Tactical workforce sizing and scheduling decisions for last-mile delivery

Minakshi Punam Mandal mandal@essec.edu[1], Alberto Santini alberto.santini@upf.edu[2,3], and Claudia Archetti archetti@essec.edu[1]

[1] Department of Information Systems, Decision Sciences and Statistics, ESSEC Business School, 3 Avenue Bernard Hirsch, 95021, Cergy, France
[2] Department of Economics and Business, Universitat Pompeu Fabra, Carrer Ramon Trias-Fargas, 25–27, 08005, Barcelona, Spain
[3] Data Science Centre, Barcelona School of Economics, Carrer Ramon Trias-Fargas, 25–27, 08005, Barcelona, Spain

October 16, 2024

Abstract

We tackle the problems of workforce sizing and shift scheduling of a logistic operator delivering parcels in the last-mile segment of the supply chain. Our working hypothesis is that the relevant decisions are affected by two main trade-offs: workforce size and shift stability. A large workforce can deal with demand fluctuations but incurs higher fixed costs; by contrast, a small workforce might require excessive outsourcing to third-party logistic providers. Stable shifts, i.e., with predictable start times and lengths, improve worker satisfaction and reduce turnover; at the same time, they might be less able to adapt to an unsteady demand. We test these assumptions through an extensive computational campaign based on a novel mathematical formulation. We find that extreme shift stability is, indeed, unsuitable for last-mile operations. At the same time, introducing a very limited amount of flexibility achieves similar effects as moving to a completely flexible system while ensuring a better work-life balance for the workers. Several recent studies in the social sciences have warned about the consequences of precarious working conditions for couriers and retail workers and have recommended—among other things—stable work schedules. Our work shows that it is possible to offer better working conditions in terms of shift stability without sacrificing the company's bottom line. Thus, companies prioritising profitability (as is often the case) can improve workers' well-being and increase retention with a negligible cost impact.

Keywords: OR in service industries; last-mile delivery; workforce scheduling; workforce sizing; shift stability.

1 Introduction

Last-mile delivery (LMD) is the final segment of the supply chain, starting at the last warehouse and ending when the goods reach the customer. With the boom of e-commerce, especially during and after the Covid-19 pandemic, LMD in large cities is dominated by home deliveries, i.e., by carriers delivering many small parcels up to the customers' doorsteps using a fleet of vehicles from legacy vans to more sustainable means such as cargo bikes (Alfonso et al. 2021). A fundamental tactical question arises for logistic operators involved in LMD in the urban environment: how many couriers should they employ? On the one hand, a more extensive workforce is associated with higher staffing costs; on the other hand, using fewer couriers degrades the quality of service or forces the operator to resort to expensive outsourcing options. Demand for home delivery

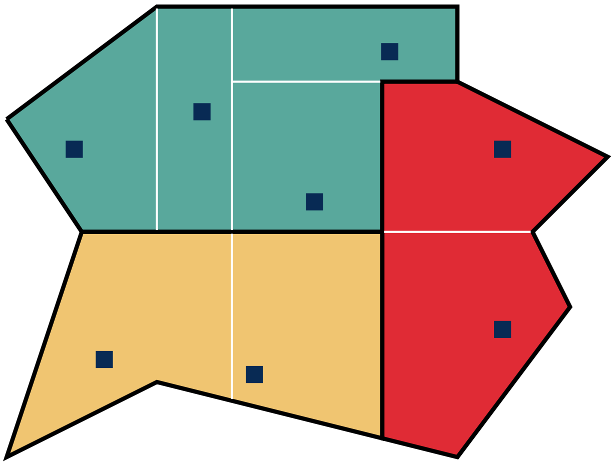

Figure 1: Example of a city with three regions (delimited with thicker black lines) subdivided into smaller areas. Blue squares indicate the position of the satellites.

is highly seasonal (throughout the year and at specific hours of the day), further complicating the challenge of choosing the correct workforce size. This paper introduces a decision support system for tactical hiring decisions, incorporating realistic constraints and demand uncertainty.

The importance of increasing the efficiency of LMD stems from its relevance in the global economy. For example, LMD is expected to grow at a compound annual rate of 6.12% from 2023 to 2030 (Contrive Datum Insights 2023). Among the optimisation problems linked with LMD, Boysen, Fedtke, and Schwerdfeger (2021) identified staffing and fleet sizing as needing attention from the operational research community because of the "lack [of] scientific decision support".

We fill this gap by considering the problem of a logistics operator who must deliver parcels throughout the day and faces the tactical problem of sizing its workforce. The operator can decide to fulfil each delivery with either a fleet of owned vehicles driven by couriers or paying a fee to an outsourcing (or crowdsourcing) provider. Maintaining a large workforce would allow the operator to avoid paying such fees at the price of high fixed staffing costs. Conversely, hiring few couriers means the operator must extensively resort to outsourcing, leading to high variable costs. The logistic operator must then balance the tactical workforce sizing and operational outsourcing decisions. From this point of view, our work contributes to a recent research stream about workforce sizing in the logistics and service industries (see, e.g., Section 2.2 and (Dai and P. Liu 2020a; Turan et al. 2022; Pandey, Gajjar, and Shah 2021)).

A central concept in our setting is that of *satellites* (Crainic et al. 2021). These are locations within the city where the couriers start and end their delivery trips. They are intermediate between large distribution centres (usually on the city's outskirts) and the customers. They are used for transhipments with little or no temporary storage capabilities. Examples of satellites are small warehouses, parking lots, mobile vehicles (Gonzalez-Feliu 2012), micro-consolidation centres (Arrieta-Prieto et al. 2022), or even public transit stops (Delle Donne, Alfandari, et al. 2023; Delle Donne, Santini, and Archetti 2024).

Each satellite is associated with a given portion of the city, called an *area*. All deliveries within a given area will occur with vehicles starting and ending their routes at the corresponding satellite. Areas are further grouped into *regions*. We assume that hiring decisions must be taken at the tactical and regional level, i.e., a courier is hired for a specific region and an extended period. Assignment of couriers to satellites (and, therefore, to areas) happens on the operational level according to the needs of the logistic operator. Figure 1 shows the example of a city in which three regions (delimited with thicker black lines) are further divided into several areas (delimited with white lines). Blue squares indicate the locations of the satellites.

The logistic operator must decide (i) how many couriers to hire in the mid-to-long term in each region (workforce sizing) and (ii) which area to assign them in the short term to minimise the combined labour costs and expected outsourcing costs (assignment and scheduling). To build a mathematical model to address this problem, we will first present intermediate models with simplifying assumptions. We consider these models not only because they simplify our exposition but also because they correspond to different levels of flexibility allowed to the decision-maker. For example, we will first assume that the logistics company can employ couriers to work shifts as short as desired. Potentially, the company could hire ten couriers to work from 19:00 to 21:00 and assign them to a different area every day. However, concerns about job quality and service level suggest that further conditions be imposed. We will then consider situations in which couriers must be hired for fixed shifts (e.g., 08:00–16:00 each day) or for flexible shifts (e.g., a period of 8 consecutive hours, but starting at any time during the day) and that couriers can move between areas, but only within the same region. Indeed, one of the contributions of our work is to initiate a discussion on the impact of shift flexibility on the company's bottom line, complementing research in the social sciences, which instead addresses the effect of shift instability on workers' well-being (see Section 2.3).

We emphasise that we are concerned with sizing and scheduling the workforce, assuming that the company already owns a fleet of vehicles. Therefore, we do not study the problem of purchasing or leasing vehicles and do not consider the corresponding sunk costs. For recent works on fleet sizing, we refer the reader to, e.g., (Franceschetti et al. 2017; Banerjee, Erera, and Toriello 2022; Shehadeh, Hai Wang, and Zhang 2021; Ertogral, Akbalik, and González 2017; Kunz and Van Wassenhove 2019; Goulart et al. 2021; Rahimi-Vahed et al. 2015; Castillo et al. 2022; Loxton and Lin 2011). We review the contributions more closely related to the present paper in Section 2.2.

Finally, we highlight the stochastic nature of our problem. The decision-maker can estimate the number of deliveries in each area but cannot know this number precisely on the timescale required to make tactical decisions. Therefore, we will introduce a subproblem to estimate the number of parcels delivered by the hired couriers in each area and the number of parcels which must instead be outsourced. To this end, we will evaluate several demand scenarios and adapt approximation formulas from the literature (Figliozzi 2008).

The rest of the paper is organised as follows. In Section 2, we position our contribution in the literature on LMD scheduling and review related topics such as fleet sizing and districting. We also review current literature on the topic of stability in workforce scheduling from both operations research and the social sciences. In Section 3, we formalise our problem and introduce several mathematical formulations, which share the same base but differ in the amount of flexibility available to the decision-maker. Because the formulations are extremely quick to solve using commercial software, in Section 4, we present the results of an extensive computational campaign. We provide managerial insights and highlight the roles of stability and flexibility on the costs and operations of an LMD logistics company. Finally, we summarise the main findings and our recommendations in Section 5.

To the best of our knowledge, this work is the first to explore the impact of stability in the context of workforce sizing and scheduling problems. In particular, we propose a new mathematical formulation that estimates the hiring and outsourcing costs of a company performing parcel deliveries in an urban environment. By extending this formulation, we can model different levels of shift stability and—through a vast computational campaign—provide insights into how the company performs under different exogenous conditions such as the demand volume, outsourcing costs, or demand patterns. Specifically, we show that shift stability can be guaranteed with a negligible sacrifice in profitability. This is a major outcome: companies which mostly focus on competitiveness can anyway act towards improving workers' well-being.

2 Related works

This section presents related contributions and positions the current work in the literature. We focus on three main areas. The first is tactical workforce scheduling, focusing exclusively on LMD. Compared with classical scheduling, LMD presents additional challenges. Most notably, demand is stochastic and seasonal; when some couriers are crowdsourced, supply is also stochastic. The second research area is fleet sizing and districting. These decisions usually happen before workforce scheduling and are better classified as strategic rather than tactical. Still, there are several points of contact with our work, especially in the methodology used to approximate operational-level costs. Finally, we add an ethical dimension to our research by considering the stability of employee shifts. Research in social science associates stable work shifts with higher job satisfaction, better work-life balance and reduced turnover. We briefly review this literature, motivating the investigation into the impact of shift stability vs. flexibility on the bottom line of logistic operators.

We remark that the above areas are not exhaustive. Indeed, recent literature on optimising LMD operations has focused on timely real-world problems at all decision-making levels. For example, at the operational level, on determining which deliveries to outsource to crowd couriers (Fatehi and Wagner 2022); at the tactical level, on balancing driver workload over a week or a month (Y. Wang et al. 2022); at the strategic level, on partitioning urban areas into regions (Carlsson et al. 2024). While these decisions are related to workforce sizing and scheduling, we only include the contributions that share methodological or motivating characteristics with our work in the following review.

2.1 Workforce scheduling for last-mile delivery

Workforce scheduling concerns the assignment of couriers to perform deliveries in given areas during specific periods. It is a critical task in all parts of the supply chain, particularly in its most labour-intensive segment: the last mile. Yildiz and Savelsbergh (2019) have identified "the importance of having the right number of couriers at the right time" in last-mile meal delivery, and indeed, their observations can be generalised to other types of LMD.

A few works in the literature tackle workforce scheduling in the LMD setting. Restrepo, Semet, and Pocreau (2019) consider the combined problems of scheduling couriers (at the tactical level) and assigning them specific orders (at the operational level). Unlike our approach, the authors assume staffing is already decided, and the workforce size is fixed. On the other hand, similarities with our setting include the possibility of outsourcing deliveries when capacity is full and the fact that the territory is divided into areas. The authors propose an exact two-stage stochastic approach. The first-stage tactical problem assigns couriers to shifts and areas, while the second-stage operational problem allocates orders to couriers (or an outsourcing provider). Using an L-shaped method (Laporte and Louveaux 1993), they achieve solutions with average gaps of 1.07% from the optimum in realistic instances with up to 150 orders, 42 couriers, 23 scenarios and a capacity of at most two orders for each courier in a given period.

Another stream of work deals with courier scheduling under supply uncertainty. Some delivery companies use a mixed workforce of scheduled couriers and occasional ones, e.g., because they partly rely on crowdsourcing (Santini et al. 2022). Depending on the number of occasional couriers available, they face the double challenge of uncertain customer demand and supply capacity. Behrendt, Savelsbergh, and He Wang (2023) and Ulmer and Savelsbergh (2020) tackle the problem of a company relying on a mix of scheduled and crowdsourced couriers. This problem differs from ours because even scheduled couriers are hired and dismissed per shift. The company, in fact, can use the crowdsourcing platform to offer both shifts and single deliveries. If a person accepts a shift, they become available for the corresponding period, during which they can be assigned multiple deliveries. Otherwise, a person can accept to perform a single delivery without committing to be available for an extended period. The objective of the problem is to determine the ideal number of scheduled couriers to hire at each period to minimise labour costs and late-delivery penalties and thus decide the start time and duration of the shifts offered on the crowdsourcing platform. In (Behrendt, Savelsbergh, and He Wang 2023), the authors use continuous approximations and value function approximation methods to estimate the number of couriers required to meet a given service level, assuming a homogeneous order arrival rate. To the same end, Ulmer and Savelsbergh (2020) use a neural network trained on an off-line dataset generated via sample average approximation (Kleywegt, Shapiro, and Homem-de-Mello 2002).

Finally, we mention the work of Dai and P. Liu (2020b), who tackle the problem of determining the correct workforce size and parcel allocation for a combined staff of in-house couriers and crowdsourced drivers. They remark that an over-reliance on crowdsourcing can provide short-term benefits but sacrifices long-term objectives such as workforce retention and system robustness to fluctuation.

2.2 Fleet sizing and districting for last-mile delivery

The problem of deciding the size and composition of a fleet of vehicles is known as fleet sizing. Because purchasing or leasing vehicles involves high capital costs or long-term contracts, fleet sizing often happens at the strategic level. Similar to our staff scheduling problem, operational decisions are usually approximated when performing fleet sizing.

