Demand Management Strategies for Operations of Shared Mobility Services

Mahsa Farhani

Demand Management Strategies for Operations of Shared Mobility Services

Mahsa FARHANI

Demand Management Strategies for Operations of Shared Mobility Services

Dissertation

for the purpose of obtaining the degree of doctor
at Delft University of Technology
by the authority of the Rector Magnificus, Prof.dr.ir. H. Bijl,
chair of the Board for Doctorates
to be defended publicly on
Thursday 5 February 2026 at 12:30

by

Mahsa FARHANI

This dissertation has been approved by the promotors.

Composition of the doctoral committee:

Rector Magnificus	Chairperson
Prof.dr. J. Rezaei	Delft University of Technology, promotor
Dr. S. Sharif Azadeh	Delft University of Technology, promotor
Dr. M.Y. Maknoon	Delft University of Technology, copromotor

Independent members:

Prof.dr.ir. L.A. Tavasszy	Delft University of Technology
Dr. B. Atasoy	Delft University of Technology
Dr. B. Farooq	Toronto Metropolitan University, Canada
Dr.ir. N. Alisoltani	Gustave Eiffel University, France
Prof.dr. R.R. Negenborn	Delft University of Technology, reserve member

TRAIL Thesis Series no. T2026/4, The Netherlands Research School TRAIL

TRAIL
P.O. BOX 5017
2600 GA Delft
The Netherlands
E-mail: info@rsTRAIL.nl

ISBN: 978-90-5584-379-4

Printed in the Netherlands

To my husband, Mohammad.

Acknowledgment

Since I have reached the end of my Ph.D. journey, I would like to take this opportunity to express my heartfelt gratitude to those who have been part of it. This path, marked by aspirations, challenges, and invaluable lessons, would not have been possible without the support of many remarkable individuals. I am deeply thankful to all who accompanied me along the way.

I am especially grateful to my supervisory team: Prof. Jafar Rezaei, Dr. Yousef Maknoon, and Dr. Shadi Sharif Azadeh, for your invaluable guidance and support throughout this journey. I feel truly fortunate to have had the chance to work under your supervision. From you, I learned to question assumptions and to approach research from different angles with greater rigor and clarity. Each of you brought unique strengths that enriched my PhD experience and contributed greatly to my development as a researcher.

Jafar, your continuous support gave me direction at crucial moments. The constructive and detailed feedback you provided pushed me to refine my ideas and remain focused. I deeply admire your intellectual rigor, sharp clarity of thought, and remarkable problem-solving skills, both in research and beyond, which have shaped both my work and my way of thinking.

Yousef, working with you has been an enriching experience. Your depth of knowledge and sharp insights challenged me to think more critically and to strengthen my work. I greatly valued the freedom you gave me to explore my own ideas, always balanced with your guidance whenever it was needed. Your generosity in sharing your expertise has contributed significantly to my growth as a researcher.

Shadi, the perspectives you brought into our discussions encouraged me to approach research with greater focus and precision. Your feedback helped me to sharpen my arguments. I greatly admire your determination, and your example as a strong and accomplished woman in science has been a source of inspiration throughout my journey.

I am also sincerely grateful to Prof. Caspar Chorus, my promotor during the first half of my PhD. Your encouragement, thoughtful feedback, and structured thinking were immensely valuable, and it was a privilege to learn from your insights. I also wish to thank my doctoral committee: Prof. Lori Tavasszy, Dr. Bilge Atasoy, Dr. Bilal Farooq, and Dr. Negin Alisoltani, for dedicating their time and expertise to reviewing my dissertation, and whose invaluable comments were instrumental in its final improvement. It has been a true honor to have you as part of this journey.

During my PhD, I had the privilege of meeting many friends and colleagues, and it was a unique experience to connect with so many inspiring people. As it is impossible to name everyone, I would like to express my gratitude to all who made TPM a vibrant and pleasant working environment.

Baiba, I owe you a very special thank you. Meeting you at the very beginning of my PhD — while you were completing your PhD and later becoming an assistant professor in the same section — was something I have always been grateful for. You have been not only an inspiring colleague but also a true friend. Thank you for the comfort and support you gave me, especially in the moments I needed it most.

Somayeh, you are a truly kind and generous friend, and I feel very fortunate that our paths crossed first at Sharif and later again at TU Delft. Thank you for being not only a supportive colleague but also a dear friend, and for always offering your thoughtful advice.

Zeinab, meeting you during your visiting research at TU Delft was a gift. From the very beginning, it felt as if we had known each other for ages. I will always cherish our deep conversations and the joyful moments we shared.

Pelin, I am especially grateful for your support during this journey. Our many conversations, about both the challenges and progress of our PhD journeys and also about life beyond research, reminded me that I was never alone in this process.

I am grateful to my officemates: Julien, Hong, Yvonne, Jerico, Ignacio, Qiuju, Steven, and Ahmad, along with my TU Delft friends and colleagues: Patrick, Sara, Shahrzad, Bahareh, Geqie, Merve, Dongyong, Alex, Mohammad Reza, Ali, and Pouria. Your academic and emotional support throughout this adventure, and the relaxed and friendly atmosphere you created, made this journey much more enjoyable.

I want to express my heartfelt appreciation to my beloved family. Although distance separates us, your constant encouragement and support have given me strength along the way. To my wonderful parents, I am deeply grateful for your unconditional love, your unwavering belief in me, the countless sacrifices, and the constant support, which have shaped who I am today. I carry the lessons of your wisdom, resilience, and kindness with me always. Words fall short of conveying the depth of my love and gratitude for you. To my kind and loving parents-in-law, your warmth has given me a true sense of belonging. To my amazing siblings and siblings-in-law: Zahra, Reyhaneh, Somayeh, Soudabeh, Sadegh, and Ali; I am endlessly thankful for having you. The joy and support you bring into my heart have been a constant blessing. I would also like to extend my gratitude to my dear uncles, Mahdi and Mostafa, and to Maryam Shapouri, for your wise advice and encouragement, which brought me comfort that meant a great deal to me during this journey. I also remember with love my grandfather, who always believed in me with pride; his faith remains a lasting source of inspiration.

Last but not least, I want to thank my husband and best friend, Mohammad. I feel incredibly fortunate to have you by my side, and I cannot imagine completing this dissertation without you. Your kindness, patience, and unwavering support have been my strength throughout this journey. In moments when my motivation was low, you lifted me with your encouragement, thoughtful perspective, can-do attitude, and strong belief in me, which inspired me to keep pursuing my dreams. You have been my anchor and a true source of resilience. This accomplishment belongs to you as much as it does to me.

Mahsa Farhani
Delft, September 2025.

Contents

Summary

Shared mobility services, such as bike-, car-, and ride-sharing, offer significant potential for reducing emissions and congestion and enhancing mobility accessibility. Realizing their potential depends not only on technological development, policy support, and high-quality service provision but also on the consideration of the interaction between the choices of providers and users. Service providers – digital platforms or fleet operators – interact directly with users and determine service quality and attractiveness through their service design and demand management decisions. Users, in turn, decide whether to adopt a service, how frequently to use it, and which option to choose in their engagement with the service (e.g., selecting among trip alternatives that differ in price, travel time, waiting time, mode, or route). These users' decisions shape service uptake and utilization, which ultimately determine whether shared mobility achieves its intended economical, environmental, and social benefits. Motivated by the interdependence between providers' decisions and users' responses, and their joint role in shaping the success of shared mobility, this dissertation focuses on the provider–user interaction, examining operations management from the provider's perspective while incorporating behavioral aspects of user decision-making.

In practice, users' decisions depend on service characteristics shaped by providers' operations management and are systematically influenced by interpersonal and intrapersonal behavioral factors. Interpersonal factors capture external influences from social interactions, such as social influence (i.e., imitating others' decisions or conforming to others' choices). Intrapersonal factors refer to behavioral phenomena arising from internal cognitive and psychological processes that shape how individuals evaluate and choose among alternatives. Among these are context effects, where the composition of the choice set influences user preferences between alternatives and their choices. Most current models either overlook the link between service design and user decisions or neglect the behavioral factors that shape how users respond to service design and demand management decisions. This dissertation aims to bridge this gap by examining how service providers can incorporate behavioral factors into operations management. To this end, three interrelated studies are conducted.

Study 1 develops a conceptual framework that links user choice behavior with providers' decisions. It reviews shared mobility operations management research from a behavioral perspective and identifies behavioral factors that are highly relevant but remain underexplored. The results highlight two complementary pathways that are investigated in the subsequent studies: (i) incorporating social influence as an interpersonal behavioral factor at the aggregated level to guide long-term service design, and (ii) using context effects as intrapersonal behavioral factors that can be leveraged in short-term choice set design decisions.

Study 2 incorporates social influence into users' adoption decisions for new services, capturing the resulting S-shaped market penetration curve: slow uptake in the early stages, rapid growth in

the intermediate phase, and eventual stabilization at saturation. We address the strategic service design problem of determining service location and fleet size by formulating a mixed-integer linear programming model for phased deployment, where fleet size and coverage expand over time as adoption grows, with the objective of maximizing total profit over a multi-period planning horizon. In the model, user adoption depends explicitly on both the number of current adopters (social influence) and service characteristics (coverage and vehicle availability, i.e., level of service) at an aggregated level. Scenario-based analyses using a real bike-sharing service dataset show that selective coverage with a high level of service outperforms broad coverage with a low level of service. Phased deployment increases profits by up to 4%, while ignoring service characteristics can reduce profits by up to 54%. This study shifts service design from static demand assumptions to behaviorally informed planning, highlighting the importance of social influence and service quality in shaping uptake. Our findings suggest that gradual service deployment that is aligned with adoption growth is critical for long-term profitability.

Study 3 examines choice set design in shared mobility platforms, with a focus on how context effects shape the composition of alternatives (i.e., the service menu) presented to users. We develop a framework for choice set selection under the Generalized Random Regret Minimization (G-RRM) choice model, which explicitly incorporates context effects. The proposed model accounts for the presence of an external mobility alternative (i.e., a competitor), thereby capturing the possibility that users may reject the offered set entirely and choose the competitor option instead. To support analysis, we introduce diagnostics that quantify compromise and attraction structures, enabling investigation of how these effects guide assortment selection. In addition, we propose an efficient optimization algorithm. In our experiments, we compare assortments generated under the context-sensitive G-RRM model with those generated under the Multinomial Logit (MNL) model, which ignores context effects and serves as the benchmark. Our scenario-based analysis shows that incorporating behavioral factors—specifically context effects—into choice set selection can improve outcomes, whether in terms of expected profit or the choice share of the set against competitors. Such improvements may result either from suppressing the appeal of competing options or from strategically eliminating options to avoid context effects that generate undesired outcomes. These findings demonstrate that context effects can function both as assets and as liabilities: in the first case, they should be leveraged when designing the choice set, while in the second case, they should be anticipated and mitigated.

This work sheds light on the interaction between service operations and user behavior in shared mobility services. Across the three studies, it demonstrates that interpersonal and intrapersonal behavioral mechanisms can be systematically embedded into providers' decisions. They conceptually and analytically show why and how user decision-making behavior must be integrated into service design and demand management. It shows that managers should adopt gradual, high-quality rollouts to secure early adoption, avoid overinvestment, and improve profitability, while leveraging context effects as behavioral levers instead of relying only on costly incentives. For policymakers, it provides a dual user–provider perspective that supports the design of promotion schemes, regulations, and subsidies aligned with user adoption dynamics while ensuring providers' financial viability to support their sustained engagement. Our models are theory-driven and tested on synthetic data, which allows controlled experimentation; future empirical research is needed to examine their generalizability.

Samenvatting

Gedeelde mobiliteitsdiensten, zoals deelfietsen, deelauto's en carpoolen, bieden aanzienlijke mogelijkheden om emissies en congestie te verminderen en de toegankelijkheid van mobiliteit te vergroten. Het realiseren van dit potentieel hangt echter niet alleen af van technologische ontwikkeling, beleidsmatige ondersteuning en hoogwaardige dienstverlening, maar ook van de interactie tussen de keuzes van aanbieders en gebruikers. Dienstaanbieders – digitale platforms of vlootexploitanten – hebben een directe interactie met gebruikers en bepalen de kwaliteit en aantrekkelijkheid van de dienst via hun beslissingen over serviceontwerp en vraagsturing. Gebruikers besluiten op hun beurt of zij gebruik maken van een dienst, hoe frequent zij deze gebruiken en welke optie zij kiezen (bijvoorbeeld het selecteren van reisalternatieven die verschillen in prijs, reistijd, wachttijd, vervoerswijze of route). Deze beslissingen beïnvloeden de mate van adoptie en gebruik, die uiteindelijk bepalen of gedeelde mobiliteit de beoogde economische, ecologische en maatschappelijke baten oplevert. Vanuit de onderlinge afhankelijkheid tussen de keuzes van aanbieders en de reacties van gebruikers, en hun gezamenlijke rol in het succes van gedeelde mobiliteit, richt dit proefschrift zich op de interactie tussen aanbieder en gebruiker. Het onderzoekt operations management vanuit het perspectief van de aanbieder, met expliciete aandacht voor de gedragsaspecten van gebruikers. In de praktijk worden gebruikersbeslissingen beïnvloed door dienstkenmerken die voortkomen uit operations management van aanbieders, en systematisch gestuurd door interpersoonlijke en intrapersoonlijke gedragsfactoren. Interpersoonlijke factoren verwijzen naar externe invloeden uit sociale interacties, zoals sociale beïnvloeding (bijvoorbeeld het imiteren of overnemen van keuzes van anderen). Intrapersoonlijke factoren hebben betrekking op cognitieve en psychologische processen die bepalen hoe individuen alternatieven evalueren en kiezen. Voorbeelden hiervan zijn contexteffecten, waarbij de samenstelling van het keuzepakket de voorkeuren en keuzes van gebruikers beïnvloedt.

De meeste bestaande modellen negeren ofwel de relatie tussen serviceontwerp en gebruikerskeuzes, ofwel de gedragsfactoren die bepalen hoe gebruikers reageren op beslissingen over serviceontwerp en vraagsturing. Dit proefschrift beoogt deze kennislacune te overbruggen door te onderzoeken hoe aanbieders gedragsfactoren kunnen integreren in operations management. Hiervoor zijn drie onderling samenhangende studies uitgevoerd.

Studie 1 ontwikkelt een conceptueel kader dat gebruikerskeuzegedrag koppelt aan beslissingen van aanbieders. Het biedt een overzicht van onderzoek naar operations management in gedeelde mobiliteit vanuit een gedragsmatig perspectief en identificeert relevante maar onderbelichte gedragsfactoren. De resultaten wijzen op twee complementaire onderzoekspaden die in de volgende studies verder worden uitgewerkt: (i) het integreren van sociale beïnvloeding als interpersoonlijke gedragsfactor op geaggregeerd niveau om langetermijnbeslissingen in serviceontwerp te ondersteunen, en (ii) het benutten van contexteffecten als intrapersoonlijke ge-

dragsfactor in kortetermijnbeslissingen over keuzesetontwerp.
Studie 2 integreert sociale beïnvloeding in gebruikersadoptie van nieuwe diensten en modelleert de resulterende S-vormige marktpenetratiecurve: trage adoptie in de beginfase, snelle groei in de tussenfase en stabilisatie bij verzadiging. Het strategische vraagstuk van serviceontwerp – locatiekeuze en vlootomvang – wordt benaderd met een mixed-integer linear programming-model voor gefaseerde uitrol, waarbij vlootgrootte en dekking groeien naarmate de adoptie toeneemt. Het doel is het maximaliseren van de totale winst over een multi-periodieke planningshorizon. In dit model is gebruikersadoptie expliciet afhankelijk van zowel het aantal huidige gebruikers (sociale beïnvloeding) als dienstkenmerken (dekking en voertuigbeschikbaarheid). Scenario-analyses met gegevens van een fietsdeeldienst tonen aan dat selectieve dekking met hoge servicekwaliteit beter presteert dan brede dekking met lage servicekwaliteit. Gefaseerde uitrol verhoogt de winst met maximaal 4%, terwijl het negeren van dienstkenmerken kan leiden tot een winstdaling van wel 54%. Deze studie verschuift de focus van statische vraagaanname naar gedrag-geïnformeerde planning en benadrukt het belang van sociale beïnvloeding en service-kwaliteit in gebruikersadoptie.
Studie 3 onderzoekt het ontwerp van keuzemenu's in gedeelde mobiliteitsplatforms, met speciale aandacht voor hoe contexteffecten de samenstelling van alternatieven (d.w.z. het servicemenu) beïnvloeden. We ontwikkelen een raamwerk voor keuze¬setselectie onder het Generalized Random Regret Minimization (G-RRM) keuzemodel, dat contexteffecten expliciet meeneemt. Het voorgestelde model houdt rekening met de aanwezigheid van een externe mobiliteitsoptie (d.w.z. een concurrent), en weerspiegelt daarmee de mogelijkheid dat gebruikers de aangeboden set volledig afwijzen en in plaats daarvan voor een concurrent kiezen. Ter ondersteuning van de analyse introduceren we diagnostieken die compromis- en aantrekkingsstructuren kwantificeren, waardoor onderzocht kan worden hoe deze effecten de assortimentselectie sturen. Daarnaast stellen we een efficiënt optimalisatie-algoritme voor. In onze experimenten vergelijken we assortimenten die worden gegenereerd onder het contextgevoelige G-RRM-model met die gegenereerd onder het Multinomial Logit (MNL)-model, dat contexteffecten negeert en als referentiepunt dient. Onze scenario-gebaseerde analyse laat zien dat het opnemen van gedragsfactoren—met name contexteffecten—in de keuze¬setselectie de uitkomsten kan verbeteren, zowel in termen van verwachte winst als van het keuzedeel ten opzichte van concurrenten. Deze verbeteringen kunnen ontstaan door het onderdrukken van de aantrekkingskracht van concurrerende opties of door strategisch opties te elimineren om contexteffecten die tot ongewenste resultaten leiden te vermijden. De bevindingen tonen aan dat contexteffecten zowel als assets als liabilities kunnen fungeren: in het eerste geval zouden ze benut moeten worden bij het ontwerpen van de keuzemenu's, terwijl ze in het tweede geval juist voorzien en vermeden moeten worden.
De bevindingen impliceren dat managers gefaseerde, hoogwaardige uitrolstrategieën zouden moeten hanteren om vroege adoptie te stimuleren, overinvestering te vermijden en winstgevendheid te vergroten, terwijl contexteffecten benut kunnen worden als gedragsmatige hefbomen in plaats van uitsluitend kostbare prikkels. Voor beleidsmakers biedt dit proefschrift een gecombineerd gebruikers- en aanbiedersperspectief dat kan bijdragen aan de vormgeving van promoties, regelgeving en subsidies die inspelen op adoptiedynamieken, terwijl tegelijkertijd de financiële duurzaamheid van aanbieders wordt gewaarborgd. De modellen zijn theorie-gedreven en getest met synthetische data, wat gecontroleerde experimenten mogelijk maakt; toekomstig empirisch onderzoek is noodzakelijk om de generaliseerbaarheid verder te toetsen.

Chapter 1

Introduction

1.1 Motivation

Shared mobility, such as bike-sharing, car-sharing, and ride-pooling, provides short-term, on-demand vehicle access without ownership burdens. These services attract users by reducing financial commitments and maintenance responsibilities, while charging only for actual use (Shaheen & Chan, 2016). They also support societal and environmental goals by reducing emissions and congestion while improving accessibility (Shaheen et al., 2016; Golbabaei et al., 2021; Zhu et al., 2023; Martinez et al., 2024).

Despite potential and strong policy support for promoting shared mobility (European Commission, 2020), many services face persistent challenges in sustained adoption, profitability, and measurable environmental impact (Lagadic et al., 2019; Boer et al., 2022). A key reason lies in how research and practice model the interaction between service providers and users. Users' and providers' decisions are interdependent. Providers manage operations through decisions on fleet deployment, pricing, and rebalancing, which determine service characteristics such as price and reliability. Users, in turn, make two types of decisions in response to the service characteristics. The first is service adoption, whether to use the service at all and how frequently to use it. The second involves short-term decisions, such as which option to select among multiple alternatives that differ in price, waiting time, travel time, mode, or route. Together, these adoption and usage decisions shape demand levels and spatial–temporal patterns of utilization, ultimately conditioning whether shared mobility can deliver its intended economic, environmental, and social benefits. However, this interdependence between providers and users is often simplified or ignored.

In practice, users' decisions are not only shaped by service attributes determined by providers' operations (Zhang & Kamargianni, 2023; Bachand-Marleau et al., 2012; Ma et al., 2023; Albiński et al., 2018), but also by a range of behavioral factors. These factors can be divided into intrapersonal and interpersonal dimensions. Intrapersonal factors arise from internal cognitive and psychological processes, such as using decision heuristics, which systematically bias perception and evaluation of alternatives from the rationality principles. Interpersonal factors emerge from social interactions, such as imitating others' choices, where the behavior or expectations of others affect an individual's decisions.

Nevertheless, much of the existing literature on shared mobility operations—including fleet sizing, facility location, pricing, and inventory planning (Agatz et al., 2012; Mourad et al., 2019; Tafreshian et al., 2020; Ho et al., 2018) represents user decisions through the lens of standard economic theory. In these models, users are typically treated as fully rational agents with stable preferences, who independently optimize utility[1]. While analytically convenient, this assumption neglects the systematic and predictable deviations from rationality documented in behavioral economics, thereby overlooking opportunities to incorporate user behavior into demand management strategies for shared mobility services.

While Behavioral Operations Management has demonstrated how behavioral factors shape operational outcomes (Bendoly et al., 2006; Gino & Pisano, 2008; Croson et al., 2013), their integration into shared mobility research remains limited. Operations broadly refer to the processes through which products and services are developed, produced, delivered, and distributed (Gino & Pisano, 2008). In this dissertation, we use the term operations to denote the full spectrum of decisions that determine how a service is designed, managed, and delivered. In the context of shared mobility, this encompasses long-term and short-term decisions, including service location, fleet sizing, pricing, resource allocation, request–supply matching, and vehicle repositioning. Accordingly, operational models are defined as analytical frameworks that address these operational decisions. Within this view, demand management seeks to align supply and demand through both long-term service design (e.g., service rollout, fleet sizing, facility location, pricing, relocation policies) and short-term actions (e.g., dynamic pricing[2], targeted incentives, and availability management).

This dissertation is motivated by the central role of users and providers in realizing the potential of shared mobility services, and by the interdependence between providers' operational decisions and users' behavioral responses. It examines shared mobility through the lens of behavioral operations management, with a focus on the provider–user interaction. It recognizes that adoption of a service and maintaining its use are systematically shaped by intrapersonal and interpersonal behavioral factors, and that these behaviors can be leveraged as demand management strategies in shared mobility operations. Positioning user behavior as a central element of demand management strategies provides a pathway to close the gap between the expected potential of shared mobility and its often limited performance in practice.

1.2 Research Questions

The goal of this dissertation is to develop effective demand management strategies in shared mobility services by incorporating user's behavioral factors into providers' operations.

Main Research Question. The overarching research question guiding this dissertation is:

How can shared mobility service providers incorporate users' behavioral factors in their demand management strategies?

To address this question, the dissertation (i) identifies and analyzes relevant behavioral fac-

[1] *Utility* represents decision-makers' subjective perceptions of value that induce a preference ordering over actions/choices.

[2] The difference between *pricing* and *dynamic pricing* is that pricing is the general practice of setting a price, while dynamic pricing is a specific, flexible strategy that uses real-time data to adjust those prices frequently.

tors that shape user behavior and interact with providers' operations in shared mobility services, (ii) develops models to capture and quantify these behavioral factors, and (iii) examines the implications of incorporating them into providers' operations. We focus on behavioral effects that are directly relevant to shared mobility operations and can be acted upon within providers' service design and demand management decisions. In particular, we are interested in effects that can be anticipated, influenced, or strategically leveraged during different stages of user interaction with the provided service.

The research is structured around the following sub-questions, each targeting a specific objective.

SQ1 *What users' behavioral factors should be integrated into the operations of shared mobility services?*
The first step is to identify relevant behavioral factors that systematically affect user decisions in shared mobility. The objective is to build a conceptual framework that links these behavioral factors to providers' operations and to identify the research gaps.

Identifying and structuring the behavioral factors most relevant for shared mobility provides the scientific foundation for embedding behavioral realism into operational models. It provides an understanding of which aspects of user behavior can be leveraged as predictable and actionable demand management levers.

Among the wide range of behavioral factors studied in the literature, this dissertation concentrates on social influence and context effects that are especially relevant in shared mobility operations. Social influence refers to changes in individuals' preferences that arise from the observed or expected behavior of others, often through imitation or conformity (Baddeley, 2010; Deutsch & Gerard, 1955; Hsu & Lu, 2004; Cialdini & Goldstein, 2004). In shared mobility, this means that the likelihood of adoption often increases as more people in one's network or community begin using the service, creating gradual rather than instantaneous demand uptake.

Context effects, by contrast, describe how the composition of the available choice set can systematically shift preferences in ways not predicted by rational choice theory. Accordingly, the probability of selecting a given option depends not only on its attributes but also on the presence or absence of alternatives (Prelec et al., 1997; Huber et al., 1982; Simonson, 1989). In shared mobility platforms, this suggests that the menu of routes, vehicles, or price–quality combinations presented to the user can influence users' decisions in a predictable way.

These two behavioral elements provide complementary pathways for linking behavioral insights with demand management through operational models, which in this research refer specifically to service deployment and menu design. Social influence, as an interpersonal factor, shapes how adoption unfolds over time and can be incorporated into service deployment through operational demand management levers. By operational demand management lever, we mean adjustments in service design and management decisions, such as rollout strategies, fleet sizing, or facility location, to better align supply with demand. Context effects, as an intrapersonal factor, alter user choices when selecting from a menu of options and can be applied as a behavioral demand management lever, where behavioral insights are intentionally used to influence users' choices.

SQ2 *How can service design be adapted to the gradual demand realization of new services as an operational demand management strategy?*

In shared mobility services, the joint consideration of service characteristics and user behavior in shaping adoption dynamics remains largely absent from operational models. Service characteristics such as quality and price (Zhang & Kamargianni, 2023; Bachand-Marleau et al., 2012; Ma et al., 2023; Albiński et al., 2018) determine the potential for adoption, yet this potential demand is realized gradually rather than instantaneously. Adoption typically follows an S-shaped trajectory: slow uptake in the early stages, rapid growth during the intermediate phase, and eventual stabilization at market saturation (Baddeley, 2010; Deutsch & Gerard, 1955; Hsu & Lu, 2004; Cialdini & Goldstein, 2004). This gradual pattern reflects social influence, where users' adoption depends not only on service attributes but also on others' behavior. Accordingly, the likelihood that an individual adopts a service increases as more people adopt the service. Incorporating these dynamics into service design enables providers to align rollout strategies, such as fleet sizing and service location expansion, with the predictable trajectory of adoption shaped by social influence. This helps to anticipate and regulate the pace of demand realization through service design and reduces the risks of overinvestment, underutilization, or premature service failure.

SQ3 *How can context effects be leveraged as a behavioral demand management strategy for shared mobility services?*

Demand management is often pursued through operational levers such as price-based incentives or vehicle repositioning. Yet providers interact with users through digital platforms that define the choice set offered at the moment of request. This interaction creates an opportunity to use the composition of options presented to users as a behavioral demand management strategy, where menus can be adapted to influence user selections without altering underlying service attributes.

Incorporating context effects in designing the service menus offers a novel and underexplored approach to guiding user choices in shared mobility. By carefully selecting the composition of the menu of offered options to each user, providers can influence decisions toward outcomes that improve system performance, for example, reallocating demand toward oversupplied areas, encouraging parking in undersupplied locations, or promoting higher-margin options. Embedding these interpersonal behavioral mechanisms into operations extends demand management strategies beyond merely adjusting service attributes, such as monetary incentives.

This dissertation adopts a theory-driven analytical modeling approach, supported by insights from a systematic literature review. SQ1 identifies relevant behavioral factors in shared mobility operations by reviewing established theories and empirical findings. Building on these insights, SQ2 and SQ3 each examine one behavioral factor within operations. SQ2 considers social influence, together with service quality, as drivers of gradual adoption dynamics, and explores how aligning service rollout with this pattern functions as an operational demand management lever. SQ3 analyzes context effects, which shape user decisions within a given choice set, and integrates them into menu design as a behavioral demand management lever. Scenario- and sensitivity-based analyses are used to evaluate the implications of these mechanisms under controlled conditions. Together, these studies provide conceptual and methodological contributions that advance the integration of behavioral realism into operational models.

1.3 Scientific and Societal Relevance

This research provides both scientific and societal contributions as described below.

Scientific contributions. This dissertation is positioned within Behavioral Operations Management, which studies how actual human behavior influences operational outcomes such as profitability and resource utilization (Bendoly et al., 2006; Gino & Pisano, 2008; Croson et al., 2013). It brings this perspective to shared mobility by linking providers' operations (service design and demand management) with behavioral factors, specifically, social influence and context effects. The central premise is that user behavior can be treated as a lever, shaping both opportunities and constraints in operational decision-making.

This dissertation reframes users as actors whose adoption and usage decisions systematically interact with providers' strategies, not as passive or fully rational agents. It develops a framework for incorporating behavioral realism into shared mobility operations by embedding interpersonal factors (social influence) and intrapersonal factors (context effects). In doing so, it anticipates adoption dynamics in long-term planning and influences user choices within menu-based decision contexts.

The dissertation develops analytical models that integrate behavioral mechanisms into optimization frameworks for shared mobility. It extends adoption modeling by linking service rollout decisions to gradual, socially influenced demand realization, and develops methods for quantifying the presence and strength of context effects in any choice sets.

Societal contributions. The transport sector accounts for nearly one-fifth of global CO_2 emissions (Ritchie, H., 2020), with private cars as the main source. Shared mobility is often promoted as a way to reduce vehicle-kilometers traveled, energy use, and emissions, while improving urban accessibility. Yet many initiatives fail to scale or sustain operations despite strong policy support (Lagadic et al., 2019; Boer et al., 2022).

This dissertation contributes to overcoming these challenges by showing how behavioral factors can be incorporated into demand management strategies. The insights help providers strengthen adoption trajectories, improve profitability, and support the long-term survival of services. Ensuring that shared mobility services can endure is a prerequisite for realizing their broader societal promise.

The models and frameworks developed in this dissertation show that user behavior is a strategic demand management lever. This perspective can inform policy design (e.g., phased deployment requirements, targeted subsidies) and guide provider practices (e.g., demand steering through behavioral levers such as menu design rather than relying solely on price incentives). By enabling more effective demand management and improving the chances of long-term profitability, the dissertation supports shared mobility services to realize their societal potential in terms of sustainability.

1.4 Dissertation Outline

The structure of this dissertation is summarized in Figure 1.1. The chapters address the main research question by conducting a literature review and a conceptual foundation (Chapter 2), incorporating adoption dynamics as part of operational demand management through phased service rollout in long-term planning (Chapter 3), and investigating how choice set composition affects users' decisions and providers' menu design as a behavioral demand management strategy (Chapter 4).

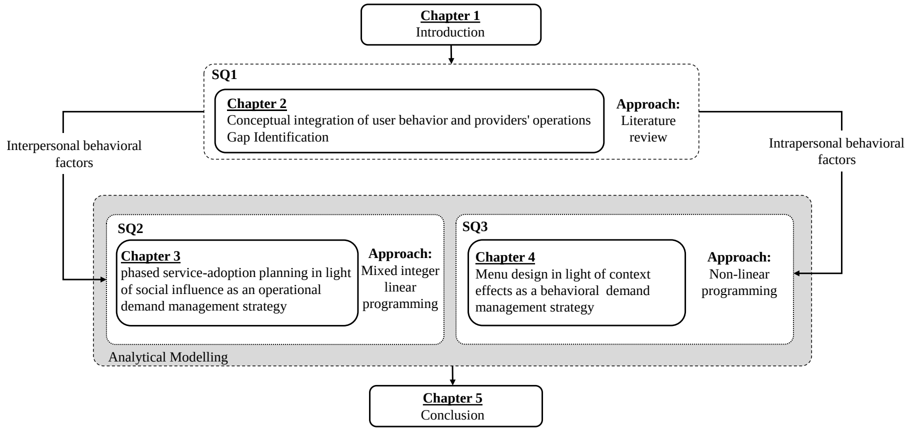


Figure 1.1: Overview of this dissertation

Chapter 2 develops the conceptual foundation and addresses the first sub-question (SQ1). It develops a framework that links user decision-making with providers' operational choices and reviews shared mobility studies that integrate user decisions into providers' operations from a behavioral perspective. The purpose is to identify behavioral factors that are most relevant but underrepresented in the existing models. This review highlights two promising directions that guide the next chapters: using social influence to better anticipate adoption dynamics in service design, and using context effects as potential levers in menu design.

