Integrating Collection-and-Delivery Points in the Strategic Design of Last-Mile E-Commerce Distribution Networks

by

Himanshu Rautela

Bachelor of Engineering (Honors), Mechanical Engineering
Birla Institute of Technology and Science (2009)

SUBMITTED TO THE PROGRAM IN SUPPLY CHAIN MANAGEMENT
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF ENGINEERING IN SUPPLY CHAIN MANAGEMENT
AT THE
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

JUNE 2019

© 2019 Himanshu Rautela. All rights reserved.
The author hereby grants to MIT permission to reproduce and to distribute publicly paper and electronic copies of this thesis document in whole or in part in any medium now known or hereafter created.
[This page is intentionally left blank]
Integrating Collection-and-Delivery Points in the Strategic Design of Last-Mile E-Commerce Distribution Networks

by

Himanshu Rautela

Submitted to the Program in Supply Chain Management on May 10, 2019 in Partial Fulfillment of the Requirements for the Degree of Master of Engineering in Supply Chain Management

ABSTRACT

The rapid growth in e-commerce volumes, coupled with customer expectations of faster, flexible and cheaper parcel deliveries is increasing the pressure on retailers to design the most efficient delivery network. Collection-and-delivery points (CDPs) allow for the aggregation of demand and enable reductions in travel time and costs. CDPs also help minimize additional tours arising due to failed deliveries or failed pickups for returns. We formulate an optimization model that integrates CDPs in the design of the overall distribution network, including the location of upstream transshipment facilities. The model accounts for changes in demand density due to the placement of CDPs. It considers demand aggregation at the CDP for both forward and return flows, and the impact of failed deliveries and failed return pickups on the routing cost. The model considers multiple different route options and solves them using extended routing cost approximation formulae thus allowing the implementation of the model on large-scale problems. We then apply the model to solve a real-world case study on the last-mile distribution network of a major Brazilian e-commerce retailer. The results demonstrate that failed deliveries and failed return pickups increase both the last-mile cost and the overall cost of distribution, and CDPs effectively reduce these costs by aggregating the demand and minimizing travel time.

Thesis Advisor: Dr. Milena Janjevic
Title: Postdoctoral Associate, Center for Transportation and Logistics
Thesis Errata Sheet

Author  Himanshu Rautela
Primary Dept.  Department of Supply Chain Management
Degree  Master  Graduation date  Jun 2019

Thesis title
Integrating Collection-and-Delivery Points in the Strategic Design of Last-Mile E-commerce Distribution Networks

Brief description of errata sheet
An additional reference to the paper by Janjevic et al. (2019) is added to the paper.

Number of pages  2  (11 maximum, including this page)

► Author: I request that the attached errata sheet be added to my thesis. I have attached two copies prepared as prescribed by the current Specifications for Thesis Preparation.

Signature of author  ____________________________________________ Date 08/31/2020

► Thesis Supervisor or Dept. Chair: I approve the attached errata sheet and recommend its addition to the student’s thesis.

Signature  ____________________________________________ Date 08/31/2020
Name  Milena Janjevic  Thesis supervisor  □ Dept. Chair

► Vice Chancellor or his/her designee:
I approve the attached errata sheet and direct the Institute Archives to insert it into all copies of the student’s thesis held by the MIT Libraries, both print and electronic.

Signature  ____________________________________________ Date 12/17/2020
Name: Ian A. Waitz
ERRATA

This thesis develops an original optimization models that is partially based on analytical formulations developed by Janjevic et al. (2019). The following changes should be made:

- Section 3.5, paragraph 3: the following sentence should be added at the end of the paragraph:
  Similarly, Janjevic, Winkenbach and Merchán (2019) integrate CDPs in network design but only consider forward flows and ignore failed deliveries and returns.

- Section 4.3, the following sentence should be added in the beginning of the section:
  In the following, we present our extension of ARCE formulation by Janjevic, Winkenbach and Merchán (2019).

- Section 4.3.3, sentences 1-4 should be replaced by:
  To estimate the the cost of those tours, $f_{ij}^{BT}(w)$ and $f_{ij}^{C}(w)$, we rely on ARCE formulae introduced by Winkenbach et al. (2016) and further extended by Janjevic, Winkenbach and Merchán (2019) to account for individual customers and CDPs. As demonstrated by Janjevic, Winkenbach and Merchán (2019), the introduction of CDPs gives rise to different types of routes (i.e., blended or customer-only) but does not impact the quality of the approximation formulae and only results in the adaptation of its parameters.

- References: the following reference should be added:
Acknowledgments

First and foremost, I would like to thank my thesis advisor, Dr. Milena Janjevic, for her unwavering support, patience, and guidance. This thesis has been possible largely due to the time and resources she has invested in this work.

I am grateful to Dr. Matthias Winkenbach for allowing me to work in the Megacity Logistics Lab and providing me access to the required data and tools. Thanks to fellow CTL student Maggie Wilson for helping me with the implementation of the research work in Python.

Thanks to my thesis committee members, Dr. Chris Caplice, Dr. Bruce Arntzen, Dr. Josue Velazquez Martinez, and Dr. Maria Jesus Saenz, for their periodic feedback and suggestions to improve the output of the research.

In addition, I would like to thank Toby Gooley and Pamela Siska for reviewing the manuscript and providing valuable feedback.

Lastly, thanks to Justin Snow for all the administrative help during the program.
Table of Contents

List of Figures .............................................................................................................................................. 9
List of Tables .................................................................................................................................................. 10
Chapter 1: Introduction .................................................................................................................................. 11
  1.1 — Multi-echelon networks .................................................................................................................. 12
  1.2 — Collection-and-delivery points ...................................................................................................... 14
  1.3 — Research objective ........................................................................................................................ 15
  1.4 — Thesis structure ................................................................................................................................ 16
Chapter 2: Problem Setting .......................................................................................................................... 18
Chapter 3 — Literature Review .................................................................................................................... 21
  3.1 — Location-routing problem (LRP) ..................................................................................................... 21
  3.2 — Two-echelon location-routing problem (2E-LRP) .......................................................................... 23
  3.3 — Continuum approximation method ............................................................................................... 25
  3.4 — Consumer behavior and adoption of collection-and-delivery points (CDPs) ......................... 26
  3.5 — Integration of collection-and-delivery points in the LRP ............................................................... 27
  3.6 — Literature gap ................................................................................................................................... 29
Chapter 4: Methodology ............................................................................................................................. 31
  4.1 — Definition of the distribution network ......................................................................................... 31
  4.2 — The optimization model ............................................................................................................... 36
  4.3 — The routing cost ............................................................................................................................. 38
    4.3.1 — Cost of dedicated CDP routes ................................................................................................. 40
    4.3.2 — Cost of delivering CDP volumes on blended routes .............................................................. 41
    4.3.3 — Regular cost of delivery on blended and customer only routes ............................................ 42
Chapter 5: Results and Analysis .................................................................................................................. 48
  5.1 — Model parameters .......................................................................................................................... 48
    5.1.1 — Demand data and operational parameters ........................................................................... 48
    5.1.2 — Cost parameters .................................................................................................................... 50
  5.2 — Definition of scenarios .................................................................................................................... 51
    5.2.1 — Scenarios without CDPs ........................................................................................................... 51
    5.2.2 — Scenarios with CDPs ............................................................................................................... 53
  5.3 — Analysis of results .......................................................................................................................... 55
5.3.1 — Analysis of distribution networks without CDPs.......................................................... 56
5.3.2 — Analysis of distribution networks with CDPs ........................................................................... 59

Chapter 6: Discussion ........................................................................................................................................ 63
6.1 — Cost savings generated by integrating CDPs in different scenarios .................................................... 63
6.2 — Estimating maximum savings from CDPs............................................................................................ 66
6.3 — Exploring the effect of alternative governance models ....................................................................... 66

Chapter 7: Conclusion ....................................................................................................................................... 70
References.......................................................................................................................................................... 72
List of Figures

Figure 1: Illustration of the distribution network and routes ................................................................. 19
Figure 2: Demand density in the Metropolitan Sao Paulo region ........................................................... 49
Figure 3: Facility location in the Metropolitan Sao Paulo region ............................................................ 50
Figure 4: Optimal SF configuration in the network without CDPs .......................................................... 56
Figure 5: Optimal SF configuration in the network with CDPs ............................................................... 59
List of Tables

Table 1: General model parameters ........................................................................................................32
Table 2: Operational parameters ..............................................................................................................32
Table 3: Demand parameters ....................................................................................................................33
Table 4: Cost parameters ..........................................................................................................................35
Table 5: Additional notations expressing demand parameters due to CDPs...........................................35
Table 6: Parameter values for scenarios without CDPs .............................................................................52
Table 7: Parameter values for scenarios with CDPs .................................................................................54
Table 8: Total cost of distribution for the network without CDPs ...........................................................57
Table 9: Total cost of distribution for the network with CDPs .................................................................60
Table 10: Cost savings in a network with CDPs compared to a network without CDPs .........................64
Table 11: Total cost of distribution for the network with outsourced CDP operations..........................67
Table 12: Cost savings in the outsourced model compared to the proprietary model .........................68
The rise of mobile technologies and the Internet has led to more and more consumers choosing the convenience of online shopping. This technological boom has led to rapid growth in the e-commerce industry in both developed and emerging markets. Global e-commerce volumes are expected to grow from US$ 2.8 trillion in 2018 to US$ 5.8 trillion in 2022 (451 Research, 2018). Most of this growth is expected from emerging markets, where there will be 3 billion internet users by 2022, which will be three times as many internet users as in the developed markets (The Boston Consulting Group, 2018). The growth in emerging markets is driven by the unprecedented rise in the urban population in large cities. Today, 55% of the world’s population lives in urban areas, and by 2050, this number is expected to go up to 68% (United Nations, Department of Economic and Social Affairs, Population Division, 2018).

While the rapid growth of urban population creates further opportunities for the growth of e-commerce, it also poses several challenges for urban planning and sustainable logistics. As retailers strive to minimize delivery times and reduce transportation costs, high population density puts pressure on road traffic in residential areas, making home delivery in large cities even more challenging (Visser, Nemoto, and Browne, 2014). At the same time, customer expectations are changing towards anytime-anywhere shopping. Today consumers not only want the flexibility to order products online, delivered at a fast pace but also want the convenience to pick up or return the products at convenient physical locations. (BigCommerce, 2018) Thus, both brick-and-mortar and online retailers are realizing that the future of the industry is not merely online, but rather in omnichannel offerings which provide a consistent experience to customers by converging their offline and online channels (ATKearney, 2015).

The need to transport goods to consumers’ homes rather than to retail stores, combined with the customers’ need for faster delivery leads to a higher demand fragmentation, which results in lower level
of vehicle utilization in the last mile. The lower vehicle utilization increases the number of freight movements and related delivery costs. The last mile refers to the final stage in the distribution network which brings the goods or parcels to the consumers' doorsteps. In 2016, the total cost of the last-mile of parcel delivery was approximately US$ 80 billion (McKinsey & Company, 2016) and it continues to be the most expensive and time-consuming step in the overall fulfilment process (Business Insider Intelligence, 2019). E-commerce players use their delivery capabilities as a source of competitive advantage, and while a large group of consumers desires faster delivery and a variety of delivery options, they remain sensitive to the delivery charges (McKinsey & Company, 2016).

The customer expectations of faster and flexible parcel deliveries, at low costs, is increasing the pressure on retailers to design the most efficient delivery networks. Retailers can design their delivery networks either as single-echelon or multi-echelon. A single-echelon system offers a direct shipment from the origin to the destination while a multi-echelon system manages the distribution through one or more intermediate facilities where each echelon refers to one level of the distribution network. In single-echelon networks, origin facilities are far from the destination, leading to suboptimal vehicle utilization and route selection, thus making the deliveries slow and expensive (Boccia, Crainic, Sforza, and Sterle, 2010). Savelsbergh and Van Woensel (2016) explain that multi-echelon distribution networks that allow the consolidation of multiple shipments at a satellite facility (SF) present an opportunity to lower logistics costs by reducing freight movements, especially in megacities (cities with population greater than 10 million).

1.1 — Multi-echelon networks

In a multi-echelon network, SFs function as transshipment nodes which offer an opportunity for consolidation of inbound shipments and deconsolidation of outbound shipments. These nodes also allow
the usage of two different transportation modes, a bigger vehicle for line-haul inbound transportation and a smaller vehicle for outbound customer delivery. Dell'Amico and Hadjidimitriou (2012) have shown that the introduction of a transshipment area along with the usage of vehicles of two different capacities for each echelon not only reduces delivery time and cost but also reduces pollution and congestion. Multi-echelon distribution systems are currently being deployed by several online retailers in both the developed and emerging markets (e.g., Amazon in the US, B2W in Brazil, and JD.com in China) (Winkenbach and Janjevic, 2018). There is a growing interest in the design of multi-echelon networks as shown by the surveys of Drexl and Schneider (2015) and Nagy and Salhi (2007). Most of the current research attempts to solve this network design problem in the form of a location-routing problem (LRP), which combines the decision on the location of facilities with the decisions on the vehicle routing. Salhi and Rand (1989) have shown that making these decisions independently generates suboptimal results.

