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 delivery locations. 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 a 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. Therefore, 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. Tailored to meet these challenges, our approximation methods aim to accurately determine avoided costs, which are a key factor in calculating optimal compensation. Our simulation study reveals that the savings generated by involving ODs in deliveries can be significantly increased through our individualized dynamic compensation policy. This approach not only excels in generating savings for the firm but also provides a utility surplus for ODs. Additionally, we demonstrate the applicability of our approach to scenarios with time windows and illustrate the trade-off that arises from time window partitioning.

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 on that day, poses a significant challenge. This uncertainty in OD availability contrasts sharply with the reliability of DDs, which offer a more predictable delivery solution. 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 practical feasibility, there remains a considerable gap in bringing 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 sequentially and stochastically arriving ODs with delivery tasks to minimize the expected costs (or equivalently to maximize expected savings through additionally employing ODs along with DDs). Each OD is offered a customized bundle, consisting of a delivery task and compensation. This customized offer is influenced by the 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 dispatch time are fulfilled by DDs. Since our focus is on investigating the cost-saving potential of engaging ODs for deliveries, we assume that DD capacity is sufficient to serve all delivery locations. However, our framework can also be applied in scenarios where DD capacity is insufficient, by incorporating an appropriate penalty term when evaluating DD delivery costs (refer to Section 3.1). This approach, akin to a reversed form of dynamic pricing, is detailed in Figure 1 and offers a more realistic and effective solution for employing independent individuals in last-mile delivery scenarios.

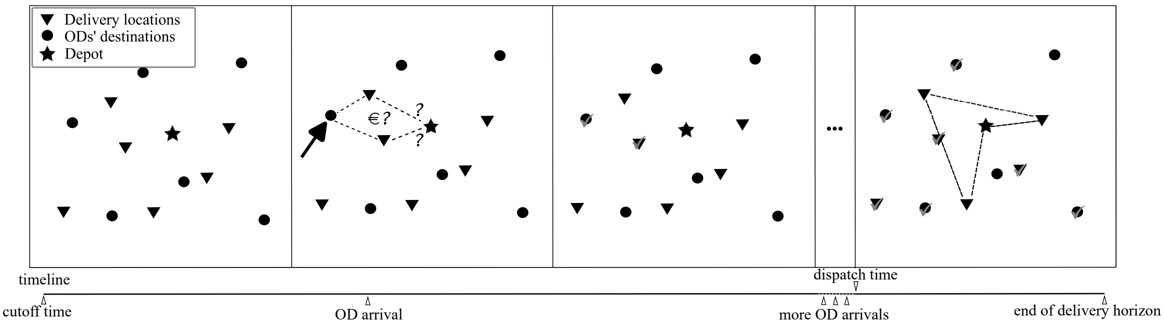

Fig. 1: Dynamic programming process. This illustration depicts the sequential decision-making process as ODs arrive over time. It focuses on the current state, emphasizing three critical elements: the current OD, identified by their destination (represented by an arrow), potential future ODs, marked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

1.1 Problem description

In the following, we describe our scenario in greater detail, beginning with the perspectives of online customers, ODs, and the firm. Furthermore, we point out that our problem characteristics are in line with current crowdshipping literature (refer also to Table 1 in Section 2.4).

- **The online customer's perspective:** Customers can place delivery orders at any time, but there is a communicated cutoff time that must be met for orders to be served in the upcoming delivery horizon. Orders placed after the cutoff time will be served at the next or a later delivery horizon. For example, an order placed before the store opens will be fulfilled on the same business day, whereas an order placed after the store opens will be fulfilled on the next business day.

- **The OD's perspective:** ODs are customers who regularly shop at the store and have, through a one-time registration, expressed their willingness to occasionally make deliveries. During registration, they also indicated their usual destinations, such as their home or workplace. Upon entering the store, each OD receives a customized offer, consisting of a specific delivery location along with the corresponding compensation. They can accept or decline this offer. On acceptance, ODs collect the order and compensation at the checkout counter and deliver the package en route to their predetermined destination.
- **The firm's perspective:** At the cutoff time, the firm knows all the orders that need to be handled in the upcoming delivery horizon (e.g., the business day). There is a latest dispatch time by which orders must leave the store to meet the promised delivery times. If the firm does not have its own fleet of DDs, the dispatch time is typically agreed upon with a third-party service provider (e.g., DHL, UPS). The firm is notified upon the arrival of registered ODs and informed of their usual destinations. This information allows the firm to make specific offers of compensation and delivery locations. Orders not handled by ODs are handed over to DDs at dispatch time. DDs are capable and willing to fulfill any assigned orders in a timely manner.

A cutoff time is widely used in practice (e.g., Amazon with next-day delivery) as well as in literature (see, e.g., Archetti et al, 2016; Boysen et al, 2022; Mousavi et al, 2022). The dispatch time can either be chosen by the firm as the time it must deploy its vehicle fleet, or it can be an agreed-upon time when a contracted third-party service provider picks up the orders from the store. Since our focus is on the optimal engagement of ODs, we will concentrate on the latter case, as it simplifies the evaluation of the final states in our dynamic programming formulation (see, Section 3.2). However, note that all our results (particularly, our structural properties and solution methods) are still applicable in a scenario where the firm uses its own vehicle fleet and solves a vehicle routing problem for evaluating the final states.

The period between cutoff and dispatch time can be used to engage any OD who enters the store during that timeframe. Clearly, crowdshipping can only have a measurable impact if this period is sufficiently long, where the required duration depends on store traffic and the ratio of ODs within this traffic. Exploiting this period for ODs naturally results in a framework where OD fulfillment precedes DD fulfillment. This approach ensures that all orders will be delivered within the promised delivery horizon, as no order will be left behind by DDs in the hope of an adequate OD arriving after the dispatch time. Consequently, this sequence of engaging ODs and having DDs handle the remaining deliveries afterward is widely observed in crowdshipping literature (see, e.g., Gdowska et al, 2018; Torres et al, 2022a; Silva et al, 2023b).

As is common in a crowdshipping context (see, e.g., Archetti et al, 2016; Ausseil et al, 2022; Barbosa et al, 2023), we focus on orders manageable by a single person. This ensures that the orders are of reasonable size and weight, eliminating the need to track these attributes. Thus, the single characterizing attribute of an order is given by its delivery location. The firm makes a single offer consisting of one delivery location and compensation to each arriving OD, which minimizes communication between the OD and the firm and reduces the need to consider ODs' capacity capabilities. This approach recognizes that ODs are individuals who do not primarily seek to earn a living through deliveries but are willing to spend a small part of their free time participating in crowdshipping. Therefore, it is crucial to reduce any inconvenience for them. Given this design choice, we do not consider the vehicle capacities of ODs to be a limiting factor. However, if ODs are unable to take on delivery for any reason (such as a fully-loaded vehicle, an unusual destination, or time constraints) they can always decline the offer. This flexibility is represented in the OD choice model by a random variable, accounting for the volatility in ODs' willingness to participate.

While our framework is most naturally suited for (same-day) unattended home delivery, which has got a lot of attention from recent crowdshipping research (see, e.g., Boysen et al, 2022; Mousavi et al, 2022; Silva et al, 2023b), it can also be adapted for scenarios with attended home delivery by shortening the delivery horizon to just a few hours instead of an entire business day. Note that the effectiveness of engaging ODs diminishes as the delivery horizon decreases. This is because ODs begin deliveries shortly after they arrive at the store, leaving a limited timeframe for the firm to identify and engage suitable ODs for a time-restricted delivery. Furthermore, smaller time windows usually necessitate that DDs cover multiple delivery horizons within a single tour. This often means that orders must commence their journey before the firm has the opportunity to offer them to an OD. Consequently, ODs are less beneficial in scenarios with very small delivery windows. Nevertheless, medium-sized delivery windows, such as a few hours or half of the business day, align well with

our approach when applied consecutively (see Section 6). This consecutive approach to dealing with time windows is comparable to the approaches covered by Arslan et al (2019) or Silva et al (2023a). In our setting, ODs are customers who shop in the store and have previously expressed a general interest in occasionally taking on delivery, along with providing their usual destinations after shopping. Consequently, the firm has a pool of ODs and their respective destinations, a scenario that is often observed in the literature (see, e.g., Feng et al, 2021; Le et al, 2021; Mancini and Gansterer, 2022).
Through individualized offers, consisting of a specific delivery location and compensation, the firm can influence ODs' willingness to participate in crowdshipping. By selecting the delivery location to offer, the firm determines the required detour for the OD. Simultaneously, it offers individual compensation to offset the inconveniences, incurred by fulfilling the delivery task. This assumption is in line with the literature, where higher compensations typically increase the willingness of ODs to accept these offers (see, e.g., Cachon et al, 2017; Taylor, 2018; Yildiz and Savelsbergh, 2019).

1.2 Contribution and outline of the paper

Our goal is to determine the optimal strategy for engaging ODs in making deliveries while modeling a realistic interaction between the firm and ODs. This approach contrasts sharply with current research, which often relies on simplifying assumptions such as deterministic or exogenous OD behavior. Additionally, many potential opportunities are overlooked in existing studies. Without incorporating a dynamic decision-making framework, firms lose the ability to swiftly respond to new information revealed through the arrivals and decisions of ODs. Moreover, by ignoring the impact of the firm's decisions on OD decision-making, firms miss a crucial lever for controlling the entire decision process. To the best of our knowledge, we are the first to develop a dynamic optimization framework where ODs arrive dynamically and stochastically over time, prompting the firm to decide on the optimal offer, consisting of a specific delivery location and individual compensation. This compensation depends on the specific OD, the current state of the system, and the offered delivery location. Furthermore, OD behavior is adequately modeled, being influenced by the delivery location offered and the compensation provided.
This paper makes several significant contributions to the field of logistics, utilizing dynamic pricing methodologies:

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 finding an optimal policy, laying the foundation to develop appropriate approximation methods.
3. **Tailored approximation methods:** In Section 5, we introduce specialized adaptations of well-known approximation methods, specifically designed to align with the inherent structure of our problem. 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:** In a detailed simulation study presented in Section 6, we apply our algorithms to demonstrate the advantages of acknowledging the uncertainties inherent in employing ODs. This study not only validates our approach but also provides valuable insights into the benefits of dynamic compensation strategies across various urban configurations and OD arrival patterns. Additionally, we analyze the effect of time windows in our crowdshipping setting.

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 (which acts as the depot) and, if conditions align with their schedule, may undertake a delivery task. An

important element in our analysis is the uncertainty associated with the arrival of ODs at the store and their decision to accept a task. We posit that compensations significantly influence ODs' acceptance decisions. 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 does 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 uncertain behaviors of ODs, regarding their arrival at the depot 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.
In Section 2.3, we briefly discuss related studies that share similarities with our work in terms of model formulation and solution methods. Finally, in Section 2.4, we conclude our review with a summary and provide an overview of how our problem characteristics align with relevant literature, as well as the unique capabilities of our model, in Table 1.