Chapter 3 examines how users' adoption behavior can be incorporated into operational demand management strategies. Shared mobility services are often rolled out under the assumption of stabilized demand, yet in practice, adoption develops gradually, shaped by service quality and social influence. This chapter explores how such gradual demand uptake can be incorporated into deployment models. An optimization framework is developed that links adoption growth to service attributes and social influence. This framework synchronizes fleet sizing and location expansion with users' adoption trajectories. The focus is not on changing adoption behavior itself, but on using adoption patterns to anticipate and regulate demand realization over time.

Chapter 4 examines behavioral demand management by focusing on context effects in menu design. When users request service via digital platforms, the choice set they see is generated by providers. In this chapter, we investigate how the composition of this set can influence user decisions, beyond what rational-choice models would predict. This chapter explores whether incorporating context effects alters optimal menu design and under what conditions such modeling changes outcomes. To support this analysis, we define general metrics to quantify context effects across diverse choice sets and propose computational methods for solving the resulting menu design problem.

Chapter 5 concludes by synthesizing the insights from the three studies, outlining their contributions, reflecting on limitations, and suggesting directions for future work.

Together, these chapters build the argument that behavioral factors, both interpersonal and intrapersonal, can be treated as levers in demand management strategies, offering a new perspective on how shared mobility operations can be aligned with user behavior.

Bibliography

Agatz, N., A. Erera, M. Savelsbergh, X. Wang (2012) Optimization for dynamic ride-sharing: A review, *European Journal of Operational Research*, 223(2), pp. 295–303.

Albiński, S., P. Fontaine, S. Minner (2018) Performance analysis of a hybrid bike sharing system: A service-level-based approach under censored demand observations, *Transportation Research Part E: Logistics and Transportation Review*, 116, pp. 59–69.

Bachand-Marleau, J., B. H. Lee, A. M. El-Geneidy (2012) Better understanding of factors influencing likelihood of using shared bicycle systems and frequency of use, *Transportation Research Record*, 2314(1), pp. 66–71.

Baddeley, M. (2010) Herding, social influence and economic decision-making: socio-psychological and neuroscientific analyses, *Philosophical Transactions of the Royal Society B: Biological Sciences*, 365(1538), pp. 281–290.

Bendoly, E., K. Donohue, K. L. Schultz (2006) Behavior in operations management: Assessing recent findings and revisiting old assumptions, *Journal of Operations Management*, 24(6), pp. 737–752.

Boer, M., O. Türetken, O. E. Adali (2022) A review of business models for shared mobility and mobility-as-a-service (maas): a research report.

Cialdini, R. B., N. J. Goldstein (2004) Social influence: Compliance and conformity, *Annual Review of Psychology*, 55(1), pp. 591–621.

Croson, R., K. Schultz, E. Siemsen, M. Yeo (2013) Behavioral operations: the state of the field, *Journal of Operations Management*, 31(1-2), pp. 1–5.

Deutsch, M., H. B. Gerard (1955) A study of normative and informational social influences upon individual judgment., *The Journal of Abnormal and Social Psychology*, 51(3), p. 629.

European Commission (2020) Sustainable and smart mobility strategy–putting european transport on track for the future, *Sustainable and Smart Mobility Strategy–putting European transport on track for the future*.

Gino, F., G. Pisano (2008) Toward a theory of behavioral operations, *Manufacturing & Service Operations Management*, 10(4), pp. 676–691.

Golbabaei, F., T. Yigitcanlar, J. Bunker (2021) The role of shared autonomous vehicle systems in delivering smart urban mobility: A systematic review of the literature, *International Journal of Sustainable Transportation*, 15(10), pp. 731–748.

Ho, S. C., W. Y. Szeto, Y.-H. Kuo, J. M. Leung, M. Petering, T. W. Tou (2018) A survey of dial-a-ride problems: Literature review and recent developments, *Transportation Research Part B: Methodological*, 111, pp. 395–421.

Hsu, C.-L., H.-P. Lu (2004) Why do people play on-line games? an extended tam with social influences and flow experience, *Information & Management*, 41(7), pp. 853–868.

Huber, J., J. W. Payne, C. Puto (1982) Adding asymmetrically dominated alternatives: Violations of regularity and the similarity hypothesis, *Journal of Consumer Research*, 9(1), pp. 90–98.

Lagadic, M., A. Verloes, N. Louvet (2019) Can carsharing services be profitable? a critical review of established and developing business models, *Transport Policy*, 77, pp. 68–78.

Ma, L., X. Zhang, F. Hao, S. Liu (2023) Service quality, perceived price fairness, and users' continuous usage intentions regarding shared bike service, *Total Quality Management & Business Excellence*, pp. 1–21.

Martinez, L. M., J. P. Pritchard, P. Crist (2024) Shared mobility's role in sustainable mobility: Past, present, and future, *Annual Review of Environment and Resources*, 49.

Mourad, A., J. Puchinger, C. Chu (2019) A survey of models and algorithms for optimizing shared mobility, *Transportation Research Part B: Methodological*, 123, pp. 323–346.

Prelec, D., B. Wernerfelt, F. Zettelmeyer (1997) The role of inference in context effects: Inferring what you want from what is available, *Journal of Consumer research*, 24(1), pp. 118–125.

Ritchie, H. (2020) Cars, planes, trains: where do co emissions from transport come from?, URL https://ourworldindata.org/co2-emissions-from-transport.

Shaheen, S., N. Chan (2016) Mobility and the sharing economy: Potential to facilitate the first- and last-mile public transit connections, *Built Environment*, 42(4), pp. 573–588.

Shaheen, S., A. Cohen, I. Zohdy, et al. (2016) Shared mobility: current practices and guiding principles, Tech. rep., United States. Federal Highway Administration.

Simonson, I. (1989) Choice based on reasons: The case of attraction and compromise effects, *Journal of Consumer Research*, 16(2), pp. 158–174.

Tafreshian, A., N. Masoud, Y. Yin (2020) Frontiers in service science: Ride matching for peer-to-peer ride sharing: A review and future directions, *Service Science*, 12(2-3), pp. 44–60.

Zhang, Y., M. Kamargianni (2023) A review on the factors influencing the adoption of new mobility technologies and services: autonomous vehicle, drone, micromobility and mobility as a service, *Transport Reviews*, 43(3), pp. 407–429.

Zhu, J., N. Xie, Z. Cai, W. Tang, X. Chen (2023) A comprehensive review of shared mobility for sustainable transportation systems, *International Journal of Sustainable Transportation*, 17(5), pp. 527–551.

Chapter 2

Integrating Users' Decision-Making into Shared Mobility Operations

Abstract: Shared mobility services, such as bike-sharing, car-sharing, and ride-sharing, are widely promoted as sustainable mobility solutions. Yet many services continue to struggle with user uptake, consistent usage, and profitability. A central reason is that analytical models of shared mobility often impose restrictive assumptions about user behavior, treating choices as independent of operations or assuming full rationality. As a result, these models fail to capture the behavioral mechanisms that shape real-world demand, limiting their ability to inform strategies for improving performance and sustainability. This study aims to bridge the gap between operational models and the behavioral realism required for understanding and influencing demand by integrating a behavioral perspective into shared mobility modeling. We develop a conceptual framework that links user decision-making and behavioral factors with providers' operational strategies, offering a structured basis for incorporating behavioral realism into operational models. Guided by this framework, we review existing studies across shared mobility domains, examine how user decisions are represented, and identify key insights and research gaps. The contributions of the study are twofold: (1) proposing a framework for integrating behavioral aspects into shared mobility operations, and (2) synthesizing the literature through this framework to outline directions for research and practice.

Keywords: Shared mobility; Car-sharing; Ride-sharing; Micromobility; Decision-making; Choice behavior, Behavioral operations

---

Research from this chapter contributed to a paper submitted to the journal of *Transportation Research Record*.

2.1 Introduction

Existing research has substantially advanced our understanding of shared mobility. Studies have examined factors influencing service usage and adoption, as well as the broader impacts of these services on travel behavior. Numerous determinants of shared mobility usage have been identified. The main categories include service attributes (e.g., cost and quality) (Zhang & Kamargianni, 2023; Bachand-Marleau et al., 2012; Ma et al., 2023; Albiński et al., 2018), user characteristics (e.g., age, gender, income, car ownership), and trip characteristics (e.g., frequency, purpose, length) (Liao & Correia, 2022; Aguilera-García et al., 2020; Faghih-Imani et al., 2017). Reviews are available for specific modes such as bike sharing (Zhang & Kamargianni, 2023), scooter sharing (Badia & Jenelius, 2023), and car sharing (Ferrero et al., 2018). Complementing this user and demand analysis perspective, another line of research addresses providers' operations, including fleet sizing, service location, and rebalancing, as reviewed by Agatz et al. (2012); Mourad et al. (2019); Tafreshian et al. (2020); Ho et al. (2018).

A missing piece in this literature is the explicit connection between user decision-making and providers' operational decisions. In particular, little attention has been paid to how users' behavioral responses interact with providers' choices in service design and demand management, shaping both long-term service adoption and real-time demand-supply dynamics. Evidence from behavioral economics shows that human decision-making systematically departs from classical rational choice theory and can influence providers' operational decisions, such as pricing, inventory management, and resource allocation (Bendoly et al., 2006; Gino & Pisano, 2008; Croson et al., 2013). These behavioral insights are relevant to users of shared mobility. However, most research that models providers' operational decisions represents users as fully rational decision-makers or treats their choices as exogenous inputs, independent of providers' operations.

To bridge this gap between behavioral science and operations research in the context of shared mobility services, this study examines how user decision-making can be systematically integrated into supply-side operations. The contributions of this study are twofold. First, drawing on theories and concepts from behavioral economics, we propose a conceptual framework that links user choice behavior with provider operations management, providing a structured basis for embedding behavioral realism into shared mobility operational modeling. We use the term operational models to denote analytical approaches that address providers' decisions on service configuration and resource management (e.g., fleet sizing, location, rebalancing). Second, guided by this framework, we synthesize studies that incorporate user decisions into providers' operational models, evaluate them from a behavioral perspective, and identify key insights and opportunities for future research.

The remainder of this study is organized as follows. Section 2.2 clarifies the interdependencies between user decision-making and supply-side operations in shared mobility services and introduces a conceptual framework. Sections 2.3 and 2.4 elaborate on user decision-making and the behavioral factors most relevant to this context, respectively. Section 2.5 synthesizes the literature on operational models that incorporate user decisions. Section 2.6 highlights research gaps and opportunities for future work. Finally, Section 2.7 concludes the study.

2.2 Interactions of User Decisions and Provider Operations

Shared mobility services operate within a complex ecosystem involving diverse stakeholders, including service providers, users, regulators, local governments, and infrastructure planners. This study focuses on the interactions between users and service providers, specifically examining providers' operations in light of users' choice behavior. Users represent the demand side, with preferences shaped by user characteristics (e.g., age, gender, income, car ownership) and choice context, elaborated in the following parts. These preferences and users' requirements, such as service expectations about travel time and price, translate into both constraints and opportunities for providers. On the supply side, service providers, including digital platforms and fleet operators, must align their operations with user preferences and requirements while also ensuring efficiency, profitability, and regulatory compliance. Providers' own operational constraints further shape the set of service alternatives they can offer, which in turn constitute the supply-driven component of the user's choice context. This user–provider interaction can be conceptualized as two interdependent decision-making problems: the service provider's operational decisions and the user's decision-making.

Figure 2.1 presents a general framework linking service providers' operations with user decisions.

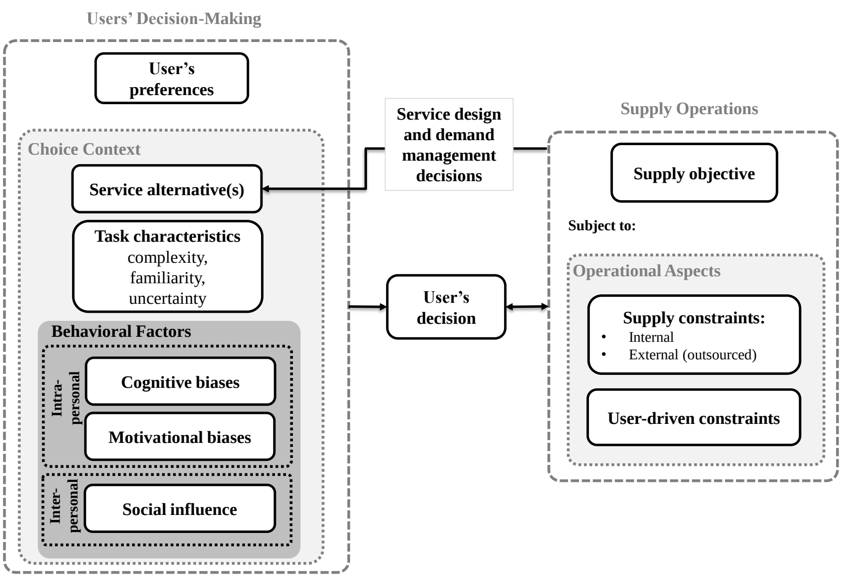


Figure 2.1: Conceptual framework integrating user behavior with shared-mobility service provider operations

Provider's Operational Decision-Making

On the supply side, operational decisions involve long-term and short-term decisions, including fleet sizing, inventory distribution, service zone coverage, service configuration (e.g., pooled vs. solo rides), and pricing schemes, as well as real-time matching, dynamic pricing, and routing. Service providers typically opt for profit maximization, demand fulfillment, or promoting specific transportation modes. These objectives are pursued under a combination of internal (e.g., budget and fleet size) and external (e.g., imposed by outsourced operators like freelance drivers) operational constraints and user-driven constraints (e.g., preferred travel times, service expectations). Providers' operational decisions determine the attributes of the service alternatives offered. These include dimensions such as service quality (e.g., fulfillment rate, spatial and temporal coverage) and price.

Users' Decision-Making

Users face a decision problem that ranges from adopting a service (e.g., selecting a provider or mode) to either accepting a single offered alternative or choosing among a set of alternatives designed and presented by the provider in response to their ride or vehicle request. Users may also need to decide whether to accept incentives, tolerate differences between offered and preferred options, wait for service, or cancel their request while considering other available service options.

Choice context. The choice context is shaped by task characteristics, service alternatives (defined by attributes such as price, reliability, and convenience), and behavioral factors, all of which influence users' decisions. The users' decision-making problem is characterized by features such as uncertainty, familiarity (e.g., repeated decisions with known options vs. novel, unfamiliar scenarios), and complexity, which we call task characteristics. The number of available options, their composition, and the number and comparability of attributes (e.g., differing units or vague descriptions) determine the task complexity for the user.

Behavioral factors. Users' choices are influenced by two broad categories of behavioral factors. Intrapersonal factors arise from internal cognitive and psychological processes that shape how individuals perceive and evaluate alternatives, often leading to systematic biases such as cognitive or motivational biases. Interpersonal factors, by contrast, are external influences stemming from social interactions, such as social influence, imitation, and conformity, where the behavior or expectations of others affect one's choices.

- **Cognitive and motivational biases.** There is a vast literature on biases that are classified into motivational and cognitive biases (Montibeller & von Winterfeldt, 2024; Montibeller & Von Winterfeldt, 2015). Motivational and cognitive biases are systematic deviations from fully rational decision-making. Motivational biases arise due to motivated reasoning, which directs individuals' decision-making processes in a way that ensures they reach their desired outcome (Molden & Higgins, 2012). For example, confirmation bias leads individuals to selectively seek and interpret evidence that supports their existing beliefs (Nickerson, 1998). Cognitive biases often arise due to simplifying reasoning, typically involving a decision heuristic (Montibeller & von Winterfeldt, 2024). Cognitive biases are categorized based on the psychological factors that drive them, including: too much information to process; not enough meaning, such as ambiguity or incompleteness in information and outcomes; the need to act quickly; and the need to remember, which reflects limitations in memory (Manoogian & Benson, 2017).
- **Social influence.** Individuals' choices can be influenced by the behavior of others, partic-

ularly in situations involving unfamiliarity and uncertainty, such as deciding whether to adopt an innovative service. A common explanation for this behavior is that individuals feel uncertain about the outcomes of choosing a new or unfamiliar option. In such cases, they are more likely to follow others, either to get social approval and acceptance (i.e, normative conformity) or to seek accurate evaluation and make correct decisions (i.e., informational conformity), which is known as social influence (Deutsch & Gerard, 1955; Hsu & Lu, 2004; Cialdini & Goldstein, 2004).

In the subsequent sections, we explain rational decision-making and human decision behavior. We then elaborate on the behavioral factors discussed above, focusing on those specifically relevant in the context of shared mobility operations.

2.3 Users Decision Behavior

Generally speaking, decision-making refers to the process of selecting one alternative from a set of feasible options, based on the decision-maker's preferences. The concept of utility provides a numerical representation of these preferences, enabling systematic comparison across options (von Neumann & Morgenstern, 1944; Debreu et al., 1954). Under the principle of utility maximization, the decision-maker is assumed to choose the alternative associated with the highest utility.

2.3.1 Rationality; Axiomatic Definition

The rational choice theory, pioneered by von Neumann & Morgenstern (1944) with their Expected Utility Maximization model, provides some necessary and sufficient conditions, referred to as axioms, to describe and model how humans would behave if they were perfectly rational.

Preference relations. Let Ω denote the set of feasible alternatives from which a decision-maker can choose. A preference relation, noted by $\succeq$, $\succ$, or $\sim$, is a binary relation on Ω. For any $X, Y \in \Omega$:

$$\begin{aligned}
X \succeq Y \quad &\Longleftrightarrow \quad X \quad \text{is at least as preferred as} \quad Y \quad \text{(weak preference)},\\
X \sim Y \quad &\Longleftrightarrow \quad X \succeq Y \text{ and } Y \succeq X \quad \text{(indifference)},\\
X \succ Y \quad &\Longleftrightarrow \quad X \succeq Y \text{ and not } (Y \succeq X) \quad \text{(strict preference)}.
\end{aligned}$$

The main rationality axioms (von Neumann & Morgenstern, 1944; Savage, 1972) are as follows, where $X, Y, Z \in \Omega$:

Completeness. Formally, for all X and Y, either $X \succeq Y$ or $Y \succeq X$ (or both).

Completeness means that a decision-maker can always compare any two options. Faced with two alternatives, they either prefer one or are indifferent between them. This axiom ensures that all decision problems can be addressed with a well-defined preference. For example, when deciding between public transport, taxi, or cycling, rational individuals should be able to form a ranking among these options and make a decision.

Transitivity. Formally, if $X \succeq Y$ and $Y \succeq Z$, then $X \succeq Z$.

In words, if an individual prefers X to Y and Y to Z, then they prefer X to Z. Without transitivity, preferences could form cycles (e.g., $X \succ Y$, $Y \succ Z$, $Z \succ X$), which undermines coherent decision-making. Transitivity requires consistency in preferences and guarantees that preferences can be arranged in a stable ordering, enabling a rational choice.

When preferences are transitive and complete, decision-makers can create a ranking of options from most to least preferred, known as a preference ordering. Completeness guarantees that there will be only one ordering, with no gaps or ambiguous comparisons. Transitivity guarantees that there will be no cycles in strict preference. For example, if there are three options (X, Y, and Z), they must be fully ranked, such as X is better than Y is better than Z, or X and Y are indifferent but both are better than Z.

Independence. Formally, if $X \succeq Y$, then for any third option Z and probability $\lambda \in (0, 1]$,

$$\lambda X + (1 - \lambda)Z \succeq \lambda Y + (1 - \lambda)Z.$$

This axiom means that if X is preferred to Y, then a compound option that yields X with probability λ and Z with probability $(1 - \lambda)$ will still be preferred to the corresponding compound option that yields Y with probability λ and Z with probability $(1 - \lambda)$.

Continuity. This axiom ensures that preferences are continuous. Formally, if $X \succeq Y \succeq Z$, there exists $\lambda \in (0, 1)$ such that $\lambda X + (1 - \lambda)Z \sim Y$.

It states that if an individual prefers X to Y and Y to Z, then there exists some probability $\lambda \in (0, 1)$ such that the compound option $\lambda X + (1 - \lambda)Z$, which yields the outcome of option X with probability λ and the outcome of option Z with probability $1 - \lambda$, is valued equally to Y. In simpler terms, this means that for any three outcomes ranked from best to worst, the middle option can be represented as an appropriate combination of the best and worst outcomes. For example, consider three options with the following outcomes: X results in winning 100$, Y results in winning 50$, and Z results in losing 100$, where $X \succ Y \succ Z$. Continuity requires that there exists a probability $\lambda \in (0, 1)$ such that the individual is indifferent between receiving 50$ for sure (i.e., option Y) and a compound option that yields 100 $ with probability λ and costs 100$ with probability $(1 - \lambda)$.

The classical axioms of rational choice are completeness, transitivity, continuity, and independence, as established in expected utility theory (von Neumann & Morgenstern, 1944; Savage, 1972). In the discrete choice tradition, Luce's (1959) Independence of Irrelevant Alternatives (IIA) condition is a rationality requirement. The IIA condition suggests that the presence of a third option like Z does not alter the original ordering between two other alternatives like X and Y. This property allows preferences to be analyzed without requiring the full choice set, and it simplifies economic models by making preference orderings more stable and predictable. It is also referred to as the presentation and procedural invariance principle because it ensures context-independence preferences. Accordingly, preferences between any two options are not affected by the strategic inclusion or exclusion of alternatives. This independence has several implications: it allows preferences to be analysed without requiring the full choice set, it simplifies economic models by making preference orderings more stable and predictable, and it assumes that decision outcomes are not affected by the strategic inclusion or exclusion of alternatives.

If all these axioms are satisfied, then the individual is considered rational. In that case, preferences can be represented by a utility function, meaning that one can assign numbers to each alternative to reflect the strength of preference (Mas-Colell et al., 1995).

Formally, a utility function $u(\cdot)$ maps alternatives to real numbers, with larger values assigned to more preferred options. These utilities are grounded in the observable characteristics of the alternatives. An attribute is a measurable feature, such as cost, time, or quality, that shapes preferences. Multi-attribute utility theory captures this by expressing overall utility as a function of the attributes that define each alternative.

The preference relation can then be expressed in terms of the utility function as follows (Mas-Colell et al., 1995):

$$X \succeq Y \iff u(X) \geq u(Y), \quad X \succ Y \iff u(X) > u(Y), \quad X \sim Y \iff u(X) = u(Y).$$

That is, option X is at least as preferred as option Y if and only if $u(X) \geq u(Y)$; X is strictly preferred to Y if and only if $u(X) > u(Y)$; and X is indifferent to Y if and only if $u(X) = u(Y)$.

A rational decision-maker is assumed to select the option with the highest utility (i.e., i^*):

$$i^* = \arg\max_i u(i).$$

Example. Consider a mobility context in which a traveler evaluates three alternatives: Option X (bus), Option Y (bike), and Option Z (car). Each option is described by two attributes: travel time and cost. Option X requires 40 minutes and costs \$2, Option Y requires 60 minutes with no cost, and Option Z requires 20 minutes and costs \$6.

The utility of option $i \in \{X, Y, Z\}$ can be specified as

$$u(i) = \omega_1 \cdot \text{Cost}_i + \omega_2 \cdot \text{Time}_i,$$

where ω_1 and ω_2 are weights capturing the traveler's sensitivity to cost and time, respectively. In this formulation, ω_1 reflects the marginal utility of cost, that is, the change in utility associated with a one-unit change in monetary cost, while ω_2 reflects the marginal utility of time, or the change in utility associated with a one-unit change in travel time. Because cost and time are expressed in different units, these weights simultaneously rescale the attributes onto a common utility scale and capture the relative importance the traveler assigns to each attribute. The rational choice is then the option with the highest utility value:

$$\text{Optimal option} = \arg\max_{i \in \{X,Y,Z\}} u(i).$$

These options are evaluated based on how travelers, as decision-makers, value different attributes. For instance, one traveler may prioritize travel time, while another may be more sensitive to costs. A time-sensitive traveler might therefore assign higher utility to faster modes, for instance $u(X) = 0.6$, $u(Y) = 0$, and $u(Z) = 0.9$ (normalized between 0 and 1); resulting in the preference ordering $Z \succ Y \succ X$. Since $u(Z)$ is the highest among the three alternatives, the rational choice is Option Z (car).

2.3.2 Deviations from Rationality

Rationality assumptions ignore humans' cognitive limitations, such as their limited information processing capacity. Individuals deal with their cognitive limitations by using simplifying rules. They often restrict their attention to a subset of alternatives and attributes and settle for an option that is "good enough" rather than the optimal one, known as bounded rationality (Simon, 1955).

Rationality axioms imply that people have a well-defined preference order of alternatives that are independent of the alternatives' composition and description. The notion of constructive preference, introduced by Payne et al. (1992), contradicts these assumptions and emphasizes the constructive nature of decision behavior. It suggests that individuals do not have a master list of preferences to evoke in their decision-making to choose the option with the highest preference rank in the list. Instead, they often construct their preferences for alternatives in tasks of any complexity, not in a consistent manner (Payne et al., 1992), which makes their evaluations subject to deviations from rationality. The constructive nature of preferences, and the resulting biases, including loss aversion, and context effects, have been widely demonstrated in empirical studies (Tversky & Simonson, 1993; Huber et al., 1982; Ariely & Wallsten, 1995; Ariely & Jones, 2008; Huber et al., 2014; Tversky & Kahneman, 1991; Bettman et al., 1998; Yoon & Simonson, 2008).

Strategies that are applied in constructing preferences depend on individuals' cognitive abilities and psychological state (e.g., need for justifying their choices (Simonson, 1989; Montgomery, 1983)), the characteristics of the decision-making task (i.e., task unfamiliarity (Bettman et al., 1998; Yoon & Simonson, 2008), uncertainty (Kahneman & Tversky, 1979; Tversky & Kahneman, 1992), complexity (Chernev, 2003; Nagpal & Krishnamurthy, 2008)), choice-set architecture (composition of alternatives) (Prelec et al., 1997; Simonson, 1989), and how they are presented (framing) (Tversky & Kahneman, 1981, 1989)). These factors motivate the use of decision heuristics, mental shortcuts that simplify decision-making, to reduce the cognitive effort required and could result in predictable deviations from idealized rational choices (i.e., cognitive biases).

2.4 Behavioral Factors in Shared Mobility Operations

The behavioral effects that are directly relevant to and actionable within service providers' operations management decisions made by providers include social influence, context, and loss aversion.

Social influence. In the context of shared mobility, the environment is dynamic and continuously evolving, with new modes of transport, vehicle types, and service providers. This dynamic environment introduces considerable uncertainty for users, who must make decisions about unfamiliar options whose quality and reliability they cannot fully anticipate. In such situations, social influence can play a key role in reducing uncertainty and cognitive effort for decision-making (Cialdini & Goldstein, 2004; Stathopoulos et al., 2017).

Social influence implies that individuals' preferences may be altered by the choices of others, factors that are irrelevant to the actual characteristics of the option being evaluated. For instance, an individual's preference for adopting a new service or product is often positively correlated with the proportion of the population already using it. As a result, after a certain point in time, people may begin to adopt a service even in the absence of any change to the service itself. Social influence reflects a behavioral effect, shaping decisions through social cues rather than changes in a service's intrinsic attributes. It has been documented in various domains, including transportation (Stathopoulos et al., 2017; Manca et al., 2019).

Loss and risk aversion. Kahneman & Tversky (1979) introduced Prospect Theory as a model of decision-making under risk, where outcomes occur with objectively known probabilities (e.g., gambling). The framework was later extended to Cumulative Prospect Theory, which accommodates both risky and uncertain choices, the latter referring to situations in which outcome probabilities are unknown (Tversky & Kahneman, 1992).

Example. Consider two urns, each containing red and black balls. Urn A contains exactly 50 red balls and 50 black balls. If drawing a red ball yields \$100, then the probability of winning is objectively known to be 0.5. This setting illustrates decision-making under risk, since the relevant probabilities are explicitly specified. In contrast, suppose Urn B contains 100 balls that are either red or black, but the exact proportion is unknown. Drawing a red ball again yields \$100, yet in this case the probability of winning cannot be determined objectively (Ellsberg, 1961). This situation exemplifies decision-making under uncertainty.

Risk aversion means preferring a sure outcome over a risky gamble with the same expected payoff. In the urn example, consider a third option, Option C, which offers a sure payoff of 50\$. Risk-averse individuals will typically prefer Option C to Urn A, even though both yield the same expected monetary value of 50\$, because the certainty of payoff is valued more highly than the gamble. In Expected Utility Theory, this is captured by assuming a concave utility function, which reflects diminishing marginal utility of wealth (Pratt, 1978).

If the urn example is modified so that drawing a black ball results in a loss of \$100 instead of receiving nothing, most individuals deviate from the predictions of expected utility theory. While the expected monetary values are symmetric between gaining \$100 and losing \$100 with an equal number of black and red balls in Urn A, participants typically require the potential gain to be substantially larger than the potential loss (e.g., 200\$ for drawing a red ball) in order to accept the gamble. This behavior is known as *loss aversion* (Kahneman & Tversky, 1979).

Loss aversion plays a central role under both risk and uncertainty. To explain this behavior, Tversky and Kahneman proposed that individuals evaluate risky or uncertain choices through an editing and evaluation process. This process involves simplifying or transforming the decision problem prior to evaluation. In the editing phase, individuals do an initial evaluation and aim to get a simpler representation of options. In the final evaluation, people's edited evaluations are processed to choose the best option.

In the editing phase, people code options as gains and losses in their features relative to a reference point. This reference-dependent assessment could be affected by loss aversion, which implies that people outweigh losses against gains in their assessments when comparing alternatives to a reference (Kahneman et al., 1991). This loss-averse evaluation contradicts the independent axiom of rationality. Under risk, another behavioral effect arises that is overweighting low probabilities (Kahneman & Tversky, 1979).

Prospect Theory and its cumulative extension capture these behavioral patterns through two key components: (i) an asymmetric concave value function that is steeper in the loss domain; and (ii) a nonlinear probability weighting function, which accounts for the tendency to overweight small probabilities and underweight large probabilities (Kahneman & Tversky, 1979; Tversky & Kahneman, 1992).

In the context of shared-mobility services, both risk aversion and loss aversion are highly relevant. Ride outcomes such as travel time, waiting time, or detours are inherently probabilistic, whereas private solo rides provide greater certainty. Consequently, risk-aversion influences user choices when they choose between shared and solo rides. Moreover, loss aversion plays a role when users compare uncertain outcomes to reference points, such as expected travel time. Negative deviations from these reference levels (e.g., longer waiting times or higher fares) weigh more heavily than equivalent positive deviations. The loss aversion is also relevant in the presence of reference levels for some attributes, such as reference prices from previous experiences or other available options.

Context effect. Context effect refers to how the presence or structure of other options (composition or architecture of the choice set) influences choice. The main context effects that are acknowledged in different applications, as well as the transportation (Chorus & Bierlaire, 2013; Guevara & Fukushi, 2016), include the compromise and attraction (decoy) effects. *compromise effect* occurs when users gravitate toward a middle-ground option that appears more justifiable relative to extremes, especially under loss aversion (Simonson, 1989). The *attraction (or decoy) effect* arises when the introduction of an *asymmetrically dominated alternative*, an option that is dominated by one alternative in the set but not by others, increases the probability of choosing the dominating option. By simplifying trade-offs, such decoys enhance the relative appeal of the dominating alternative (Huber et al., 1982).

Context effects can be explained by the two main principles: trade-off contrasts and extremeness aversion. The former asserts that the choice between two items depends on their relative performance as well as the trade-offs between them and other available options. The attraction effect is the direct result of this principle. Huber et al. (1982) utilize Parducci's Range-Frequency theory (Parducci, 1974) to explain how the trade-off contrasts result in the attraction effect. Accordingly, the choice probability of a targeted option increases when the decoy option extends the range of attributes where the target is outperformed by the option that is compared to (i.e., range principle) or increases the frequency of its best attributes (i.e., frequency principle). In contrast, the extremeness aversion suggests the tendency to an option that has intermediate losses and gains compared to the other options, which is equivalent to the compromise effect (Simonson & Tversky, 1992).

Decisions in shared mobility, such as selecting a mode, provider, route, or specific option within a platform, are often non-repetitive. Available choices depend on variables like time of day, day of the week, season, and the specific origin and destination of the trip. This variability contributes to decision complexity and makes it difficult for users to form stable preferences. Additionally, in some situations, users need to choose between multiple options (e.g., choosing between multiple possible ride-matching options or bike-sharing to pick up). As a result, users are more susceptible to biases such as context effects.

2.5 Modeling Users' Decision-Making in Shared Mobility

This section is not intended to be a comprehensive literature review. Rather, our aim is to provide insight into how user decision-making is incorporated into providers' operations of shared mobility services. We focus on studies that: (i) explicitly model user decisions, (ii) address supply-side decisions (e.g., pricing, fleet management, service area design) within an optimization-based framework, and (iii) establish a direct link between user behavior and supply-side decision-making.