Traditionally, retailers partnered with existing carriers such as postal service operators for last-mile delivery of online orders. These carriers generally have a delivery infrastructure to consolidate parcel volumes and then deliver to the end-customers. This model was effective when e-commerce volumes were small, and demand was not dynamic. However, it is not designed to serve the on-demand requirement of modern-day consumers who often want products delivered within a day, if not in hours. In most cases, this infrastructure is also not equipped to cater to perishable goods which require special handling and temperature control. Retailers now see faster and flexible delivery options as a source of competitive advantage and designing a better last-mile delivery network as the final frontier to fight the e-commerce battle. While satellite facilities in a multi-echelon system offer a solution for consolidating freight and lowering the distribution cost, they do not offer customers the flexibility to pick up parcels from a convenient physical location at a convenient time.
1.2 — Collection-and-delivery points

In e-commerce distribution, collection-and-delivery points (CDPs) are facilities to which a carrier may deliver parcels for later independent pickup by the customer. CDPs can be either unattended, such as Amazon’s parcel lockers in the U.S, or attended, such as Flipkart’s partnership in India with a local pharmacy (Apollo) that serves as CDPs. The former are mostly common in developed markets, while the latter are common in emerging markets (Winkenbach and Janjevic, 2018). CDPs not only allow for the aggregation of demand and enable reductions in travel time, travel distance, emissions, and traffic congestion but also reduce the number of failed deliveries or theft of packages delivered at the doorstep and left unattended.

The aggregation of many points of customer demand into a single point of delivery not only reduces the cost for the carrier, but it also saves costs linked to multiple delivery attempts originating from failed deliveries, especially for payment-on-delivery orders, a common practice in emerging markets (Winkenbach and Janjevic, 2018). Additionally, CDPs can also offer an option to aggregate the demand for returned products and multiple failed pickups for returned merchandise. Thus, flexible delivery and return options for the consumers and rising cost pressure for retailers due to multiple failed deliveries have led to a rise in the popularity of collection-and-delivery points across the world (World Economic Forum, 2018).

The financial benefit of multi-echelon network design combined with the financial benefit of CDPs and the flexibility offered to customers for pickups and returns opens new avenues for efficient network design. Existing works on multi-echelon network design generally assume that product exchange, or the delivery of the product to the end user, takes place at the point of the demand. Also, the literature on the use of CDPs in the context of urban last-mile distribution focuses largely on considerations of consumer preferences, choices, and behavior rather than carriers’ concerns of integrating these new types of
product exchange points including returns, into the overarching design of large-scale distribution networks.

Morganti, Dablanc, and Fortin (2014) have shown the benefits of automated CDPs in reducing traffic congestion and environmental pollution in France and Germany. The extant literature also covers location-routing problems that can consider multiple options for location, transportation modes, and routing. However, the existing models do consider servicing of the demand at the CDPs, and routing of returned products and failed deliveries through the CDPs in the strategic design of networks. These factors, if included in the decision-making framework, can aggregate many points of customer demand into a single point and lower the last-mile distribution cost, thus, changing the strategic design of e-commerce distribution networks.

1.3 — Research objective

The objective of this thesis is to propose a framework to integrate CDPs in the design of an optimal last-mile e-commerce distribution network and assess the implication of CDPs in changing the retailer’s overall distribution cost. This is done by formulating a multi-echelon location-routing problem (LRP) that includes CDPs in a broader network design problem, including location decision for upstream satellite facilities. The CDPs aggregate demand for both forward and return flows within the demand zone where they are placed and help minimize additional tours in the case of failed deliveries. The model thus accounts for changes in demand patterns for both forward and return flow due to the placement of CDPs. Additionally, it considers the impact of failed deliveries and multiple pickup attempts on the last-mile distribution cost.

The scalability of the solution has also been considered in formulating the problem. To develop a model which can be applied to solve large-scale problems with several thousand consumers, we extend
the routing cost approximation formulae proposed by Winkenbach, Kleindorfer, and Spinler (2016) to estimate the vehicle routing cost. These routing costs formulae constitute the routing component of the location-routing problem that we are modeling.

To analyze the real-world application of the model, we present a case study where the model is applied to the operations of a major Brazilian e-commerce player which services the Sao Paulo metropolitan region. We model and analyze several different scenarios covering delivery with and without the presence of CDPs and consider various parameters for failed deliveries and failed pickup attempts. These scenarios help us understand how CDPs can reduce the overall distribution costs by consolidating demand, streamlining return flows, and minimizing multiple attempts for failed home deliveries and return pickups.

The results of the analysis demonstrate how failed delivery attempts and failed pickup attempts increase the overall delivery costs and how CDPs can help reduce these costs as well. Additionally, the analysis supports the quantitative and qualitative comparison of CDP deployment scenarios and clearly show the cost reduction achieved due to the introduction of a CDP in a demand zone. The results thus help us understand the possible use cases for CDP deployment and the specific conditions, including demand and cost parameters, in which their use can yield cost savings.

1.4 — Thesis structure

This thesis is structured into seven chapters. In Chapter 2, we present the problem setting that we are modeling in this study and explain the various delivery models. Chapter 3 provides an extensive review of the literature relevant to the proposed problem setting and methodology. In Chapter 4, we develop and discuss the methodology used to integrate CDPs in the strategic design of last-mile distribution networks, including the optimization model, and the routing cost approximation formulae. In
Chapter 5, we present and discuss the results of implementing the model in a real-world case study. In Chapter 6, we discuss the cost savings achieved due to the introduction of CDPs and explore the different governance models to run a CDP network. Finally, in Chapter 7, we conclude our analysis and discuss potential paths for modeling and analysis of more advanced systems.
Chapter 2: Problem Setting

We consider a two-echelon distribution network with two levels of logistic facilities. This system is shown in Figure 1. A central distribution hub serves several satellite facilities using first-echelon vehicles that perform dedicated trips to each satellite facility (SF). The hub does not serve the customers directly. The SFs make the last-mile delivery in two different ways, either directly to individual customers or to CDPs. This second-echelon transportation is performed by second-echelon vehicles which have a lower capacity than the first echelon-vehicles. The deliveries in the second-echelon are to be performed within a given maximum service time.

The service-zone is split into multiple demand zones. Each demand zone has a given customer density (number of customers per square kilometer). The introduction of CDPs in a particular demand zone changes the demand patterns in the service area. CDPs attract part of the individual customer delivery and returns in their surroundings, resulting in a decrease in the density of individual customers. We assume that customer demand density decreases linearly with the distance from the CDP in that demand zone.
Figure 1: Illustration of the distribution network and routes

There can be three types of routes from the SFs to the demand zones. First, we have dedicated CDP routes, these constitute a direct route from the SF to the CDP without making any delivery to individual customers. These are the most economical routes and are selected for CDPs which attract a heavy demand. In the dedicated route, a fully-loaded vehicle picks parcels at the SF, drops them at the CDP, picks up returned parcels at the CDP and drops them at the SF on the direct backhaul route. This is illustrated as the route between SF1 and CDP1 in Figure 1.

The second type of routes combine deliveries to CDPs and individual customers and are referred to as blended routes. A blended route can service multiple customers and CDPs. These routes are applicable when the demand attracted by a CDP is either insufficient to justify a fully-loaded dedicated route, or the demand is large enough to require a partially loaded truck over and above the dedicated delivery of a fully loaded truck. In case of a blended route, the vehicle picks parcels at the SF for delivering
both at the CDP and at individual customer locations. It first services the individual customers, delivering the parcels and picking return items in the same tour. It then visits the CDP to deliver all parcels originally attracted at the CDP, and additional parcels rerouted to the CDP due to failed home deliveries. On the return route back to the SF, returned parcels aggregated at the CDP are delivered at the SF. The design of the route, such that the CDP is serviced after servicing the customers, ensures that the failed customer deliveries can be delivered at the CDP if the customer chooses that option. This route is illustrated between SF1 and CDP2, and SF2 and CDP4 and CDP5 in Figure 1.

The last type of routes are the customer-only routes. These can be of two sub-types. In the case, where there is no CDP in the demand zone, the vehicles deliver parcels to individual customers and pick returned products from customer locations and deliver them and failed deliveries back to the SF. These are illustrated as routes originating from SF1 in Figure 1. In an alternative scenario, where there is a CDP in the demand zone, the vehicles deliver parcels to individual customers, pick returned products from customer locations to be delivered back to the SF. However, the vehicles make an additional stop at the CDP to drop parcels from failed customer deliveries, for customers who chose the option to route failed deliveries at the CDP for pickup later. These routes are illustrated as routes originating from SF2 and making the intermittent stop at CDP3 in Figure 1. While this set of routes look very similar to the blended route originating from SF1 in Figure 1, the key difference is that in a customer-only route there is no parcel delivery scheduled from the SF to the CDP.
Chapter 3 – Literature Review

Today, e-commerce customers desire faster home deliveries and the flexibility to pick up and return items at a physical location, yet they remain highly price sensitive. For retailers, fulfilling customer demand requires an efficient distribution network, which will also offer them a source of competitive advantage in the growing e-commerce market. The integration of collection-and-delivery points (CDPs) in the design of a multi-echelon distribution network offers a solution to minimize the last-mile distribution cost by streamlining both forward and return flow and minimizing the number of failed delivery and failed pickup attempts for returned products.

This is done by modeling a two-echelon capacitated location-routing formulation which includes CDPs. This model determines the most cost-effective network configuration for last-mile delivery. To develop this model, this chapter draws on literature in four key areas: (i) location-routing problem to solve for strategic network design in last-mile distribution; (ii) two-echelon location-routing problem to solve for a two-tiered distribution network; (iii) continuum approximation method and its application to solve location-routing problem at scale; (iv) CDP adoption and consumer behavior, and (v) integrating CDPs in the design of the distribution network. Finally, the chapter concludes by identifying gaps in the existing literature.

3.1 — Location-routing problem (LRP)

The design of a distribution network involves two key decisions: (i) long-term strategic decisions about facility location and (ii) operational decisions about vehicle routing. The interdependence between distribution center location and transportation cost was first highlighted by Maranzana (1964), though he incorporated the shortest-path algorithm, instead of vehicle routing into the location allocation problem.
Laporte, Nobert, and Taillefer (1988) made the first classification of an LRP and examined a multi-depot, vehicle routing, and location-routing problem. Salhi and Rand (1989), demonstrated that making strategic location decision and tactical routing decisions independent of each other leads to suboptimal results. This is applicable even when the planning horizon is extremely long and vehicle routes are allowed to change (Salhi and Nagy, 1999). This publication led to subsequent research attempting to solve what is known as the location-routing problem (LRP). In the LRP, the primary objective is still to solve a facility location problem, while simultaneously considering the vehicle routing. The survey conducted by Nagy and Salhi (2007) documents the problem variants and the existing algorithms to solve the LRP. While most of the early research focused on single-echelon LRP with deterministic demand, recent research considers multi-echelon LRP with some variants considering stochastic demand and different types of vehicles across each echelon.

The location-routing problems are an integration of two classical problems, the location allocation problem and the vehicle routing problem. Mathematically, these problems are usually modelled as a combinatorial optimization problem. Given that both these problems individually are NP-hard problems, the LRP is also an NP-hard problem to solve. The extant literature shows the application of the LRP in solving real-world problems. Wasner and Zäpfel (2004) applied the LRP to a parcel delivery service provider in Austria, where the vehicles perform both deliveries and pickups to-and-from customer locations. Depot locations and vehicle routes were determined heuristically in this case.

Single-echelon location-routing problems (LRPs) can be further sub-classified into several types, these can be LRP with uncapacitated vehicles, LRP with uncapacitated depots, capacitated LRP, multi-period LPR, or Inventory LRP. While most of the initial research tries to solve for locations with no capacity constraints, the survey by Prodhon and Prins (2014) focuses on research with capacity constraints on locations and vehicles, which they refer to as the Capacitated Location-Routing problem (CLRP). As with other LRPs, CLRP includes two problems that are independently NP-hard, making CLRP an NP-hard
problem. This poses a challenge in implementing the existing models in large-cities with several thousand customers. This challenge has limited the number of publications on an exact solution to the problem, and even the existing publications have tried to solve the problem for up to a few hundred customers only (Prodhon and Prins, 2014).

Drexl and Schneider (2015) further document the variants of the LRP, which include multi-echelon LRPs, LRPs with stochastic data and LRPs with pickups and deliveries. While most of the extant research focuses on deterministic customer demand, Albareda-Sambola, Fernández, and Laporte (2007) have attempted to solve a stochastic location-routing problem where the customer demand is random. In this two-step model, the set of locations and a priori routes are determined without factoring the customer demand. The second step adapts the routes using the customers to be served, which are modeled using a Bernoulli random variable. A lower-bound and heuristic-based hybrid model is used to solve the problem.

3.2 — Two-echelon location-routing problem (2E-LRP)

The two-echelon location-routing problem (2E-LRP) is a subset of the multi-echelon LRP, with two echelons. In the first echelon, one or several hubs serve the satellite facilities in full-truckload vehicles. The routes from the satellite facilities to customer locations represent the second echelon. Jacobsen and Madsen (1980) developed the first 2E-LRP model and used it model newspaper distribution. The newspapers are delivered from the factory to several transshipment points. The transshipment points are chosen from a set of candidate locations. In the second-echelon, the vehicles distribute the newspapers from these transshipment points to the end-customers. The transshipment points, in this case, are uncapacitated.
Crainic, Sforza, and Sterle (2011) propose a two-echelon location routing problem (2E-LRP), which optimizes the number and the location of two different kinds of facilities – platforms and satellites. The platforms are the hub locations which serve as origins while satellites serve as transshipment points where freight coming from platforms is consolidated and placed into smaller vehicles. The problem models the size of two different vehicle fleets and the related routes and is solved using mixed integer linear programming (MILP). However, the scale of the problem is limited to 3 platforms, 25 satellites, and 25 customers.