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 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 (see, e.g., Dahle et al, 2019; Boysen et al, 2022). The most distinct difference between their setting and our work lies in the absence of uncertainty. In their setting, the firm knows exactly which ODs are or will be available and can perfectly predict and rely on their decisions. In contrast, our work models an OD as an independent individual who may or may not arrive and may or may not be willing to accept an offered task.
Diverging from the fixed compensation schemes which are applied by Archetti et al (2016), 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 aligned with ODs' acceptance thresholds. These studies introduce comprehensive OD-specific compensation schemes, aligning with our approach in this aspect. However, they rely on ODs communicating their desired compensation and assume that the firm has access to numerous ODs at the moment of decision. This implies that ODs are either waiting to be employed or that the firm knows with certainty which ODs will be available throughout the day. These assumptions differ significantly from our setting, where the firm is only informed of an OD's arrival, necessitating an immediate decision on the specific task and corresponding compensation to offer without certainty regarding the OD's response.
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). Examining a dynamic setting is a core similarity to our work. However, our approach explicitly anticipates future uncertain arrivals of ODs, and these anticipations influence current decisions. Moreover, we do not assume the ability to predict ODs' decisions, a common assumption in deterministic settings where ODs are expected to act as planned by the firm.
In conclusion, the deterministic literature shares many similarities in problem characteristics with our work. However, they all lack the possibility of anticipating uncertain OD behavior, which is one of the core capabilities of our approach.

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 to consider ODs' acceptance uncertainty were Gdowska et al (2018) who introduced a two-stage optimization model. In this model, tasks are initially assigned to ODs and DDs, including solutions to vehicle routing problems, while anticipating potential rejections by ODs. The likeliness of rejection is thereby assumed to be influenced by compensation, albeit compensations are not part of the firm's decision problem. Tasks rejected by ODs are reassigned to DDs in the second stage. Naturally, studies like Gdowska et al (2018), which account for the inherent uncertainties when dealing with independent individuals, are more closely related to our setting than deterministic studies mentioned in the previous section. However, there are still vast differences. For instance, we allow the firm to influence ODs' decision-making by choosing an appropriate OD-, state-, and task-specific compensation. Moreover, through our dynamic setting, the firm can react to every new piece of information it acquires through the arrival of ODs and their decisions.

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. Hou et al (2023) determine a detour-minimizing bipartite matching between OD destinations and delivery locations in the first stage, before setting OD-specific compensations in the second stage. 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, respectively. Santini et al (2022) introduce a TSP variant with uncertain node availability in a setting with ODs. None of these two-stage models can swiftly react to new information revealed over time, a capability our dynamic programming formulation provides. However, Silva et al (2023a) makes some progress in this direction by repeatedly solving a two-stage problem to account for information updates and meet time windows. Other similarities to our research can be found in Silva and Pedroso (2022), Hou et al (2023), and Barbosa et al (2023), where compensations are part of the firm's decision-making process, albeit at a static level rather than in an OD-, state-, and task-specific manner as we do.

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. These studies acknowledge the inherent uncertainty regarding OD decisions and provide a basis for OD-specific offers, aligning with our approach. However, we use compensations as a primary mechanism to influence OD decisions, simplifying the interaction to a straightforward offer and a binary response from the OD. This method reduces the complexity of communication between the firm and the OD, focusing on immediate acceptance or rejection of the task based on the compensation offered. Moreover, it allows for a higher degree of offer customization, where every OD receives an individual offer consisting of a specific delivery task and corresponding compensation.

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. Similar to our approach, these models account for the dynamic and stochastic arrival of ODs, with each new arrival providing information that can be utilized in the firm's decision-making. However, a key difference in our approach is our focus on determining optimal compensations to influence OD decisions. On the contrary, except for Ausseil et al (2022) via offer sets, none of these studies explore the potential for firms to influence the uncertain decisions of ODs through their offers.

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 investigating 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 savings 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 contrast to the predominantly aggregated decision-making in the aforementioned studies, our model leverages the interplay between individual ODs' destinations and delivery locations. Since these studies vastly differ in their perspective on ODs and in methodology from our approach, we have excluded them from Table 1.
In conclusion, the stochastic literature showcases an evolving landscape of methodologies and models that effectively engage the inherent uncertainties in OD participation, similar to our approach. 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. However, none of these studies combine all the capabilities offered by our framework, where uncertain OD behavior is anticipated while dynamically optimizing OD-, state-, and task-specific compensations.

2.3 Methodological similarities

From a modeling standpoint, our dynamic programming formulation can be compared to a multiproduct dynamic pricing setting. This becomes evident when we compare delivery locations to products, compensations to prices, and ODs to customers. ODs' demand is stochastic and the firm is facing a finite selling horizon with scarce product-dependent inventory (see, e.g., Zhang and Cooper, 2009; Dong et al, 2009; Akçay et al, 2010). However, a significant difference between existing literature and our setting arises from the pool of specific ODs: each OD has a known and unique characterization (their destinations) and they vary greatly in their suitability for specific delivery locations. The uniqueness of each OD makes it impossible to use vertical differentiation in which a global ranking of preferences is assumed. This interplay between ODs and delivery locations creates a distinct optimization problem that has not been addressed in multiproduct dynamic pricing studies so far. Additionally, as the same OD does not arrive multiple times at the store, the program must keep track of which ODs have arrived and which might still arrive. This adds another dimension to the state space, increasing the difficulties from the infamous curse of dimensionality and necessitating the development of an efficient solution approach.
Since we use a stochastic dynamic programming model, the decision (as will be discussed in detail in Section 3) in each stage depends on the value of the consecutive stages. For business-relevant instance sizes, finding the true value of being in a certain state becomes intractable because of the curses of dimensionality. Powell (2011) deals with the problem of finding suitable approximations for the value of consecutive stages. One approach to approximate the value function is the parametric value function approximation (VFA), widely used in related literature (e.g., Koch and Klein, 2020; Silva and Pedroso, 2022). Another often observed approach is to use a fluid approximation (FA) (e.g., Gallego and Van Ryzin, 1997; Maglaras and Meissner, 2006). In this work, we follow both approaches, tailoring them to meet the specific problem structure, including the complex interplay between OD destinations and delivery locations.

2.4 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; Silva and Pedroso, 2022). 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 uncertain 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. Table 1 presents a summary of the related literature, focusing on the problem characteristics and the capabilities of the model. It is evident from this summary that while our characteristics align well with the current literature, our model has a unique combination of capabilities.

Authors (Year)	Problem characteristics					Model capabilities			
	Time windows	Dispatch time for DD delivery	Known OD destinations	Known set of customer orders	Single delivery per OD	Anticipation of uncertain future arrival	Anticipation of uncertain OD acceptance	Compensation decision to influence uncertain OD behaviour	Dynamic decision making
Archetti et al (2016)		-	X	X	X				
Kafle et al (2017)	X		X	X					
Gdowska et al (2018)		X		X	X		X	X	
Allahviranloo and Baghestani (2019)	X	-	X	X					X
Arslan et al (2019)	X		X	X					X
Dahle et al (2019)	X	-	X	X	X				
Mofidi and Pazour (2019)				X			X		
Dayarian and Savelsbergh (2020)	X		X			X			X
Archetti et al (2021)	X	X	X						X
Feng et al (2021)		-	X	X					
Horner et al (2021)		-		X			X		
Le et al (2021)	X	-	X	X			X		
Ausseil et al (2022)	X	-			X	X	X		X
Boysen et al (2022)		-	X	X					
Mancini and Gansterer (2022)		X		X					
Mousavi et al (2022)			X	X		X			
Santani et al (2022)		X		X			X		
Silva and Pedroso (2022)	X	X			X		X	X	
Torres et al (2022a)		X	X	X		X			
Torres et al (2022b)	X	X		X		X			
Barbosa et al (2023)		X		X	X		X	X	
Hou et al (2023)		X	X	X	X		X	X	
Silva et al (2023a)	X	X			X	X			X
Silva et al (2023b)	X	X			X	X			
This work	X	X	X	X	X	X	X	X	X

Table 1: Summary of problem characteristics and model capabilities. The category "dispatch time for DD deliveries" is marked (X) for all studies that have a dedicated point in time in which all remaining tasks are outsourced to a fleet of DDs. Studies that do not incorporate DDs into their model (pure crowdshipping models) or do not specify the delivery time are excluded (-) from this category.

3 The dynamic compensation problem for occasional drivers

In this section, we present the dynamic optimization model along with the corresponding notation. Section 3.1 introduces the general notation, laying the foundation for a detailed explanation of the essential components of our dynamic programming formulation, in accordance with Powell (2011), which is discussed in Section 3.2. Finally, Section 3.3 presents the Bellman equation, representing the optimization problem after decomposing the entire dynamic programming formulation into state-dependent optimization problems.

3.1 General setting and notation

At the cutoff time, all orders that have to be fulfilled in the upcoming delivery horizon are predetermined and known. We denote the set of delivery locations as $\mathcal{C} = \{1, 2, \ldots, C\}$. Please note that, for any $c \in \mathcal{C}$, we use the terms customer, (delivery) location, (delivery) task, and order synonymously, as each order corresponds to a customer at a specific location. In this context, C represents the number of orders. Additionally, the set of potential OD destinations is denoted as $\mathcal{O} = \{C+1, C+2, \ldots, C+O\}$. Similarly, for any $o \in \mathcal{O}$, the terms OD and destination are used interchangeably, as ODs have previously registered, providing their usual destinations. This registration makes ODs a heterogeneous group, explicitly distinguishable from each other by their destination. Since the exact number of arriving ODs is unknown beforehand, O refers to the number of registered ODs that might arrive during this delivery horizon. The depot location is indexed as 0. Together, $\mathcal{C}, \mathcal{O}$, and $\{0\}$ define the nodes of a complete graph $G = (V, E)$ with $E = \mathcal{C} \cup \mathcal{O} \cup \{0\}$. The distances between the depot, locations, and destinations are denoted by d_{uv} $\forall\ u, v \in \mathcal{C} \cup \mathcal{O} \cup \{0\}$. This forms the basis for calculating the detour of an OD o to a delivery location c: $u_{oc} = d_{co} + d_{0c} - d_{0o} \forall o \in \mathcal{O}, c \in \mathcal{C}$.
The time horizon between cutoff and dispatch time is divided into discrete periods $t \in \mathcal{T} = \{1, 2, ..., T+1\}$, with $t = 1$ and $t = T+1$ marking the first period after cutoff and dispatch time, respectively. Each time period $t \in \mathcal{T} \setminus \{T+1\}$ is chosen to be sufficiently small, rendering the likelihood of two OD arrivals within the same period negligible. This practice aligns with established conventions in standard revenue management literature (Strauss et al, 2018).