Mobility context. In this study, we examine the broad spectrum of shared mobility services, spanning both car-based and micromobility modes, and including electric as well as non-electric vehicles. Within this scope, we consider vehicle-sharing systems that operate either as station-based (docked) models, where vehicles must be picked up and returned to fixed locations, or as free-floating (dockless) models, where vehicles can be accessed and left anywhere within a defined service area. We also consider ride-sharing platforms, defined broadly as services enabling multi-passenger trips (Furuhata et al., 2013; Shaheen et al., 2016). Within this category, ride-sourcing (also referred to as ride-hailing) involves a traveler requesting a ride via a digital platform such as Uber or Lyft, whereas ride-pooling/ride-splitting refers to cases in which independent riders with overlapping routes are matched to share a vehicle. Together, these services are often situated in the literature under the broader concept of Mobility-on-Demand (MOD), also referred to as Demand Responsive Transit (DRT). MOD serves as an umbrella term for technology-enabled platforms and services that provide users with on-demand access to diverse transport options, including vehicle sharing, ride-sourcing, pooling, micromobility, and other flexible, connected modes (Shaheen & Cohen, 2020).

Classification of User Decisions. We classify the reviewed studies based on the specific user decisions they incorporate into supply-side modeling. These decisions fall into two main categories:

1. **Service selection (adoption):** This refers to a user's decision on whether to adopt a particular mobility service (e.g., bike-sharing, ride-hailing, or ride-sharing). Adoption determines the aggregated service demand.

2. **Alternative selection:** These are decisions users make during interaction with the service, often in response to dynamically presented options. This category includes:

    - **Single offer acceptance:** Deciding whether to accept a specific match proposed by the platform (e.g., an assigned driver or ride).
    - **Unfiltered selection:** Choosing freely among all feasible options (e.g., selecting any nearby bike in a free-floating bike-sharing system), rather than from a platform-curated subset.
    - **Menu-based selection (assortment):** Selecting from a limited, curated set of alternatives offered by the platform (e.g., different ride types or pickup time windows).
    - **Post-selection decisions:** Deciding to cancel a previously accepted service or wait longer when faced with delays or new options.
    - **Decisions under unavailability:** Responding to service failure in serving users (e.g., encountering a full or empty station), including walking to another station or abandoning the trip.

2.5.1 Service Adoption

Service adoption in shared mobility systems captures the user's decision to utilize a service, shaped by factors such as availability, pricing, and spatial coverage. These individual decisions are typically aggregated to estimate demand, which in turn informs provider-side planning, including fleet sizing, pricing, and rebalancing strategies.

Service attributes, such as service costs and convenience and reliability of access, play a central role in users' decisions to adopt a service (Zhang & Kamargianni, 2023; Ma et al., 2023). In practice, convenience is often measured through accessibility, i.e., the proximity of service stations to users' origin or destination, and reliability of access is typically measured through availability, i.e., the probability of finding a vehicle nearby. Several studies model adoption as a direct function of these attributes.

In the context of one-way electric car sharing, He et al. (2017) focus on location accessibility as the key service attribute. They assume users follow a satisficing rule, adopting the service if the total utility of their accessible trips exceeds a threshold. This threshold-based adoption model is embedded in a robust optimization framework that determines fleet size, service area, and repositioning strategies to maximize provider profit based on the expected number of adopters across locations. Moreover, fleet size is chosen as the minimum number of vehicles required to satisfy a predefined availability level, measured as the proportion of requests that can be served. They benchmark their integrated optimization model against heuristic fleet sizing and service area designs, showing that the integrated design is more profitable and resilient.

For free-floating bike-sharing systems, Wang et al. (2023) incorporate service availability, defined as the probability of finding a nearby bike, as a central attribute, with availability shaped by provider rebalancing actions. Aggregated location-specific demand is modeled as a linear function of availability and location characteristics such as population density and proximity to points of interest. This demand model is then embedded in a profit-maximizing optimization that determines vehicle routing and rebalancing volumes. By comparing models with and without availability and accessibility effects, they show that incorporating these factors improves operator profit, expands demand coverage, and reduces emissions.

In ride-hailing services, Chen et al. (2020) emphasize service cost and time attributes, fare, waiting time, and in-vehicle time, by formulating a nonlinear adoption model based on disutility. Disutility is expressed in monetary terms and depends on fare, waiting time, and in-vehicle time. Demand is then estimated as spatio-temporal arrival rates derived from these disutilities. This demand model feeds into a dynamic optimization of pricing and driver incentives, which simultaneously affects driver participation through expected income, while the platform seeks to maximize profit or social welfare subject to minimum income constraints. By comparing their model with static pricing and incentive schemes, Chen et al. (2020) show that considering a dynamic disutility-based adoption model in pricing decisions leads to higher profit and more reliable service.

Service attributes not only shape adoption in isolation but also determine demand when services compete with other modes or platforms. In competitive settings, users' adoption decisions are modeled by comparing the cost or utility of alternatives, each defined as a function of service attributes such as fare, travel time, waiting time, and availability. Thus, competition requires explicit modeling of how attributes across options are evaluated and traded off in user choice.

In dockless bike-sharing, Zheng et al. (2024) model user mode choice as the selection among shared bikes, walking, and motorized vehicles, where users choose the option with the lowest generalized travel cost. This cost aggregates service attributes such as fare, travel time, and access time. Provider decisions on fleet size, rebalancing, and pricing are then optimized under objectives such as profit maximization, ridership maximization (with profit constraints), or social welfare maximization, with demand modeled under a market equilibrium condition. This study compares two policies for influencing the market share of each mode: pricing strategies and fleet expansion. The analysis shows that pricing is more effective in raising demand and profit, whereas fleet growth yields diminishing returns.

Competition is also examined in the context of ride-sourcing services. Naumov & Keith (2023) develop a logit-based demand model to capture user mode choice among private rides, pooled rides, and outside options such as public transit or personal vehicles. User utility depends on fare, travel time, and the uncertainty associated with pooled travel, which is incorporated as an additional term in the utility function. On the supply side, providers adopt long-term pricing strategies to influence demand distribution across modes. The results of this study indicate that long-term pricing strategies increase pooled ride shares, sustain revenue, and reduce vehicle miles traveled.

Competition across ride-sourcing platforms is analyzed by Karamanis et al. (2021), who employ a logit-based model in which users choose between providers based on comparative attributes including price, waiting time, and travel time that are shaped by providers' operational decisions. Each provider optimizes their vehicle allocation across trip-serving, repositioning, and reserve functions to minimize operational costs while meeting the captured demand. They compare redistribution policies that account for platform competition with those that ignore it, and find that competition-aware models reduce inefficiencies in vehicle allocation.

While most studies link adoption directly to service attributes, some take a different approach by considering the role of social influence in service adoption decisions. Villa et al. (2024) study car-sharing adoption as a socially driven process, where an individual's likelihood of adopting increases with the number of peers (e.g., friends and neighbors) who have already done so. This is captured through a threshold-based function: an individual adopts the service at a specific time if the fraction of adopting peers exceeds a given threshold. This threshold can be reduced through pricing incentives. Instead of modifying service attributes directly, the provider intervenes through informational campaigns or subsidies that accelerate adoption. These interventions are optimized over time using an optimal control framework, with the objective of maximizing adoption under budget constraints. Their finding show that in contrast to uniform or untargeted incentives, targeted interventions exploiting peer effects accelerate adoption at lower cost.

Together, these studies provide valuable insights by formalizing how user adoption can be modeled and embedded within provider optimization frameworks. They highlight the diverse ways in which service attributes, behavioral mechanisms, and competitive dynamics can be captured to inform strategic decisions on pricing, fleet management, and service design.

However, these studies generally overlook the joint impact of service attributes and social influence on user adoption behavior. Service attributes define the quality of the service and determine whether it is attractive and feasible for individuals, while social influence shapes the adoption process, driving a gradual market penetration and adoption growth (Miranda et al., 2024; Manca et al., 2019). He et al. (2017); Wang et al. (2023); Chen et al. (2020); Zheng et al. (2024); Naumov & Keith (2023); Karamanis et al. (2021) design services based on the total expected demand, implicitly assuming that all demand is realized immediately. Villa et al. (2024), in contrast, incorporates social influence but neglects the role of service attributes in adoption decisions.

2.5.2 Alternative Selection

This section reviews studies that focus on users' responses during their interaction with the service, such as accepting a match, choosing between multiple offerings, or responding to incentives.

Single Offer Acceptance

When users request a ride or mobility service, the provider typically generates an option that is acceptable to the user.

One group of studies considers single options that match the user's request, but user acceptance is contingent on certain conditions. In ride-hailing, Li & Liu (2021) model acceptance through a satisficing rule: the user accepts if the total cost, including fare, detour time, and value of time, remains below a personalized aspiration threshold. This condition is embedded directly into the provider's dynamic ride-matching optimization, which only assigns rides that satisfy the acceptance threshold and jointly optimizes driver–rider matching, routing, and discount pricing over a rolling horizon to maximize profit. Compared with baseline ride-matching, their aspiration-threshold model shows that incorporating user acceptance thresholds improves profitability in real-world operations.

In shared rides, uncertainty about service attributes complicates the acceptance decision model. Actual travel time may differ from the user's expectation due to detours for serving multiple users who share their rides. To capture this, Guan et al. (2019) model acceptance decision using Cumulative Prospect Theory, where users consider some possible trip outcomes (e.g., trip time levels) and evaluate the outcomes of choosing the shared option relative to a reference point/level such as expected trip time. Users are assumed to be influenced by the loss aversion in these evaluations. Accordingly, an acceptance probability is computed for the shared option. The platform solves an inverse optimization problem to dynamically set fares such that acceptance probabilities meet system-wide targets. They benchmark prospect-theory-based acceptance against rational models, showing that accounting for reference dependence and loss aversion sustains acceptance under uncertain detours.

The other approach is to offer an option that deviates from the user's original request while remaining acceptable. Users may, for example, tolerate a delayed pickup time or a longer walk to the pickup location. Such flexibility can enhance operational efficiency for providers. In car-sharing, Ströhle et al. (2019) incorporate this flexibility and model user reaction to deviations in time and location. It is assumed that users accept the offered option (with deviations) if its utility loss relative to the original request remains below a threshold. This study uses a multinomial mixed logit framework to estimate sensitivities to price, walking distance, and temporal deviation, which is calibrated to stated preference data. The provider incorporates users' acceptance thresholds into a mixed-integer programming framework that minimizes the fleet size and involves rebalancing and matching constraints. Relative to models without flexibility, incorporating acceptance flexibility enables reassignment and rebalancing that reduce fleet size by up to 30% in their numerical experiments.

User loss aversion becomes particularly relevant when deviations from the original request are introduced, since the initial request could serve as a reference point. Wang et al. (2022) incorporate this behavior in the context of demand-responsive transit, where users evaluate the single offered option with a delay relative to their original requests. User decisions are modeled using Prospect Theory, which captures asymmetric time sensitivity: late departures are penalized more heavily than early ones. A delayed departure is accepted if the reference-dependent perceived cost remains below the user's acceptance threshold. This model is integrated into the provider's routing, discount design, and scheduling optimization, formulated as a multi-objective framework that optimizes profit and user cost. They show that including the slack departures with discounts improves operator profit (+63%), increases vehicle load (+30%), and reduces user travel cost (–12%) compared with a no slack case.

Even if uncertainty about service attributes is neglected or deviations from the user's original request are not considered, loss aversion may still be a relevant behavioral effect, as individuals' previous trip experiences or their service expectations could act as a reference point and shape their decisions, a factor not addressed in the reviewed studies. Moreover, in single-match studies, user choice is typically framed as a binary accept–reject decision: the user accepts if the option satisfies their aspirations or requirements, and rejects otherwise. In practice, however, rejection may not simply mean abandoning the service but instead opting for a competing alternative.

Unlike the studies discussed above, where users are presented with a single option, the following studies consider settings in which users choose from a set of options. These studies fall into two categories: those where the option set is exogenously given, and those where the service provider actively decides which subset or menu of options to present to users.

Unfiltered Selection

In this group, the set of available options is not controlled or filtered by the provider; rather, users face all feasible options and make a choice among them. The provider must therefore predict user choice behavior and incorporate it into operational planning. Providers may also adjust attributes of the alternatives, most commonly through pricing, to influence user decisions and optimize system objectives such as balancing vehicle availability, increasing mode share, or managing energy load in electric services.

In free-floating vehicle-sharing systems, users choose among available nearby vehicles, and their choices can be affected by real-time pricing decisions aimed at improving vehicle availability and fleet utilization. Müller et al. (2023) study this problem by modeling users' choices using a Multinomial Logit (MNL) framework, where the utility of each reachable vehicle depends on walking distance and rental price. These utilities are then translated into probabilistic demand, which feeds directly into an origin-based dynamic pricing framework. The provider employs Approximate Dynamic Programming (ADP) to adjust prices per vehicle at each decision epoch, using nonparametric value function approximations trained on historical data to anticipate future vehicle distribution and demand. Against operator-centric pricing, their logit-based demand model supports customer-centric pricing that increases profit by up to 8%.

Users of free-floating bike-sharing services may face multiple feasible mode–route combinations (i.e., path). Zhang et al. (2019) model these decisions using a user equilibrium framework, where users choose their path, including cycling, walking, or bus, that minimizes their generalized travel cost. In this study, the provider offers location- and time-specific monetary incentives to influence choices, motivating users to take bikes from oversupplied areas and return them to undersupplied ones to improve the supply–demand balance. These incentives are adjusted in real time based on observed imbalances, linking user route choices to vehicle availability. Introducing location- and time-based incentives in this study induces voluntary rebalancing and improves bike distribution.

In multimodal transportation systems, users face simultaneous choices across access and mainline modes. Fournier et al. (2021) model this case for a single multimodal provider that sets prices for both access modes (walk, bike, drive) and mainline modes (train, highway) to manage the demand between modes. The objective is to design coordinated pricing policies (for access and mainline modes) that shift the aggregate modal split toward an equilibrium minimizing total travel time. Users make two sequential choices: first, how to access the mainline (access mode); second, which mainline mode to use. Users minimize their generalized disutility, calculated as the sum of travel time and monetary cost, with time converted to monetary terms using user-specific value-of-time parameters. According to the numerical experiments in this study, coordinated pricing reduces generalized travel time (–22%) and inequality (–30%) compared with flat or uncoordinated fares.

In electric mobility services, users also face multiple feasible charging options, deciding either when or where to charge, and providers must anticipate these decisions to balance demand across space and time. Vuelvas et al. (2020) model temporal charging choices in response to day-ahead time-of-use (TOU) electricity prices. Users are grouped into clusters based on charging schedules and are assumed to minimize their charging costs over a 24-hour horizon. The provider sets TOU prices to shift aggregate charging patterns, formulated as a bi-level optimization problem in which the provider maximizes expected profit under uncertainty while users minimize individual costs. Dynamic TOU pricing, compared with flat rates, raises aggregator profit (+15.6%) and shifts charging to off-peak hours. Yang et al. (2021) focus on spatial charging choices across available locations. In this study, users evaluate all feasible charging stations, with utilities depending on price, charger availability, congestion, and personal preferences. User choice is modeled with an MNL framework, while the provider dynamically adjusts zone-level prices in a rolling-horizon framework to steer demand. The objective is to maximize system utility, defined as revenue net of congestion and unmet demand penalties. They benchmark dynamic zone pricing against static tariffs and show it reduces waiting times, balances demand, and increases welfare.

Menu-Based Selection

In this group of studies, service providers do not expose users to all feasible service alternatives but instead construct a curated menu of options based on real-time constraints and anticipated user choices.

Providers need to generate feasible alternatives and select the best subset to present via an online platform, which requires operational decisions on routing, pricing, matching, and assignment. User behavior is typically modeled using discrete choice frameworks, where utilities depend on alternatives' attributes such as fare, travel time, and are translated into selection probabilities. Atasoy et al. (2015) study this problem for mobility-on-demand platforms, where each user request is responded to by a menu of options, including solo taxi, shared taxi, and mini-bus, based on the user's origin, destination, and request time. The problem is formulated as a two-step optimization: first, feasible service options are generated through real-time routing and vehicle assignment; then, an assortment optimization problem is solved to select the subset of options that maximizes expected profit. User choice from the menu is modeled using an MNL framework, where utility depends on fare, in-vehicle time, access and egress time, schedule delay, and detour. Compared with static menus, optimized dynamic menus raise profit by up to 74%.

Building on this concept for ride-sharing platforms, Azadeh et al. (2022) incorporate pickup delays in generating feasible options subject to the condition that any offered option must meet a minimum acceptance probability. Ride options are defined by fare, pickup time windows, expected delay, and in-vehicle travel time. As in Atasoy et al. (2015), user choice is modeled with an MNL framework, but the optimization model is a choice-based Dial-a-Ride Problem (DARP) that jointly considers routing, pricing, and menu selection. Delays are compensated by price adjustments to maintain option acceptability, improving vehicle utilization while preserving user acceptance. Relative to classic DARP, their choice-based DARP with assortment optimization for menu selection serves more customers, reduces routing costs, and increases profit.

Other studies treat the provider's menu and external alternatives as separate choice nests, where users first select a nest (e.g., solo ride vs. shared ride) and then choose an option within that nest. In this category, Guo et al. (2025) examine a high-capacity ride-sharing system where user decisions are modeled using a Nested Logit model with three nests: shared options (e.g., buspooling, customized buses), solo rides, and opt-out alternatives such as other modes or providers. The provider generates the menu of shared options (the shared nest), while solo and opt-out options form the other nests. Utilities are defined over fare, travel time, and comfort. The platform dynamically constructs shared ride alternatives and adjusts pricing, for both new and on-board passengers, to compensate for detours or reward emissions savings. The objective is to maximize marginal profit within a rolling-horizon optimization framework. Compared with traditional ride-hailing, high-capacity pooling reduces operator costs (–10.4%), lowers fares, and cuts emissions.

As mentioned earlier, in shared-mobility services, users often face uncertainty in service attributes, such as walking distance to the vehicle, trip duration, and final price, due to dynamic vehicle locations, variable traffic conditions, and system-driven pricing policies. Wu et al. (2022) model this uncertainty in modeling user behavior for a free-floating car-sharing context by treating these attributes as random variables, represented through triangular distributions. Utilities are formulated using nonlinear functions to reflect risk aversion behavior, and embedded in an MNL model to capture probabilistic user responses. Anticipating these responses, the provider optimizes prices across time, location, and user types within a rolling-horizon framework, aiming to maximize revenue or welfare. Relative to rational (risk-neutral) assumptions, modeling risk aversion avoids revenue losses of up to 43% and yields more robust pricing policies.

Jacob & Roet-Green (2021) add the perspective of freelance drivers in analyzing ride-sharing platforms. The study examines how the platform jointly designs a service menu (solo, pooled, or both) and sets prices to maximize revenue, accounting for passenger and driver responses in a static equilibrium model. Drivers are modeled as homogeneous agents who decide whether to participate by comparing expected earnings with a reservation wage. Drivers do not choose among requests and they serve whichever requests are assigned. Passenger behavior is modeled through deterministic utility maximization, with each user choosing among solo, pooled, or opting out. Utility consists of four components: (i) a mode-specific constant, (ii) delay cost determined via a queueing model, (iii) an inconvenience cost of sharing, and (iv) network effects. The last component creates interdependence across users and drivers: as more passengers choose pooling or more drivers participate, delays and match probabilities change, which in turn affects users' utilities and driver participation. The platform's decision problem is to select the service menu and prices that maximize revenue, subject to the joint equilibrium of passenger and driver decisions. Benchmarking alternative menu designs, they show that including pooling options increases their share and revenue under the joint passenger–driver equilibrium.

This body of work provides an essential foundation for platforms that design a menu of service options. Selecting a menu of service options based on estimated user preferences is already a valuable tool for platforms. Compared with presenting a single option, a menu gives users more flexibility; compared with offering all feasible options, it allows the platform to retain control over demand.

User choice in shared mobility is typically modeled with rationality-based discrete choice frameworks, most commonly Multinomial or Nested Logit. Beyond this rational optimization perspective, menu design can also act as a behavioral lever. The composition of alternatives in the menu may influence user decisions through context effects, yet such effects have not been systematically explored in the shared mobility literature. Unlike pricing adjustments, which reduce already low profit margins, careful design of the service menu could exploit context effects to steer demand. Moreover, these effects could be combined with pricing strategies or integrated into alternative generation, offering platforms additional tools for influencing user behavior without relying solely on fare manipulation.

Post-Selection Decision; Cancel or Wait

When users are waiting for the service after placing their order, they face the decision of whether to continue waiting or drop out, as they may have other alternatives (e.g., public transit or a taxi when waiting for their vehicle).

Accepting users likely to cancel leads to wasted capacity and poor system performance. For on-demand shared mobility services, Dong et al. (2022) develop a chance-constrained filtering mechanism that accepts only those user requests for which the expected utility of the shared service exceeds that of a private option (as an alternative when waiting for the shared ride) with high probability. Utility depends on fare, in-vehicle time, detours, and schedule delay, and the platform solves a mixed-integer linear program that jointly optimizes pricing, routing, and acceptance, with the objective of maximizing expected profit while avoiding low-satisfaction users. Their chance-constrained model filters low-utility requests, reducing cancellations and improving profit and reliability.

User cancellations can be reduced through pricing strategies. Wang et al. (2019) examine this approach in ride-sourcing, focusing on real-time cancellations where users may switch to a street-hail taxi while waiting for the assigned vehicle. User behavior is modeled with a threshold rule: a cancellation occurs if the perceived cost of switching, including fare difference, waiting time, psychological and safety factors, and any cancellation penalty, is lower than the cost of continuing with the original ride. Rather than filtering demand, the provider uses this threshold-rule model to design pricing schemes and cancellation penalties, optimizing either platform profit or social welfare by discouraging costly cancellations. They compare cancellation-aware versus cancellation-ignorant policies, showing that threshold-based modeling guides pricing and penalties that reduce cancellations and increase profit/welfare.

In ride-sharing platforms with freelance drivers, variation in driver working patterns influences user waiting time and, consequently, cancellation rates. User cancellations, in turn, affect the likelihood of successful matches and shape drivers' income and working patterns. Wang et al. (2024) incorporate this interdependency between users and drivers by modeling user decisions through a discrete choice framework that compares ride-sharing with public transit based on fare, waiting time, and travel duration. Driver behavior is represented using a Gaussian Mixture Model, which clusters their working patterns (e.g., part-time vs. full-time). Cancellation probability then feeds back into drivers' working patterns. These dynamics are integrated into a reinforcement learning framework that jointly optimizes dispatching, repositioning, and incentive strategies to reduce cancellations and improve revenue. This study shows that reinforcement learning that incorporates cancellation behavior raises matching rates (2–6%), increases revenue (3–9%), and reduces cancellations.

These studies provide valuable operational strategies for platforms by incorporating user cancellation behavior into service design, thereby mitigating its negative impacts. Yet, behavioral aspects such as risk aversion and loss aversion are also highly relevant, as they influence how users perceive waiting uncertainty and evaluate available alternatives. While waiting for a vehicle, an alternative option such as a passing taxi or a nearby bus is a certain, immediately available choice. In contrast, the expected waiting time for the assigned ride remains uncertain, making continued waiting a risky prospect. Risk-averse users may therefore prefer switching to the certain alternative rather than facing the uncertainty of extended waiting. At the same time, available alternatives can serve as reference points in users' evaluations.

Decision in the Case of Service Unavailability; Leave or Switch

Service unavailability, such as encountering an empty origin station or a full destination station, poses a critical challenge in station-based bike-sharing systems. From the platform's perspective, failure to anticipate user responses to such disruptions can lead to ineffective rebalancing and unmet demand. Integration of user behavior under these conditions could improve providers' resource allocation and inventory planning decisions.

When a user arrives at an empty origin station or a full destination station, they must decide whether to walk to a nearby station or abandon the trip. These decisions directly influence the distribution of resources across the system, making it essential to anticipate user behavior.

Datner et al. (2019) model user choices using a deterministic framework, where individuals select the option that minimizes additional travel time relative to an ideal journey, a measure termed journey dissatisfaction. These behavioral rules are integrated into an overnight inventory planning model that employs Guided Local Search to optimize bike allocations, with the objective of minimizing total system dissatisfaction. By comparing rebalancing with and without user continuation (walking to nearby stations), Datner et al. (2019) show that accounting for continuation reduces failed trips and improves reliability. Instead of a deterministic method, Costa Affonso et al. (2021) represent user decisions probabilistically, modeling the likelihood of searching for a bike or dock at a nearby station as a function of walking distance. These probabilities, estimated from stated preference data, are incorporated into a static rebalancing optimization model that minimizes lost demand by aligning inventory decisions with user expectations over a single planning horizon. Relative to ignoring spillover, probabilistic continuation modeling improves coverage and reduces lost demand in station-based bike sharing.

2.5.3 Summary

Table 2.1 summarizes the reviewed references by mobility context, behavioral modeling approach, provider objectives, decisions, and key findings. The reviewed studies show that incorporating user decisions consistently improves provider outcomes, by raising profit, improving coverage, reducing cancellations, or lowering emissions. Nevertheless, representations of user behavior remain limited, as most studies rely on rational choice or simple threshold rules, with only a few explicitly capturing behavioral effects such as loss aversion, risk aversion, or social influence.

Table 2.1: Summary of reviewed references by behavioral model, effect, objective, formulation, and findings.

Reference	Mobility context	User decision model	Behavioral effect	Objective	Formulation & decisions	Findings
User Decision: *Service selection (adoption)*						
He et al. (2017)	EV-sharing (FF)	Threshold satisficing: adopt if access utility $\geq$ aspiration	None	Maximize profit	MISOCP; fleet sizing, service area, repositioning	Integrated design improves profitability.
Wang et al. (2023)	Bike-sharing (FF)	Linear demand function (availability & accessibility)	None	Maximize profit	MILP; routing, rebalancing	Considering availability and accessibility increases profit, improves coverage, and lowers emissions.
Chen et al. (2020)	Ride-hailing	Disutility-based arrival rates (fare, wait, ride time)	None	Maximize profit / welfare	DP; dynamic pricing, driver incentives	Incorporating disutility-based adoption in pricing decisions improves profitability.
Zheng et al. (2024)	Bike-sharing (FF)	Deterministic cost minimization (fare, travel time, access)	None	Maximize profit / ridership / welfare	Market equilibrium; pricing, fleet sizing	Pricing is more effective than fleet growth.
Naumov & Keith (2023)	Ride-sourcing	Logit mode choice (fare, time, uncertainty)	None	Maximize revenue and pooled share	DP; long-term pricing	Long-term pricing increases pooling adoption, sustains revenue, and reduces VMT.
Karamanis et al. (2021)	Ride-sourcing (autonomous)	Logit platform choice (fare, waiting, and travel time)	None	Minimize operating cost	Min-cost flow; vehicle allocation	Competition-aware redistribution improves allocation efficiency compared with ignoring competition.
Villa et al. (2024)	Car-sharing	Peer adoption threshold	Social influence	Minimize intervention cost	Optimal control; interventions	Targeted policy interventions informed by social effect accelerate adoption at lower cost compared with uniform incentives.
User Decision: *Single offer acceptance*						
Li & Liu (2021)	Ride-hailing	Threshold rule: accept if cost $\leq$ aspiration	None	Maximize profit	Rolling-horizon MILP; matching, routing, pricing	Incorporating users acceptance decision improves profitability.
Guan et al. (2019)	Ride-pooling	Cumulative Prospect Theory; reference-dependent valuation of uncertainty	Loss aversion	Maintain acceptance probability	Inverse optimization; dynamic fare adjustment	Modeling loss aversion sustains user acceptance compared with rational assumptions.

Reference	Mobility context	User decision model	Behavioral effect	Objective	Formulation & decisions	Findings
Ströhle et al. (2019)	Car-sharing	Mixed logit (fare, walk, deviation)	None	Minimize fleet size	MIP; reassignment, rebalancing	Considering users' time/location flexibility reduces fleet size.
Wang et al. (2022)	MOD	Prospect Theory; lateness sensitivity	Loss aversion	Maximize profit + minimize delay cost	Multi-objective MILP; routing, scheduling, discounts	Including the slack departures with discounts increases profit, load, and reduces user cost.
User Decision: *Selection among all alternatives (Unfiltered selection)*						
Müller et al. (2023)	Vehicle-sharing (FF)	Logit (walk distance, fare)	None	Maximize profit	ADP; dynamic pricing	Customer-centric pricing improves operator profit compared with operator-centric pricing.
Zhang et al. (2019)	Bike-sharing (FF)	Deterministic disutility minimization (fare, time, comfort)	None	Minimize imbalance	User equilibrium; relocation incentives	Incentives induce voluntary rebalancing compared with no incentives.
Fournier et al. (2021)	Multimodal transit	Deterministic disutility minimization (fare, travel time)	None	Minimize travel time and inequality	Convex optimization; coordinated pricing	Coordinated pricing reduces travel time and inequality compared with flat pricing.
Vuelvas et al. (2020)	EV charging	Deterministic cost minimization	None	Maximize profit	Bi-level MILP; aggregator pricing and scheduling	TOU pricing increases profit and shifts demand off-peak compared with flat tariffs.
Yang et al. (2021)	EV charging	Logit (price, congestion, availability)	None	Maximize welfare	Rolling-horizon optimization; zone pricing	Zone-based pricing reduces waiting and increases welfare compared with static pricing.
Wu et al. (2022)	Car-sharing (FF)	Logit with nonlinear utility (fare, walk, travel time)	Risk aversion	Maximize profit / welfare	Rolling-horizon dynamic pricing	Modeling risk aversion prevents revenue loss compared with rational demand assumptions.
User Decision: *Choosing from the provider's selected menu (Menu-based selection)*						
Atasoy et al. (2015)	MOD	Logit (fare, time, detour)	None	Maximize profit	Assortment optimization with routing	Dynamic menu selection increases profit.
Azadeh et al. (2022)	Ride-sharing	Logit (fare, pickup window, delay, ride time)	None	Maximize profit	MILP; choice-based DARP	Choice-based dial-a-ride increases served demand and profit compared with classic models.

Reference	Mobility context	User decision model	Behavioral effect	Objective	Formulation & decisions	Findings
Jacob & Roet-Green (2021)	Ride-sharing	Deterministic utility (fare, delay, sharing cost, network effects)	None	Maximize revenue	Static equilibrium; pricing and menus	Offering pooling in menus increases their usage and revenue compared with solo-only menus.
Guo et al. (2025)	High-capacity ride-sharing	Nested logit (are, time, comfort)	None	Maximize profit	Rolling-horizon matching and pricing	High-capacity pooling reduces cost, fares, and emissions compared with solo rides.
User Decision: *Post-selection; wait or cancel*						
Dong et al. (2022)	MOD	Chance-constrained threshold: accept service if utility $\geq$ private option with a confidence	None	Maximize profit	MILP; request filtering, routing, pricing	Chance-constrained acceptance modeling reduces cancellations and increases profit.
Wang et al. (2019)	Ride-sourcing	Threshold (leave vs wait); cancel if switching is cheaper	None	Maximize profit / welfare	Pricing and penalty design	Cancellation-aware pricing and penalties reduce cancellations and improve outcomes.
Wang et al. (2024)	Ride-hailing	Logit cancel vs transit	None	Maximize profit	Multi-agent RL; dispatch, repositioning, incentives	Reinforcement learning with cancellation modeling increases matching rate and revenue.
User Decision: *choose to leave or switch under unavailability*						
Datner et al. (2019)	Bike-sharing (SB)	Threshold rule: walk if extra time $\leq$ tolerance	None	Minimize dissatisfaction	GLS; overnight inventory optimization	Considering user decision in the case of unavailability reduces failed trips and improves reliability.
Costa Affonso et al. (2021)	Bike-sharing (SB)	Probabilistic rule; walk vs abandon by distance	None	Minimize lost demand	Static rebalancing optimization	Considering user decision in the case of unavailability reduces lost demand and improves coverage.

FF = free-floating; SB = station-based; EV = Electric vehicle; MILP = Mixed-Integer Linear Program; MIP = Mixed-Integer Program; MISOCP = Mixed-Integer Second-Order Cone Program; DP = Dynamic Programming; ADP = Approximate Dynamic Programming; RL = Reinforcement Learning; GLS = Guided Local Search; TOU = Time-of-Use; VMT = Vehicle Miles Traveled.

2.6 Research Gaps and Directions

Our review underscores that incorporating user decisions is essential, as it improves service efficiency and performance in terms of profit, coverage, number of served users, and emissions reduction. These outcomes directly address the persistent challenges of adoption, profitability, and environmental impact that shared mobility services continue to face (Lagadic et al., 2019; Boer et al., 2022). Extending models that explicitly incorporate behavioral effects in users' decisions offers even greater potential. Yet, despite their demonstrated relevance, behavioral effects such as loss aversion, risk aversion, and social influence have been largely overlooked in operational modeling, leaving a critical research gap between behavioral insights and operational models in the shared mobility context.

Incorporating behavioral effects into shared mobility modeling introduces both conceptual and algorithmic challenges for supply-side optimization. Supply-side decision-making in shared mobility systems is inherently complex. Generating feasible service options involves solving combinatorial problems, such as vehicle routing, dynamic scheduling, and real-time pricing, that are typically NP[1]-hard and must be solved under time constraints. To manage these complexities, many existing models rely on rational choice frameworks, particularly the MNL model, which offer analytical convenience due to their assumptions of context-independent, transitive preferences. However, these assumptions often fail to capture empirically observed behaviors such as context effects, social influence, loss and risk aversion, leading to suboptimal operational decisions.