Lin and Lei (2009) model a two-echelon LRP with uncapacitated satellite depots. Their model includes a set of plants, a set of big customers and a set of small customers. The model attempts to locate uncapacitated satellite depots (referred to as distribution centers), determine the subset of big customers served in the first routing level, and build the routes for both levels. They used a hybrid genetic algorithm embedded with a routing heuristic to solve the 2E-LRP but due to the computational complexity, limited the problem to 131 potential clients. The use of Tabu search heuristics (Boccia et al., 2010) gives similar results on a smaller problem with 200 customers but becomes computationally hard when the size of the problem gets large.

All these variants of the location-routing problem are solved using either mixed integer linear programming (exact method), heuristics, or more advanced genetic algorithm; however, due to the computational complexity of the underlying problem, the number of data points is restricted to a few hundred at max. Therefore, these results cannot directly be scaled to real-world applications to solve for urban last-mile distribution networks, where we have thousands of customers and corresponding data points.
3.3 — Continuum approximation method

Real-world application of the location-routing problem requires solving large-scale optimization problems in finite time, a problem often faced by exact methods. This can be achieved by the use of continuum-approximation (CA) methods which can not only reduce the complexity of the problem but also generate near-optimal solutions (Smilowitz and Daganzo, 2007). CA methods approximate the routing cost of the vehicle based on simple analytical forms for the expected total route based on the area of the service region and density of demand in the area (Daganzo, 1984). Here multiple points of demand are aggregated into service areas and the routing costs are approximated based on the aggregated demand properties in these service areas.

The use of CA-based routing-cost approximation functions allows the model to approximate operational routing decisions sufficiently well while leading to a significant reduction in computation time for the LRP (Nagy and Salhi, 2007). Within an LRP, CA-based approximation functions are used to approximate the vehicle routing component, while the location-allocation decisions are solved using the traditional exact method.

Bruns, Klose, and Stähly (2000) solve the parcel delivery problem in the Swiss postal service to decide on the number, location, size, and service areas of transshipment points to serve customers. The routing cost was estimated using CA-based routing cost approximation function. However, they did not consider return flows in their model.

Winkenbach, Kleindorfer, and Spinler (2016) introduce an augmented routing-cost estimation (ARCE) function based on the analytical route-length estimation proposed by Daganzo (1984). They incorporate the estimation in a two-echelon capacitated location routing problem to design a last-mile parcel distribution system in a dense city environment. They divide the city into multiple rectangular segments, and the demand is assumed to be uniformly distributed within each of these city segments.
The ARCE formula, applied to each segment, incorporates a maximum service time constraint, vehicle capacity constraints, vehicle access, and positioning restrictions, and routing with simultaneous pickups and deliveries by a heterogeneous fleet of vehicles. This approximation method significantly reduces the total computational time while still maintaining the quality and precision of the solution. Additionally, they integrate the ARCE function into a 2E-CLRP to design a last-mile distribution system for parcel delivery.

Merchán and Winkenbach (2018) use the ARCE to develop a transferable method to increase the resolution of tour distance approximations in large-scale urban logistics applications (16,000 customer locations). They leverage large traffic datasets and prove that network circuity does not significantly impact the decisions regarding the location of facilities and definition of service area, but can have significant impact in quantifying the size of the vehicle fleet. Snoeck, Winkenbach, and Mascarino (2018) apply a CA-based segmentation approach to model a mixed integer linear programming model with stochastic demand, which lowers the total cost of operations.

All the applications of the continuum approximation function demonstrate that it not only makes the problem scalable but also generates near optimal results. Additionally, in the retail business-to-consumer (B2C) environment, as the customer delivery location changes daily, the exact delivery location is not known a priori. In such a case, using exact method for vehicle routing is bound to give inaccurate results.

3.4 — Consumer behavior and adoption of collection-and-delivery points (CDPs)

Collection-and-delivery points (CDPs) are facilities to which a carrier may deliver parcels for later independent pickup by the customer. Morganti, Dablanc, and Fortin (2014) have shown the benefits of automated CDPs in reducing the fragmentation of deliveries and thus reducing traffic congestion and
environmental pollution. This benefit translates to an increase in successful first-time deliveries while lowering operational costs. Morganti, Seidel, Blanquart, Dablanc, and Lenz (2014) present a survey of existing CDP solutions in France and Germany and the socio-demographic contexts in which these are located.

Weltevreden (2008) conducted a consumer survey in The Netherlands to understand consumer behavior around CDP adoption. He concludes that online shoppers are willing to actively use CDPs if they are within a five-minute drive time from their homes. Morganti, Dablanc, et al. (2014) have shown that customers are willing to accept CDPs as alternative locations if they are located approximately 5 km from their homes.

Collins (2015) analyzes the utility that CDPs offer to customers and evaluates the tradeoff between delivery and pickup offerings. He concludes that by changing the location of the CDPs and by offering incentives, customers can choose to use environmentally friendly means to pick products from the CDPs. These means include cycling, walking and even a detour to a CDP in an existing car trip. Additionally, small differences in distances does not affect customer preferences for those driving to the CDPs.

3.5 — Integration of collection-and-delivery points in the LRP

The extant literature on locating CDPs considers both LRPs and stand-alone location-allocation models. Guerrero-Lorente, Ponce-Cueto, and Blanco (2017) propose a mixed-integer linear programming model to integrate both forward and return flows in the consumer e-commerce environment. However, they do not consider routing decisions in the model. Karaoglan, Altiparmak, Kara, and Dengiz (2011) consider simultaneous customer pickup and delivery in the LRP. They model it using an exact vehicle routing formula and solve it using a mixed-integer linear programming formulation. However, given the
computational complexity of the problem, the solution is not scalable, and they restrict the number of customers to a hundred only.

Brown and Guiffrida (2014) propose a CA-based method to model CDPs. They compare carbon emissions, resulting from carrier travel times and distances, in a home delivery scenario against a CDP enabled pickup scenario. However, they do not consider integrating CDPs with upstream distribution facilities.

Savelsbergh and Van Woensel (2016) have discussed the benefits of CDPs in reducing the number of potentially failed deliveries by aggregating demand at the CDP, but do not discuss the benefit of routing failed deliveries to the CDP, which can further reduce the multiple delivery attempts. Aksen and Altinkemer (2008) consider the possibility for customer pick-up at brick-and-mortar locations but do not consider product returns or failed deliveries. McLeod, Cherrett, and Song (2006) analyze the impact of CDPs on both carrier and customer mileage if the CDPs are used as an alternative location for failed home deliveries. They demonstrate that CDPs can help save mileage for customers. They, however, do not consider product returns in this analysis. Ambrosino and Grazia Scutellà (2005) model a two-echelon distribution network with different products. However, they restrict their model to consider only forward flows.

The existing research in this field considers a lot of factors independently but does not offer a comprehensive study on the benefits offered by CDPs. The existing studies do not encompass a network model with CDPs including simultaneous forward and return flows. Additionally, while there is some study on the implications of failed deliveries, it is not done in conjunction with the aggregation of overall demand at the CDPs. Lastly, none of the extant literature models CDPs, with these parameters, as part of a broader network design problem, considering location and routing decisions together. However, as CDPs
aggregate demand and streamline return flows and re-delivery attempts for previously failed home deliveries, this can have implications in the design of upstream transshipment facilities.

3.6 — Literature gap

We reviewed the literature on two-echelon location-routing problems, using both exact method and continuum approximation approaches to estimate route lengths. We further reviewed the literature on the integration of CDPs in the design of last-mile distribution networks and conclude that none of the existing research offers a comprehensive solution to jointly consider strategic decisions related to network design while incorporating all the functions of CDPs.

Existing work on LRPs using CA-based approaches generally assumes that product exchange takes place at the customer location, and do not consider simultaneous deliveries to customers and CDPs in the same tour. This limits the direct application of the ARCE function for our research. Additionally, the literature on CDPs does not consider the benefits of integrating these demand aggregation points in optimizing the design of the overall network. Finally, while the extant literature talks about the benefits of CDPs in aggregating demand for forward flow, it does not consider the impact of CDPs in preventing multiple failed delivery attempts and failed pickup attempts by giving the customers an option to route such parcels to the CDP.

We address these gaps in the literature by proposing a framework that:

- models a two-echelon location-routing problem (2E-LRP) to select sites for SF locations from a set of candidate locations
- considers demand aggregation at the CDP for both forward delivery and return flows
• considers the change in customer demand density (for both forward and return flows) due to the introduction of a CDP in a demand zone

• considers the impact of failed deliveries and failed return pickups on the routing cost, and the aggregation benefit that CDPs offer for these

• integrates a routing cost estimation function to make the model scalable

• designs a comprehensive distribution network including the location for upstream transshipment facilities.

This framework is mathematically modeled using a non-linear optimization model.
Chapter 4: Methodology

As mentioned in the problem setting, we have two different types of points-of-demand to be serviced, the CDPs and the individual customer locations, through three different types of routes – dedicated CDP routes, blended routes, and customer-only routes. We formulate an optimization model with the objective of minimizing the total cost of the network encompassing the three possible routes for the entire city area. This model solves over a discrete solution space. We adapt the method developed by (Winkenbach et al., 2016) and refined by (Merchán and Winkenbach, 2018) to develop continuum approximation (CA) based augmented routing-cost estimation (ARCE) function. This method considers unevenly distributed demand zones and divides them into small rectangular segments with uniformly distributed demand. The adapted formulation calculates the cost for the three different kinds of routes using a combination of explicit cost model and the ARCE function.

In this chapter, we formulate the optimization model to solve the problem as defined in Chapter 2. In Chapter 4.1, we first define the distribution network parameters relevant to our problem setting. We then present the optimization model in Chapter 4.2. In Chapter 4.3, we extend the ARCE function to our problem.

4.1 — Definition of the distribution network

In this section, we formally describe the distribution network relevant to our problem setting. All relevant model parameters are listed in Table 1 to 5. Let $J$ be the set of candidate SFs. Let $I$ be the demand segments such that the demand area is divided into a set of sufficiently small segments $I$. 

31
The general model parameters are listed in Table 1. $r_{ij}$ represents the rectilinear distance between candidate satellite facility $j$ and segment $i$. $d_j$ represents the rectilinear distance between the hub and the candidate SF $j$. Each segment $i \in I$ has an area $A_i$. $T_m$ denotes the maximum service time restriction for delivery in the second-echelon.

The operational parameters are listed in Table 3. $Z_j$ represents the maximum capacity (in terms of the number of items) of an SF $j \in J$. Individual customer demand and CDP demand are served from SFs using second-echelon vehicles. Each segment is served by a unique SF through one or several vehicle routes, depending on the volume of demand items. Each route departs from a single SF and can serve multiple segments. The volumetric carrying capacity of first-echelon vehicles is represented by $\xi^a$ and for second-echelon vehicles, it is represented by $\xi^b$. Movement of vehicles in different parts of the network is described by vehicle speed and average circuity factor.
Circuity captures the effect of road network features on travel directness and is usually measured as the ratio between the network distance and the Euclidean distance (Merchan and Winkenbach, 2018). For a given area, average circuity factors are computed by considering circuity factors between random point-to-point trips within the area. For first-echelon vehicles, operating between the distribution hub and the SFs, \( s^d \) represents the average speed and \( \kappa^d \) the average circuity. For second-echelon vehicles, we distinguish between (i) the line-haul journey between SFs and demand segments, characterized by average speed \( s^{\text{B}_{1}} \) and average circuity \( \kappa^{\text{B}_{1}} \); and (ii) the inter-stop journey within a segment, characterized by average speed \( s^{\text{B}_{3}} \), and average circuity \( \kappa^{\text{B}_{3}} \). To account for differences in the road network configuration for each \( i \in I \) we define \( \kappa^{\text{B}_{1}}_i \) for each segment.

\( t^{i,v} \) represents the variable time for vehicle loading or unloading at the satellite facilities. \( t^{p,v} \) represents the variable time for loading or unloading at the CDP. Both \( t^{i,v} \) and \( t^{p,v} \) are variable with respect to the number of items. \( t^{c,f} \) represents the fixed time for each stop at the customer location, this encompasses a stop for both product delivery or return pickup.

### Table 3: Demand parameters

| \( \gamma^F_{i,R} \) | Initial demand density for forward delivery in segment \( i \) [customers/km\(^2\)] |
| \( \gamma^R_{i,R} \) | Initial demand density for returns in segment \( i \) [customers/km\(^2\)] |
| \( \rho_i \) | Average dropsize in the the segment \( i \) [items/customer] |
| \( \theta^C_{i} \) | Average physical volume per item for deliveries towards customers in segment \( i \) [m\(^3\)/item] |
| \( \tau^F \) | Intercept of the linear decay function for delivery items defining the decline of \( \psi_{di} \) in \( r_{di} \) |
| \( \tau^R \) | Intercept of the linear decay function for return items defining the decline of \( \psi_{di} \) in \( r_{di} \) |
| \( \eta^F \) | Slope of the linear decay function for delivery items defining the decline of \( \psi_{di} \) in \( r_{di} \) |
| \( \eta^R \) | Slope of the linear decay function for return items defining the decline of \( \psi_{di} \) in \( r_{di} \) |
| \( \delta \) | Proportion of returns per delivery |
| \( \lambda^F \) | Probability of failed delivery at customer location |
| \( \lambda^R \) | Probability of failed return pickup at customer location |
| \( \beta \) | Proportion of customers specifying a failed delivery to be routed to a CDP for self pickup |
Table 3 lists the demand parameters. The demand in each segment comprises the demand for product delivery in the forward flow and the demand for product returns in the reverse flow. The demand in each segment is described by four parameters. The initial demand density in a segment $i$ before the placement of a CDP is represented by $y_{iF}^{0}$ for the forward flow and $y_{iR}^{0}$ for the return flow. This is defined as the number of customers in a segment divided by its area. The average drop-size in the segment $i$ is represented by $p_i$. It is defined as the average number of items per customers. The average physical volume per item is represented by $\theta_i^c$. Additionally, the proportion of returns for each delivery is defined by the parameter $\delta$. This establishes the relationship between the initial demand densities $y_{iF}^{0}$ and $y_{iR}^{0}$.

