Optimizing Last-Mile Delivery: A Dynamic Compensation Strategy for Occasional Drivers

Rouven Schur[1*†] and Kai Winheller[1†]

[1]Mercator School of Management, University of Duisburg-Essen, Lotharstraße, Duisburg, 47084, North Rhine-Westphalia, Germany.

*Corresponding author(s). E-mail(s): rouven.schur@uni-due.de;
Contributing authors: kai.winheller@uni-due.de;
†These authors contributed equally to this work.

Abstract

Amid the rapid growth of online retail, last-mile delivery faces significant challenges, including the cost-effective delivery of goods to all customers. Accordingly, the development and improvement of innovative approaches thrive in current research. Our work contributes to this stream by applying dynamic pricing techniques to effectively model the possible involvement of the crowd in fulfilling delivery tasks. The use of occasional drivers (ODs) as a viable, cost-effective alternative to traditional dedicated drivers (DDs) prompts the necessity to focus on the inherent challenge posed by the uncertainty of ODs' arrival times and willingness to perform deliveries.

We introduce a dynamic programming framework that offers individualized bundles of delivery task and compensation to ODs as they arrive. This model, akin to a reversed form of dynamic pricing, accounts for ODs' decision-making by treating their acceptance thresholds as a random variable. Thereby, our model addresses the dynamic and stochastic nature of OD availability and decision-making.

We analytically solve the stage-wise optimization problem, outline inherent challenges such as the curses of dimensionality, and present structural properties. Designed to cope with these challenges, our approximation methods, a parametric value function approximation and a fluid approximation, aim to accurately determine avoided costs, which are a key factor in calculating optimal compensation.

A comprehensive simulation study compares our algorithms with benchmark strategies, and shows the advantages of dynamic compensation across a range of scenarios. We conclude our work with managerial insights and a summary of our findings, offering significant implications for last-mile delivery operations.

Keywords: Dynamic Pricing, Last-Mile Delivery, Occasional Drivers, Approximate Dynamic Programming

1 Introduction

Fueled by the rapid growth of online retail and brick-and-mortar stores expanding into online sales, last-mile delivery has become increasingly important in the logistics sector. This trend has further accentuated the cost-intensive nature of urban transportation operations, necessitating innovative approaches to mitigate these challenges. One of these innovative approaches is the concept of using occasional drivers (ODs). ODs are (in-store) customers or individuals willing to divert their planned route to deliver online orders for monetary compensation. Thereby, they present a cost-effective alternative to dedicated drivers (DDs), who are employed or contracted through a third-party logistics provider. This OD approach, leveraging existing store traffic, offers not only economic benefits (Archetti et al, 2016; Dayarian and Savelsbergh, 2020) but also aids in reducing pollution and urban traffic, while enhancing community ties (Buldeo Rai et al, 2018; Hutter and Neumann, 2023). However, the unpredictability of ODs, influenced by factors like the frequency of store visits and their fluctuating willingness to undertake deliveries at that day, poses a significant challenge. This uncertainty in OD availability, contrasts sharply with the reliability of DDs, who offer a more predictable delivery solution, albeit at a higher operational cost. A key factor influencing ODs' willingness to participate is compensation, which effectively offsets the additional effort and time ODs invest in fulfilling delivery tasks (Archetti et al, 2016; Qi et al, 2018).
While current implementations of similar concepts on platforms such as Walmart's "Spark" or Shopopop have demonstrated the practical feasibility, there remains a considerable gap to bring out the intended potential of ODs. Specifically, current research has mainly focused on deterministic models with fixed acceptance criteria for ODs (Boysen et al, 2021), with little room for customized offers, while stochastic models typically ignore the dynamic nature of ODs arrivals or the role of compensation in decision-making (Savelsbergh and Ulmer, 2022).
Our research addresses these gaps by proposing a dynamic model that matches arriving ODs with delivery tasks. Each OD is offered a customized bundle, consisting of a delivery task and a compensation. This customized offer considers ODs' destination, the remaining time, and the pool of potential future ODs. The acceptance threshold of the OD is considered a random variable, reflecting the unpredictable nature of their decision-making. Unassigned deliveries at the end of the planning horizon are managed by DDs. This approach, akin to a reversed form of dynamic pricing, is detailed in Figure 1 and offers a more realistic and effective solution for last-mile delivery scenarios.
This paper makes several significant contributions to the field of logistics, utilizing dynamic pricing methodologies:

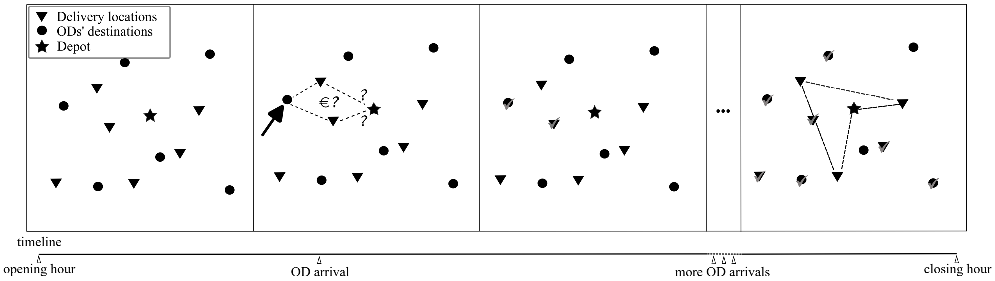

Fig. 1 Dynamic programming process. This illustration depicts the sequential decision-making process as ODs arrive over time. It focuses on the current state, emphasizing three critical elements: the current OD, identified by their destination (represented by an arrow), potential future ODs, marked by their respective destinations (represented by triangles), and open delivery tasks, characterized by their delivery locations (represented by circles).

1. **Introduction of the dynamic compensation problem for occasional drivers:** We propose a novel framework that accounts for the sequential arrival of ODs and incorporates uncertainties related to their arrival and acceptance decisions in Section 3.
2. **Analytical solution to the stage-wise optimization problem:** Utilizing a Bellman equation, we solve the stage-wise optimization problem. Additionally, Section 4 discusses the structural properties and challenges to find an optimal policy, laying the foundation to develop appropriate approximation methods.
3. **Development of advanced algorithms:** In Section 5, we introduce sophisticated algorithms for approximating the value function and determining optimal compensation and delivery task policies for ODs. These include a parametric value function approximation and a fluid approximation, each designed to adapt to the dynamic nature of OD arrivals and reflect structural properties.
4. **Comprehensive simulation study:** Our algorithms are rigorously tested against benchmark strategies in a detailed simulation study, as presented in Section 6. This study not only validates our model but also offers valuable insights into the benefits of dynamic compensation strategies under various urban configurations and OD arrival patterns.

The paper is structured as follows: We begin with a literature review in Section 2. This is followed by a formal problem description in Section 3, and an in-depth analysis of the problem in Section 4. Our proposed solution approaches are detailed in Section 5, and the results of our simulation study are presented in Section 6. We conclude with a discussion of managerial insights and a summary of our key findings in Section 7.

2 Literature review

In the preceding discussion, we conceptualized an OD as an individual who may visit the store (pickup location) and, if conditions align with their schedule, may undertake a delivery task. An important element in our analysis is the unpredictability associated with the arrival of ODs at the store and their decision to accept a task.

We posit that compensations significantly influence ODs' acceptance decision. Based on this definition, we structure our literature review as follows.
Initially, in Section 2.1, we provide a concise overview of existing literature focused on a deterministic framework, which do not explicitly anticipate uncertainties regarding ODs' arrival and task acceptance. Such studies have little in common with our research focus. Moving forward to Section 2.2, our attention turns to scenarios where the arrival and/or acceptance of ODs are stochastic. This section examines models that accommodate the unpredictable behaviors of DS, regarding their arrival at the pickup location or their decision to accept a delivery task. We specifically investigate studies suggesting that these behaviors can be effectively influenced or altered through the compensation strategies. This section of our review aligns more closely with our study.

2.1 Deterministic models

This section is dedicated to deterministic models, concentrating on scenarios where firms possess complete information about ODs, including their constraints such as start and end locations, minimum compensation expectations, and time restrictions. It assumes a constant willingness among ODs to accept delivery tasks if their specified constraints are met, with compensation determined by a predefined scheme.
A cornerstone in this domain is the work by Archetti et al (2016). Their model represents one of the first in this domain, assigning ODs to single delivery tasks within a specified detour limit. This seminal study highlights the potential for cost savings through crowdsourced delivery and emphasizes the importance of a structured compensation scheme. It sets the stage for further research that introduces additional layers of complexity.
Further research has built upon this initial model, integrating time-window constraints and transshipment nodes (Chen et al, 2018; Macrina et al, 2020), individualized minimum compensation thresholds (Dahle et al, 2019), multi-package delivery capabilities (Chen et al, 2018; Boysen et al, 2022), and merging OD delivery with item-sharing (Behrend et al, 2019). Further, Dai and Liu (2020) differentiates between part-time and full-time ODs within these frameworks.
Diverging from fixed compensation schemes, several studies have explored models wherein ODs propose their desired compensation, leaving firms to decide on these proposals. Notable contributions in this area include works by Kafle et al (2017), Allahviranloo and Baghestani (2019), Feng et al (2021), and Mancini and Gansterer (2022). A notable instance is Le et al (2021), who developed an integrated routing and matching approach featuring various compensation schemes, including those aligning with ODs' acceptance thresholds.
Some studies transition to a dynamic setting, albeit maintaining deterministic assumptions by not forecasting future arrivals within the optimization process. This dynamic approach typically employs a rolling horizon method, constantly updating decisions. This approach is exemplified in the works of Allahviranloo and Baghestani (2019), Arslan et al (2019), and Archetti et al (2021).
In conclusion, the deterministic literature demonstrates a clear trajectory towards

enhanced complexity and adaptability. These models are progressively addressing intricate operational demands, varying compensation strategies, and integrating dynamic components while retaining a deterministic essence. This evolution lays a solid foundation for exploring stochastic models, crucial for addressing the uncertainties in more complex operational settings.

2.2 Stochastic models

This segment explores stochastic models that tackle the uncertainties surrounding OD availability and their willingness to participate. This area is predominantly characterized by two-stage problems, where tasks are initially assigned to ODs (first stage), with the anticipation that some of these tasks may remain unfulfilled, requiring alternative solutions, typically involving DDs (second stage). Additionally, several studies focus on an assortment optimization framework, where firms decide on a selection of tasks to offer ODs to influence their participation decisions. Notably, there is also research incorporating dynamic frameworks and, particularly intriguing, studies where firms strategically determine compensations to influence ODs' decisions.

Among the first in considering ODs' acceptance uncertainty were Gdowska et al (2018) who introduce a two-stage optimization model. In this model, tasks are initially assigned to ODs and DDs, including vehicle routing problem solutions, while anticipating potential OD rejections. The chance of a rejection is thereby assumed to be influenced by compensation. Tasks rejected by ODs are reassigned to DDs in the second stage. Further two-stage models have been developed and extended in various studies, including Mousavi et al (2022), which introduces mobile depots, Silva and Pedroso (2022), and Barbosa et al (2023), where compensation decisions are part of the first stage, influencing ODs' availability. Torres et al (2022b) incorporate time-windows, Torres et al (2022a) focus on specific OD destinations, while Silva et al (2023b) and Silva et al (2023a) explore task-specific uncertainties and dynamic arrivals of tasks and ODs, necessitating a dynamic recourse response.

In another research stream, assortment optimization is used to influence ODs' decisions. Mofidi and Pazour (2019) and Horner et al (2021) determine a set of delivery tasks to present to ODs, who then signal their availability for specific tasks, with the firm finalizing the assignment. Ausseil et al (2022) expand this approach to a dynamic setting, accounting for the ongoing arrival of tasks and ODs.

Apart from Silva et al (2023a) and Ausseil et al (2022), Dayarian and Savelsbergh (2020) also introduce a dynamic component. Here, tasks and ODs stochastically arrive over time and are integrated into the optimization system. The study focuses on optimal task assignments at each point in time, considering the uncertain future arrivals of tasks and ODs.

The final category comprises research that employs compensation strategies to enhance OD availability. These studies generally propose that offering higher compensation can effectively increase OD availability. This relationship is explored in an aggregated context, often without delving into the specifics of individual OD behaviors. Key contributions in this area include works by Cachon et al (2017), Kung and Zhong (2017), Taylor (2018), Qi et al (2018), Yildiz and Savelsbergh (2019) and Cao

et al (2020). Notably, Cao et al (2020) speculate on the potential for significant cost reductions through an online compensation framework, particularly when combined with their dynamic task assignment framework. This body of research underscores the critical role of compensation as a lever for optimizing OD participation in last-mile delivery operations.
In conclusion, the stochastic literature showcases an evolving landscape of methodologies and models that effectively engage the inherent uncertainties in OD participation. Notably, several of these studies underscore the strategic use of compensation as a crucial tool to influence and manage the stochastic availability of ODs, demonstrating its critical role in optimizing delivery logistics.

2.3 Literature review summary

The compensation framework emerges as a critical facet in leveraging ODs within various models, as delineated in several studies (e.g., Le et al, 2021; Gdowska et al, 2018). These frameworks typically employ compensation strategies to manage OD availability effectively (e.g., Barbosa et al, 2023; Qi et al, 2018). Moreover, some research (e.g., Arslan et al, 2019; Ausseil et al, 2022) introduce dynamic elements to more accurately mirror OD arrival volatility. However, a comprehensive model that combines these elements is notably absent. Current literature lacks a framework where ODs dynamically and stochastically arrive, exhibit decision-making influenced by variable compensation offers, and where an optimization system enables firms to set dynamic, individual compensations that adjust over time based on the system's current state. This gap mirrors familiar concepts in revenue management and dynamic pricing. Addressing this, our research aims to bridge this void in crowdshipping literature by applying dynamic pricing methodologies. Our research aspires to fill this gap within the crowdshipping domain by applying dynamic pricing methodologies, envisaging a reversed pricing mechanism wherein firms propose compensatory offers to secure delivery services.

3 The dynamic compensation problem for occasional drivers

In this section, we present the model along, with the corresponding notation and underlying assumptions.

3.1 General setting and notation

Our analysis focuses on a retailer that operates both a brick-and-mortar store and an online shop. At the commencement of store operations, all online orders scheduled for fulfillment on that day are predetermined and known. We assume the retailer possesses the flexibility to assign any online order for delivery to a fleet of DDs, who operate as a homogeneous group. The time horizon is modeled as a set of discrete periods $\mathcal{T} = \{1, 2, ..., T + 1\}$, whereas in each period $t \in \mathcal{T} \setminus \{T + 1\}$ at most one OD may arrive. We denote the set of delivery locations as $\mathcal{C} \subseteq \{1, 2, \ldots, C\}$ and the set of

ODs' destinations as $\mathcal{O} \subseteq \{C+1, C+2, \ldots, C+O\}$, where C and O are the number of total delivery locations and ODs, respectively. In period t, the remaining delivery locations are given by $\mathcal{C}_t$, while $\mathcal{O}_t$ contains the destinations of all ODs who may still arrive. Index 0 is exclusively used to refer to the depot location. The distances between locations are denoted by $d_{uv}\ \forall\, u, v \in \mathcal{C} \cup \mathcal{O} \cup \{0\}$, building the foundation for calculating the detour of an OD o to a delivery location c: $u_{oc} = d_{co} + d_{0c} - d_{0o} \forall o \in \mathcal{O}, c \in \mathcal{C}$. Upon their arrival, ODs are recognized by the retailer, making ODs' destinations exploitable information. The retailer decides on a compensation/delivery location combination to offer to the OD. This offer is either accepted or declined by the OD. On acceptance, the OD receives the compensation and delivers the parcel to the delivery location. When declining the offer, the OD departs without taking on a delivery task. At the terminal period $T+1$, all delivery tasks that have not been served by ODs are served by DDs.

3.2 Formulation as a dynamic program

The remaining problem description will be structured by the essential components of a dynamic program in line with Powell (2011).

3.2.1 Periods

If each time period $t \in \mathcal{T}$ is chosen to be sufficiently small, the likelihood of experiencing two OD arrivals within the same period becomes negligible. This practice aligns with established conventions in standard revenue management literature (Strauss et al, 2018). For an OD $o \in \mathcal{O}$, the probability of arrival is denoted as λ_{ot} for all $t \in \mathcal{T} \setminus \{T+1\}$. The probability of no arrival in period t is represented by $\lambda_{0t} = 1 - \sum_{o \in \mathcal{O}} \lambda_{ot}$, indicating that $\sum_{o \in \mathcal{O}} \lambda_{ot} + \lambda_{0t} = 1$. To maintain generality in the formulation, the probability of OD arrival may vary for each OD and could also depend on t.

3.2.2 States

The progression through each period involves three primary types of states: pre-arrival (S_t^A), pre-decision (S_t^X), and post-decision state (S_t^P). Each state structure varies, encapsulating information specific to its context. A visual representation of the dynamic program is illustrated in Figure 2 as a decision tree, where circles represent random nodes, and squares denote decision nodes.

1. *Pre-arrival states* S_t^A contain information about potential OD arrivals, denoted as $\mathcal{O}_t$, and unassigned delivery tasks, denoted as $\mathcal{C}_t$. Formally, $S_t^A = (\mathcal{O}_t, \mathcal{C}_t)$. The pre-arrival state transitions into the pre-decision state with the arrival of an OD in period t. If no OD arrives during the pre-arrival state in period t, the subsequent states are skipped, and the next state becomes the pre-arrival state of period $t+1$.
2. *Pre-decision states* S_t^X include all information from the preceding pre-arrival state and the specific OD arrival, denoted as o_t, in period t. In mathematical terms, $S_t^X = (\mathcal{O}_t, \mathcal{C}_t, o_t)$. With the knowledge of OD o_t arriving, the firm must decide on the offer bundle (c, r), comprising a delivery task $c \in \mathcal{C}_t$ and a compensation r (refer to 3.2.4).

3. *Post-decision states* S_t^P encompass all information from the preceding pre-decision state and the bundle (c, r) offered to OD o_t, i.e., $S_t^P = (\mathcal{O}_t, \mathcal{C}_t, o_t, c, r)$. After OD o_t makes a decision, the state transitions into the pre-arrival state of the next period by updating $\mathcal{O}_t$ and $\mathcal{C}_t$ accordingly, based on the OD's decision.

The transition from states over periods t is terminated when period $T + 1$ is reached. For a comprehensive description of the transition function see 3.2.3.