In focusing on behavioral effects that are directly relevant and actionable for providers' operations, we identify two key gaps. These gaps, and their implications, are discussed below.

1. **From potential to realized demand.** Shared mobility service design, including decisions about station locations, fleet size, and pricing, is typically guided by potential demand estimates, assuming that users will adopt the service immediately upon its launch. In practice, however, adoption occurs gradually, as users become aware of the service, evaluate its value, and observe adoption by others. This progression is influenced by both individual preferences and social influence (Rogers et al., 2014; Bass, 1969; Stathopoulos et al., 2017).

   Service design plays a crucial role in shaping this adoption path: ineffective design or early service failures can discourage early adopters, slow down diffusion, and ultimately reduce realized demand. In contrast, well-designed services that align with user expectations can accelerate uptake and amplify positive social influence. Incorporating adoption behavior and gradual market penetration into planning models allows for more effective resource allocation and improved long-term profitability. This perspective shifts transportation planning from a static view of demand to one that explicitly optimizes for dynamic, time-dependent growth.

2. **Modeling context effects in menu design.** In digital mobility platforms, users are often presented with a menu of service alternatives, such as solo or shared rides, flexible pickup times, or different price levels. Behavioral research shows that the structure and composition of this choice set can systematically influence preferences, making certain options

[1]Non-polynomial time

more attractive not only because of their intrinsic attributes but also because of their relative position within the set. Accordingly, strategic inclusion or exclusion of alternatives can therefore steer users toward a particular option.

Although such context effects (e.g., decoy/attraction and compromise effects) have been empirically demonstrated in mobility settings (Chorus & Bierlaire, 2013; Guevara & Fukushi, 2016; Ding et al., 2025), they remain largely absent from operational models.

A promising modelling framework for capturing these effects is assortment optimization, which enables embedding user choice models that are sensitive to the offered set (e.g., incorporating decoy, compromise, or dominance effects). Assortment optimization allows planners to design the composition of menus aligned with supply-side goals such as profit maximization, load balancing, or sustainability. In the digital era of mobility platforms, assortment optimization offers a mechanism for using behavioral effects to motivate users toward shared, off-peak, or low-emission options without needing to alter service attributes.

2.7 Conclusion

This study highlights that user decision-making plays a central role in determining the operational outcomes of shared mobility systems, including profitability, adoption, and sustainability. To capture this interdependence between user behavior and provider operations, we proposed a conceptual framework that links the two and provides a structured lens for analyzing existing research from both behavioral and operational perspectives. Drawing on behavioral studies, we identified the factors most relevant to provider operations and examined which user decisions are considered and how they are modeled in the literature. The analysis revealed critical gaps, most notably the reliance on rational choice and the limited integration of behavioral effects such as loss aversion, context effects, and social influence. These findings provide a foundation for advancing behaviorally richer operational models. They also offer actionable guidance on how behavioral factors can be used to improve service performance, reduce inefficiencies, and align user choices with policy goals such as sustainability.

Bibliography

Agatz, N., A. Erera, M. Savelsbergh, X. Wang (2012) Optimization for dynamic ride-sharing: A review, *European Journal of Operational Research*, 223(2), pp. 295–303.

Aguilera-García, Á., J. Gomez, N. Sobrino (2020) Exploring the adoption of moped scooter-sharing systems in spanish urban areas, *Cities*, 96, p. 102424.

Albiński, S., P. Fontaine, S. Minner (2018) Performance analysis of a hybrid bike sharing system: A service-level-based approach under censored demand observations, *Transportation Research Part E: Logistics and Transportation Review*, 116, pp. 59–69.

Ariely, D., S. Jones (2008) *Predictably irrational*, HarperCollins New York.

Ariely, D., T. S. Wallsten (1995) Seeking subjective dominance in multidimensional space: An explanation of the asymmetric dominance effect, *Organizational Behavior and Human Decision Processes*, 63(3), pp. 223–232.

Atasoy, B., T. Ikeda, X. Song, M. E. Ben-Akiva (2015) The concept and impact analysis of a flexible mobility on demand system, *Transportation Research Part C: Emerging Technologies*, 56, pp. 373–392.

Azadeh, S. S., B. Atasoy, M. E. Ben-Akiva, M. Bierlaire, M. Maknoon (2022) Choice-driven dial-a-ride problem for demand responsive mobility service, *Transportation Research Part B: Methodological*, 161, pp. 128–149.

Bachand-Marleau, J., B. H. Lee, A. M. El-Geneidy (2012) Better understanding of factors influencing likelihood of using shared bicycle systems and frequency of use, *Transportation Research Record*, 2314(1), pp. 66–71.

Badia, H., E. Jenelius (2023) Shared e-scooter micromobility: review of use patterns, perceptions and environmental impacts, *Transport reviews*, 43(5), pp. 811–837.

Bass, F. M. (1969) A new product growth for model consumer durables, *Management Science*, 15(5), pp. 215–227.

Bendoly, E., K. Donohue, K. L. Schultz (2006) Behavior in operations management: Assessing recent findings and revisiting old assumptions, *Journal of Operations Management*, 24(6), pp. 737–752.

Bettman, J. R., M. F. Luce, J. W. Payne (1998) Constructive consumer choice processes, *Journal of Consumer Research*, 25(3), pp. 187–217.

Boer, M., O. Türetken, O. E. Adali (2022) A review of business models for shared mobility and mobility-as-a-service (maas): a research report.

Chen, X. M., H. Zheng, J. Ke, H. Yang (2020) Dynamic optimization strategies for on-demand ride services platform: Surge pricing, commission rate, and incentives, *Transportation Research Part B: Methodological*, 138, pp. 23–45.

Chernev, A. (2003) When more is less and less is more: The role of ideal point availability and assortment in consumer choice, *Journal of Consumer Research*, 30(2), pp. 170–183.

Chorus, C. G., M. Bierlaire (2013) An empirical comparison of travel choice models that capture preferences for compromise alternatives, *Transportation*, 40(3), pp. 549–562.

Cialdini, R. B., N. J. Goldstein (2004) Social influence: Compliance and conformity, *Annual Review of Psychology*, 55(1), pp. 591–621.

Costa Affonso, R., F. Couffin, P. Leclaire (2021) Modelling of user behaviour for static rebalancing of bike sharing system: Transfer of demand from bike-shortage stations to neighbouring stations, *Journal of Advanced Transportation*, 2021(1), p. 8825521.

Croson, R., K. Schultz, E. Siemsen, M. Yeo (2013) Behavioral operations: the state of the field, *Journal of Operations Management*, 31(1-2), pp. 1–5.

Datner, S., T. Raviv, M. Tzur, D. Chemla (2019) Setting inventory levels in a bike sharing network, *Transportation Science*, 53(1), pp. 62–76.

Debreu, G., et al. (1954) Representation of a preference ordering by a numerical function, *Decision Processes*, 3, pp. 159–165.

Deutsch, M., H. B. Gerard (1955) A study of normative and informational social influences upon individual judgment., *The Journal of Abnormal and Social Psychology*, 51(3), p. 629.

Ding, J., E.-J. Kim, V. Maksimenko, P. Bansal (2025) Can decoy effects nudge ride-hailing drivers' preferences for electric vehicles?, *Transportation Research Part A: Policy and Practice*, p. 104470.

Dong, X., J. Y. Chow, S. T. Waller, D. Rey (2022) A chance-constrained dial-a-ride problem with utility-maximising demand and multiple pricing structures, *Transportation Research Part E: Logistics and Transportation Review*, 158, p. 102601.

Ellsberg, D. (1961) Risk, ambiguity, and the savage axioms, *The Quarterly Journal of Economics*, 75(4), pp. 643–669.

Faghih-Imani, A., R. Hampshire, L. Marla, N. Eluru (2017) An empirical analysis of bike sharing usage and rebalancing: Evidence from barcelona and seville, *Transportation Research Part A: Policy and Practice*, 97, pp. 177–191.

Ferrero, F., G. Perboli, M. Rosano, A. Vesco (2018) Car-sharing services: An annotated review, *Sustainable Cities and Society*, 37, pp. 501–518.

Fournier, N., E. Christofa, E. J. Gonzales (2021) A continuous model for coordinated pricing of mixed access modes to transit, *Transportation Research Part C: Emerging Technologies*, 128, p. 103208.

Furuhata, M., M. Dessouky, F. Ordóñez, M.-E. Brunet, X. Wang, S. Koenig (2013) Ridesharing: The state-of-the-art and future directions, *Transportation Research Part B: Methodological*, 57, pp. 28–46.

Gino, F., G. Pisano (2008) Toward a theory of behavioral operations, *Manufacturing & Service Operations Management*, 10(4), pp. 676–691.

Guan, Y., A. M. Annaswamy, H. E. Tseng (2019) Cumulative prospect theory based dynamic pricing for shared mobility on demand services, in: *2019 IEEE 58th Conference on Decision and Control (CDC)*, IEEE, pp. 2239–2244.

Guevara, C. A., M. Fukushi (2016) Modeling the decoy effect with context-rum models: Diagrammatic analysis and empirical evidence from route choice sp and mode choice rp case studies, *Transportation Research Part B: Methodological*, 93, pp. 318–337.

Guo, R., X. Liu, Y. Sun, X. Yan, W. Guan, S. S. Azadeh (2025) From ride-hailing to high-capacity ride-sharing: a user-centric shared mobility service design, *Transportmetrica A: Transport Science*, pp. 1–31.

He, L., H.-Y. Mak, Y. Rong, Z.-J. M. Shen (2017) Service region design for urban electric vehicle sharing systems, *Manufacturing & Service Operations Management*, 19(2), pp. 309–327.

Ho, S. C., W. Y. Szeto, Y.-H. Kuo, J. M. Leung, M. Petering, T. W. Tou (2018) A survey of dial-a-ride problems: Literature review and recent developments, *Transportation Research Part B: Methodological*, 111, pp. 395–421.

Hsu, C.-L., H.-P. Lu (2004) Why do people play on-line games? an extended tam with social influences and flow experience, *Information & Management*, 41(7), pp. 853–868.

Huber, J., J. W. Payne, C. Puto (1982) Adding asymmetrically dominated alternatives: Violations of regularity and the similarity hypothesis, *Journal of Consumer Research*, 9(1), pp. 90–98.

Huber, J., J. W. Payne, C. P. Puto (2014) Let's be honest about the attraction effect, *Journal of Marketing Research*, 51(4), pp. 520–525.

Jacob, J., R. Roet-Green (2021) Ride solo or pool: Designing price-service menus for a ride-sharing platform, *European Journal of Operational Research*, 295(3), pp. 1008–1024.

Kahneman, D., J. L. Knetsch, R. H. Thaler (1991) Anomalies: The endowment effect, loss aversion, and status quo bias, *Journal of Economic Perspectives*, 5(1), pp. 193–206.

Kahneman, D., A. Tversky (1979) Prospect theory: An analysis of decision under risk, in: *Handbook of the fundamentals of financial decision making: Part I*, World Scientific, pp. 99–127.

Karamanis, R., E. Anastasiadis, M. Stettler, P. Angeloudis (2021) Vehicle redistribution in ride-sourcing markets using convex minimum cost flows, *IEEE Transactions on Intelligent Transportation Systems*, 23(8), pp. 10287–10298.

Lagadic, M., A. Verloes, N. Louvet (2019) Can carsharing services be profitable? a critical review of established and developing business models, *Transport Policy*, 77, pp. 68–78.

Li, Y., Y. Liu (2021) Optimizing flexible one-to-two matching in ride-hailing systems with boundedly rational users, *Transportation Research Part E: Logistics and Transportation Review*, 150, p. 102329.

Liao, F., G. Correia (2022) Electric carsharing and micromobility: A literature review on their usage pattern, demand, and potential impacts, *International Journal of Sustainable Transportation*, 16(3), pp. 269–286.

Luce, R. D., et al. (1959) *Individual choice behavior*, vol. 4, Wiley New York.

Ma, L., X. Zhang, F. Hao, S. Liu (2023) Service quality, perceived price fairness, and users' continuous usage intentions regarding shared bike service, *Total Quality Management & Business Excellence*, pp. 1–21.

Manca, F., A. Sivakumar, J. W. Polak (2019) The effect of social influence and social interactions on the adoption of a new technology: The use of bike sharing in a student population, *Transportation Research Part C: Emerging Technologies*, 105, pp. 611–625.

Manoogian, J., B. Benson (2017) Cognitive bias codex, *URL: https://www. sog. unc. edu/sites/www. sog. unc. edu/files/course_materials/Cognitive% 20Biases% 20Codex. pdf (accessed 30.08. 2025)*.

Mas-Colell, A., M. D. Whinston, J. R. Green, et al. (1995) *Microeconomic theory*, vol. 1, Oxford University Press, New York.

Miranda, M., M. Pereda, A. Sánchez, E. Estrada (2024) Indirect social influence and diffusion of innovations: An experimental approach, *PNAS nexus*, 3(10), pp. 1–11.

Molden, D. C., E. T. Higgins (2012) Motivated thinking, *The Oxford handbook of thinking and reasoning*, p. 390.

Montgomery, H. (1983) Decision rules and the search for a dominance structure: Towards a process model of decision making, in: *Advances in Psychology*, vol. 14, Elsevier, pp. 343–369.

Montibeller, G., D. Von Winterfeldt (2015) Cognitive and motivational biases in decision and risk analysis, *Risk Analysis*, 35(7), pp. 1230–1251.

Montibeller, G., D. von Winterfeldt (2024) Behavioral decision research: Descriptive and prescriptive perspectives, in: *Behavioral Decision Analysis*, Springer, pp. 15–40.

Mourad, A., J. Puchinger, C. Chu (2019) A survey of models and algorithms for optimizing shared mobility, *Transportation Research Part B: Methodological*, 123, pp. 323–346.

Müller, C., J. Gönsch, M. Soppert, C. Steinhardt (2023) Customer-centric dynamic pricing for free-floating vehicle sharing systems, *Transportation Science*, 57(6), pp. 1406–1432.

Nagpal, A., P. Krishnamurthy (2008) Attribute conflict in consumer decision making: The role of task compatibility, *Journal of Consumer Research*, 34(5), pp. 696–705.

Naumov, S., D. Keith (2023) Optimizing the economic and environmental benefits of ride-hailing and pooling, *Production and Operations Management*, 32(3), pp. 904–929.

Nickerson, R. S. (1998) Confirmation bias: A ubiquitous phenomenon in many guises, *Review of General Psychology*, 2(2), pp. 175–220.

Parducci, A. (1974) Contextual effects: A range-frequency analysis, *EC Carterette, MP (Eds.), Handbook of Perception*, 2, pp. 127–141.

Payne, J. W., J. R. Bettman, E. J. Johnson (1992) Behavioral decision research: A constructive processing perspective., *Annual Review of Psychology*, pp. 87–131.

Pratt, J. W. (1978) Risk aversion in the small and in the large, in: *Uncertainty in economics*, Elsevier, pp. 59–79.

Prelec, D., B. Wernerfelt, F. Zettelmeyer (1997) The role of inference in context effects: Inferring what you want from what is available, *Journal of Consumer Research*, 24(1), pp. 118–125.

Rogers, E. M., A. Singhal, M. M. Quinlan (2014) Diffusion of innovations, in: *An integrated approach to communication theory and research*, Routledge, pp. 432–448.

Savage, L. J. (1972) *The foundations of statistics*, Courier Corporation.

Shaheen, S., A. Cohen (2020) Similarities and differences of mobility on demand (mod) and mobility as a service (maas), *Institute of Transportation Engineers. ITE Journal*, (6), pp. 29–35.

Shaheen, S., A. Cohen, I. Zohdy, et al. (2016) Shared mobility: current practices and guiding principles, Tech. rep., United States. Federal Highway Administration.

Simon, H. A. (1955) A behavioral model of rational choice, *The Quarterly Journal of Economics*, pp. 99–118.

Simonson, I. (1989) Choice based on reasons: The case of attraction and compromise effects, *Journal of Consumer Research*, 16(2), pp. 158–174.

Simonson, I., A. Tversky (1992) Choice in context: Tradeoff contrast and extremeness aversion, *Journal of Marketing Research*, 29(3), pp. 281–295.

Stathopoulos, A., C. Cirillo, E. Cherchi, E. Ben-Elia, Y. T. Li, J. D. Schmöcker (2017) Innovation adoption modeling in transportation: New models and data, *Journal of Choice Modelling*, 25, pp. 61–68.

Ströhle, P., C. M. Flath, J. Gärttner (2019) Leveraging customer flexibility for car-sharing fleet optimization, *Transportation Science*, 53(1), pp. 42–61.

Tafreshian, A., N. Masoud, Y. Yin (2020) Frontiers in service science: Ride matching for peer-to-peer ride sharing: A review and future directions, *Service Science*, 12(2-3), pp. 44–60.

Tversky, A., D. Kahneman (1981) The framing of decisions and the psychology of choice, *Science*, 211(4481), pp. 453–458.

Tversky, A., D. Kahneman (1989) Rational choice and the framing of decisions, in: *Multiple criteria decision making and risk analysis using microcomputers*, Springer, pp. 81–126.

Tversky, A., D. Kahneman (1991) Loss aversion in riskless choice: A reference-dependent model, *The Quarterly Journal of Economics*, 106(4), pp. 1039–1061.

Tversky, A., D. Kahneman (1992) Advances in prospect theory: Cumulative representation of uncertainty, *Journal of Risk and Uncertainty*, 5, pp. 297–323.

Tversky, A., I. Simonson (1993) Context-dependent preferences, *Management Science*, 39(10), pp. 1179–1189.

Villa, E., V. Breschi, C. Ravazzi, M. Tanelli, F. Dabbene (2024) Can control aid in attaining sustainable goals? an improved data-informed framework to promote shared mobility, *Control Engineering Practice*, 153, p. 106106.

von Neumann, J., O. Morgenstern (1944) Theory of games and economic behaviour princeton univ, *Press, Princeton*, p. 44.

Vuelvas, J., F. Ruiz, G. Gruosso (2020) Energy price forecasting for optimal managing of electric vehicle fleet, *IET Electrical Systems in Transportation*, 10(4), pp. 401–408.

Wang, H., H. Guan, H. Qin, W. Li, J. Zhu (2022) A slack departure strategy for demand responsive transit based on bounded rationality, *Journal of Advanced Transportation*, 2022(1), p. 9693949.

Wang, X., W. Liu, H. Yang, D. Wang, J. Ye (2019) Customer behavioural modelling of order cancellation in coupled ride-sourcing and taxi markets, *Transportation Research Procedia*, 38, pp. 853–873.

Wang, X., H. Sun, S. Zhang, Y. Lv (2023) Bike-sharing rebalancing problem by considering availability and accessibility, *Transportmetrica A: Transport Science*, 19(3), p. 2067262.

Wang, Y., H. Sun, Y. Lv, X. Chang, J. Wu (2024) Reinforcement learning-based order-dispatching optimization in the ride-sourcing service, *Computers & Industrial Engineering*, 192, p. 110221.

Wu, C., S. Le Vine, A. Sivakumar, J. Polak (2022) Dynamic pricing of free-floating carsharing networks with sensitivity to travellers' attitudes towards risk, *Transportation*, 49(2), pp. 679–702.

Yang, T., Q. Guo, L. Xu, H. Sun (2021) Dynamic pricing for integrated energy-traffic systems from a cyber-physical-human perspective, *Renewable and Sustainable Energy Reviews*, 136, p. 110419.

Yoon, S.-O., I. Simonson (2008) Choice set configuration as a determinant of preference attribution and strength, *Journal of Consumer Research*, 35(2), pp. 324–336.

Zhang, J., M. Meng, Z. David (2019) A dynamic pricing scheme with negative prices in dockless bike sharing systems, *Transportation Research Part B: Methodological*, 127, pp. 201–224.

Zhang, Y., M. Kamargianni (2023) A review on the factors influencing the adoption of new mobility technologies and services: autonomous vehicle, drone, micromobility and mobility as a service, *Transport reviews*, 43(3), pp. 407–429.

Zheng, H., K. Zhang, Y. Nie, P. Yan, Y. Qu (2024) How many are too many? analyzing dockless bike-sharing systems with a parsimonious model, *Transportation Science*, 58(1), pp. 152–175.

Chapter 3

Service Design for Shared Micromobility: Long-Term Planning under Gradual User Adoption

Abstract: Service design for shared micromobility services, such as e-mopeds, involves decisions on service location, fleet size, pricing, and resource allocation, which are typically based on potential demand estimates. In practice, however, potential users adopt the service gradually, leading to market penetration over time. Ignoring this gradual penetration can lead to suboptimal decisions and service failures. We propose a phased optimization model to maximize profitability over a planning horizon divided into discrete periods. The model jointly determines service locations, fleet size, and spatio-temporal resource allocation while ensuring a targeted service level. To model gradual adoption, we apply an extended Bass diffusion framework in which adoption depends on service quality, defined by accessibility and availability, and social influence, reflecting the impact of current users on others' adoption. The phased deployment strategy aligns service expansion and investment decisions with adoption growth (market penetration) over multiple periods. Applied to a bike-sharing case, the proposed approach improves resource allocation and increases profitability by up to 54% compared to an upfront rollout that is based on potential rather than realized demand. The results shed light on the value of synchronizing service design decisions with user adoption dynamics.

Keywords: Service design, Shared micromobility, Phased planning, Bass diffusion, Social influence

---

Research from this chapter contributed to a paper submitted to the *Research in Transportation Business & Management*.

3.1 Introduction

Sustained user adoption and profitability are persistent challenges for many shared mobility services, including micromobility (Lagadic et al., 2019; Boer et al., 2022; Zhao et al., 2021). These challenges are closely tied to how services are designed and operated, as service design and user adoption are interconnected and jointly shape demand uptake. Shared micromobility service design is shaped by operations planning decisions such as service location, fleet size, pricing, and resource allocation (Laporte et al., 2018). These decisions are typically based on potential demand, which defines the maximum expected number of service users. Relying on potential demand estimation results in sub-optimal service design decisions, as it overlooks the gradual market penetration of the service. Market penetration, defined as the proportion of potential users who adopt the service over time, follows an S-shaped trend: slow uptake in the early stages of service introduction, rapid growth in the intermediate phase, and eventual saturation when the market is stabilized and reaches saturation. This curve reflects the cumulative outcome of adoption behavior (Rogers et al., 2014; Bass, 1969; Stathopoulos et al., 2017).

Adoption behavior (how potential users become actual users of an innovation, i.e., a new service or product) is influenced by service quality, such as convenient and reliable access (Zhang & Kamargianni, 2023; Ma et al., 2023), and social influence. Social influence is the effect of others' choices on an individual's decision. In the context of service adoption, as more people adopt (use) a service, the probability that an individual will also decide to adopt increases (Miranda et al., 2024; Manca et al., 2019). Adoption occurs only when service quality meets user expectations. Some users adopt based on service quality alone, regardless of others' choices. Others adopt when both service quality is adequate and they observe adoption by peers, where social influence plays a role. In this case, each additional adopter increases the likelihood that others will adopt, thereby accelerating overall adoption growth. Ignoring adoption dynamics has repeatedly led to service failures. For example, Beijing's early bike-share and Melbourne's public scheme both expanded capacity before adoption was sufficient, resulting in underutilization and financial losses (Liu et al., 2012; Fishman et al., 2012). Similarly, China's dockless bike-share sector scaled fleets rapidly ahead of adoption growth, creating severe oversupply and leading to failures (Zhao et al., 2021). These cases demonstrate that premature expansion, without aligning deployment to adoption growth, can undermine both operational efficiency and financial viability. Recent research confirms that calibrating fleet size, distribution, and pricing to actual adoption rates improves utilization and financial performance (Fuady et al., 2024).

The contribution of this research lies in integrating user adoption behavior into service design by proposing a long-term planning framework for free-floating micromobility. Conceptually, this shifts transportation planning from designing for static, stabilized demand to adoption-aware planning that explicitly considers demand uptake dynamics. To address this adoption-driven service design problem, this study formulates a mathematical model that explicitly represents adoption behavior in service design decisions and evaluates its implications using scenario-based analysis. We develop a phased optimization model to maximize profitability over a planning horizon divided into discrete periods, incorporating S-shaped market penetration through an adjusted Bass diffusion model. The framework supports phased deployment strategies and explicitly accounts for interdependencies between decisions across different periods. The model determines service locations, fleet size, and spatio-temporal resource allocation needed to maintain target service quality, measured by accessibility and availability, within each period.

The proposed model is applied to a bike-sharing case. Simulation results show that aligning service deployment with the adoption curve improves resource allocation, avoids over- and under-supply, and determines the timing of service location and fleet expansion across periods to match projected demand growth. The phased deployment framework supports continuous monitoring of adoption and allows adjustments to service design when deviations from projections occur. Scenario-based analyses are employed both to examine how variations in adoption behavior and demand patterns affect outcomes and to explore cases where initial adoption estimates are not correct. In the latter, an external belief-updating mechanism (Bayesian updating) is used to revise demand estimates and adjust service design decisions accordingly. Performance is benchmarked against designs based on stabilized demand and against a classical Bass model without service quality integration. The results demonstrate that accounting for demand growth across planning periods improves service profitability.

The remainder of this paper is structured as follows: Section 3.2 reviews the related literature, identifies research gaps, and outlines our contributions. Section 3.3 describes the problem setting and presents the dynamic programming formulation. Section 3.4 introduces the mixed-integer linear reformulation of the model. Section 3.5 details the data and scenario development for the bike-sharing case, followed by results and discussion in Section 3.6 and Section 3.7. Finally, Section 4.8 concludes the paper.

3.2 Relevant Literature

This study contributes to behavioral operations management in shared mobility by integrating a behavioral aspect (user adoption behavior) into operations management (long-term service design) models for free-floating services. To position this contribution within the literature, the review is structured to first provide an overview of shared free-floating micromobility service design examined in the literature in Section 3.2.1. Then, in Section 3.2.2, we elaborate on the adoption behavior and gradual market penetration of shared micromobility services, an aspect largely overlooked in prior research. Finally, we highlight the specific contributions of the current study.

3.2.1 Free-Floating Micromobility Service Design

The literature on shared micromobility service optimization is broad. Various papers such as Laporte et al. (2018); Zheng et al. (2024); Ricci (2015); Ataç et al. (2021) summarize the literature. This section is only intended to provide an overview of the service design problems that have been addressed for free-floating shared micromobility services.

Service design for free-floating micromobility involves three primary decisions: (1) determining the service area, (2) fleet sizing and resource allocation, and (3) repositioning and resource management (Laporte et al., 2018; Vogel & Vogel, 2016).

Service area determines the locations where users can take or return shared vehicles, known as resources hereafter. In free-floating services, pick-up and drop-off points are not restricted to fixed stations, as in docked systems. Instead, users can pick up and return resources (e.g., bikes or scooters) at any location within the defined service area. The service area design can be treated as a facility location problem that is used for station-based services by discretizing the target urban area, known as the operating area, into small zones, and selecting between the zones as service locations.

Most existing studies on free-floating micromobility systems treat the service region as given. Çelebi et al. (2018) and Giordano & Chow (2024) determine the services by choosing between candidate service zones for bike and scooter-sharing services, respectively. Çelebi et al. (2018) minimize the total unsatisfied demand and Giordano & Chow (2024) maximize the provider profit.

Fleet sizing is a tactical decision that determines the total resources (i.e., micromobility vehicles, such as bikes or scooters) required to develop the service. It depends on the time- and location-specific allocation of resources across service zones, known as the resource allocation problem. Caggiani et al. (2017) propose a strategic-level model to determine fleet size for a bike-sharing service that minimizes the bicycle equity index, and Jiang et al. (2020) study the fleet size problem with the objective of maximizing a provider's profit while incorporating a competing service provider. Both of these studies assume that the potential demand is given independently of the service design. On the contrary, Zheng et al. (2024) model the interdependence between demand and service design by jointly optimizing fleet size and pricing under a mode-choice framework with a profit-maximizing objective. In their model, users choose between transport modes based on trip time and cost, both of which depend on fleet size and repositioning operations.

In order to guarantee high levels of service quality and increased resource utilization rates, it is necessary to monitor and replenish the inventory level at each service location by periodically repositioning bicycles from one station to another.

Service providers implement either operator-based or user-based rebalancing strategies. In operator-based rebalancing, the service actively relocates resources using dedicated staff and rebalancing vehicles such as trucks and vans (Kim et al., 2023; Pal & Zhang, 2017; Reiss & Bogenberger, 2015; Caggiani et al., 2018; Wang et al., 2023). In user-based approaches, users are encouraged through incentives, such as discounts or credits, to end trips in under-supplied areas or begin in over-supplied ones (Jin et al., 2022; Mahmoodian et al., 2022; Zhang et al., 2019; Cheng et al., 2021). Since rebalancing directly influences resource availability across space and time, it must be coordinated with fleet sizing and resource allocation to maintain service quality, specifically availability, and efficiency.

Hybrid approaches that combine user-based and operator-based strategies have also emerged. Jin et al. (2023) develop a two-stage stochastic optimization model for rebalancing in free-floating micromobility systems, integrating outsourced third-party logistics (3PL) with user incentives.

Research Gap

In all the above studies, demand is treated as fixed, either independent of or dependent on service design, without considering gradual market penetration. As a result, demand is often modeled as if it has already reached its potential level. To the best of our knowledge, no prior work integrates service design with gradual user adoption in the context of micromobility services, while explicitly capturing the interdependence between service design and demand. To address this gap, we develop a phased optimization framework for free-floating micromobility services that explicitly links adoption behavior to service design through an adjusted Bass diffusion model. The following section describes the adoption behavior and the Bass diffusion model, followed by a detailed discussion of our contributions.

3.2.2 Adoption Behavior

An individual's decision to adopt a service is affected by the choices of others, i.e., social influence. A common explanation for this effect is that individuals face uncertainty when evaluating new or unfamiliar options, such as whether to adopt a new micromobility service (Manca et al., 2019). In such situations, they may follow others either to gain social approval and acceptance (normative conformity) or to obtain more accurate judgments and make informed decisions (informational conformity) (Deutsch & Gerard, 1955; Hsu & Lu, 2004; Cialdini & Goldstein, 2004).

The Bass diffusion model (Bass, 1969) is one of the most widely used frameworks for mathematically modeling adoption behavior. It distinguishes between two adopter types: *innovators*, who adopt independently of others, and *imitators*, who are influenced by other adopters. The model captures the cumulative impact of social influence through its imitation parameter and has been successfully applied in mobility contexts such as electric vehicles, carsharing, and micromobility (Jensen et al., 2017; Zhang et al., 2020; Ghasri & Vij, 2021; Massiani & Gohs, 2015; Dan et al., 2018; Brdulak et al., 2021).

The original Bass model does not account for the effects of service quality on adoption. To address this limitation, several extensions have been proposed. For example, Zhang et al. (2020) introduce an adjusted Bass model for car-sharing services that incorporates service-related effects. Their model distinguishes between hesitant and fast adopters and predicts station-level demand over time. Hesitant and fast adopters differ in their timing of adoption after being informed about a new service station. Their model captures network synergies, where the presence of one station increases the potential demand at others by enhancing connectivity and travel options. Similarly, Luo et al. (2024) extended the Bass model to study electric vehicle (EV) adoption in which the availability of charging stations influences user adoption.

Hybrid approaches have also been explored. Jensen et al. (2017) and El Zarwi et al. (2017) combine discrete choice models (DCMs) with the Bass framework to integrate socially driven adoption dynamics with utility-based user decisions. Along similar lines, Rotaris & Scorrano (2023) embed the Bass model within an agent-based simulation to capture both social influence and individual behavioral heterogeneity in car-sharing adoption.

The Bass model is a parsimonious model with a clear behavioral interpretation of its parameters, the distinction between innovators and imitators. This paper extends it to incorporate both service quality and continued usage behavior. In micromobility services, continued usage is critical to success, whereas the original Bass model primarily represents first-time adoption or purchase.

3.2.3 Contributions

The main contributions of this study lie in developing a phased optimization framework that integrates gradual demand realization (service adoption) into service design of shared micromobility systems, and are three-fold:

I. Explicitly modeling the joint impacts of service quality and social influence on the adoption of micromobility services during the early deployment stages.

II. Developing a phased service-adoption operations planning framework for micromobility services. This framework considers the full adoption trajectory rather than skipping this phase and planning only for the ultimate mature market share. By accounting for gradual adoption, service deployment, including service area and fleet size expansion, is aligned with actual realized demand at each stage, avoiding the risks of over- or under-supply and service failure.

III. Proposing an adjusted Bass diffusion model adapted for service design, in which service quality (i.e., service availability and accessibility) changes over time and directly influences adoption decisions. Potential users adopt the service if the quality meets their expectations, and adoption is further shaped by the influence of existing adopters. The adjusted Bass model provides a mathematical representation of adoption that captures initial engagement, continued usage, and the possibility of re-joining by previous users when service quality improves. This model reflects the distinctive nature of mobility services, where adoption requires not just first-time use but sustained participation.