The probability of failed delivery at the customer location is represented by $\lambda_i^f$, similarly, the probability of failed return pickup at the customer location is represented by $\lambda_i^r$. The proportion of customers specifying a failed delivery to be routed to a CDP for a self-pickup later is represented by $\beta$.

Let $D \subset i$ be the subset of segments where uncapacitated CDPs can be placed. When a CDP is placed in a segment, it attracts customer demand from the neighboring segments. The number of items attracted to a CDP in segment $i$ from a neighboring segment $k$ decreases linearly in distance between segments $r_{ik}$. We model this linear decay functions with the help of parameters $\tau_i^f$, $\tau_i^r$, $\eta_i^f$, and $\eta_i^r$. $\tau_i^f$ and $\tau_i^r$ represent the intercepts of the linear decay function for the forward and return flows respectively. They define the maximum share of items that can be attracted from a segment $k$ to a CDP located in $i$. $\eta_i^f$ and $\eta_i^r$ represent the slope of the linear decay function for the forward and return flows respectively. They define the rate at which the share of attracted items is decreasing with $r_{ik}$. We assume that customers always chose the CDP closest to their location.
Table 4: Cost parameters

| $c_{j,F}$ | fixed cost to enable a SF $j$ | [$] |
| $c_{j,v}$ | per-item handling cost at SF $j$ | [$/item] |
| $c_{i,F}$ | fixed cost to enable a CDP in segment $i$ | [$] |
| $c_{i,v}$ | per-item cost for a CDP in segment $i$ | [$/item] |
| $c_{i,h}$ | first-echelon transportation cost per time-unit to SF $j$ | [$/hour] |
| $c_{i,d}$ | fixed cost per delivery at customers location | [$/delivery] |
| $w$ | wage cost per hour | [$/hour] |

Table 4 lists the cost parameters. For every $j \in J$, $c_{j,F}$ as the fixed cost to enable a CDP and $c_{j,v}$ as the variable cost per-item at the SF. For every $i \in D$, we define $c_{i,F}$ as the fixed cost to enable a CDP and $c_{i,v}$ as the variable cost per-item at the CDP. We define the first-echelon transportation cost per unit of time to SF $j$ as $c_{i,h}$. The fixed cost per delivery at the customer location is represented as $c_{i,d}$. The wage cost per hour is represented as $w$.

Table 5: Additional notations expressing demand parameters due to CDPs

| $\psi_{ik}^f(w)$ | number of items attracted to a CDP located at a segment $i \in D$ from segment $k \in I$ | [items] |
| $\psi_{ik}^r(w)$ | number of return items attracted to a CDP located at segment $i \in D$ from segment $k \in I$ | [items] |
| $\gamma_i^f(w)$ | adjusted customer density for delivery in the segment $i$ | [customers/km2] |
| $\gamma_i^r(w)$ | adjusted customer density for returns in the segment $i$ | [customers/km2] |
| $\theta_i^f(w)$ | average physical volume per item for requests at CDP located in segment $i$ | [m3] |

In Table 5, we list the additional notations expressing demand parameters under CDP configuration $w$. $\psi_{ik}^f(w)$ represents the number of delivery items attracted from segment $k \in I$ to a CDP located at a segment $i \in D$. Similarly, $\psi_{ik}^r(w)$ represents the number of return items attracted to a CDP located at a segment $i \in D$. $\gamma_i^f(w)$ and $\gamma_i^r(w)$ represent the adjusted customer density for forward delivery and return items, respectively, in segment $i$. $\theta_i^f(w)$ represents the average physical volume per-item for items at the CDP.
4.2 — The optimization model

To design a last-mile distribution network using CDPs, we formulate an optimization model with the objective of minimizing the total cost of the network. All costs are modeled as time-based costs. This model solves over a discrete solution space and considers: (i) satellite facility (SF) location decisions, with vector \( y \) representing a given SF configuration and binary variable \( y, j \in J \) indicating if SF \( j \) is active and (ii) allocation decisions, with vector \( x \) representing a given allocation configuration and binary variable \( x, i \in I, j \in J \) indicating if segment \( i \) is served from SF \( j \). The CDP locations denoted here by vector \( w \) are fixed and introduced as a parameter in the model. The various parameters used in the model are defined in the preceding chapter. The resulting model can be formulated as:

\[
\min_{y, w, x} K(y, w, x) = K^F(y) + K^P(w) + K^T(x, w) + K^R(x, w) 
\]

where:

\[
K^F(y) = \sum_{j \in J} y_j c^F_j + \sum_{j \in J} \sum_{i \in I} x_{ij} \left( \gamma^F_i(w) + \gamma^R_i(w) \right) A_i + \sum_{k \in I} \psi^F_i(w) + \sum_{k \in I} \psi^R_i(w),
\]

\[
K^P(w) = \sum_{i \in I} c_i^P w_i + \sum_{i \in I} c_i^P, \left( \sum_{k \in I} \psi^R_i(w) + \sum_{k \in I} \gamma^F_i(w) \right) A_i + \sum_{k \in I} \psi^R_i(w),
\]

\[
K^T(x, w) = 2 \frac{\sum_{j \in J} x_{ij} f_{ij}(w)}{\sum_{j \in J} x_{ij}},
\]

\[
K^R(x, w) = \sum_{j \in J} \sum_{i \in I} x_{ij} f_{ij}(w),
\]

\[
\psi^F_i(w) = \gamma^0_i A_i \rho_i \max[\tau^F - \eta^F r_{ki}, 0], \quad \forall i \in D, k \in I,
\]

\[
\psi^R_i(w) = \gamma^0_i A_i \rho_i \max[\tau^R - \eta^R r_{ki}, 0], \quad \forall i \in D, k \in I,
\]
\[
\gamma^F_k(w) = \left( \gamma^0_k - \sum_{i \in D} \frac{\psi^F_{ik}(w)}{p_k A_k} \right) (1 + \lambda^F (1 - \beta w)), \quad \forall k \in I, \tag{8}
\]

\[
\gamma^R_k(w) = \left( \gamma^0_k - \sum_{i \in D} \frac{\psi^R_{ik}(w)}{p_k A_k} \right) (1 + \lambda^R), \quad \forall k \in I, \tag{9}
\]

\[
\gamma^0_{k,R} = \delta \gamma^0_{k,D}, \quad \forall k \in I, \tag{10}
\]

\[
\theta^0_i(w) = \frac{\sum_{k \in I} \psi^F_{ik}(w) \theta^C_k}{\sum_{k \in I} \psi^F_{ik}(w)}, \quad \forall i \in D, \tag{11}
\]

subject to

\[
\sum_{j \in J} x_{ij} = 1, \quad \forall i \in I, \tag{12}
\]

\[
\sum_{i \in I} x_{ij} \left( (\gamma^F_i(w) + \gamma^R_i(w)) \rho_i A_i + \sum_{k \in I} \psi^F_{ik}(w) + \sum_{k \in I} \psi^R_{ik}(w) \right) \leq Z_j y_j, \quad \forall j \in J, \tag{13}
\]

\[y_j \in \{0, 1\}, \quad \forall j \in J, \tag{14}\]

\[x_{ij} \in \{0, 1\}, \quad \forall i \in I, j \in J, \tag{15}\]

\[w_i \in \{0, 1\}, \quad \forall i \in D. \tag{16}\]

The objective function is given by Equation (1). \(K^F\) is the cost of operating SFs and its expression is provided by Equation (2). \(K^C\) is the cost of CDP operation and its expression is provided by Equation (3). \(K^T\) is the first-echelon transportation cost and its expression is provided by Equation (4). \(K^D\) is the last-mile distribution cost and its expression is given by Equation (5).

Equations (6) to (11) provide expressions of demand parameters under a given CDP configuration \(w\). Equation (6) provides the number of delivery items attracted to a CDP located at a segment \(i\) from
segment $k$, $\psi_{ik}(w)$, proportional to the number of items initially present in the segment $k$ (i.e., $y_{ik}^0 A_k \rho_k$) and linearly decreasing in the distance $r_k$ between the segment and CDP. Similarly, equation (7) provides the number of return items attracted to a CDP located at a segment $i$ from segment $k$, $\psi_{ik}(w)$, proportional to the number of items initially present in the segment $k$ (i.e., $y_{ik}^0 A_k \rho_k$) and linearly decreasing in the distance $r_k$ between the segment and CDP.

Equation (8) provides the customer density for forward delivery in segment $k$, $y_{ik}^f(w)$, obtained by adjusting the initial customer density $y_{ik}^0$ to account for items attracted to CDPs. Similarly, equation (9) provides the customer density for returns in segment $k$, $y_{ik}^R(w)$, obtained by adjusting the initial customer density $y_{ik}^0$ to account for items attracted to CDPs. Equation (10) establishes the relationship between the initial customer density for forward and return flows. Equation (11) provides the average physical volume per item for requests at a CDP located in segment $i$, $\Theta_f^f(w)$.

In Equation (5), $f_{ij}(w)$ represents the routing cost linked to deliveries to segment $i$ from SF $j$ under a CDP configuration $w$. This cost is computed using expressions presented in Chapter 4.3.

Equations (12) to (16) provide constraints of the optimization model. Constraint (12) signifies that every segment must be allocated to exactly one SF. Constraint (13) enforces capacity limitations at each SF and links location and allocation variables. Constraints (14) to (16) define the domain of the decision variables.

4.3 — The routing cost

The placement of CDPs in a segment changes the demand pattern of the segment. The delivery in each of the segments with CDPs can be performed via three possible routes. Therefore, for $i \in D$, the expression of routing cost includes three components. For other segments, $i \in I \setminus D$, routing cost is
expressed through a single component representing customer-only routes. We formulate the resulting expression of \( f_i(w) \) as:

\[
f_{ij}(w) = \begin{cases} 
  f^D_{ij}(w) + f^B_{ij}(w) + f^C_{ij}(w), & \text{for } i \in D, \forall i \in I, j \in J, \\
  f^C_{ij}(w), & \text{otherwise,}
\end{cases}
\]

In Equation (17), superscripts \( D, B \) and \( C \) refer to the route type (i.e., dedicated, blended and customer-only routes, respectively). The cost of each route type depends on CDP configuration and demand characteristics and can be equal to zero. In the absence of CDPs in a segment, dedicated and blended routes are not deployed, and the associated cost components are equal to zero. Depending on the characteristics of the individual customer and CDP demand, segments with active CDPs can be served through (1) a combination of three route types; (2) a combination of dedicated and customer-only routes; (3) a combination of blended and customer-only routes; and (4) through blended routes only. Analytical expressions of cost of each route type allow establishing the number of routes of each type for given CDP configuration and demand properties. Furthermore, they reflect differences in the organization of the three route types. The cost of dedicated CDP routes, \( f^D_i(w) \), is explicitly modeled as a cost of return trips from SFs to CDPs.

As shown in Equation (18), cost of blended routes, \( f^B_i(w) \), incorporates two subcomponents: (i) \( f^B_{ij}^{BP}(w) \) associated to time required for loading and delivering CDP items, and (ii) \( f^B_{ij}^{BT}(w) \) associated to regular delivery, i.e., the realization of a vehicle tour visiting individual customers and the CDP. The first subcomponent is explicitly modeled. The second cost component is computed through the application of augmented routing cost estimation (ARCE) with adapted parameters. The cost of customer-only routes,
\( f_i^c(w) \), is computed as a cost of regular delivery through the application of ARCE formulae with adapted parameters.

### 4.3.1 — Cost of dedicated CDP routes

The cost of dedicated CDP routes is given by:

\[
\begin{align*}
    f_{ij}^D(w) &= wc_{ij}^D(w) \left( \mu_{ij}^{D,F}(w) \left( t_{L,v} + t_{P,v} \right) + \frac{2r_{ij}R_{\beta,l}}{s^{\beta,l}} + \mu_{ij}^{D,R}(w) \left( t_{P,v} + t_{L,v} \right) \right), \quad \forall i \in D, j \in J, \quad (19)
\end{align*}
\]

\[
\mu_{ij}^{D,\tau}(w) = \frac{\xi^{\beta}}{\theta_{ij}^D(w)}, \quad \forall i \in D., \quad (20)
\]

\[
c_{ij}^D(w) = \left[ \frac{\sum_{k \in D} \psi_{ik}^{D,F}(w)}{\mu_{ij}^{D,F}(w)} \right], \quad \forall i \in D, j \in J, \quad (21)
\]

\[
\mu_{ij}^{D,R}(w) = \min \left[ \frac{\sum_{k \in D} \psi_{ik}^{D,R}(w)}{c_{ij}^D(w), \mu_{ij}^{D,F}(w)} \right], \quad \forall i \in D, j \in J. \quad (22)
\]

The cost of dedicated routes in Equation (19) is formulated as a time-based cost, with \( \omega \) representing the wage cost, \( c_{ij}^D(w) \) the number of routes and the term in parentheses the duration of one route. The loading and unloading time at the SF and CDP is modeled as a function of the number of items. The term \( \mu_{ij}^{D,F}(w) \) represents the number of items per route. The product of the number of items per route and the variable time for loading the vehicle at SF gives the time at the SF. The second term represents the duration of line-haul travel between the SF and the CDP. The third term represents the service time for unloading the items at the CDP. The fourth term represents the time for loading the return items at the CDP. The fifth term represents the variable time for unloading the returned items at the SF.
The values of \( c_i^{D}(w) \) and \( \mu_i^{DF}(w) \) are obtained through Equations (20) to (21), accounting for vehicle capacity, \( \xi^D \). Equation (20) provides the number of items that could be delivered per route while considering vehicle capacity limitation only. Here we assume that the average physical volume per item for returned items is the same as the average physical volume for the items in the forward flow. Additionally, we assume that only vehicle capacity constrains the CDP deliveries and the maximum service time is enough to not impact the number of items per route.