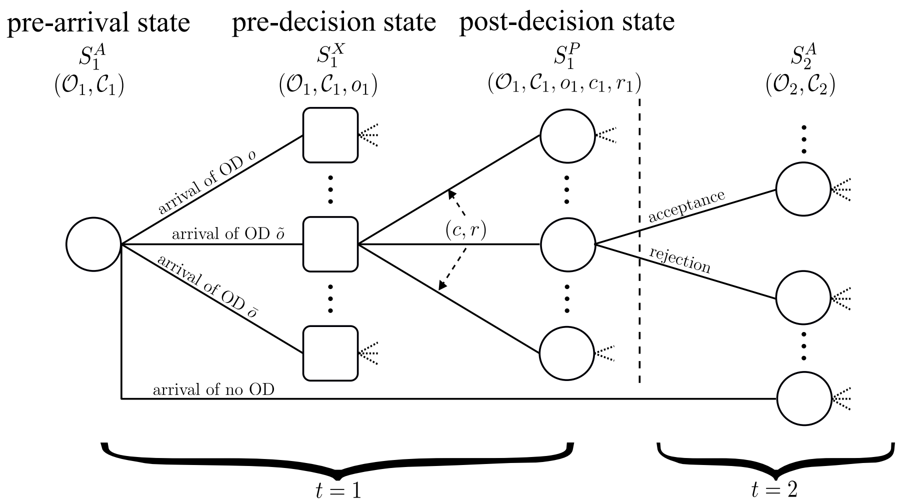

Fig. 2 Tree representation of the Dynamic Program. Circles represent states in which the transition is determined by a random process, while squares represent states in which the subsequent state is determined deterministically by the firms decision

3.2.3 Transitions

We move through states within a period and from one period t to the next, integrating new information as the scenario evolves. Transitioning from pre-arrival to pre-decision state involves a stochastic process, where additional information is based on the stochastic arrival of up to one OD $o_t \in \mathcal{O}_t$. The transition from pre-decision to post-decision state is deterministic, dependent on the firm's decision. Moving from post-decision to pre-arrival state in the subsequent period follows a stochastic process, reflecting the OD's unknown decision regarding the firm's offer. Specifically, OD o_t is excluded from $\mathcal{O}_t$, assuming departed ODs don't return within the considered horizon. This mirrors their behavior as non-employed individuals who complete their shopping in one visit and return only when supplies are depleted. Furthermore, based on OD o_t accepting or rejecting the offered task c, c is either removed from or retained in $\mathcal{C}_t$. In summary, the transition from one period to another can be characterized by:

$$S_{t+1}^{A} = \begin{cases} (\mathcal{O}_t \setminus \{o_t\}, \mathcal{C}_t \setminus \{c_t\}) & \text{if OD } o_t \text{ arrives and serves delivery location } c_t \\ (\mathcal{O}_t \setminus \{o_t\}, \mathcal{C}_t) & \text{if OD } o_t \text{ arrives and rejects the offer} \\ (\mathcal{O}_t, \mathcal{C}_t) & \text{if no OD arrives in } t \end{cases} \quad (1)$$

3.2.4 The firm's decisions

The decision space, denoted as $\mathcal{X} = (\mathcal{X}_C, \mathcal{X}_R)$, represents the entire collection of all possible decisions and depends on the current pre-decision state S_t^X. Consequently, we write $\mathcal{X}(S_t^X) = (\mathcal{X}_C(S_t^X), \mathcal{X}_R(S_t^X))$. A decision is represented as $(c, r) \in \mathcal{X}(S_t^X)$, where c denotes a delivery location, and r indicates the compensation paid to the OD for fulfilling this delivery. In our framework, we assume that the firm offers at most one delivery location at a time, eliminating the need to consider trunk space or weight limitations. This assumption emphasizes that occasional drivers are viewed as opportunistic individuals rather than employed personnel, reflecting their preference for minimal commitment. Additionally, we do not impose further limitations on offers, making $\mathcal{X}_C(S_t^X)$ equivalent to the most general case, $\mathcal{C}_t$. $\mathcal{X}_R(S_t^X)$ can be modeled as either a discrete set or a continuous range of possible compensations. In our work, we allow for the most general case, where $\mathcal{X}_R(S_t^X) = \mathbb{R}_+$.

3.2.5 The OD's decision

Each OD possesses a unique indifference compensation (IC), which marks the minimum compensation that leads to accepting a delivery task to a specified delivery location. The prevailing trend in existing literature often assumes or implicitly suggests, that the ICs of ODs are known to the firm. In an effort to deviate from this trend, a crucial step is to establish a suitable representation of the IC. In our study, we propose that the IC is influenced by two factors: one encompasses observable information (such as detour, traffic, parking availability at the delivery location, weather, etc.), expressed as a_{oct} indicating the monetary compensation required for known inconveniences; the other involves unobservable information (like the OD's time constraints and mood on a specific day), represented by the random variable ω, denoting the additional (unknown) amount required to persuade the OD to accept today's offer. Formally, it holds that

$$IC_{oct} = a_{oct} + \omega. \quad (2)$$

We assume that ω follows a continuous distribution with positive support on $[0, b_{oct}]$, characterized by its cumulative distribution function denoted as F and its probability density function as f. While the choice of distribution is flexible and may vary across ODs, delivery locations, and periods (although we don't explicitly specify this by using f_{oct}), we restrict ourselves to a specific class of distributions characterized by the condition that f/F is decreasing[1]. This condition ensures that our decision problem has a unique solution (refer to Proposition 1). Notably, every distribution with a decreasing f satisfies this criterion. Furthermore, this condition is akin to requiring

[1] We use decreasing/increasing and lower/higher in a weak sense.

that the reflected distribution, given by $\bar{f}(x) = f(-x)$ over the support $[-b_{oct}, 0]$, exhibits an increasing failure rate $\bar{h}(x) = \bar{f}(x)/(1 - \bar{F}(x))$.
Adopting the assumption of an increasing failure rate aligns with common practices in dynamic pricing literature. Random variables exhibiting increasing failure rates have a growing generalized failure rate, as discussed in Lariviere (2006). This choice is consistent with one of the three standard assumptions outlined in Ziya et al (2004). Furthermore, it is compatible with numerous probability distributions, including but not limited to the uniform, triangular, normal, exponential, Weibull, Gumbel, gamma distributions, and their truncated variants (some of them with restrictions regarding parameter choice) as documented in Banciu and Mirchandani (2013). Each of these distributions can serve as $\bar{f}$, further extending the possibilities for the selection of f.

3.2.6 Cost of delivery by DDs

As recommended, among others, by Boysen et al (2022), a common approach to modeling delivery costs by DDs is to assume a fixed fee, denoted as κ_c, for delivering to any remaining delivery location $c \in \mathcal{C}_{T+1}$. Consequently, the total costs are computed using $\Theta(\mathcal{C}_{T+1}) = \sum_{c \in \mathcal{C}_{T+1}} \kappa_c$. This assumption aligns with the established business practice of third-party logistics providers such as DHL and UPS, which typically charge a flat rate per package delivered. For retailers without a dedicated delivery fleet or those opting to outsource deliveries, this cost structure is commonly encountered. Importantly, this assumption eliminates the need for route building, allowing us to concentrate on identifying optimal task offers and corresponding compensations. However, it's worth noting that our proposed framework is not contingent on a fixed fee, providing flexibility in structuring cost evaluations at the end of the planning horizon.

3.3 The Bellman equation

The objective is to minimize the total expected costs, denoted as $V_1(\mathcal{O}_1, \mathcal{C}_1)$. Costs are incurred by fulfilling all delivery tasks $c \in \mathcal{C}_1 = \mathcal{C}$, either by employing DDs or leveraging the potential arrivals of ODs $o \in \mathcal{O}_1 = \mathcal{O}$. Consequently, they consist of all compensations paid to ODs during the planning horizon $t = 1, \dots, T$ and to DDs at the end of the planning horizon $t = T + 1$ for all remaining delivery tasks $\mathcal{C}_{T+1}$. Given the dynamic nature of this decision problem, we formulate it through a Bellman equation. The expected costs in a pre-arrival state $S_t^A = (\mathcal{O}_t, \mathcal{C}_t)$ with remaining ODs $\mathcal{O}_t$ and remaining delivery tasks $\mathcal{C}_t$ can be represented as follows:

$$\begin{aligned} V_t(\mathcal{O}_t, \mathcal{C}_t) = E_o[min_{(c,r) \in \mathcal{X}(\mathcal{O}_t, \mathcal{C}_t, o)} \{F(r - a_{oct}) \cdot (r + V_{t+1}(\mathcal{O}_t \setminus \{o\}, \mathcal{C}_t \setminus \{c\})) \\ + (1 - F(r - a_{oct})) \cdot V_{t+1}(\mathcal{O}_t \setminus \{o\}, \mathcal{C}_t)\}] \end{aligned} \tag{3}$$

with the boundary condition $V_{T+1}(S_{T+1}^A) = \Theta(\mathcal{C}_{T+1})$. The expectation E_o covers all potential arrivals of ODs $o \in \mathcal{O}_t$ as well as the event of no arrival. In the latter case, the firm has no decision to make and faces expected future costs of $V_{t+1}(\mathcal{O}_t, \mathcal{C}_t)$.
To minimize expected costs, our approach involves a unified decision-making process following the arrival of OD o. This entails the dual determination of selecting the

delivery task c presented to OD o and simultaneously deciding on the compensation r offered for completing this task. The computation of expected costs for any given combination of o, c, and r considers two distinct outcomes: the OD's acceptance of the proposed bundle, consisting of the delivery task and compensation, or their rejection. Acceptance incurs immediate costs r and future expected costs $V_{t+1}(\mathcal{O}_t \setminus \{o\}, \mathcal{C}_t \setminus \{c\})$ for delivering the remaining tasks contained in $\mathcal{C}_t \setminus \{c\}$, with potential assistance from the remaining ODs given in $\mathcal{O}_t \setminus \{o\}$. Conversely, rejection has no immediate costs but leads to future expected costs $V_{t+1}(\mathcal{O}_t \setminus \{o\}, \mathcal{C}_t)$ for fulfilling all deliveries in $\mathcal{C}_t$ with potential ODs from $\mathcal{O}_t \setminus \{o\}$. Notably, OD o accepts the offered bundle if and only if her indifference compensation $IC_{oct} = a_{oct} + \omega$ is below the offered compensation r. By offering a bundle (c, r) to OD o, the firm hopes to reduce expected costs. Consequently, the firm avoids offers with $r > V_{t+1}(\mathcal{O}_t \setminus \{o\}, \mathcal{C}_t) - V_{t+1}(\mathcal{O}_t \setminus \{o\}, \mathcal{C}_t \setminus \{c\})$. We define

$$\Delta V_t(\mathcal{O}_t, \mathcal{C}_t, o, c) = V_{t+1}(\mathcal{O}_t \setminus \{o\}, \mathcal{C}_t) - V_{t+1}(\mathcal{O}_t \setminus \{o\}, \mathcal{C}_t \setminus \{c\}) \tag{4}$$

as avoided costs, representing the difference between the expected costs of the next period's pre-arrival states with and without the offered delivery location c. Furthermore, based on avoided costs, we define expected savings from offering the bundle (c, r) to OD o as:

$$Sav_t(\mathcal{O}_t, \mathcal{C}_t, o, c, r) = F(r - a_{oct}) \cdot (\Delta V_t(\mathcal{O}_t, \mathcal{C}_t, o, c) - r). \tag{5}$$

With expected savings, we formulate an equivalent decision problem to the one integrated in (3). The following decision problem highlights the importance of avoided costs in our decision-making process and will be a cornerstone for our analysis of the problem and the creation of solution methods. It holds that

$$\begin{aligned} & min_{(c,r)\in\mathcal{X}(\mathcal{O}_t,\mathcal{C}_t,o)}\{F(r - a_{oct}) \cdot (r + V_{t+1}(\mathcal{O}_t \setminus \{o\}, \mathcal{C}_t \setminus \{c\})) \\ & \qquad\qquad + (1 - F(r - a_{oct})) \cdot V_{t+1}(\mathcal{O}_t \setminus \{o\}, \mathcal{C}_t)\} \\ & = min_{(c,r)\in\mathcal{X}(\mathcal{O}_t,\mathcal{C}_t,o)}\{V_{t+1}(\mathcal{O}_t \setminus \{o\}, \mathcal{C}_t) - F(r - a_{oct}) \cdot (\Delta V_t(\mathcal{O}_t, \mathcal{C}_t, o, c) - r)) \\ & = V_{t+1}(\mathcal{O}_t \setminus \{o\}, \mathcal{C}_t) - max_{(c,r)\in\mathcal{X}(\mathcal{O}_t,\mathcal{C}_t,o)}\{Sav_t(\mathcal{O}_t, \mathcal{C}_t, o, c, r)\} \end{aligned} \tag{6}$$

Equation (6) demonstrates that maximizing expected savings is equivalent to minimizing expected costs, as $V_{t+1}(\mathcal{O}_t \setminus \{o\}, \mathcal{C}_t)$ is independent of the decisions. In the remainder of this work, we focus on finding the optimal bundle $(c(o_t), r(o_t))$ for OD o_t that maximizes expected savings in a given pre-decision state $S_t^X = (\mathcal{O}_t, \mathcal{C}_t, o_t)$:

$$max_{(c,r)\in\mathcal{X}(\mathcal{O}_t,\mathcal{C}_t,o_t)}\{Sav_t(\mathcal{O}_t, \mathcal{C}_t, o_t, c, r)\} \tag{7}$$

Observing the decision problem posed by equation (7), we encounter two crucial questions: First, does a unique optimal solution exist for every state, and can we effectively

determine it? Second, if we can identify the optimal solution, is this sufficient to efficiently minimize the total expected costs $V_1(\mathcal{O}_1, \mathcal{C}_1)$ throughout the entire planning horizon?
The first question is addressed in Section 4, where we establish the uniqueness of the optimal solution and provide a sufficient optimality condition for any state. However, the second question demands more nuanced consideration: By solving the decision problem for each pre-decision state $S_t^P = (\mathcal{O}_t, \mathcal{C}_t, o_t)$ in period t, we can compute the expected costs associated with every pre-arrival state $S_t^A = (\mathcal{O}_t, \mathcal{C}_t)$ in the same period (refer to equations (3) and (6)). This process, in turn, yields the avoided costs $\Delta V_{t-1}(\mathcal{O}_{t-1}, \mathcal{C}_{t-1}, o_{t-1}, c)$ for each pre-decision state $S_{t-1}^P = (\mathcal{O}_{t-1}, \mathcal{C}_{t-1}, o_{t-1})$. By iteratively applying this methodology, we aim to minimize the total expected costs $V_1(\mathcal{O}_1, \mathcal{C}_1)$. However, the sheer volume of potential pre-arrival and pre-decision states over the entire planning horizon grows exponentially, rendering a conventional roll-back procedure practically infeasible, even for relatively modest instances.
Indeed, this dynamic problem is afflicted by the "curses of dimensionality" (refer to Powell (2011)). Without constraining the number of states through rules, the cardinality of $|\mathcal{S}_t^A|$ becomes 2^{C+O} in each period t. This prompts the need for a method that provides an approximation of the value function $V_t(\mathcal{O}_t, \mathcal{C}_t)$ or the avoided costs $\Delta V_t(\mathcal{O}_t, \mathcal{C}_t, o, c)$. We address this issue in Section 5.

4 Optimal solution and structural properties

In the first part of this section, we establish the existence of a unique solution to our decision problem outlined in equation (7). Additionally, we derive the optimal solution, showcasing its closed-form expression in the event of a uniformly distributed ω. Moving on to the second part of this section, we delve into an analysis of structural properties. We aim to glean insights into the model's inherent characteristics, improving our capability to adequately approximate the value function or avoided costs in Section 5.

4.1 Optimal state-dependent solution

In this section, we operate within an arbitrary pre-decision state $S_t^X = (\mathcal{O}_t, \mathcal{C}_t, o)$, aiming to identify the optimal offer bundle $(c(o), r(o)) \in \mathcal{X}(\mathcal{O}_t, \mathcal{C}_t, o)$ that maximizes 7. Our exploration begins by focusing on a fixed delivery task $c \in \mathcal{C}_t$, and subsequently determining the optimal compensation r based on the interplay between OD o and the chosen delivery task c.

Proposition 1. *In a pre-decision state $S_t^X = (\mathcal{O}_t, \mathcal{C}_t, o)$, with a given $c \in \mathcal{C}_t$, there is a unique $r \in [a_{oct}, b_{oct} + a_{oct}]$ that maximizes $Sav_t(\mathcal{O}_t, \mathcal{C}_t, o_t, c, r)$. This r either fulfills $r + \frac{F(r-a_{oct})}{f(r-a_{oct})} = \Delta V_t(\mathcal{O}_t, \mathcal{C}_t, o, c)$ or is on the bounds of $[a_{oct}, b_{oct} + a_{oct}]$.*

Proof. Since f has positive support on $[0, b_{oct}]$, $F(r - a_{oct})$ establishes a bijective function for $r \in [a_{oct}, b_{oct} + a_{oct}]$. Consequently, we can introduce $\theta = F(r - a_{oct})$ as an alternative decision variable, where uniqueness carries over between the optimal r and the optimal θ. Utilizing θ offers the advantage that the second derivative is independent of $\Delta V_t(\mathcal{O}_t, \mathcal{C}_t, o, c)$, allowing us to generally prove the concavity of the savings function in θ. With equation (5) and $r(\theta) = F^{-1}(\theta) + a_{oct}$, we can formulate the first derivative of the savings function:

$$\begin{aligned} \frac{d}{d\theta} Sav_t(\mathcal{O}_t, \mathcal{C}_t, o, c, r(\theta)) &= (\Delta V_t(\mathcal{O}_t, \mathcal{C}_t, o, c) - F^{-1}(\theta) - a_{oct}) - \frac{\theta}{f(F^{-1}(\theta))} \\ &= (\Delta V_t(\mathcal{O}_t, \mathcal{C}_t, o, c) - r(\theta)) - \frac{F(r(\theta) - a_{oct})}{f(r(\theta) - a_{oct})}. \end{aligned} \tag{8}$$

The second derivative is expressed as:

$$\frac{d^2}{d\theta^2} Sav_t(\mathcal{O}_t, \mathcal{C}_t, o, c, r(\theta)) = -\frac{d}{d\theta} r(\theta)(1 + \frac{d}{dr} \frac{F(r - a_{oct})}{f(r - a_{oct})}|_{r=r(\theta)}). \tag{9}$$

Given that $r(\theta)$ is increasing with θ and $\frac{F(r-a_{oct})}{f(r-a_{oct})}$ is increasing with r (as indicated in Section 3.2.5), the second derivative is negative. Consequently, the savings function is (strictly) concave in θ. This leads to the existence of a unique $\theta \in [0, 1]$ that maximizes the savings function for a given c. This uniqueness also extends to $r \in [a_{oct}, b_{oct} + a_{oct}]$. $\square$

Remark 1. *With a uniformly distributed ω, i.e., $\omega \sim U_{[0,b_{oct}]}$, the optimal compensation r for a given pre-decision state $S_t^X = (\mathcal{O}_t, \mathcal{C}_t, o)$ and a given delivery location $c \in \mathcal{C}_t$ can be derived by the following closed-form expression:*

$$r = \begin{cases} a_{oct} & \text{if } \Delta V_t(\mathcal{O}_t, \mathcal{C}_t, o, c) \leq a_{oct} \\ \dfrac{\Delta V_t(\mathcal{O}_t, \mathcal{C}_t, o, c) + a_{oct}}{2} & \text{if } a_{oct} < \Delta V_t(\mathcal{O}_t, \mathcal{C}_t, o, c) < 2b_{oct} + a_{oct} \\ b_{oct} + a_{oct} & \text{if } \Delta \geq 2b_{oct} + a_{oct} \end{cases} \tag{10}$$

Knowing the optimal compensation for a specific delivery location offers the advantage of evaluating various combinations of delivery locations and their corresponding optimal compensations to identify the most favorable combination. However, the computational effort increases with the number of remaining delivery tasks, making it more challenging to find the best delivery location. Fortunately, we can introduce a criterion to simplify this process. To do so, we first need two lemmas to prepare ourselves for demonstrating that this criterion indeed identifies the optimal delivery location.