3.3 Problem Description and Formulation

We study the strategic deployment of free-floating micromobility services in urban areas in light of social influence on potential users' decisions about adopting the service. Strategic service design decisions include service region selection and fleet[1] size, which together determine service quality. Service quality is characterized by two key dimensions (Zhang & Kamargianni, 2023; Bachand-Marleau et al., 2012; Pan, 2022; Ma et al., 2023; Albiński et al., 2018): service *accessibility* (i.e., coverage of users' desired origins and destinations) and *availability* (i.e., likelihood of finding resources when needed). The provider decides about service design so that their profit is maximized over a considered planning horizon.

The service under consideration is new to the target population, either as an innovative form of micromobility or as a recent market entrant. In either case, adoption does not occur immediately; individuals need time to assess service quality, through personal use or by observing others, before committing to continued use. We assume that all potential users are already aware of the service; therefore, we do not model awareness diffusion.

User adoption is shaped by service quality and reinforced by social influence, which itself grows with the number of adopters. As service design determines quality, it actively guides the pace and scale of adoption, making demand a gradual, design-driven process.

[1] By fleet, we refer to micromobility resources such as bikes or mopeds

The service provider faces a strategic decision: invest in high-quality service upfront based on the estimated ultimate number of adopters or expand the service gradually aligned with demand growth. An upfront commitment to high-quality service imposes unnecessary investments and underutilization of resources. Alternatively, a gradual expansion allows the provider to align the investment and service capacity expansions with demand growth while observing the actual user adoption. The observation can be utilized to update the demand estimations to reduce demand over- or under-estimations that result in investment inefficiency and poor service quality, respectively.

To address these challenges, we develop a phased service-adoption planning framework to maximize profit for micromobility providers over a fixed horizon. This horizon is divided into H periods indexed by $h \in \mathcal{H}$, where $\mathcal{H}$ denotes the set of all periods, and it contains two main phases: (1) implementation phase, and (2) post-implementation phase. The provider deploys the service during H_1 implementation periods and stops expanding the fleet or service locations after that. But the system continues to work, for H_2 periods after the implementation stage. The set of implementation and post-implementation periods are denoted by $\mathcal{H}_1$ and $\mathcal{H}_2$, respectively. Accordingly, $\mathcal{H} = \mathcal{H}_1 \cup \mathcal{H}_2$. At each period, the provider could also decide to withdraw the service.

Time and location discretizations. To approximate the operational costs and trip revenue, we consider a nominal day as the operational planning horizon and divide it into multiple intervals of the same length of a unit time (e.g, one hour). The unit time is the time basis of operational planning and is denoted by T^{unit}. The shorter we consider the unit time, the more accurate our estimations will be. Each period h contains $\mathscr{N}_h^{\text{unit}}$ unit times, considering the number of days in each period and the number of intervals that we divide a day into. We also discretize the targeted urban area into N identical hexagonal grids, indexed by $i, j \in \mathcal{I}$, where $\mathcal{I}$ is the set of all grids.

Profit of period h, noted by $\mathscr{P}_h$, contains revenue and costs. Revenue of period h is denoted by $\mathscr{R}_h$, and it contains revenue from the served trips, subscription fees, and salvage of extra resources. Costs for the micromobility service are categorized into two main components: investment and operational costs. Investment costs (i.e., $\mathscr{C}_h^{\text{I}}$) encompass expenses related to the purchase and maintenance of resources as well as periodic rental costs of covered grids in period h. Operational costs (i.e., $\mathscr{C}_h^{\text{O}}$) represent expenses associated with rebalancing vehicles across zones to meet spatial demand. These costs are influenced by the selected grids, fleet size, resource distribution, and service availability decisions.

The problem can be conceptually viewed as a dynamic program as follows, where decisions are made sequentially and each decision influences future system states.

$$
\begin{aligned}
V(\mathcal{W}_h) = \max_{\boldsymbol{\omega}_h \in \Omega(\mathcal{W}_h)} \quad & \mathscr{P}_h + (1+\Upsilon)^{-1}[V(\mathcal{W}_{h+1}|\mathcal{W}_h, \boldsymbol{\omega}_h)] \\
\text{s.t.} \quad & \mathscr{P}_h = \mathscr{R}_h(\mathcal{D}_h, \boldsymbol{\omega}_h, \mathcal{W}_h) - \mathscr{C}_h^{\text{O}}(\mathcal{D}_h, \boldsymbol{\omega}_h, \mathcal{W}_h) - \mathscr{C}_h^{\text{I}}(\mathcal{D}_h, \boldsymbol{\omega}_h, \mathcal{W}_h) \\
& \Omega_h(\mathcal{W}_h) = \Big\{ \boldsymbol{\omega}_h \mid (3.1) - (3.25) \Big\}
\end{aligned}
\tag{DP}
$$

In (DP), $\mathcal{D}_h$ denotes the estimated demand in period h, comprising adopters from the previous periods and newly engaged users in h. $\mathcal{W}_h$ is the system state in period h, including the number of adopters and fleet size at the beginning of the period. The provider's decisions (covering service region, fleet size, resource distribution, service availability, and rebalancing operations) are collectively denoted by $\boldsymbol{\omega}_h$, with feasible set $\Omega(\mathcal{W}_h)$. Υ is the discount rate, which accounts for the current value of money. Finally, $V(\mathcal{W}_h)$ is the maximum expected total profit from period h onward, given state $\mathcal{W}_h$.

In the subsequent sections, we formulate the problem. We first elaborate on the demand modeling approach and the connection of demand with service design and social influence. Then, we compute the investment and operational costs, revenue, and profit.

3.3.1 Demand Model

Demand is a function of the population of potential users, service quality, users' expectations and trip patterns, and social influence. Potential users adopt the service based on the service quality, their expectations (i.e., aspiration levels), and social influence.

Adoption definitions. We define adoption as the continued use of the service, rather than a one-time engagement, which is essential for service profitability. This aligns with perspectives in the sharing economy literature, where sustained participation is considered as part of adoption behavior (Wang et al., 2021). Accordingly, adoption comprises two key decisions: the *engagement* decision, where a potential user first tries the service, and the *confirmation* decision, where a user chooses to continue using the service.

We define *potential users* as individuals who are aware of the service and are willing to adopt it if their expectations for service quality are met. When a potential user tries the service for the first time, which is driven by satisfaction with service accessibility, they become a *new user*. If the new user is satisfied with service availability and continues to use the service, they become an *adopter*. We refer to both new users and adopters as *active users*. Those who discontinue usage due to availability dissatisfaction are considered *lost adopters*.

User segments. Potential users are grouped into G segments, indexed by $g \in \mathcal{G}$, where $\mathcal{G}$ represents the set of all segments. Users in each segment have the same usage frequency (periodic number of total trips), trip pattern, and service quality expectations. The maximum number of potential users in each segment, denoted by M_g, is considered an exogenous input.

Spatio-temporal demand (trip pattern). We assume a time-dependent and segment-specific origin-destination (OD) matrix that represents the potential number of trips between grids during each daily interval by each user segment. This potential demand describes the expected number of trips when all locations are covered and all potential users are captured as adopters. We model the number of requests between each pair of grids i and j as a Poisson process, which is a common model used for representing demand rates in micromobility services where users' arrivals are occurring independently over time (Chiariotti et al., 2020; Paul et al., 2023). The parameter of these Poisson processes for each segment and OD pair (i.e., $D_{i,j,g}$) is obtained from the average of the OD matrix over the time intervals to capture trip patterns for each segment.

In the following sections, we first define the service quality and its connection to users' adoption decisions. Then, develop the adoption model that simultaneously incorporates service quality and social influence. A summary of key definitions and notation used in this paper is provided in Table 3.1.

Table 3.1: Main definitions and mathematical notations

Adoption definitions	
Potential user	An individual aware of the service and willing to adopt it if quality expectations are met
New user	A potential user who tries the service for the first time because of satisfactory accessibility (engagement)
Adopter	A user who confirms the service availability and continues using it (confirmation)
Active user	A user currently using the service (includes both new users and adopters)
Lost adopter	A user who discontinues usage due to availability dissatisfaction
Sets and indices	
$\mathcal{I}$	Index set of grids, $i, j \in \mathcal{I} = \{1, \ldots, N\}$
$\mathcal{G}$	Index set of user segments, $g \in \mathcal{G} = \{1, \ldots, G\}$
$\mathcal{H}$	Index set of planning periods, $h, h', h'' \in \mathcal{H} = \{1, \ldots, H_1 + H_2\}$; $\mathcal{H} = \mathcal{H}_1 \cup \mathcal{H}_2$
$\mathcal{H}_1$	Index set of implementation periods, $h \in \mathcal{H}_1 = \{1, \ldots, H_1\}$
$\mathcal{H}_2$	Index set of post-implementation periods, $h \in \mathcal{H}_1 = \{H_1 + 1, \ldots, H_1 + H_2\}$
$\mathcal{A}$	Index set of service availability levels, $a \in \mathcal{A} = \{0, \ldots, A\}$
$\mathcal{B}$	Index set of service accessibility satisfaction levels, $b \in \mathcal{B} = \{0, \ldots, B\}$
Main parameters	
(p_g, q_g)	Innovation and imitation coefficients for user segment g
$\mathscr{A}_g^{\text{ac}}$	Random variable representing the accessibility aspiration of users in segment g
$\mathscr{A}_g^{\text{av}}$	Random variable representing the availability aspiration of users in segment g
M_g	Total population of user segment g
$\Pi_{b,g}^{\text{av}}$	Candidate accessibility satisfaction level b for user segment g
A_a^{av}	Candidate availability level a
T^{unit}	Length of the unit time used as the basis for operational planning (e.g., one-hour intervals).
$\mathscr{N}_h^{\text{unit}}$	Number of unit times within period h
$D_{i,j,g}$	Total number of trips from grid i to j by user segment g within one time unit
$(R_{h,g}^{+}, P_g^{\text{lock}} P_g)$	Periodic subscription fee, per-trip unlock fee, and per-minute riding cost for user segment g
Decision variables	
x_{ih}	Binary variable; 1 if grid i is selected at period h, 0 otherwise
$\chi_{i,j,h}$	Binary variable; 1 if both grids i and j are selected at period h, 0 otherwise
z_h	Total fleet size in period h
z_h^{+}	Number of purchased resources in period h
z_h^{-}	Number of salvaged resources in period h
r_h	Total number of repositioned resources by trucks in each unit time (i.e., T^{unit}) of period h
$y_{i,h}$	Inventory level of resources in gird i in each unit time of period h
ϕ_{ah}	Binary variable; 1 if availability level a is selected at period h, 0 otherwise.
$\theta_{b,g,h}$	Binary variable; 1 if accessibility satisfaction level b is selected at period h for user segment g, 0 otherwise.
$\alpha_{g,h}$	Total proportion of active users of segment g in period h
$\beta_{g,h,h'}$	Total proportion of lost adopters of segment g in period h' that rejoin in period h

Service Quality

Service quality is formulated by service accessibility and availability as follows.

Service accessibility. A trip is accessible if both its origin and destination lie within service locations. The provided service accessibility for user segment g during period h is formulated as the total utility of accessible trips. We define the binary decision variable $x_{i,h}$ that is one if grid i is selected as a service location in period h, collectively represented by $\boldsymbol{x}_h$. We also define the binary decision variable $\chi_{i,j,h}$ that is one if both i and j are service locations in period h. Therefore,

$$\chi_{i,j,h} \leq x_{i,h} \qquad \forall i, j \in \mathcal{I}, h \in \mathcal{H} \tag{3.1}$$

$$\chi_{i,j,h} \leq x_{j,h} \qquad \forall i, j \in \mathcal{I}, h \in \mathcal{H} \tag{3.2}$$

$$\chi_{i,j,h} \geq x_{i,h} + x_{j,h} - 1 \qquad \forall i, j \in \mathcal{I}, h \in \mathcal{H}. \tag{3.3}$$

We assume that the network expands over the implementation phase, meaning no covered location is omitted after inclusion. This can be formulated as below:

$$x_{i,h} \geq x_{i,h-1} \qquad \forall i \in \mathcal{I}, 2 \leq h \leq H_1 \tag{3.4}$$

$$x_{i,h} = x_{i,h-1} \qquad \forall i \in \mathcal{I}, h \in \mathcal{H}_2. \tag{3.5}$$

Let $u_{i,j,g} \in [0,1]$ represent the normalized utility of a trip from i to j for user segment g, which is a function of trips' frequency and value. Thus, the provided service accessibility for user segment g during period h is $\sum_{(i,j)} u_{i,j,g}\chi_{i,j,h}$.

Service availability. We define service availability as the proportion of trip requests for which at least one resource is available, representing the share of demand that can be successfully served. Service availability is selected from a finite set of A levels A_a^{av} for each period. For $a = 0$, $A_0^{\text{av}} = 0$ that corresponds to withdrawing the service. We define the binary variable $\phi_{a,h}$ that is one if we provide the availability rate of A_a^{av} in period h. At each period, we select one of the candidate availability levels that applies for all segments using the following constraints:

$$\sum_{a=0}^{A} \phi_{a,h} = 1 \qquad \forall h \in \mathcal{H}. \tag{3.6}$$

Service quality and users' expectations. We assume that potential users follow a satisficing rule in their adoption decisions, both engagement and confirmation. They adopt the service if its quality exceeds their expectations. This threshold-based adoption modeling aligns with Expectancy-Confirmation Theory (Oliver, 1980), which posits that individuals evaluate services by comparing perceived quality to their expectations, adopting when expectations are met (Pandita et al., 2025).

Potential users in segment g engage with the service if service accessibility (i.e., $\sum_{(i,j)} u_{i,j,g}\chi_{i,j,h}$) exceeds their aspiration threshold (i.e., $\tilde{\mathscr{A}}_g^{\text{ac}}$). We model $\tilde{\mathscr{A}}_g^{\text{ac}}$ as a random variable that accounts for aspiration variation across the population. For segment g, the probability of accessibility satisfaction (i.e., engagement) during period h, i.e., $\pi_{g,h}^{\text{a}}$, is:

$$\pi_{g,h}^{\text{a}} = \text{Prob}\left(\sum_{(i,j)} u_{i,j,g}\chi_{i,j,h} \geq \tilde{\mathscr{A}}_g^{\text{ac}}\right) \qquad \forall h \in \mathcal{H}, g \in \mathcal{G}. \tag{3.7}$$

Users confirm the service quality and continue to use the service based on their experienced service availability and their aspiration levels, noted by $\tilde{\mathscr{A}}_g^{\text{av}}$. We incorporate the variability of availability aspiration levels across the population. Therefore, $\tilde{\mathscr{A}}_g^{\text{av}}$ is a random variable. The probability of availability satisfaction (i.e., confirmation) for segment g in period h, i.e.,π_{gh}^{c}, is:

$$\pi_{gh}^{c} = \text{Prob}\left(\sum_{a} \phi_{a,h} A_a^{\text{av}} \geq \tilde{\mathscr{A}}_g^{\text{av}}\right) \qquad \forall h \in \mathcal{H}, g \in \mathcal{G}. \tag{3.8}$$

Adoption Model

Original Bass model. Let $F(h)$ denote the proportion of total adopters by time h. The instantaneous adoption probability at time h is given by the derivative of the cumulative adoption function $F(h)$ with respect to time, i.e., $\frac{dF(h)}{dh}$ as the following equation (Bass, 1969):

$$f(h) = \frac{dF(h)}{dh} = \big(p + qF(h)\big)(1 - F(h)). \tag{A1}$$

Here, p and q represent the innovation and imitation coefficients, respectively. These coefficients correspond to innovators and imitators. p reflects the likelihood that potential users adopt the service independently of others, driven by intrinsic preferences and external factors such as marketing efforts. By contrast, the imitation coefficient, q, represents the extent to which the potential users are influenced by others adoption.

Equation (A1) leads to the following distribution:

$$F(h) = \frac{1 - \mathrm{e}^{-(p+q)h}}{1 + \left(\frac{q}{p}\right)\mathrm{e}^{-(p+q)h}} \tag{A2}$$

For a discrete model, as in our phased planning framework, we can define the cumulative number of adopters by the beginning of period h as N_{h-1} and n_h as the number of new users joining during (by the end of) period h. Moreover, M represents the total potential users. Accordingly, $F(h) = N_h/M$ and $f(h) = (N_h - N_{h-1})/M$. The number of new users at time h for a discrete setting is given by (Mahajan et al., 1990):

$$\begin{aligned} n_h = \frac{dN_{h-1}}{d_h} &= \overbrace{p(M - N_{h-1})}^{\text{New innovators}} + \overbrace{q\frac{1}{M}N_{h-1}(M - N_{h-1})}^{\text{New imitators}} \\ &= pM - pN_{h-1} + q\frac{1}{M}N_{h-1}M - q\frac{1}{M}N_{h-1}^2 \end{aligned} \tag{A3}$$

In (A3), the term pM denotes the number of innovators that could have joined by time h, and the term pN_{h-1} denotes innovators who have already joined. The term $q\frac{1}{M}N_{h-1}M$ denotes the population of the imitators that could have joined by h which are determined based on the population of the current adopters (i.e., N_{h-1}). The final term denotes the imitators that have joined already.

As the Bass model was originally developed for products with fixed quality and does not account for disconfirmation, we propose an adjusted model that incorporates the impact of changing service quality, as well as disconfirmation resulting from service quality dissatisfaction. To this end, we first introduce our assumptions about the adoption behavior of innovators and imitators, then propose the adjusted Bass model.

Innovators and imitators adoption. We assume that innovators immediately engage with the service if their expectation about service accessibility is met. However, imitators engage gradually over the planning horizon, influenced by service accessibility and the number of adopters. After their engagement, both innovators and imitators need one period to evaluate service availability and confirm their adoption decision. Additionally, we assume that both groups reassess service quality dynamically throughout the planning horizon. They may decide to use the service if not yet engaged, rejoin if previously rejected, or reject if currently adopted.

Adjusted Bass model. For the sake of brevity, we leave out the segment index g for constructing our adoption model. Based on our definitions, the maximum portion of the potential users that could adopt the service is a function of π_h^a which implies the expected proportion of the potential users whose accessibility aspirations is satisfied at h. Therefore, we multiply pM and $q\frac{1}{M}N_hM$ by π_h^a in our proposed model.

Based on our adoption definition, users' engagement follows with a confirmation based on the service availability, which determines the proportion of adopters. We multiply the imitation terms (those including q) in (A3) by π_{h-1}^c, ensuring that only the proportion of users who have experienced the service and confirmed its quality (i.e., adopters) contribute to further adoption in period t. As mentioned earlier, π_h^c represents the proportion of users whose availability aspiration is satisfied during period h.

The updated number of new users at period h is then given by:

$$\begin{aligned} n_h &= pM\pi_h^a - pN_{h-1} + q\frac{1}{M}\pi_{h-1}^c N_{h-1}\pi_h^a M - q\frac{1}{M}\pi_{h-1}^c N_{h-1}^2 \\ &= M\left(\overbrace{p(\pi_h^a - N_{h-1}/M)}^{\text{New innovators}} + \overbrace{(q\frac{1}{M}N_{h-1}\pi_{h-1}^c)(\pi_h^a - N_{h-1}/M)}^{\text{New imitators}}\right) \\ &= M\left((p + q\frac{1}{M}N_{h-1}\pi_{h-1}^c)(\pi_h^a - N_{h-1}/M)\right). \end{aligned} \tag{A4}$$

In Equation (A4), the term $N_{h-1}\pi_{h-1}^c$ represents the number of adopters at the start of period h. The cumulative number of active users at period h is then computed by adding this quantity to the number of new users entering during the same period:

$$\begin{aligned} N_h &= \pi_{h-1}^c N_{h-1} + n_h \\ &= M\left(\pi_{h-1}^c N_{h-1}/M + (p + q\frac{1}{M}N_{h-1}\pi_{h-1}^c)(\pi_h^a - N_{h-1}/M)\right). \end{aligned} \tag{A5}$$

Defining $\alpha_h = N_h/M$ as the cumulative proportion of active users during period h, we obtain:

$$\alpha_h = \pi_{h-1}^c\alpha_{h-1} + (p + q\alpha_{h-1}\pi_{h-1}^c)(\pi_h^a - \alpha_{h-1}). \tag{A6}$$

The proportion of new adopters, known as adoption growth, can also be obtained by the formula $\Delta\alpha_h = \alpha_h - \alpha_{h-1}$.

Re-joining. In each period h', a fraction of active users ($\alpha_{h'}$) become lost adopters due to dissatisfaction with availability, quantified as $1 - \pi_{h'}^c$. If availability improves in a future period h, these lost adopters may rejoin the service with probability J, starting from period $h+1$ onward to account for a one-period lag for users to become aware of the improvement. Therefore, rejoining in period $h \geq 3$ is only possible for lost adopters from periods $h' \leq h-2$, and only if $\pi_h^c \geq \pi_{h'}^c$.

Let $(\pi_h^c - \pi_{h'}^c)^+ = \max(0, \pi_h^c - \pi_{h'}^c)$ denote the improvement in availability in period h relative to period h'. We define $\pi_{h,h'}^{+c} = \alpha_{h'}(\pi_h^c - \pi_{h'}^c)^+$ as the total proportion of lost adopters from period h' that are eligible to rejoin in period h due to this improvement.

We define $\beta_{h,h'}$ as the fraction of lost adopters from period $h' \leq h-2$ who rejoin in period h. This is modeled as a fixed fraction J of the eligible pool $\pi_{h,h'}^{+c} \in [0,1]$, excluding those who have already rejoined in earlier improvements (i.e., $\sum_{h''=h'+1}^{h-2} \beta_{h'',h'}$). Hence:

$$\alpha_h = \alpha_{h-1}\pi_{t-1}^c + \Big(p_+ q(\pi_{h-1}^c \alpha_{h-1})\Big)(\pi_h^a - \alpha_{h-1}) + \sum_{h'} \beta_{h,h'} \tag{A7}$$

$$\beta_{h,h'} = J\left(\pi_{h,h'}^{+c} - \sum_{h''=h'+1}^{h-2} \beta_{h'',h'}\right). \tag{A8}$$

The segment-based adoption dynamics are formulated as follows:

$$\alpha_{gh} = \alpha_{g,h-1}\pi_{g,h-1}^c + \Big(p_g + q_g(\pi_{g,h-1}^c \alpha_{g,h-1})\Big)(\pi_{g,h}^a - \alpha_{g,h-1}) + \sum_{h'=1}^{h-2} \beta_{g,h,h'} \qquad \forall h \in \mathcal{H}, g \in \mathcal{G} \tag{3.9}$$

$$\pi_{g,h,h'}^{+c} \leq \beta_{g,h,h'}^{+} \qquad \forall h \geq 3, h' \leq h-2, g \in \mathcal{G} \tag{3.10}$$

$$\pi_{g,h,h'}^{+c} \leq \alpha_{g,h'}(\pi_{g,h}^c - \pi_{g,h'}^c) + (1 - \beta_{g,h,h'}^{+}) \qquad \forall h \geq 3, h' \leq h-2, g \in \mathcal{G} \tag{3.11}$$

$$\pi_{g,h,h'}^{+c} \geq \alpha_{g,h'}(\pi_{g,h}^c - \pi_{g,h'}^c) \qquad \forall h \geq 3, h' \leq h-2, g \in \mathcal{G} \tag{3.12}$$

$$\beta_{g,h,h'} = J\left(\pi_{g,h,h'}^{+c} - \sum_{h''=h'+1}^{h-1} \beta_{g,h'',h'}\right) \qquad \forall h \geq 3, h' \leq h-2, g \in \mathcal{G} \tag{3.13}$$

$$\beta_{g,h,h'} = 0 \qquad \forall h \leq 2, g \in \mathcal{G}, \tag{3.14}$$

where $\beta_{g,h,h'}^{+}$ is a binary variable that equals 1 if $\pi_{g,h}^c - \pi_{g,h'}^c \geq 0$. Additionally, the initial condition is defined as:

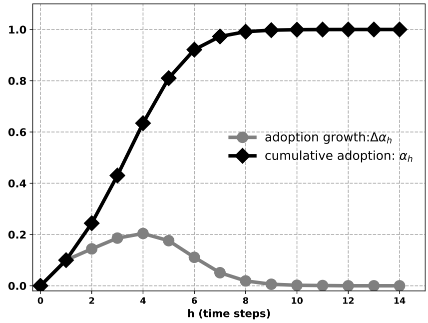
$$\pi_{g,1}^a = \pi_{g,1}^c = \beta_{g,1} = 0,$$

which indicates that no users have engaged or confirmed usage at the beginning of the planning horizon.

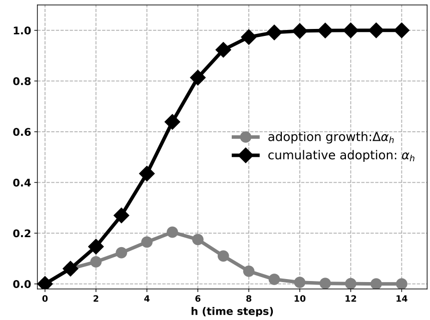
Illustrative example. Figure (3.1) displays cumulative adoption, adoption growth, and rejoining proportions over 15 periods, with parameters $p = 0.1$, $q = 0.6$, and $J = 0.5$. In Figure (3.1a), full accessibility and availability ($\pi_h^a = \pi_h^c = 1$) are assumed in all periods. Figure (3.1b) presents a case where accessibility satisfaction starts at 0.5 and increases by 0.1 per period. In both cases, cumulative adoption eventually reaches 1, but the speed of adoption is influenced by accessibility levels.

Figures (3.1c) and (3.1d) show adoption and rejoining proportions, respectively, under gradually increasing accessibility and availability satisfaction, both starting at 0.5 and increasing by 0.1 per period. Rejoining peaks at $h = 5$, when availability satisfaction reaches 1, and then declines as no new lost adopters are generated from $h' \geq 5$. Rejoining ceases entirely at $h = 12$ when full adoption is reached ($\alpha_h = 1$).

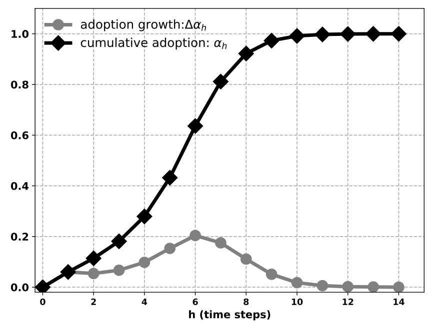
These results show that gradual improvements can lead to full adoption within 15 periods but delay the peak adoption rate. The S-shaped cumulative adoption curve (α_h) aligns with classic diffusion patterns, with the inflection point matching the peak in adoption growth ($\Delta\alpha_h = \alpha_h - \alpha_{h-1}$).



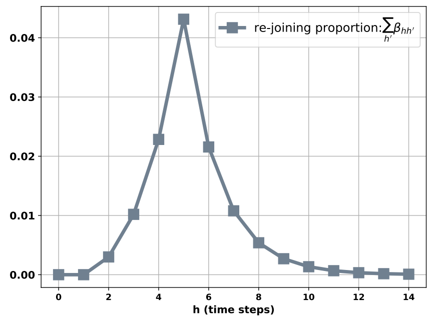
(a) Accumulated adoption and adoption growth for fixed full availability and accessibility satisfaction; $\pi_h^a = \pi_h^c = 1$ in all periods



(b) Accumulated adoption and adoption growth for fixed availability satisfaction $\pi_h^c = 1$ and growing accessibility of $\pi_h^a = 0.5 + 0.1h$



(c) Accumulated adoption and adoption growth for gradually increasing availability and accessibility satisfactions; $\pi_h^a = 0.5 + 0.1h$ and $\pi_h^c = 0.7 + 0.1h$



(d) Rejoin proportion for gradually increasing availability and accessibility satisfaction, $\pi_h^a = 0.5 + 0.1h$ and $\pi_h^c = 0.5 + 0.1h$

Figure 3.1: Cumulative adoption, adoption growth, and re-joined proportions for $p = 0.1$, $q = 0.6$, and $J = 0.5$ under various settings for periodic availability and accessibility satisfactions.

3.3.2 Investment Costs

To approximate the investment costs, we first need to determine the required fleet size, which is elaborated in the following section.

Fleet Size

To approximate the required fleet size that meets a targeted service availability (i.e., $\sum_a \phi_{a,h} A_a^{\text{av}}$), we consider that each resource can be in one of the following three states: (i) in use (rented) by users, (ii) in transit due to repositioning operations, and (iii) idle.

Resources in the in-use and in-transit states are approximated by modeling them as $./G/\infty$ queues (Whitt, 1984). This queuing model assumes general service time (no specific distribution) and an infinite number of servers (resources and reposition trucks in our problem). The service times for in-use and in-transit queues represent rental and repositioning duration, respectively.

As discussed earlier, users arrive at grid i to take resources following a Poisson process with a rate of $\sum_{j,g} D_{i,j,g}$ per unit time. This rate determines the rate at which resources arrive in the in-use state. Let $T_{i,j,g}$ denote the trip time from grid i to grid j via segment g. Therefore, $L^{\mathrm{u}}_{i,j,g} = \frac{T_{i,j,g}}{T^{\mathrm{unit}}}$ represents the portion of each unit time that a resource is in use by users, corresponding to the service time for in-use queues.

The arrival rate of resources into the transit state at grid i is determined by resource surplus caused by imbalances in trip flows:

$$\max\left\{0, \sum_{j,g}\left(D_{i,j,g} - D_{j,i,g}\right)\right\}.$$

This expression ensures that surplus resources in each grid enter the transit state, thus maintaining flow balance across grids. Only those origin-destination pairs (ODs) that are covered within a given period (i.e., those for which $\chi_{i,j,h}$ is non-zero) can contribute to flow imbalances. Thus, we define:

$$z^{\mathrm{extra}}_{i,h} = \max\left\{0, \sum_{j,g}\left(D_{i,j,g} - D_{j,i,g}\right)\chi_{i,j,h}\right\},$$

where $z^{\mathrm{extra}}_{i,h}$ is positive only if the total inflow exceeds the total outflow.

For the in-transit state, we assume that repositioning operations occur in fixed-duration shifts of length T^{shift} (e.g., one hour). Thus, $L^{\mathrm{shift}} = \frac{T^{\mathrm{shift}}}{T^{\mathrm{unit}}}$ represents the portion of each time interval that resources are in transit by trucks, corresponding to the in-transit service time.

Applying Little's Law, the expected number of resources in use by users, $z^{\text{in-use}}_{ih}$, and the expected number of resources in transit, $z^{\text{in-transit}}_{ih}$ during each unit time of period h at grid i are the product of the expected service time and the total arrival rate of resources to that state as below:

$$z^{\text{in-use}}_{i,h} = \sum_{j} L^{\mathrm{u}}_{ijg} D_{i,j,g} \chi_{i,j,h}$$
$$z^{\text{in-transit}}_{i,h} = L^{\mathrm{shift}} z^{\mathrm{extra}}_{i,h}$$

The number of idle resources, those available to be rented, is modeled using a $./\mathrm{M}/1$ queue. User arrivals follow a Poisson process; thus, the service time, which is the time between two consecutive rentals, is exponentially distributed (Markovian). Each user rents exactly one resource during their transaction.

We define the service level $\phi_{a,h}$ as the long-run proportion of service requests that are met, i.e., the probability that at least one idle resource is available. Under the $./\mathrm{M}/1$ framework, this service level requires the following number of resources in the idle state in grid i during period

h, $z^{\text{idle}}_{i,h}$:

$$z^{\text{idle}}_{i,h} = \frac{\sum_a A^{\text{av}}_a \phi_{a,h}}{1 - \sum_a A^{\text{av}}_a \phi_{a,h}}.$$

The inventory level of fleets in grid i during each unit time (i.e., T^{unit}) of period h, noted by $y_{i,h}$, must satisfy:

$$y_{i,h} \geq z^{\text{idle}}_{i,h} + z^{\text{in-use}}_{i,h} + z^{\text{in-transit}}_{i,h}$$

Finally, the total fleet size, z_h, and grid-specific fleet inventory level per unit time, $y_{i,h}$, during each period h are determined by the following set of constraints:

$$y_{i,h} \geq \overbrace{\frac{\sum_a A^{\text{av}}_a \phi_{a,h}}{1 - \sum_a A^{\text{av}}_a \phi_{a,h}} x_{i,h}}^{\text{idle}} + \overbrace{\sum_g \alpha_{g,h}(1-\phi_{0,h}) \sum_j L^{\text{u}}_{i,j,g} \chi_{i,j,h} D_{i,j,g}}^{\text{in-use}} + \overbrace{L^{\text{shift}} z^{\text{extra}}_{i,h}}^{\text{in-transit}} \quad \forall i \in \mathcal{I}, h \in \mathcal{H} \quad (3.15)$$

$$z^{\text{extra}}_{i,h} \geq \overbrace{\sum_{g,j} (1-\phi_{0,h}) \alpha_{g,h} (D_{j,i,g} - D_{i,j,g}) \chi_{i,j,h}}^{\text{aggregated trip flow imbalances for grid i}} \quad \forall i \in \mathcal{I}, h \in \mathcal{H} \quad (3.16)$$

$$z_h \geq \sum_i y_{i,h} \quad \forall h \in \mathcal{H}, \quad (3.17)$$

where $\alpha_{g,h}$ and $(1 - \phi_{0,h})$ adjust for the adopted proportion and service stop.

