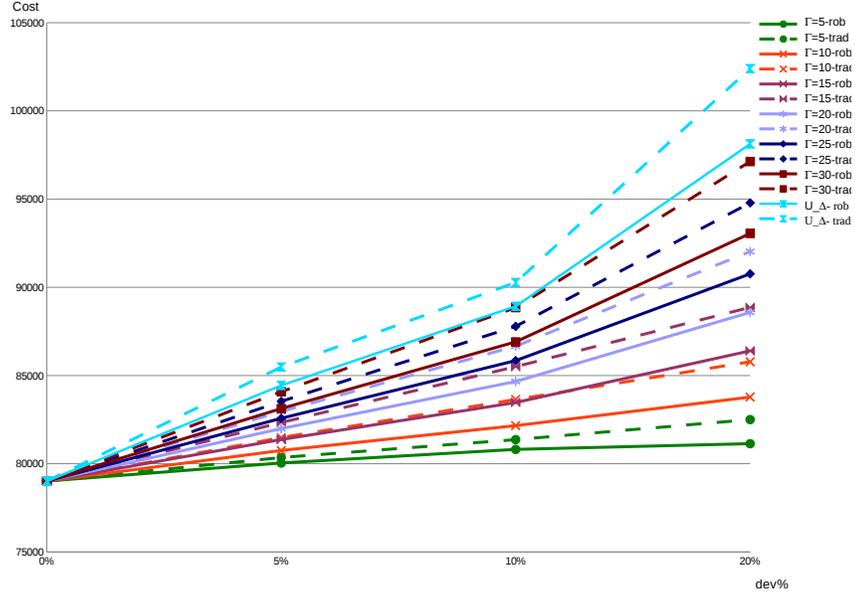


Figure 5: Γ values when $\Delta = 22$

Figure 6: Γ values when $\Delta = 45$

its worst case value $W_{\mathbf{x}}(U)$ must be considered in order to be protected against all realizations. In the robust method, both of the cost components are included in the solution to problem (9), since $W_{\mathbf{x}^*}(U) = \lambda^*$. For the traditional approach, \mathbf{x} values along with the associated shifts cost come from the solution of model (2). The cost of flexibility for the worst case realization is computed ex-post by solving problem (10) for the given \mathbf{x} .

In figures 3 and 4 the overall cost is reported as a function of Γ ($\Gamma = 0$ corresponds to nominal problem (2)). Each figure corresponds to one Δ value and contains six functions: three representing the robust cost and three representing the cost of the traditional approach for the different $dev\%$ values. In the same way, figures 5 and 6 show the costs as a function of $dev\%$ with lines corresponding to Γ values ($dev\% = 0$ is again the nominal problem). The graphs give evidence to the fact that the reallocation cost in the traditional approach (dashed lines) grows significantly with the deviations. This issue is often underestimated as the urgency of covering understaffing situations at the front office is perceived more critical at real-time level. Here, the remarkable fact is that a very high level of protection is accomplished by increasing the cost with respect to the nominal value by less than 10% and a safe protection is achieved even restricting to a 5% additional budget. This extra budget is considered acceptable or even profitable. In fact, the figures highlight that the traditional approach can be by far more expensive than the corresponding robust method (solid lines).

6. Conclusions

We investigated a two-stage robust optimization model for shift scheduling in flexible call centers. The distinguishing feature of our model is that it conveys all uncertainty sources into random staffing levels. This gives rise to a two-stage robust model with right-hand-side uncertainty. In contrast to the general case, we showed that the separation problem associated to the constraints of a Benders-like reformulation can be solved rather efficiently for practically

relevant uncertainty set formulations. This opens the way to effective algorithms which may impact on practice.