For each period t, we define the open delivery tasks and remaining potential ODs as $\mathcal{C}_t \subseteq \mathcal{C}$ and $\mathcal{O}_t \subseteq \mathcal{O}$, respectively. The exact composition of these sets depends on the stochastic process unfolding between the cutoff time and period t, driven by the stochastic arrivals of ODs over time as well as the firm's and ODs' decisions. In each period $t \in \mathcal{T} \setminus \{T+1\}$, at most one of the remaining ODs of $\mathcal{O}_t$ arrives. Thereby, the probability of a specific OD $o \in \mathcal{O}_t$ arriving is denoted by λ_{ot}. Consequently, the probability of no arrival is given by $\overline{\lambda}_t = 1 - \sum_{o \in \mathcal{O}_t} \lambda_{ot}$.
Upon their arrival, ODs are recognized by the firm, making ODs' destinations exploitable information. The firm decides on a delivery location and a corresponding individual compensation 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 order to the delivery location. When declining the offer, the OD departs without taking on a delivery task. In either case, this OD is removed from the set of potential ODs. Consequently, in period $t \in \mathcal{T} \setminus \{T+1\}$, the firm incurs costs $r_t(\mathcal{O}_t, \mathcal{C}_t, o, c)$ if it offers location $c \in \mathcal{C}_t$ to OD $o \in \mathcal{O}_t$ for compensation $r_t(\mathcal{O}_t, \mathcal{C}_t, o, c)$ and if this OD accepts this delivery task. Note that the compensation $r_t(\mathcal{O}_t, \mathcal{C}_t, o, c)$ is part of the firm's decision and depends on the current period, the remaining potential ODs, the open delivery tasks, the specific OD, and the offered location. In the remainder of the paper, we will simplify notation by writing r without explicitly expressing all these dependencies.
In the dispatch period $T+1$, any open order $c \in \mathcal{C}_{T+1}$ is handled by DDs to ensure timely delivery within the promised delivery horizon. The cost of delivery by DDs is denoted by $\Theta(\mathcal{C}_{T+1})$. As recommended by Boysen et al (2022), among others, a common approach to modeling delivery costs by DDs is to assume a fixed fee, denoted as κ_c, for delivering to a 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 firms 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 optimally engaging ODs. 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 delivery horizon. This includes solving a vehicle routing problem to determine delivery costs in the dispatch period for firms employing their fleet of DDs.
The objective of the firm is to minimize expected total costs $V_1(\mathcal{O}, \mathcal{C})$, incurred by ensuring the delivery of any order $c \in \mathcal{C}$ in the delivery horizon $\mathcal{T}$. These costs consist of paid compensations to ODs and costs of delivery by DDs. We will formulate this optimization problem using the Bellman equation in Section 3.3, following the introduction of crucial components of our dynamic programming formulation in Section 3.2.

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 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 open 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 offered bundle (c, r), comprising a delivery task $c \in \mathcal{C}_t$ and a compensation r (refer to 3.2.3).
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 3.2.2.

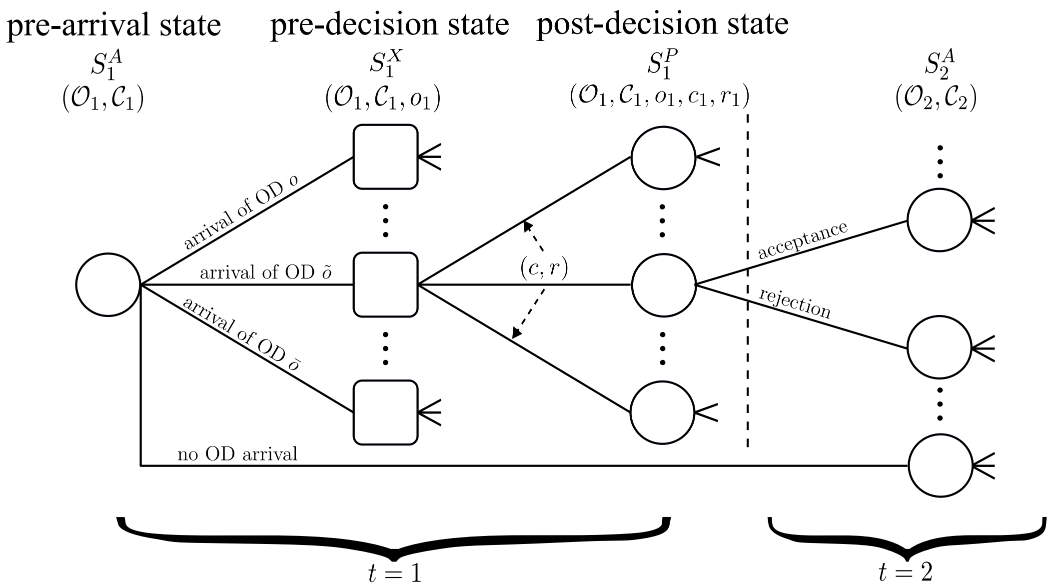

Fig. 2: Tree representation of the dynamic program. Circles represent states in which the transition is determined by a random process, while squares represent states in which the subsequent state is determined by the firm's decision.

3.2.2 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 and 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:

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

3.2.3 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 ODs 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 $\mathcal{X}_R(S_t^X) = \mathbb{R}_+$.

3.2.4 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. 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

$$IC_{oct} = a_{oct} + \omega. \tag{2}$$

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 that the reflected distribution, given by $\overline{f}(x) = f(-x)$ over the support $[-b_{oct}, 0]$, exhibits an increasing failure rate $\overline{h}(x) = \overline{f}(x)/(1 - \overline{F}(x))$.

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, and 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 $\overline{f}$, further extending the possibilities for the selection of f.

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 $t = 1, \ldots, T$ and to DDs at the dispatch time $t = T+1$ 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

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

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

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

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 present 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 3.2.4), 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:*

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

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 3.2.4). 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 from 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 in Section 5. Consequently, we commence with 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. The proof of Proposition 3 can be found in the Appendix (Section C.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 show that there is no concave/convex structure present in our setting. Specifically, we 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. The proof of Proposition 4 can be found in the Appendix (Section C.2). □

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, instead, the impact is contingent on the interplay of various factors at that moment in time.

5 Approximation 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, 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 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 post-arrival state S^X to a corresponding vector of expected avoided cost $x_c \in \mathbb{R}\ \forall c \in \mathcal{C}_t$. As the exact avoided costs are computationally impractical to determine, a parametric VFA approximates these by employing a linear combination of features, also known as basis functions. These basis functions are scaled by a set of weights, which are learned during a dedicated learning phase. Once the learning phase concludes, the approximated avoided costs can be computed by building the linear combination of basis functions. To be effective, the basis functions must encapsulate sufficient information about the current state (remaining periods, delivery locations, and ODs) for reliable predictions regarding a delivery location's avoided costs. The challenge lies in identifying a set of functions that adequately represent the structure of the state. The advantage of employing a parametric VFA 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.

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), following 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 realized 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 $\overline{r}_{ct} = \sum_{q \in Q} r_{qct}$.
- 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 $\overline{h}_{ct} = \sum_{q \in Q} h_{qct}$. In instances where delivery location c is not served in a particular run, $\overline{h}_{cT+1}$ 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.

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

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. For a pseudo-code representation of the algorithm, we refer to Appendix B.1.
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 that 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 in the corresponding weights, highlighting the increased competition in serving this delivery location and the reduced reliance on only a few ODs. We illustrate the effectiveness with a small example in Appendix B.2.

5.2 Fluid approximation

In this section, 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 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 (see, e.g., Maglaras and Meissner, 2006). The FA for approximating $V_t(\mathcal{O}_t, \mathcal{C}_t)$ is given by:

$$\overline{V}_t^{FA}(\mathcal{O}_t, \mathcal{C}_t) = \min_{r_{oc}} \sum_{c \in \mathcal{C}_t} \sum_{o \in \mathcal{O}_t} P_{\mathrm{Ar}}(o) \cdot P_{\mathrm{Ac}}(o, c, r_{oc}) \cdot r_{oc} + \sum_{c \in \mathcal{C}_t} \left(1 - \sum_{o \in \mathcal{O}_t} P_{\mathrm{Ar}}(o) \cdot P_{\mathrm{Ac}}(o, c, r_{oc})\right) \cdot \kappa_c \tag{22}$$

$$\text{s.t.} \quad \sum_{o \in \mathcal{O}_t} P_{\mathrm{Ar}}(o) \cdot P_{\mathrm{Ac}}(o, c, r_{oc}) \leq 1 \qquad \forall c \in \mathcal{C}_t \tag{23}$$

$$\sum_{c \in \mathcal{C}_t} P_{\mathrm{Ac}}(o, c, r_{oc}) \leq 1 \qquad \forall o \in \mathcal{O}_t \tag{24}$$

$$r_{oc} \geq 0 \qquad \forall c \in \mathcal{C}_t, o \in \mathcal{O}_t \tag{25}$$

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 a 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) minimizes the approximate expected costs for the remaining time horizon. Constraint (23) ensures that the demand of a delivery location 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:

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

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 variable assignments 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 firm 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 delivery locations can be assigned avoided costs of $\Delta\overline{V}_t^{SP}(S_t^X, c) = \kappa_c$. 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

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

Here, ζ_c represents the shadow prices 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 Appendix B.3, 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. The pseudo-code for simulating the decision policy to measure its performance is depicted in Algorithm 2 in Appendix B.1. The algorithm uses discrete periods but can easily be adapted to a continuous time setting.
Since our goal is to approximate the costs of being in state S_t^A, 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 $S_t^X = (\mathcal{O}_t, \mathcal{C}_t)$, it holds:*

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

Proof. The proof of Lemma 3 can be found in Appendix C.3. □

6 Simulation study

In this simulation study, we explore the cost-saving potential of engaging ODs to perform delivery tasks on their way home from shopping. Beyond cost savings, we also incorporate metrics that analyze OD utilization, such as the average compensation paid to ODs, the number of delivery orders outsourced to ODs, and the accumulated utility of ODs. To evaluate the capabilities of our model, we introduce three benchmark methods in Section 6.1, inspired by related literature. Additionally, we conduct a sensitivity analysis to examine the impact of various parameters on savings and the utilization of ODs. In Section 6.2, we examine our approach in a scenario involving time windows, demonstrating the trade-off between offering short time windows and effectively utilizing ODs.

6.1 Performance evaluation

In this section we explain our instance generation procedure in Section 6.1.1, introduce the methods that were created to serve as benchmarks for our algorithms in Section 6.1.2, and present the results of the simulation study in Section 6.1.3.

6.1.1 Instance generation

The set of parameters that need to be specified includes the number of delivery locations (C), the number of potential OD destinations (O), arrival rates (λ_{ot}), costs for the delivery by DDs at the dispatch time (κ_c), the number of periods (T) and the distribution parameters of the IC (IC_{oct}).
We define a base set of parameters with $C = 50$, $O = 50$ and $T = 50$. The arrival rate is defined by $\lambda_{ot} = \frac{1}{O}$ if $o \in \mathcal{O}_t$ else $\lambda_{ot} = 0$. Every delivery location not served by an OD before the dispatch time incurs costs of $\kappa_c = \kappa = 10$ at period $T + 1$. The IC is distributed according to a continuous uniform distribution within the bounds $[a_{oct}, a_{oct} + b_{oct}]$, where $a_{oct} = u_{oc}$ and $b_{oct} = \frac{\kappa}{2}$.