To establish the number of dedicated routes required to service segment \( c_i^D(w) \), we consider the total number of items at the CDP and the previously established number of items per route (Equation 21). In this Equation, a floor function is applied to consider only ‘full’ dedicated CDP routes. Consequently, if the number of items at the CDP is smaller than \( \mu_i^{DF}(w) \), the number of dedicated routes \( c_i^D(w) \) is equal to zero.

Equation (22) represents the number of items returned by customers at the CDP that will be shipped back to the SF. As the number of items for forward flow will always be greater than the return flow, the total number of items per return route is calculated based on the number of routes for forward flow. While calculating the return flow, we can either ship full truckload quantities in some shipments and empty trucks in the remaining ones or ship partial truckloads in all return shipments. The number of items per return route is given by equation 22.

4.3.2 — Cost of delivering CDP volumes on blended routes

Once the dedicated routes towards the CDP in segment \( i \) have been formed, we can express \( \mu_i^{BF}(w) \), the number of remaining items for forward flow and \( \mu_i^{BR}(w) \) the number of return items that will be serviced at the CDP through the blended routes.
\[ \mu_{ij}^{B,F}(w) = \sum_{k \in I} \psi_{ik}^{D,F}(w) - c_{ij}^{D}(w) \mu_{ij}^{D,F}(w), \quad \forall i \in D, j \in J, \quad (23) \]

\[ \mu_{ij}^{B,R}(w) = \sum_{k \in I} \psi_{ik}^{D,R}(w) - c_{ij}^{D}(w) \mu_{ij}^{D,R}(w), \quad \forall i \in D, j \in J, \quad (24) \]

This number corresponds to the total number of items at the CDP, decreased by items delivered through dedicated routes, \( c_i^D(w), \mu_i^{D,F}(w) \) and \( c_j^D(w), \mu_j^{D,R}(w) \) respectively for forward and return flow. The cost of the CDP component of the blended routes, \( f_i^{BP}(w) \) is explicitly modeled as

\[ f_{ij}^{BP}(w) = w \left( \mu_{ij}^{B,F}(w) (t_{L,w} + t_{P,w}) + \mu_{ij}^{B,R}(w) (t_{P,w} + t_{L,w}) \right), \quad \forall i \in D, j \in J. \quad (25) \]

The cost is expressed as a time-based cost, which is the product of the variable time per item for each loading/unloading action, the number of items, and the wage rate per-hour. Here, the first term is the sum of the time to load items at an SF and the time to deliver/unload items at the CDP. The second term corresponds to the loading and unloading of return items at the CDP and SF respectively. The time at both the SF and CDP is modeled as a function of the variable time per item. At the customer location, the time is modeled as a fixed time per delivery, as the number of items delivered per customer is a small number and the time does not change based on the number of items. This is considered in the computation of the second cost sub-component of blended routes, \( f_i^{BT}(w) \) which is described in the next section.

4.3.3 — Regular cost of delivery on blended and customer only routes

Both blended and customer-only routes involve consolidated vehicle tours visiting multiple points of demand (POD). The cost of those tours, \( f_i^{BT}(w) \) and \( f_i^{C}(w) \), respectively, is estimated through the application of ARCE formulae introduced by Winkenbach et al. (2016), with adjusted parameters.
Winkenbach et al. (2016) consider only routes servicing individual customers with static demand density. However, in the current problem setting, the demand density changes according to the CDPs configuration. Furthermore, the type of route (i.e., blended or customer-only) impacts parameters of different formulae. We, therefore, use the notation $f^g_{ij}(w)$ to refer to the cost of regular delivery routes departing from SF $j$ and servicing segment $i$. Superscript $X \in \{B, C\}$ refers to a type of regular delivery route. Equations (26) to (33) provide a generalization of ARCE formulae proposed by Winkenbach et al. (2016).

\[
f^X_{ij}(w) = wc^X_{ij}(w)\left(\frac{2r_{ij}}{s_{ij}^2} + n^X_{ij}(w)t^{C,J} + (n^X_{ij}(w) + \pi^X(w))\frac{\kappa^X_{ij} \beta^X_{ij}}{s_{ij}^2 \sqrt{(\gamma^F_{ij}(w) + \gamma^R_{ij}(w))A_i + \pi^X(w)}}ight) + \frac{n^X_{ij}(w)}{1 + \delta}(\lambda^{F\beta\beta}w_i)t^{P,v}, \quad \forall i \in S^X, j \in J \tag{26}
\]

where

\[
\zeta^X_{ij}(w) = \frac{\xi^X_{ij}(w)}{\theta_i^C \rho_i}, \quad \forall i \in S^X, \tag{27}
\]

\[
T^{X,J}_{ij} = \frac{2r_{ij}}{s_{ij}^2}, \quad \forall i \in S^X, j \in J, \tag{28}
\]

\[
T^{X,v}_{i} = t^{C,J} + \frac{\kappa_i \kappa^X_{ij}}{s_{ij}^2 \sqrt{(\gamma^F_{ij}(w) + \gamma^R_{ij}(w))A_i + \pi^X(w)}} + \frac{(\lambda^{F\beta\beta}w_i)t^{P,v}}{1 + \delta}, \quad \forall i \in S^X, \tag{29}
\]

\[
n^X_{ij}(w) = \begin{cases} 
\zeta^X_{ij}(w), & \text{for } T^{m,X}_{ij}(w) \geq T^{X,J}_{ij} + \zeta^X_{ij}(w)T^{X,v}_{i}(w), \\
T^{m,X}_{ij}(w) - T^{X,J}_{ij} & T^{X,v}_{i}(w)), & \text{otherwise,}
\end{cases} \quad \forall i \in S^X, j \in J, \tag{30}
\]

\[
m^X_{ij}(w) = \begin{cases} 
T^{m,X}_{ij}(w) & T^{m,X}_{ij}(w) \geq T^{X,F}_{ij}(w) + \zeta^X_{ij}(w)T^{X,v}_{i}(w), \\
1, & \text{otherwise,}
\end{cases} \quad \forall i \in S^X, j \in J, \tag{31}
\]
Equation (26) provides the total cost of a route, which is expressed as a time-based cost. The term $c_{ij}^X(w)$ corresponds to the number of routes for a regular delivery route of type X. The term in parentheses corresponds to the duration of a route. Here, $n_{ij}^X(w)$ is the total number of individual customers for both delivery and return pickups, and $\pi^X(w)$ the number of non-customer stops. Consequently, the first term in parentheses is the duration of line-haul travel between the SF and the segment. The second term is the service time at the customer location. The third term is the inter-stop travel duration within a segment, with $n_{ij}^X(w) + \pi^X(w)$ representing the total number of stops. For establishing this duration, the formulation by Winkenbach et al. (2016) is generalized to allow for non-customer stops, without affecting the validity of the approximation. The last term allows the estimation of the time to unload the failed delivery items at the CDP. The ratio $n_{ij}^X(w)/(1+\delta)$ denotes the number of customer deliveries, the product $\zeta_j^{F,X}(w)$ denotes the proportion of failed deliveries which are routed to the CDP, if it exists, and $t^{P_X}$ is the time to unload each item at the CDP. In this equation, SX refers to the set of segments where a route of type X is performed.

Equations (27) to (33) estimate the number of routes $c_{ij}^X(w)$ and the number of forward and return customers per route, $n_{ij}^X(w)$ while accounting for available vehicle capacity $\zeta_j^{P,X}(w)$ and maximum service time $T_j^{m,X}(w)$. Equation (27) provides the number of customers that can be served on a route while considering available vehicle capacity only. Note that in calculating the available vehicle capacity we have not factored in return flow as we are assuming that when the vehicle departs for delivery, it does not require additional space for pickup items. Instead, the space created after the delivery of outbound items will be utilized for return pickups, $\zeta_j^P(w)$. 

\[
q_{ij}^X(w) = \frac{\nu_{ij}^X(w)}{n_{ij}^X(w)m_{ij}^X(w)}, \quad \forall i \in S^X, j \in J, (32)
\]

\[
c_{ij}^X(w) = m_{ij}^X(w)q_{ij}^X(w), \quad \forall i \in S^X, j \in J. (33)
\]
Equations (28) to (30) adjust $\xi(w)$ by considering $T_{ij}^{nX}(w)$. First, we define a fixed duration of a route, $T_{ij}^{X}(w)$ (Equation (28)). Second, we define a variable duration per customer, $T_{ij}^{Xv}(w)$ (Equation (29)). Third, Equation (30) compares $T_{ij}^{nX}(w)$ with the total duration of a route servicing $\xi(w)$ customers. If the total duration of a route respects maximum service time limitation, the total number of customers effectively served through a route, $n_{ij}^{X}(w)$ is equal to previously established $\xi(w)$. In the other case, $n_{ij}^{X}(w)$ is obtained by estimating the number of customers that can be serviced within $T_{ij}^{nX}(w)$ while accounting for $T_{ij}^{X}(w)$ and $T_{ij}^{Xv}(w)$.

Equation (31) provides the number of tours that a single vehicle can perform within $T_{ij}^{nX}(w)$. Equation (32) provides the fleet size, $q_{ij}^{X}(w)$, or the total required number of vehicles. In this equation, $\nu_{ij}^{X}(w)$ represents the number of individual customers served with tours of type $X$. Equation (33) provides the total number of vehicle tours for a route of type $X$, $c_{ij}^{X}(w)$.

**4.3.3.1 — Expression for blended routes**

In case of blended routes ($X = B^T$), the regular cost of delivery, given by Equation (26), only applies to all $i \in D$, so that $S^{B^T} = D$. In this case, we can further write,

\[
\pi_{ij}^{B_i^T}(w) = 1, \quad (34)
\]

\[
\xi_{ij}^{B_i^T}(w) = \xi - \mu_{ij}^{B,F}(w) \theta_{ij}^{D}(w), \quad \forall i \in D, j \in J, \quad (35)
\]

\[
T_{ij}^{m,B_i^T}(w) = T_{ij}^{m} - \left( \mu_{ij}^{B,F}(w) (t_{LV} + t_{P,v}) + \mu_{ij}^{B,R}(w) (t_{P,v} + t_{LV}) \right), \quad \forall i \in D, j \in J, \quad (36)
\]

\[
\nu_{ij}^{B_i^T} = \min \left[ \left( \gamma_{ij}^{F}(w) + \gamma_{ij}^{R}(w) \right) A_{ij}, \left( \nu_{ij}^{B_i^T,F}(w) \right) \right], \quad \forall i \in D, j \in J. \quad (37)
\]
Equation (34) indicates that an additional stop is performed at a CDP for blended routes. Equation (35) gives the adjusted available capacity. This is obtained by subtracting the volume occupied by blended CDP items from the original vehicle capacity. We do not account for the volume occupied by returned items as they are placed after the items for delivery are unloaded and thus do not limit the space in the vehicle.

Equation (36) indicates that the maximum service time is adjusted to account for the duration of loading and delivery of CDP items. Equation (37) gives the customers served on the route. This expression tells us that the maximum number of customers served through regular deliveries for blended routes is equal to the number of customers that can be served by a single route, \( n^{B,T}(w) \). This ensures that a maximum of one blended route is deployed per segment.

4.3.3.2 — Expression for customer-only routes

Similarly, for customer-only routes \((X = C, S^X = I)\), we can write:

\[
\pi^{X=C}(w) = w_i, \quad (38)
\]

\[
\xi^{B,X=C}(w) = \zeta^3, \quad \forall i \in I, j \in J, \quad (39)
\]

\[
T^{m,X=C}(w) = T^m, \quad \forall i \in I, j \in J, \quad (40)
\]

\[
\nu^{X=C}_{ij} = max [\left( \gamma_i^F(w) + \gamma_i^R(w) \right) - c_{ij}^{X=B^T}(w) n_{ij}^{X=B^T,F}(w), 0], \quad \forall i \in I, j \in J. \quad (41)
\]

Equation (38) indicates that non-customer routes = 0 if there is no CDP in the zone or = 1 if there is a CDP in zone. This 1 non-customer stop corresponds to the stop made at the CDP to drop the failed
delivery items. Equations (39) and (40) indicate that the available vehicle capacity and maximum service time are equal to their initial values. Equation (41) expresses that the number of customers served through customer-only routes is obtained by subtracting the customers already served through blended routes from the total number of individual customers in the segments.
Chapter 5: Results and Analysis

In this chapter, we apply the model formulated in Chapter 4, on the distribution network of a leading Brazilian e-commerce retailer which operates in the Sao Paulo metropolitan region. The model is implemented in Python and uses Gurobi Optimizer as the solver engine.

5.1 — Model parameters

In this section, we define the various demand, operational and cost parameters used in our case study.