Lemma 1. *In a pre-decision state $S_t^X = (\mathcal{O}_t, \mathcal{C}_t, o)$, with $c, \tilde{c} \in \mathcal{C}_t$ and corresponding optimal compensations $r, \tilde{r}$, respectively, the following implication holds:*

$$\Delta V_t(\mathcal{O}_t, \mathcal{C}_t, o, c) - a_{oct} \geq \Delta V_t(\mathcal{O}_t, \mathcal{C}_t, o, \tilde{c}) - a_{o\tilde{c}t} \implies F(r - a_{oct}) \geq F(\tilde{r} - a_{o\tilde{c}t}) \tag{11}$$

Proof. Following a similar approach as in the proof of Proposition 1, we transition to alternative decision variables $\theta = F(r - a_{oct})$ and $\tilde{\theta} = F(\tilde{r} - a_{o\tilde{c}t})$. It is noteworthy that θ and $\tilde{\theta}$ are optimal solutions for their respective savings functions. Utilizing equation

(8) and the optimality of θ, we derive that:

$$\begin{aligned} 0 &= (\Delta V_t(\mathcal{O}_t, \mathcal{C}_t, o, c) - a_{oct} - F^{-1}(\theta)) - \frac{\theta}{f(F^{-1}(\theta))} \\ &\geq (\Delta V_t(\mathcal{O}_t, \mathcal{C}_t, o, \tilde{c}) - a_{o\tilde{c}t} - F^{-1}(\theta)) - \frac{\theta}{f(F^{-1}(\theta))}. \end{aligned} \tag{12}$$

This implies that θ is greater than or equal to the optimal solution for the savings function with $\tilde{c}$, confirming $\theta \geq \tilde{\theta}$. Consequently, $F(r - a_{oct}) \geq F(\tilde{r} - a_{o\tilde{c}t})$. □

Lemma 1 indicates that our optimal solution is selected in a manner where delivery locations exhibiting a larger difference between avoided costs and known inconveniences are paired with an optimal compensation that yields a higher probability of acceptance from the OD. The subsequent Lemma 2 conveys a related concept, emphasizing a higher realized saving rather than a higher probability.

Lemma 2. *In a pre-decision state $S_t^X = (\mathcal{O}_t, \mathcal{C}_t, o)$, with $c, \tilde{c} \in \mathcal{C}_t$ and corresponding optimal compensations $r, \tilde{r}$, respectively, the following implication holds:*

$$\begin{aligned} &\Delta V_t(\mathcal{O}_t, \mathcal{C}_t, o, c) - a_{oct} \geq \Delta V_t(\mathcal{O}_t, \mathcal{C}_t, o, \tilde{c}) - a_{o\tilde{c}t} \\ &\implies \Delta V_t(\mathcal{O}_t, \mathcal{C}_t, o, c) - r \geq \Delta V_t(\mathcal{O}_t, \mathcal{C}_t, o, \tilde{c}) - \tilde{r} \end{aligned} \tag{13}$$

Proof. We set $\hat{r} = r + \Delta V_t(\mathcal{O}_t, \mathcal{C}_t, o, \tilde{c}) - \Delta V_t(\mathcal{O}_t, \mathcal{C}_t, o, c)$ and check whether this compensation is below or above the optimal compensation $\tilde{r}$. Utilizing Proposition 1, we observe:

$$\begin{aligned} &\Delta V_t(\mathcal{O}_t, \mathcal{C}_t, o, \tilde{c}) - \hat{r} - \frac{F(\hat{r} - a_{o\tilde{c}t})}{f(\hat{r} - a_{o\tilde{c}t})} \\ &= \Delta V_t(\mathcal{O}_t, \mathcal{C}_t, o, c) - r - \frac{F(r + \Delta V_t(\mathcal{O}_t, \mathcal{C}_t, o, \tilde{c}) - \Delta V_t(\mathcal{O}_t, \mathcal{C}_t, o, c) - a_{o\tilde{c}t})}{f(r + \Delta V_t(\mathcal{O}_t, \mathcal{C}_t, o, \tilde{c}) - \Delta V_t(\mathcal{O}_t, \mathcal{C}_t, o, c) - a_{o\tilde{c}t})} \\ &\geq \Delta V_t(\mathcal{O}_t, \mathcal{C}_t, o, c) - r - \frac{F(r - a_{oct})}{f(r - a_{oct})} = 0. \end{aligned} \tag{14}$$

The inequality arises from $\Delta V_t(\mathcal{O}_t, \mathcal{C}_t, o, \tilde{c}) - \Delta V_t(\mathcal{O}_t, \mathcal{C}_t, o, c) - a_{o\tilde{c}t} \leq -a_{oct}$ and the observation that $f(x)/F(x)$ is decreasing with x (as discussed in Section 3.2.5). The final equality is a consequence of the optimality of r for c. This implies that $\hat{r}$ is greater than or equal to $\tilde{r}$. Consequently, $\tilde{r} \leq \hat{r} = r + \Delta V_t(\mathcal{O}_t, \mathcal{C}_t, o, \tilde{c}) - \Delta V_t(\mathcal{O}_t, \mathcal{C}_t, o, c)$, thus affirming the assertion of Lemma 2. □

Lemma 1 and Lemma 2 together carry a significant implication: A delivery location with a higher difference between avoided costs and known inconvenience results in an optimal compensation providing a higher probability of acceptance from the OD and greater realized savings in case of acceptance. This directly leads to the following proposition.

Proposition 2. *In a pre-decision state $S_t^X = (\mathcal{O}_t, \mathcal{C}_t, o)$, $c(o) = argmax_{c \in \mathcal{C}_t}\{\Delta V_t(\mathcal{O}_t, \mathcal{C}_t, o, c) - a_{oct}\}$ is the optimal delivery location to offer to OD o.*

Proof. The proof immediately follows by Lemmas 1 and 2, in conjunction with the definition of the savings function (refer to equation (5)). By definition, $\Delta V_t(\mathcal{O}_t, \mathcal{C}_t, o, c(o)) - a_{oc(o)t} \geq \Delta V_t(\mathcal{O}_t, \mathcal{C}_t, o, \tilde{c}) - a_{o\tilde{c}t}$ for any $\tilde{c} \in \mathcal{C}_t$. Moreover, according to Proposition 1, let r and $\tilde{r}$ be the optimal compensation for $c(o)$ and $\tilde{c}$, respectively. Then,

$$\begin{aligned} Sav_t(\mathcal{O}_t, \mathcal{C}_t, o, c(o), r) &= F(r - a_{oc(o)t}) \cdot (\Delta V_t(\mathcal{O}_t, \mathcal{C}_t, o, c(o)) - r) \\ &\geq F(\tilde{r} - a_{o\tilde{c}t}) \cdot (\Delta V_t(\mathcal{O}_t, \mathcal{C}_t, o, \tilde{c}) - \tilde{r}) = Sav_t(\mathcal{O}_t, \mathcal{C}_t, o, \tilde{c}, \tilde{r}). \end{aligned} \tag{15}$$

□

Proposition 2 provides a simple criterion for determining the optimal delivery location. Given that $\mathcal{C}_t$ is discrete, a_{oct} is a parameter, and $\Delta V_t(\mathcal{O}_t, \mathcal{C}_t, o, c)$ is derived from the expected costs in the subsequent period $t+1$ through the earlier step of a roll-back procedure, identifying the maximizer becomes a straightforward task. Following the identification of the optimal delivery location, the corresponding optimal compensation can be obtained using Proposition 1. These propositions collectively provide an immediate solution to determining the optimal offer bundle $(c(o), r(o))$ in a given pre-decision state $S_t^X = (\mathcal{O}_t, \mathcal{C}_t, o)$. The primary challenge lies in the computation of each $\Delta V_t(\mathcal{O}_t, \mathcal{C}_t, o, c)$, considering these avoided costs depend on $V_{t+1}(\mathcal{O}_{t+1}, \mathcal{C}_{t+1})$ for every possible subsequent pre-arrival state. The number of potential states grows exponentially with C and O, rendering the calculation of each instance beforehand a daunting task. Hence, there arises a necessity to approximate the value function or the avoided costs in an efficient manner. To minimize structural or systematic errors, our goal is to devise an approximation mechanism that faithfully captures the inherent structure of the value function and the avoided costs. Consequently, we delve into the identification of such structural properties in the subsequent section.

4.2 Structural properties

We commence our investigation into the structural properties by scrutinizing the monotonicities of the value function. Specifically, we draw inspiration from frequently observed structures in the realm of dynamic pricing, where concave-increasing expected revenues in remaining time and capacity are common. In our context, we will demonstrate that our objective function exhibits a similar monotonic increasing/decreasing behavior. Moreover, we will highlight the distinctive feature of our setting, where we typically do not observe a concave or convex behavior, setting it apart from standard dynamic pricing scenarios.

Initially, we establish that expected costs increase with $\mathcal{C}_t$ and with t, while they decrease with $\mathcal{O}_t$.

Proposition 3. *For any pre-arrival state $S_t^X = (\mathcal{O}_t, \mathcal{C}_t)$, it holds:*

1. $V_t(\mathcal{O}_t, \tilde{\mathcal{C}}_t) \leq V_t(\mathcal{O}_t, \mathcal{C}_t)$ *with* $\tilde{\mathcal{C}}_t \subset \mathcal{C}_t$
2. $V_t(\tilde{\mathcal{O}}_t, \mathcal{C}_t) \geq V_t(\mathcal{O}_t, \mathcal{C}_t)$ *with* $\tilde{\mathcal{O}}_t \subset \mathcal{O}_t$
3. $V_t(\mathcal{O}_t, \mathcal{C}_t) \leq V_{t+1}(\mathcal{O}_t, \mathcal{C}_t)$ *with* $t \leq T$ *and time-homogeneous arrivals*

Proof. We will prove each statement individually. To establish the validity of the first and second statements, it is sufficient to demonstrate the assertions for any subsets $\tilde{\mathcal{C}}_t \subset \mathcal{C}_t$ and $\tilde{\mathcal{O}}_t \subset \mathcal{O}_t$, where $\mathcal{C}_t \setminus \tilde{\mathcal{C}}_t = \{\tilde{c}_t\}$ and $\mathcal{O}_t \setminus \tilde{\mathcal{O}}_t = \{\tilde{o}_t\}$ for any $\tilde{c}_t$ and $\tilde{o}_t$, respectively. By consistently applying the arguments outlined below for the first and second statements, we can verify the more general assertion of the proposition. We will establish the validity of the first two statements through induction over t. The statements are evidently true for the base case with $t = T + 1$, as $V_t(\mathcal{O}_t, \tilde{\mathcal{C}}_t) = \sum_{c \in \tilde{\mathcal{C}}_t} \kappa_c \leq \sum_{c \in \tilde{\mathcal{C}}_t} \kappa_c + \kappa_{\tilde{c}_t} = V_t(\mathcal{O}_t, \mathcal{C}_t)$, $V_t(\mathcal{O}_t, \mathcal{C}_t) = \sum_{c \in \mathcal{C}_t} \kappa_c = V_t(\tilde{\mathcal{O}}_t, \mathcal{C}_t)$, allowing us to focus on the induction step.

1. Induction step, $t + 1 \to t$:
   We examine two pre-arrival states $(\mathcal{O}_t, \tilde{\mathcal{C}}_t)$ and $(\mathcal{O}_t, \mathcal{C}_t)$. Both states contain the same set of remaining ODs, giving rise to related pre-decision states $(\mathcal{O}_t, \tilde{\mathcal{C}}_t, o)$ and $(\mathcal{O}_t, \mathcal{C}_t, o)$, respectively, with identical arrival probabilities λ_{ot}. In the following, we will prove that the expected costs associated with the pre-decision states $(\mathcal{O}_t, \tilde{\mathcal{C}}_t, o)$ are lower than the expected costs stemming from $(\mathcal{O}_t, \mathcal{C}_t, o)$ for any $o \in \mathcal{O}_t$. Let $(c(o), r(o))$ denote the optimal offer bundle in any pre-decision state $(\mathcal{O}_t, \mathcal{C}_t, o)$.
   Now, we manipulate the policy in the pre-decision states $(\mathcal{O}_t, \tilde{\mathcal{C}}_t, o)$. Instead of applying the optimal policy, we offer $(c(o), r(o))$ if $c(o) \neq \tilde{c}$, and $(c, 0)$ for an arbitrary $c \in \mathcal{C}_t$ if $c(o) = \tilde{c}$. Through the subsequent case analysis, we demonstrate that even with this suboptimal policy, we can still achieve lower expected costs originating from the state $(\mathcal{O}_t, \tilde{\mathcal{C}}_t, o)$ for any $o \in \mathcal{O}_t$, implying expected costs in $(\mathcal{O}_t, \tilde{\mathcal{C}}_t)$ are lower than in $(\mathcal{O}_t, \mathcal{C}_t)$.
   Two cases can apply:
   <u>Case 1:</u> If we encounter an OD o with $c(o) \neq \tilde{c}$, we have the same immediate costs transitioning from both states $(\mathcal{O}_t, \tilde{\mathcal{C}}_t, o)$ and $(\mathcal{O}_t, \mathcal{C}_t, o)$ to the subsequent pre-arrival states. However, our induction hypothesis reveals that the expected future costs from the subsequent pre-arrival states are lower for $(\mathcal{O}_t \setminus \{o\}, \tilde{\mathcal{C}}_t \setminus \{c(o)\})$ and $(\mathcal{O}_t \setminus \{o\}, \tilde{\mathcal{C}}_t)$ compared to $(\mathcal{O}_t \setminus \{o\}, \mathcal{C}_t \setminus \{c(o)\})$ and $(\mathcal{O}_t \setminus \{o\}, \mathcal{C}_t)$, respectively. Therefore, in this case, the expected costs are lower for the pre-decision state $(\mathcal{O}_t, \tilde{\mathcal{C}}_t, o)$ than for $(\mathcal{O}_t, \mathcal{C}_t, o)$.
   <u>Case 2:</u>. If we encounter an OD o with $c(o) = \tilde{c}$, there are no immediate costs when transitioning from $(\mathcal{O}_t, \tilde{\mathcal{C}}_t, o)$ to $(\mathcal{O}_t \setminus \{o\}, \tilde{\mathcal{C}}_t)$. Conversely, transitioning from $(\mathcal{O}_t, \mathcal{C}_t, o)$, leads to two possibilities: either incurring immediate costs by transitioning to $(\mathcal{O}_t \setminus \{o\}, \mathcal{C}_t \setminus \{\tilde{c}\})$, or incurring no immediate costs by transitioning to $(\mathcal{O}_t \setminus \{o\}, \mathcal{C}_t)$. In the first scenario, both pre-decision states $(\mathcal{O}_t, \tilde{\mathcal{C}}_t, o)$ and $(\mathcal{O}_t, \mathcal{C}_t, o)$ result in the same subsequent pre-arrival state $(\mathcal{O}_t \setminus \{o\}, \tilde{\mathcal{C}}_t)$. However, the latter transition involves immediate costs. In the second scenario, both transitions are free of immediate costs, but transitioning from $(\mathcal{O}_t, \mathcal{C}_t, o)$ results in a subsequent pre-arrival state with higher expected future costs than transitioning from $(\mathcal{O}_t, \tilde{\mathcal{C}}_t, o)$, as indicated by our induction hypothesis.
2. Induction step, $t + 1 \to t$:
   We analyze two pre-arrival states, namely $(\tilde{\mathcal{O}}_t, \mathcal{C}_t)$ and $(\mathcal{O}_t, \mathcal{C}_t)$. It is noteworthy that, in this scenario, the arrival of $o \in \tilde{\mathcal{O}}_t$ is equally likely in both states. However, the arrival of $\tilde{o}$ can only occur in the state $(\mathcal{O}_t, \mathcal{C}_t)$. Therefore, based on our assumptions, we have $\lambda_{\tilde{o}t} + \lambda_{0t} = \tilde{\lambda}_{0t}$, where λ_{0t} and $\tilde{\lambda}_{0t}$ represent the no-arrival

probability of ODs in the states $(\mathcal{O}_t, \mathcal{C}_t)$ and $(\tilde{\mathcal{O}}_t, \mathcal{C}_t)$, respectively.

Let $(\tilde{c}(o), \tilde{r}(o))$ denote the optimal offer bundle in any pre-decision state $(\tilde{\mathcal{O}}_t, \mathcal{C}_t, o)$. Now, we manipulate the policy in the pre-decision states $(\mathcal{O}_t, \mathcal{C}_t, o)$. Instead of applying the optimal policy, we offer $(\tilde{c}(o), \tilde{r}(o))$ when OD $o \in \tilde{O}_t$ arrives, and $(c, 0)$ for an arbitrary $c \in \mathcal{C}_t$ when OD $\tilde{o}$ arrives.

With this suboptimal policy, we constructed a scenario in which pre-decision states $(\tilde{\mathcal{O}}_t, \mathcal{C}_t, o)$ and $(\mathcal{O}_t, \mathcal{C}_t, o)$ transition with the same immediate costs and probabilities to subsequent pre-arrival states, for each $o \in \tilde{O}_t$, leading to expected future costs that are higher for the subsequent state of $(\tilde{\mathcal{O}}_t, \mathcal{C}_t, o)$ than for the subsequent state of $(\mathcal{O}_t, \mathcal{C}_t, o)$. In the absence of any OD or the arrival of OD $\tilde{o}$, both pre-decision states transition without any immediate costs to subsequent pre-arrival states, where $(\tilde{\mathcal{O}}_t, \mathcal{C}_t, o)$ once again finds itself in a subsequent state with higher expected future costs than its counterpart.

3. This proof diverges from the induction approach, relying instead on policy manipulation across the entire planning horizon. To begin, it is crucial to highlight that both pre-arrival states share an identical set of remaining ODs. Consequently, any OD o has the same probability of arriving in either state. Moreover, both states encompass the same remaining delivery locations. The sole distinction lies in the time period: while $V_t(\mathcal{O}_t, \mathcal{C}_t)$ has $T+1-t$ periods left to acquire ODs, $V_{t+1}(\mathcal{O}_t, \mathcal{C}_t)$ has only $T-t$ periods left.

   Let $\pi_{t'}$, with $t+1 \leq t' \leq T$, represent the optimal policy employed to compute $V_{t+1}(\mathcal{O}_t, \mathcal{C}_t)$, where $\pi_{t'}$ maps a pre-decision state in t' with a bundle (c, r). We proceed to replicate this policy to calculate expected costs for $(\mathcal{O}_t, \mathcal{C}_t)$ in period t. Therefore, we apply the policy $\pi_{t'}$ in period $t'-1$ for every $t+1 \leq t' \leq T$. This ensures that we encounter the same immediate costs and undergo identical transitions, regardless of whether we initiate the process in t or $t+1$. Nevertheless, commencing in t provides an extra period after the planning horizon, allowing us to further minimize expected costs when starting from t instead of $t+1$.

□

Besides monotonicities in the value function, an often observed pattern in dynamic pricing is that additional capacity and time have a diminishing effect with every additional unit, i.e. the value function is concave in these state variables. We will now show that there is no concave/convex structure present in our setting. Specifically, we will show that avoided costs do not generally increase/decrease with $\mathcal{C}_t$, $\mathcal{O}_t$, or t.