Fleet sizing happens before staff scheduling because the number of available vehicles determines how many couriers can work simultaneously. It can also happen before, after or simultaneously with districting, i.e., the problem of partitioning a given geographical region into fixed areas and distributing the vehicles among the areas. While, in principle, an operator could skip districting and solve a large routing problem each day, real-life practice has shown that geographical partitions drastically simplify operations and increase service quality (see, e.g., Boysen, Fedtke, and Schwerdfeger 2021; S. Liu, He, and Zuo-Jun 2021; Monteiro Ferraz, Cappart, and Vidal 2022). In the following, we review two contributions closest to our approach.

Franceschetti et al. (2017) consider the problem of partitioning a rectangular city into rectangular areas. There are several differences between their approach and the problem we tackle in this paper. (a) In Franceschetti et al. (2017)'s settings, there is only one central depot from where all vehicles are deployed. From this point of view, their work is tailored more towards classical delivery vans than zero-emission vehicles. (b) The authors also consider the problem of designing the areas. However, one vehicle operates in each area; therefore, the operational problem is a Travelling Salesman Problem, compared to a Vehicle Routing Problem (VRP) in our case. (c) The couriers fulfil all requests with no possibility of outsourcing. (d) Because they focus on classical delivery vehicles, the authors also consider the cost of owning or leasing such vehicles and the transportation costs. Similar to our work, they use continuous approximation formulas to estimate operational routing decisions. The authors consider the case of a heterogeneous fleet with some vehicle types subject to access restrictions (i.e., they cannot enter certain areas during some parts of the day). After analysing optimal solutions obtained via Dynamic Programming and a Mixed-Integer model, they conclude that access restrictions can sometimes be counterproductive, increasing the total number of vehicles on the road. Their computational results also show that the advantage of having a heterogeneous fleet is minor compared with the corresponding increase in operational complexity.

Banerjee, Erera, and Toriello (2022) tackle a similar problem of designing distribution areas and determining the correct fleet size to deploy in a same-day-delivery system. Similar to Franceschetti et al. (2017), all vehicles start and end their routes from the depot, each area is allocated one vehicle, and no outsourcing is possible. However, the city and its areas are not limited to being rectangular. The problem characteristics hint at a strong correspondence between minimising the fleet size and maximising the area covered by each vehicle. The authors exploit this link to develop area-maximising policies and apply a fleet-size minimisation algorithm. Both in Franceschetti et al. (2017) and Banerjee, Erera, and Toriello (2022), similar to our work, the authors use a continuous approximation model to estimate operational cost. However, the problem setting fundamentally differs from the one we consider. First, the focus of the above two papers is on districting decisions; moreover, both works assume that a single vehicle operates in each area while we allow deploying multiple couriers.

2.3 Geographical and temporal stability in workforce scheduling

While, on paper, extremely flexible and volatile schedules might appear the most suited to meet a highly dynamic demand, real-world practice reveals the importance of stability and planning at the tactical level.

Regarding *geographical* stability, a stable assignment of couriers to areas or even to customers leads to shorter routing and service times. In classical supply chains where drivers visit a few large customers, consistency leads to quicker operations and increased customer satisfaction. In a seminal work, Groër, B. Golden, and Wasil (2009) introduced the Consistent Vehicle Routing Problem (ConVRP), a multi-period routing problem rewarding stability in assigning drivers to customers and visiting the same customer at similar times. This problem has garnered considerable attention: we refer the reader to a survey by Kovacs et al. (2014) and to the work of Smilowitz, Nowak, and Jiang (2013) for links between the ConVRP and workforce management. Regarding recent contributions published after the survey, see, e.g., (Rodríguez-Martín, Salazar-González, and Yaman 2019; M. Schneider 2016; Goeke, Roberti, and M. Schneider 2019). Stable assignment of drivers to areas is also beneficial in modern last-mile settings. For example, in a recent keynote presentation at the 12$^{\text{th}}$ DIMACS implementation challenge, Werneck (2022) has emphasised the importance of consistency and driver familiarity for last-mile delivery at Amazon.

The topic of *temporal* stability falls in the broader literature of workforce scheduling with workers' preferences (see, e.g., (Ruiz-Torres et al. 2015; Mohan 2008; Yura 1994) and the survey of Van den Bergh et al. (2013)).

The issues of shift instability and low wages have been identified among the most critical aspects of the modern logistics industry, especially in the last mile.

From the couriers' point of view, unstable or unreliable shifts cause a sensible decrease in happiness and worsen the work-life balance. In a survey of workers in Illinois, United States, Dickson, L. Golden, and Bruno (2018) showed that, in 2018, 35% knew their schedule at least one week before and 10% only knew it 24 hours in advance. Furthermore, the average gap between the minimum and maximum weekly hours during the six months before the survey was 14 hours, suggesting large fluctuations from one week to another. Part-time workers, who are largely represented in LMD, are particularly affected: Dickson, L. Golden, and Bruno (2018) report that "the incidence of unpredictable or varying shift times […] falls disproportionately on part-time workers—12 percent of part-timers experience irregular shift times". Furthermore, the downsides of erratic shifts affect some categories, such as single parents, more than others (Ananat and Gassman-Pines 2021; Harknett, D. Schneider, and Luhr 2022). Carrillo et al. (2017) highlights that: "as dual-earner couples, single parent families, and irregular work schedules have risen in prevalence, the logistics of arranging child care to match work shifts have grown increasingly complex". This complexity comes at a significant cost for children; quoting again Carrillo et al. (2017): "Instability and unpredictability at work were reproduced in […] child-care arrangements at home. This scramble led to inconsistency in children's care and also imposed a heavy psychological burden on parents as they reconciled the difficulty of finding care for their children with the imperative to keep an open availability for work and catch the shifts that became available to them."

From the point of view of logistic operators, excessive shift instability causes higher employee turnover and lower performance, producing a net detrimental effect for the firm. Chung (2022) studies the impact of variable work schedules (VWS) on quick-service restaurant chains and concludes that "despite the common assumption that their [VWS] use helps firms achieve higher performance by matching the supply of labor to demand fluctuations [...] this study demonstrates otherwise", and that "scholars and practitioners should reconsider the general assumption that staffing flexibility helps organisations adapt to uncertain environments". Although the role of schedule volatility on employee turnover has been studied before (see, e.g., Henly and Lambert 2014; Henly, Shaefer, and Waxman 2006), its impact has increased after the Covid-19 pandemic (see Choper, D. Schneider, and Harknett 2022; Bergman, David, and Song 2023; Rhee, Park, and Lee 2020).

The above considerations clarify that shift flexibility in workforce scheduling in LMD is worth researching and that the literature on this topic can be enriched considering an ethical dimension (Le Menestrel and Van Wassenhove 2009; Ormerod and Werner 2013; Bellenguez, Brauner, and Tsoukiàs 2023).

3 Problem setting and formulation

In this section, we formalise the problem we are studying and provide several mathematical models to make workforce sizing and scheduling decisions. Section 3.1 provides a mathematical description of the considered problem. We present a base mathematical model in Section 3.3 and extend it in Section 3.4.

3.1 Problem description

We consider a logistic operator working in the last mile of the supply chain in an urban environment. The city is divided into a set A of areas, usually corresponding to districts or neighbourhoods, with exactly one satellite in each area. Areas are grouped into regions $\mathcal{R}$, such that $\mathcal{R}$ forms a partition of A.

The daily planning horizon is discretised into a set Θ of periods. The planner chooses the length of the periods in a way that (a) is long enough to ensure that couriers complete their tours and (b) is suitable for accurately estimating the demand based on historical data. In our main application setting, couriers employ sustainable vehicles with limited capacities, such as cargo bikes. Because vehicle capacities limit tour durations, the duration of a period is usually limited to a few hours.

Couriers start their tours at the beginning of each period and return to the starting satellite before the end of the period. If a courier changes his/her assigned satellite between two consecutive periods, we assume the transfer time to be negligible. This assumption is justified because transfers are only possible between satellites in the same region.

Each area has an associated demand distribution, which determines how many deliveries are required during each period. We adopt a scenario-based approach and consider a set S of scenarios. Each scenario $s \in S$ determines, for each area $a \in A$ and period $\theta \in \Theta$, the demand $n^s_{a\theta} \in \mathbb{N}$, i.e., the number of deliveries to perform.

The decision variable is the number of couriers assigned to each area during each period and is denoted with $x_{a\theta} \in \mathbb{N}$ ($a \in A$, $\theta \in \Theta$). This decision is taken at the tactical level; therefore, the assignment persists across all scenarios. Eventually, we will introduce constraints on this assignment, which guarantee, e.g., that couriers work for a minimum number of consecutive periods (a shift). For the moment, we only note that the number of couriers assigned to an area must not necessarily ensure they can deliver all parcels under all scenarios. Indeed, the cost incurred by the planner is the sum of the couriers' labour cost and the expected outsourcing cost. We denote with $c_{a\theta} > 0$ the unit cost of employing a courier in area a for period θ. The labour cost associated with a and θ is thus $c_{a\theta}x_{a\theta}$. We denote with $\omega_{a\theta}(x_{a\theta}) \geq 0$ the random variable representing the outsourcing cost for area a and period θ when employing $x_{a\theta}$ couriers. In our approach, we estimate the expected value of $\omega_{a\theta}$ by computing the average outsourcing cost over all scenarios:

$$\mathbb{E}\big[\omega_{a\theta}(x_{a\theta})\big] \simeq \frac{1}{|S|}\sum_{s\in S}\omega^s_{a\theta}(x_{a\theta}),$$

where $\omega^s_{a\theta}(x_{a\theta})$ is the deterministic outsourcing cost incurred under scenario s. With the above notation, the objective function of our problem is

$$\min_{x_{a\theta}\in\mathbb{N}} \quad \sum_{a\in A}\sum_{\theta\in\Theta}\left(c_{a\theta}x_{a\theta} + \frac{1}{|S|}\sum_{s\in S}\omega^s_{a\theta}(x_{a\theta})\right). \tag{1}$$

Determining the value of $\omega^s_{a\theta}(x_{a\theta})$ is not straightforward. To know how many deliveries can be performed by $x_{a\theta}$ couriers (and, thus, how many must be outsourced), we would have to solve an instance of the $\mathcal{NP}$-complete Capacitated Vehicle Routing Problem (CVRP) for each area, period and scenario. However, knowing the exact value of $\omega^s_{a\theta}$ is unnecessary at the tactical planning level. Therefore, in Section 3.2, we devise a method to approximate this value. In the rest of this section, we introduce constraints which, together with the objective function (1), will model realistic staff sizing problems faced by LMD operators.

3.2 Approximation of the outsourcing costs

We assume that the outsourcing cost depends linearly on the number of outsourced deliveries. Specifically, let $C_{\text{out}} > 0$ be the cost to outsource one delivery and $m^s_{a\theta} \in \mathbb{N}$ the number of couriers needed to fulfil all deliveries in area $a \in A$ during period $\theta \in \Theta$ according to scenario $s \in S$.

Under the assumption that all couriers deliver roughly the same number of parcels, we can write the outsourcing cost function as

$$\omega^s_{a\theta}(x_{a\theta}) = \begin{cases} 0 & \text{if } x_{a\theta} \geq m^s_{a\theta}, \\ (m^s_{a\theta} - x_{a\theta})\frac{n^s_{a\theta}}{m^s_{a\theta}}C_{\text{out}} & \text{otherwise.} \end{cases} \tag{2}$$

Equation (2) states that the planner does not incur outsourcing costs if a sufficient number of couriers is hired. Otherwise, a fraction of $1 - x_{a\theta}/m^s_{a\theta}$ of the deliveries must be outsourced at unit cost C_{out}.

The problem of calculating $\omega^s_{a\theta}(x_{a\theta})$ then reduces to the computation of $m^s_{a\theta}$. As mentioned above, its exact value is given by the solution of a CVRP. In the following, we propose to compute an approximation $\hat{m}^s_{a\theta}$. To this end, denote with $\alpha_a > 0$ the surface of area a, with $\bar{r}_a > 0$ the average distance between a point in a and the satellite, with $Q > 0$ and $v > 0$ respectively the capacity and the speed of the courier vehicles, with $\tau > 0$ the service time at the customer's, and with $T > 0$ the duration of a period. We note that capacity Q is expressed in the number of deliveries (thus assuming that parcels are not too dissimilar in size) and that the unit of measure of v is derived from those of $\bar{r}_a$ and T (i.e., units of space over units of time).