5.1.1 — Demand data and operational parameters

The retailer has around 15,500 daily customers in the Sao Paulo region, spread across an area of 2,400 km². We divide the city area into square segments i of size \( A_i = 1 \) km²; \( \forall i \in I \), resulting in 2,400 segments. This translates to an approximate customer density of 6.5 customer/km². Figure 2 illustrates the distribution of the demand density in the service area and the location of the hub and SFs. The darker colors represent a higher demand density, with the maximum value of 93 customer/km². The company has a two-echelon distribution network, with a hub that has a capacity of 20,000 daily orders and is located in the outskirts of the city. The hub services the 5 candidate SF locations, each with a capacity of 7,500 daily orders. The hub does not service the customer demand directly. There are 85 potential CDP locations available for operations. In the first echelon, we have a fleet of trucks, with a volumetric capacity of 23 m³ each, servicing the SFs. In the second-echelon we have a fleet of minivans, with a volumetric capacity of 2.4 m³ each, servicing the customers and the CDPs. Figure 3 illustrates the location of the hub, SFs, and CDPs in the city segment.
Figure 2: Demand density in the Metropolitan Sao Paulo region
5.1.2 — Cost parameters

The operating cost of the retailer is confidential and not publicly available. Thus, to maintain confidentiality around the operating cost, we have indexed all costs to the baseline cost calculated in the next section. This helps us objectively compare various scenarios against a reference number, without revealing the true cost of operations. We would refer to the total distribution costs in the baseline scenario as 100. The fixed cost of operating the hub and each SF is obtained from the retailer and indexed to the total cost of distribution.

The cost of operating the CDP is broken down into a fixed cost per CDP and a variable cost per item. The variable time for handling each item at the CDP is the same as the variable time for handling
each item at the SF. The fixed cost of operating the CDP is assumed to be in proportion to the fixed cost of operating the SF. The variable cost of handling an item at the CDP is the same as the variable cost of handling an item at the SF. The time to load or unload an item at the CDP is the same as the time to load or unload an item at the SF.

5.2 — Definition of scenarios

We compare multiple scenarios to analyze the differences in the total cost under various settings. We consider four broad scenarios: (i) Forward flow only; (ii) Forward flow with failed deliveries; (iii) Forward and return flow without failed deliveries; and (iv) Forward and return flow with failed deliveries and failed return pickups. For each of these scenarios, we first consider a network without CDPs and then a network with CDPs. This helps us in evaluating the cost savings due to the introduction of CDPs in the network. Additionally, within each scenario, we consider different parameter values. This further helps us understand the impact of each of these parameters on the total cost.

5.2.1 — Scenarios without CDPs

We first discuss the various scenarios without considering CDPs in the network. The various scenarios are listed in Table 6.
Table 6: Parameter values for scenarios without CDPs

<table>
<thead>
<tr>
<th>Definition</th>
<th>Scenario #</th>
<th>Forward Demand (as % of original demand)</th>
<th>Return Demand (as % of original demand)</th>
<th>Percentage of Failed Deliveries</th>
<th>Percentage of Failed Pickups</th>
</tr>
</thead>
<tbody>
<tr>
<td>Forward flow only</td>
<td>1.A</td>
<td>100%</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Forward flow with failed deliveries</td>
<td>1.B.a</td>
<td>100%</td>
<td>5%</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>1.B.b</td>
<td>100%</td>
<td>10%</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Forward flow + Return flow. No failed deliveries</td>
<td>1.C.a</td>
<td>85%</td>
<td>15%</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>1.C.b</td>
<td>100%</td>
<td>15%</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Forward flow + Return flow + Failed deliveries + Failed pickups</td>
<td>1.D.a</td>
<td>85%</td>
<td>15%</td>
<td>5%</td>
<td>5%</td>
</tr>
<tr>
<td></td>
<td>1.D.b</td>
<td>85%</td>
<td>15%</td>
<td>5%</td>
<td>10%</td>
</tr>
<tr>
<td></td>
<td>1.D.c</td>
<td>85%</td>
<td>15%</td>
<td>10%</td>
<td>10%</td>
</tr>
<tr>
<td></td>
<td>1.D.d</td>
<td>85%</td>
<td>15%</td>
<td>10%</td>
<td>20%</td>
</tr>
</tbody>
</table>

The first scenario with forward flow only (Scenario 1.A) is considered the baseline scenario. This has all flow directed as customer deliveries and does not consider failed deliveries.

Next, we consider the impact of failed deliveries. Based on estimates by Post and Parcel (2018) approximately 5% of all e-commerce deliveries fail the first time. We use this as the base case for failed deliveries (Scenario 1.B.a) and additionally consider another scenario with a 10% failed delivery rate (Scenario 1.B.b).

The e-commerce industry (Shopify, 2019) has an average product return rate of 15%. We consider this as the percentage to analyze the impact of product returns on the distribution costs. Additionally, we model returns in two different ways. In the first, we assume that the overall demand does not change and is redistributed between forward and return shipments. This means, in cases with the product return rate of 15% (Scenario 1.C.a), the original demand of 100 is adjusted to 85 units for forward flow and 15 units...
for return flow. In the second case (Scenario 1.C.b), we keep the original forward demand intact as 100 and add the return demand of 15 units on top of that. This brings the overall demand to 115 units.

In the last set of scenarios, we consider the consolidated impact of forward flow, failed deliveries, return flow and failed pickups. This is modeled from Scenario 1.D.a to 1.D.d, where we vary the percentage of failed deliveries and failed pickups to consider four different combinations. Scenario 1.D.b and 1.D.d have a higher percentage of failed pickups than that of failed deliveries because a return pickup from a customer's home requires the presence of a person to make that return, which is often not the requirement for home delivery. This additional constraint can result in more failures and hence has been analyzed separately.

5.2.2 — Scenarios with CDPs

In the second set of scenarios, we study the impact due to the introduction of CDPs. The demand attracted to the CDPs is defined by a linear decay function as specified in Equation 6 and Equation 7 in Chapter 4. As discussed in Chapter 3.4, Morganti, Dablanc, et al. (2014) have shown that customers are willing to accept CDPs as alternative locations if they are located approximately 5 km from their homes. Thus, we consider two different scenarios of CDPs. In the first, we consider that the CDPs attract 10% of the total demand. This is modeled with $\tau = 0.2$ and $\eta = 0.045$, which means that the customers within a 4.5 km radius would be interested in routing their parcels to the CDP. In the second scenario, we consider that the CDPs attract 20% of the total demand. This is modeled with $\tau = 0.3$ and $\eta = 0.04$, which means that customers within a 7.5 km radius of a CDP would be interested in routing their parcel to the CDP. We have a network of 85 CDPs, whose location is pre-selected by the retailer. These locations are based on demand levels within the various demand zones and on the availability of a facility location to station a CDP.
Table 7: Parameter values for scenarios with CDPs

<table>
<thead>
<tr>
<th>Definition</th>
<th>Scenario #</th>
<th>Forward Demand (as % of original demand)</th>
<th>Return Demand (as % of original demand)</th>
<th>Percentage of Failed Deliveries</th>
<th>Percent age of Failed Pickups</th>
<th>Demand attracted by CDP</th>
<th>Percentage of customers choosing a CDP as an alternate location for failed deliveries</th>
</tr>
</thead>
<tbody>
<tr>
<td>Forward flow only</td>
<td>2.A.a</td>
<td>100%</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>10%</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>2.A.b</td>
<td>100%</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>20%</td>
<td>-</td>
</tr>
<tr>
<td>Forward flow with failed deliveries</td>
<td>2.B.a</td>
<td>100%</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>20%</td>
<td>40%</td>
</tr>
<tr>
<td></td>
<td>2.B.b</td>
<td>100%</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>20%</td>
<td>80%</td>
</tr>
<tr>
<td></td>
<td>2.B.c</td>
<td>100%</td>
<td>-</td>
<td>-</td>
<td>10%</td>
<td>20%</td>
<td>40%</td>
</tr>
<tr>
<td></td>
<td>2.B.d</td>
<td>100%</td>
<td>-</td>
<td>-</td>
<td>10%</td>
<td>20%</td>
<td>80%</td>
</tr>
<tr>
<td>Forward flow + Return flow. No failed deliveries</td>
<td>2.C.a</td>
<td>85%</td>
<td>15%</td>
<td>-</td>
<td>-</td>
<td>20%</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>2.C.b</td>
<td>100%</td>
<td>15%</td>
<td>-</td>
<td>-</td>
<td>20%</td>
<td>-</td>
</tr>
<tr>
<td>Forward flow + Return flow + Failed deliveries + Failed pickups</td>
<td>2.D.a</td>
<td>85%</td>
<td>15%</td>
<td>5%</td>
<td>5%</td>
<td>20%</td>
<td>40%</td>
</tr>
<tr>
<td></td>
<td>2.D.b</td>
<td>85%</td>
<td>15%</td>
<td>5%</td>
<td>10%</td>
<td>20%</td>
<td>40%</td>
</tr>
<tr>
<td></td>
<td>2.D.c</td>
<td>85%</td>
<td>15%</td>
<td>10%</td>
<td>10%</td>
<td>20%</td>
<td>40%</td>
</tr>
<tr>
<td></td>
<td>2.D.d</td>
<td>85%</td>
<td>15%</td>
<td>10%</td>
<td>20%</td>
<td>20%</td>
<td>40%</td>
</tr>
</tbody>
</table>

All the scenarios considered in Table 7 are similar to those without CDPs (Table 6). Scenario 2.A.a considers forward flow with 10% demand attracted to CDPs, while Scenario 2.A.b considers forward flow with 20% demand attracted to CDPs. In all subsequent scenarios, we use 20% demand attracted as the base case while we vary the other parameters.

As discussed in Chapter 1.4, in cases where there are failed deliveries, CDPs offer customers the option to select the nearest CDP as an alternative location to temporarily hold the parcel. The customer can then collect the parcel from the CDP at a convenient time. This is typically done in the case of high-value items, where customers don't want the product to be left at the doorstep and are not sure if they
would be home to receive the product in-person during a subsequent delivery attempt. In the next set of scenarios, we consider the impact of failed delivery on the distribution cost. Scenario 2.B.a considers a failed delivery rate of 5%. We also introduce an additional parameter which decides the customers’ response in case of a failed delivery. In this scenario, 40% of the customers opt for the product to be routed to and held at the CDP and, correspondingly, 60% would expect the delivery provider to make another attempt to deliver on another day. In Scenario 2.B.b we increase the proportion of customers choosing failed deliveries to be routed to the CDPs to 80%. In Scenario 2.B.c and 2.B.d, we repeat the above two scenarios with a failed delivery rate of 10% to compare it with the “No CDP” scenario (1.B.b) discussed in Chapter 5.2.1.

Next, we consider the impact of return flows in a network with CDPs. In Scenario 2.C.a, we keep the total demand intact at 100% and reduce the demand in the forward flow to 85% to accommodate the return demand of 15%. In Scenario 2.C.b we keep the forward flow demand at 100% and add the return flow demand of 15%, thus increasing the overall flow in the demand zone to 115% of the original value.

In the last set of scenarios, we consider the consolidated impact of forward flow, failed deliveries, return flow and failed pickups. This is modeled in Scenario 2.D.a to 2.D.d, where we vary the percentage of failed deliveries and failed pickups to consider 4 different combinations, which are similar to the four cases considered in Scenario 1.D.a to 1.D.d.

5.3 — Analysis of results

In this section, we discuss the results of various scenarios, modeled with different combinations of returns, failed deliveries and failed pickups, both with and without CDPs in the distribution network.
5.3.1 — Analysis of distribution networks without CDPs

The optimal network configuration in the different scenarios is the same. All scenarios result in the same facility locations, with 3 of the 5 SFs getting activated. This can be seen in Figure 4.

Figure 4: Optimal SF configuration in the network without CDPs
Table 8: Total cost of distribution for the network without CDPs

<table>
<thead>
<tr>
<th>Scenario #</th>
<th>Parameters*</th>
<th>Hub + SF Fixed cost</th>
<th>SF Variable cost</th>
<th>First Tier Transport cost</th>
<th>CDP Fixed cost</th>
<th>CDP Variable cost</th>
<th>Last-mile distribution cost</th>
<th>Total cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>Forw ard flow only</td>
<td>[100-0-0-0-0-0]</td>
<td>12.5</td>
<td>3.8</td>
<td>1.1</td>
<td>0</td>
<td>0</td>
<td>82.6</td>
<td>100.0</td>
</tr>
<tr>
<td>Forw ard flow with failed deliveries</td>
<td>[100-0-5-0-0-0]</td>
<td>12.5</td>
<td>4.1</td>
<td>1.1</td>
<td>0</td>
<td>0</td>
<td>85.6</td>
<td>103.3</td>
</tr>
<tr>
<td>1.D.a</td>
<td>[85-15-5-0-0-0]</td>
<td>12.5</td>
<td>3.9</td>
<td>1.1</td>
<td>0</td>
<td>0</td>
<td>85.7</td>
<td>103.3</td>
</tr>
<tr>
<td>1.D.b</td>
<td>[85-15-5-10-0-0]</td>
<td>12.5</td>
<td>3.9</td>
<td>1.1</td>
<td>0</td>
<td>0</td>
<td>86.2</td>
<td>103.7</td>
</tr>
<tr>
<td>1.D.c</td>
<td>[85-15-10-10-0-0]</td>
<td>12.5</td>
<td>4.1</td>
<td>1.1</td>
<td>0</td>
<td>0</td>
<td>88.9</td>
<td>106.6</td>
</tr>
<tr>
<td>1.D.d</td>
<td>[85-15-10-20-0-0]</td>
<td>12.5</td>
<td>4.2</td>
<td>1.1</td>
<td>0</td>
<td>0</td>
<td>89.7</td>
<td>107.5</td>
</tr>
</tbody>
</table>

*Parameters: [Forward Demand – Return Demand – % Failed Deliveries – % Failed Pickups – Demand Attracted by CDP – % Customers choosing a CDP as an alternate location for failed deliveries]

Table 8 gives a breakdown of all the costs incurred in the downstream supply chain. As mentioned earlier, the total distribution costs obtained in Scenario 1.A are denoted as 100 and considered as the baseline scenario. All other costs are indexed to this number. The fixed cost of operating the hub and the SF is 12.5 and remains the same in all scenarios. This is because the network configuration does not change in any of the scenarios. The SF variable cost also increases with a corresponding increase in the last-mile cost. This is due to the higher number of items handled at the SFs due to failed deliveries. However, the change is not significant compared to the overall distribution cost. The first-tier transport cost does not change unless the overall demand changes (Scenario 1.C.b).
The table clearly shows that the last mile is the most expensive leg of the overall distribution. In the baseline scenario (1.A), the last-mile costs 82.6. This cost goes up to 85.6 (Scenario 1.B.a) and 88.8 (Scenario 1.B.b) as we increase the proportion of failed deliveries. This is accompanied by a corresponding increase in the total distribution cost to 103.3 and 106.7, respectively. This clearly shows that failed deliveries can significantly increase the distribution cost for retailers.