Proposition 4. *For any pre-decision state $S_t^X = (\mathcal{O}_t, \mathcal{C}_t, o)$, it holds:*

1. *Neither $\Delta V_t(\mathcal{O}_t, \tilde{\mathcal{C}}_t, o, c) \leq \Delta V_t(\mathcal{O}_t, \mathcal{C}_t, o, c)$ nor $\Delta V_t(\mathcal{O}_t, \tilde{\mathcal{C}}_t, o, c) \geq \Delta V_t(\mathcal{O}_t, \mathcal{C}_t, o, c)$ is generally true for $\tilde{\mathcal{C}}_t \subset \mathcal{C}_t$*
2. *Neither $\Delta V_t(\tilde{\mathcal{O}}_t, \mathcal{C}_t, o, c) \leq \Delta V_t(\mathcal{O}_t, \mathcal{C}_t, o, c)$ nor $\Delta V_t(\tilde{\mathcal{O}}_t, \mathcal{C}_t, o, c) \geq \Delta V_t(\mathcal{O}_t, \mathcal{C}_t, o, c)$ is generally true for $\tilde{\mathcal{O}}_t \subset \mathcal{O}_t$ is not true*
3. *Neither $\Delta V_t(\mathcal{O}_t, \mathcal{C}_t, o, c) \leq \Delta V_{t+1}(\mathcal{O}_t, \mathcal{C}_t, o, c)$ nor $\Delta V_t(\mathcal{O}_t, \mathcal{C}_t, o, c) \geq \Delta V_{t+1}(\mathcal{O}_t, \mathcal{C}_t, o, c)$ is generally true for $t \leq T$*

Proof. In order to validate the proposed propositions, we present counterexamples for each, utilizing a specially designed instance as depicted in Figure 3. This instance

is crafted to elucidate specific effects that invalidate potential monotonicities. While maintaining consistent coordinates, we vary the arrival probabilities and the presence of certain nodes across different scenarios.
A key element of our analysis focuses on the detours associated with various delivery locations. Consider OD2, which is located at a distance r from the depot. It follows that any delivery location situated on a circle centered at OD2 with radius r would have a detour distance equivalent to its distance from the depot. Consequently, the detour from OD2 to C1 is quantified as 3 units, and from OD2 to C2 as 8 units.
Additionally, we chose the location of OD3 along the orthogonal bisector of the line connecting C1 and the depot. This geometric alignment ensures that the detour for a path from OD3 to C1 is also 3 units. Furthermore, OD1 is strategically placed on the direct line between C1 and the depot. This implies that the detour for OD1 involves a double traversal of the segment between its location and C1, amounting to a total detour distance of 3 units (twice the x-coordinate difference of OD1 from C1). The detour values for each OD delivery location combination can be found in table 1. For clarity, we introduce a conceptual dummy OD4, characterized by an infinitely long detour to every delivery location and a guaranteed arrival in period 1. Throughout all instances we assume a uniformly distributed IC_{oct} which is independent of t and has the bounds $a_{oct} = u_{oc}$ and $b_{oct} = 2$. Every delivery location is associated with a cost of $\kappa_c = 10$ if it is not served by an OD until the time horizon is over.

u_{oc}	**C1**	**C2**	**C3**
OD1	3	13.2	10.7
OD2	3	8	12
OD3	3	12.9	0

Table 1 Detour Matrix

1. In the first example, we analyze the change in avoided costs that occurs when adding an OD to $\mathcal{O}_t$ in a fixed pre-decision state $V_t(\mathcal{O}_t, \mathcal{C}_t, o, c)$. Specifically, we show that the avoided costs $\Delta V_1(\{1,2,4\},\{1,2\},4,1)$ increase when introducing OD3 while the avoided costs $\Delta V_1(\{1,2,4\},\{1,2\},4,2)$ decrease. The arrival probabilities for the ODs are as follows: OD1 is certain to arrive in period 2 (probability of 1) with no arrival probability in other periods. OD2 and OD3, when the latter exists, have an arrival probability of 0.5 in periods 3 and 4. It is important to note that OD3 services no other delivery location than C1, as the lower bound of the IC for other delivery locations lies above the threshold κ_c. This renders offering these to OD3 economically irrelevant for the firm. The introduction of OD3 increases the likelihood of delivery location C1 being served in periods 3 or 4 due to the availability of two potential ODs for this location, subsequently influencing the firm's willingness to spend on serving delivery location C1 in period 2 when OD1 arrives.
   In Scenario 1, where OD3 is absent, the uncertainty of OD2's arrival prompts the firm to offer OD1 a compensation of 4.8125 in period 2 for serving C1, which is accepted with a probability of 0.90625. This means that there is a high probability that delivery location C2 is the only one remaining to be serviced during

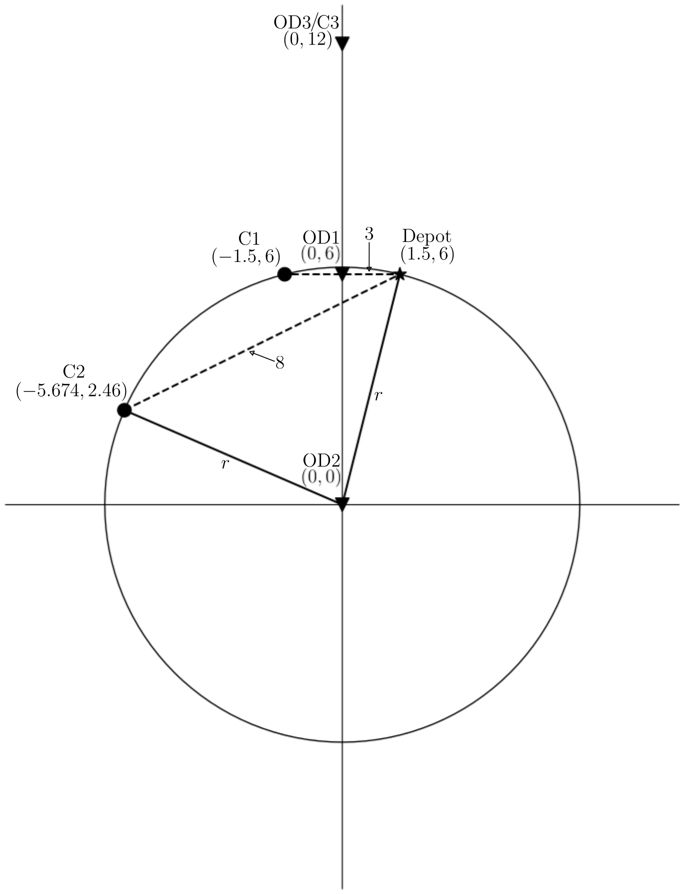

Fig. 3 Graphical representation of the created example instance

the final two periods (3 and 4). Subsequently, should OD2 make an appearance in either of these periods, the firm offers delivery location C2, coupled with a compensation of 9. This offer has a 0.5 chance of acceptance. Consequently, the overall serving probability for delivery location C2 in this case is calculated as $0.90625 \cdot (1 - 0.5^2) \cdot 0.5 \approx 0.34$. If OD1 declines the offer to serve C1 in period 2, C2 is left without the possibility of being served by an OD. This outcome stems from the fact that the remaining driver, OD2, would face a significantly smaller detour to serve location C1. Consequently, from an economic standpoint, it becomes more beneficial for the firm to allocate OD2 to C1 rather than to C2. Building on the optimal policy in periods 2 to 4, we can calculate the avoided costs in period 1, resulting in $\Delta V_1(\{1, 2, 4\}, \{1, 2\}, 4, 2) \approx 9.68$, while $\Delta V_1(\{1, 2, 4\}, \{1, 2\}, 4, 1) \approx 4.98$. A tree representation of this scenario can be seen in figure 4.

Scenario 2 introduces OD3, providing an additional viable option to serve delivery location C1 in later periods. This increased competition for serving delivery location C1, due to more service opportunities, leads to avoided costs of

$\Delta V_1(\{1,2,3,4\},\{1,2\},4,1) \approx 4.62$, a decrease compared to the avoided costs without the additional driver, OD3. The firm, therefore, reduces the amount it is willing to spend on serving C1 in early periods, which is reflected in a compensation of 4.125 for OD1 on arrival in period 2, leading to an acceptance probability of 0.5625. This decreased acceptance probability ultimately also reduces the total probability of delivery location C2 being served: Under this scenario, delivery location C2 is served in two cases. First, if OD1 accepts the offer for serving C1 in period 2 and subsequently, OD2 arrives in a later period to serve delivery location C2 (probability of occurrence: $0.5625 \cdot (1 - 0.5^2) \cdot 0.5 \approx 0.211$). Second, if OD1 declines the offer, followed by OD3's arrival in period 3, coupled with the acceptance of serving C1, and OD2's arrival in period 4 to serve delivery location C2 (probability of occurrence: $(1 - 0.5625) \cdot 0.5 \cdot 1 \cdot 0.5 \cdot 0.5 \approx 0.0547$). This results in a total serving probability for delivery location C2 of approximately 0.266 in Scenario 2, explaining why the avoided costs for delivery location C2 increase to 9.74 with the addition of OD 3, while the avoided costs for C1 decrease to 4.62. A tree representation of this scenario can be seen in figure 5.

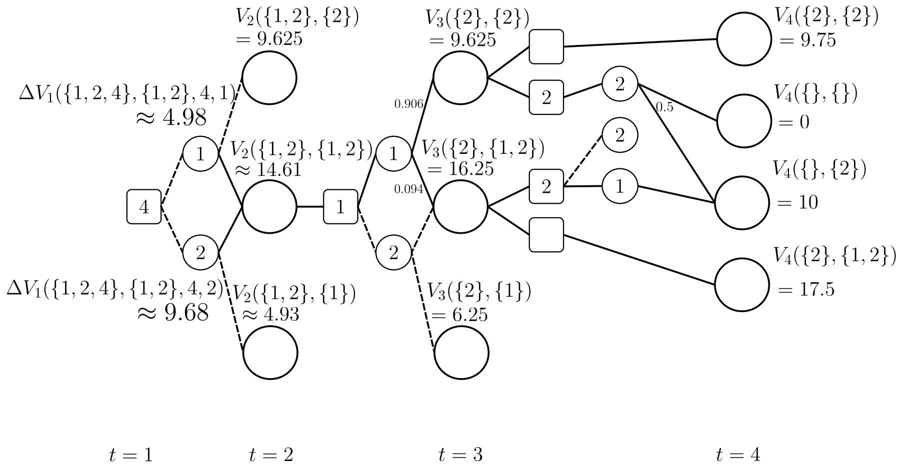

Fig. 4 Tree diagram of the example instance of Example 1 (without the additional OD)

2. In the second example, we analyze the change in avoided costs that occurs when adding a delivery location to $\mathcal{C}_t$ in a fixed pre-decision state $V_t(\mathcal{C}_t, \mathcal{O}_t, o, c)$. Specifically, we show that the avoided costs $\Delta V_1(\{1,2,3,4\},\{1,2\},4,1)$ increase when introducing delivery location 3, while the avoided costs $\Delta V_1(\{1,2,3,4\},\{1,2\},4,2)$ decrease. Considering that the scenario without the existence of delivery location C3 has been addressed previously (refer to 1., scenario 2, where the avoided costs for C1 and C2 were 4.62 and 9.74, respectively), we will focus our attention directly on the scenario where delivery location C3 is included.
   By comparing this new scenario with the first scenario from the previous example (where C3 and OD3 are absent), we observe notable similarities leading to the same avoided costs for delivery location C1 and C2. These similarities arise from

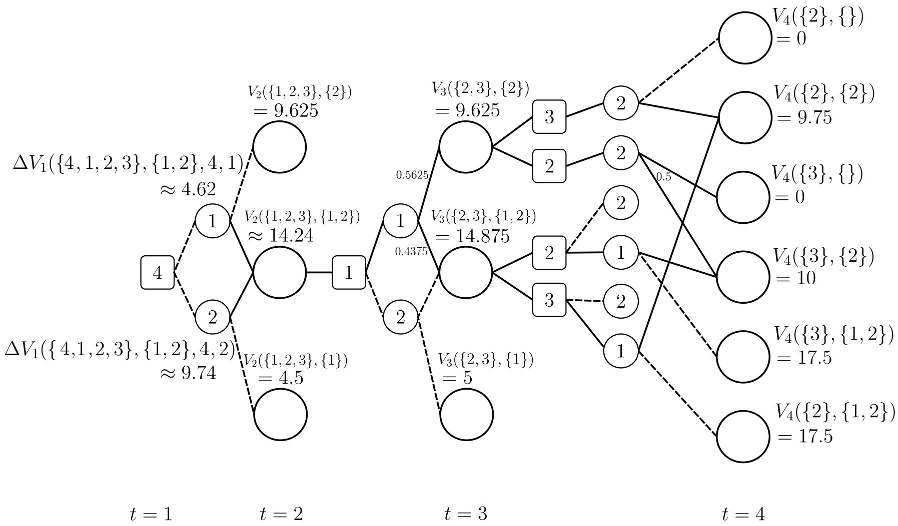

Fig. 5 Tree diagram of the example instance of Example 1 (with the additional OD)

the unique association between OD3 and C3, both situated in the same location. As a result, OD3 is consistently offered delivery location C3 for a compensation of 2 upon arrival, an offer that is accepted with a probability of 1. This arrangement effectively isolates OD3 to delivery location C3, as he becomes irrelevant for other nodes due to the strong match with delivery location C3. In the absence of delivery location C3, OD3 would have been a potential service provider for delivery location C1, resulting in the same effects already discussed in Scenario 2 of the previous example.

The introduction of delivery location C3 and the consequent exclusive pairing with OD3 reduces the competition for serving delivery location C1. This reduction in competition leads to an increase in the avoided costs for delivery location C1, akin to a scenario where neither OD3 nor delivery location C3 exists (refer to 1., scenario 1). Consequently, this alteration in the service landscape enhances the probability of delivery location C2 being served, as OD1 now serves delivery location C1 with a certainty of approximately 0.906 instead of only 0.5625 in the second period. As a result of these shifts in service probabilities and competition dynamics, the avoided costs for delivery location C2 decrease, aligning back to approximately 9.68, as observed in the scenario without OD3 and C3. Analogously, the avoided costs for C1 increase, returning to approximately 4.98. A tree diagram with the effect can be seen in figure 6. To simplify, the choices illustrated in the pre-decision states (represented as squares) have been reduced to only the optimal choices.

3. In the third example, we analyze the effect of decreasing the remaining time until the store closes. For this we hold the pre-decision state $V_t(\mathcal{C}_t, \mathcal{O}_t, o, c)$ fixed and observe how $\Delta V_t(\{2,3,4\},\{1,2\},4,c)$ changes with t for each c.

   In this scenario, avoided costs can increase and decrease with period t. This example is set within a framework of only three periods, excluding OD1 and delivery location C3 for clarity. To facilitate a comparison of avoided costs across stages, dummy

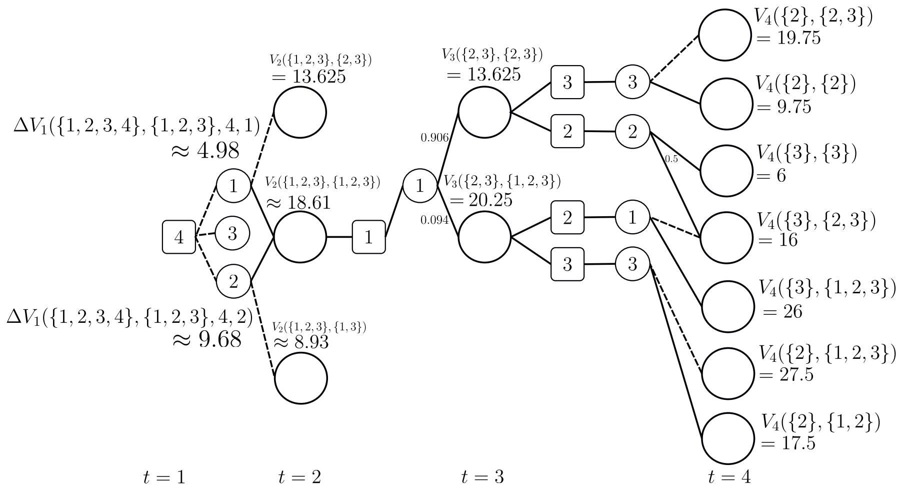

Fig. 6 Tree diagram of the example instance of Example 2 (with the additional delivery location)

OD4 might arrive in periods 1 and 2 respectively. The arrival probabilities are defined as follows: OD2 has a probability of 0.5 in periods 2 and 3, while OD3's probabilities are 0.1 in period 2 and 0.5 in period 3.

Notably, the avoided costs for delivery location C2 in period 1 are relatively high at 9.975. This outcome arises from the only viable service scenario for delivery location C2: OD3 must arrive in period 2, which occurs with a probability of 0.1. Upon arrival, OD3 is presented with an offer to serve C1, which he accepts with certainty (probability of 1), due to the compensation of 5. Following this, OD2's arrival in period 3 (with a probability of 0.5) is necessary. OD2 must then accept an offer of 9 to serve C2 (which occurs with a probability of 0.5). The cumulative probability of these events leading to the service of C2 is calculated as $0.1 \cdot 1 \cdot 0.5 \cdot 0.5 = 0.025$. This implies that there's a 2.5% chance for this sequence to unfold, resulting in a cost saving of 1 for serving C2 under these specific conditions. Advancing to the next period without altering the state renders it impossible for delivery location C2 to be served in the next period, as both ODs are offered delivery location C1 upon their arrival in period 3. As a result, the avoided costs for serving delivery location C2 increase to 10 in period 2.

The avoided costs for delivery location C1 in period 1 amount to 5.35, reflecting primarily the immediate benefit of eliminating the need to serve C1 in later states. A portion of these avoided costs, although smaller, is attributed to the enhanced potential of serving C2 in the last two periods 3 and 4, assuming C1 has been serviced already. Moving on to period 2, we observe a decline of C1's avoided costs to 5.25. This decline is primarily due to the reduced opportunity for serving C2 in absence of C1, given the diminishing time frame. With only period 4 remaining, the probability and, consequently, the strategic value of serving C2 decrease. A tree diagram with the effect can be seen in figure 7.

□

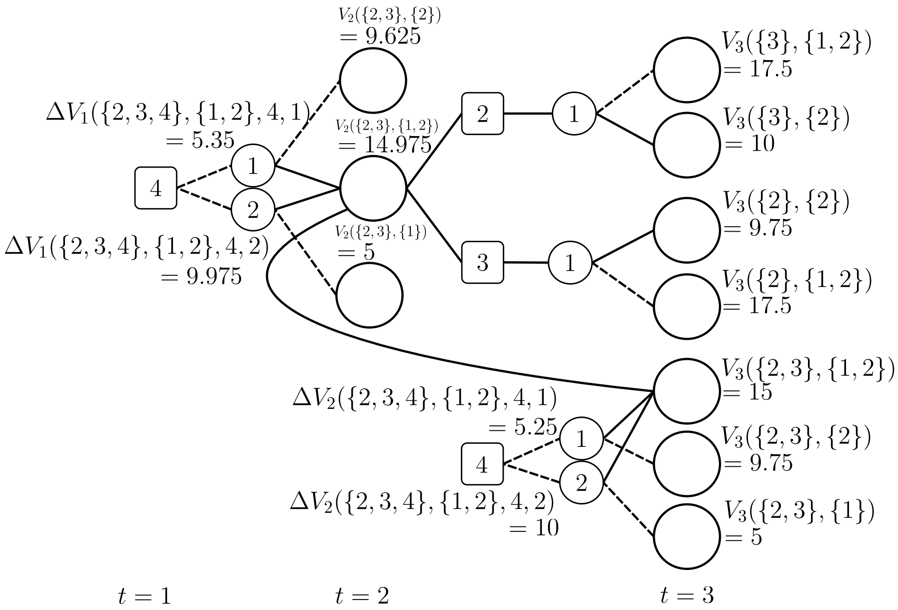

Fig. 7 Tree diagram of the example instance of Example 3 (progressing t)

Proposition 4 underscores the complexity and context-specific nature of making operational decisions. This complexity arises because the value of avoided costs does not have a simple relationship with the state variables such as the available delivery locations, the remaining ODs, and the number of periods yet to come. The valuation is influenced by a multitude of factors, including but not limited to, the spatial distribution of locations, the timing and sequence of ODs, and the direct and indirect relationship between various delivery locations and ODs. As such, reducing state variables does not straightforwardly lead to higher or lower avoided costs, but instead, the impact is contingent on the interplay of various factors at that moment in time.