Availability Requirements. To ensure a desired level of availability in free-floating services, users should be able to find resources within a convenient walking distance from their location (Kabra et al., 2020; Reck et al., 2021; Wang et al., 2018). To ensure users can find at least one resource within a convenient walking area of A(W) with a minimum probability of $\mathcal{P}$, the inventory level of resources in each covered grid must satisfy (Soppert et al., 2023):

$$y_{i,h} \geq \frac{\ln(1-\mathcal{P})}{\ln\left(1 - \frac{\text{A(W)}}{\text{A(Z)}}\right)} \quad \forall h \in \mathcal{H}. \quad (3.18)$$

Where A(Z) is the grid area.

Resource purchase and salvage. The number of purchased and salvaged resources of period h are denoted by z^+_h and z^-_h, respectively. We formulate the connection between the fleet size, purchasing, and salvage decisions as follows:

$$z^+ = 0 \quad \forall h \in \mathcal{H}_2 \quad (3.19)$$

$$z_h = z_{h-1} + z^+_h - z^-_{h-1} \quad \forall h \in \mathcal{H}_1 \quad (3.20)$$

Investment costs. Let B^-_h and M^-_h denote the purchase and maintenance cost per resource, respectively. Moreover, the provider must pay a periodic cost of C^{grid}_h to the authorities to operate in each grid as their service grid in period h. Therefore, the investment cost for period h can be computed as follows:

$$\mathscr{C}_h^{\mathrm{I}} = B_h^- z_h^+ - M_h^- z_h - C_h^{\mathrm{grid}}(1 - \phi_{0,h}) \sum_i x_{i,h}$$

The term $(1 - \phi_{0,h})$ sets the grid costs to zero when the service stops in period h.

3.3.3 Operational Costs

To estimate operational costs, we consider identical trucks with a capacity of K resources to reposition them during each time interval. Repositioning trips consist of two components: line-haul trips, where trucks move resources between excess and deficit locations, and peddling trips, which involve loading and unloading resources.

To ensure the targeted inventory level $y_{i,h}$ is maintained over time, the flow imbalance of grids caused by trip flow imbalances must be repositioned. This value for grid i equals $z_{i,h}^{\mathrm{extra}}$, as discussed above. Therefore, we have:

$$r_h = \sum_i z_{i,h}^{\mathrm{extra}} \qquad \forall h \in \mathcal{H}. \tag{3.21}$$

Repositioning Time Estimation

To estimate the total time required for repositioning operations by trucks, we approximate the time needed for three main tasks: (i) Travel Time Between Grids, the average time required for trucks to travel between each pair of grids they visit, (ii) Resource Search and Load/Unload Time, the time required to find resources scattered within each grid, including load and unload time, and (iii) Empty Truck Travel Time, the time required for trucks to travel empty to fill up to their capacity.

Travel time between grids. The centroids of the service grids are considered as visiting points for trucks. Let $\mathcal{E}_{i,j}$ represent the Euclidean distance between the centroids of grids i and j. To estimate the total travel time between consecutive visiting points, we adopt the method from Daganzo & Smilowitz (2004).

In a fully covered urban area $\mathcal{U}$ with N service locations, the total pairwise distance is $\sum \mathcal{E}_{i,j}$, and the number of points per unit area is $\frac{N}{\mathcal{U}}$. The expected distance between two random points is given by:

$$\frac{1.1}{\sqrt{N/\mathcal{U}}}.$$

For partial coverage, where only a subset of locations is covered, we adjust the formula to account for the spatial distribution of the grids. The expected distance becomes $\frac{1.1}{\delta_h\sqrt{N/\mathcal{U}}}$, where δ_h is the dispersion indicator, which adjusts the expected distance based on the distribution of selected grids in period h. In period h, $\sum_i x_{i,h}$ grids are covered out of N, and the total pairwise distance between covered grids is $\sum_{i,j} \mathcal{E}_{i,j}\chi_{i,j,h}$. The dispersion indicator is thus given by:

$$\delta_h = \frac{\sum_i x_{i,h}}{\sum_{i,j} \mathcal{E}_{i,j}\chi_{i,j,h}} \times \frac{N}{\sum_{i,j} \mathcal{E}_{i,j}}.$$

Assuming trucks travel at a constant speed V^{truck}, the total travel time is the product of the expected travel time and the number of trips $\sum_i x_{i,h}$, giving the total time during period h:

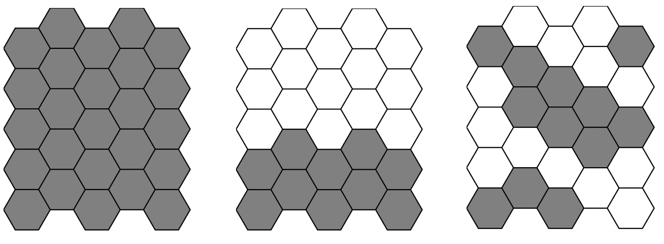
$$t_h^{\text{ped}} = \frac{1}{V^{\text{truck}}} \sum_i x_{i,h} \frac{1.1}{\delta_h \sqrt{N/\mathcal{U}}}.$$

The total travel time, considering all trips, is given by:

$$\begin{aligned} t_h^{\text{ped}} &= (1-\phi_{0,h}) \left[\frac{1.1N\sqrt{\mathcal{U}}}{V^{\text{truck}} \sum_i x_{i,h} \sqrt{N} \sum_{i,j} \mathcal{E}_{i,j}} \sum_i x_{i,h} \sum_{i,j} \mathcal{E}_{i,j} \chi_{i,j,h} \right] \\ &= (1-\phi_{0,h}) \left[\frac{1.1\sqrt{N\mathcal{U}}}{V^{\text{truck}} \sum_{i,j} \mathcal{E}_{i,j}} \sum_{i,j} \mathcal{E}_{i,j} \chi_{i,j,h} = T^{\text{CA}} \sum_{i,j} \mathcal{E}_{i,j} \chi_{i,j,h} \right] \qquad \forall h \in \mathcal{H} \end{aligned} \tag{3.22}$$

In (3.22), the term $(1-\phi_{0,h})$ ensures that the travel time is zero if the service is stopped during period h.

Example. Figure 3.2 illustrates different cases, where gray hexagons denote covered grids (trucks' visiting points) and white hexagons denote non-covered grids. Each grid covers an area of 1 km^2



(a) Full coverage *(b) Compact coverage* *(c) Scattered coverage*

Figure 3.2: Various distributions of visiting points (service locations) within the urban area.

The left plot (3.2a) represents full coverage with a density of 1 per km^2 (i.e., $\frac{N}{\mathcal{U}} = \frac{25}{25}$). In this case, the expected distance between each pair of grids is 1.1 km. The middle plot (3.2b)shows a concentrated coverage, where service grids are adjacent and form a continuous region. For this case, we can limit the whole area to the covered grids, and this scenario also maintains a density of 1 per km^2. The right plot (3.2c) represents randomly scattered covered grids, requiring a density adjustment due to the irregular distribution of the grids.

If we do not adjust for the spatial distribution with the dispersion indicator, and instead consider the total number of grids divided by the sum of their area, it would approximate the area as in the middle plot, where all covered grids are adjacent with no non-service grids between them. However, the proposed formula takes into account the ratio of the distance between covered grids to the total distance between grids in the fully covered scenario. This adjustment reflects the actual distribution and spacing of the covered grids.

Resource search and load/unload time. Each unit of repositioned resource requires a time of T^{service}. This time includes the time required for loading/unloading each resource, T^{rep}, and the time required for searching within each grid to find resources, T^{grid}. Thus, the total service time during each unit time of period h can be computed as follows:

$$(T^{\text{rep}} + T^{\text{grid}})r_h = T^{\text{service}} r_h$$

Empty truck travel time (line-haul). As trucks have a limited capacity K, they may need to make multiple trips to fully load or unload bikes. For each repositioning operation, the number of empty trips is proportional to the number of repositioning items r_h divided by the truck's capacity (i.e., r_h/K). This number is doubled because each empty trip is followed by a full trip. Let E_{LH} represent the average distance of each such trip, depending on the urban area. The total time spent on filling the capacity (starting a new repositioning path or by a new truck) trips during period h is:

$$\frac{2E_{LH}}{V^{\text{truck}}K} r_h = T^{\text{lh}} r_h.$$

Operational costs. Let C_h^{truck} denote costs of using one truck per shift during period h. Therefore, the operational costs during period h is computed as follows:

$$\mathscr{C}_h^{\text{O}} = \mathscr{N}_h^{\text{unit}} \frac{C_h^{\text{truck}}}{T^{\text{shift}}} \left[t_h^{\text{ped}} + r_h (T^{\text{lh}} + T^{\text{service}}) \right],$$

where $\left[t_h^{\text{ped}} + r_h (T^{\text{lh}} + T^{\text{service}}) \right]$ represents the total time required for repositioning of resources within each time interval. Dividing this value by the shift length T^{shift} gives the total number of truck-shifts within each time interval, each costing C_h^{truck}. Multiplying this by the total number of intervals in a period gives the total repositioning costs.

3.3.4 Periodic Profit

To compute the profit obtained during period h, we first need to compute the revenue. To this end, we define $R^+_{h,g}$ as the subscription fee for user segment g for period h. Users are also charged P_g per time unit of riding and an unlock fee of P_g^{lock} per trip. Moreover, the provider can salvage extra resources at the beginning of each period h that yield a revenue of B_h^+ per resource. Therefore, the revenue is written as below:

$$\mathscr{R}_h = \overbrace{(1-\phi_{0,h}) \sum_{g\in\mathcal{G}} R^+_{h,g} M_g \alpha_{g,h}}^{\text{Subscription revenue}} + \overbrace{\mathscr{N}_h^{\text{unit}} \sum_a \phi_{a,h} A_a^{\text{av}} \sum_{g\in\mathcal{G}} \alpha_{g,h} \sum_{i,j} (P_g T_{i,j,g} + P_g^{\text{lock}}) D_{i,j,g} \chi_{i,j,h}}^{\text{Trip revenue}} + \overbrace{B^+ z_h^-}^{\text{Salvage}}$$

The first term of the revenue represents the subscription revenue from active users, given by $M_g \alpha_{g,h}$, if the service is not stopped (i.e., $1-\phi_{0,h}$). The second term is the total trip revenue, including unlock and trip fee, which depends on the portion of served demand (selected availability level A_a^{av}). The trip demand is influenced by the proportion of active users ($\alpha_{g,h}$) and the covered locations ($D_{i,j,g}\chi_{i,j,h}$). This demand is per time unit, so it must be multiplied by the total number of time units per period. The last term represents the revenue from selling extra resources during period h.

Finally, the profit of period h is formulated as:

$$\mathscr{P}_h = \overbrace{\left[(1-\phi_{0,h})\sum_{g\in\mathcal{G}} R^{+}_{h,g} M_g \alpha_{g,h} + \mathscr{N}^{\text{unit}}_h \sum_a \phi_{a,h} A^{\text{av}}_a \sum_{g\in\mathcal{G}} \alpha_{g,h} \sum_{i,j} (P_g T_{i,j,g} + P^{\text{lock}}_g) D_{i,j,g} \chi_{i,j,h} + B^{+} z^{-}_h\right]}^{\text{Revenue}} -$$
$$\overbrace{\left[B^{-}_h z^{+}_h - M^{-}_h z_h - C^{\text{grid}}_h (1-\phi_{0,h}) \sum_i x_{i,h}\right]}^{\text{Investment costs}} - \overbrace{\mathscr{N}^{\text{unit}}_h \frac{C^{\text{truck}}_h}{T^{\text{shift}}} \left[t^{\text{ped}}_h + r_h (T^{\text{lh}} + T^{\text{service}})\right]}^{\text{Operational costs}} \tag{3.23}$$

Domain constraints. The domains of the defined decision variables are:

$$x_{i,h}, \chi_{i,j,h}, \phi_{a,h}, \beta^{+}_{g,h,h'} \in \{0,1\} \tag{3.24}$$
$$\alpha_{g,h}, \beta_{g,h,h'}, \pi^{a}_{gh}, \pi^{c}_{g,h}, \pi^{+c}_{g,h,h'} \in [0,1] \quad \& \quad z_h, z^{+}_h, z^{-}_h, z^{\text{extra}}_{i,h}, y_{i,h}, r_h \in \mathbb{Z}^{+} \quad \& \quad t^{\text{ped}}_h \in \mathbb{R}^{+} \tag{3.25}$$

3.4 Mixed-Integer Linear Formulation

In this section, we reformulate the dynamic programming problem (DP) as a deterministic Mixed-Integer Linear Programming (MILP), referred to as the *Phased Service-Adoption Model (PSAM)*. Specifically, we replace the recursive formulation of total profit with an equivalent objective that maximizes the sum of period-wise profits $\mathscr{P}_h$ over the planning horizon. We reformulate the probabilistic constraints (3.7) and (3.8), linearize the nonlinear expressions, and express the decision process as a deterministic sequence of period-by-period decisions.

Probabilistic constraints. Let accessibility satisfaction probability $\pi^{\text{a}}_{g,h}$ denotes the confidence level for satisfying accessibility aspirations of user segment g in period h. This way, we could deal with (3.7) as chance constraints wherein the confidence level is a decision variable:

$$\pi^{\text{a}}_{g,h} \leq \text{Prob}\left(\sum_{(i,j)} u_{i,j,g} \chi_{i,j,h} \geq \tilde{\mathscr{A}}^{\text{ac}}_g\right)$$

We assume that $\pi^{\text{a}}_{g,h}$ is selected from a finite set of B levels $\Pi^{a}_{b,g}$ for each period. $b = 0$ denotes zero accessibility satisfaction (i.e., $\Pi^{a}_{0,g} = 0$) and $b = B$ represents full accessibility satisfaction (i.e., $\Pi^{a}_{B,g} = 1$). We define the binary variable θ_{bgh} that is one if we decide to satisfy the accessibility expectations with the probability of $\Pi^{a}_{b,g}$ for segment g in period h.

We also define:

$$\Pi^{c}_{a,g} = \text{Prob}(A^{\text{av}}_a \geq \tilde{\mathscr{A}}^{\text{av}}_g) \qquad \forall a \in \{1,\ldots,A\}$$

For normally distributed aspiration levels $\tilde{\mathscr{A}}^{\text{ac}}_g \sim \mathcal{N}(\mu^{\text{ac}}_g, (\sigma^{\text{ac}}_g)^2)$ and $\tilde{\mathscr{A}}^{\text{av}}_g \sim \mathcal{N}(\mu^{\text{av}}_g, (\sigma^{\text{av}}_g)^2)$, constraints (3.7) and (3.8) can be reformulated using the following set of constraints:

$$\sum_{b=0}^{B} \theta_{b,g,h} = 1 \qquad \forall h \in \mathcal{H}, g \in \mathcal{G} \tag{3.26}$$

$$\pi^{\mathrm{a}}_{g,h} = \sum_{b=0}^{B} \theta_{b,g,h} \Pi^{a}_{b,g} \qquad \forall h \in \mathcal{H}, g \in \mathcal{G} \tag{3.27}$$

$$\sum_{i,j} u_{i,j,g} \chi^{h}_{i,j} \geq \mu^{\mathrm{ac}}_{g} + \sigma^{\mathrm{ac}}_{g} \mathbf{\Phi}^{-1}(\Pi^{a}_{b,g}) - (1 - \theta_{b,g,h}) \qquad \forall h \in \mathcal{H}, g \in \mathcal{G}, b \in \{0, \ldots, B\} \tag{3.28}$$

$$\sum_{i,j} u_{i,j,g} \chi^{h}_{i,j} \leq \mu^{\mathrm{ac}}_{g} + \sigma^{\mathrm{ac}}_{g} \mathbf{\Phi}^{-1}(\Pi^{a}_{b+1,g}) + (1 - \theta_{b,g,h}) \qquad \forall h \in \mathcal{H}, g \in \mathcal{G}, b \in \{0, \ldots, B-1\} \tag{3.29}$$

$$\pi^{c}_{g,h} = \sum_{a=0}^{A} \phi_{a,h} A^{\mathrm{av}}_{a} \Pi^{c}_{a,g} \qquad h \in \mathcal{H}, g \in \mathcal{G}. \tag{3.30}$$

Here, $\mathbf{\Phi}^{-1}(\cdot)$ denotes the inverse of the standard normal cumulative distribution function.

Non-linear terms. Constraints (3.9) contain the quadratic term $\alpha^2_{g,h-1}$. We define $\eta_{g,h} = (\alpha_{g,h})^2$ and use the piecewise linear approximation to model $\eta_{g,h}$. Let the breakpoints be $O_1 < O_2 < \cdots < O_M$, and define auxiliary variables $f_{m,h} \in [0, 1]$ such that:

$$\sum_{m=1}^{M} f_{m,h} = 1, \qquad \forall h \in \mathcal{H} \tag{3.31}$$

$$\alpha_{g,h} = \sum_{m=1}^{M} f_{m,h} O_m \qquad \forall h \in \mathcal{H}, g \in \mathcal{G} \tag{3.32}$$

Then the piecewise linear approximation of $\eta_{g,h} := (\alpha_{g,h})^2$ is[2]:

$$\eta_{g,h} \approx \sum_{m=1}^{M} f_{m,h} O^2_m \qquad \forall h \in \mathcal{H}, g \in \mathcal{G} \tag{3.33}$$

We also define the auxiliary variables $\kappa_{a,g,h} := \phi_{a,h}\eta_{g,h}$, $\zeta^{h}_{a,i,h} := \phi_{ah}x_{ih}$, $\nu_{a,g,h,h'} := \alpha_{g,h}\phi_{a,h'}$, $\dot{\nu}_{g,h} := \alpha_{g,h}(1 - \phi_{0,h})$, $\dot{\chi}_{i,j,h} = \chi_{i,j,h}(1 - \phi_{0,h})$, $\gamma_{a,g,i,j,h} := \nu_{a,g,h,h}\chi_{i,j,h}$, $\dot{\gamma}_{g,i,j,h} := \dot{\nu}_{g,h}\chi_{i,j,h}$, and $\tau_{a,b,g,h} := \nu_{a,g,h-1,h-1}\theta_{b,g,h}$ to linearize the corresponding bilinear using the following set of constraints that are collectively noted by $\mathcal{S}$:

$$\begin{aligned}
&\zeta_{a,i,h} \leq \phi_{a,h}, \quad \zeta_{a,i,h} \leq x_{i,h}, \quad \zeta_{a,i,h} \geq \phi_{ah} + x_{i,h} - 1 && \forall i \in \mathcal{I}, a \in \mathcal{A}, h \in \mathcal{H}\\
&\dot{\chi}_{i,j,h} \leq \chi_{i,j,h}, \quad \dot{\chi}_{i,j,h} \leq 1 - \phi_{0,h}, \quad \dot{\chi}_{i,j,h} \geq \chi_{i,j,h} - \phi_{0,h} && \forall i, j \in \mathcal{I}, h \in \mathcal{H}\\
&\tau_{a,b,g,h} \leq \theta_{b,g,h}, \quad \tau_{a,b,g,h} \leq \nu_{a,g,h-1,h-1}, \quad \tau_{a,b,g,h} \geq \theta_{b,g,h} + \nu_{a,g,h-1,h-1} - 1 && \forall a \in \mathcal{A}, 2 \leq h \leq H, g \in \mathcal{G}\\
&\nu_{a,g,h,h'} \leq \alpha_{g,h}, \quad \nu_{a,g,h,h'} \leq \phi_{a,h'}, \quad \nu_{a,g,h,h'} \geq \alpha_{g,h} + \phi_{a,h'} - 1 && \forall a \in \mathcal{A}, h, h' \in \mathcal{H}, g \in \mathcal{G}\\
&\dot{\nu}_{g,h} \leq \alpha_{g,h}, \quad \dot{\nu}_{g,h} \leq 1 - \phi_{0,h}, \quad \dot{\nu}_{g,h} \geq \alpha_{g,h} - \phi_{0,h} && \forall h \in \mathcal{H}, g \in \mathcal{G}\\
&\gamma_{a,g,i,j,h} \leq \nu_{a,g,h,h}, \quad \gamma_{a,g,i,j,h} \leq \chi_{i,j,h}, \quad \gamma_{a,g,i,j,h} \geq \nu_{a,g,h,h} + \chi_{i,j,h} - 1 && \forall i, j \in \mathcal{I}, a \in \mathcal{A}, h \in \mathcal{H}, g \in \mathcal{G}\\
&\dot{\gamma}_{g,i,j,h} \leq \dot{\nu}_{g,h}, \quad \dot{\gamma}_{g,i,j,h} \leq \chi_{i,j,h}, \quad \dot{\gamma}_{g,i,j,h} \geq \dot{\nu}_{g,h} + \chi_{i,j,h} - 1 && \forall i, j \in \mathcal{I}, h \in \mathcal{H}, g \in \mathcal{G}\\
&\kappa_{a,g,h} \leq \phi_{a,h}, \quad \kappa_{a,g,h} \leq \eta_{g,h}, \quad \kappa_{a,g,h} \geq \phi_{a,h} + \eta_{g,h} - 1 && \forall a \in \mathcal{A}, h \in \mathcal{H}, g \in \mathcal{G}
\end{aligned}$$

The MILP reformulation of the problem (DP) is given by model PSAM, defined as follows:

[2] In our numerical implementation, we divide the interval [0,1] into 20 equal segments using a step size of 0.05.

$$\max \sum_h (1+\Upsilon)^{-h} \mathscr{P}_h \tag{3.34}$$

s.t.

(3.1) − (3.6), (3.10), (3.13), (3.17) − (3.21), (3.26) − (3.33)

$$\mathscr{P}_h = \sum_{g\in\mathcal{G}} R^+_{h,g} M_g \dot{\nu}_{g,h} + \tag{3.35}$$

$$\mathscr{N}_h^{\text{unit}} \left[\sum_a A_a^{\text{av}} \sum_{i,j,g} (P_g T_{i,j,g} + P_g^{\text{lock}}) D_{i,j,g} \gamma_{a,g,i,j,h} \right] +$$

$$B^+ z_h^- - \left[B_h^- z_h^+ - M_h^- z_h - C_h^{\text{grid}} \sum_i \zeta_{0,i,h} \right] -$$

$$\left[\mathscr{N}_h^{\text{unit}} \frac{C_h^{\text{truck}}}{T^{\text{shift}}} \left(t_h^{\text{ped}} + r_h (T^{\text{lh}} + T^{\text{service}}) \right) \right] \qquad \forall h \in \mathcal{H}$$

$$\alpha_{g,h} = \sum_a \Pi^c_{ag} \nu_{a,g,h-1,h-1} + p_g \left(\sum_b \Pi^a_{b,g} \theta_{b,g,h} - \alpha_{g,h-1}\right) + \tag{3.36}$$

$$q_g \sum_{a,b} \Pi^c_{a,g} \Pi^a_{b,g} \tau_{a,b,g,h} - q_g \sum_{a,g} \Pi^c_{ag} \kappa_{a,g,h} + \sum_{h'} \beta_{g,h,h'} \qquad \forall h \geq 2, g \in \mathcal{G} \tag{3.37}$$

$$\alpha_{g,0} = p_g \sum_b \Pi^a_{b,g} \theta_{b,g,0} \qquad \forall g \in \mathcal{G} \tag{3.38}$$

$$\pi^{+c}_{g,h,h'} \leq \sum_a \Pi^c_{a,g} (\nu_{a,g,h',h} - \nu_{a,g,h,h}) + (1 - \beta^+_{h,h',g}) \qquad \forall h \geq 3, h' \leq h-2, g \in \mathcal{G} \tag{3.39}$$

$$\pi^{+c}_{g,h,h'} \geq \sum_a \Pi^c_{a,g} (\nu_{a,g,h',h} - \nu_{a,g,h,h}) \qquad \forall h \geq 3, h' \leq h-2, g \in \mathcal{G} \tag{3.40}$$

$$y_{i,h} \geq \sum_{a=0}^{A} \zeta_{a,i,h} \frac{A_a^{\text{av}}}{1 - \sum_a A_a^{\text{av}}} + \sum_{g,j} L^{\text{u}}_{i,j,g} D_{i,j,g} \dot{\gamma}_{g,i,j,h} + L^{\text{shift}} z_{ih}^{\text{extra}} \qquad \forall i \in \mathcal{I}, h \in \mathcal{H} \tag{3.41}$$

$$z_{i,h}^{\text{extra}} \geq \sum_{g,j} (D_{j,i,g} - D_{i,j,g}) \dot{\gamma}_{g,i,j,h} \qquad \forall i \in \mathcal{I}, h \in \mathcal{H} \tag{3.42}$$

$$t_h^{\text{ped}} = T^{\text{CA}} \sum_{i,j} \mathcal{E}_{i,j} \dot{\chi}_{i,j,h} \qquad \forall h \in \mathcal{H} \tag{3.43}$$

$$(\dot{\chi}_{i,j,h}, \nu_{a,g,h,h'}, \dot{\nu}_{g,h}, \gamma_{a,g,i,j,h}, \dot{\gamma}_{g,i,j,h}, \tau_{a,b,g,h}, \kappa_{a,g,h}) \in \mathcal{S} \tag{3.44}$$

$$x_{i,h}, \chi_{i,j,h}, \phi_{a,h}, \beta^+_{h,h',g}, \theta_{b,g,h} \in \{0,1\} \tag{3.45}$$

$$(\alpha_{g,h}, \beta_{g,h,h'}, \pi^a_{g,h}, \pi^{+c}_{g,h,h'}, \nu_{a,g,h,h'}, \dot{\nu}_{g,h}, \gamma_{a,g,i,j,h}, \dot{\gamma}_{g,i,j,h}, \tag{3.46}$$

$$\tau_{a,b,g,h}, \eta_{g,h}, \kappa_{a,g,h}, f_{m,h}, \zeta_{a,i,h}) \in [0,1]$$

$$z_h, z_h^+, z_h^-, z_{i,h}^{\text{extra}}, y_{i,h} \in \mathbb{Z}^+, r_h, t_h^{\text{ped}} \in \mathbb{R}^+ \tag{3.47}$$

In the objective function (3.34), the present value of profit in period h is computed by multiplying $\mathscr{P}_h$ with the discount factor $(1+\Upsilon)^{-h}$.

3.5 Numerical Study

In this section, we describe the dataset, define the baseline parameter settings, and construct scenarios and performance indicators to evaluate the proposed model using a bike-sharing service design as a showcase application.

3.5.1 Data Description

We use open-source trip data from CitiBike[3], a station-based bike-sharing service in New York and New Jersey, focusing on the Manhattan core (25 km^2) from January 1 to March 31, 2024. The area is divided into 26 hexagonal grids, each with a 460-meter edge length, corresponding to a walkable range of 300-500 meters (Kabra et al., 2020). This grid layout is shown in Figure 3.3.

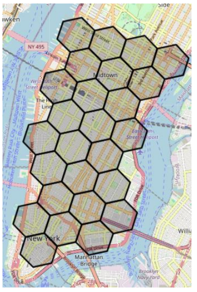


Figure 3.3: Hexagonal zone discretization of Manhattan core.

In the data, users are categorized as *members* (annual subscribers) and *casual users* (single-ride or day pass users). We assume all casual users are single-ride users, as fewer than 2% use day passes [4].

We use trip data filtered for the Manhattan core from January 1 to March 31, 2024, during the 6:00 AM to 6:00 PM period. The average trip duration is 10 minutes for members and 14 minutes for casual users. Since user identifiers are unavailable, we estimate segment sizes (M_g) using CitiBike's monthly operational statistics. Members take average 15 trips/month. This value is not available in the data for casual users. To ensure the single-trip fare is more economical than a monthly membership, casual users are assumed to take 5 trips/month. This gives an estimated 3,065 members and 1,036 casual users in the Manhattan core.

Pricing scheme. To ensure consistency with real-world practices, our pricing structure closely aligns with CitiBike's publicly available policy, as outlined in Table (3.2).

Table 3.2: Pricing scheme

User segment	**Subscription Fee**	**Unlock Fee**	**Per Minute Fee**
Member	20.00$/month	-	0.25 $
Casual (Single ticket)	-	5.00$/ride	0.35 $(after 30 min)

Segment-specific OD trips. Each day is divided into 3-hour time intervals, which are sufficient for trip completions, between 6:00 AM and 6:00 PM. Trip origin and destination stations are assigned to grids using GPS coordinates, which gives segment-specific spatio-temporal OD trip distributions. For each interval, we compute the average segment-specific OD trips between grid pairs (i.e., $D_{i,j,g}$). To reduce computational complexity, we aggregate adjacent grid pairs,

[3]https://citibikenyc.com/system-data
[4]From CitiBike monthly reports: January 2024, February 2024, March 2024.

resulting in a total of $N = 13$ candidate service grids. Based on Equation (3.18), ensuring a bike is available within 300 meters with probability $\mathcal{P} = 0.99998$ requires placing at least 12 randomly scattered bikes in each merged zone.

Utility of OD trips. We use the normalized segment-specific OD trips by the total number of trips per segment as the proxy of trips' utilities for each segment:

$$u_{i,j,g} = \frac{D_{i,j,g}}{\sum_{i',j} D_{i',j,g}}$$

The trip share is a suitable proxy for utility because the number of OD trips, $D_{i,j,g}$, is determined by the trip-generating capacity of origins ($\mathrm{TG}_{i,g}$), the trip-attracting capacity of destinations ($\mathrm{TA}_{j,g}$), and the ease of travel between i and j ($\mathrm{TE}_{i,j}$). This relationship is expressed as:

$$D_{i,j,g} = \mathcal{F}(\mathrm{TG}_{i,g}, \mathrm{TA}_{j,g}, \mathrm{TE}_{i,j}),$$

where $\mathcal{F}(\cdot)$ models the interactions among these factors. Therefore, $\frac{D_{i,j,g}}{\sum_{i',j} D_{i',j,g}}$ inherently captures user preferences and encapsulates the frequency and value of OD trips for users. With these definitions, the aspiration threshold, $\mathscr{A}_g^{\mathrm{ac}}$, is interpreted as the fraction of covered OD trips.

Costs. We assume a purchase price of \$3,000 per bike and a salvage value of \$500. Annual city costs include a permit fee of \$1,000 per grid and \$30 per bike, paid to municipal authorities[5]. Each bike also incurs an annual maintenance cost of 200 \$. The annual interest rate Υ is set to 5%.

Repositioning parameters. We consider homogeneous repositioning trucks, each with a capacity of 20 bikes and an operational cost of \$30 per hour, operating in 1-hour shifts. Trucks are assumed to travel at an average speed of 60 km/h.

We assume a fixed time of 10 minutes for searching and collecting bikes within merged grid areas ($T^{\mathrm{grid}} = 10$ minutes), and 1 minute for each loading or unloading operation. We also set $\mathrm{E_{LH}} = 6.8$ km, approximated as half of the total distance between all grids.

Trip duration $T_{i,j,g}$ is estimated using centroid-to-centroid distance at 20 km/h, with 8-minute and 10.5-minute adjustments for members and casual users, respectively, reflecting 75% of their average observed durations in the data.

Candidate Levels for Availability and Accessibility. The availability level $\phi_{a,h}$ is chosen from $\{0.85, 0.90, 0.92, 0.94, 0.96, 0.98\}$, and the target accessibility probability $\theta_{b,g,h}$ from $\{0, 0.35, 0.55, 0.65, 0.75, 0.85, 1\}$.

3.5.2 Baseline Parameter Settings

The Bass diffusion parameters, innovation (p) and imitation (q), can be estimated from surveys or analogous services. Empirical studies show p typically lies between 0.01 and 0.03, and q between 0.3 and 0.7 (Mahajan et al., 1995).

[5]These values are informed by data from https://micromobility.mitre.org/resources/city-survey/, which reports location-specific costs across various cities. We use this as a reference to define realistic values.

Time granularity. The Bass model parameters depend on the time scale of estimation. For a coarse time step (e.g., one year), let p and q denote the innovation and imitation coefficients, and p' and q' represent corresponding values over a finer interval (e.g., one month). A commonly used scaling heuristic (Putsis Jr, 1996) relates (p, q) to (p', q') via:

$$p \approx n \cdot p', \quad q \approx n \cdot q'.$$

For discrete three-month periods, we use values from Dan et al. (2018) for Mobike: $p' = 0.0138$, $q' = 0.2727$, yielding:

$$p = 0.0138 \times 3 = 0.04, \quad q = 0.2727 \times 3 = 0.8.$$

We define $\tilde{\mathscr{A}}_g^{\text{ac}}, \tilde{\mathscr{A}}_g^{\text{av}} \in [0, 1]$ and $\sum_{(i,j)} u_{ijg}\chi_{ijh} \in [0, 1]$. Thus, $\theta_{Bgh} = 1$ when all locations are covered, and $\theta_{0gh} = 1$ when none are. Similarly, for 100% availability, $\Pi_{ag}^{c} = 1$, and for 0% availability, $\Pi_{ag}^{c} = 0$. To ensure this, we model $\tilde{\mathscr{A}}_g^{\text{ac}}$ and $\tilde{\mathscr{A}}_g^{\text{av}}$ using truncated Normal distributions:

$$\mathbf{\Phi}^{-1}(\Pi_{0g}^{a} = 0) := -\mu_g^{\text{ac}}, \quad \mathbf{\Phi}^{-1}(\Pi_{Bg}^{a} = 1) := 1 - \mu_g^{\text{ac}},$$
$$\mathbf{\Phi}^{-1}(A_0^{\text{av}} = 0) := 0, \quad \mathbf{\Phi}^{-1}(A_A^{\text{av}} = 1) := 1.$$

The baseline adoption and aspiration parameters are summarized in Table 3.3. We assume members have higher expectations than casual users due to their greater commitment through subscription.