Figliozzi (2008) proposed a closed-form approximation of the cost of the optimal solution of a VRP with $n^s_{a\theta}$ customers and m vehicles:

$$k_a \cdot \frac{n^s_{a\theta} - m}{n^s_{a\theta}}\sqrt{\alpha_a \cdot n^s_{a\theta}} + 2\bar{r}_a \cdot m, \tag{3}$$

where k_a is a coefficient which depends on the shape of area a and must be learned, e.g., via regression. We extend this formula to use it as an approximation of the average time a courier needs to complete a route, including the service time at the customers:

$$\text{route time} = \frac{1}{m}k_a\frac{n^s_{a\theta} - m}{v \cdot n^s_{a\theta}}\sqrt{\alpha_a \cdot n^s_{a\theta}} + 2\frac{\bar{r}_a}{v} + \frac{n^s_{a\theta}}{m} \cdot \tau. \tag{4}$$

The first term approximates the travel time between customers, the second approximates the round trip from the satellite, and the third term accounts for the service time at the customers. Because each courier must respect both the capacity and the route duration constraint, we look for the smallest integer m such that

$$\begin{aligned} & m \geq \frac{n^s_{a\theta}}{Q}, \text{ and} \\ & \frac{1}{m} k_a \frac{n^s_{a\theta} - m}{v \cdot n^s_{a\theta}} \sqrt{\alpha_a \cdot n^s_{a\theta}} + 2\frac{\bar{r}_a}{v} + \frac{n^s_{a\theta}}{m} \cdot \tau \leq T. \end{aligned}$$

Simple algebraic manipulations yield

$$\hat{m}^s_{a\theta} = \left\lceil \max \left\{ \frac{n^s_{a\theta}}{Q}, \frac{\frac{k_a}{v} \sqrt{\alpha_a n^s_{a\theta}} + n^s_{a\theta} \cdot \tau}{T + \frac{k_a}{v \cdot n^s_{a\theta}} \sqrt{\alpha_a n^s_{a\theta}} - \frac{2\bar{r}_a}{v}} \right\} \right\rceil. \tag{5}$$

3.3 Base model

We introduce a base optimisation model which uses objective function (1) and the framework described in Section 3.2 to solve a loosely constrained version of our problem. Indeed, the only basic constraints we introduce are: (a) a global upper bound $u \in \mathbb{N}$ on the number of couriers that the logistic operator can employ during any given period and (b) a regional-level upper bound $u_R \in \mathbb{N}$ on the number of couriers that can work in region $R \in \mathcal{R}$ during any given period. These bounds can derive from real-life considerations, such as a staffing budget (for bound u) or the number of available vehicles (for bounds u_R). The model, denoted MBase, reads as follows.

$$\begin{aligned} \min \quad & \sum_{a \in A} \sum_{\theta \in \Theta} \left(c_{a\theta} x_{a\theta} + \frac{1}{|S|} \sum_{s \in S} \hat{\omega}^s_{a\theta}(x_{a\theta}) \right) && && \text{(6a)} \\ \text{s.t.} \quad & \sum_{a \in R} x_{a\theta} \leq u_R && \forall R \in \mathcal{R},\ \forall \theta \in \Theta && \text{(6b)} \\ & \sum_{a \in A} x_{a\theta} \leq u && \forall \theta \in \Theta && \text{(6c)} \\ & x_{a\theta} \in \mathbb{N} && \forall a \in A,\ \forall \theta \in \Theta. && \text{(6d)} \end{aligned}$$

In (6b) and (6c), the bounds are enforced for all periods $\theta \in \Theta$ to ensure that the maximum workforce size is not exceeded at any period of the planning horizon. Function $\hat{\omega}^s_{a\theta}$ denotes the approximate outsourcing cost $\omega^s_{a\theta}$ in which $m^s_{a\theta}$ is replaced by $\hat{m}^s_{a\theta}$ in eq. (2). Indeed, directly using (2), we obtain the following formulation for MBase:

$$\begin{aligned} \min \quad & \sum_{a \in A} \sum_{\theta \in \Theta} \left(c_{a\theta} x_{a\theta} + \frac{1}{|S|} \sum_{s \in S} \Omega^s_{a\theta} \right) && && \text{(7a)} \\ \text{s.t.} \quad & \sum_{a \in R} x_{a\theta} \leq u_R && \forall R \in \mathcal{R},\ \forall \theta \in \Theta && \text{(7b)} \\ & \sum_{a \in A} x_{a\theta} \leq u && \forall \theta \in \Theta && \text{(7c)} \\ & \Omega^s_{a\theta} \geq (\hat{m}^s_{a\theta} - x_{a\theta}) \frac{n^s_{a\theta}}{\hat{m}^s_{a\theta}} C_{\text{out}} && \forall a \in A,\ \forall \theta \in \Theta,\ \forall s \in S && \text{(7d)} \\ & x_{a\theta} \in \mathbb{N} && \forall a \in A\ \forall \theta \in \Theta && \text{(7e)} \\ & \Omega^s_{a\theta} \geq 0 && \forall a \in A,\ \forall s \in S\ \forall \theta \in \Theta. && \text{(7f)} \end{aligned}$$

In model (7a)–(7f), we introduced new variables $\Omega^s_{a\theta}$ to hold the value of $\hat{\omega}^s_{a\theta}(x_{a\theta})$. We remark that the above formulation is decomposable by period θ; however, when we add further constraints in Section 3.4, this will no longer be the case.

The complete solution of model MBase requires two steps. First, we compute the approximate values $\hat{m}^s_{a\theta}$ using (5), for each s, a and θ. Then, we use these values to solve model MBase and to find the optimal values of $x_{a\theta}$ and their corresponding cost.

3.4 Shift linking constraints

Model MBase potentially allows employing couriers for just one period or intermittent periods during the day. However, job quality and service level considerations forbid such practices in most real-life situations. To this end, we introduce the concept of shifts: a set of consecutive periods such that if a courier works during one of them, they must work during all of them.

While the concept of a work shift is almost universally employed, its implementation changes from company to company. In what follows, we propose three types of shifts with different levels of flexibility. Each will correspond to new variables and constraints extending model MBase; their impact will be evaluated in Section 4.

The first type is fixed shifts: partitioning the set of periods Θ into contiguous non-overlapping sets. For example, if the working day is from 9 AM to 9 PM and each period spans two hours, we would have $\Theta = \{1, \ldots, 6\}$. Two fixed shifts could be 9 AM to 3 PM (periods 1 to 3) and 3 PM to 9 PM (periods 4 to 6). The start period and the duration of each shift are given in advance by labour regulations or local uses, and are not decision variables. Each courier is assigned to one of these preset shifts.

The second type is flexible shifts. Each courier has an associated shift, i.e., a set of contiguous periods of fixed total duration. However, the start time of each courier's shift is not given in advance and is a decision variable: different couriers can have shifts starting at different times. Unlike fixed shifts, flexible shifts do not need to partition the set of periods and can overlap.

The third type is partially flexible shifts, which provide intermediate flexibility between fixed and flexible shifts. In this case, we limit the number of possible distinct shift start times. For example, we might aim to create four possible shifts (i.e., selecting four possible start times) and then assign one shift to each courier. If the number of possible start times equals the number of periods, then we are in the special case of flexible shifts. On the other hand, limiting the number of shift start times to only a few possibilities decreases the system's flexibility and creates stabler rosters for the couriers.

Figure 2 shows three examples of shifts which can be devised for the same demand pattern displayed at the bottom. The blue shifts at the top are fixed with 9 AM and 3 PM start times. The red shifts in the middle are flexible, and the yellow ones at the bottom are partially flexible.

In the remainder of this section, we introduce the necessary notation, shift-type-specific variables, and constraints that we add to model MBase.

3.4.1 Fixed shifts

Let $\mathcal{P}$ be the set of shifts, i.e., a partition of Θ such that each shift $P \in \mathcal{P}$ is a contiguous set of periods. We denote with θ^{s}_P and θ^{e}_P the first and last periods of shift P.

We introduce new variables $y_{a_1 a_2 \theta} \in \mathbb{N}$ denoting the number of couriers moving from area a_1 to area a_2 between periods $\theta - 1$ and θ ($a_1 \neq a_2$, a_1 and a_2 belong to the same region, θ is not the first period of the day).

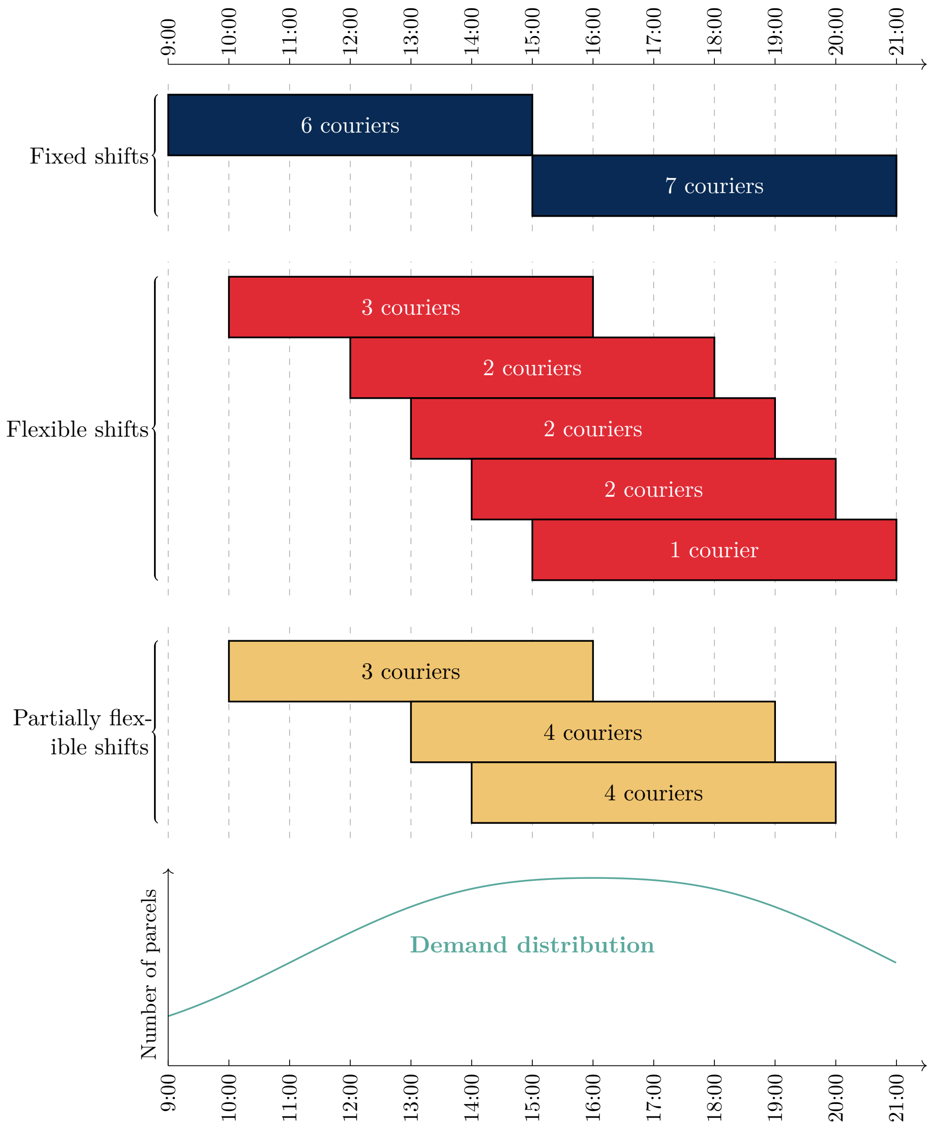

Figure 2: Example of fixed (blue), flexible (red) and partially flexible (yellow) shifts for a 12-hour working day. The demand distribution at the bottom shows that the afternoon is busier than the morning.

We add the following constraints to formulation (7a)–(7f) to model fixed shifts:

$$\sum_{a_2 \in R} y_{a_1 a_2 \theta} \leq x_{a_1 \theta} \qquad \forall R \in \mathcal{R},\ \forall a_1 \in R,\ \forall \theta \in \Theta \tag{8a}$$

$$\sum_{a \in R} x_{a\theta} = \sum_{a \in R} x_{a\theta^{\mathrm{s}}_P} \qquad \forall R \in \mathcal{R},\ \forall P \in \mathcal{P},\ \forall \theta \in P \setminus \{\theta^{\mathrm{s}}_P\} \tag{8b}$$

$$x_{a_1\theta} = x_{a_1,\theta-1} + \sum_{a_2 \in R \setminus \{a_1\}} y_{a_2 a_1 \theta} - \sum_{a_2 \in R \setminus \{a_1\}} y_{a_1 a_2 \theta} \qquad \forall R \in \mathcal{R},\ \forall a_1 \in R,\ \forall P \in \mathcal{P},\ \forall \theta \in P \setminus \{\theta^{\mathrm{s}}_P\}. \tag{8c}$$

Constraint (8a) ensures that no more couriers move away from each area a_1 than there are working in a_1. Constraint (8b) ensures that the number of employed couriers stays constant within each region for the duration of each shift, thus forbidding hiring or dismissing couriers in the middle of a shift. Constraint (8c) states that the number of couriers working in area a_1 during period θ is given by the number of couriers working in a_1 during the previous period, plus couriers who move into a_1, minus couriers who move out of a_1. We denote with Fixed the model obtained adding (8b)–(8c) to MBase.

3.4.2 Flexible shifts

To model flexible shift, we add to the already introduced x and y new variables $z^-_{a\theta} \in \mathbb{N}$, denoting the number of couriers starting their shift in area a at the beginning of period θ, and $z^+_{a\theta} \in \mathbb{N}$, denoting the number of couriers ending their shift in area a at the end of period θ.

Denoting with $\ell \in \mathbb{N}$ the shift length, we observe that variables $z^-_{a\theta}$ must be set to zero for all areas $a \in A$ and for periods $\theta \in \Theta$ such that $\theta > |\Theta| - \ell$. Indeed, a shift has to start before $|\Theta| - \ell$ in order to satisfy shift duration ℓ. Analogously, variables $z^+_{a\theta}$ are set to zero for all areas $a \in A$ and for periods $\theta \in \Theta$ such that $\theta < \ell$.

To obtain a model for the flexible shifts, we add constraint (8a) and the following constraints to formulation (7a)–(7f):

$$\sum_{a \in R} z^-_{a\theta} = \sum_{a \in R} z^+_{a,\theta+\ell-1} \qquad \forall R \in \mathcal{R},\ \forall \theta \in \Theta,\ \theta \leq |\Theta| - \ell + 1 \tag{9a}$$

$$\begin{aligned} x_{a_1\theta} = x_{a_1,\theta-1} & \\ + \sum_{a_2 \in R \setminus \{a_1\}} & y_{a_2 a_1 \theta} \\ - \sum_{a_2 \in R \setminus \{a_1\}} & y_{a_1 a_2 \theta} \\ + z^-_{a_1\theta} - & z^+_{a_1,\theta-1} \end{aligned} \qquad \forall R \in \mathcal{R},\ \forall a_1 \in R,\ \forall \theta \in \Theta,\ \theta > 1 \tag{9b}$$

$$x_{a1} = z^-_{a1} \qquad \forall a \in \mathcal{A}. \tag{9c}$$

Constraint (9a) makes sure that all couriers starting a shift at the beginning of period θ complete it at the end of period $\theta + \ell - 1$. Constraint (9b) extends (8c) by considering couriers who start or end their shift. Because (9b) is defined for $\theta > 1$, constraint (9c) addresses the special case of the beginning of the planning horizon, stating that all workers who start a shift during the first period are working in the respective areas. We denote with Flex the model obtained by adding (9a)–(9c) to MBase.