5 Solution methods

In this Section, we develop two algorithms that are used to approximate the avoided costs $\Delta V_t(S_t^X, c)$, associated with the costs that are avoided by removing a single delivery location c from the pre-arrival state of the next period. The first algorithm employs parametric value function approximation (VFA), rendering it a learning-based approach suitable for a broad spectrum of IC distributions discussed in section 3. The second algorithm, in contrast, adopts a fluid approximation (FA). It demonstrates effective predictive capabilities particularly when the IC follows a uniform distribution. Both algorithms have undergone rigorous testing in the simulation study outlined in Section 6.

5.1 Parametric value function approximation

We propose a parametric VFA utilizing basis functions to effectively approximate avoided costs. The objective is to map a state S to a corresponding value $x \in \mathbb{R}$ using a VFA. As the exact value of the state S is computationally impractical to determine, a parametric VFA approximates the state by employing a linear combination of features, also known as basis functions. These basis functions are scaled by a set of parameters, commonly referred to as weights, which are learned during a dedicated learning phase. Once the learning phase concludes, the approximated value of the state S can be computed by building the linear combination of basis functions. To be effective, the basis functions must encapsulate sufficient information about the state for reliable predictions regarding its true value. The challenge lies in identifying a set of functions that adequately represent the structure of all states. The advantage of employing a parametric value function approximation is the ability to leverage the inherent structure of the state. This not only facilitates efficient memory usage but also results in shorter computation times.

Based on the analytical insights discussed in Section 4.1, we are primarily interested in the estimation of avoided costs. This estimation can be approached through two viable methods. The first method involves estimating the value functions as an initial step, followed by the calculation of avoided costs based on these estimations. The second method bypasses the preliminary estimation of value functions, focusing directly on estimating the avoided costs themselves. In our VFA, we found the latter approach more suitable.

5.1.1 The basis functions

Selecting appropriate basis functions is crucial for the effective learning of meaningful weights and achieving a robust estimation of the avoided costs. In Section 4.2, we observed that avoided costs result from a complex interplay of various factors. Among these factors were the spatial distribution of delivery locations and the destinations of ODs. Moreover, we found that the dynamics and arrival probabilities of ODs play a significant role. Motivated by these insights, we define our basis functions to capture the availability of ODs, while the weights are later designed to accurately reflect the relationship between delivery locations and ODs.

For each remaining OD and for every period within the planning horizon, a specific basis function is dedicated to reflect the total arrival probability of this OD over the remaining horizon. Specifically, we define the basis functions as

$$\phi_o(t) = \mathbb{P}_{Arrival}(o, t+1). \tag{16}$$

While the basis function is not dependent on the delivery location c, our selection of weights, denoted by η, ensures that a distinct weight, denoted by η_{oc}, corresponds to each combination of a delivery location c and an OD o. Combining basis functions, weights, and the end-of-horizon delivery fee κ_c results in our formulation of estimated avoided costs:

$$\overline{\Delta V}_t(S_t^A, \tilde{o}, c|\eta) = \kappa_c - \sum_{o \in \mathcal{O} \setminus \{\tilde{o}\}} \eta_{oc} \cdot \phi_o(t). \tag{17}$$

It's important to note that $\tilde{o}$ in this context refers to the OD that arrived in the pre-decision state S_t^X. These formulations of estimated avoided costs and basis functions ensure that the avoided costs in the last period equal κ_c, accurately reflecting the real avoided costs and thereby avoiding systematical errors.

5.1.2 Iterative learning-phase

The set of weights is initialized with an initial belief, denoted as η_0. To evaluate the current set of weights η_n in iteration n, the entire arrival process is simulated a total of Q times. Decisions are made throughout the simulation based on the current set of weights. After completing Q simulation runs, a regression model is employed to update the weights, aiming at a better alignment with observed costs. Updated weights η_{n+1} are subsequently evaluated in iteration $n+1$, until the final iteration N ends with the final set of weights.

During a single simulation run, the avoided costs for each delivery location c are computed using the current set of weights η_n and the relevant basis functions $\phi_o(t)$ (which depend on the remaining ODs and the current period), in accordance with equation (17). Subsequently, the optimal delivery location to offer in state S_t^X can be determined with $c^* = \text{argmax}_{c \in \mathcal{X}_C(S_t^X)} \{\overline{\Delta V}_t(S_t^X, c|\eta_t) - a_{oct}\}$, as outlined in Proposition 2. Following the optimality condition provided in Proposition 1, the optimal compensation r^* can be efficiently determined using numerical methods.

After each simulation run q, a cost is associated with each delivery location. This cost is either κ_c for locations not served by an OD, or the compensation paid to an OD o for the delivery to location c. Other crucial aspects for accurately updating the weights are the period during which a delivery location is served, the specific OD who served c, and the remaining ODs at that state. The service period is either the one in which the OD accepted the delivery location or $T+1$ if the location was not served by an OD during that run.

To incorporate this comprehensive data into our subsequent regression model, we introduce additional notations:

- Let r_{qct} denote the costs observed for delivery location c during simulation run q at period t, and calculate the total costs across all simulations as $\bar{r}_{ct} = \sum_{q \in Q} r_{qct}$.
- Let h_{qct} denote an auxiliary variable that is 1 if delivery location c was served in period t in a simulation run q. The total number of servings of c in t during one iteration is denoted by $\bar{h}_{ct} = \sum_{q \in Q} h_{qct}$. In instances where delivery location c is not served in a particular run, $\bar{h}_{cT+1}$ is increased by 1.
- Let τ_{qc} denote the period in which delivery location c was served during simulation run q.
- Let $\mathcal{O}_{qt}$ denote the set of remaining ODs in run q and period t.
- Let o_{qt} denote the OD who arrived in run q at period t.

To update the weights for a better alignment with the observed costs after Q simulation runs, we introduce the following regression model, inspired by a similar regression model that has been used by Koch and Klein (2020) in the context of dynamic pricing in attended home delivery.

$$
\begin{aligned}
min \quad & 0.5 \cdot \sum_{q=1}^{Q} \sum_{c \in \mathcal{C}} \epsilon_{qc}^2 && (18) \\
s.t. \quad & \epsilon_{qc} = \frac{\sum_{t=\tau_{qc}+1}^{T+1} \bar{r}_{ct}}{\sum_{t=\tau_{qc}+1}^{T+1} \bar{h}_{ct}} - \kappa_c + \sum_{o' \in \mathcal{O}_{q\tau_{qc}} \setminus \{o_{qc}\}} \eta_{oc} \cdot \phi_o(\tau_{qc}) \quad \forall\, q \in \{1, ..., Q\}, c \in \mathcal{C} && (19) \\
& \epsilon_{qc} \in \mathbb{R} \qquad \forall\, q \in \{1, ..., Q\}, c \in \mathcal{C} && (20) \\
& \eta_{oc} \geq 0 \qquad \forall\, o \in \mathcal{O}, c \in \mathcal{C} && (21)
\end{aligned}
$$

The regression model, as outlined in (18)-(21), treats the weights η_{oc} as decision variables. These weights are optimized to minimize the squared error between observed and predicted costs (refer to (19)). This minimization process ensures that the estimated costs align closely with actual outcomes.

After conducting Q simulation runs with a fixed set of weights, the regression model is employed to align the weights with the current observations. In the subsequent iteration, the updated weights are used in place of the previous weights to achieve a more accurate approximation of the avoided costs. This procedure iterates N times to progressively obtain improved weights. Following the Nth iteration, the current weights are applied to approximate the avoided costs in an online scenario.

A higher value of Q provides more information on the efficacy of our set of weights η_n. However, higher values of Q also extend the completion time of the learning process. Similarly, higher values of N entail more evaluations of η, while also increasing the learning time.

5.1.3 Interpretation of the weights and the relationship to the basis functions

In Section 4.2, we discussed the complex relationship between delivery locations and ODs. This relationship is effectively captured by the set of weights η, where an often observed effect in our numerical study was the strong linkage between a detour (an OD had to make for serving a delivery location) and the assigned weight. Particularly, a smaller detour usually resulted in a higher weight, emphasizing the cost-reducing nature of ODs who encounter only a small detour while serving a delivery location.

Furthermore, these weights also reflect the presence or absence of other ODs who share a similar effectiveness in serving the same delivery location. Our algorithm learns from instances where multiple ODs frequently visit certain delivery locations, and it adjusts the weights accordingly. Usually, this leads to a decrease of the corresponding weights, highlighting the increased competition in serving this delivery location and the reduced reliance on only a few ODs.

Our basis functions, defined as the remaining arrival probabilities of ODs, reflect the importance of specific OD availability, outlined in Section 4.2. Additionally, they ensure that the last period is associated with avoided costs of κ_c. This is a desirable feature as it correctly captures that no further savings can be made by waiting for another OD to arrive, and it thereby avoids to be burdened with systematical errors. Lastly, the structure of the approximated avoided costs mimics the structure of the true avoided costs. To show that, we utilize an example instance where coordinates are randomly selected from integers ranging between 0 and 10. The parameters set for this analysis include $T = 20$, $\kappa_c = 10\ \forall c$, and depot coordinates at $(5, 5)$. The IC is uniformly distributed with bounds $a = 0.5u_{oc} + 1$ and $b = 0.5u_{oc} + 2$ and the arrival probabilities are 0.1 in each period for each OD. In our graphical analysis, shown in Figure 8, the avoided costs over time are depicted on the left with the underlying instance coordinates on the right. On the left graph, the darker lines represent the true avoided costs calculated through a recursive computation with full enumeration. In contrast, the lighter lines indicate the approximated avoided costs derived from the VFA algorithm, which underwent 12 iterations with 2000 runs each. Extensive computational studies have revealed a "normal" structure of the avoided costs, which we aim to mimic through the VFA. This "normal" pattern typically shows an increasing trend over time. By selecting basis functions that progressively decrease with the remaining probability of arrival, our approximate avoided costs successfully mimic this normal structure.

However, our example and the proof of Proposition 4 illustrate that the structure of the true avoided costs can deviate from the norm. This deviation is particularly evident at delivery location 1, where the avoided costs tend to increase at a decreasing rate. The graph illustrates that for most parts, the structure of the true avoided costs is well captured by our VFA. This is especially clear for delivery locations 3 and 4. While the true avoided costs at delivery location 4 exhibit a slight deviation towards the end of the observed time horizon, the overall structure aligns well with that of the approximated avoided costs. The same holds for delivery location 3. Although the approximation does not perfectly mirror the true avoided costs, it proves to be a valuable tool. After a brief learning phase, it can be applied in an online operation to offer adequate compensations for emerging ODs, regardless of the instance's coordinates' structure.

In essence, our decision to utilize VFA for estimating avoided costs in each state is founded on its ability to capture relationships between delivery locations and ODs, account for the presence of competition, ensure desirable properties at the last period, and maintain structural congruence with observed data in smaller instances. These factors collectively make VFA an effective tool within our analytical framework. Consequently, we will test this approach in our numerical study in Section 6 and compare its performance with the performance of other mechanisms.

5.2 Fluid approximation

In this subsection, we introduce an algorithm that utilizes a substitute program to approximate the avoided costs $\Delta V_t(S_t^X, c)$. The procedure involves solving a simplified version of the Bellman equation. In this simplified version, the remaining time

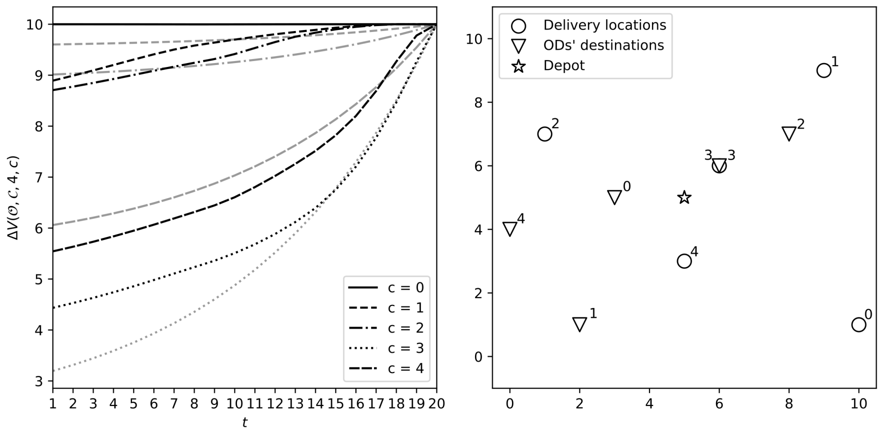

Fig. 8 The graph on the left compares the avoided costs and the approximate avoided costs in an instance with 5 ODs and 5 delivery locations. The brighter lines represent the approximate avoided costs, while the darker lines depict the true avoided costs calculated by full enumeration. The different line types depict the different delivery locations, which are to be served from the full set of delivery locations. The right graph shows the instance's coordinates

---

Algorithm 1 trainBasisFunctions()

Require: $\mathcal{C}, \mathcal{O}, T, \lambda_o, N, Q$
  Initialize weights $\eta_{oc} \leftarrow 0 \ \forall \ o, c$
  for $n \in \{0, ..., N\}$: **do**
    $\hat{r}_{ct}$, $\hat{h}_{ct} \leftarrow 0 \ \forall \ o, t$
    for $q \in \{1, ..., Q\}$: **do**
      r_{qct}, h_{qct}, τ_{qc}, $\mathcal{O}_{qt}, o_{qc} \leftarrow$ received by Q simulation runs with weights η_n
      $\hat{r}_{ct} \leftarrow \hat{r}_{ct} + r_{qct} \forall \ c, t$
      $\hat{h}_{ct} \leftarrow \hat{h}_{ct} + h_{qct} \forall \ c, t$
    end for
    Update η_{oc} values by calculating (18)-(21) with $\hat{r}_{ct}$, $\hat{h}_{ct}$, τ_{qc}, o_{qc} and $\mathcal{O}_{qt}$
  end for
  return $\eta_{oc} \ \forall \ o, c$

---

horizon is condensed to a single period and probabilities are represented by continuous fractions of the corresponding events. This adjustment allows ODs to both arrive and serve multiple delivery locations partially. As a consequence, the optimization model adopts a deterministic and continuous nature, which is why it is often referred to as fluid approximation (FA) (see, e.g., Maglaras and Meissner, 2006). The FA for approximating $V_t(\mathcal{O}_t, \mathcal{C}_t)$ is given by:

$$\begin{aligned}
\bar{V}_t^C(\mathcal{O}_t, \mathcal{C}_t) = \quad & \min_{r_{oc}} \sum_{c\in\mathcal{C}_t} \sum_{o\in\mathcal{O}_t} P_{\mathrm{Ar}}(o) \cdot P_{\mathrm{Ac}}(o,c,r_{oc}) \cdot r_{oc} \\
& + \sum_{c\in\mathcal{C}_t} \left(1 - \sum_{o\in\mathcal{O}_t} P_{\mathrm{Ar}}(o) \cdot P_{\mathrm{Ac}}(o,c,r_{oc})\right) \cdot \kappa_c && (22) \\
\text{s.t.} \quad & \sum_{o\in\mathcal{O}_t} P_{\mathrm{Ar}}(o) \cdot P_{\mathrm{Ac}}(o,c,r_{oc}) \leq 1 \qquad \forall c \in \mathcal{C}_t && (23) \\
& \sum_{c\in\mathcal{C}_t} P_{\mathrm{Ac}}(o,c,r_{oc}) \leq 1 \qquad \forall o \in \mathcal{O}_t && (24) \\
& r_{oc} \geq 0 \qquad \forall c \in \mathcal{C}_t, o \in \mathcal{O}_t && (25)
\end{aligned}$$

The parameter $P_{\mathrm{Ar}}(o)$ models the fraction of OD o that arrives within the remaining time, designed to resemble the corresponding arrival probability. Consequently, $P_{\mathrm{Ar}}(o)$ can be substituted with $1 - \mathbb{P}(\text{o does not arrive}) = 1 - (1 - \lambda_o)^{T+1-t}$, assuming the arrival probability of o is independent of the arrival of other ODs and time-homogeneous. When the arrival probabilities are time-heterogeneous, $P_{\mathrm{Ar}}(o)$ can be calculated by $1 - \prod_{s=t}^{T}(1 - \lambda_{os})$. The function $P_{\mathrm{Ac}}(o, c, r_{oc})$ represents the fraction of OD o that serves delivery location c for compensation r_{oc}. This function mirrors the acceptance probability of OD o and can be substituted by $F_{IC_{oct}}(r_{oc} - a_{oct})$. When applying the continuous uniform distribution with lower bound greater than 0, the model transforms into a quadratic program and can be solved efficiently with standard algorithms. The decision variable r_{oc} denotes the compensation offered to OD o for serving delivery location c. Given that the model does not determine a specific pairing of delivery locations with ODs, this variable is established for each possible OD-delivery location pair. The objective function (22) aims to minimize the approximate expected costs for the remaining time horizon. The first term represents the expected compensations paid to ODs. The second term accounts for the expected costs of serving the delivery locations remaining at $T + 1$. Constraint (23) ensures that the demand of a customer cannot be over-served. Constraint (24) limits an OD o to a total delivery of one, summing up all (fractional) deliveries to different delivery locations. Constraint (25) allows only non-negative compensations.

The avoided costs of serving delivery location c in state S_t^X can be approximated using different methods. One variation, proposed by Bertsimas and Popescu (2003), involves calculating the difference of the expected costs via the FA with and without delivery location c:

$$\Delta\bar{V}_t^{FA}(S_t^X, c) = \bar{V}_{t+1}^{FA}\left(\mathcal{O}_t \setminus \{o\}, \mathcal{C}_t\right) - \bar{V}_{t+1}^{FA}\left(\mathcal{O}_t \setminus \{o\}, \mathcal{C}_t \setminus \{c\}\right). \tag{26}$$

However, a drawback of this method is that to calculate the avoided costs of a specific delivery location, the FA substitute program needs to be solved separately for each delivery location. This significantly increases the computational effort, especially when dealing with a large number of potential delivery locations.