Each parameter of the base set is changed individually in our study to analyze their influence on the savings. For the parameters C, O, T, a_{oct} and b_{oct}, we generate two further parameter sets for each: one with a 50%-decrease and another with a 50%-increase in the respective parameter, while keeping all other parameters unchanged. We analyze each set of parameters by first drawing random locations for each customer and random destinations for each potential OD. The coordinates for each delivery location in $\mathcal{C}$ are drawn uniformly from the coordinate set of the Solomon instance R-101 (see, Solomon (1987)), as also observed in other studies (e.g., Archetti et al, 2016; Macrina et al, 2020; Torres et al, 2022a). The coordinates of each destination in $\mathcal{O}$ are represented by whole numbers which are generated randomly using a discrete uniform distribution with bounds that mirror the respective Solomon instance, i.e., $x_o \in \{0, ..., 70\}$ and $y_o \in \{0, ..., 80\}$. To analyze the effect of a city structure with clustered districts, we additionally generate one parameter set with clustered delivery locations and ODs' destinations (sampled from C-101 data). The distance measure used is the Euclidean distance. We generated five graphs for each combination of C and O for both analyzed city structures that differ in the locations of their nodes. We combined these with the other parameters to create five distinct instances for each set of parameters to mitigate outliers. The performance of each method (see, Section 6.1.2) was evaluated on 100 sampled OD arrival streams for each instance, and the average results of 500 simulation runs were analyzed for each set of parameters.

6.1.2 Comparative methods

In line with the main incentive of our study, which is to analyze the advantages of optimizing individual dynamic compensations while anticipating uncertain OD arrivals and acceptance, we adopt a step-wise approach to analyze these model capabilities separately. To achieve this, we developed three methods, that draw inspiration from the literature and build progressively upon each other. The methods become gradually more complex as we introduce the capabilities of our model (refer to Table 2):

1. *Anticipation of uncertain dynamic arrivals of ODs*
2. *Anticipation of uncertain OD acceptance*
3. *Optimizing individual dynamic compensations*

In the first method, decisions are predetermined, disregarding the uncertain and dynamic nature of the arrival process. Delivery locations and ODs are matched by solving an assignment problem, that maximizes savings while ignoring uncertainty. This approach is comparable to the setting in Archetti et al (2016), where OD arrival and acceptance are deterministic. As deterministic models often replace random variables with their expected values, we replace the IC with its expectation in this method. Consequently, the resulting compensation scheme mirrors the expected ICs, given by $u_{oc} + \frac{\kappa}{4}$. We refer to this method as IA (initial assignment).

The second method addresses uncertain dynamic arrivals using a dynamic myopic assignment strategy, similar to approaches observed in related literature (e.g., Arslan et al, 2019; Dayarian and Savelsbergh, 2020; Ausseil et al, 2022). In this method, the firm determines the optimal matching upon the arrival of each OD, focusing on the immediate effectiveness of a match, independent of uncertain future arrivals. This approach ensures that only arriving ODs are matched with delivery locations, and each OD receives the most suitable location. The compensation is chosen according to the expected IC. We refer to this method as DYN, reflecting its dynamic assignment strategy.

The third method introduces the additional capability of anticipating uncertain OD acceptance and the ability to influence ODs' decision-making through an optimized static compensation scheme. This method is inspired by similar approaches in literature (see, e.g., Silva and Pedroso, 2022; Barbosa et al, 2023). Specifically, this method offers an optimized static compensation to all incoming ODs, mimicking the approach of Barbosa et al (2023). In their work, the authors calculate the expected costs of the delivery process by offering a static compensation to ODs and using a full enumeration of all possible scenarios. However, due to the high complexity introduced by our dynamic approach, this brute-force evaluation is infeasible for relevant instance sizes. Instead, we replace full enumeration with a simulation of 100 arrival processes and use the average savings from these runs as the evaluation criterion. Consistent with Barbosa et al (2023), we apply the golden ratio search algorithm to find the best static compensation. Matching decisions remain dynamic and myopic, in accordance with method DYN. We refer to this method as OSCS (optimized static compensation strategy).

Our approach combines all previously mentioned capabilities, optimizing individualized and dynamic compensations while anticipating uncertain dynamic arrivals and OD decision-making. The optimal solution is derived from Propositions 1 and 2, with avoided costs approximated using our methods from Sections 5.1 and 5.2. Both the VFA and FA, when combined with our analytical solution, demonstrated similarly strong performances across all instances, as illustrated in Figure 3 for the base set of parameters. To condense the results of our sensitivity analysis, we used the VFA to represent our approach that incorporates all model capabilities.

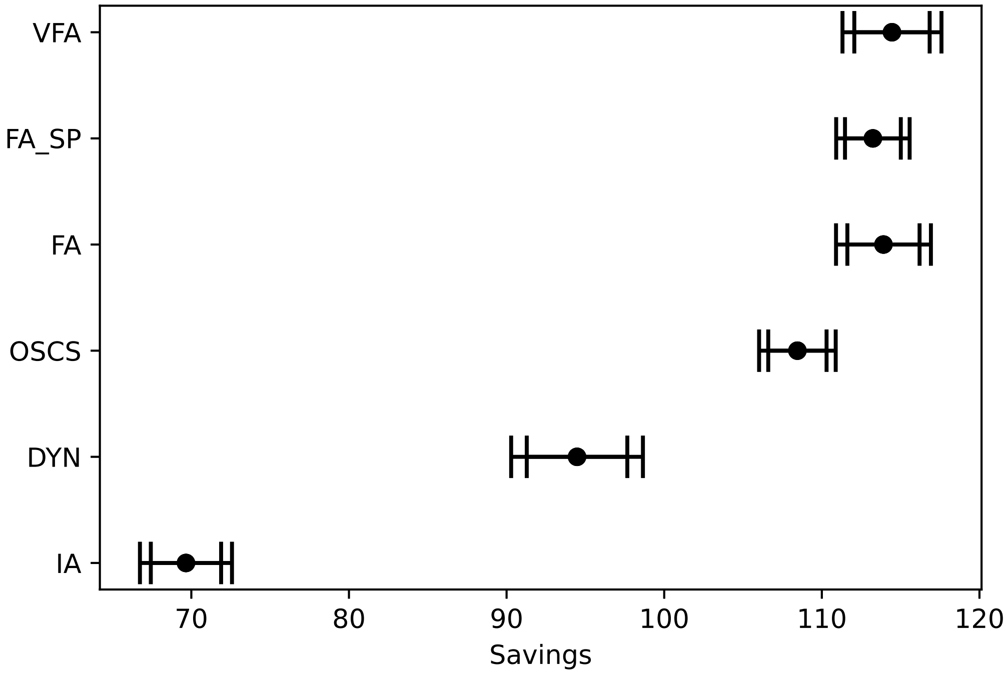

Fig. 3: Comparison of the average savings generated by each method, along with the respective 95% and 99% confidence intervals, for the base set of parameters.

	Incorporating:		
Methods	**Anticipation of uncertain dynamic arrivals of ODs**	**Anticipation of uncertain OD acceptance**	**Optimizing individual dynamic compensations**
IA	X	X	X
DYN	✓	X	X
OSCS	✓	✓	X
VFA	✓	✓	✓

Table 2: Summary of the model capabilities of the methods under comparison.

Research object	Method	Base	$C = 25$	$C = 75$	Clustered	$O = 25$	$O = 75$	$T = 25$	$T = 75$	$a = 0.5u_{oc}$	$a = 1.5u_{oc}$	$b = 0.25\kappa$	$b = 0.75\kappa$
Total Savings	IA	69.66	44.63	95.18	74.09	61.59	69.18	43.11	84.85	77.11	44.35	70.34	41.94
	DYN	94.47	67.15	104.62	104.47	64.90	107.01	62.73	111.66	103.52	88.72	113.21	77.14
	OSCS	108.45	71.36	124.45	121.83	78.17	120.15	75.20	123.80	116.43	102.19	148.00	78.52
	VFA	114.45	79.39	135.63	126.83	83.99	130.69	77.40	136.76	127.57	109.23	171.17	82.74
Relative Savings[1]	IA	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
	DYN	135.62%	150.44%	109.91%	141.00%	105.37%	154.68%	145.50%	131.61%	134.25%	200.07%	160.94%	183.92%
	OSCS	155.69%	159.87%	130.75%	164.44%	126.91%	173.67%	174.43%	145.91%	150.99%	230.44%	210.40%	187.21%
	VFA	164.31%	177.86%	142.50%	171.18%	136.36%	188.91%	179.55%	161.19%	165.43%	246.31%	243.33%	197.28%
Compensations to ODs	IA	101.04	48.21	124.82	95.49	83.97	97.98	62.03	122.61	68.17	51.71	52.18	75.84
	DYN	54.81	54.45	51.14	54.39	38.86	62.15	33.31	65.66	52.64	50.88	37.41	70.50
	OSCS	98.65	51.78	103.67	111.89	66.07	92.19	69.10	86.92	93.17	88.57	53.66	74.22
	VFA	103.11	68.07	129.67	118.81	84.95	112.35	72.70	119.02	113.55	98.13	82.41	77.32
Relative OD costs[2]	IA	23.48%	23.47%	19.06%	22.42%	19.15%	22.74%	13.58%	29.53%	16.12%	11.35%	12.14%	16.56%
	DYN	13.52%	29.78%	7.92%	13.75%	8.93%	15.81%	7.62%	16.91%	13.28%	12.37%	9.67%	16.67%
	OSCS	25.19%	28.99%	16.57%	29.59%	15.66%	24.27%	16.27%	23.10%	24.29%	22.26%	15.24%	17.61%
	VFA	26.74%	39.90%	21.11%	31.84%	20.42%	30.42%	17.20%	32.77%	30.49%	25.11%	25.06%	18.53%
Number of served locations	IA	17.07	9.28	22	16.95	14.55	16.71	10.51	20.74	14.52	9.60	12.25	11.77
	DYN	14.92	12.16	15.57	15.88	10.37	16.916	9.60	17.73	15.61	13.96	15.06	14.76
	OSCS	20.71	12.31	22.81	23.37	14.42	21.23	14.43	21.07	20.96	19.07	20.16	15.27
	VFA	21.75	14.74	26.53	24.56	16.89	24.30	15.01	25.57	24.11	20.73	25.35	16.00
Relative Number of Served Locations	IA	34.14%	37.14%	29.33%	33.92%	29.11%	33.43%	21.03%	41.49%	29.06%	19.21%	24.50%	23.56%
	DYN	29.86%	48.64%	20.77%	31.77%	20.75%	33.83%	19.21%	35.46%	31.23%	27.92%	30.12%	29.53%
	OSCS	41.42%	49.26%	30.42%	46.74%	28.85%	42.47%	28.86%	42.14%	41.92%	38.15%	40.33%	30.55%
	VFA	43.51%	58.98%	35.37%	49.13%	33.79%	48.61%	30.02%	51.16%	48.22%	41.47%	50.72%	32.01%
OD's utility surplus	IA	56.41	28.47	67.68	54.33	43.13	53.13	34.67	68.46	58.03	46.41	53.83	46.60
	DYN	40.41	45.05	34.96	38.26	29.09	45.92	23.44	48.61	36.58	37.58	32.46	25.17
	OSCS	56.91	39.51	66.77	64.11	42.15	56.42	40.53	61.89	58.78	48.38	47.38	30.95
	VFA	64.61	47.53	77.04	75.39	50.55	69.97	44.55	74.61	69.87	61.33	61.49	45.67
Avg. compensation per OD	IA	5.92	5.19	5.67	5.63	5.77	5.86	5.90	5.91	4.69	5.38	4.26	6.44
	DYN	3.67	4.48	3.28	3.42	3.75	3.67	3.47	3.70	3.37	3.64	2.48	4.78
	OSCS	4.76	4.21	4.54	4.79	4.58	4.34	4.79	4.12	4.45	4.64	2.66	4.86
	VFA	4.74	4.62	4.89	4.84	5.03	4.62	4.84	4.65	4.71	4.73	3.25	4.83

Table 3: Results of the sensitivity analysis. Each entry is the average value resulting from 500 simulation runs.