3.4.3 Partially flexible shifts

To model partially flexible shifts, we introduce variables $w_\theta \in \{0, 1\}$ ($\theta \in \Theta$, $\theta \leq |\Theta| - \ell + 1$) taking value 1 if and only if a shift starts at the beginning of period θ. Let $\mu \in \mathbb{N}^+$ be the maximum number of shifts to create. A model for partially flexible shift uses constraints (9a)–(9c) together with the following inequalities:

$$\sum_{a \in R} z^{-}_{a\theta} \leq u_R \cdot w_\theta \qquad \forall R \in \mathcal{R},\ \forall \theta \in \Theta,\ \theta \leq |\Theta| - \ell + 1 \tag{10a}$$

$$\sum_{\theta=1}^{|\Theta|-\ell+1} w_\theta \leq \mu. \tag{10b}$$

Constraint (10a) links the z and w variables allowing couriers to start their shift only when such a shift is created (value u_R acts as a "big-M" constant). Constraint (10b) limits the number of created shifts. We denote with PartFlex the model obtained by adding (9a)–(9c) and (10a)–(10b) to MBase.

3.4.4 Dealing with symmetry

Models using variables $y_{a_1 a_2 \theta}$ suffer from symmetry. For example, increasing by one the value of both $y_{a_1 a_2 \theta}$ and $y_{a_2 a_1 \theta}$ yields a new solution with the same cost and corresponding to an unrealistic scenario (two couriers swapping areas without reason). To break this symmetry, we add to the objective function (7a) the following term:

$$\varepsilon \cdot \sum_{R \in \mathcal{R}} \sum_{a_1 \in R} \sum_{a_2 \in R \setminus \{a_1\}} \sum_{\theta \in \Theta} y_{a_1 a_2 \theta}, \tag{11}$$

where $\varepsilon > 0$ is a small constant. The term (11) penalises unnecessarily large values of variables $y_{a_1 a_2 \theta}$ and prevents situations such as the one described above. In the above example, increasing by one the value of $y_{a_1 a_2 \theta}$ and $y_{a_2 a_1 \theta}$ would cause the objective function to increase by $2\varepsilon > 0$ making the resulting solution sub-optimal.

4 Results

In this section, we present the results of our computational experiments. First, we describe how we generated our instances based on realistic data from four large European cities. Second, we show that the optimisation problems presented in Section 3 are fast to solve on commonly available computers, making them particularly suitable as decision support tools. Third, exploiting this computational efficiency, we perform a large experimental campaign aimed at deriving managerial insights and assessing the role of shift flexibility on the company's bottom line.

4.1 Instances

We generate instances based on Paris, Lyon (France), Berlin and Frankfurt (Germany). There are three main components to instance generation: the geographical subdivision of each city into regions and areas, the demand distribution, and the parameters related to couriers (costs, bounds on workforce size, shift lengths, etc). All parameters are summarised in Table 1.

4.1.1 City geography

Each area in the cities corresponds to a postcode. We obtained the corresponding data under the Open Database License from OpenStreetMap (2023). Cities were subdivided into four regions, grouping areas to form compact groups of roughly equal population. We obtained population data from the EU's Global Human Settlement dataset (Schiavina et al. 2023). Figure 3 depicts the four cities and how they are divided into regions and areas. Paris has 20 areas, Lyon has 16, Berlin 59 and Frankfurt 32. Satellites are located around the centroids of the areas, ensuring that each satellite falls inside the area. The regression coefficient k_a of the cost approximation used in (5) is set to 0.77 $\forall a \in A$, as suggested by Figliozzi (2008) for areas with a central depot.

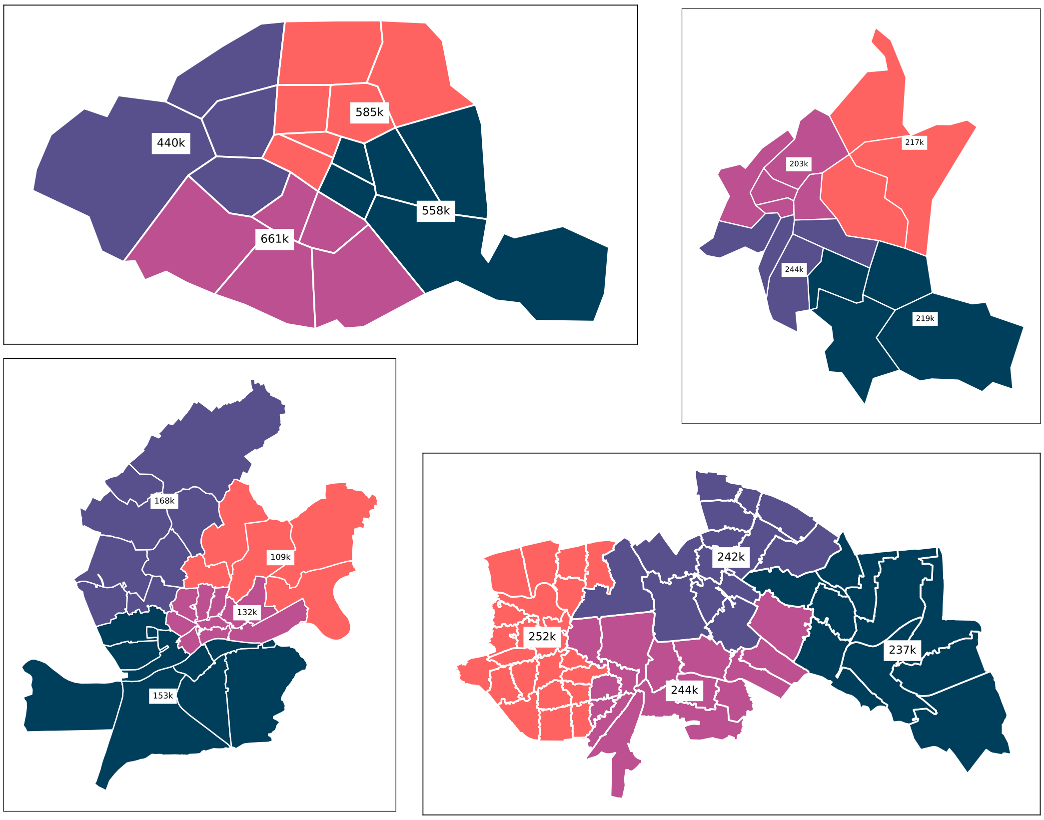

Figure 3: The four considered cities and their subdivision into areas (white boundaries) and regions (coloured). The numbers indicate the number of people living in each region. Top left: Paris, top right: Lyon, bottom right: Frankfurt, bottom left: Berlin.

Notation	Value(s)	Description
—	Berlin, Frankfurt, Lyon, Paris	City.
`DB`	0.5, 1, 2, 4	Number of parcels per 1000 inhabitants and day.
`DT`	Uniform, Peak, DoublePeak, AtEnd	Demand type.
`OC`	1.2, 1.5, 1.8, 2	Per-delivery outsourcing cost multiplier.
`RM`	0.75, 1, 1.5, 3, 5	Multiplier to determine the regional courier upper bound u_R.
`GM`	0.6, 0.7, 0.8, 0.9, 1	Multiplier to determine the global courier upper bound u.
μ	2, 3, 4	Maximum number of shifts for PartFlex.
—	16	Daily planning horizon in hours.
—	2	Period duration in hours.
k_a	0.77	Regression coefficient for VRP cost estimation.
Q	5	Courier capacity in number of parcels.
v	21	Courier speed in km/h.
τ	5	Courier service time in min.
$c_{a\theta}$	1	Courier labour cost per period.

Table 1: Instance generation parameters.

4.1.2 Demand distribution

Demand is proportional to area population via a Demand Baseline (`DB`) parameter expressed in the number of parcels per thousand inhabitants and day. When we generate an instance, the daily demand of each area is chosen uniformly at random in the interval $[0.75\texttt{DB}, 1.25\texttt{DB}]$. Once the total daily demand is determined, we distribute it among the periods making up the time horizon. Our instances use eight periods of two hours each for a daily planning horizon of 16 hours (6 AM to 10 PM). We consider four ways of assigning demand to each period based on a demand type parameter (`DT`). We describe them below and visualise them in Figure 4.

- For Uniform demand, we distribute the total number of parcels uniformly throughout the planning horizon. Although home deliveries rarely occur steadily throughout the day, we use this demand type as a baseline.
- Peak demand resembles the histogram of a truncated normal distribution with mean $|\Theta|/2$, standard deviation $|\Theta|/6$, left extreme 0 and right extreme $|\Theta|$. It corresponds to a peak in deliveries in the central hours of the day.
- DoublePeak demand follows the histogram of a mixture of two truncated normal distributions. They are similar to the distribution used for Peak demand, but their means are at $|\Theta|/3$ and $2|\Theta|/3$, and their standard deviation is $|\Theta|/10$. It simulates two peak hours: a morning one (around 11 AM) for workplace deliveries and an evening one (around 5 PM) for home deliveries.
- AtEnd is similar to Peak demand, but the mean of the truncated normal distribution is at $2|\Theta|/3$. It corresponds to a situation where most deliveries are at people's homes at the end of the workday (around 5 PM).

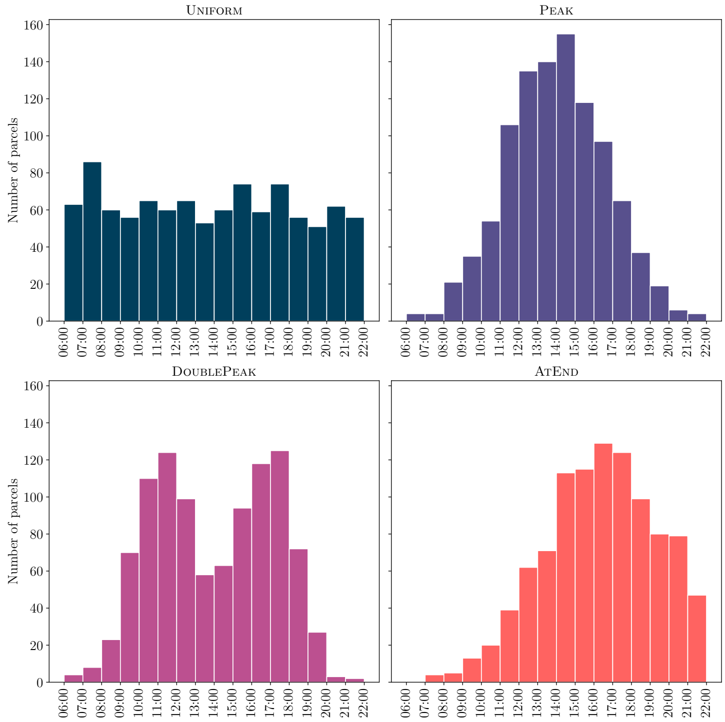

Figure 4: Example of hourly parcel demand according to each of the four demand types (DT) used in instance generation. The daily demand in the given area is 1000.

4.1.3 Courier parameters

We consider a uniform fleet of bike couriers with capacity $Q = 5$ parcels and speed $v = 21$ km/h (Romanillos and Gutiérrez 2020). We normalise the labour costs to $c_{a\theta} = 1\ \forall a \in A,\ \forall \theta \in \Theta$. If couriers travel at full capacity, i.e., carrying five parcels, the courier per-delivery cost is then 0.2. We obtain the per-delivery outsourcing cost C_{out} by multiplying this figure by a multiplier OC. For example, when $\texttt{OC} = 1.2$, we set $C_{\text{out}} = 0.2 \cdot 1.2 = 0.24$.

Recall that $\hat{m}^s_{a\theta}$, defined in Section 3.2, is the approximate number of couriers required to deliver all parcels in area a during period θ according to scenario s. Then, the average number of couriers required to deliver all parcels in a during θ across all scenarios is $\frac{1}{|S|} \sum_{s \in S} \hat{m}^s_{a\theta}$. Averaging over all periods, we obtain values $\hat{m}_a = \frac{1}{|\Theta|} \sum_{\theta \in \Theta} \hat{m}^s_{a\theta}$ and $\hat{m}_R = \sum_{a \in R} \hat{m}_a$ (for $R \in \mathcal{R}$). This is a rough approximation of the number of couriers per period required in each region to serve the entire demand. Especially for non-Uniform demand types, this average will not be a good approximation, and more couriers will be required during peak periods and fewer during valley periods. Indeed, we only use $\hat{m}_R$ as a baseline to choose parameters u_R and u, i.e., the per-region and global upper bounds on the number of couriers we can employ. We set the regional upper bounds as $u_R = \texttt{RM} \cdot \hat{m}_R$ and the global upper bound $u = \texttt{GM} \cdot \sum_{R \in \mathcal{R}} u_R$, where RM and GM are parameters. When GM takes value 1, global bound u is moot, and only the regional bounds can be tight.