Another method, which requires only one solution of the FA to obtain an approximation of the avoided costs for all remaining delivery locations, is to utilize shadow prices. These shadow prices represent the optimal variables of the dual problem corresponding to (22) - (25). Particularly valuable are the shadow prices associated with constraints (23). In this application, the values of the shadow prices can be interpreted as the extent to which the retailer prioritizes assigning the associated delivery location to OD's destinations. delivery locations that are separated from the OD's direct routes from the depot tend to incur higher supply costs. In the optimal solution of the FA, the constraints of these delivery locations are associated with a shadow price of 0 since constraint (23) is not binding. These customers can be assigned avoided costs of $\Delta\bar{V}_t^{SP}(S_t^X, c) = \kappa_c$. The delivery locations for which the constraint is indeed binding have positive shadow prices associated with them. The approximate avoided costs of these delivery locations are

$$\Delta\bar{V}_t^{SP}(S_t^X, c) = \kappa_c - \zeta_c. \tag{27}$$

Here, ζ_c represents the shadow price of constraint (23) associated with delivery location c.

The inputs of (22) - (25) are subsets $\mathcal{O}' \subseteq \mathcal{O}_t$ and $\mathcal{C}' \subseteq \mathcal{C}_t$. Instead of utilizing the entire set of ODs and delivery locations as the input for the FA, one might opt for a subset of the mentioned sets. This approach offers the advantage of reducing the complexity of the optimization problem and eliminating redundancy. In the next subsection, we present a straightforward method for dividing the complete set of remaining ODs and delivery locations into a more manageable subset, while retaining relevant knowledge.

5.2.1 Generation of relevant subsets

It is improbable that an arriving OD o is willing to make a detour to delivery locations, resulting in unreasonably long detours. At least, unless the vendor does not provide an unreasonably high compensation, which renders the offer unprofitable. We reduce computational complexity by excluding delivery locations and ODs that are likely to never be affected by the absence of OD o and the offered delivery location c. This is particularly crucial when calculating the approximated avoided costs with equation (26). When using the shadow prices, calculating (22)-(25) once suffices to compute the avoided costs for all delivery locations. As a result, the need for complexity reduction is less pronounced in the latter approach.

In Proposition 4 and its proof, we have discussed the complex interplay between different ODs and delivery locations. Besides the immediate effect ODs and delivery locations may exert on each other, we also have observed the indirect impact the presence of a specific OD or delivery location can have on others, showcasing a competition effect. Therefore, we opted not to base these subsets solely on the detour distance to a specific OD, as such an approach could overlook ODs that, despite not being in close proximity to the delivery location, are nonetheless willing to serve it in the absence of better alternatives. For clarity, we introduce the notation $c_o^i(\mathcal{C})$ for the delivery location for which OD o has the ith shortest detour among all delivery locations in $\mathcal{C}$.

When deciding on the delivery location to offer during a pre-decision state, the subset of relevant ODs $\mathcal{O}'$ and relevant delivery locations $\mathcal{C}'$ is created by identifying neighboring sets of ODs' destinations and delivery locations. For this, we define the set of neighboring OD destinations of a delivery location c as the subset $\mathcal{O}'(\mathcal{O}, \mathcal{C}, c) = \{o' \in \mathcal{O} | \, c^1_{o'}(\mathcal{C}) = c\}$. This set can be described as the set of all OD destinations, that have delivery location c as their shortest detour delivery location. The set of neighboring delivery locations of a delivery location c is defined as $\mathcal{C}'(\mathcal{O}, \mathcal{C}, c) = \{c' \in \mathcal{C} | \, \exists o \in \mathcal{O}'(\mathcal{O}, \mathcal{C}, c) \text{ such that } (c^2_o(\mathcal{C}) = c') \vee (c^3_o(\mathcal{C}) = c')\}$. This set can be described as the set of delivery locations, that are the second or third priority of any OD o that has delivery location c as their priority considering the detour. We define the (first order) neighborhood of a delivery location c as $\mathcal{N}_1(\mathcal{O}, \mathcal{C}, c) = (\mathcal{O}'(\mathcal{O}, \mathcal{C}, c), \mathcal{C}'(\mathcal{O}, \mathcal{C}, c) \cup \{c\})$. The first input set of the neighborhood represents the neighboring ODs including o and the second one the neighboring delivery locations including c.

Depending on the setting, the neighborhood might be small or even empty. The size of the neighborhood can be increased by considering a higher-degree neighborhood. The second-order neighborhood is created by evaluating the neighborhood of every delivery location in $\mathcal{N}_1(\mathcal{O}, \mathcal{C}, c)$ and combining all the resulting neighborhoods into one large neighborhood. One could further increase the size of the neighborhood by considering neighborhoods of any degree k, adding the first-order neighborhood of delivery locations included in the last iteration iteratively, i.e. $\mathcal{N}_k(\mathcal{O}, \mathcal{C}, c) = \bigcup_{c':(\cdot,c') \in \mathcal{N}_{k-1}(\mathcal{O},\mathcal{C},c)} \mathcal{N}_1(\mathcal{O}, \mathcal{C}, c')$. However, we found the second-order neighborhood to be sufficient to achieve high accuracy in predicting the avoided costs. A visual representation of the neighborhood generation is depicted in Figure 9. In this example, the second-degree neighborhood is created.

The pseudo-code for the simulation of the algorithm to measure its performance is depicted in Algorithm 2. The algorithm uses discrete periods but can easily be adapted to a continuous time setting.

5.2.2 Monotonicity properties

Since our goal is to approximate the costs of being in state S^A_t, the model should depict characteristics of our true value function. The following lemma states that the same monotonicities as displayed in Proposition 3 hold.

Lemma 3. *For any pre-arrival state $S^X_t = (\mathcal{O}_t, \mathcal{C}_t)$, it holds:*

1. $\bar{V}_t(\mathcal{O}_t, \tilde{\mathcal{C}}_t) \leq \bar{V}_t(\mathcal{O}_t, \mathcal{C}_t)$ *with* $\tilde{\mathcal{C}}_t \subset \mathcal{C}_t$
2. $\bar{V}_t(\tilde{\mathcal{O}}_t, \mathcal{C}_t) \geq \bar{V}_t(\mathcal{O}_t, \mathcal{C}_t)$ *with* $\tilde{\mathcal{O}}_t \subset \mathcal{O}_t$
3. $\bar{V}_t(\mathcal{O}_t, \mathcal{C}_t) \leq \bar{V}_{t+1}(\mathcal{O}_t, \mathcal{C}_t)$ *with* $t \leq T$

Proof. We will prove each statement individually. To establish the validity of the first and second statements, it is sufficient to demonstrate the assertions for any subsets $\tilde{\mathcal{C}}_t \subset \mathcal{C}_t$ and $\tilde{\mathcal{O}}_t \subset \mathcal{O}_t$, where $\mathcal{C}_t \setminus \tilde{\mathcal{C}}_t = \{\tilde{c}_t\}$ and $\mathcal{O}_t \setminus \tilde{\mathcal{O}}_t = \{\tilde{o}_t\}$ for any $\tilde{c}_t$ and $\tilde{o}_t$, respectively. By consistently applying the arguments outlined below for the first and second statements, we can verify the more general assertion of the proposition.

1. Expanding $\tilde{\mathcal{C}}_t$ to $\mathcal{C}_t$ by incorporating $\tilde{c}_t$ alters the formulation of the fluid approximation as follows: Firstly, we introduce the decision variables $r_{o\tilde{c}_t} \geq 0$ to the

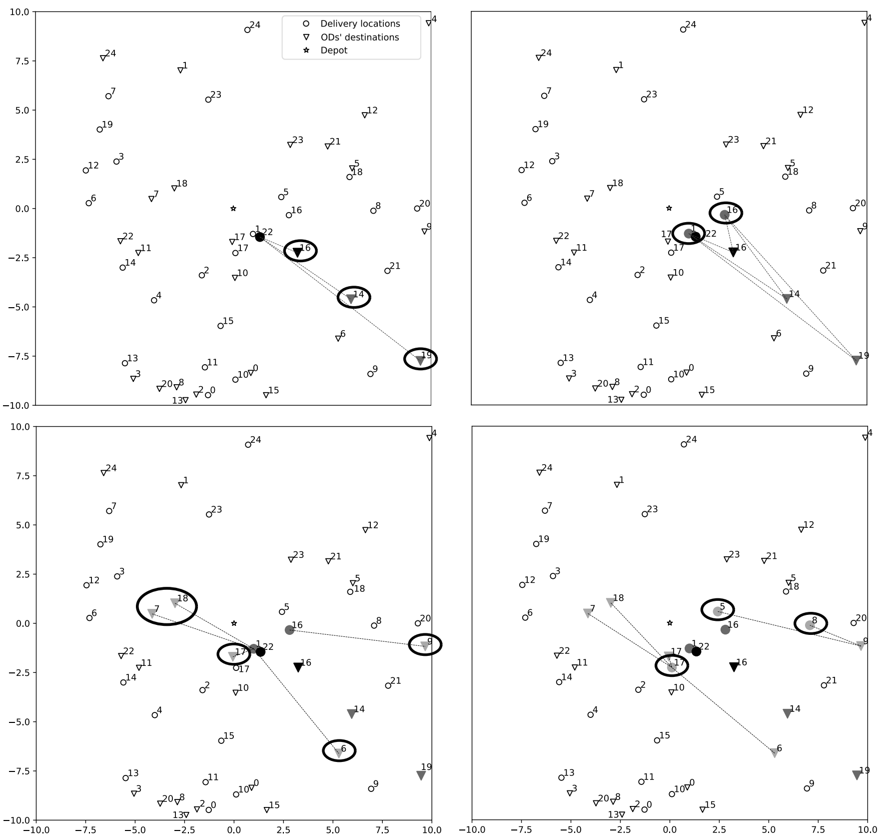

Fig. 9 The subsets of relevant delivery locations ($\bar{\mathcal{C}}$) and the subset of relevant OD destinations ($\bar{\mathcal{O}}$) are generated iteratively, as demonstrated through a series of figures. The first figure (upper left) illustrates delivery location 1 and the three ODs (namely, 14, 16, and 19) for whom delivery location 1 represents the shortest detour. The second figure (upper right) identifies delivery locations 16 and 22 as significant because they represent the second and third most favorable detour options for the ODs identified in the first step. In this figure, the filled symbols collectively represent the first-order neighborhood of delivery location 1. The third figure (lower left) continues this iterative process by beginning the construction of the first-order neighborhood for delivery locations 16 and 22, which were added in the previous step. The final figure (lower right) showcases the completion of the first-order neighborhood creation for delivery locations 16 and 22. The aggregation of these first-order neighborhoods, along with that of delivery location 1, forms the second-order neighborhood of delivery location 1, which is depicted in the last figure with filled symbols once more.

optimization model. Secondly, a non-negative term is added to the objective function 22. Specifically,

$$
\begin{aligned}
&\sum_{c\in\mathcal{C}_t}\sum_{o\in\mathcal{O}_t} P_{\mathrm{Ar}}(o)\cdot P_{\mathrm{Ac}}(o,c,r_{oc})\cdot r_{oc} + \sum_{c\in\mathcal{C}_t}\left(1-\sum_{o\in\mathcal{O}_t} P_{\mathrm{Ar}}(o)\cdot P_{\mathrm{Ac}}(o,c,r_{oc})\right)\cdot\kappa_c \\
=&\sum_{c\in\tilde{\mathcal{C}}_t}\sum_{o\in\mathcal{O}_t} P_{\mathrm{Ar}}(o)\cdot P_{\mathrm{Ac}}(o,c,r_{oc})\cdot r_{oc} + \sum_{c\in\tilde{\mathcal{C}}_t}\left(1-\sum_{o\in\mathcal{O}_t} P_{\mathrm{Ar}}(o)\cdot P_{\mathrm{Ac}}(o,c,r_{oc})\right)\cdot\kappa_c \\
&\quad+\sum_{o\in\mathcal{O}_t} P_{\mathrm{Ar}}(o)\cdot P_{\mathrm{Ac}}(o,\tilde{c}_t\, r_{o\tilde{c}_t})\cdot r_{o\tilde{c}_t} + \left(1-\sum_{o\in\mathcal{O}_t} P_{\mathrm{Ar}}(o)\cdot P_{\mathrm{Ac}}(o,\tilde{c}_t\, r_{o\tilde{c}_t})\right)\cdot\kappa_{\tilde{c}_t}
\end{aligned}
\tag{28}
$$

Thirdly, the set of constraints specified in (23) is enriched with the additional constraint $\sum_{o\in\mathcal{O}_t} P_{\mathrm{Ar}}(o)\cdot P_{\mathrm{Ac}}(o,\tilde{c}_t,r_{o\tilde{c}_t}) \leq 1$. Lastly, the non-negative term $P_{\mathrm{Ac}}(o,\tilde{c}_t\, r_{o\tilde{c}_t})$ is added to the left-side of constraints (24). Consequently, the optimal solution that yields $\bar{V}_t(\mathcal{O}_t,\mathcal{C}_t)$ remains feasible, upon removing the decision variable $r_{o\tilde{c}_t}$, for the optimization problem pertaining to state $(\mathcal{O}_t,\tilde{\mathcal{C}}_t)$. Furthermore, we can infer that the objective value $\bar{V}_t(\mathcal{O}_t,\mathcal{C}_t)$ exceeds the value this feasible solution yields for the state $(\mathcal{O}_t,\tilde{\mathcal{C}}_t)$. Since this suboptimal but feasible solution is below $\bar{V}_t(\mathcal{O}_t,\mathcal{C}_t)$, it follows that the optimal solution is also below this value by definition.

2. This proof follows a similar logic to the one discussed above. By expanding the set of ODs from $\tilde{\mathcal{O}}_t$ to $\mathcal{O}_t$ by including $\tilde{o}_t$, the fluid approximation undergoes analogous changes as outlined previously. Decision variables $r_{\tilde{o}_t c}\geq 0$ are introduced into the model. A term, which is non-positive in the optimal solution, is added to the objective function: $\sum_{c\in\mathcal{C}_t} P_{\mathrm{Ar}}(\tilde{o}_t)\cdot P_{\mathrm{Ac}}(\tilde{o}_t,c,r_{\tilde{o}_t c})\cdot(r_{\tilde{o}_t c}-\kappa_c)$. Additionally, the constraints in 23 and 24 are adjusted to accommodate the new set of decision variables. It is crucial to note that the optimal solution for the state $(\tilde{\mathcal{O}}_t,\mathcal{C}_t)$ can be transformed into a feasible solution for the state $(\mathcal{O}_t,\mathcal{C}_t)$ by setting $r_{\tilde{o}_t c}=0$ and retaining any other value over. Given that $P_{\mathrm{Ac}}(\tilde{o}_t,c,0)=0$, this feasible solution yields the same objective value for the state $(\mathcal{O}_t,\mathcal{C}_t)$ as the optimal objective value $\bar{V}_t(\tilde{\mathcal{O}}_t,\mathcal{C}_t)$. Since we are dealing with a minimization problem, the optimal objective value is below this value. Thus, the statement holds true.
3. In this proof, we examine the implications of lowering $t+1$ to t. Employing the same methodology as before, we modify the optimal solution for $t+1$ to be feasible for t while resulting in a lower objective value than $\bar{V}_{t+1}(\mathcal{O}_t,\mathcal{C}_t)$. As per its definition, $P_{\mathrm{Ar}}(o)$ increases with a decrease in t. This gives rise to a dual impact on the optimization model: firstly, the value function decreases for any solution $r_{oc}\leq\kappa_c$. Secondly, constraints 23 become more stringent. So, how should we adapt the optimal solution of $t+1$?
   To adhere to the tightened constraints in 23, adjustments must be made to some of the r_{oc}. This entails lowering certain $r_{o\tilde{c}}\leq\kappa_{\tilde{c}}$ values when the constraint is violated for $\tilde{c}$. More specifically, we can select these decision variables in a manner that reconstructs the same probabilities $P_{\mathrm{Ar}}(o)\cdot P_{\mathrm{Ac}}(o,\tilde{c},r_{o\tilde{c}})$ for t as we initially had for $t+1$. Consequently, this particular customer location now carries lower expected costs, leading to a reduction in the objective value. The decision variables

concerning customer locations where the constraint is not violated can remain the same and still result in lower expected costs (due to the increase in the arrival probability). Overall, these alterations lead to a feasible solution for t, yielding lower expected costs than for $t+1$.

□

Algorithm 2 FA

Require: $\mathcal{C}, \mathcal{O}, T, \lambda_{ot}, IC_{oct}, k$
  Initialize total costs $\mathcal{G} \leftarrow 0$
  Simulate an arrival sequence $(o_t)_{t\in\mathcal{T}}$ ▷ o_t might be a no-show
  for $t \in \mathcal{T} \setminus \{T+1\}$ **do**
    $h \leftarrow o_t$
    for $c \in \mathcal{C}$ **do**
      $(\bar{\mathcal{O}}, \bar{\mathcal{C}}) \leftarrow \mathcal{N}_k(\mathcal{O}, \mathcal{C}, c)$
      $\Delta\bar{V}_t^{FA}(\bar{\mathcal{O}} \cup \{h\}, \bar{\mathcal{C}}, h, c) \leftarrow \bar{V}_{t+1}^{FA}\left(\bar{\mathcal{O}}, \bar{\mathcal{C}}\right) - \bar{V}_{t+1}^{FA}\left(\bar{\mathcal{O}}, \bar{\mathcal{C}} \setminus \{c\}\right)$
    end for
    $c^* \leftarrow \text{argmax}_{c\in\mathcal{C}}\{\Delta\bar{V}_t^{FA}(\bar{\mathcal{O}} \cup \{h\}, \bar{\mathcal{C}}, h, c) - a_{hct}\}$
    Determine r^* by equation (10)
    Offer (c^*, r^*) and observe whether h accepts
    if Yes **then**
      $\mathcal{G} \leftarrow \mathcal{G} + r^*; \mathcal{C} \leftarrow \mathcal{C} \setminus \{c^*\}$
      if $\mathcal{C} = \emptyset$ **then**
        return $\mathcal{G}$
      end if
    end if
    $\mathcal{O} \leftarrow \mathcal{O} \setminus \{h\}$
  end for
  $\mathcal{G} \leftarrow \mathcal{G} + \sum_{c\in\mathcal{C}} \kappa_c$
  return $\mathcal{G}$

6 Simulation study

The structure of our simulation study is as follows. We aim primarily to evaluate the performance of our algorithms through simulations based on instances that are typically encountered in the literature. This evaluation is conducted by comparing the performance, specifically average costs, of the algorithms in relation to each other and against benchmarks commonly used in the literature.
We then perform a more detailed analysis of the assignment decisions made by our algorithms and compare their decisions to those made by a myopic assignment strategy. This gives a better insight into how our algorithms make decisions, in relation to a benchmark strategy.

6.1 Instance generation

The instances on which we test our algorithms are varied across several categories to assess the robustness of the compensation-setting strategies and to identify conditions under which performance excels. The instance parameters differ in the categories shown below:

Number of ODs

In all of our test instances, we set the number of delivery locations, C, equal to the number of OD destinations, O. We set the number of periods, T, to be equal to the number of ODs so that each OD may arrive, that is, $T = O$. We classify our instances into three different sizes: large instances with $O = 100$, medium-sized instances with $O = 50$, and small instances with $O = 25$. This allows us to systematically explore instances with varying numbers of ODs, enabling a nuanced examination of algorithmic performance across different instance scales.