[1] Relative savings compared to IA.

[2] Relative OD costs refer to the share of costs that is attributed to compensations to ODs.

6.1.3 Results

Table 3 presents the results of our simulation study. The savings generated by utilizing ODs vary significantly depending on the parameter set. However, a clear ranking of the analyzed methods emerges, where IA consistently produces the lowest savings and VFA yields the highest savings. For example, under the base set of parameters, the average savings compared to no OD utilization are 13.9% for IA and 22.9% for VFA. Figure 4 illustrates the difference in savings between IA and VFA, as well as the proportion of costs attributed to OD compensations. To examine the cost-saving potential of each model capability, we consider the savings of IA as a lower bound and evaluate the relative savings when additional model capabilities are incorporated.

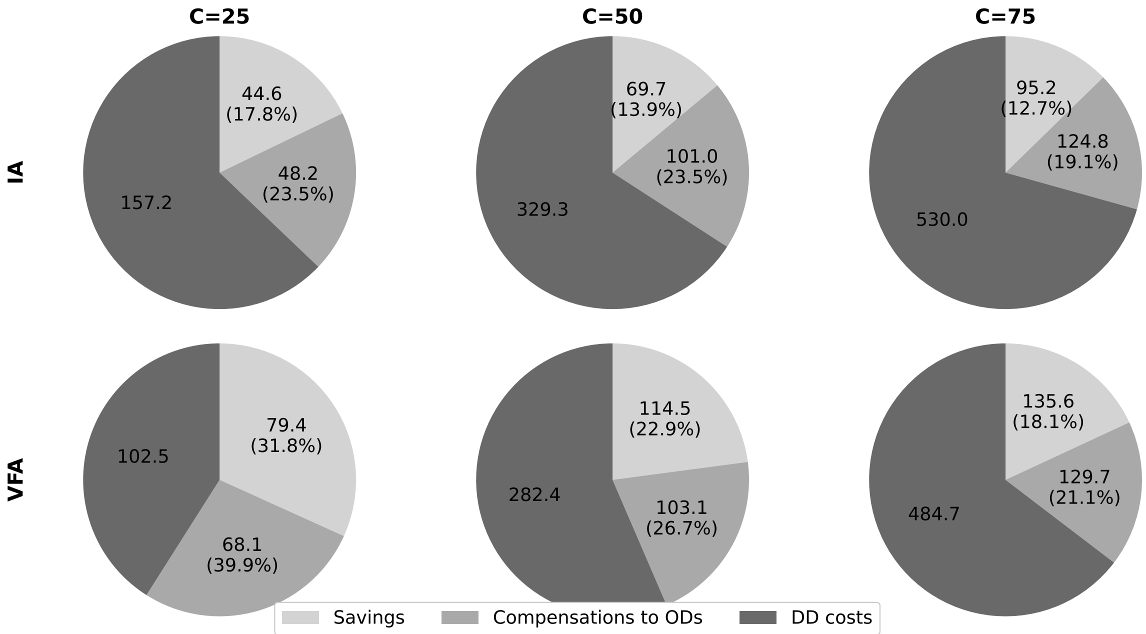

Fig. 4: Comparison of saved costs, compensations to ODs, and costs of DD delivery. The full circle represents the costs of a delivery without OD utilization. The middle column represents the base set of parameters, while the left and right columns represent the instances where C is decreased and increased, respectively.

The largest savings leap is observed when dynamic assignments are introduced in DYN. This alone accounts for 5%-100% of additional savings, depending on the set of parameters. The savings increase by an additional 3%-50% when an instance-optimized static compensation in OSCS is used. By further considering individual dynamic compensations, our methodology could generate an additional savings of 5%-33%. Overall, our sensitivity analysis has identified three main factors that impact the savings and other metrics, which we will elaborate on below.

Clustering tends to create more and better pairings (shorter detours) between ODs and delivery locations, leading to lower compensations and ultimately higher savings. Clustering also has a positive impact on the number of delivery orders served by ODs. The higher service rate also led to an increase in total paid compensation. The advantage of individual dynamic compensations (VFA) is persistent, but slightly lower compared to the base case. This is largely due to the more homogeneous IC resulting from shorter detours, which allows well-chosen static compensations to achieve decent results.

The ratio of the number of arriving ODs to delivery locations is an important factor influencing savings. Increasing the number of periods (T) or the number of potential ODs (O) positively affects the average number of arriving ODs. In general, a higher average number of arriving ODs results in higher savings and higher utilization of ODs. At the same time, compensations per delivery decreased on average using OSCS and VFA in comparison to the instances with the base set of parameters. These methods benefit from the implicit and explicit anticipation of expected avoided costs. The advantages of individual dynamic compensations rose in the $T = 75$ and the $O = 75$ scenario with

increasing average OD arrivals. This means the advantages of individual dynamic compensations are accelerated by a high OD emergence. A high supply of delivery orders (C) results in on average lower relative savings compared to the instances with the base set of parameters. However, total savings increase with C as more orders are accepted by ODs for lower compensations.
The structure of the IC is crucial for the potential savings. It can be noted that reducing the impact of the detour on the IC through a reduction of the detour scalar a_{oct} leads to higher average savings. We expected that with the decrease of a_{oct}, the additional relative savings generated by individual dynamic compensations (VFA) would decrease since the IC becomes more homogeneous. Surprisingly, this was not the case, indicated by a higher relative savings gap between OSCS and VFA, compared to the base case. Further investigation revealed that not only compensations but also the additional non-myopic assignment strategy played a decisive role. When the interval length of the IC b_{oct} was reduced, the advantage of individual dynamic compensations increased significantly. The main reason for this is that OSCS had a structural disadvantage in this scenario because the ICs were more heterogeneous on average.
Inspired by Le et al (2021), we also tracked the average sum of OD utility surpluses, which is the difference between the realized IC and the compensation of an accepted delivery request. This can be seen as an indicator of the OD's satisfaction with the participation and the chances of them participating again. The most important aspect, that was contradictory to what we were expecting is that the overall OD's utility surplus was highest in VFA over all instances. This shows that individual compensations do not necessarily contradict a high OD satisfaction because of IC exploitation. Quite to the contrary, both the store and the ODs can benefit from individual compensations.

6.2 Impact of time windows

In this section, we demonstrate that our framework, while primarily designed for unattended home delivery, can also be adapted to scenarios involving time windows, making it applicable to attended home delivery. For example, a customer may have the option to be delivered in the morning (8:00-12:00), the afternoon (12:00-16:00), or the evening (16:00-20:00). Each of these time windows can be treated as a delivery horizon, allowing us to consecutively apply our approach to effectively engage ODs during these periods (for similar approaches, see, e.g., Arslan et al, 2019; Silva et al, 2023a). However, we demonstrate the trade-off between offering short time windows and the potential savings achieved through OD utilization. This trade-off is not specific to our model but is inherent in engaging ODs. While narrow delivery time windows may be important for customer satisfaction, they reduce the likelihood of finding a suitable OD to carry out the delivery. In our computational study, we experimentally evaluate this trade-off, providing valuable insights for companies considering the integration of ODs into their delivery processes.
We examine this trade-off by dividing the time horizon of our base case into several delivery horizons. For easier divisibility, we start with 60 delivery requests, 60 potential ODs, and 60 periods, representing, for example, a business day. The arrival rate is defined by $\lambda_{ot} = 0.7 \cdot \frac{1}{|\mathcal{O}_t|}$ for $o \in \mathcal{O}_t$. For the remainder of this section, all other parameters are set according to the base set of parameters introduced in Section 6.1. In each scenario, 10% of each delivery horizon is dedicated to DDs, while the remaining periods are available for engaging ODs. Figure 5 illustrates the design of these delivery horizons, each consisting of a timeframe for OD engagement and DD deliveries, separated by the dispatch time. In all three scenarios we investigated, the 60 delivery tasks are evenly distributed across the time windows. To ensure comparability of simulated savings, all three scenarios share the same stream of ODs. This is crucial because different time window designs allocate varying periods for DD deliveries, affecting the availability of ODs who can only be engaged outside these periods. For example, if OD 63 arrives in period 28, they can only perform delivery in the first and third scenarios, as period 28 is dedicated to DD deliveries in scenario 2. If an OD arrives in any period (including those dedicated to DD deliveries), they are excluded from the set $\mathcal{O}$ of potential future ODs in the current and subsequent delivery horizon, reflecting the assumption that ODs go shopping at most once a day. The following results are generated by using our analytical solution (refer to Propositions 1 and 2) in combination with FA-SP.

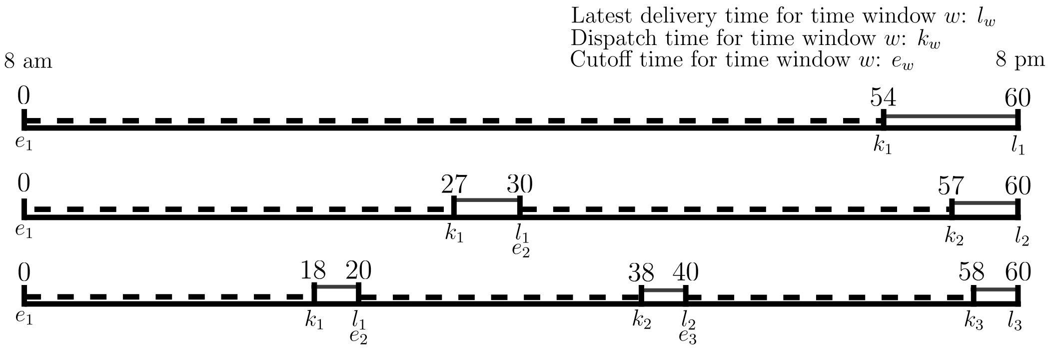

Fig. 5: Illustration of the time windows in each of the three scenarios.

# of TW	Savings	# Utilized ODs	Avg. Detour	Avg. Compensation
1	154.12	29	0.59	4.69
2	120.36	24	1.28	4.99
3	99.09	19	1.18	4.78

Table 4: Results of the Computational Study on Time Windows

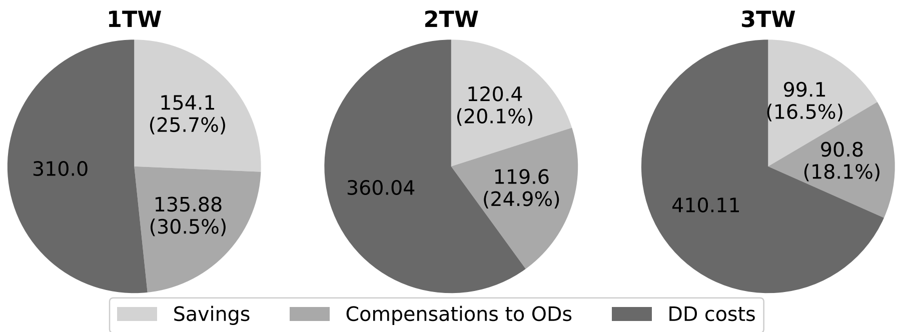

Fig. 6: Visualization of the differences in savings and OD utilization with different time window granularities. The full circle depicts the costs that would have occurred without ODs.