4.1.4 Instance availability

Varying the parameters introduced in this Section (city, DB, DT, OC, RM, GM), the model (MBase, Flex, PartFlex, Fixed) and, for model PartFlex, the value of μ, we obtain a large set of 8 000 instances and 48 000 experiments. We generate 90 scenarios per each instance by repeatedly drawing from the relevant random distributions. Preliminary experiments, however, determined that reducing the number of scenarios to 30 does not significantly affect the quality of the cost

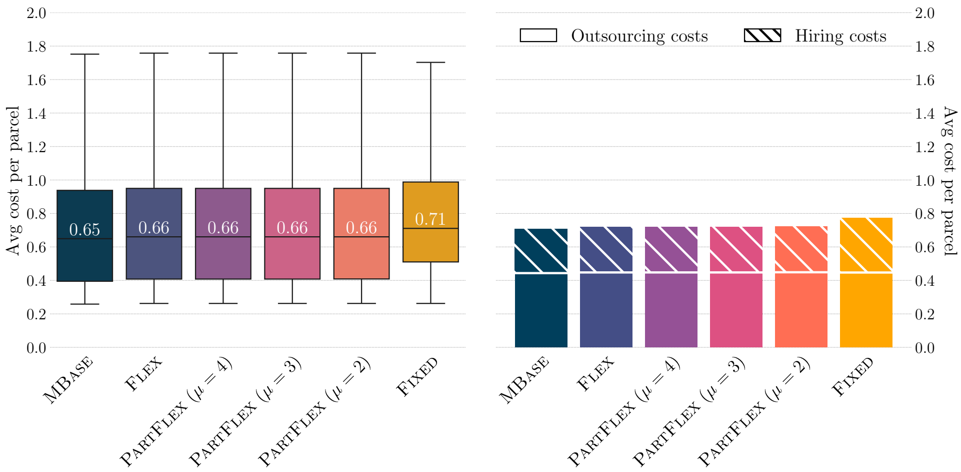

Figure 5: Cost per parcel vs. model. The left box plot summarises the cost distribution over all instances. The bar plot on the right shows the average over all instances and splits the cost into its hiring and outsourcing components.

approximation and, perhaps more importantly, the overall solution. We provide in repository (Mandal, Santini, and Archetti 2024) the instances that we use in this study. The repository also includes the scripts used to generate the instances, the solver, and the scripts used to produce tables and figures.

4.2 Insights

We implemented the models in Python using Gurobi 9 as the MIP solver. Gurobi solves each instance in a fraction of a second, ranging from an average of 0.06 s for model MBase to 0.13 s for PartFlex with $\mu = 2$. This allowed us to run an extensive computational campaign on our large instance set and to draw the managerial insights described in the rest of this section. The main research question is to understand the impact of shift stability on the logistic provider in terms of costs and operational complexity.

4.2.1 High-level impact of flexibility on costs

The main high-level result about costs is depicted in Figure 5. The left figure shows the cost per parcel when using the different models. This cost is defined as the objective value of the optimal solution divided by the total number of parcels to deliver over all areas and periods. Each box represents the cost distribution over all the instances and, therefore, refers to 8000 observations. The central line in the box is the median; its value is also written inside the box. The top and bottom borders are the third and first quartiles, respectively. Whiskers extend to the rest of the distribution except for outliers, i.e., observations more extreme than twice the interquartile range. Because our hiring costs are normalised and both hiring and outsourcing costs can vary considerably in different markets, the reader should consider their relative differences rather than their absolute values.

The figure shows two important effects. On the one hand, using fixed shifts results in noticeably higher costs, justifying the assumption that some degree of flexibility is required in an industry with unsteady demand. On the other hand, the difference between the base model (workers can be hired and dismissed at each period), the flexible model (shifts can start at any period), and the partially flexible model (shifts can only start in μ different periods) is small. In particular, moving away from MBase causes a marginal increase in the median cost per parcel (from 0.65 to 0.66, i.e., +1.54%) and the difference between the flexible and the partially flexible

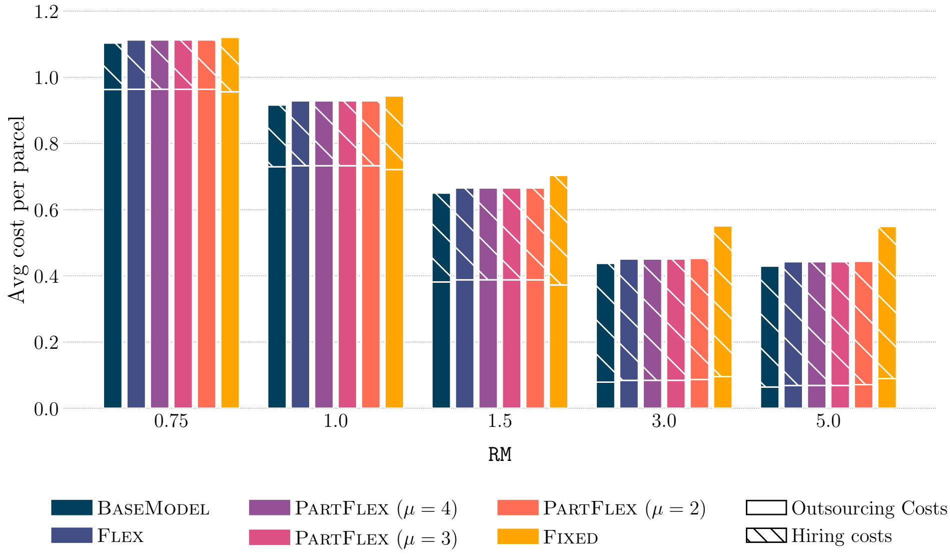

Figure 6: Impact of the `RM` parameter on the cost per parcel.

models is so small that the costs are identical up to the second decimal digit. Indeed, in the vast majority of the instances, the solutions obtained by Flex are identical to those obtained, e.g., by PartFlex ($\mu = 3$). This observation supports the conclusion that limiting shift instability is a viable strategy to reconcile the company's bottom line with the workers' well-being, because high shift stability can be achieved with a negligible cost increase..

The right plot in Figure 5 also refers to the costs per parcel. The height of each bar corresponds to the average cost over all instances, which we split into its hiring and outsourcing components. As we will see in the following, the ratio of each component in the total costs depends on many factors, the main one being the unit outsourcing cost. This is an exogenous market characteristic, and, in our instances, we only assume that outsourcing a delivery is more expensive than performing it in-house. We control the unit outsourcing cost more precisely through parameter `OC`. Because, in this plot, each bar shows the average over all instances, we cannot appreciate the impact of `OC`. Still, this figure shows that—even in aggregate—using fixed shifts results in sensibly higher hiring costs compared with the other models. When scheduling flexibility is very limited, optimal solutions use a larger workforce and incur higher hiring costs but do not significantly differ in terms of outsourcing costs.

4.2.2 Detailed cost analysis

In the following subsection, we evaluate the effect of the instance parameters that most impact the cost structure of the logistic provider.

Figure 6 shows how the costs change with the regional bound parameter `RM`. Recall that this bound limits the workforce size at a regional level and models external constraints, such as the fleet size, that prevent a decision-maker from hiring too many couriers. As expected, relaxing this bound results in lower costs, and the cost decrease is significant. Therefore, decision-makers who find themselves bound by fleet capacity should look into mid or long-term fleet expansion rather than consistently relying on outsourcing.

We also remark that the cost structure also changes when `RM` changes. When the bound is lax (high `RM`), Fixed gives relatively larger costs compared to the other models. When the bound is tight (low `RM`), on the contrary, the cost per parcel is high but similar for all the models. A tight bound means that the operator is subject to structural constraints such as a small fleet or a workforce shortage (i.e., it is understaffed) that are commonly associated with tight or even negative profit margins. Contrary to common intuition, such an operator would

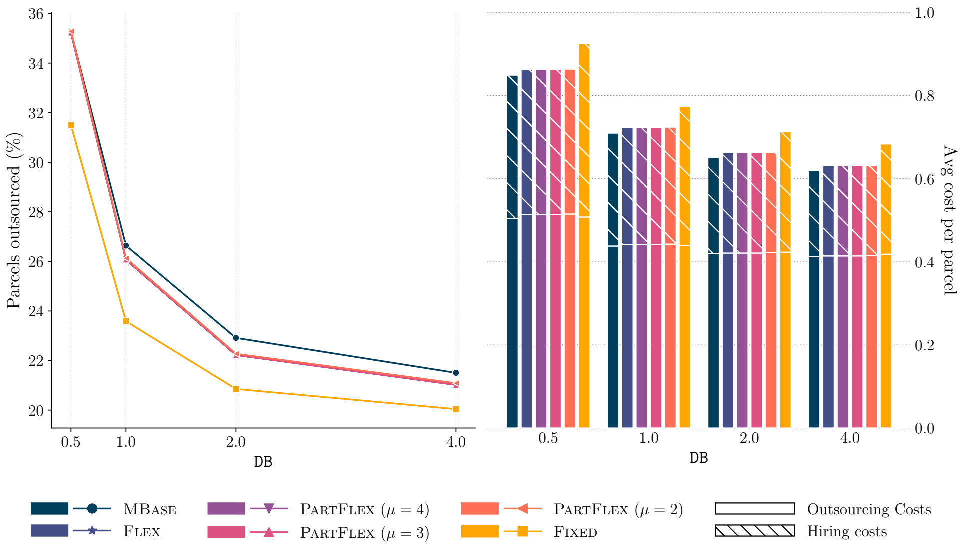

Figure 7: Impact of the DB parameter on the number of outsourced parcels and costs.

not get a large advantage by moving to more flexible schedules; instead, it would compound burnout from understaffing and overwork with decreased job quality due to shift instability. The relationship between overwork, instability, burnout, and mental health has been studied in the medical and social science literature, especially in relation to nurses, doctors, and healthcare workers. These categories have been historically known to be subject to long and intense work hours, demanding both on the physical and emotional levels. Still, the results shown in Figure 6 suggest that future studies should not neglect LMD workers. More specifically, minor cost increases can achieve better working conditions. Thus, companies that do not want to sacrifice their bottom line could still act in favour of workers' well-being, given the negligible impact on profitability. We believe this is a strong argument and a call to action towards a more conscious relationship between service sector companies and their workers. Indeed, some recent research has started to focus on these categories (see, e.g., Pyo et al. 2023; Wei et al. 2023; Couve, Lam, and Verlinghieri 2023).

Figure 7 shows the impact of the baseline demand parameter DB. The left figure reports the percentage of outsourced parcels, and the right one gives a breakdown of the costs. When demand is low, optimal solutions tend to use more outsourcing (left figure). In this scenario, in fact, a large workforce would be idle for a considerable portion of the time, and outsourcing becomes a more attractive option. This consideration holds for all models, including MBase. When demand grows, the logistic operator outsources fewer deliveries, up to the point when the capacity of the in-house delivery system is reached and the curves in the left figure start flattening. The Fixed model tends to keep a large workforce and, therefore, requires less outsourcing compared to the other models.

The right figure supports three main points. First, as in most other businesses, LMD shows economies of scale, and the cost per parcel decreases when the volume increases. Second, these efficiency gains incur diminishing returns; e.g., doubling DB from 0.5 from 1 produces a larger cost decrease than doubling from 2 to 4. Third, the difference in costs is usually larger between instances with different values of DB than it is between models for a given value of DB. Model Fixed displays higher costs per parcel even for larger volumes (for example, Fixed's costs when $\mathtt{DB} = 4$ are higher than MBase's costs when $\mathtt{DB} = 2$), but this is the exception rather than the norm. The other models tend to have more similar costs, hinting at the fact that increasing flexibility can only partially mitigate underlying problems such as low demand.

The impact of the outsourcing costs, controlled by parameter OC, is reported in Figure 8. The higher the costs, the lower the number of outsourced parcels; the left plot in the figure,

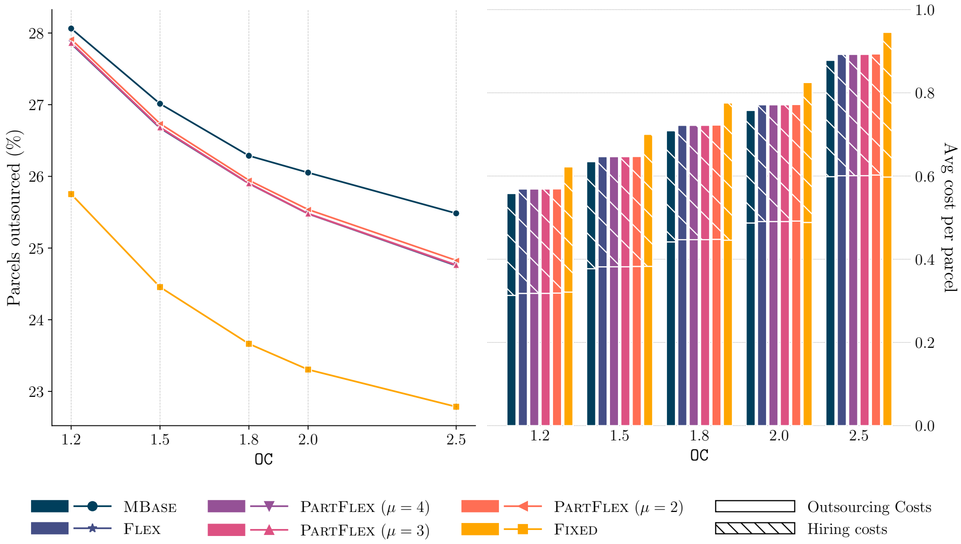

Figure 8: Impact of the `OC` parameter on the number of outsourced parcels and costs.

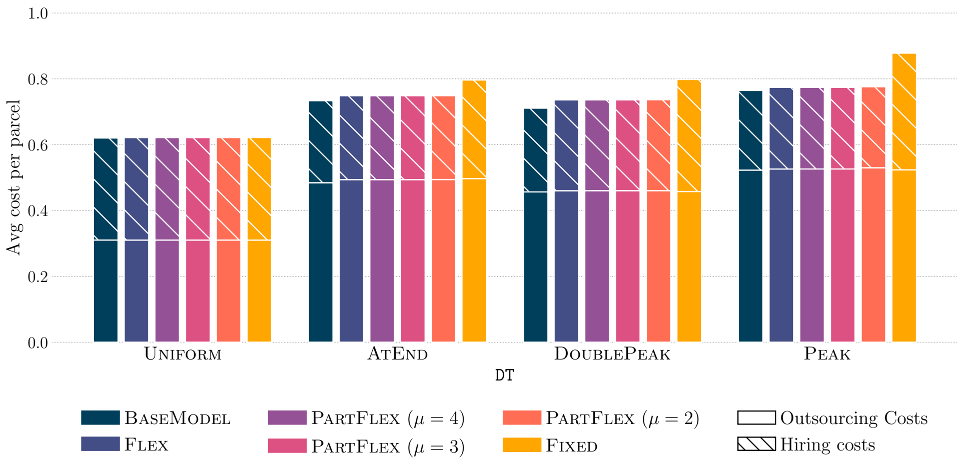

Figure 9: Impact of the `DT` parameter on the costs.

however, shows that the relation is not linear. Although the effect is less pronounced than in Figure 7, we see that the curves representing the percentage of outsourced parcels tend to flatten when reaching the limit of parcels that in-house couriers can deliver. The right figure shows that increasing the outsourcing costs causes a slight increase in the hiring costs (because it is convenient to employ more in-house couriers) and a large increase in the outsourcing costs.