City structure

For the present study, the delivery location coordinates are derived from the Solomon instances, specifically, the instances denoted as C-101 and R-101. Originally designed for the capacitated vehicle routing problem with time windows, the Solomon instances are used to model the structure of a city in related literature, most noteworthy by Archetti et al (2016).
To ensure a congruent structure in our study, the geographical coordinates of the OD's destinations are generated using a random distribution, replicating the inherent structure of the chosen instances. Specifically, in instances characterized by clustering, such as C-101, the OD's destinations are exclusively drawn from regions containing existing delivery locations. This approach ensures that the simulated OD pairs accurately mirror the spatial characteristics of the underlying Solomon instances.
To reduce the effect of outliers, we generate 5 instances for each instance size and city structure combination, by sampling random coordinates out of the coordinate sets. In our analysis we refer to the average results.

Arrival rates

The determination of arrival rates is critical to the predictive modeling of future OD arrivals. Given the constraint that the sum of arrival rates within a single period should not exceed 1, the arrival rate becomes directly linked to the number of ODs. To investigate the impact of different expected arrival scenarios (AR), we conduct tests with both high ($\lambda_o = \frac{1}{O}$) and low ($\lambda_o = \frac{0.5}{O}$) arrival rates. In our study, the arrival rate for all ODs are equal; this reflects scenarios where there is general awareness of the overall number of expected OD arrivals during the time horizon, but limited knowledge of the arrival probability associated with each OD.

Other parameters

The fixed costs, denoted as κ_c, associated with the delivery to a delivery location by a dedicated driver are maintained at a constant value of 10 for each location c. The

IC is modeled as a uniform distribution. Furthermore, the parameters for the IC are as follows: $a_{oct} = 1 + 0.5u_{oc}$ and $b_{oct} = 2 + 0.5u_{oc}$.

6.2 Compensation setting strategies

To demonstrate the significant advantages of the proposed algorithms over standard methods, we compare our algorithms with four benchmark strategies frequently used in the literature. To enhance distinguishability, we categorize all approaches into two distinct groups, based on their compensation strategy. The predictive group makes decisions by approximating the avoided costs using the methods discussed in sections 5.1 and 5.2. Specifically, we compare the value function approximation algorithm (subsequently labeled as VFA) and the fluid approximation algorithms (subsequently labeled as FA and FA-SP, with the latter being based on shadow prices). We use the second-degree neighborhood as relevant subsets for decision-making (refer to 5.2.1) in FA. The myopic group contains benchmark rule-based decision-making policies commonly employed in the literature.

Benchmark 1: Fixed compensation

As the name suggests, the first benchmark relies on a compensation that is determined at the start of the process and is offered to any arriving OD.

Benchmark 2: Distance-based compensation

The second benchmark's compensation strategy is to offer $\rho \cdot d_{0c}$ to any OD to serve delivery location c, where $\rho > 0$. Using the scaled distance is proposed in Archetti et al (2016) and has the advantage that the compensation is independent of the OD's destination.

Benchmark 3: Detour-based compensation

The third benchmark's compensation strategy is to offer $\rho \cdot u_{oc}$ to the OD o for serving delivery location c, where $\rho > 0$. This strategy, referred to as 'scaled detour' also incorporates knowledge about the OD's destinations to create a compensation that aims to closely resemble the OD's perceived true cost. In a deterministic setting, Archetti et al (2016) used this compensation scheme with the assumption that all ODs arrive and accept the offer when their detour does not exceed a predefined threshold.

Benchmark 4: Combination of fixed and detour-based compensation

The fourth benchmark's compensation strategy involves offering $\nu + \rho \cdot u_{oc}$ to OD o for serving delivery location c, where $\rho > 0$ and $\nu > 0$. This strategy is a combination of benchmarks 1 and 3 that combines a fixed with a detour-based component. This compensation strategy best aligns with the assumptions we selected for our IC and is therefore expected to outperform the other benchmark algorithms in the simulated settings.

Implementing a myopic strategy poses the challenge that the firm must determine precise parameters. Identifying optimal parameters in real-world scenarios can be inherently challenging. To address this issue in our simulation study, we conducted a thorough parameter search to identify a well-fitting set of parameters before simulating the arrival process. This approach ensures a fair comparison with the algorithms in the predictive group.

Myopic approaches calculate the compensation for an OD for a given delivery location using straightforward rules. However, the choice of which delivery location to propose is still to be decided. Our results, as per Proposition 2, show that the optimal delivery location to offer is the one with the largest difference between avoided costs and known inconveniences. Since myopic compensation strategies refrain from approximating avoided costs, the decision about the delivery location is made purely based on the minimal known inconvenience. Therefore, in our scenarios, the offered delivery location is always the one with the shortest detour for the arriving OD. To provide a fair comparison, the compensation is lowered to the known upper bound of the IC when the myopic compensation is above that upper bound. Table 2 provides a summary of all compared compensation strategies.

Table 2 Summary Table of the tested compensation-setting strategies

Comp. Strategy	**Group**	**Compensation**	**ΔV Appr.**	**Ref.**
FA	Predictive	(10)	(26)	-
FA-SP	Predictive	(10)	(27)	-
VFA	Predictive	(10)	(17)	-
B1	Myopic	$min(\rho, b_{oct} + a_{oct}), \rho > 0$	-	-
B2	Myopic	$min(\rho \cdot d_{0c}, b_{oct} + a_{oct})$	-	Archetti et al (2016), Le et al (2021)
B3	Myopic	$min(\rho \cdot u_{oc}, b_{oct} + a_{oct})$	-	Archetti et al (2016), Boysen et al (2021)
B4	Myopic	$min(\nu + \rho \cdot u_{oc}, b_{oct} + a_{oct})$	-	Dayarian and Savelsbergh (2020)

6.3 Algorithm performance

In evaluating algorithmic performance within our simulation study, a comprehensive set of indicators has been established to measure and compare the efficiency and effectiveness of different algorithms. The primary focus of these indicators is to quantify the economic impact and operational involvement of ODs in the delivery process. The indicators are defined as follows:

1. **Average costs:** Defined as AC(Algorithm, Instance), this metric represents the average total cost incurred across all runs for instances with identical characteristics. The average cost includes the sum of the compensation paid to ODs for each run and the cost of dedicated drivers. Since the firm aims at minimizing the cost, lower average costs are generally desirable.

Table 3 Comparison of the AC

CS	Size	AR	FA	FA-SP	VFA	B1	B2	B3	B4
C	25	1	183.72	183.79	182.04	185.04	210.63	241.02	184.62
C	25	0.5	199.80	199.30	199.78	202.70	232.42	241.88	202.42
C	50	1	337.23	338.35	336.47	337.09	433.52	457.86	336.41
C	50	0.5	387.83	387.87	389.24	389.66	452.37	467.03	389.27
C	100	1	627.78	626.87	627.17	629.19	979.25	1000.00	628.51
C	100	0.5	742.56	743.72	752.74	745.70	984.15	1000.00	745.42
R	25	1	184.67	183.44	181.55	182.76	213.44	243.29	182.25
R	25	0.5	199.35	199.45	199.17	203.49	214.56	230.14	203.23
R	50	1	342.26	342.72	340.44	343.31	375.82	437.08	342.70
R	50	0.5	385.73	386.04	384.81	388.62	474.14	484.43	388.15
R	100	1	656.74	654.97	640.85	659.17	982.44	1000.00	658.22
R	100	0.5	751.42	750.11	749.62	756.53	986.95	1000.00	756.32

2. **Average payment (compensation) per OD:** Denoted as AP(Algorithm, Instance), this metric measures the average compensation received by an OD for successful deliveries. To minimize costs, a firm may prefer lower average compensation, yet this approach entails trade-offs in certain contexts. On one hand, higher average payments result in more ODs accepting the compensation. That also means that the fixed cost κ_c is avoided for more delivery locations. The compensation strategy is more risk-averse towards the OD's acceptance decision. On the other hand, lower compensations might be sufficient to convince the OD to accept an offer. This leads to a more risk-affine compensation strategy, in the spirit of exploiting the OD's willingness to participate more aggressively.
3. **Average number of deployed ODs:** This indicator assesses the involvement level of ODs in serving delivery locations, providing the total average number of deployed ODs (TN). When only considering the firm's goal of minimizing the costs, the number of deployed ODs is a neutral indicator, since it does not say anything about the reduction of costs achieved by the avoided fixed costs κ_c. A higher number of OD deployments could however benefit the environment and reduce traffic in the city, which is generally desirable.

To facilitate a comparative analysis between different algorithms, the gap metric $Param\text{-}G(A1, A2, I) = 1 - \frac{Param(A1,I)}{Param(A2,I)}$ is introduced. This formula quantifies the relative performance discrepancy between two algorithms ($A1$ and $A2$) for a given parameter ($Param$) and instance (I). For instance, the gap in average costs between two algorithms is denoted as $AC\text{-}G(A1, A2, I) = 1 - \frac{AC(A1,I)}{AC(A2,I)}$. In cases where the instance is omitted as a parameter and no additional information is provided, the reference value defaults to the average across all instances.

6.3.1 Predictive group

In our examination of algorithmic performance within clustered instances (the first six lines in Tables 3-5), it becomes evident that there's no definitive frontrunner among the algorithms. This observation is quantitatively substantiated by the narrow average $AC\text{-}G$ over all clustered instances, whose average values do not surpass 0.1%. This impression shifts when analyzing scenarios with randomly distributed locations

Table 4 Comparison of the AP

CS	Size	AR	FA	FA-SP	VFA	B1	B2	B3	B4
C	25	1	4.59	4.71	4.46	4.37	4.35	8.75	4.29
C	25	0.5	4.29	4.34	4.25	4.31	4.54	6.37	4.30
C	50	1	4.22	4.30	4.00	4.14	4.38	6.32	4.15
C	50	0.5	4.11	4.14	4.05	4.03	4.23	5.33	4.08
C	100	1	3.93	3.94	3.67	3.85	3.00	nan	3.80
C	100	0.5	3.60	3.62	3.78	3.64	3.00	nan	3.63
R	25	1	4.40	4.42	4.03	4.19	4.39	8.95	4.15
R	25	0.5	4.05	4.11	3.99	4.29	4.72	7.34	4.22
R	50	1	4.04	4.13	3.88	3.78	4.22	7.11	3.87
R	50	0.5	3.85	3.89	3.77	3.77	3.83	4.90	3.78
R	100	1	3.84	3.83	3.49	3.71	3.00	nan	3.63
R	100	0.5	3.54	3.54	3.53	3.49	3.00	nan	3.51

Table 5 Comparison of the TN

CS	Size	AR	FA	FA-SP	VFA	B1	B2	B3	B4
C	25	1	12.26	12.52	12.27	11.55	6.96	7.20	11.45
C	25	0.5	8.79	8.96	8.73	8.31	3.22	2.24	8.34
C	50	1	28.14	28.36	27.26	27.82	11.83	11.46	27.96
C	50	0.5	19.06	19.14	18.60	18.49	8.26	7.05	18.69
C	100	1	61.35	61.58	58.94	60.25	2.96	0.00	59.93
C	100	0.5	40.21	40.15	39.76	39.96	2.26	0.00	39.94
R	25	1	11.66	11.92	11.46	11.57	6.52	6.39	11.59
R	25	0.5	8.51	8.58	8.46	8.15	6.71	7.48	8.09
R	50	1	26.48	26.77	26.08	25.20	21.50	21.76	25.67
R	50	0.5	18.58	18.64	18.48	17.88	4.19	3.06	17.98
R	100	1	55.69	55.88	55.13	54.19	2.51	0.00	53.63
R	100	0.5	38.47	38.68	38.72	37.38	1.86	0.00	37.53

(the last six lines in Tables 3-5). Here, the VFA consistently emerges as the superior algorithm across all six instances, further highlighted by an average $AC\text{-}G(\text{VFA}, \text{FA})$ of 0.87% and $AC\text{-}G(\text{VFA}, \text{FA-SP})$ of 0.73% over all randomly spread instances. This marks the VFA as particularly adept in managing dispersed delivery locations efficiently.

However, when delving into the comparative performance of the FA and FA-SP algorithms, the picture becomes less sharp. The FA-SP algorithm marginally outperforms the FA algorithm, as evidenced by an average $AC\text{-}G(\text{FA-SP}, \text{FA})$ of 0.06%. Although this is just a slight advantage, it gains significance in light of the FA algorithm's prolonged run time, which is 34.65% longer on average compared to the FA-SP (refer to table A1 in the appendix). This considerable delay positions the FA-SP algorithm favorably, particularly in time-sensitive operational environments.

In terms of average costs, the VFA emerged as a notable performer. In the following, we turn our attention on the next indicator, the average payment per OD. I this metric, VFA proves to offer the lowest average compensation for ODs in eleven of twelve scenarios. Contrary, the FA-SP exhibits in most scenarios the highest average compensation paid per OD. Upon closer examination of the average difference in compensation between VFA and FA-SP, denoted as $AC\text{-}G(\text{VFA}, \text{FA-SP})$, we observed a reduction in the average compensation per OD by 4.04% across all instances. This gap widened in non-clustered instances, where the difference in average compensation increased to approximately 5%. These findings underscore the effectiveness of VFA's more aggressive compensation and matching decisions, particularly in non-clustered instances, leading to lower average costs for the firm. However, it is crucial to note that this reduced average compensation under the VFA corresponded to a 1.65% decrease in the number of delivery locations served by ODs. As evidenced by the advantage in AC compared to the FA-SP algorithm, this trade-off highlights the VFA's superior balance between the number of served delivery locations and the height of compensation. Interestingly, while the FA-SP algorithm consistently resulted in higher average payments per OD compared to the FA, the overall average costs did not show a significant difference between these two algorithms. This observation suggests that the increased payments in the FA-SP do not necessarily translate into proportionally higher overall costs.

6.3.2 Myopic group

In every instance evaluated, algorithm B4 consistently secured the lowest average costs at 418.13, closely followed by B1 at 418.60 and then B2 at 544.98. Notably, B3 lagged significantly and had the least favorable average costs. Despite B4's leading performance in all cases, it's important to note that the average $AC\text{-}G(\mathrm{B4}, \mathrm{B1})$ was small, at just 0.19% across all instances. This proximity in the average costs across B1 and B4 is a result of their compensation structures, which tie them at a tight bound by the parameters of the IC framework. For example, B1 sets an average compensation of 6.22 for each OD, while B4 chooses a fixed compensation of 4.85 (refer to table A3), supplemented by an average detour scalar of 0.26 (refer to table A5). Although there are differences, the actual compensations are capped at the upper bound of the IC, when initial calculations exceed this limit. The operational difference between both approaches only begins to emerge when the ODs see substantial detours, especially those that are 3 or more, raising the bound of ICs considerably. Here, the different compensation models of B1 and B4 reveal their impact, introducing variability in the decision-making. However, these are exceptional cases, making the working environment predominantly homogeneous in most instances for the two mentioned benchmarks. The other two benchmarks, based solely on detour or distance, cannot address the fixed payment expectation component of the ODs due to their structure, which is why their performance falls short in comparison to B1 and B4. The consideration of a fixed payment expectation is especially important in instances with short detours, which are often found in densely populated areas.
For B3, there is a noticeable trend where no OD is deployed in larger settings, where arriving ODs are facing a rich bouquet of delivery locations. With many possibilities, there is usually a delivery location that can be served by this OD with only a short detour. Consequently, this benchmark offers a small compensation, too small to offset the fixed payment expectation component of the IC. Despite the challenges, B1 and B4 perform well due to their compensation strategies that align better with the assumptions on the IC in this study. This emphasizes the significance of designing algorithms that are aligned with real-world conditions in which they operate. In terms of the total number of ODs deployed, B1 and B4 are equally consistent, each

averaging 26.73. In contrast, B2 and B3 deploy only 6.56 and 5.55 ODs on average, respectively. Accompanied by the higher average payments per OD, B2 and B3 are conclusively non-competitive when compared to B1 and B4. B4's flexibility suggests superior adaptability in decision-making across instances, yet this does not significantly affect the average compensation per OD relative to B1, indicated by an average $AP\text{-}G(\mathrm{B4}, \mathrm{B1}) = 0.33\%$ over all instances.

6.3.3 Comparison of the predictive and myopic algorithms

In our comparative analysis of the predictive and myopic groups, we focused on the algorithms with the lowest average costs within each group. The analysis shows that the benefit of using the predictive algorithm is quantifiable, with $AC\text{-}G$ ranging from -0.02% to 2.64% across different instances. In particular, the occurrence of a negative gap indicates that myopic decisions can occasionally lead to a reduction in average costs, underscoring their potential utility in certain scenarios.
On a broader scale, the predictive algorithm shows an average benefit of 0.95% across all instances. This is especially significant given that the myopic algorithms underwent extensive parameter optimization before deployment to ensure that they were operating at optimal performance levels. However, performance varies significantly across different city structures. For example, in clustered instances, the average $AC\text{-}G$ is 0.66%. In random instances, it increases significantly to 1.24%. This variation is largely attributed to the superior performance of VFA within the predictive group, which effectively exploits the heterogeneous nature of the IC and the dynamics of competition among different ODs.
Interestingly, the $AC\text{-}G$ of 0.89% in high arrival rate instances is unexpectedly lower than the 1.01% observed in low arrival rate instances. This observation suggests that the impact of arrival rates on $AC\text{-}G$ is complex and may require further investigation through a larger simulation study. Moreover, the instance size significantly influences the performance benefits of the predictive algorithm. In clustered instances with a size of 25, the $AC\text{-}G$ reaches 1.47%, a stark contrast to the 0.25% in clustered instances with a different number of ODs. This increased benefit is primarily due to the algorithm's strategic allocation of delivery locations, a factor that becomes more important when the number of ODs, and therefore profitable matches, is limited. Conversely, in random instances, the $AC\text{-}G$ remains relatively stable across different instance sizes, indicating the algorithm's consistent performance in such city structures.
In our analysis of the VFA and B4 algorithms, which stand as representatives of their respective groups, we observe an average $AP\text{-}G(\mathrm{VFA}, \mathrm{B4})$ of 1.03%. This indicates that VFA not only yields the lowest average payments per OD among all examined predictive algorithms but also outperforms B4 in this metric. A plausible reason for this outcome is that while assigning ODs to the shortest detour locations might result in lower compensations in the short term, it regularly leads to scenarios where the only option is to offer exceptionally high compensations due to poor OD-delivery location pairings. Furthermore, the average $TN\text{-}G(\mathrm{VFA}, \mathrm{B4})$ is -1.71%, suggesting that, on average, the predictive algorithm engages a higher proportion of ODs. This outcome,

particularly in combination with AP, clearly demonstrates that VFA achieves significantly better OD-deliver location matchings compared to B4, explaining its efficiency in minimizing costs.

6.3.4 Simulation study summary

The analysis in this chapter demonstrates that utilizing a predictive algorithm is advantageous compared to using a myopic algorithm. This advantage is significant, indicating a robust benefit of using VFA in various operational contexts. The feasibility of implementing this algorithm is emphasized by its adaptability and the significant cost reduction it offers, especially in rural areas with a more dispersed population.
When evaluating benchmark algorithms, it is essential to thoroughly examine the IC structure. This fundamental understanding is critical for effectively tailoring compensation strategies. When the IC exhibits a relatively fixed nature, a fixed compensation or a hybrid model that integrates both fixed and detour-based compensations can be a viable and simplistic alternative to the predictive approach.
Lastly, it is important to note the limitations of compensation models that rely solely on distance or detour metrics. Under the conditions and structures examined in this study, these approaches are insufficient to provide the adaptability and efficiency required for optimal performance in our dynamic setting. Overall, benchmark strategies seem to be more sensitive regarding the considered setting than predictive algorithms.