Table 3.3: Baseline Adoption Parameters (3-Month Period)

Segment	p_g	q_g	J	$\tilde{\mathscr{A}}_g^{\text{ac}}$	$\tilde{\mathscr{A}}_g^{\text{av}}$
Members (g=1)	0.04	0.8	0	$\sim \mathcal{N}(0.50, 0.03^2)$	$\sim \mathcal{N}(0.85, 0.03^2)$
Casual users (g=2)	0.04	0.8	0	$\sim \mathcal{N}(0.45, 0.03^2)$	$\sim \mathcal{N}(0.80, 0.03^2)$

In the following sections, we vary (p_g, q_g) and aspiration levels to evaluate their impact on the results.

3.5.3 Scenarios and Indicators

We simulate different scenarios of adoption behavior (e.g., baseline setting, faster due to an aggressive marketing campaign, or slower due to strong influence of uncertainty about the service quality on the population) and trip patterns to evaluate the system performance under various conditions in terms of: time to the maximum penetration and stabilized adoption as well as the operational and financial metrics like fleet size, revenue, and profit.

Adoption Scenarios

In the Bass diffusion model, the innovation coefficient defines the steepness of the adoption curve during the early stages of the service introduction, and the imitation coefficient plays a significant role in shaping the adoption curve as time elapses in the later stages. Accordingly, we define four distinct adoption behavior scenarios by varying these coefficients to examine the impact of adoption dynamics on service design. Each scenario reflects a theoretically grounded behavioral pattern:

- **Low-Innovation Adoption:** The innovation effect is weakened by halving the coefficient p_g while maintaining the baseline imitation coefficient q_g. This scenario captures a context where external influences, such as marketing and media exposure, have a limited impact compared to social influence. We set $p_1 = p_2 = 0.02$, $q_1 = q_2 = 0.8$.
- **Low-Imitation Adoption:** The imitation effect is reduced by halving the coefficient q_g, while the innovation coefficient remains unchanged. This represents environments where social influence is weaker. We use $p_1 = p_2 = 0.04$, $q_1 = q_2 = 0.4$.
- **Slow Adoption:** Both innovation and imitation coefficients are reduced by half, resulting in a slow pace of adoption across all user segments. This configuration models conservative markets with limited external stimuli and social influence. Here, $p_1 = p_2 = 0.02$, $q_1 = q_2 = 0.4$.
- **Segment-Specific Adoption:** Members adopt more slowly ($p_1 = 0.01$, $q_1 = 0.4$), while casual users adopt rapidly ($p_2 = 0.1$, $q_2 = 0.8$), reflecting heterogeneity in user behavior.

Demand Patterns

We consider four demand scenarios that vary in origin-destination (OD) trip patterns and/or total trip volume, as follows:

- **Baseline Demand (BD):** OD patterns are derived directly from CitiBike trip data, serving as the reference case.
- **Concentrated Demand (CD):** For each user segment, OD pairs are ranked by trip volume. The top $2N = 26$ arcs (where N is the number of grids) retain their original demand, while all remaining arcs have their demand reduced to 25% of the baseline level, i.e., $0.25 \times D_{i,j,g}$.
- **Decentralized Demand (DD):** Total outbound demand per grid is maintained as in the baseline; however, OD flows are redistributed. Grids are partitioned into two disjoint subsets, one subset is composed of the high-demand grids, with no trips allowed between them. Outbound demand from each grid is randomly reassigned within its own subset.
- **Low-Decentralized Demand (LD):** This scenario mirrors the decentralized pattern in structure but halves the total outbound demand of each grid compared to the decentralized setting.

Performance Indicators

The main indicators in our results are: time to maximum market penetration (h_g^*), time to adoption steady state ($\dot{h}_g$), profitability time threshold (h^+), fleet size, and total profit.

Time to maximum market penetration (h_g^*). Adoption follows an S-shaped curve, peaking at h_g^*, which is approximated by (Satoh, 2001):

$$h_g^* = \arg\max_h \{\alpha_{g,h} - \alpha_{g,h-1}\} = -\frac{1}{p_g + q_g} \ln\left(\frac{p_g}{q_g}\right) \tag{3.48}$$

For full accessibility satisfaction, the peak typically occurs after the fourth period (end of year one). Reduced accessibility or availability may delay this peak.

Table 3.4 shows the estimated periods to reach maximum market penetration growth, calculated using Equation (3.48), for each adoption scenario.

Table 3.4: Time (number of 3-month periods) to maximum market penetration (h^*) for different adoption scenarios.

Adoption setting	p_g	q_g	h_g^*
Baseline	0.04	0.8	$\sim$ 4–5
Slow innovation	0.02	0.8	$\sim$ 5–6
Slow imitation	0.04	0.4	$\sim$ 5–6
Slow adoption-I	0.02	0.4	$\sim$ 7–8
Slow adoption-II	0.01	0.4	$\sim$ 9–10
Fast adoption	0.10	0.8	$\sim$ 2–3

Time to stabilized adoption ($\dot{h}_g$). This is the period when adoption reaches a steady state, meaning the marginal adoption increase is negligible. It satisfies:

$$\dot{h}_g = \arg\min_h \{\alpha_{g,h} - \alpha_{g,h-1} = 0\}.$$

Profitability time threshold (h^+). This is the first period when cumulative revenue exceeds costs, yielding positive profit:

$$h^+ = \arg\min_{h'} \left\{ \sum_{h=1}^{h'} \mathscr{P}_h \geq 0 \right\}.$$

3.6 Results

We solve the PSAM model using Gurobi 11.0.1 on a system equipped with 10 GB of RAM to obtain the optimal service design across various defined scenarios. The results are presented in the following sections.

Calibration and planning horizon. Time plays a critical role in shaping both adoption dynamics and system profitability. Using baseline adoption parameters (Table 3.3) and the origin-destination (OD) pattern extracted from CitiBike trip data for Manhattan core, we first calibrate an appropriate planning horizon for system evaluation.

We solved the PSAM model under varying planning durations, from short-term (1–11 months) to longer-term horizons (1–4 years). Results show that the service does not achieve profitability within the first year. The capital investment required for fleet acquisition and deployment is too large to be offset by early-stage adoption, even under full spatial coverage and high availability.

Based on these calibration tests, we selected a 3-year planning horizon, divided into a one-year implementation phase (4 periods) and a two-year post-deployment phase (8 periods), i.e., $H_1 = 4$, $H_2 = 8$, and $H = 12$. Within this horizon, the service reaches near-complete adoption, making it sufficient for comparing performance across strategies.

3.6.1 Adoption Behavior

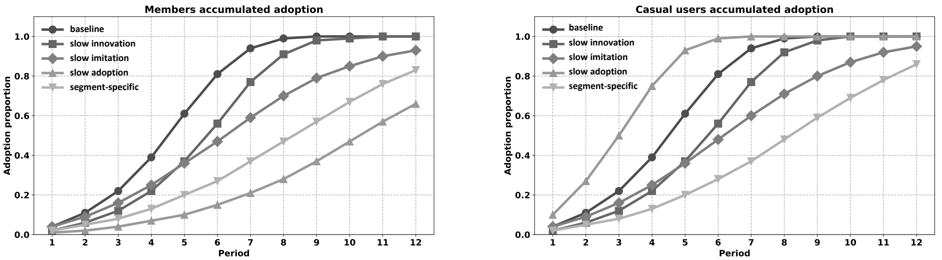
Table 3.5: Service design under different adoption scenarios over a three-year planning horizon ($H = 12$, $H_1 = 4$)

Adoption pattern	Grids No.	Fleet size	Availability level	(h_1^*, h_2^*)	$(\dot{h}_1, \dot{h}_2)$	h^+	Profit($)
Baseline	[13, 13, 13, 13]	[160, 225, 240, 332 *]	[0.92, 0.94, 0.94, 0.94, 0.94, 0.94, 0.94, 0.94, 0.94, 0.94, 0.94, 0.92]	(4, 4)	(9, 9)	5	8,194,789
Slow innovation	[11, 13, 13, 13]	[179, 218, 226, 332 *]	[0.94, 0.94, 0.94, 0.94, 0.94, 0.94, 0.94, 0.94, 0.94, 0.94, 0.94, 0.92]	(6, 6)	(11,10)	6	6,837,272
Slow imitation	[13, 13, 13, 13]	[214, 224, 231, 270 *]	[0.94, 0.94, 0.94, 0.94, 0.94, 0.94, 0.92*]	(6, 5)	$(\times,\times)$	5	6,005,536
Slow adoption	[13, 13, 13, 13]	[157, 162, 221, 250, 250, 250, 250, 250, 250, 250, 250, 226]	[0.92, 0.92, 0.94, 0.94, 0.94, 0.94, 0.92, 0.92, 0.92, 0.92, 0.92, 0.9]	(6,7)	$(\times,\times)$	6	4,063,003
Segment-specific	[13, 13, 13, 13]	[216, 223, 235, 281, 281, 281, 281, 281, 281, 252*]	[0.94, 0.94, 0.94, 0.94, 0.94, 0.94, 0.94, 0.94, 0.94, 0.92, 0.92, 0.92]	(9,3)	$(\times,7)$	4	5,453,661

$\star$ shows the fixed value until the end of the planning horizon.
$\times$ shows that the adoption does not reach the steady state before the end of the planning horizon.

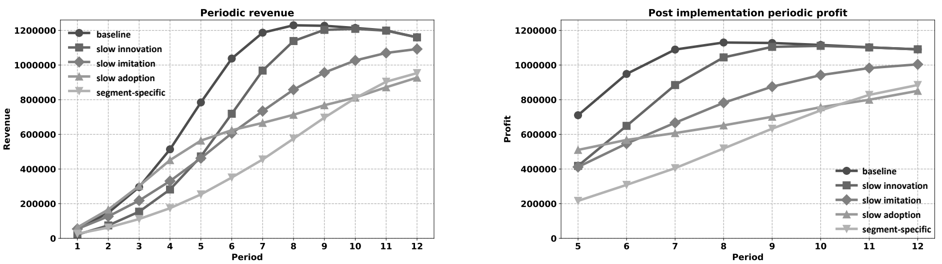
Table 3.5 summarizes the outcomes across different adoption scenarios. The first column lists the scenarios described in Section 3.5. “Grid No.” reports the number of grids selected as service locations (i.e., service grids) during the implementation phase ($h \in \{1, 2, 3, 4\}$). No expansion occurs beyond this phase, in accordance with the problem description. Subsequent columns present the fleet size, availability levels, segment-specific values h_g^* and $\dot{h}_g$, and h^+. The final column, “Profit ($),” shows the total profit (summed over all periods) in dollars. The accumulated adoption curves are displayed in Figure 3.4. Periodic revenue and profit are also displayed in Figure 3.5.

In all scenarios except slow innovation, the optimal design starts with full coverage of the urban area. In contrast, under slow innovation, where early demand is particularly low, the model begins with partial coverage, focusing on the 11 out of 13 locations that fully meet users’ accessibility aspirations (i.e., $\pi_{g,h}^a = 1$). This targeted approach minimizes initial grid and fleet costs without hindering adoption, as users’ accessibility expectations are fully satisfied. Partial coverage under low-demand cases is also observed in the analysis of demand pattern scenarios presented in the next section, highlighting accessibility as a cost-control lever rather than a limiting factor on adoption.



(a) Members accumulated adoption ($\alpha_{1,h}$) *(b) Casual users accumulated adoption ($\alpha_{2,h}$)*

Figure 3.4: Members and casual users accumulated adoption for various adoption scenarios.



(a) Periodic revenue $\mathscr{R}_h$ *(b) Post implementation periodic profit $\mathscr{P}_h$*

Figure 3.5: Periodic revenue and profit for various adoption scenarios.

Availability levels range between 0.92 and 0.94, both sufficient to fully satisfy active users (i.e., confirmation probability $\pi^c_{g,h} \approx 1$). While 0.94 yields more served trips (and higher revenue), it requires a larger fleet and increases costs. For the slow adoption, slow imitation, and segment-specific adoption scenarios, availability is strategically reduced from 0.94 to 0.92 shortly after the peak adoption growth period h_1^* for members, who account for the majority of trips compared to casual users according to CitiBike reports. Since user satisfaction remains unaffected, the reduction in availability serves as a cost-saving measure, avoiding overinvestment in fleet capacity at a stage when the cost of high availability is no longer justified by corresponding gains in revenue.

The results, shown in Figure 3.4 and Table 3.5, indicate that slower innovation and/or imitation delays both the onset and magnitude of adoption growth. Faster adoption leads to an earlier peak in adoption growth, but this does not necessarily mean the service reaches saturation sooner. For example, in the slow innovation case, casual users' adoption growth peaks ($h_2^* = 6$) later than in a slow innovation scenario ($h_2^* = 5$), yet still eventually reaches saturation ($\dot{h}_2 = 11$). In contrast, in the slow imitation case, saturation does not occur.

When innovator adoption is slow, the early-stage uptake is limited, which in turn suppresses imitation effects throughout the horizon, see Figure 3.4. In the case of slow imitation, even with moderate initial adoption, the weaker social influence results in flatter and slower adoption growth. Both patterns reduce the cumulative number of adopters over time and postpone saturation, affecting long-term demand and system profitability as discussed below.

Figure 3.5 shows that the baseline scenario consistently generates the highest revenue across all periods, driven by a balanced adoption pace (p_g and q_g) among both user segments. In contrast, the slow adoption scenario yields the lowest revenue due to the slowest adoption growth across the scenarios.

The slow innovation scenario starts with lower revenue compared to the baseline case, reflecting the delayed adoption by innovators, but its revenue grows fast in later periods and reaches the baseline revenue after adoption saturation. While the slow imitation scenario generates relatively high revenue in the early stages due to fast innovator adoption. Yet, its revenue growth flattens in later periods because of the limited imitation effect.

The segment-specific case shows the highest early revenue, driven by rapid adoption among casual users. However, it experiences a steep revenue decline in the second half of the horizon. This drop stems from extremely slow member adoption, who account for the majority of trips.

It is worth mentioning that a part of the observed late-stage revenue decline across scenarios is due to discounting, as revenue is reported in present value terms.

The differences in profit trajectories across adoption scenarios, as shown in Figure 3.5, are primarily driven by revenue patterns and variations in fleet-related costs. In the post-implementation phase, where no additional capital investments are made, profit trends closely follow the revenue patterns discussed above.

Notably, a smaller value of h^+, indicating an earlier point at which profit becomes positive, does not necessarily correspond to higher total profit. For example, the segment-specific scenario reaches profitability the fastest but yields the second-lowest total profit due to slow member adoption. The reason is that members account for the majority of trips as reported in the data description.

3.6.2 Demand Pattern

This section examines how demand patterns, including the number of potential trips (i.e., demand size) and their distribution, and availability aspirations shape the optimal system configuration and outcomes. Three factors are varied: (i) demand pattern (baseline (i.e., B) vs. concentrated (i.e., CD) vs. decentralized (i.e., DD) vs. low demand (i.e., LD)), (ii) higher availability expectations, and (iii) the pace of adoption (baseline vs. slow adoption).

Table 3.6 summarizes the results. The first column ("OD") refers to the demand patterns introduced earlier.The second column indicates the adoption scenario: "Base" corresponds to the baseline case with $(p_g, q_g) = (0.04, 0.8)$, while "Slow" represents the slow adoption case with $(p_g, q_g) = (0.02, 0.4)$. The third column shows the availability aspiration level. If the aspirations match those reported in Table 3.3, we label them as "Base." The label $\mathscr{A}^{\text{av}\uparrow}$ indicates that the mean of $\tilde{\mathscr{A}}^{\text{av}}$ has been increased to 0.95 for members and 0.9 for casual users. The remaining columns are consistent with those in the previous tables.

Scenarios with higher availability aspirations ($\mathscr{A}^{\text{av}\uparrow}$) lead to significantly larger fleet sizes and increased operating costs. To manage this, the system adopts a gradual expansion strategy: rather than starting with full network coverage, it begins with partial coverage that ensures full accessibility ($\pi^a_{g,h} = 1$), effectively balancing cost control with user satisfaction. In low-demand scenarios (i.e., LD and CD), with the baseline availability aspiration levels, the service also be-

Table 3.6: Service design under different demand patterns

OD	Adoption	Availability aspiration	Grid No.	Fleet size	Availability level	h^+	Profit($)
BD	Base	$\mathscr{A}^{\text{av}\uparrow}$	[9, 13, 13, 13]	[450, 656, 670, 754, 754, 754, 754, 754, 754, 754, 754, 265]	[0.98, 0.98, 0.98, 0.98, 0.98, 0.98, 0.98, 0.98, 0.98, 0.98, 0.98,0.92]	7	6,092,264
CD	Base	Base	[12, 12, 13, 13]	[147, 151, 175, 239, 239, 239, 239, 239, 239, 239, 239, 213]	[0.92, 0.92, 0.92, 0.92, 0.92, 0.92, 0.92, 0.92, 0.92, 0.92, 0.92, 0.9]	6	3,104,499
	Slow	Base	[8, 8, 13, 13]	[96, 97, 163, 212, 212, 212, 212, 212, 212, 212, 201, 201]	[0.92, 0.92, 0.92, 0.92, 0.92, 0.92, 0.92, 0.92, 0.92, 0.92, 0.9, 0.9]	8	1,336,861
	Base	$\mathscr{A}^{\text{av}\uparrow}$	[7, 7, 13, 13]	[351, 352, 352 ,369, 369, 369, 369, 369, 369, 369, 369, 208]	[0.98, 0.98, 0.96, 0.96, 0.96, 0.96, 0.96, 0.96, 0.96, 0.96, 0.96, 0.92]	8	1,452,878
DD	Base	Base	[13, 13, 13, 13]	[163, 227, 241, 346, 346, 346, 346, 346, 346, 346, 346, 292]	[0.92, 0.94, 0.94, 0.94, 0.94, 0.94, 0.94, 0.94, 0.94, 0.94, 0.94, 0.92]	5	7,373,502
	Slow	Base	[13, 13, 13, 13]	[215, 216, 223, 316, 316, 316, 316, 316, 316, 316, 316, 270]	[0.94, 0.94, 0.94, 0.94, 0.94, 0.94, 0.94, 0.94, 0.94, 0.94, 0.94, 0.92]	7	3,569,689
	Base	$\mathscr{A}^{\text{av}\uparrow}$	[9, 9, 13, 13]	[451, 456, 671, 766, 766, 766, 766, 766, 766, 766, 766, 277]	[0.98, 0.98, 0.98, 0.98, 0.98, 0.98, 0.98, 0.98, 0.98, 0.98, 0.98, 0.92]	7	5,388,344
LD	Base	Base	[7, 13, 13, 13]	[84, 165, 170, 224, 224, 224, 224, 224, 224, 224, 224, 195]	[0.92, 0.92, 0.92, 0.92, 0.92, 0.92, 0.92, 0.92, 0.92, 0.92, 0.92, 0.9]	6	3,305,023
	Slow	Base	[7, 7, 13, 13]	[84, 85, 164, 207, 207, 207, 207, 207, 207, 207, 207, 186]	[0.92, 0.92, 0.92, 0.92, 0.92, 0.92, 0.92, 0.92, 0.92, 0.92, 0.92, 0.9]	8	1,454,210
	Base	$\mathscr{A}^{\text{av}\uparrow}$	[7, 7, 13, 13]	[350, 351, 351, 362, 362, 362, 362, 362, 362, 362, 362, 175]	[0.98, 0.98, 0.96, 0.96, 0.96, 0.96, 0.96, 0.96, 0.96, 0.96, 0.96, 0.9]	8	1,626,919

gins with partial coverage and operates with smaller fleets and reduced availability levels. These configurations lead the model to prioritize high-demand zones that guarantee full accessibility satisfaction and expand gradually. In the decentralized demand case (DD), although the total demand matches that of the baseline (BD), the uneven trip distribution increases rebalancing needs, requiring a larger fleet to maintain service. Despite using similar design parameters (accessibility and availability), DD yields lower profits, approximately 10% lower than BD under baseline adoption and aspirations, and up to 12% lower under high availability expectations. To offset these costs, the model again opts for partial early-stage coverage while preserving full accessibility, unlike BD, which maintains full coverage from the outset.

3.6.3 Adoption Modeling Strategies

In this section, we compare our proposed adjusted Bass adoption model in Eq. (A7), which incorporates the effects of service quality and social influence, and enables alignment between service design and adoption dynamics through the PSAM service design model, with the following two simplified models:

- **Static Adoption Model.** This model assumes the adoption process has already stabilized, commonly used in service design studies. Demand is estimated based on the ultimate number of adopters whose accessibility and availability aspirations are satisfied. Service design (service grids, fleet, and availability) is decided upfront and held fixed. The adoption proportion is computed as:

$$\alpha_{g,h} = \pi^{\mathrm{a}}_{g,h}\pi^{\mathrm{c}}_{g,h}, \quad p_g = 1, \quad q_g = 0,$$

  implying immediate adoption by potential users whose expectations are met. This model ignores the social influence and overestimates early demand.

- **Original Bass Model.** This model accounts for social influence while ignoring the impacts of service quality on adoption proportion. It is formulated as below:

$$\alpha_{g,h} = (p_g + q_g\alpha_{g,h-1})(1 - \alpha_{g,h-1}).$$

Table 3.7: Comparison of adoption modeling strategies

Adoption model	Service quality	Social influence
Proposed adjusted Bass	✓	✓
Static adoption	✓	×
Original Bass	×	✓

In all modeling strategies, user adoption is influenced by availability and accessibility. The static and original Bass adoption models differ from our proposed model in how the adoption is estimated from the service provider's perspective. For comparison, the service designs generated by each model are evaluated using our PSAM model to compute realized adoption, costs, and profits under quality and social-driven adoption behavior.

We evaluated the models under three representative conditions: (i) baseline adoption and availability settings, (ii) slow adoption pattern ($p_g = 0.02, q_g = 0.4$) with baseline service aspirations, and (iii) high availability expectations ($\mathscr{A}^{\mathrm{av}\uparrow}$ represents increased mean of availability aspiration to 0.95 for members and to 0.9 for casual users) with baseline adoption pattern.

Table 3.8 summarizes the resulting system designs. The first column lists the adoption models, as described above. The next two columns follow the same structure and definitions as in Table 3.6, covering adoption scenarios and availability aspiration levels. The final column reports profit losses relative to the service design using the adjusted Bass adoption model. The remaining columns are consistent with those presented in the previous tables.

The results highlight the value of gradual deployment aligned with the adoption growth enabled by the proposed PSAM model in this study. The static adoption model assumes immediate full adoption and plans for steady-state demand from the beginning. This leads to unnecessary early investment, especially when adoption is slow. Under the slow adoption scenario, the static model results in a 3.8$\%$ profit loss. Even under high availability expectations, where large fleets are eventually needed, PSAM outperforms by deferring costs until justified by demand.

The importance of service quality inclusion in adoption formulation and designing the service accordingly is also evident when comparing the original Bass model against our proposed formulation. Neglecting the service quality impacts on adoption has a limited impact under

Table 3.8: Service design under alternative adoption models compared with our proposed adjusted Bass formulation

Model	Adoption	Availability aspiration	Gird No.	Fleet size	Availability level	h^+	Profit loss
Static	Base	Base	13*	332*	0.94*	5	0.4%
	Slow	Base	13*	332*	0.94*	7	3.8%
	Base	$\mathscr{A}^{\mathrm{av}\uparrow}$	13*	758*	0.98*	7	3.5%
Original Bass	Base	Base	13*	[214, 225, 240, 332*]	0.94*	5	0.02%
	Slow	Base	13*	[157, 162, 168, 260*]	0.92*	6	2%
	Base	$\mathscr{A}^{\mathrm{av}\uparrow}$	13*	[214, 225, 240, 332*]	0.94*	6	54%

* indicates fixed values for the remainder of the planning horizon.

baseline conditions (0.02% profit loss) and a moderate impact under slow adoption (2% profit loss). However, when user expectations are high, ignoring the effect of service quality on adoption leads to significant adopter loss and a 54% drop in total profit.

3.6.4 Value of Phased Implementation with Belief-Updating Mechanism

In earlier sections, we showed that accounting for gradual adoption yields up to 4% higher profit compared to strategies that assume full, stabilized adoption from the outset. While this improvement may appear modest under ideal parameter knowledge, its value becomes significantly amplified when adoption behavior is uncertain or misestimated.

To examine the value of gradual implementation with a belief-updating mechanism and adjustment of service design in response to adoption growth, we compare two design strategies under parameter uncertainty:

- **Belief-updating strategy:** Gradual implementation with periodic re-optimization based on observed adoption outcomes. Adoption coefficients are treated as random variables and are sequentially updated via Bayesian inference.
- **Upfront strategy:** Full service design is fixed at the beginning based on initial beliefs about adoption rates, without adjustment or belief-updating.

In the following sections, we first describe the belief-updating framework and the simulation setup used to evaluate it. We then present the results under both the updating-based and upfront strategies to assess the impact of decision adjustments enabled by phased implementations.

Bayesian Belief-Updating Framework

We assume that the real adoption coefficients p_g and q_g are unknown. Thus, the innovation and imitation coefficients are modeled as random variables, denoted by a tilde sign, following normal distributions:

$$\tilde{p}_g \sim \mathcal{N}\left(\mu_g^p, (\sigma_g^p)^2\right), \quad \tilde{q}_g \sim \mathcal{N}\left(\mu_g^q, (\sigma_g^q)^2\right).$$

At each period h, we use our current estimate of the expected value of $\tilde{p}_g$ and $\tilde{q}_g$, which are denoted by $p_{g,h}$ and $q_{g,h}$, receptively. We also note the observed values by the hat sign (e.g., $\hat{p}_g$, $\hat{q}_g$).

We define $\xi^p_{g,h} = 1/(\sigma^p_g)^2$ and $\xi^q_{g,h} = 1/(\sigma^q_g)^2$ as the precision of our belief at iteration (period) h. Moreover, ϵ denotes the precision of observations. At each period h, we observe the adoption growth:

$$\Delta\hat{\alpha}_{gh} = \hat{\alpha}_{g,h} - \hat{\alpha}_{g,h-1} = (p_g + q_g\hat{\alpha}_{g,h-1})\pi^c_{g,h-1}(\pi^a_{g,h} - \hat{\alpha}_{g,h-1}),$$

Given $\pi^c_{g,h-1}$, $\pi^a_{g,h}$, and $\Delta\hat{\alpha}_{g,h-1}$, we define

$$\Xi_{g,h} = \frac{\Delta\hat{\alpha}_{g,h}}{\pi^c_{g,h-1}(\pi^a_{g,h} - \hat{\alpha}_{g,h-1})},$$

and perform conditional updates:

- Update μ^p_g given the current belief on $\mu^q_g = q_{g,h}$,

$$\hat{p}_g = \Xi_{g,h} - \mu_{qg}\hat{\alpha}_{g,h-1}, \quad p_{g,h+1} = \frac{\epsilon\hat{p}_g + \xi^p_{g,h}p_{g,h}}{\epsilon + \xi^p_{g,h}}, \quad \xi^p_{g,h} = \xi^p_{g,h-1} + \epsilon.$$

- Update μ^q_g given $p_{g,h+1}$,

$$\hat{q}_g = \frac{\Xi_{g,h} - p_{g,h+1}}{\hat{\alpha}_{g,h-1}}, \quad q_{g,h+1} = \frac{\epsilon\hat{q}_g + \xi^q_{g,h}q_{g,h}}{\epsilon + \xi^q_{g,h}}, \quad \xi^q_{g,h} = \xi^q_{g,h-1} + \epsilon.$$

This belief updating mechanism follows the Bayesian sequential belief-updating framework Powell (2022) and ensures that our parameters' estimations adapt dynamically to new information.

We initialize the prior precisions as $\xi^p_{g,0} = \xi^q_{g,0} = 1$, and set the observation precision to $\epsilon = 10$ to reflect high trust in observations during early iterations, with increasing the confidence in the updated beliefs over time. This belief-updating and re-optimization process is summarized in Algorithm 1.

Algorithm 1 Belief Updating and Re-Optimization Process

```
Initialize: prior beliefs μ_g^{p0}, μ_g^{q0}; precision ξ^p_{g,0}, ξ^q_{g,0} = 1; observation precision ϵ = 10
for each period h = 1, ..., H do
    Solve the model using current belief parameters
    Observe adoption growth: Δα̂_{g,h}
    Compute Ξ_{g,h} and update estimates of p_{g,h+1}, q_{g,h+1}
    Update belief precisions ξ^p_{g,h+1} and ξ^q_{g,h+1}
    Re-optimize the PSAM model for the remaining periods by fixing the decision variables
      for all periods up to and including period h.
end for
```

Simulation Setup and Belief Updates

Belief-updating occurs at the end of each implementation period, triggering a new iteration where adoption coefficients are updated and future decisions can be adjusted based on improved parameter estimates. In this section, we adopt a 3-year planning horizon as before but extend the implementation phase from 4 to 6 periods, with $\mathcal{H}_1 = 6$ and $\mathcal{H}_2 = 6$. The extended implementation phase allows the model to make adjustments over more periods (6 rather than 4) as new information becomes available.

We simulate two types of estimation errors:

- **Over-estimation:** Prior belief overstates real adoption; observed values of p_g, q_g decline over time.
- **Under-estimation:** Prior belief underestimates adoption; observed values increase over time.

Table 3.9 summarizes the initial beliefs about the adoption coefficients, the observed values, and the updated beliefs using the belief-updating framework described in Algorithm 1 for the two underestimation and overestimation scenarios discussed above.

Table 3.9: Prior, observed, and updated values of p_g and q_g in over- and under-estimation scenarios

	Innovation coefficient (p_g)	Imitation coefficient (q_g)
Over-estimation: prior	0.04	0.80
Observed values ($\hat{p}_g$)	[0.03, 0.025, 0.022, 0.02, 0.015, 0.01]	[0.5, 0.55, 0.6, 0.6, 0.65, 0.65]
Updated beliefs ($p_{g,h}$)	[0.031, 0.028, 0.026, 0.025, 0.023, 0.021]	[0.53, 0.54, 0.56, 0.57, 0.58, 0.60]
Under-estimation: prior	0.02	0.40
Observed values ($\hat{p}_g$)	[0.04, 0.035, 0.035, 0.032, 0.032, 0.03]	[0.7, 0.7, 0.75, 0.8, 0.8, 0.8]
Updated beliefs ($p_{g,h}$)	[0.038, 0.037, 0.036, 0.035, 0.035, 0.034]	[0.67, 0.68, 0.71, 0.73, 0.74, 0.75]

Results

Table 3.10 reports the service design decisions and profit estimates under the belief-updating strategy. The first column indicates the simulation scenario, either overestimation or underestimation of the adoption coefficients. The second column, labeled "It.", denotes the iteration number of the updating process. The remaining columns follow the same format as in the previous tables. In all cases, all grids are selected; therefore, grid selection details are omitted to avoid repetition.

The results show that in the overestimation scenario, both profit and fleet size are initially overestimated. As belief-updating progresses, profit estimates decline, and the time to profitability increases due to fleet oversupply and demand overestimation, reflecting the model's correction of optimistic adoption assumptions. In contrast, the underestimation scenario shows increasing profits and earlier profitability.