In Scenario 1.C.a, with return flows, the last-mile and the total cost are the same as the baseline scenario. This is because the overall demand in the region does not change, and we have assumed the same cost parameters for forward flow and return flow. In Scenario 1.C.b, with return flows, we see a sharp increase in the last-mile cost to 92.0. This change is due to the increase in the overall demand from 100% to 115%, as in this case, we add the return demand over the original forward demand.

Lastly, we combine returns, failed deliveries, and failed pickups. This results in the cost of last-mile distribution rising to 85.7 (Scenario 1.D.a) along with the total cost going up to 103.3. In Scenario 1.D.b, when we change the proportion of failed pickups compared to Scenario 1.D.a, the last-mile cost increases further to 86.2 and the total cost goes up to 103.7. In Scenario 1.D.c and Scenario 1.D.d, the increase in the proportion of failed deliveries to 10% increases the last-mile cost to 88.9 and 89.7, respectively. The total cost of distribution in these two cases subsequently increases to 106.6 and 107.5, respectively.

The various scenarios clearly demonstrate that failed deliveries and failed return pickups increase both the last-mile cost and the overall cost of distribution, even when the overall demand distribution does not change. Thus, failed deliveries and failed pickups can significantly impact the profitability of retailers.
5.3.2 — Analysis of distribution networks with CDPs

As we add CDPs in the network, there is no change in the overall distribution network. We have the same 3 SF locations active in all cases in Scenario 2. This distribution network can be seen in Figure 5.

Figure 5: Optimal SF configuration in the network with CDPs
Table 9: Total cost of distribution for the network with CDPs

<table>
<thead>
<tr>
<th>Scenario #</th>
<th>Parameters*</th>
<th>Hub + SF Fixed cost</th>
<th>SF Variable cost</th>
<th>First Tier Transport cost</th>
<th>CDP Fixed cost</th>
<th>CDP Variable cost</th>
<th>Last-mile distribution cost</th>
<th>Total cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>Forward flow only</td>
<td>2.A.a [100-0-0-0-10-0]</td>
<td>12.5</td>
<td>3.8</td>
<td>1.1</td>
<td>4.6</td>
<td>0.7</td>
<td>76.2</td>
<td>98.9</td>
</tr>
<tr>
<td></td>
<td>2.A.b [100-0-0-0-20-0]</td>
<td>12.5</td>
<td>3.8</td>
<td>1.1</td>
<td>4.6</td>
<td>1.5</td>
<td>71.0</td>
<td>94.5</td>
</tr>
<tr>
<td>Forward flow with failed deliveries</td>
<td>2.B.a [100-0-5-0-20-40]</td>
<td>12.5</td>
<td>3.9</td>
<td>1.1</td>
<td>4.6</td>
<td>1.7</td>
<td>72.6</td>
<td>96.4</td>
</tr>
<tr>
<td></td>
<td>2.B.b [100-0-5-0-20-80]</td>
<td>12.5</td>
<td>3.8</td>
<td>1.1</td>
<td>4.6</td>
<td>1.8</td>
<td>71.6</td>
<td>95.4</td>
</tr>
<tr>
<td></td>
<td>2.B.c [100-0-10-0-20-40]</td>
<td>12.5</td>
<td>3.9</td>
<td>1.1</td>
<td>4.6</td>
<td>1.8</td>
<td>74.2</td>
<td>98.1</td>
</tr>
<tr>
<td></td>
<td>2.B.d [100-0-10-0-20-80]</td>
<td>12.5</td>
<td>3.8</td>
<td>1.1</td>
<td>4.6</td>
<td>1.8</td>
<td>72.1</td>
<td>95.9</td>
</tr>
<tr>
<td>Forward flow + Return flow. No Failed deliveries</td>
<td>2.C.a [85-15-0-0-20-0]</td>
<td>12.5</td>
<td>3.8</td>
<td>1.1</td>
<td>4.6</td>
<td>1.5</td>
<td>71.0</td>
<td>94.5</td>
</tr>
<tr>
<td></td>
<td>2.C.b [100-15-0-0-20-0]</td>
<td>12.5</td>
<td>4.3</td>
<td>1.2</td>
<td>4.6</td>
<td>1.7</td>
<td>79.1</td>
<td>103.4</td>
</tr>
<tr>
<td>Forward flow + Return flow + Failed deliveries + Failed Pickups</td>
<td>2.D.a [85-15-5-5-20-40]</td>
<td>12.5</td>
<td>3.8</td>
<td>1.1</td>
<td>4.6</td>
<td>1.6</td>
<td>72.8</td>
<td>96.4</td>
</tr>
<tr>
<td></td>
<td>2.D.b [85-15-5-10-20-40]</td>
<td>12.5</td>
<td>3.8</td>
<td>1.1</td>
<td>4.6</td>
<td>1.6</td>
<td>73.1</td>
<td>96.7</td>
</tr>
<tr>
<td></td>
<td>2.D.c [85-15-10-10-20-40]</td>
<td>12.5</td>
<td>3.9</td>
<td>1.1</td>
<td>4.6</td>
<td>1.8</td>
<td>74.5</td>
<td>98.4</td>
</tr>
<tr>
<td></td>
<td>2.D.d [85-15-10-20-20-40]</td>
<td>12.5</td>
<td>3.9</td>
<td>1.1</td>
<td>4.6</td>
<td>1.8</td>
<td>75.1</td>
<td>99.0</td>
</tr>
</tbody>
</table>

*Parameters: [Forward Demand – Return Demand – % Failed Deliveries – % Failed Pickups – Demand Attracted by CDP – % Customers choosing a CDP as an alternate location for failed deliveries]

As can be seen from Table 9, there is no change in the cost of first-tier transport or the fixed cost of operating the hub and satellite facilities. The CDPs however, impact the cost of distribution. While we incur a fixed cost in operating the CDPs, it is constant, irrespective of the demand attracted to the CDPs. The handling cost at the CDPs varies with the change in the total forward and return demand they attract. This cost, however, is very small compared to the cost of the last-mile distribution. Additionally, we see
that the last-mile distribution cost in all cases is significantly lower than the corresponding scenarios without CDPs. This also translates to a lower total cost of distribution in all the scenarios discussed.

In Scenario 2.A.a, with forward flow only, where the demand attracted at the CDPs is 10%, the last-mile cost falls to 76.2, compared to 82.6 in Scenario 1.A (Table 8). We discuss the cost savings generated due to CDPs in more detail in Chapter 6. In Scenario 2.A.b, when the demand attracted at the CDPs goes up to 20%, the last-mile distribution cost falls even further, to 71.0. At the same time, the variable cost of operating the CDPs doubles to 1.5. However, as this cost is very small in magnitude compared to the last-mile distribution cost, the total cost of distribution reduces to 94.5.

As we introduce failed deliveries into the system, the last-mile delivery costs go up from 71.0 to 72.6 (Scenario 2.B.a). However, as we double the proportion of customers choosing CDPs as alternative locations to hold the parcels from 40% to 80%, the last-mile distribution cost reduces to 71.6 (Scenario 2.B.b). Next, as we increase the proportion of failed deliveries to 10%, the last-mile distribution cost goes up to 74.2 (Scenario 2.B.c) and 72.1 (Scenario 2.B.d), respectively, for cases with 40% and 80% customers choosing CDPs as alternative delivery locations for failed deliveries. This set of scenarios clearly demonstrates that, while failed deliveries increase the distribution cost, as more customers choose CDPs as alternative pickup locations to hold their products, the overall distribution cost goes down.

In Scenario 2.C.a, with return flows, the last-mile and the total cost are the same as in the Scenario 2.A.b with CDPs. This is because the overall demand in the region does not change, and we have assumed the same cost parameters for forward flow and return flow. In Scenario 2.C.b, with return flows, we see an increase in the last-mile cost to 79.1. This change, however, is due to the increase in the overall demand from 100% to 115%, similar to the case we had observed in Scenario 1.C.b.

In the last set of scenarios, we combine returns, failed deliveries, and failed pickups. This results in the cost of last-mile distribution to increase to 72.8 (Scenario 2.D.a) along with the total cost going up.
to 96.4. In Scenario 2.D.b, when we change the proportion of failed return pickups compared to Scenario 2.D.a, the last-mile cost increases further to 73.1 and the total cost goes up to 96.7. In Scenario 2.D.c and Scenario 2.D.d, the increase in the proportion of failed deliveries to 10% increases the last-mile cost to 74.5 and 75.1, respectively. The total cost of distribution in these two cases subsequently increases to 98.4 and 99.0, respectively. It is interesting to note that in all these scenarios, the costs are not only lower than the corresponding scenarios without CDPs but also lower than the baseline scenario (1.A) where there are no failed deliveries or failed pickups.

These scenarios demonstrate that CDPs not only reduce the total distribution cost due to the aggregation of demand but, they can also help mitigate the impact of an increase in costs due to failed deliveries.
Chapter 6: Discussion

In this chapter, we discuss the implications of our results. First, we discuss the total cost savings in a network with CDPs, then we assess the maximum potential savings that can be generated in a fully utilized network. Lastly, we discuss various governance models for setting up CDPs.

6.1 — Cost savings generated by integrating CDPs in different scenarios

CDPs help reduce both the overall cost of distribution and the last-mile delivery cost. The cost savings in the last-mile is directly related to the overall demand attracted to the CDPs, while the total cost savings is a tradeoff between the increase in the fixed cost of operating the CDPs and the cost savings incurred in the last-mile. As the demand attracted goes up, the cost savings also increase. Table 10 shows a comparison of cost savings generated in a network with CDPs relative to a network without CDPs. This table directly compares each scenario in Table 9 with a corresponding scenario in Table 8.
Table 10: Cost savings in a network with CDPs compared to a network without CDPs

<table>
<thead>
<tr>
<th>Scenario #</th>
<th>Parameters*</th>
<th>Total cost savings</th>
<th>Last-mile cost savings</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Forward flow only</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.A.a</td>
<td>[100-0-0-0-10-0]</td>
<td>1.1%</td>
<td>7.7%</td>
</tr>
<tr>
<td>2.A.b</td>
<td>[100-0-0-0-20-0]</td>
<td>5.5%</td>
<td>14.0%</td>
</tr>
<tr>
<td><strong>Forward flow with failed deliveries</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.B.a</td>
<td>[100-0-5-0-20-40]</td>
<td>6.7%</td>
<td>15.2%</td>
</tr>
<tr>
<td>2.B.b</td>
<td>[100-0-5-0-20-80]</td>
<td>7.6%</td>
<td>16.4%</td>
</tr>
<tr>
<td>2.B.c</td>
<td>[100-0-10-0-20-40]</td>
<td>8.1%</td>
<td>16.4%</td>
</tr>
<tr>
<td>2.B.d</td>
<td>[100-0-10-0-20-80]</td>
<td>10.1%</td>
<td>18.8%</td>
</tr>
<tr>
<td><strong>Forward flow + Return flow. No Failed deliveries</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.C.a</td>
<td>[85-15-0-0-20-0]</td>
<td>5.6%</td>
<td>14.0%</td>
</tr>
<tr>
<td>2.C.b</td>
<td>[100-15-0-0-20-0]</td>
<td>6.1%</td>
<td>14.0%</td>
</tr>
<tr>
<td><strong>Forward flow + Return flow + Failed deliveries + Failed Pickups</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.D.a</td>
<td>[85-15-5-5-20-40]</td>
<td>6.6%</td>
<td>15.1%</td>
</tr>
<tr>
<td>2.D.b</td>
<td>[85-15-5-10-20-40]</td>
<td>6.7%</td>
<td>15.2%</td>
</tr>
<tr>
<td>2.D.c</td>
<td>[85-15-10-10-20-40]</td>
<td>7.7%</td>
<td>16.2%</td>
</tr>
<tr>
<td>2.D.d</td>
<td>[85-15-10-20-20-40]</td>
<td>7.9%</td>
<td>16.3%</td>
</tr>
</tbody>
</table>

*Parameters: [Forward Demand – Return Demand – % Failed Deliveries – % Failed Pickups – Demand Attracted by CDP – % Customers choosing a CDP as an alternate location for failed deliveries]

The CDPs offer several benefits which lead to this cost savings. Firstly, the CDPs help aggregate demand in the forward flow. This is shown in the cost savings in Scenario 2.A.a and Scenario 2.A.b. As the demand attracted goes up from 10% to 20%, the last-mile cost savings double from 7.7% in Scenario 2.A.a to 14.0% to Scenario 2.A.b, and the total cost savings increase from 1% to 5.5%. The total cost savings in Scenario 2.A.a are lower because the fixed cost of operating the CDPs is the same as in Scenario 2.A.b and thus partly offsets the cost savings in the last-mile distribution.
In case of failed deliveries, as CDPs offer an alternative location to route multiple failed deliveries, the cost savings go further up. In Scenario 2.B.a, with 5% failed deliveries and 40% opting for CDPs as alternative delivery locations for failed deliveries, the last-mile cost savings go up to 15.2%. In Scenario 2.B.b, when 80% of the customers opt for CDPs as an alternative delivery location for failed deliveries, the last-mile cost savings increase further, to 16.4%. The corresponding savings in total cost in these two scenarios are 6.7% and 7.6%, respectively. In Scenarios 2.B.c and 2.B.d, we increase the proportion of failed deliveries to 10% for the preceding two scenarios. As a result, the last-mile cost savings come in at 16.4% (Scenario 2.B.c) for 40% of the customers choosing CDPs as alternative delivery locations and 18.8% (Scenario 2.B.d) for 80% of the customers choosing CDPs as alternative delivery locations for failed deliveries. The corresponding savings in total cost in these two scenarios is 8.1% and 10.1%, respectively. This shows that cost savings go up with an increase in both, the number of failed deliveries, and the proportion of customers opting failed deliveries to be routed to the CDPs.