6.4 Comparison of the assignment strategies

The decision in each pre-decision state involves determining the compensation and assigning a delivery location. In Figure 10, we illustrate the assignment decisions made by the myopic group in contrast to those made by FA. The thickness of a line in the visualization corresponds to the number of matches between the connected OD and the delivery location. The thicker the line, the more frequently the OD has visited that delivery location in 1000 simulation runs. The comparison is based on the R-101 instance with 25 fixed ODs and delivery locations, but varying OD arrival orders. This visualization demonstrates that predicting future OD arrivals is taken into account by the FA algorithm, leading to more sophisticated assignment decisions.
In Figure 10, we see some notable differences in the assignment decisions, highlighted by two examples, labeled as A and B. In example A, the encircled pair is matched about half as frequently in the myopic assignment strategy compared to the predictive strategy. This difference arises due to a delivery location with a lower detour for the highlighted OD, which the myopic approach strongly prefers, as indicated by the thick line between them. Given that other ODs could also efficiently serve this centrally located delivery location, the FA algorithm strategically reserves it for such occasions, thus matching the highlighted pair more frequently to reduce total costs.
In example B, the encircled delivery location offers minimal detours for two nearby ODs. However, with only one OD able to be assigned to this delivery location, the myopic approach assigns it to the OD that arrives first. The randomness of OD arrival orders leads to several potential pairings in this approach, as evidenced by the four

connecting lines in the left figure. Yet, this method does not guarantee the most cost-effective matches. In contrast, the predictive assignment strategy forms only two distinct pairings, reserving the encircled delivery location for the non-encircled OD when the encircled OD arrives first, optimizing the overall costs.

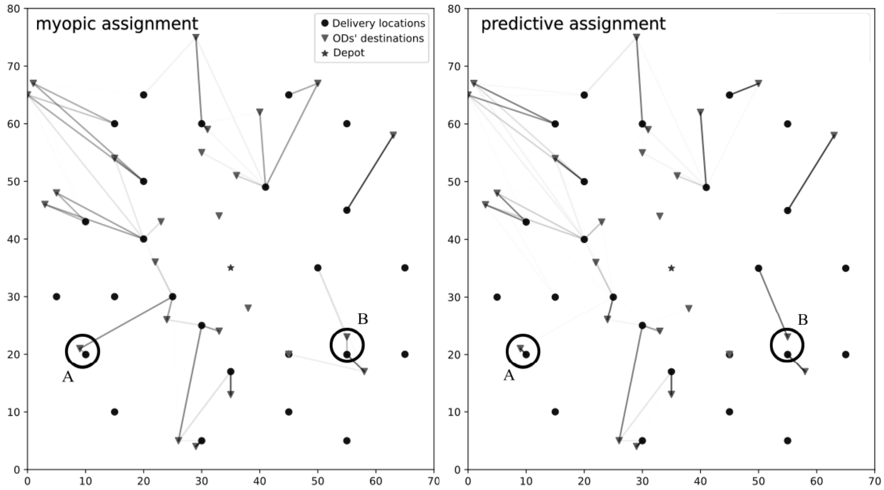

Fig. 10 Visual comparison of the assignment strategies of the myopic strategy (left) and the predictive strategy of the FA-Algorithm (right). The thickness of the lines represents the frequency in which a delivery location is visited by an OD.

7 Conclusion

In our article, we introduce an innovative approach to leveraging in-store customers, who may be willing to divert their planned route to deliver online orders for monetary compensation. While these occasional drivers present a cost-effective alternative to traditional dedicated drivers, they also pose additional challenges, arising from their unpredictive nature, manifested in their arrival time and decision-making. Our approach meets these characteristics by dynamically matching arriving occasional drivers with delivery tasks, incorporating individualized compensations that consider each occasional driver's specific circumstances. This approach, akin to a reverse dynamic pricing model, explicitly considers the stochasticity inherent in occasional drivers' availability and decision-making.

We prove several properties of the optimization problem, with a special focus on the optimal solution and avoided costs: First, we establish the existence of a unique optimum in the step-wise optimization. Furthermore, we provide a closed-form solution for scenarios where the IC is uniformly distributed, improving both computational efficiency and interpretability of the results. Second, we outline an efficient and simple method to find the optimal matching between OD and delivery locations, thereby refining the decision-making process. Finally, we shed light on the monotonic behavior

Table A1 Average time (in seconds) of one simulation run. Since one simulation run includes multiple decisions, the number needs to be divided by the number of OD arrivals to get the average time it took to generate the offer bundle

CS	Size	AR	FA	FA-SP	VFA	B1	B2	B3	B4
C	25	1	0.197	0.134	0.003	0.000	0.000	0.000	0.001
C	25	0.5	0.126	0.084	0.003	0.000	0.000	0.000	0.000
C	50	1	0.793	0.582	0.011	0.001	0.001	0.001	0.001
C	50	0.5	0.797	0.500	0.009	0.001	0.001	0.001	0.001
C	100	1	12.400	7.267	0.040	0.002	0.002	0.002	0.002
C	100	0.5	7.538	4.002	0.030	0.002	0.002	0.002	0.002
R	25	1	0.214	0.148	0.004	0.001	0.000	0.000	0.000
R	25	0.5	0.166	0.113	0.003	0.000	0.000	0.000	0.000
R	50	1	1.262	0.939	0.011	0.001	0.001	0.001	0.001
R	50	0.5	1.004	0.656	0.008	0.001	0.001	0.001	0.001
R	100	1	12.265	8.273	0.042	0.003	0.002	0.002	0.003
R	100	0.5	9.006	5.261	0.030	0.002	0.002	0.002	0.002

Table A2 Average OD arrivals

CS	Size	AR	FA	FA-SP	VFA	B1	B2	B3	B4
C	25	1	16.16	16.16	16.16	16.16	16.16	16.16	16.16
C	25	0.5	10.3	10.3	10.3	10.3	10.3	10.3	10.3
C	50	1	32.16	32.16	32.16	32.16	32.16	32.16	32.16
C	50	0.5	20.42	20.42	20.42	20.42	20.42	20.42	20.42
C	100	1	63.58	63.58	63.58	63.58	63.58	63.58	63.58
C	100	0.5	40.48	40.48	40.48	40.48	40.48	40.48	40.48
R	25	1	16.16	16.16	16.16	16.16	16.16	16.16	16.16
R	25	0.5	10.3	10.3	10.3	10.3	10.3	10.3	10.3
R	50	1	32.16	32.16	32.16	32.16	32.16	32.16	32.16
R	50	0.5	20.42	20.42	20.42	20.42	20.42	20.42	20.42
R	100	1	63.58	63.58	63.58	63.58	63.58	63.58	63.58
R	100	0.5	40.48	40.48	40.48	40.48	40.48	40.48	40.48

of the avoided costs, paving the way for the development of effective approximation algorithms.

In addition, we introduce two algorithms that exploit the gained structural insights. Their effectiveness is validated through a comprehensive simulation study, where they collectively demonstrate superior performance compared to existing benchmarks. Further results from the simulation show that the predictive algorithms have a higher cost saving effect than conventional compensation setting methods when the spatial distribution of delivery locations is dispersed.

Future research should focus on examining the IC of an OD population to better assess algorithm performance in different settings. Additionally, exploring the impact of variable cost structures for DDs, particularly those based on mileage, could be of significant benefit for companies that manage their own fleet.

Appendix A Other material

Table A3 Average fixed compensation

CS	Size	AR	FA	FA-SP	VFA	B1	B2	B3	B4
C	25	1	0	0	0	6.28	0	0	4.52
C	25	0.5	0	0	0	6.36	0	0	5.08
C	50	1	0	0	0	6.1	0	0	4.84
C	50	0.5	0	0	0	6.08	0	0	5.08
C	100	1	0	0	0	6.04	0	0	4.74
C	100	0.5	0	0	0	6.34	0	0	5.08
R	25	1	0	0	0	6.08	0	0	4.52
R	25	0.5	0	0	0	6.66	0	0	5.14
R	50	1	0	0	0	5.7	0	0	4.76
R	50	0.5	0	0	0	6.44	0	0	4.84
R	100	1	0	0	0	6.38	0	0	4.58
R	100	0.5	0	0	0	6.2	0	0	5

Table A4 Average distance scalar

CS	Size	AR	FA	FA-SP	VFA	B1	B2	B3	B4
C	25	1	0	0	0	0	0.24	0	0
C	25	0.5	0	0	0	0	0.2	0	0
C	50	1	0	0	0	0	0.24	0	0
C	50	0.5	0	0	0	0	0.24	0	0
C	100	1	0	0	0	0	0.22	0	0
C	100	0.5	0	0	0	0	0.18	0	0
R	25	1	0	0	0	0	0.22	0	0
R	25	0.5	0	0	0	0	0.46	0	0
R	50	1	0	0	0	0	0.44	0	0
R	50	0.5	0	0	0	0	0.26	0	0
R	100	1	0	0	0	0	0.24	0	0
R	100	0.5	0	0	0	0	0.26	0	0

Table A5 Average detour scalar

CS	Size	AR	FA	FA-SP	VFA	B1	B2	B3	B4
C	25	1	0	0	0	0	0	15.7	0.3
C	25	0.5	0	0	0	0	0	19	0.24
C	50	1	0	0	0	0	0	39.5	0.26
C	50	0.5	0	0	0	0	0	40.3	0.24
C	100	1	0	0	0	0	0	1	0.26
C	100	0.5	0	0	0	0	0	1	0.26
R	25	1	0	0	0	0	0	27.2	0.28
R	25	0.5	0	0	0	0	0	70.6	0.22
R	50	1	0	0	0	0	0	76.8	0.24
R	50	0.5	0	0	0	0	0	20.3	0.28
R	100	1	0	0	0	0	0	1	0.26
R	100	0.5	0	0	0	0	0	1	0.24

References

Allahviranloo M, Baghestani A (2019) A dynamic crowdshipping model and daily travel behavior. Transportation Research Part E: Logistics and Transportation Review 128:175–190. https://doi.org/10.1016/j.tre.2019.06.002

Archetti C, Savelsbergh M, Speranza MG (2016) The vehicle routing problem with occasional drivers. European Journal of Operational Research 254(2):472–480. https://doi.org/10.1016/j.ejor.2016.03.049

Archetti C, Guerriero F, Macrina G (2021) The online vehicle routing problem with occasional drivers. Computers Operations Research 127:105144. https://doi.org/10.1016/j.cor.2020.105144

Arslan AM, Agatz N, Kroon L, et al (2019) Crowdsourced delivery—a dynamic pickup and delivery problem with ad hoc drivers. Transportation Science 53(1):222–235. https://doi.org/10.1287/trsc.2017.0803

Ausseil R, Pazour JA, Ulmer MW (2022) Supplier menus for dynamic matching in peer-to-peer transportation platforms. Transportation Science 56(5):1304–1326. https://doi.org/10.1287/trsc.2022.1133

Banciu M, Mirchandani P (2013) Technical note—new results concerning probability distributions with increasing generalized failure rates. Operations Research 61(4):925–931. https://doi.org/10.1287/opre.2013.1198

Barbosa M, Pedroso JP, Viana A (2023) A data-driven compensation scheme for last-mile delivery with crowdsourcing. Computers & Operations Research 150:106059. https://doi.org/10.1016/j.cor.2022.106059

Behrend M, Meisel F, Fagerholt K, et al (2019) An exact solution method for the capacitated item-sharing and crowdshipping problem. European Journal of Operational Research 279(2):589–604. https://doi.org/10.1016/j.ejor.2019.05.026

Bertsimas D, Popescu I (2003) Revenue management in a dynamic network environment. Transportation Science 37(3):257–277. https://doi.org/10.1287/trsc.37.3.257.16047

Boysen N, Fedtke S, Schwerdfeger S (2021) Last-mile delivery concepts: a survey from an operational research perspective. OR Spectrum: Quantitative Approaches in Management 43(1):1–58. https://doi.org/10.1007/s00291-020-00607-8

Boysen N, Emde S, Schwerdfeger S (2022) Crowdshipping by employees of distribution centers: Optimization approaches for matching supply and demand. European Journal of Operational Research 296(2):539–556. https://doi.org/10.1016/j.ejor.2021.04.002

Buldeo Rai H, Verlinde S, Macharis C (2018) Shipping outside the box. environmental impact and stakeholder analysis of a crowd logistics platform in belgium. Journal of Cleaner Production 202:806–816. https://doi.org/10.1016/j.jclepro.2018.08.210

Cachon GP, Daniels KM, Lobel R (2017) The role of surge pricing on a service platform with self-scheduling capacity. Manufacturing & Service Operations Management 19(3):368–384. https://doi.org/10.1287/msom.2017.0618

Cao J, Olvera-Cravioto M, Shen ZJ (2020) Last-mile shared delivery: A discrete sequential packing approach. Mathematics of Operations Research 45(4):1466–1497. https://doi.org/10.1287/moor.2019.1039

Chen W, Mes M, Schutten M (2018) Multi-hop driver-parcel matching problem with time windows. Flexible Services and Manufacturing Journal 30(3):517–553. https://doi.org/10.1007/s10696-016-9273-3

Dahle L, Andersson H, Christiansen M, et al (2019) The pickup and delivery problem with time windows and occasional drivers. Computers Operations Research 109:122–133. https://doi.org/10.1016/j.cor.2019.04.023

Dai H, Liu P (2020) Workforce planning for o2o delivery systems with crowdsourced drivers. Annals of Operations Research 291:219–245. https://doi.org/10.1007/s10479-019-03135-z

Dayarian I, Savelsbergh M (2020) Crowdshipping and same-day delivery: Employing in-store customers to deliver online orders. Production and Operations Management 29(9):2153–2174. https://doi.org/10.1111/poms.13219

Feng X, Chu F, Chu C, et al (2021) Crowdsource-enabled integrated production and transportation scheduling for smart city logistics. International Journal of Production Research 59(7):2157–2176. https://doi.org/10.1080/00207543.2020.1808258

Gdowska K, Viana A, Pedroso JP (2018) Stochastic last-mile delivery with crowdshipping. Transportation Research Procedia 30:90–100. https://doi.org/10.1016/j.trpro.2018.09.011

Horner H, Pazour J, Mitchell JE (2021) Optimizing driver menus under stochastic selection behavior for ridesharing and crowdsourced delivery. Transportation Research Part E: Logistics and Transportation Review 153:102419. https://doi.org/10.1016/j.tre.2021.102419

Hutter L, Neumann S (2023) Crowdshopping: An empirical analysis of a novel concept for grocery shopping. Working Paper https://doi.org/http://dx.doi.org/10.2139/ssrn.4459488

Kafle N, Zou B, Lin J (2017) Design and modeling of a crowdsource-enabled system for urban parcel relay and delivery. Transportation Research Part B: Methodological

99:62–82. https://doi.org/10.1016/j.trb.2016.12.022

Koch S, Klein R (2020) Route-based approximate dynamic programming for dynamic pricing in attended home delivery. European Journal of Operational Research 287(2):633–652. https://doi.org/10.1016/j.ejor.2020.04.002

Kung LC, Zhong GY (2017) The optimal pricing strategy for two-sided platform delivery in the sharing economy. Transportation Research Part E: Logistics and Transportation Review 101:1–12. https://doi.org/10.1016/j.tre.2017.02.003

Lariviere MA (2006) A note on probability distributions with increasing generalized failure rates. Operations Research 54(3):602–604. https://doi.org/10.1287/opre.1060.0282

Le TV, Ukkusuri SV, Xue J, et al (2021) Designing pricing and compensation schemes by integrating matching and routing models for crowd-shipping systems. Transportation Research Part E: Logistics and Transportation Review 149:102209. https://doi.org/10.1016/j.tre.2020.102209

Macrina G, Di Puglia Pugliese L, Guerriero F, et al (2020) Crowd-shipping with time windows and transshipment nodes. Computers & Operations Research 113:104806. https://doi.org/10.1016/j.cor.2019.104806

Maglaras C, Meissner J (2006) Dynamic pricing strategies for multiproduct revenue management problems. Manufacturing & Service Operations Management 8(2):136–148. https://doi.org/10.1287/msom.1060.0105

Mancini S, Gansterer M (2022) Bundle generation for last-mile delivery with occasional drivers. Omega 108:102582. https://doi.org/10.1016/j.omega.2021.102582

Mofidi SS, Pazour JA (2019) When is it beneficial to provide freelance suppliers with choice? a hierarchical approach for peer-to-peer logistics platforms. Transportation Research Part B: Methodological 126:1–23. https://doi.org/10.1016/j.trb.2019.05.008

Mousavi K, Bodur M, Roorda MJ (2022) Stochastic last-mile delivery with crowd-shipping and mobile depots. Transportation Science 56(3):612–630. https://doi.org/10.1287/trsc.2021.1088

Powell WB (2011) Approximate dynamic programming: Solving the curses of dimensionality (2nd). Wiley, Hoboken, New Jersey, https://doi.org/10.1002/9781118029176

Qi W, Li L, Liu S, et al (2018) Shared mobility for last-mile delivery: Design, operational prescriptions, and environmental impact. Manufacturing & Service Operations Management 20(4):737–751. https://doi.org/10.1287/msom.2017.0683

Savelsbergh MW, Ulmer MW (2022) Challenges and opportunities in crowd-sourced delivery planning and operations. 4OR 20(1):1–21. https://doi.org/10.1007/s10288-021-00500-2

Silva M, Pedroso JP (2022) Deep reinforcement learning for crowdshipping last-mile delivery with endogenous uncertainty. Mathematics 10(20):3902. https://doi.org/10.3390/math10203902

Silva M, Pedroso JP, Viana A (2023a) Deep reinforcement learning for stochastic last-mile delivery with crowdshipping. EURO Journal on Transportation and Logistics 12:100105. https://doi.org/10.1016/j.ejtl.2023.100105

Silva M, Pedroso JP, Viana A (2023b) Stochastic crowd shipping last-mile delivery with correlated marginals and probabilistic constraints. European Journal of Operational Research 307(1):249–265. https://doi.org/10.1016/j.ejor.2022.10.039

Strauss AK, Klein R, Steinhardt C (2018) A review of choice-based revenue management: Theory and methods. European Journal of Operational Research 271(2):375–387. https://doi.org/10.1016/j.ejor.2018.01.011

Taylor TA (2018) On-demand service platforms. Manufacturing & Service Operations Management 20(4):704–720. https://doi.org/10.1287/msom.2017.0678

Torres F, Gendreau M, Rei W (2022a) Crowdshipping: An open vrp variant with stochastic destinations. Transportation Research Part C: Emerging Technologies 140:103677. https://doi.org/10.1016/j.trc.2022.103677

Torres F, Gendreau M, Rei W (2022b) Vehicle routing with stochastic supply of crowd vehicles and time windows. Transportation Science 56(3):631–653. https://doi.org/10.1287/trsc.2021.1101

Yildiz B, Savelsbergh M (2019) Service and capacity planning in crowd-sourced delivery. Transportation Research Part C: Emerging Technologies 100:177–199. https://doi.org/10.1016/j.trc.2019.01.021

Ziya S, Ayhan H, Foley RD (2004) Relationships among three assumptions in revenue management. Operations Research 52(5):804–809. https://doi.org/10.1287/opre.1040.0134