Finally, in Figure 9, we focus on the impact of the demand type (parameter `DT`) on the costs. When the demand is Uniform, costs are low, no flexibility is required, and all models yield roughly the same costs. Furthermore, in a uniform demand scenario, the optimal number of couriers to hire does not change with the period. Therefore, the problem reduces to find *the* best workforce size. This scenario is similar to the classical newsvendor problem: the workforce size is analogous to the order quantity, the number of couriers required to deliver all parcels is the stochastic demand, and the cost difference between outsourcing and in-house delivery is the opportunity cost. Indeed, the cost structure shown in Figure 9 also shows a remarkable similarity with the optimal newsvendor solution: optimal solutions are characterised by equal outsourcing and staffing costs, i.e., the solid and the hatched bars have the same height.

The more challenging demand type is Peak, i.e., when the highest parcel volume occurs in the middle of the day. Model Fixed uses two shifts dividing the daily planning horizon into two

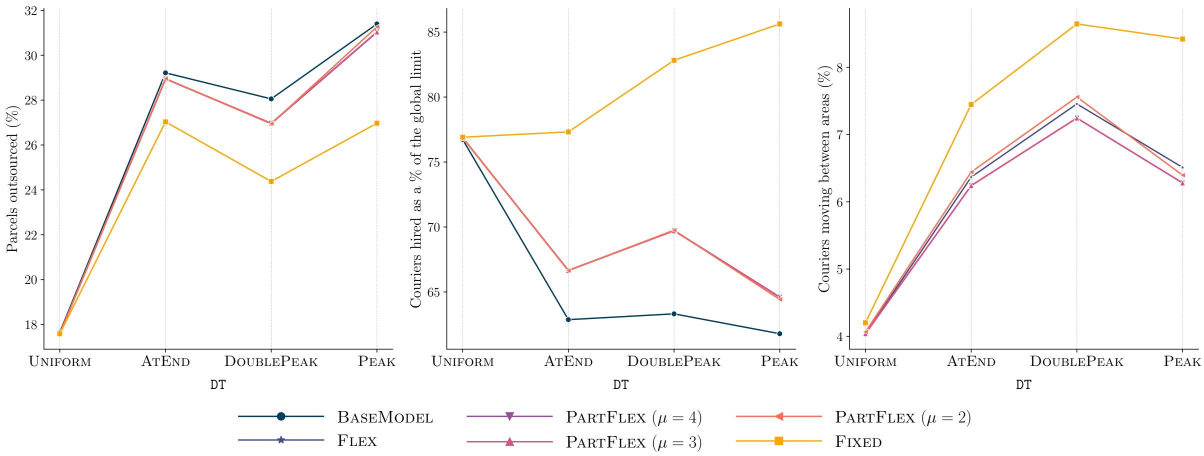

Figure 10: Impact of the DT parameter on three company operations metrics.

halves and, as a result, is the least suitable to deal with this demand pattern. Still, we remark that the cost difference between the three types, AtEnd, DoublePeak, and Peak, is small and that the cost structure is similar. These facts suggest that the conclusions that we draw in this analysis are valid for a diverse range of demand types and could be generalised beyond the patterns that we study in this paper.

4.2.3 Impact of flexibility on operations

In this section, we study the impact that the instance generation parameters and the different models have on key indicators of the company's operational practices. The first indicator, which we already presented in Figures 7 and 8, is the percentage of outsourced parcels. The second is the number of couriers hired as a percentage of the global limit u. The third is the percentage of couriers who change area at the end of each period. This last indicator is used as a proxy of the operational complexity and is related to the geographical stability concept discussed in Section 2.3.

Figure 10 shows the impact of the demand type on the three metrics mentioned above. When the demand is Uniform, as we have seen for the costs, all models show similar characteristics. However, there are considerable differences between the other demand types. When using Fixed shifts, the percentage of employed couriers increases with non-Uniform demand types. The Flex and PartFlex models, on the other hand, can take advantage of the fact that demand concentrates during some periods of the day (and is much lower during the other periods) and require hiring overall fewer couriers. In particular, the Flex model keeps the number of hired couriers significantly lower compared with the PartFlex models.

In general, all models only require a modest amount of area changes; in this respect, demand type DoublePeak is the most demanding. Still, this metric should be considered with care because, in our instance generation procedure, the number of parcels to deliver in each area is only proportional to its population. Logistic operators might have access to more detailed data, which could exacerbate the demand difference between areas. For example, residential areas might require more deliveries in the late afternoon, while commercial areas could have a higher demand during the mornings.

This figure also shows that the Fixed model is an outlier that not only yields higher costs but also requires significantly different operational choices. On the other hand, the Flex and PartFlex models are not too dissimilar, especially when considering the number of parcels outsourced and the area movements required.

Figure 11 reports the percentage of outsourced parcels as a function of the global (left) and regional (right) multipliers of the workforce size upper bounds. First, we note that the

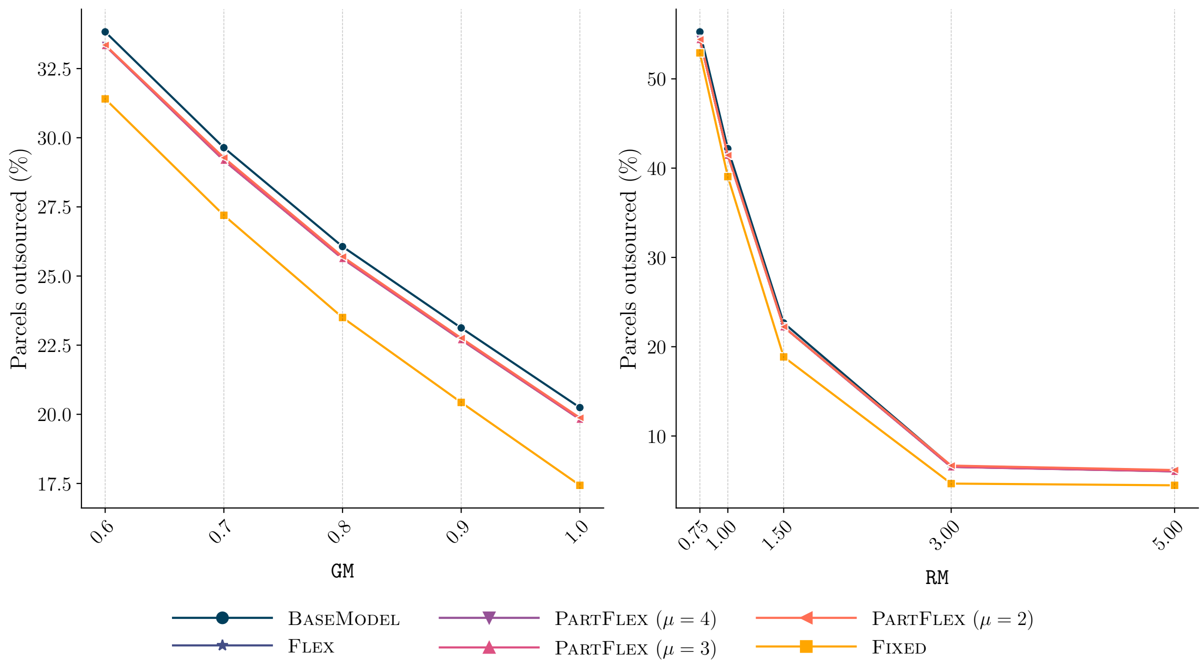

Figure 11: Impact of the GM and RM parameters on the number of outsourced parcels.

two multipliers impact this metric differently. The parameter GM decreases the number of outsourced parcels almost linearly, while the relationship between parameter RM and the number of outsourced parcels is non-linear. Furthermore, even when the bounds are large, the optimal amount of outsourcing is strictly positive (around 20% for the highest value of GM and around 7% for the highest value of RM). Indeed, we repeated our computational experiments by completely removing constraints (7b) and (7c), and we found that the average percentage of outsourced parcels ranged between 4.48% for Fixed and 6.18% for PartFlex ($\mu = 2$). This shows that outsourcing can be economically convenient to balance fixed and variable costs, even when there is no tight bound on the workforce size.

Finally, Figure 12 shows how the model and the outsourcing costs affect the couriers' mobility between areas at each period. We do not report results relative to model MBase because it does not include the y variables necessary to track courier movements between areas. The distribution of this indicator is skewed and shows large right tails: whereas the medians are all low and similar to each other, the left figure shows that a part of the distribution reaches values of over 25%. The right figure, which is on a different scale compared to the left one, further shows that the means exhibit a larger variation, with model Fixed requiring more courier movement compared to the other models, outlining a trade-off between temporal and geographical stability. Furthermore, when outsourcing is more expensive, in the majority of cases, the company responds by increasing the geographical mobility to better adapt to the demand and outsource less. The Fixed model is an exception in this respect because when OC is high, this model uses a larger workforce that is sometimes idle and requires less repositioning. On the other hand, models that use the courier workforce more also need more movement between areas.

4.3 Robustness to changes in demand types

At the operational level, the demand distribution (identified by the parameter DT in our instances) can change on specific days. For example, different patterns can be observed during weekdays and weekends or during the holiday season. If these differences are predictable, a decision-maker can solve an instance of our problem per each expected demand pattern. Operational decisions, such as rostering, can help reconcile the different solutions. For example, if the weekday problem

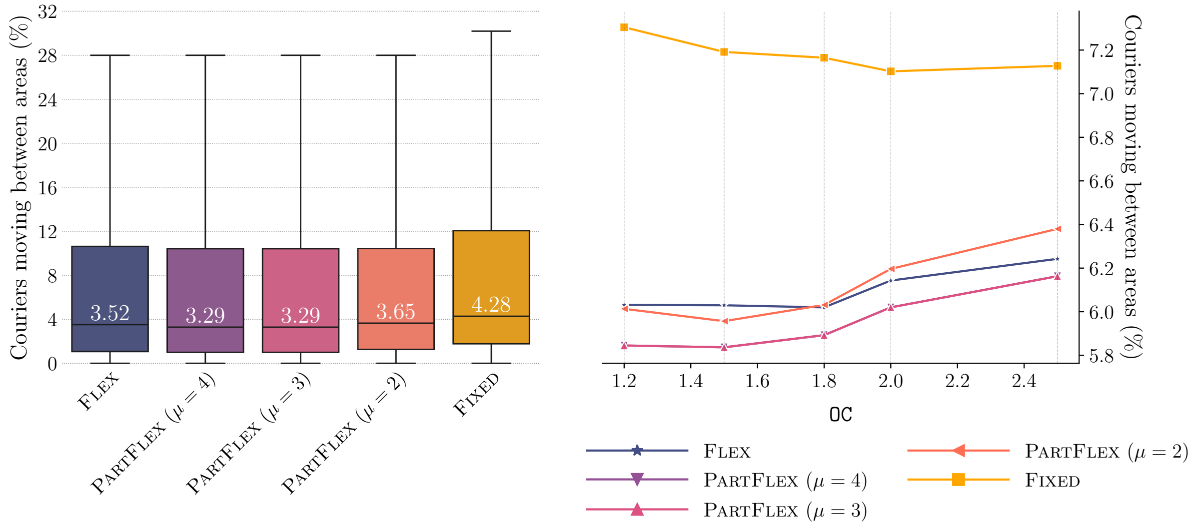

Figure 12: Impact of the model and outsourcing costs (`OC`) on the average number of couriers changing areas at the end of each period. *Note:* the right and the left plot have different y-axis limits.

requires 20 couriers and the weekend one only requires 10, the couriers' roster can exploit this difference to schedule appropriate weekly rest days.

If the variations in demand patterns are more unpredictable, the logistic operator will sometimes have to use a strategic solution devised for a given demand type with a different realised demand type. In this section, we investigate how robust solutions are to changes of the `DT` parameter. More precisely, we want to estimate how much efficiency a decision-maker would lose if they sized and scheduled their workforce for a given value of `DT`, but then a different demand type was realised.

To answer this question, we ran the following experiment. For each combination of parameters, we fix the corresponding solution and we evaluate its cost on instances with the same parameters—city, `DB`, `OC`, `RM` and `GM`—but different demand type `DT`. We do so by fixing variables x_{at} to the value they take in the optimal solution of an instance with a given demand type $\mathtt{DT}_1$, and, keeping these variables fixed, we recompute the solution cost on instances with different demand types $\mathtt{DT}_2 \neq \mathtt{DT}_1$.

Figure 13 reports the results of this experiment. The figure shows a heatmap for each of the six models. The value $\mathtt{DT}_1$ used to fix variables x_{at} is on the x-axis, and the value $\mathtt{DT}_2$ used to evaluate the solution cost is on the y-axis. The values reported in the heatmap are the average percent cost increases over all instances sharing the same values for parameters city, `DB`, `OC`, `RM` and `GM`. For example, the value corresponding to PEAK on the x-axis and ATEND on the y-axis reports the average cost increase incurred when using a solution devised for the PEAK demand type when the actual demand distribution follows the ATEND pattern and all other instance generation parameters are the same.