The CDPs offer the benefit of aggregation in the return flow as well. The savings in Scenario 2.C.a is same as those obtained in Scenario 2.A.b as the overall demand remains the same. In Scenario 2.C.b, when the return demand of 15% is added on top of the original forward flow, the total cost savings go up to 6.1%.

When the impact of returns, failed deliveries, and failed return pickups are collectively applied, we achieve the maximum cost savings seen in Scenario 2.D.a to Scenario 2.D.d. In Scenario 2.D.a, with 5% failed deliveries and 5% failed pickups, we achieve total cost savings of 6.6% compared to a network without CDPs. When the number of failed pickups increases to 10% (Scenario 2.D.b), the total cost savings rise to 6.7%. In Scenario 2.D.c, when the number of failed deliveries goes up to 10%, the cost savings jump to 7.7%. While retaining the percentage of failed deliveries at 10%, if the proportion of failed pickups increases to 20%, the cost savings jump to 7.9% (Scenario 2.D.d).
We see that the cost savings due to an increase in failed delivery are significantly larger than an increase in failed returns. This is because the actual number of failed return pickups remain small relative to the number of failed deliveries, as the total returns itself is 15% of the overall demand. Also, customers have an additional option to route failed deliveries to the nearest CDP location, which significantly reduces the cost of multiple delivery attempts.

6.2 — Estimating maximum savings from CDPs

To assess the maximum savings potential from CDPs, we simulate a case with the maximum practical limits of all parameters. This scenario includes 30% of the demand attracted at the CDPs, 20% failed deliveries, 40% failed return pickups, and 80% of the customers choosing CDP as an alternative location to temporarily hold failed deliveries for collection later. This can be denoted as a [85-15-20-40-30-80] scenario using the nomenclature used in Table 8 to 10. This results in a 12% savings in the total distribution costs compared to a similar scenario without CDPs. The higher cost saving in this scenario is due to the cumulative effect of all benefits discussed earlier. This clearly proves that CDPs offer significant cost savings through the aggregation of both the regular demand and the attempts for failed deliveries. Thus, the CDPs offer an attractive option for retailers to save on the overall distribution cost. If the fixed cost of operating the CDPs can be lowered, and if the overall demand routed to the CDPs can be increased, the cost savings can be further increased.

6.3 — Exploring the effect of alternative governance models

In the above analysis, we have assumed that CDPs are owned by the retailer, and thus incur a fixed cost per CDP and a variable cost per item handled. We call this the proprietary ownership model. An
alternative operating model is to outsource the operation of the CDP altogether to a third party. This model of governance will incur no fixed cost in operating the CDPs, however, the variable cost of handling items at the CDP will be significantly higher. In the following discussion, we assess the impact of the two different governance models.

Table 9 discuss the cost incurred in running proprietary CDPs, owned and operated by the retailer. For outsourced CDPs, we assume that there will be no fixed cost charged to the retailer, while the variable cost of handling each item would be as much as four times the cost incurred during in-house handling. Simulating the results with these parameters gives us the costs as incurred in Table 11. Note that all costs other than the CDP fixed cost and CDP variable cost remain the same as in Table 9.

### Table 11: Total cost of distribution for the network with outsourced CDP operations

<table>
<thead>
<tr>
<th>Scenario #</th>
<th>Parameters</th>
<th>Hub Fixed cost</th>
<th>SF Fixed cost</th>
<th>SF Variable cost</th>
<th>First Tier Transport cost</th>
<th>CDP Fixed cost</th>
<th>CDP Variable cost</th>
<th>Last-mile distribution cost</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Forward flow only</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.A.a</td>
<td>[100-0-0-0-10-0]</td>
<td>12.5</td>
<td>3.8</td>
<td>1.1</td>
<td>0.0</td>
<td>2.8</td>
<td></td>
<td>76.2</td>
<td>96.4</td>
</tr>
<tr>
<td>2.A.b</td>
<td>[100-0-0-0-20-0]</td>
<td>12.5</td>
<td>3.8</td>
<td>1.1</td>
<td>0.0</td>
<td>6.0</td>
<td></td>
<td>71.0</td>
<td>94.4</td>
</tr>
<tr>
<td><strong>Forward flow with failed deliveries</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.B.a</td>
<td>[100-0-5-0-20-40]</td>
<td>12.5</td>
<td>3.9</td>
<td>1.1</td>
<td>0.0</td>
<td>6.6</td>
<td></td>
<td>72.6</td>
<td>96.7</td>
</tr>
<tr>
<td>2.B.b</td>
<td>[100-0-5-0-20-80]</td>
<td>12.5</td>
<td>3.8</td>
<td>1.1</td>
<td>0.0</td>
<td>7.2</td>
<td></td>
<td>71.6</td>
<td>96.2</td>
</tr>
<tr>
<td>2.B.c</td>
<td>[100-0-10-0-20-40]</td>
<td>12.5</td>
<td>3.9</td>
<td>1.1</td>
<td>0.0</td>
<td>7.2</td>
<td></td>
<td>74.2</td>
<td>98.9</td>
</tr>
<tr>
<td>2.B.d</td>
<td>[100-0-10-0-20-80]</td>
<td>12.5</td>
<td>3.8</td>
<td>1.1</td>
<td>0.0</td>
<td>7.2</td>
<td></td>
<td>72.1</td>
<td>96.7</td>
</tr>
<tr>
<td><strong>Forward flow + Return flow, No Failed deliveries</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.C.a</td>
<td>[85-15-0-0-20-0]</td>
<td>12.5</td>
<td>3.8</td>
<td>1.1</td>
<td>0.0</td>
<td>6.0</td>
<td></td>
<td>71.0</td>
<td>94.4</td>
</tr>
<tr>
<td>2.C.b</td>
<td>[100-15-0-0-20-0]</td>
<td>12.5</td>
<td>4.3</td>
<td>1.2</td>
<td>0.0</td>
<td>6.8</td>
<td></td>
<td>79.1</td>
<td>103.9</td>
</tr>
<tr>
<td><strong>Forward flow + Return flow + Failed deliveries + Failed Pickups</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.D.a</td>
<td>[85-15-5-5-20-40]</td>
<td>12.5</td>
<td>3.8</td>
<td>1.1</td>
<td>0.0</td>
<td>6.5</td>
<td></td>
<td>72.8</td>
<td>96.7</td>
</tr>
<tr>
<td>2.D.b</td>
<td>[85-15-5-10-20-40]</td>
<td>12.5</td>
<td>3.8</td>
<td>1.1</td>
<td>0.0</td>
<td>6.5</td>
<td></td>
<td>73.1</td>
<td>97.0</td>
</tr>
<tr>
<td>2.D.c</td>
<td>[85-15-10-10-20-40]</td>
<td>12.5</td>
<td>3.9</td>
<td>1.1</td>
<td>0.0</td>
<td>7.0</td>
<td></td>
<td>74.5</td>
<td>99.0</td>
</tr>
<tr>
<td>2.D.d</td>
<td>[85-15-10-20-20-40]</td>
<td>12.5</td>
<td>3.9</td>
<td>1.1</td>
<td>0.0</td>
<td>7.0</td>
<td></td>
<td>75.1</td>
<td>99.6</td>
</tr>
</tbody>
</table>
Table 12 shows the cost savings of operating outsourced CDPs compared to operating proprietary CDPs.

Table 12: Cost savings in the outsourced model compared to the proprietary model

<table>
<thead>
<tr>
<th>Scenario #</th>
<th>Parameters*</th>
<th>Percentage cost savings in the outsourced model compared to the proprietary model</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Forward flow only</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.A.a</td>
<td>[100-0-0-10-0]</td>
<td>2.5%</td>
</tr>
<tr>
<td>2.A.b</td>
<td>[100-0-0-20-0]</td>
<td>0.1%</td>
</tr>
<tr>
<td><strong>Forward flow with failed deliveries</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.B.a</td>
<td>[100-0-5-20-40]</td>
<td>-0.4%</td>
</tr>
<tr>
<td>2.B.b</td>
<td>[100-0-5-20-80]</td>
<td>-0.8%</td>
</tr>
<tr>
<td>2.B.c</td>
<td>[100-0-10-0-20-40]</td>
<td>-0.8%</td>
</tr>
<tr>
<td>2.B.d</td>
<td>[100-0-10-0-20-80]</td>
<td>-0.8%</td>
</tr>
<tr>
<td><strong>Forward flow + Return flow. No Failed deliveries</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.C.a</td>
<td>[85-15-0-0-20-0]</td>
<td>0.1%</td>
</tr>
<tr>
<td>2.C.b</td>
<td>[100-15-0-0-20-0]</td>
<td>-0.5%</td>
</tr>
<tr>
<td><strong>Forward flow + Return flow + Failed deliveries + Failed Pickups</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.D.a</td>
<td>[85-15-5-20-40]</td>
<td>-0.3%</td>
</tr>
<tr>
<td>2.D.b</td>
<td>[85-15-5-20-40]</td>
<td>-0.3%</td>
</tr>
<tr>
<td>2.D.c</td>
<td>[85-15-10-20-40]</td>
<td>-0.7%</td>
</tr>
<tr>
<td>2.D.d</td>
<td>[85-15-10-20-40]</td>
<td>-0.7%</td>
</tr>
</tbody>
</table>

In Scenario 2.A.a, where the demand attracted by the CDP is only 10%, the outsourced model results in significantly lower costs. In Scenario 2.A.b and 2.C.a, where the demand attracted by the CDPs is 20%, the outsourced model is slightly more efficient. However, in all the other scenarios, the failed deliveries routed towards the CDPs increase the net demand at the CDPs, thus making the outsourced model less efficient.

While the exact benefits in the two governance models are sensitive to the cost parameters selected, this analysis shows how the tradeoffs can influence the selection of the operating model. A practical way to implement this is to assess the demand density, return proportion, and failed delivery
percentages attracted to each of the candidate CDP locations, and then outsource the operations of CDPs with lower net demand while owning the CDP sites which attract higher demand. This can help the retailer gain the benefit of both the models and further lower the cost of operations.
Chapter 7: Conclusion

In this thesis, we developed a framework for integrating collection-and-delivery points (CDPs) in the design of last-mile distribution networks. This was done using an optimization model that allows to solve a two-echelon location-routing problem. We considered simultaneous deliveries to customers and CDPs in the same tour through three possible route options- dedicated CDP routes, blended routes, and customer-only routes. The model considers the change in the demand due to the introduction of CDPs in the network and the increase in cost due to failed deliveries and failed return pickups. We used a continuum-approximation based augmented routing-cost estimation function to estimate the routing cost, which helps solve the problem efficiently in finite time.

We applied the framework on the distribution network of a leading Brazilian retailer and simulated several alternative scenarios. The results clearly demonstrate how the integration of CDPs results in lowering the overall distribution cost. There is a tradeoff in the fixed cost of a CDP and benefit derived from aggregating demand at the CDP, and as the demand attracted at the CDP increases, the cost savings increase significantly. Additionally, the results demonstrated the impact of CDPs in cost reduction due to preventing multiple delivery attempts for previously failed deliveries by giving the customers an option to route such parcels to the CDP. As the number of failed deliveries and failed customer pickups go up, the potential savings increase even further.

We also evaluated two possible governance models, proprietary CDPs which are owned by the retailer, and outsourced CDPs where there would be no fixed operating cost but a significantly higher handling cost per item. In demand zones, where the overall demand density, including returns and failed deliveries is low, outsourced CDPs offer a more cost-efficient operating model. Retailers can thus choose how to operate the downstream distribution channel.
Retailers can also use the overall framework to design the upstream distribution network. In our example, while the location of active SFs did not change due to the integration of CDPs, the overall cost of distribution decreased. Additionally, if the SF locations had more stringent capacity constraints, the resulting network might have looked different. The primary factor in deciding the cost savings is the amount of demand attracted to the CDPs, the more the demand, the higher the savings. The retailers should thus find ways to incentivize customers to choose CDPs as delivery options. A possible way to do this is in the form of a pickup discount, whereby the retailers pass-on some of the cost savings to the customers in the form of an additional discount when opting for a pickup from a CDP.

While we have proposed a way to integrate CDPs in the overall network design, there is scope to refine the methodology further. Future work should consider extensions to incorporate customer preferences such as instantaneous deliveries and delivery windows. The research can also be extended to consider scenarios such as capacitated CDPs which constrain the amount of demand that can be attracted at each CDP. Keeping the number and location of CDPs as a decision variable in the optimization model can help in minimizing the distribution cost even further. Future work can additionally capture other benefits from the integration of CDPs. For instance, we haven’t quantified the impact of the reduction in traffic congestion and the resultant impact on saving travel time due to the reduced number of customer deliveries. Alternative models, attempting to minimize parameters such as total carbon emissions, can also be considered.
References