The results of the computational study (refer to Table 4) demonstrate that dividing the time horizon into multiple segments has a significant impact on savings. The highest savings, amounting to 154, were achieved without any division of the time horizon. When the time horizon was divided into two equal parts, the savings decreased to 120, which is around 22% less than the savings achieved without division. Further dividing the time horizon into three equal parts resulted in even lower savings of approximately 100, which is around 16.6% less than the savings from the two-part division (see Figure 6).

This indicates that the division of the time horizon into multiple segments diminishes the cumulative savings that could be achieved over the entire business day without segmentation. The underlying reason for this reduction is that when delivery tasks are confined to specific time windows, certain pairings that were previously valid become invalid, necessitating the search for alternative pairings. Consequently, this leads to generally less profitable pairings and thus fewer assignments.

This trend is also reflected in the number of ODs utilized, which decreases from 29 without any time window division to 24 with a two-part division, and further down to 19 with a three-part division. However, the average detour taken by ODs does not exhibit a clear pattern. The lowest average detour, approximately 0.59, was observed in the scenario without any division. Contrary to expectations, the average detour in the two-part division scenario was higher at around 1.28, compared to 1.18 in the three-part division scenario. This anomaly can be explained by the reduced number of delivery tasks fulfilled by ODs in the segmented scenarios, leading to the exclusion of some less favorable pairings.

In general, the choices become more limited as the number of time windows increases. In Figure 7 the assignments in each scenario are visualized. Each graph represents a full business day with 54 arrival periods. Despite the obvious trade-off, even when dividing the business day into three-time windows, there are still significant benefits in utilizing ODs, indicated by savings of about $\frac{1}{6}$ of the total costs without ODs.

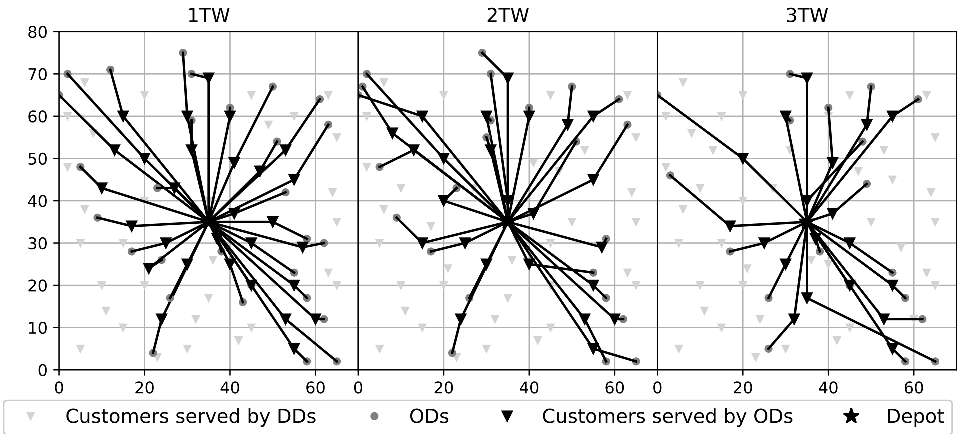

Fig. 7: Visualization of the assignment policy observed in each scenario.

7 Conclusion

In our work, we introduce an innovative approach to leverage 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 unpredictable 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 destination. This approach, similar to a reverse dynamic pricing model, explicitly accounts for the inherent unpredictability in the availability and decision-making of occasional drivers.

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 of the avoided costs, paving the way for tailoring well-known approximation methods to our problem structure.

In a comprehensive simulation study, we evaluate policies derived from our analytical state-dependent solution in combination with our approximation methods and compare resulting cost savings with those from methods inspired by recent literature. In these comparative methods, we incrementally introduce the capabilities of our approach: the anticipation of uncertain dynamic arrivals of ODs, the anticipation of uncertain OD acceptance, and the optimization of individual dynamic compensations for ODs. Our results suggest that significant savings can be achieved when all these capabilities are combined. However, the extent of the savings varies significantly depending on the set of parameters. For instance, scenarios with a high participation rate (i.e., a large number of potential ODs) tend to yield higher savings. Moreover, our approach not only demonstrated the highest savings potential across all instances but also proved to be the most beneficial for ODs in terms of their total utility surplus. Finally, we demonstrate the applicability of our framework in scenarios involving time windows, where it effectively generates significant savings. We also analyze the impact of time window length, offering firms a valuable decision-making tool for balancing the trade-off between OD engagement and time window design.

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

Appendix A List of notation

Table A1: Notation and description

Symbol	Description
Sets and Collections	
$(c \in)\mathcal{C}$, $(c_t \in)\mathcal{C}_t$	Index set of delivery locations (remaining in period t)
$(o \in)\mathcal{O}$, $(o_t \in)\mathcal{O}_t$	Index set of ODs' destinations (remaining in period t)
$(t \in)\mathcal{T}$	Index set of periods
$\mathcal{X}_C(S)$	Set of delivery locations valid for an offer in state S
$\mathcal{X}_R(S)$	Set of compensations valid for an offer in state S
$\mathcal{X}(S)$	Collection of all possible decisions $(\mathcal{X}_C(S), \mathcal{X}_R(S))$
Parameters	
C	Number of delivery locations at the beginning of the horizon
O	Number of potential ODs' destinations at the beginning of the horizon
d_{uv}	Distance from node u to node v
u_{oc}	Detour for OD o to serve delivery location c
λ_{ot}	Arrival probability of OD o in period t
$\overline{\lambda}_t$	No-show probability in period t
κ_c	Fix cost incurred when delivery location c is delivered by a DD
a_{oct}	Lower bound of the indifference compensation for OD o serving delivery location c in period t
$a_{oct} + b_{oct}$	Upper bound of the indifference compensation for OD o serving delivery location c in period t
States	
S_t^A	Pre-arrival state $(\mathcal{O}_t, \mathcal{C}_t)$
S_t^X	Pre-decision state $(\mathcal{O}_t, \mathcal{C}_t, o_t)$
S_t^P	Post-decision state $(\mathcal{O}_t, \mathcal{C}_t, o_t, c_t)$
Functions	
$\Theta(S)$	Costs of serving the delivery locations in state S with DDs
$V_t(S)$	Costs of being in state S in period t
$Sav_t(S^P)$	One-period cost reduction (savings) achieved in post-decision state S^P in period t
$\Delta V(\mathcal{O}_t, \mathcal{C}_t, o, c)$	Expected cost difference of entering the next stage with and without delivery location c and without OD o
IC_{oct}	Random variable, that resembles the minimal compensation an OD o is willing to accept in period t to serve delivery location c (indifference compensation)
$f_{oct}(x), f(x)$	Density function of the IC
$F(x)$	Cumulative density function of the IC
$h(x)$	Increasing failure rate
Decisions	
$r_t(S) \in \mathcal{X}_R(S), r$	Compensation (offered in a state S in period t)
$c_t(S) \in \mathcal{X}_C(S), c_t, c$	delivery location (offered in state S in period t)
Solution method specific Functions	
$\phi_o(t)$	Arrival probability of OD o in any period after t
$\Delta\overline{V}_t(\mathcal{O}_t, \mathcal{C}_t, o, c \mid \eta)$	Approximated avoided costs in a pre-decision state for delivery location c using the VFA with weights η
$\overline{V}_t^{FA}(\mathcal{O}_t, \mathcal{C}_t)$	Approximated value of being in a pre-arrival state in period t using the FA algorithm
$P_{Ar}(o)$	The fraction of OD o that arrives within the remaining time, designed to resemble the corresponding arrival probability.
$P_{Ac}(o, c, r_{oc})$	The fraction of OD o that serves delivery location c for compensation r_{oc}
$\Delta\overline{V}_t^{FA}(\mathcal{O}_t, \mathcal{C}_t, o, c)$	Approximated avoided costs of delivery location c using the FA algorithm
$\Delta\overline{V}_t^{SP}(\mathcal{O}_t, \mathcal{C}_t, o, c)$	Approximated avoided costs of delivery location c using the FA-SP algorithm
$\mathcal{N}_k(\mathcal{O}, \mathcal{C}, c)$	kth-order neighborhood of a pre-arrival state in which customer c is served
Algorithm components	
η_{oc}	Weight of delivery location c and OD o
Q	Number of simulation runs in one iteration
N	Number of iterations
r_{qct}	Observed delivery costs for delivery location c during simulation run q
h_{qct}	Auxiliary variable that is 1 if delivery location c was served in period t in simulation run q
τ_{qc}	Period in which delivery location c was served during simulation run q
$\mathcal{O}_{qt}$	Set of remaining ODs in run q at period t
o_{qt}	OD who arrived in simulation run q at period t
ζ_c	Shadow price of delivery location c (optimal dual of c specific constraint (23))

Appendix B Additional information about the approximation methods

B.1 Pseudo-code

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 $\eta_{oc} \; \forall \; o, c$

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**
      $(\overline{\mathcal{O}}, \overline{\mathcal{C}}) \leftarrow \mathcal{N}_k(\mathcal{O}, \mathcal{C}, c)$
      $\Delta \overline{V}_t^{FA}(\overline{\mathcal{O}} \cup \{h\}, \overline{\mathcal{C}}, h, c) \leftarrow \overline{V}_{t+1}^{FA}(\overline{\mathcal{O}}, \overline{\mathcal{C}}) - \overline{V}_{t+1}^{FA}(\overline{\mathcal{O}}, \overline{\mathcal{C}} \setminus \{c\})$
    end for
    $c^* \leftarrow \operatorname{argmax}_{c \in \mathcal{C}} \{\Delta \overline{V}_t^{FA}(\overline{\mathcal{O}} \cup \{h\}, \overline{\mathcal{C}}, h, c) - a_{hct}\}$
    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 $\mathcal{G}$

B.2 Interpretation of the weights and the relationship to the basis functions

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 being 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 B1, 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 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 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 3, 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 particularly evident for delivery locations 4 and 5. While the actual avoided costs at the delivery location 5 exhibit a slight deviation towards the end of the observed time horizon, the overall pattern remains closely aligned with the approximated avoided costs. The same holds for delivery location 4. 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' relationships.