Figure 13 prompts the following observations. On the one hand, when the realised demand is UNIFORM, solutions obtained according to other demand types are particularly ineffective (first row of each heatmap). Indeed, for a given area, any solution deploying couriers in a pattern that deviates from a constant value throughout the day is inefficient because demand *is* constant throughout the day. On the other hand, PEAK and DOUBLEPEAK cause the smallest cost increases when the realised demand is one of these `DT`s and the solution was obtained using the other. Furthermore, and most important for our analysis, we remark that the average percent cost increases for the FIXED model are generally the lowest. (The only exception is when using a solution for PEAK on a DOUBLEPEAK realised demand, in which case the BASEMODEL

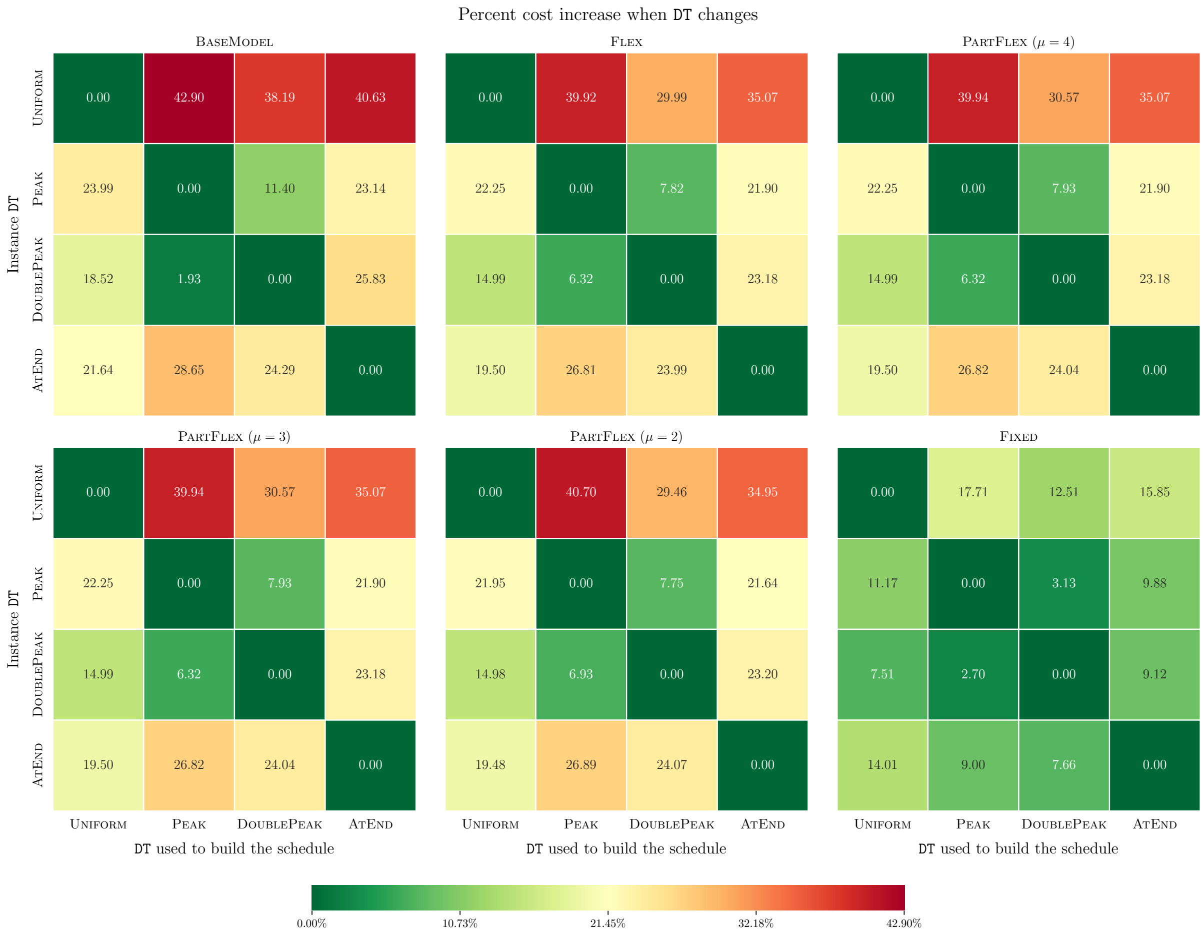

Figure 13: Impact of changing the demand type of an instance (parameter DT) on the cost of solutions obtained optimising for a different demand type.

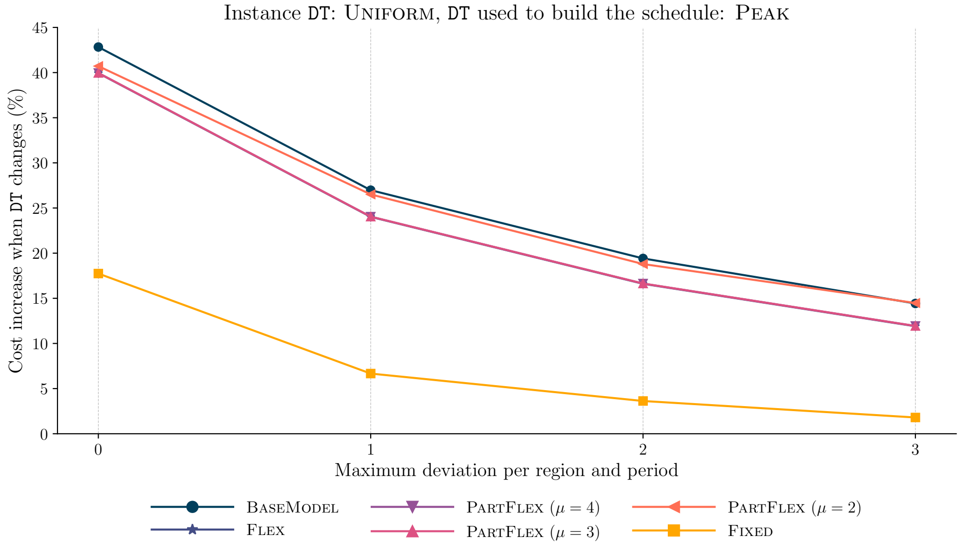

Figure 14: Impact of allowing more flexibility to adjust a predetermined solution when the demand type DT changes at the operational level. In this case study, the operational DT is UNIFORM and the DT used to obtain the solution is PEAK.

shows a smaller cost increase.) In other words, using completely fixed shifts is more expensive but more resilient to changes in demand patterns. The excess courier capacity that is usually present in the solutions of the FIXED model (see the central plot in Figure 10) acts as a buffer that helps deal with demand changes.

The above analysis relies on the assumption that the solution devised for a given value of DT must be used without modifications when the realised DT is different. In practical circumstances, the decision-maker might deviate, on the operational level, from the strategic problem solution. For example, they might increase the number of couriers working during some periods by using overtime. In this case, couriers might be subject to sudden roster variation, a practice that negatively affects their work-life balance. While the impact of deviating from the schedule on workers' well-being is clear, it is less clear whether these deviations help reduce costs. To answer this question, we conducted a small case study based on the demand pattern variation that seems most challenging for all models, i.e., using a solution devised for DT PEAK when the realised DT is UNIFORM. In the case study, we allow the value of variables x_{at} to deviate by at most $\delta \in \{0, \dots, 3\}$ from the value taken in the optimal solution for DT $=$ PEAK. When $\delta = 0$, the decision maker cannot deviate from the solution, and we recover the case studied in Figure 13. For $\delta > 0$, we allow the decision maker to deploy more or fewer couriers in each area and period, up to a difference of $\pm\delta$. Figure 14 shows the results of this experiment. Whereas allowing more deviations reduces costs, the FIXED model remains the most competitive in terms of resiliency. Furthermore, the most flexible models (BASEMODEL and FLEX) with an allowed deviation of $\delta = 3$ produce cost increases barely smaller than the FIXED model with $\delta = 0$. And when $\delta = 2$, all other models produce a larger relative cost increase compared with the FIXED model with $\delta = 0$.

5 Conclusions

In this paper, we have tackled the problem of tactical hiring and scheduling decisions for a company performing parcel deliveries in the last-mile segment of the supply chain. In particular, we have developed mathematical models to determine the correct number of couriers to hire to balance salary costs and outsourcing costs. These latter are paid when the company do not hire enough couriers to deliver all parcels. We have put a particular emphasis on the topic of shift stability, i.e., devising shifts with a predictable start time and duration. In doing so, we wanted to explore if flexible shifts, which decrease job satisfaction and disrupt the couriers' work-life balance, are justified by large savings. Our main conclusions are the following:

- Using completely fixed shifts that start at two predetermined times during the day results in significantly higher costs. Over all instances, the average per-parcel cost obtained using fixed shifts is 9.36% higher than the one obtained using extremely flexible schedules, in which couriers can be called into (and out of) at each two-hour period.
- A partially flexible model (PartFlex) that uses two shifts—but allows their start times to be a decision variable—incurs costs that are only 1.89% higher than those obtained with extremely flexible schedules.
- The advantage of flexible schedules compared to fixed ones is more significant when the company can hire a large workforce and has a large fleet. When the company cannot hire many couriers (because it does not have enough vehicles to operate or because market conditions make labour scarce), flexible and fixed schedules yield almost the same costs. The conclusion is that stable shifts are a viable strategy for a company that has trouble finding couriers. Stable shifts do not significantly increase costs, and they provide better working conditions that help attract potential employees.
- As in many industries, we observed economies of scale. When the volume increases, the cost per parcel decreases, although with diminishing returns.
- Predictable demand patterns, such as having one or two daily peak times, do not have a large impact on the total costs, but they significantly change some aspects of the company's operations. For example, instances with two daily peaks require more couriers and more courier movements between geographical areas, compared with instances with a single peak (either in the central or the later part of the day).
- When demand patterns are less predictable, the logistic operator will have to use a schedule optimised for a given pattern in days when the realised pattern is different. Solutions featuring stable shifts are particularly suitable in this case and provide the smallest cost increases. In other words, stable shifts are generally more expensive to implement but more resilient to sudden changes in demand patterns.

We conclude by remarking that our work relies on a number of assumptions and that our conclusions are based on computational results over synthetic instances. At the same time, we tested our approach on a variety of instance generation parameters, and we observed significant results consistently over a large number of instances. This suggests that a compromise approach that features limited schedule flexibility deserves further analysis, especially from larger logistic providers using scientific management approaches to optimise their tactical and operational planning decisions.

Concerning future research directions, a natural continuation of this work involves rostering decisions. Rostering involves specifying detailed working hours for each worker and day based on the shift schedules determined with the methodology presented in this work. In addition, it might be interesting to integrate districting decisions with shift scheduling. This would lead to a more sophisticated problem, but the increase in complexity might be compensated by benefits in terms of economic convenience and work stability. Finally, we argue that there is a need to put more emphasis on social welfare by considering workers' well-being as the main objective in planning decisions, extending our work to problems arising in fields beyond LMD.

Acknowledgements

The work of Alberto Santini was partly funded by the Ministry of Science and Innovation of Spain under the "Ramon y Cajal" grant with number MCIN/AEI/10.13039/501100011033 and by the European Social Fund Plus (FSE+).

References

Alfonso, Viviana et al. (2021). *E-commerce in the pandemic and beyond.* Bulletin 36. Bank for International Settlements. URL: `https://web.archive.org/web/20210126163639/https://www.bis.org/publ/bisbull36.htm` (visited on 08/02/2023).

Ananat, Elizabeth and Anna Gassman-Pines (2021). "Work Schedule Unpredictability: Daily Occurrence and Effects on Working Parents' Well-Being". In: *Journal of Marriage and Family* 83 (1), pp. 10–26. DOI: `10.1111/jomf.12696`.

Arrieta-Prieto, Mario et al. (2022). "Location of urban micro-consolidation centers to reduce the social cost of last-mile deliveries of cargo: A heuristic approach". In: *Networks* 79.3, pp. 292–313. DOI: `10.1002/net.22076`.

Banerjee, Dipayan, Alan L Erera, and Alejandro Toriello (2022). "Fleet sizing and service region partitioning for same-day delivery systems". In: *Transportation Science* 56 (5), pp. 1327–1347. DOI: `10.1287/trsc.2022.1125`.

Behrendt, Adam, Martin Savelsbergh, and He Wang (2023). "A Prescriptive Machine Learning Method for Courier Scheduling on Crowdsourced Delivery Platforms". In: *Transportation Science* 57.4, pp. 889–907. DOI: `10.1287/trsc.2022.1152`.

Bellenguez, Odile, Nadia Brauner, and Alexis Tsoukiàs (2023). "Is there an ethical operational research practice? And what this implies for our research?" In: *EURO Journal on Decision Processes* 11. DOI: `https://doi.org/10.1016/j.ejdp.2023.100029`.

Bergman, Alon, Guy David, and Hummy Song (2023). ""I Quit"': Schedule Volatility as a Driver of Voluntary Employee Turnover". In: *Manufacturing & Service Operations Management* 25.4, pp. 1416–1435. DOI: `10.1287/msom.2023.1205`.

Boysen, Nils, Stefan Fedtke, and Stefan Schwerdfeger (2021). "Last-mile delivery concepts: a survey from an operational research perspective". In: *OR Spectrum* 43, pp. 1–58. DOI: `10.1007/s00291-020-00607-8`.

Carlsson, John Gunnar et al. (2024). "Provably Good Region Partitioning for On-Time Last-Mile Delivery". In: *Operations Research* 72 (1), pp. 91–109. DOI: `10.1287/opre.2021.0588`.

Carrillo, Dani et al. (2017). "Instability of Work and Care: How Work Schedules Shape Child-Care Arrangements for Parents Working in the Service Sector". In: *Social Service Review* 91.3, pp. 422–455. DOI: `10.1086/693750`.

Castillo, Vincent et al. (2022). "Hybrid last-mile delivery fleets with crowdsourcing: A systems view of managing the cost-service trade-off". In: *Journal of Business Logistics* 43 (1), pp. 36–61. DOI: `10.1111/jbl.12288`.