References

- [1] <http://www.nextstepwfp.com/2012/01/how-much-is-over-staffing-costing-your-contact-centre>.
- [2] R. Ahuja, T. Magnanti, and J. Orlin, *Network Flows: Theory, Algorithms and Applications*. Prentice Hall, 1993.
- [3] Z. Aksin, M. Armony, and V. Mehrotra, “The modern call center: A multi-disciplinary perspective on operations management research,” *Production and Operations Management*, vol. 16, no. 6, pp. 665–688, 2007.
- [4] J. Atlason, M. Epelman, and S. Henderson, “Optimizing call center staffing using simulation and analytic center cutting-plane methods,” *Management Science*, vol. 54, no. 2, pp. 295–309, 2008.
- [5] A. Avramidis, A. Deslauriers, and P. P. L’Ecuyer, “Modeling daily arrivals to a telephone call center,” *Manage. Sci.*, vol. 50, no. 7, pp. 896–908, 2004.
- [6] A. Ben-Tal and A. Nemirovski, “Robust convex optimization,” *Math. Oper. Res.*, vol. 23, pp. 769–805, 1998.
- [7] D. Bertsimas and M. Sim, “The price of robustness,” *Oper. Res.*, vol. 52, no. 1, pp. 35–53, 2004.
- [8] M. Bodur and J. Luedtke, “Mixed-integer rounding enhanced benders decomposition for multiclass service system staffing and scheduling with arrival rate uncertainty,” *Optimization Online*, 2014.
- [9] T. Cezik and P. L’Ecuyer, “Staffing multiskill call centers via linear programming and simulation,” *Manage. Sci.*, vol. 54, no. 2, pp. 310–323, 2008.
- [10] N. Gans, G. Koole, and A. Mandelbaum, “Telephone call centers: Tutorial, review and research prospects,” *Manufacturing and Service Operations Management*, vol. 5, no. 2, pp. 79–141, 2003.
- [11] L. Green, P. Kolesar, and J. Soares, “Improving the sipp approach for staffing service systems that have cyclic demands,” *Operations Research*, vol. 49, no. 4, pp. 549–564, 2001.
- [12] M. Grötschel, L. Lovász, and A. Schrijver, “The ellipsoid method and its consequences in combinatorial optimization,” *Combinatorica*, vol. 1, no. 2, pp. 169–197, 1981.
- [13] I. Gurvich, J. Luedtke, and T. Tezcan, “Staffing call centers with uncertain demand forecasts: A chance-constrained optimization approach,” *Management Science*, vol. 56, no. 7, pp. 1093–1115, 2010.
- [14] J. Kelley, “The cutting-plane method for solving convex programs,” *Journal of the Society for Industrial and Applied Mathematics*, vol. 8, no. 4, pp. 703–712, 1960.

- [15] G. Koole, *Call center mathematics. A scientific method for understanding and improving contact centers*. MG books, 2013.
- [16] G. Koole and E. van der Sluis, “Optimal shift scheduling with a global service level constraint,” *IIE Transactions (Institute of Industrial Engineers)*, vol. 35, no. 11, pp. 1049–1055, 2003.
- [17] S. Liao, G. Koole, C. van Delft, and O. Jouini, “Staffing a call center with uncertain non-stationary arrival rate and flexibility,” *OR Spectrum*, vol. 34, no. 3, pp. 691–721, 2012.
- [18] S. Liao, C. van Delft, and J.-P. Vial, “Distributionally robust workforce scheduling in call centres with uncertain arrival rates,” *Optimization Methods and Software*, vol. 28, no. 3, pp. 501–522, 2013.
- [19] F. Margot, “Testing cut generators for mixed-integer linear programming,” *Math. Prog. Comp.*, vol. 1, no. 1, pp. 69–95, 2009.
- [20] S. Mattia, F. Rossi, M. Servilio, and S. Smriglio, “Robust shift scheduling in call centers,” in *Combinatorial Optimization - Third International Symposium, ISCO 2014, Lisbon, Portugal, March 5-7, 2014*, pp. 336–346, 2014.
- [21] K. Mehlhorn and M. Ziegelmann, “Resource constrained shortest paths,” in *Algorithms - ESA 2000* (M. Paterson, ed.), vol. 1879 of *LNCS*, pp. 326–337, 2000.
- [22] M. Minoux, “Robust linear programming with right-hand-side uncertainty, duality and applications,” in *Encyclopedia of Optimization* (C. Floudas and P. Pardalos, eds.), pp. 3317–3327, Springer US, 2009.
- [23] M. Minoux, “On 2-stage robust lp with rhs uncertainty: complexity results and applications,” *J. Global Optim.*, vol. 49, no. 3, pp. 521–537, 2011.
- [24] J. Orlin, “A faster strongly polynomial algorithm for the minimum cost flow problem,” *Oper. Res.*, vol. 41, pp. 338–350, 1993.
- [25] M. Segal, “The operator-scheduling problem: A network-flow approach,” *Operations Research*, vol. 22, no. 4, pp. 808–823, 1974.
- [26] S. Steckley, S. Henderson, and V. Mehrotra, “Forecast errors in service systems,” *Probability in the Engineering and Informational Sciences*, vol. 23, no. 2, pp. 305–332, 2009.
- [27] A. Thiele, T. Terry, and M. Epelman, “Robust linear optimization with recourse,” *Rapport technique*, pp. 4–37, 2009.
- [28] R. Thomas and T. Harrison, “A stochastic programming model for scheduling call centers with global service level agreements,” *European Journal of Operational Research*, vol. 207, no. 3, pp. 1608–1619, 2010.
- [29] W. Whitt, “Staffing a call center with uncertain arrival rate and absenteeism,” *Production and Operations Management*, vol. 15, no. 1, pp. 88–102, 2006.