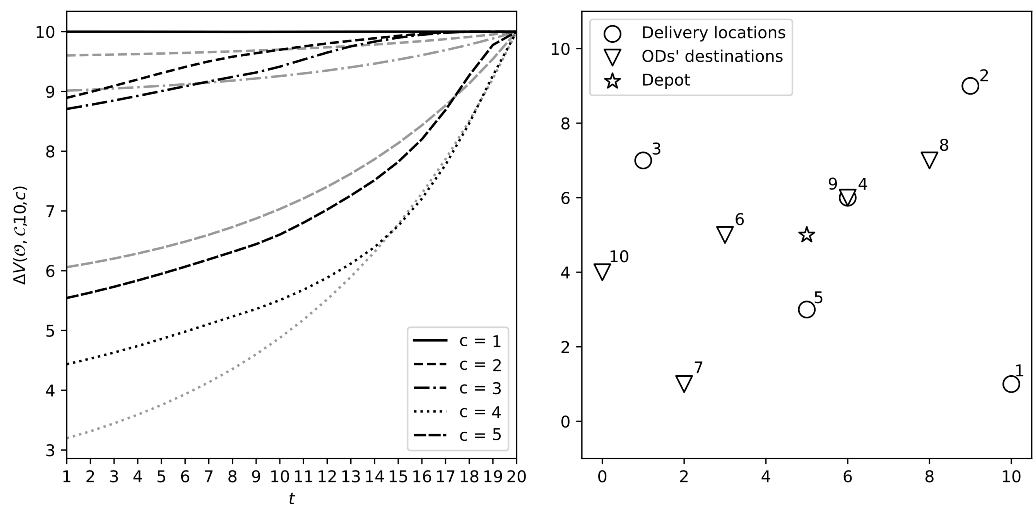

Fig. B1: The graph on the left compares the avoided costs and the approximate avoided costs in an instance with 5 ODs and 5 delivery locations. The brighter lines represent the approximate avoided costs, while the darker lines depict the true avoided costs calculated by full enumeration. The different 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.

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 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 tested this approach in our simulation study in Section 6 and compared its performance with the performance of other mechanisms.

B.3 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, the offer is unprofitable. We reduce computational complexity by excluding delivery locations and ODs that are likely never to 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 near 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 $\mathcal{O}'(\mathcal{O}, \mathcal{C}, c) = \{o' \in \mathcal{O} | \, c_{o'}^1(\mathcal{C}) = c\}$. This set can be described as the set of all OD destinations with delivery location c as their shortest detour delivery location. The set of neighboring delivery locations of a delivery location c is defined as $\mathcal{C}'(\mathcal{O}, \mathcal{C}, c) = \{c' \in \mathcal{C} | \, \exists o \in \mathcal{O}'(\mathcal{O}, \mathcal{C}, c)$ such that $(c_o^2(\mathcal{C}) = c') \vee (c_o^3(\mathcal{C}) = c')\}$. 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 B2. In this example, the second-degree neighborhood is created.

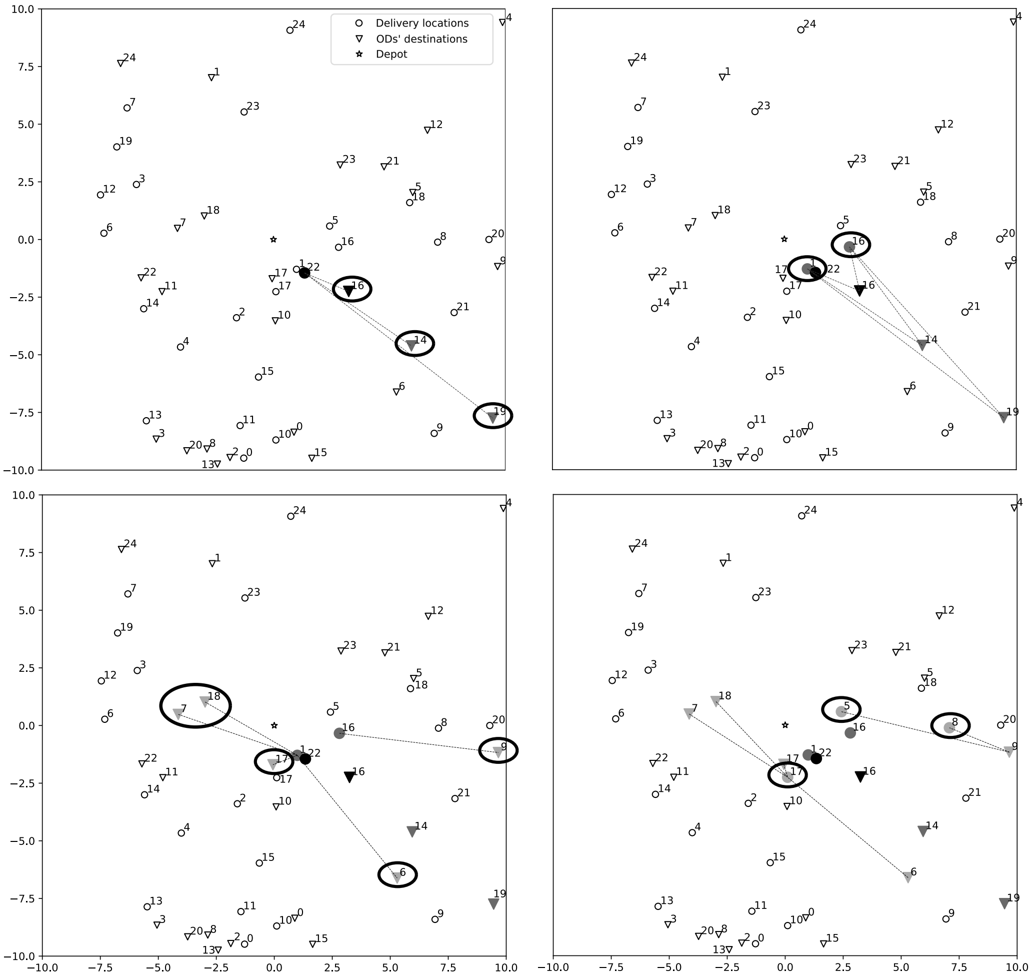

Fig. B2: The subsets of relevant delivery locations ($\overline{\mathcal{C}}$) and the subset of relevant OD destinations ($\overline{\mathcal{O}}$) are generated iteratively, as demonstrated through a series of figures. The first figure (upper left) illustrates delivery location 1 and the three ODs (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.

Appendix C Proofs

C.1 Proofs of Proposition 3

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 $\lambda_{\tilde{o}t} + \overline{\lambda}_t = \tilde{\lambda}_{0t}$, where $\overline{\lambda}_t$ 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 $t + 1$.

□

C.2 Proof of Proposition 4

Proof. In order to validate the proposed propositions, we present counterexamples for each, utilizing a specially designed instance as depicted in Figure C1. 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 C1.

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.

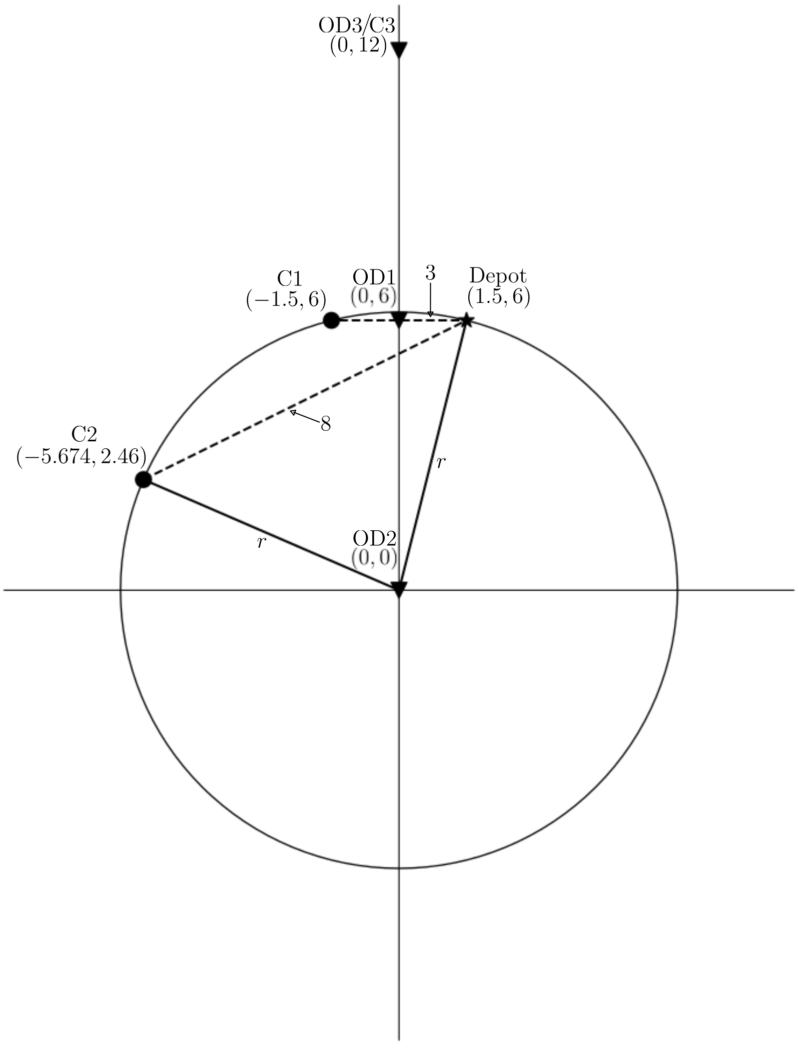

Fig. C1: Graphical representation of the created example instance.

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

Table C1: 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 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 C2.
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 C3.
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.

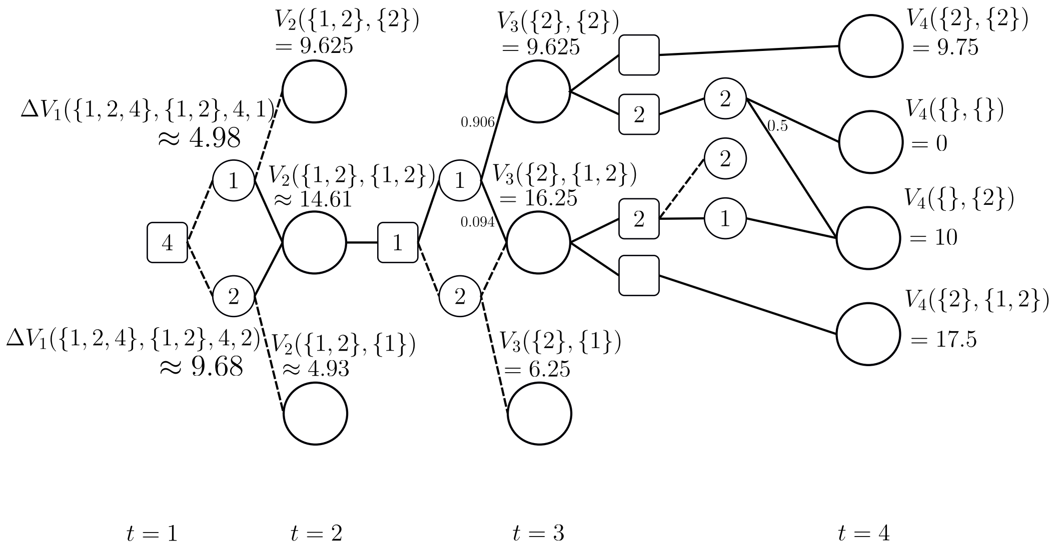

Fig. C2: Tree diagram of the example instance of Example 1 (without the additional OD).