Choper, Joshua, Daniel Schneider, and Kristen Harknett (2022). "Uncertain Time: Precarious Schedules and Job Turnover in the US Service Sector". In: *ILR Review* 75.5, pp. 1099–1132. DOI: `10.1177/00197939211048484`.

Chung, Hyesook (2022). "Variable work schedules, unit-level turnover, and performance before and during the COVID-19 pandemic". In: *Journal of Applied Psychology* 107 (4), pp. 515–532. DOI: `10.1037/apl0001006`.

Contrive Datum Insights (2023). *First and Last Mile Delivery Market Size, Share & Trends Estimation Report.* Tech. rep. 44470. Contrive Datum Insights. URL: `https://www.contrivedatuminsights.com/product-report/first-and-last-mile-delivery-market-44470` (visited on 08/02/2023).

Couve, Charlie, Tiffany Lam, and Ersilia Verlinghieri (2023). *Delivering Good Work: Labour, employment and wellbeing in London's cargo bike sector*. Tech. rep. University of Westminster. URL: https://blog.westminster.ac.uk/ata/wp-content/uploads/sites/60/2023/06/221222_Delivering-good-work_Londons-Cargo-Bike-secton_V6.pdf.

Crainic, Teodor Gabriel et al. (2021). *Operations research for planning and managing city logistics systems*. Tech. rep. 45. CIRRELT. URL: https://inria.hal.science/hal-03464029/document (visited on 08/02/2023).

Dai, Hongyan and Peng Liu (2020a). "Workforce planning for O2O delivery systems with crowdsourced drivers". In: *Annals of Operations Research* 291, pp. 219–245. DOI: 10.1007/s10479-019-03135-z.

Dai, Hongyan and Peng Liu (2020b). "Workforce planning for O2O delivery systems with crowdsourced drivers". In: *Annals of Operations Research* 291, pp. 219–245. DOI: 10.1007/s10479-019-03135-z.

Delle Donne, Diego, Laurent Alfandari, et al. (2023). "Freight-on-Transit for urban last-mile deliveries: A Strategic Planning Approach". In: *Transportation Research Part B: Methodological* 169, pp. 53–81. DOI: 10.1016/j.trb.2023.01.004.

Delle Donne, Diego, Alberto Santini, and Claudia Archetti (2024). "Integrating Public Transport in Sustainable Last-Mile Delivery: Column Generation Approaches". In: *Optimization Online*. URL: https://optimization-online.org/?p=25263.

Dickson, Alison, Lonnie Golden, and Robert Bruno (2018). *Scheduling Stability: The Landscape of Work Schedules and Potential Gains From Fairer Workweeks in Illinois and Chicago*. Tech. rep. 3172354. Social Science Research Network. DOI: 10.2139/ssrn.3172354.

Ertogral, Kadir, Ayse Akbalik, and Silvia González (2017). "Modelling and analysis of a strategic fleet sizing problem for a furniture distributor". In: *European Journal of Industrial Engineering* 11.1, pp. 49–77. DOI: 10.1504/EJIE.2017.081428.

Fatehi, Soraya and Michael Wagner (2022). "Crowdsourcing Last-Mile Deliveries". In: *Manufacturing & Service Operations Management* 24 (2), pp. 791–809. DOI: 10.1287/msom.2021.0973.

Figliozzi, Miguel Andres (2008). "Planning approximations to the average length of vehicle routing problems with varying customer demands and routing constraints". In: *Transportation Research Record* 2089.1 (1), pp. 1–8. DOI: 10.3141/2089-01.

Franceschetti, Anna et al. (2017). "Strategic fleet planning for city logistics". In: *Transportation Research Part B: Methodological* 95, pp. 19–40.

Goeke, Dominik, Roberto Roberti, and Michael Schneider (2019). "Exact and Heuristic Solution of the Consistent Vehicle-Routing Problem". In: *Transportation Science* 53.4, pp. 1023–1042. DOI: 10.1287/trsc.2018.0864.

Gonzalez-Feliu, Jesus (2012). "Freight distribution systems with cross docking: A multidisciplinary analysis". In: *Journal of the Transportation Research Forum* 51.1. DOI: 10.5399/osu/jtrf.51.1.2821.

Goulart, Pablo et al. (2021). "Programação linear para dimensionamento de frota de un operador logístic (Linear programming for fleet sizing of a logistic operator)". In: *Revista Eletrônica de Estratégia e Negócios* 14 (1), pp. 60–84. DOI: 10.19177/reen.v14eI202160-84.

Groër, Chris, Bruce Golden, and Edward Wasil (2009). "The Consistent Vehicle Routing Problem". In: *Manufacturing & Service Operations Management* 11.4, pp. 630–643. DOI: 10.1287/msom.1080.0243.

Harknett, Kristen, Daniel Schneider, and Sigrid Luhr (2022). "Who Cares if Parents have Unpredictable Work Schedules?: Just-in-Time Work Schedules and Child Care Arrangements". In: *Social Problems* 69 (1), pp. 164–183. DOI: 10.1093/socpro/spaa020.

Henly, Julia and Susan Lambert (2014). "Unpredictable Work Timing in Retail Jobs: Implications for Employee Work–Life Conflict". In: *ILR Review* 67.3, pp. 986–1016. DOI: 10.1177/0019793914537458.

Henly, Julia, Luke Shaefer, and Elaine Waxman (2006). "Nonstandard Work Schedules: Employer- and Employee-Driven Flexibility in Retail Jobs". In: *Social Service Review* 80.4, pp. 609–634. DOI: 10.1086/508478.

Kleywegt, Anton, Alexander Shapiro, and Tito Homem-de-Mello (2002). "The Sample Average Approximation Method for Stochastic Discrete Optimization". In: *SIAM Journal on Optimization* 12.2, pp. 479–502. DOI: 10.1137/S1052623499363220.

Kovacs, Attila et al. (2014). "Vehicle routing problems in which consistency considerations are important: A survey". In: *Networks* 64 (3), pp. 192–213. DOI: 10.1002/net.21565.

Kunz, Nathan and Luk Van Wassenhove (2019). "Fleet sizing for UNHCR country offices". In: *Journal of Operations Management* 65 (3), pp. 282–307. DOI: 10.1002/joom.1013.

Laporte, Gilbert and Fran cois Louveaux (1993). "The integer L-shaped method for stochastic integer programs with complete recourse". In: *Operations Research Letters* 13.3, pp. 133–142. DOI: 10.1016/0167-6377(93)90002-X.

Le Menestrel, Marc and Luk Van Wassenhove (2009). "Ethics in Operations Research and Management Sciences: A never-ending effort to combine rigor and passion". In: *Omega* 37.6, pp. 1039–1043. DOI: https://doi.org/10.1016/j.omega.2008.12.009.

Liu, Sheng, Long He, and Max Sheng Zuo-Jun (2021). "On-Time Last-Mile Delivery: Order Assignment with Travel-Time Predictors". In: *Management Science* 67.7, pp. 4095–4119. DOI: 10.1287/mnsc.2020.3741.

Loxton, Ryan and Quin Lin (2011). "Optimal fleet composition via dynamic programming and golden section search". In: *Journal of Industrial and Management Optimization* 7 (4), pp. 875–890. DOI: 10.3934/jimo.2011.7.875.

Mandal, Minakshi Punam, Alberto Santini, and Claudia Archetti (2024). *Instance generator for the Tactical LMD Staffing Problem.* DOI: 10.5281/zenodo.10561003. URL: https://github.com/alberto-santini/lmd-sizing-scheduling (visited on 01/24/2024).

Mohan, Srimathy (2008). "Scheduling part-time personnel with availability restrictions and preferences to maximize employee satisfaction". In: *Mathematical and Computer Modelling* 48.11, pp. 1806–1813. DOI: 10.1016/j.mcm.2007.12.027.

Monteiro Ferraz, Artur, Quentin Cappart, and Thibaut Vidal (2022). "Deep Learning for Data-Driven Districting and Routing". In: *Proceedings of the 11^{th} Triennial Symposium on Transportation Analysis*, pp. 1–4. URL: https://tristan2022.org/Papers/TRISTAN_2022_paper_0727.pdf.

OpenStreetMap (2023). *Copyright and license.* URL: https://www.openstreetmap.org/copyright (visited on 10/09/2023).

Ormerod, Richard and Ulrich Werner (2013). "Operational research and ethics: A literature review". In: *European Journal of Operational Research* 228.2, pp. 291–307. DOI: https://doi.org/10.1016/j.ejor.2012.11.048.

Pandey, Peeyush, Hasmukh Gajjar, and Bhavin Shah (2021). "Determining optimal workforce size and schedule at the retail store considering overstaffing and understaffing costs". In: *Computers & Industrial Engineering* 161. Article ID 107656. DOI: 10.1016/j.cie.2021.107656.

Pyo, Jeehee et al. (2023). "How has COVID-19 affected the work environment of delivery workers?: An interpretative phenomenological analysis". In: *Plos One* 18 (9). DOI: 10.1371/journal.pone.0290403.

Rahimi-Vahed, Alireza et al. (2015). "Fleet-sizing for multi-depot and periodic vehicle routing problems using a modular heuristic algorithm". In: *Computers & Operations Research* 53, pp. 9–23. DOI: 10.1016/j.cor.2014.07.004.

Restrepo, María, Frédéric Semet, and Thomas Pocreau (2019). "Integrated Shift Scheduling and Load Assignment Optimization for Attended Home Delivery". In: *Transportation Science* 53 (4), pp. 917–1212. DOI: 10.1287/trsc.2018.0857.

Rhee, Min-Kyoung, Soo Kyung Park, and Chung-Kwon Lee (2020). “Pathways from workplace flexibility to turnover intention: Role of work–family conflict, family–work conflict, and job satisfaction”. In: *International Journal of Social Welfare* 29.1, pp. 51–61. DOI: https://doi.org/10.1111/ijsw.12382.

Rodríguez-Martín, Inmaculada, Juan-José Salazar-González, and Hande Yaman (2019). “The periodic vehicle routing problem with driver consistency”. In: *European Journal of Operational Research* 273.2, pp. 575–584. DOI: 10.1016/j.ejor.2018.08.032.

Romanillos, Gustavo and Javier Gutiérrez (2020). “Cyclists do better. Analyzing urban cycling operating speeds and accessibility”. In: *International Journal of Sustainable Transportation* 14 (6), pp. 448–464. DOI: 10.1080/15568318.2019.1575493.

Ruiz-Torres, Alex et al. (2015). “Scheduling to maximise worker satisfaction and on-time orders”. In: *International Journal of Production Research* 53.9, pp. 2836–2852. DOI: 10.1080/00207543.2015.1005764.

Santini, Alberto et al. (2022). “The probabilistic Travelling Salesman Problem with crowdsourcing”. In: *Computers & Operations Research* 142. Article ID 105722. DOI: 10.1016/j.cor.2022.105722.

Schiavina, Marcello et al. (2023). *GHS-POP R2023A - GHS population grid multitemporal (1975-2030).* DOI: 10.2905/2FF68A52-5B5B-4A22-8F40-C41DA8332CFE. URL: https://ghsl.jrc.ec.europa.eu/ghs_pop2023.php (visited on 10/09/2023).

Schneider, Michael (2016). “The vehicle-routing problem with time windows and driver-specific times”. In: *European Journal of Operational Research* 250 (1), pp. 101–119. DOI: 10.1016/j.ejor.2015.09.015.

Shehadeh, Karmel, Hai Wang, and Peter Zhang (2021). “Fleet sizing and allocation for on-demand last-mile transportation systems”. In: *Transportation Research Part C: Emerging Technologies* 132. DOI: 10.1016/j.trc.2021.103387.

Smilowitz, Karen, Maciek Nowak, and Tingting Jiang (2013). “Workforce Management in Periodic Delivery Operations”. In: *Transportation Science* 47.2, pp. 214–230. DOI: 10.1287/trsc.1120.0407.

Turan, Hasan Hüseyin et al. (2022). “A joint problem of strategic workforce planning and fleet renewal: With an application in defense”. In: *European Journal of Operational Research* 296 (2), pp. 615–634. DOI: 10.1016/j.ejor.2021.04.010.

Ulmer, Martin and Martin Savelsbergh (2020). “Workforce Scheduling in the Era of Crowdsourced Delivery”. In: *Transportaion Science* 54 (4), pp. 1113–1133. DOI: 10.1287/trsc.2020.0977.

Van den Bergh, Jorne et al. (2013). “Personnel scheduling: A literature review”. In: *European Journal of Operational Research* 226.3, pp. 367–385. DOI: 10.1016/j.ejor.2012.11.029.

Wang, Yang et al. (2022). “Multi-Period Workload Balancing in Last-Mile Urban Delivery”. In: *Transportation Science* 56 (5), pp. 1348–1368. DOI: 10.1287/trsc.2022.1132.

Wei, Hua et al. (2023). *Understanding technostress in the gig economy — A Job Demands-Resources analysis of Chinese couriers.* DOI: 10.1101/2022.09.13.22279887.

Werneck, Renato (Apr. 7, 2022). *Last Mile Deliveries at Amazon.* Keynote at the 12th DIMACS Implementation Challenge Workshop. online: Center for Discrete Mathematics and Theoretical Computer Science. URL: https://www.youtube.com/watch?v=CKJ6zuyq9JU (visited on 08/05/2023).

Yildiz, Baris and Martin Savelsbergh (2019). “Provably High-Quality Solutions for the Meal Delivery Routing Problem”. In: *Transportation Science* 53 (5), pp. 1372–1388. DOI: 10.1287/trsc.2018.0887.

Yura, Kenji (1994). “Production scheduling to satisfy worker’s preferences for days off and overtime under due-date constraints”. In: *International Journal of Production Economics* 33.1, pp. 265–270. DOI: 10.1016/0925-5273(94)90139-2.