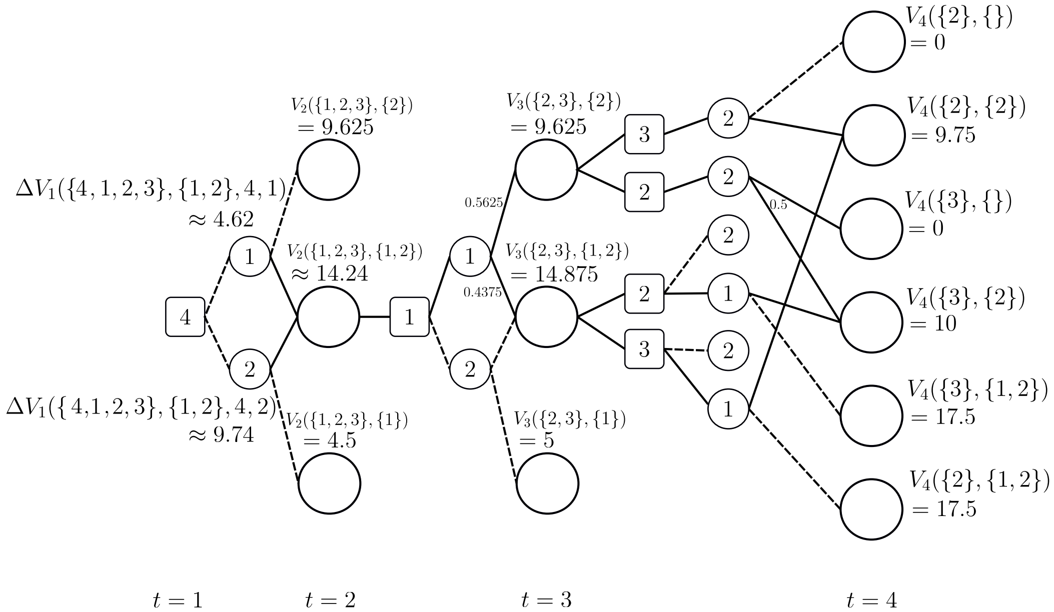

Fig. C3: Tree diagram of the example instance of Example 1 (with the additional OD).

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 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 C4. To simplify, the choices illustrated in the pre-decision states (represented as squares) have been reduced to only the optimal choices.

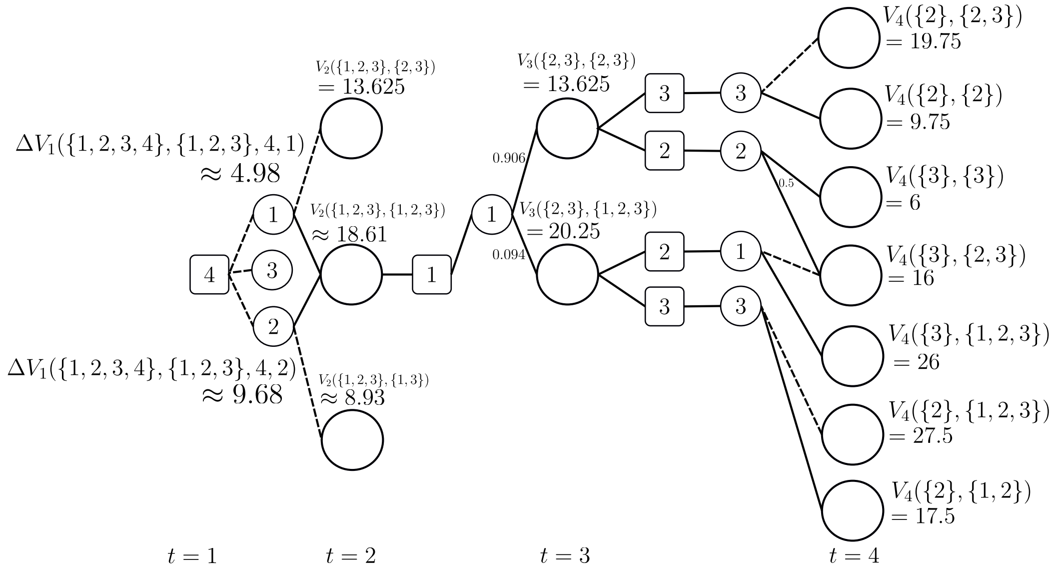

Fig. C4: Tree diagram of the example instance of Example 2 (with the additional delivery location).

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 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 the 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 C5.

□

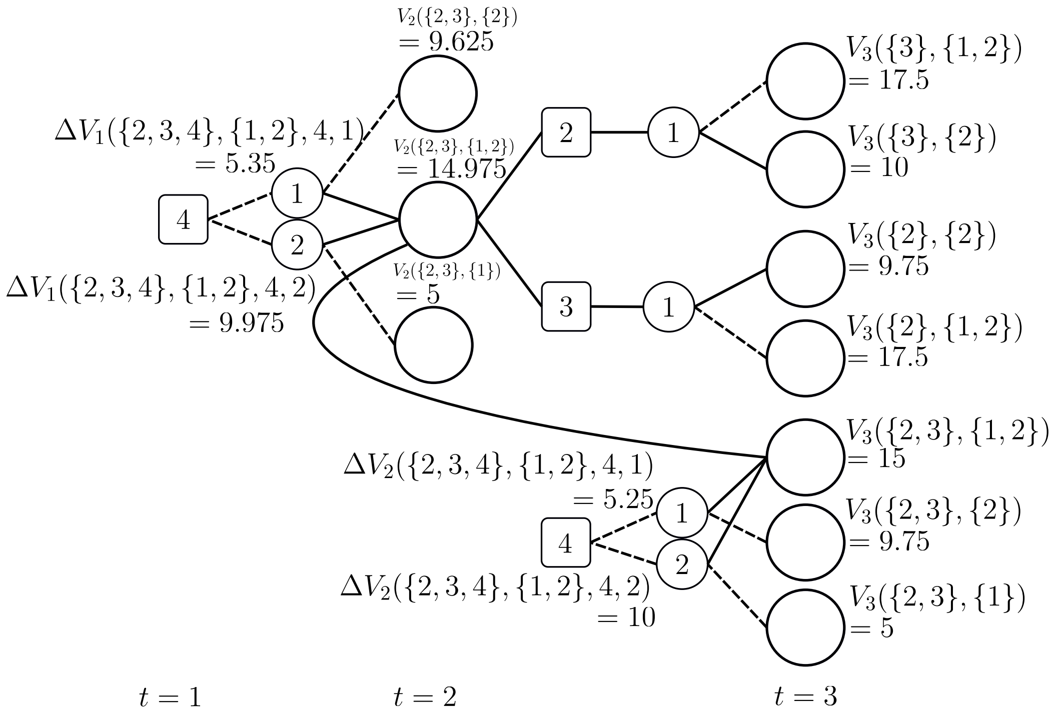

Fig. C5: Tree diagram of the example instance of Example 3 (progressing t).

C.3 Proof of Lemma 3

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 FA as follows: Firstly, we introduce the decision variables $r_{o\tilde{c}_t} \geq 0$ to the optimization model. Secondly, a non-negative term is added to the objective function (22). Specifically,

$$
\begin{aligned}
&\sum_{c\in\mathcal{C}_t}\sum_{o\in\mathcal{O}_t} P_{\mathrm{Ar}}(o)\cdot P_{\mathrm{Ac}}(o,c,r_{oc})\cdot r_{oc} + \sum_{c\in\mathcal{C}_t}\left(1-\sum_{o\in\mathcal{O}_t} P_{\mathrm{Ar}}(o)\cdot P_{\mathrm{Ac}}(o,c,r_{oc})\right)\cdot\kappa_c \\
&= \sum_{c\in\tilde{\mathcal{C}}_t}\sum_{o\in\mathcal{O}_t} P_{\mathrm{Ar}}(o)\cdot P_{\mathrm{Ac}}(o,c,r_{oc})\cdot r_{oc} + \sum_{c\in\tilde{\mathcal{C}}_t}\left(1-\sum_{o\in\mathcal{O}_t} P_{\mathrm{Ar}}(o)\cdot P_{\mathrm{Ac}}(o,c,r_{oc})\right)\cdot\kappa_c \\
&\quad + \sum_{o\in\mathcal{O}_t} P_{\mathrm{Ar}}(o)\cdot P_{\mathrm{Ac}}(o,\tilde{c}_t\, r_{o\tilde{c}_t})\cdot r_{o\tilde{c}_t} + \left(1-\sum_{o\in\mathcal{O}_t} P_{\mathrm{Ar}}(o)\cdot P_{\mathrm{Ac}}(o,\tilde{c}_t\, r_{o\tilde{c}_t})\right)\cdot\kappa_{\tilde{c}_t}
\end{aligned}
\tag{C1}
$$

   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 $\overline{V}_t(\mathcal{O}_t,\mathcal{C}_t)$ 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 $\overline{V}_t(\mathcal{O}_t,\mathcal{C}_t)$ 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 $\overline{V}_t(\mathcal{O}_t,\mathcal{C}_t)$, 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 FA 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 $\overline{V}_t(\tilde{\mathcal{O}}_t, \mathcal{C}_t)$. 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 $\overline{V}_{t+1}(\mathcal{O}_t, \mathcal{C}_t)$. 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 delivery location now carries lower expected costs, leading to a reduction in the objective value. The decision variables concerning delivery 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 $t+1$.

□

References

Akçay Y, Natarajan HP, Xu SH (2010) Joint dynamic pricing of multiple perishable products under consumer choice. Management Science 56(8):1345–1361. https://doi.org/10.1287/mnsc.1100.1178

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, Zuidwijk R (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

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

Dahle L, Andersson H, Christiansen M, Speranza MG (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

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

Dong L, Kouvelis P, Tian Z (2009) Dynamic pricing and inventory control of substitute products. Manufacturing & Service Operations Management 11(2):317–339. https://doi.org/10.1287/msom.1080.0221

Feng X, Chu F, Chu C, Huang Y (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

Gallego G, Van Ryzin G (1997) A multiproduct dynamic pricing problem and its applications to network yield management. Operations Research 45(1):24–41. https://doi.org/10.1287/opre.45.1.24

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

Hou S, Gao J, Wang C (2023) Optimization framework for crowd-sourced delivery services with the consideration of shippers' acceptance uncertainties. IEEE Transactions on Intelligent Transportation Systems 24(1):684–693. https://doi.org/10.1109/TITS.2022.3215512, URL https://ieeexplore.ieee.org/document/9945652/

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, Van Woensel T (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, Laporte G (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, 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, Shen ZJM (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

Santini A, Viana A, Klimentova X, Pedroso JP (2022) The probabilistic travelling salesman problem with crowdsourcing. Computers & Operations Research 142:105722. https://doi.org/10.1016/j.cor.2022.105722

Savelsbergh MW, Ulmer MW (2022) Challenges and opportunities in crowdsourced 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

Solomon MM (1987) Algorithms for the vehicle routing and scheduling problems with time window constraints. Operations Research 35(2):254–265. https://doi.org/10.1287/opre.35.2.254

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

Zhang D, Cooper WL (2009) Pricing substitutable flights in airline revenue management. European Journal of Operational Research 197(3):848–861. https://doi.org/10.1016/j.ejor.2006.10.067

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