Figure 3.6 illustrates fleet size evolution across updating iterations, showing the relative deviation between each iteration's fleet size and the final value (at iteration $h = 6$). For overestimation, the figure reports relative oversupply, while for underestimation, it shows relative

Table 3.10: Service design using PSAM combined with belief-updating approach over a three-year horizon with $\mathcal{H}_1 = 6$

Scenario	It.	Fleet size	Availability level	h^+	Profit ($)
Over-estimated	1	[214, 225, 240, 260, 286, 332, 332, 332, 332, 332, 332, 277]	[0.94, 0.94, 0.94, 0.94, 0.94, 0.94, 0.94, 0.94, 0.94, 0.94, 0.94, 0.92]	4	8,205,047
	2	[214, 221, 229, 240, 256, 329, 329, 329, 329, 329, 329, 276]	[0.94, 0.94, 0.94, 0.94, 0.94, 0.94, 0.94, 0.94, 0.94, 0.94, 0.94, 0.92]	5	6,280,897
	3	[214, 221, 227, 239, 250, 329, 329, 329, 329, 329, 329, 276]	[0.94, 0.94, 0.94, 0.94, 0.94, 0.94, 0.94, 0.94, 0.94, 0.94, 0.94, 0.92]	5	6,112,960
	4	[214, 221, 227, 237, 249, 329, 329, 329, 329, 329, 329, 276]	[0.94, 0.94, 0.94, 0.94, 0.94, 0.94, 0.94, 0.94, 0.94, 0.94, 0.94, 0.92]	5	6,074,267
	5	[214, 221, 227, 237, 248, 329, 329, 329, 329, 329, 329, 276]	[0.94, 0.94, 0.94, 0.94, 0.94, 0.94, 0.94, 0.94, 0.94, 0.94, 0.94, 0.92]	5	6,001,892
	6	[214, 221, 227, 237, 248, 282, 282, 276, 276, 276, 276, 276]	[0.94, 0.94, 0.94, 0.94, 0.94, 0.94, 0.94, 0.92, 0.92, 0.92, 0.92, 0.92]	6	5,922,785
Under-estimated	1	[157, 216, 222, 228, 236, 256, 256, 251, 251, 251, 251, 227]	[0.92, 0.92, 0.94, 0.94, 0.94, 0.94, 0.94, 0.92, 0.92, 0.92, 0.92, 0.9]	6	4,075,718
	2	[157, 225, 236, 250, 273, 332, 332, 332, 332, 332, 332, 277]	[0.9*, 0.94, 0.94, 0.94, 0.94, 0.94, 0.94, 0.94, 0.94, 0.94, 0.94, 0.92]	5	7,590,411
	3	[157, 225, 233, 249, 272, 332, 332, 332, 332, 332, 332, 277]	[0.9, 0.94, 0.94, 0.94, 0.94, 0.94, 0.94, 0.94, 0.94, 0.94, 0.94, 0.92]	5	7,560,650
	4	[157, 225, 233, 249, 273, 332, 332, 332, 332, 332, 332, 277]	[0.9, 0.94, 0.94, 0.94, 0.94, 0.94, 0.94, 0.94, 0.94, 0.94, 0.94, 0.92]	5	7,622,124
	5	[157, 225, 233, 249, 274, 332, 332, 332, 332, 332, 332, 277]	[0.9, 0.94, 0.94, 0.94, 0.94, 0.94, 0.94, 0.94, 0.94, 0.94, 0.94, 0.92]	5	7,659,147
	6	[157, 225, 233, 249, 274, 332, 332, 332, 332, 332, 332, 277]	[0.9, 0.94, 0.94, 0.94, 0.94, 0.94, 0.94, 0.94, 0.94, 0.94, 0.94, 0.92]	5	7,680,297

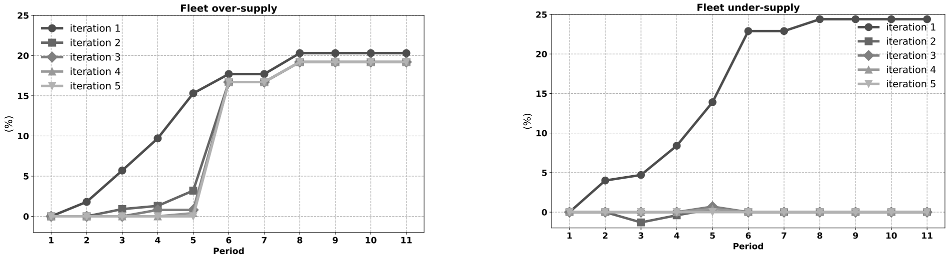
Note: In all cases, all grids are selected in the first period.
* Based on the updated values of p_g and q_g at the end of period 1, the actual availability provided by the selected fleet size in the first period is 0.90, not 0.92 as initially selected.

undersupply. The decreasing trend in deviations indicates the effectiveness of iterative belief-updating in reducing decision errors.

The patterns of observed values, shown in Table 3.9, significantly influence fleet size deviations. In the underestimation scenario, initial beliefs differ from observed values, but subsequent observations are stable, enabling quick adjustment. A significant under-supply is seen in the first iteration, but deviations drop near zero by the second iteration.

In the overestimation scenario, observed values converge more gradually with larger differences between successive observations. Initial overestimation results in substantial oversupply, and while deviations decrease over time, they do not fully vanish. After period 5, fleet size deviations increase due to the restriction on new fleet acquisitions. The model sets a higher fleet size before period 6 to meet the expected adoption growth. As shown in Table 3.10, at iteration 5, the fleet size is set to 329 for periods 6–11, but after updates in iteration 6, it is reduced to 282 for periods 6-7 and 276 in period 8.

To examine the value of belief-updating to adjust decisions, we compare the performance of the belief-updating strategy with the upfront strategy that relies solely on initial parameter estimates, without any updates or decision adjustments. Table 3.11 presents the results under the upfront design with incorrect initial beliefs. The first column indicates the simulation scenario.



(a) Overestimation scenario: percentage of fleet over-supply

(b) Underestimation scenario: percentage of fleet under-supply

Figure 3.6: Fleet size deviation over updating iterations, expressed as a percentage relative to the final fleet size based on the updated parameters at the last iteration (iteration 6).

The last column reports the profit loss, calculated as the difference (in percentage) between the total profit achieved by the belief-updating strategy and that of the upfront design. The remaining columns are consistent with those described previously.

Table 3.11: Upfront service design for under and over-estimations over a one-year implementation phase

Adoption coef.	Grid No.	Fleet size	Availability level	h^+	Profit loss(%)
Over-estimated	13	332	0.94	6	39.3
Under-estimated	13	315	[0.94,0.94,0.94,0.94,0.92*]	5	47

* According to the updated values of p and q, the selected fleet size is not enough for providing the availability level of 0.94 and it reduces to 0.92 after the fourth period

Ignoring the belief-updating opportunity to adjust decisions based on the observations results in significant profit losses. In the overestimation scenario, the upfront strategy leads to resource oversupply, resulting in a 39.3% reduction in total profit compared to the belief-updating strategy. In the underestimation scenario, the upfront design fails to meet service quality requirements and leads to a 47% profit loss.

3.7 Discussion

We synthesize the results of our numerical experiments, using a free-floating bike-sharing service design as a representative application, into the following managerial insights.

Service design choices (such as fleet size and service coverage) directly determine service quality, which in turn sets the upper bound of potential user uptake by defining who finds the service satisfactory. Within this bound, the pace and pattern of user adoption drive demand growth and profitability. Accurately capturing this interplay between service design and adoption dynamics is essential for realistic demand estimation and effective long-term planning.

Initial service quality is crucial for attracting and retaining innovators, thereby encouraging adoption among imitators. Even under cost pressures (such as low demand or highly imbalanced trip patterns) the optimal strategy maintains high availability and accessibility for the targeted coverage area. This often translates into offering selective, high-quality service rather than broad, lower-quality coverage. Doing so preserves user satisfaction, reduces churn, and lays the foundation for future demand growth.

Allowing for phased implementation aligned with adoption trajectories enables more efficient scaling of operations. By explicitly incorporating social influence and service quality in demand modeling, the system better matches supply with demand, avoiding the inefficiencies of over-supply or unmet demand.

While modeling adoption dynamics alone offers performance benefits, incorporating a belief-updating mechanism substantially enhances adaptability under uncertainty. By updating beliefs and adjusting future decisions based on observed adoption data, the belief-updating strategy protects against the long-term costs of misestimation. In contrast, a static upfront strategy, locked into initial assumptions, risks significant inefficiencies.

Faster convergence of observations allows earlier and more accurate updates, improving the alignment between supply decisions and actual demand. This highlights the importance of collecting reliable early-stage data and ensuring mechanisms for timely updates in system planning.

Overall, our findings highlight the importance of jointly modeling adoption behavior and service quality within a phased service deployment. This approach enables the gradual implementation of decisions aligned with the evolving adoption. It also offers the flexibility to revise design choices, such as fleet size and service coverage, as more is learned about actual user behavior. By phasing implementation, the model inherently supports a belief-updating process, allowing the system to update demand estimates based on observed adoption patterns and improve future decisions. This adaptability leads to more efficient resource allocation, sustained service quality, and greater profitability.

3.8 Conclusion

This paper proposes a phased service-adoption planning framework for shared micromobility operations that explicitly links user adoption dynamics to service design. By extending the Bass diffusion model to incorporate service quality and embedding it within a phased deployment structure, the framework synchronizes fleet sizing and service location decisions with projected adoption trends. By incorporating the interdependence between service design and adoption behavior, the proposed framework enables management of demand uptake through periodic adjustments in service availability and accessibility. The results of our numerical tests show that accounting for adoption behavior improves resource utilization and profitability compared with assuming immediate, socially independent, or service-quality-independent uptake. By scaling service quality over time, the model avoids the pitfalls of overinvestment in low-demand or slow-adoption settings and under-provision in high-demand or high-expectation cases, which can occur when the interdependence between demand, social influence, and service quality is ignored. While the model itself does not include endogenous learning, our numerical experiments illustrate that its phased structure supports decision adjustments when paired with external belief-updating processes. In particular, belief-updating simulations reveal how periodic updates to adoption forecasts can significantly reduce losses under parameter uncertainty by guiding mid-course adjustments in service design.

The proposed model can be extended to various emerging shared mobility services, beyond micromobility, as the core components of service quality are common across different modes. However, it needs to be adapted to incorporate resource-specific considerations, such as charging infrastructure for electric vehicles and repositioning operations for different vehicle types. This study opens several avenues for future work. First, embedding learning mechanisms, such as Bayesian inference or reinforcement learning, directly into the optimization process would enable real-time adaptation to observed user behavior. Second, incorporating uncertainty through stochastic or robust optimization could improve decision robustness under incomplete data or uncertainty around adoption. Third, extending the model to capture market competition or to integrate promotion strategies, such as marketing efforts and targeted incentives, would enhance its applicability to early-stage deployment planning.

Bibliography

Albiński, S., P. Fontaine, S. Minner (2018) Performance analysis of a hybrid bike sharing system: A service-level-based approach under censored demand observations, *Transportation Research Part E: Logistics and Transportation Review*, 116, pp. 59–69.

Ataç, S., N. Obrenović, M. Bierlaire (2021) Vehicle sharing systems: A review and a holistic management framework, *EURO Journal on Transportation and Logistics*, 10, p. 100033.

Bachand-Marleau, J., B. H. Lee, A. M. El-Geneidy (2012) Better understanding of factors influencing likelihood of using shared bicycle systems and frequency of use, *Transportation Research Record*, 2314(1), pp. 66–71.

Bass, F. M. (1969) A new product growth for model consumer durables, *Management Science*, 15(5), pp. 215–227.

Boer, M., O. Türetken, O. E. Adali (2022) A review of business models for shared mobility and mobility-as-a-service (maas): a research report.

Brdulak, A., G. Chaberek, J. Jagodziński (2021) Bass model analysis in “crossing the chasm” in e-cars innovation diffusion scenarios, *Energies*, 14(11), p. 3216.

Caggiani, L., R. Camporeale, M. Ottomanelli (2017) Planning and design of equitable free-floating bike-sharing systems implementing a road pricing strategy, *Journal of Advanced Transportation*, 2017(1), p. 3182387.

Caggiani, L., R. Camporeale, M. Ottomanelli, W. Y. Szeto (2018) A modeling framework for the dynamic management of free-floating bike-sharing systems, *Transportation Research Part C: Emerging Technologies*, 87, pp. 159–182.

Çelebi, D., A. Yörüsün, H. Işık (2018) Bicycle sharing system design with capacity allocations, *Transportation Research Part B: Methodological*, 114, pp. 86–98.

Cheng, Y., J. Wang, Y. Wang (2021) A user-based bike rebalancing strategy for free-floating bike sharing systems: A bidding model, *Transportation Research Part E: Logistics and Transportation Review*, 154, p. 102438.

Chiariotti, F., C. Pielli, A. Zanella, M. Zorzi (2020) A bike-sharing optimization framework combining dynamic rebalancing and user incentives, *ACM Transactions on Autonomous and Adaptive Systems (TAAS)*, 14(3), pp. 1–30.

Cialdini, R. B., N. J. Goldstein (2004) Social influence: Compliance and conformity, *Annual Review of Psychology*, 55(1), pp. 591–621.

Daganzo, C. F., K. R. Smilowitz (2004) Bounds and approximations for the transportation problem of linear programming and other scalable network problems, *Transportation Science*, 38(3), pp. 343–356.

Dan, J., Y. Dan, Y. Liu, L. Yalan (2018) Research on product diffusion of shared bicycle based on bass model, *Innovation and Manaegement*, p. 1520.

Deutsch, M., H. B. Gerard (1955) A study of normative and informational social influences upon individual judgment., *The Journal of Abnormal and Social Psychology*, 51(3), p. 629.

El Zarwi, F., A. Vij, J. L. Walker (2017) A discrete choice framework for modeling and forecasting the adoption and diffusion of new transportation services, *Transportation Research Part C: Emerging Technologies*, 79, pp. 207–223.

Fishman, E., S. Washington, N. Haworth (2012) Barriers and facilitators to public bicycle scheme use: A qualitative approach, *Transportation Research Part F: Traffic Psychology and Behaviour*, 15(6), pp. 686–698.

Fuady, S. N., P. C. Pfaffenbichler, Y. O. Susilo (2024) Bridging the gap: Toward a holistic understanding of shared micromobility fleet development dynamics, *Communications in Transportation Research*, 4, p. 100149.

Ghasri, M., A. Vij (2021) The potential impact of media commentary and social influence on consumer preferences for driverless cars, *Transportation Research Part C: Emerging Technologies*, 127, p. 103132.

Giordano, M., J. Y. Chow (2024) An e-scooter service region and fleet allocation design problem with elastic demand, *Transportation Research Part D: Transport and Environment*, 130, p. 104153.

Hsu, C.-L., H.-P. Lu (2004) Why do people play on-line games? an extended tam with social influences and flow experience, *Information & Management*, 41(7), pp. 853–868.

Jensen, A. F., E. Cherchi, S. L. Mabit, J. d. D. Ortúzar (2017) Predicting the potential market for electric vehicles, *Transportation Science*, 51(2), pp. 427–440.

Jiang, Z., C. Lei, Y. Ouyang (2020) Optimal investment and management of shared bikes in a competitive market, *Transportation Research Part B: Methodological*, 135, pp. 143–155.

Jin, H., S. Liu, K. C. So, K. Wang (2022) Dynamic incentive schemes for managing dockless bike-sharing systems, *Transportation Research Part C: Emerging Technologies*, 136, p. 103527.

Jin, Z., Y. Wang, Y. F. Lim, K. Pan, Z.-J. M. Shen (2023) Vehicle rebalancing in a shared micromobility system with rider crowdsourcing, *Manufacturing & Service Operations Management*, 25(4), pp. 1394–1415.

Kabra, A., E. Belavina, K. Girotra (2020) Bike-share systems: Accessibility and availability, *Management Science*, 66(9), pp. 3803–3824.

Kim, S., G. Lee, S. Choo (2023) Optimal rebalancing strategy for shared e-scooter using genetic algorithm, *Journal of Advanced Transportation*, 2023(1), p. 2696651.

Lagadic, M., A. Verloes, N. Louvet (2019) Can carsharing services be profitable? a critical review of established and developing business models, *Transport Policy*, 77, pp. 68–78.

Laporte, G., F. Meunier, R. Wolfler Calvo (2018) Shared mobility systems: an updated survey, *Annals of Operations Research*, 271(1), pp. 105–126.

Liu, Z., X. Jia, W. Cheng (2012) Solving the last mile problem: Ensure the success of public bicycle system in beijing, *Procedia-Social and Behavioral Sciences*, 43, pp. 73–78.

Luo, X., M. J. Kuby, Y. Honma, M. Kchaou-Boujelben, X. S. Zhou (2024) Innovation diffusion in ev charging location decisions: Integrating demand & supply through market dynamics, *Transportation Research Part C: Emerging Technologies*, 165, p. 104733.

Ma, L., X. Zhang, F. Hao, S. Liu (2023) Service quality, perceived price fairness, and users' continuous usage intentions regarding shared bike service, *Total Quality Management & Business Excellence*, pp. 1–21.

Mahajan, V., E. Muller, F. M. Bass (1990) New product diffusion models in marketing: A review and directions for research, *Journal of Marketing*, 54(1), pp. 1–26.

Mahajan, V., E. Muller, F. M. Bass (1995) Diffusion of new products: Empirical generalizations and managerial uses, *Marketing Science*, 14(3_supplement), pp. G79–G88.

Mahmoodian, V., Y. Zhang, H. Charkhgard (2022) Hybrid rebalancing with dynamic hubbing for free-floating bike sharing systems, *International Journal of Transportation Science and Technology*, 11(3), pp. 636–652.

Manca, F., A. Sivakumar, J. W. Polak (2019) The effect of social influence and social interactions on the adoption of a new technology: The use of bike sharing in a student population, *Transportation Research Part C: Emerging Technologies*, 105, pp. 611–625.

Massiani, J., A. Gohs (2015) The bass diffusion model: A review of its applications in forecasting new product adoption, *Technological Forecasting and Social Change*, 99, pp. 77–93.

Miranda, M., M. Pereda, A. Sánchez, E. Estrada (2024) Indirect social influence and diffusion of innovations: An experimental approach, *PNAS Nexus*, 3(10), p. pgae409.

Oliver, R. L. (1980) A cognitive model of the antecedents and consequences of satisfaction decisions, *Journal of Marketing Research*, 17(4), pp. 460–469.

Pal, A., Y. Zhang (2017) Free-floating bike sharing: Solving real-life large-scale static rebalancing problems, *Transportation Research Part C: Emerging Technologies*, 80, pp. 92–116.

Pan, M. (2022) *Behavior Science Methods for Active Transportation Demand Management*, Ph.D. thesis, The University of Arizona.

Pandita, S., H. G. Mishra, A. A. Bhat (2025) Consumer adoption behaviour in ride-sharing economy: an integrated tam-ecm framework, *International Journal of Emerging Markets*, 20(6), pp. 2491–2523.

Paul, A., K. Flynn, C. Overney (2023) Estimating censored spatial-temporal demand with applications to shared micromobility, *arXiv preprint arXiv:2303.09971*.

Powell, W. B. (2022) Sequential decision analytics and modeling: Modeling with python-part ii, *Foundations and Trends® in Technology, Information and Operations Management*, 16(1-2), pp. 1–176.

Putsis Jr, W. P. (1996) Temporal aggregation in diffusion models of first-time purchase: Does choice of frequency matter?, *Technological Forecasting and Social Change*, 51(3), pp. 265–279.

Reck, D. J., H. Haitao, S. Guidon, K. W. Axhausen (2021) Explaining shared micromobility usage, competition and mode choice by modelling empirical data from zurich, switzerland, *Transportation Research Part C: Emerging Technologies*, 124, p. 102947.

Reiss, S., K. Bogenberger (2015) Gps-data analysis of munich's free-floating bike sharing system and application of an operator-based relocation strategy, in: *2015 IEEE 18th international conference on intelligent transportation systems*, IEEE, pp. 584–589.

Ricci, M. (2015) Bike sharing: A review of evidence on impacts and processes of implementation and operation, *Research in Transportation Business & Management*, 15, pp. 28–38.

Rogers, E. M., A. Singhal, M. M. Quinlan (2014) Diffusion of innovations, in: *An integrated approach to communication theory and research*, Routledge, pp. 432–448.

Rotaris, L., M. Scorrano (2023) Insights into peer-to-peer carsharing: Modelling and scenario analysis via a bass diffusion agent-based model, *Research in Transportation Economics*, 97, p. 101261.

Satoh, D. (2001) A discrete bass model and its parameter estimation, *Journal of the Operations Research Society of Japan*, 44(1), pp. 1–18.

Soppert, M., C. Steinhardt, C. Müller, J. Gönsch, P. M. Bhogale (2023) Matching functions for free-floating shared mobility system optimization to capture maximum walking distances, *European Journal of Operational Research*, 305(3), pp. 1194–1214.

Stathopoulos, A., C. Cirillo, E. Cherchi, E. Ben-Elia, Y. T. Li, J. D. Schmöcker (2017) Innovation adoption modeling in transportation: New models and data, *Journal of Choice Modelling*, 25, pp. 61–68.

Vogel, P., P. Vogel (2016) *Service network design of bike sharing systems*, Springer.

Wang, X., X. Lin, Z. Liu (2021) Understanding consumers' post-adoption behavior in sharing economy services, *Journal of Computer Information Systems*.

Wang, X., H. Sun, S. Zhang, Y. Lv (2023) Bike-sharing rebalancing problem by considering availability and accessibility, *Transportmetrica A: Transport Science*, 19(3), p. 2067262.

Wang, Y., M. A. Douglas, B. T. Hazen, M. Dresner (2018) Be green and clearly be seen: How consumer values and attitudes affect adoption of bicycle sharing, *Transportation Research Part F: Traffic Psychology and Behaviour*, 58, pp. 730–742.

Whitt, W. (1984) Open and closed models for networks of queues, *AT&T Bell Laboratories Technical Journal*, 63(9), pp. 1911–1979.

Zhang, C., J.-D. Schmöcker, M. Kuwahara, T. Nakamura, N. Uno (2020) A diffusion model for estimating adoption patterns of a one-way carsharing system in its initial years, *Transportation Research Part A: Policy and Practice*, 136, pp. 135–150.

Zhang, J., M. Meng, D. Wang, et al. (2019) A dynamic pricing scheme with negative prices in dockless bike sharing systems, *Transportation Research Part B: Methodological*, 127(C), pp. 201–224.

Zhang, Y., M. Kamargianni (2023) A review on the factors influencing the adoption of new mobility technologies and services: autonomous vehicle, drone, micromobility and mobility as a service, *Transport Reviews*, 43(3), pp. 407–429.

Zhao, S., J. Xiong, S.-Y. Kim, L. Xu, J. Yan (2021) When bike-sharing crashed in china: a bumpy ride, *Journal of Business Strategy*, 42(5), pp. 295–303.

Zheng, H., K. Zhang, Y. Nie, P. Yan, Y. Qu (2024) How many are too many? analyzing dockless bike-sharing systems with a parsimonious model, *Transportation Science*, 58(1), pp. 152–175.

Chapter 4

Assortment Design under Context Effects: A Regret-Based Optimization Framework for Transportation Platforms

Abstract: This paper develops a methodological framework for context-aware assortment design in platform-based transportation systems. User preferences are modeled using the Generalized Random Regret Minimization (G-RRM) model, which accounts for context-dependent choice behavior such as attraction and compromise effects. We formulate the assortment optimization problem under G-RRM, derive structural properties of optimal solutions, and introduce diagnostic indicators that measure the influence of individual alternatives on these effects. To address computational complexity, we propose an efficient algorithm that embeds behavioral sensitivity directly into the optimization process. Numerical experiments demonstrate when context effects meaningfully shift optimal assortments, and when utility-maximization models yield similar outcomes despite the presence of context-driven distortions. These findings offer practical guidance for integrating behavioral modeling into platform design and operational decision-making.

Keyword: Context effects, Assortment optimization, Context-RUM models, Platform-based mobility systems

---

Research from this chapter contributed to a paper submitted to the *Transportation Research Part B: Methodological*.

4.1 Introduction

Urban mobility systems—like ride-hailing, micromobility, and on-demand transit—are increasingly shaped by platforms that algorithmically manage user-facing choices. A central mechanism in this design is the construction of travel menus, or assortments: the sets of alternatives offered to users at the point of decision. These menus shape individual choices and aggregate demand, with downstream effects on vehicle utilization, congestion, and service reliability. As such, assortment design plays a pivotal role in aligning platform goals—profitability, operational efficiency, and user satisfaction (Yang et al., 2016; Atasoy et al., 2015; Ulmer, 2020; Azadeh et al., 2022; Ausseil et al., 2022).

Assortment optimization is commonly formulated using discrete choice models grounded in Random Utility Maximization (RUM), such as the Multinomial Logit (MNL) and Nested Logit (NL) models (Ben-Akiva & Lerman, 1985). These frameworks assume that users evaluate alternatives independently based on latent utility. However, behavioral research has shown that preferences are context-dependent—shaped by the presence and structure of competing options (Simonson, 1989; Huber & Puto, 1983). Two canonical examples are the *attraction effect*, where a dominated decoy boosts the appeal of a similar alternative, and the *compromise effect*, where users disproportionately favor intermediate options. Such behaviors violate core RUM assumptions, including regularity and the independence of irrelevant alternatives (IIA) (Luce, 1977).

Contextual random utility models (Context-RUMs) have been developed to capture these dynamics. Yet most are designed for explanation or prediction in controlled, low-dimensional settings. Their relevance for real-world optimization—particularly in transportation platforms, which must generate assortments over large, dynamic alternative spaces—remains unclear. In these environments, multiple context effects may interact, and design decisions directly influence user behavior. Failing to account for such effects risks suboptimal system performance. Clarifying when and how context sensitivity alters optimal assortments is thus essential for aligning behavioral modeling with operational decision-making.

This paper develops a context-aware framework for assortment optimization in platform-based transportation systems. We focus on behavioral distortions—such as the attraction and compromise effects—that violate standard assumptions in Random Utility Maximization (RUM) models. While these effects are well-documented, their implications for design remain under-explored. We address this gap by integrating behavioral realism into the optimization process itself.

Our framework builds on the Generalized Random Regret Minimization (G-RRM) model, which captures context effects through pairwise, regret-based comparisons. We formulate the assortment optimization problem under this model, derive structural indicators that quantify the influence of individual alternatives, and analyze how assortment composition modulates user responses. These insights inform a scalable algorithm that embeds behavioral dynamics into assortment construction, balancing computational feasibility with behavioral fidelity.

Numerical experiments compare assortments produced by G-RRM and RUM-based models. The results show that context-aware designs significantly outperform standard methods in competitive or sensitivity-driven settings, where behavioral effects are pronounced. In contrast, both models converge when alternatives are similar or user preferences are stable. This yields a practical criterion: behavioral models matter most when design outcomes are sensitive to subtle shifts in user preferences.

This study makes four contributions. First, it formulates an assortment optimization framework grounded in regret-based modeling. Second, it develops diagnostic tools to expose how behavioral effects shape design. Third, it introduces a scalable algorithm that integrates these effects into operational logic. Finally, it identifies the conditions under which behavioral modeling changes optimal outcomes, clarifying where behavioral complexity adds value—and where it can be omitted without loss.

The remainder of the paper is organized as follows. Section 4.2 reviews related work on transportation assortment design and behavioral choice models. Section 4.3 formulates the regret-based assortment optimization problem. Section 4.4 formalizes attraction and compromise effects as geometric constructs and illustrates their behavior under G-RRM. Section 4.5 develops diagnostic measures for structural configurations, behavioral gains, competitor suppression, and centrality. Section 4.6 presents the solution method. Section 4.7 reports computational experiments, including comparisons of G-RRM and MNL. Section 4.8 concludes with implications and directions for future research.

4.2 Literature Review

Assortment optimization involves selecting a subset of alternatives to maximize expected outcomes, such as revenue or user welfare. This problem is fundamental to the design and operation of platform-based transport and logistics services. For instance, it underpins travel menu construction in on-demand transport systems, ride-hailing platforms, and the configuration of delivery timeframes in last-mile logistics services (e.g., Azadeh et al. (2022); Atasoy et al. (2015)). The accuracy of customer choice modeling, typically through behavioral models, is critical to the effectiveness of assortment strategies. These models influence customer decisions, thereby shaping platform profitability, service efficiency, and user satisfaction.

Classical assortment models are typically grounded in the Random Utility Maximization (RUM) framework, with Multinomial Logit (MNL) and Nested Logit (NL) models as dominant examples (Davis et al., 2014; Talluri & Van Ryzin, 2004). For a recent review of assortment optimization under RUM and a comparative empirical evaluation, see Heger & Klein (2024) and Berbeglia et al. (2022). However, a growing body of behavioral research highlights systematic violations of key RUM assumptions, such as regularity and the independence of irrelevant alternatives (IIA) (Luce, 1977). These violations emerge in phenomena such as the attraction effect—where a dominated decoy increases preference for a similar superior option—and the compromise effect, where intermediate alternatives are disproportionately favored (Simonson, 1989; Huber & Puto, 1983). To address such context-dependent choice behaviors, a range of Context-RUM models have been developed. Table 4.1 summarizes these models. In the remainder of this section, we review the main variants, assess their applicability to assortment design, and evaluate empirical evidence—particularly in transportation settings, as in Guevara & Fukushi (2016).

Emergent Value (EV) models capture context effects by using set-specific dummy variables to adjust utility based on the alternatives included in each choice set (Simonson, 1989; Rooderkerk et al., 2011). This allows the model to reflect how preferences shift depending on which options are presented together. In a comparative study of transportation mode choice, Guevara & Fukushi (2016) report that EV achieved the highest predictive accuracy for both attraction and compromise effects using stated and revealed preference data. However, because EV lacks a general utility function that applies across choice sets, it must re-estimate preferences for each new set. In algorithmic assortment optimization, where many sets are evaluated iteratively, this makes EV impractical.

Weight-Change (WC) and *Value Shift (VS)* models capture context effects by adjusting the importance of attributes based on the composition of the choice set (Huber et al., 1982; Wedell, 1991). In WC models, attribute weights are recalibrated when alternatives with extreme values are present, which can amplify preferences for compromise options. Empirically, WC has shown strong performance in predicting compromise effects, though VS models have received less consistent support Guevara & Fukushi (2016). However, both approaches often yield parameters that are difficult to interpret or generalize across sets which limits their applicability in assortment optimization.

The *Contextual Concavity Model (CCM)* draws on Prospect Theory to explain compromise behavior through reference dependence and loss aversion (Kivetz et al., 2004). The model assumes that individuals evaluate alternatives relative to a perceived reference point, with losses weighted more heavily than equivalent gains. In a comparative study using transportation mode choice data, Guevara & Fukushi (2016) reported that CCM was less effective in capturing decoy effects relative to other Context-RUM models. A key limitation is that reference points must be defined for each choice set, yet in mobility platforms, assortments vary across trips and users. This variation makes reference points unstable and difficult to define consistently, limiting CCM's applicability for assortment design in dynamic transportation environments.

The *Random Regret Minimization (RRM)* model conceptualizes choice behavior as the minimization of anticipated regret, computed through pairwise comparisons among alternatives (Loomes & Sugden, 1982). Unlike utility-maximization models, RRM emphasizes relative disadvantage rather than absolute utility. In a comparative analysis using transportation choice data, Guevara & Fukushi (2016) found that RRM effectively captured compromise effects and yielded predictive accuracy competitive with standard MNL. The *Generalized Random Regret Minimization (G-RRM)* model extends this framework by introducing a sensitivity parameter that modulates responsiveness to context. This additional flexibility enables G-RRM to represent both attraction and compromise effects within a unified structure. In the same study, the authors report that G-RRM achieved higher predictive accuracy than RRM and provided strong empirical fit for both context effects. Importantly, G-RRM retains closed-form choice probabilities and does not require recalibration for each assortment, making it especially suitable for assortment optimization.

Table 4.1: Empirical and Theoretical Comparison of Context-RUM Models

Model	Behavioral Mechanism	Attraction Effect Prediction	Compromise Effect Prediction	Limitations
EV	Set-specific utility shifts; justification heuristics	Strongest empirical support	High empirical accuracy	Requires dummy variables; lacks generalizability; needs recalibration for each set
WC	Attribute reweighting based on salience	Strong empirical fit	Best empirical fit	Parameters difficult to interpret; no closed-form solution; forecasting unstable
VS	Salience varies by dominance configuration	Strong theoretical basis	Not evaluated empirically	Identifiability issues; omitted from empirical validation
RRM	Regret from pairwise attribute comparisons	Moderate predictive power	Strong fit	Parsimonious but may oversimplify multidimensional trade-offs
G-RRM	Regret sensitivity scaling; generalizes MNL	Moderate theoretical support	Strong theoretical support	Not directly tested in Guevara & Fukushi (2016); moderate computational overhead; interpretation of parameters needed
CCM	Reference dependence with loss aversion	Weak support	Weak support	Empirical misalignment; unstable reference points; inconsistent performance

Behavioral models play a central role in platform operations—shaping not just demand forecasts but also the structure of design decisions. A model may fit user choices accurately yet yield similar assortments, or conversely, a simpler model may lead to effective designs. This study reframes behavioral modeling as a tool for guiding optimization. We integrate the Generalized Random Regret Minimization (G-RRM) model into an assortment framework to assess how context effects influence platform design. At the *assortment level*, we evaluate how context sensitivity alters the choice set itself. At the *within-assortment level*, we examine how behavioral asymmetries redirect user attention toward targeted options. This two-layer analysis provides operational clarity on when and where behavioral modeling matters.

4.3 Assortment Optimization Under Context-Dependent Preferences

Consider a universal set of alternatives $\mathcal{U} = \{1, \ldots, N\}$, where N denotes the total number of available items. An assortment is defined as a subset $\mathcal{S} \subseteq \mathcal{U}$ offered to consumers. Given an assortment $\mathcal{S}$, the probability that a consumer selects alternative $i \in \mathcal{S}$ is denoted by $\pi_i(\mathcal{S})$. These probabilities are assumed to be influenced by the *compromise effect* and the *attraction effect*.

Let $r_i \in \mathbb{R}$ denote the platform's profit from offering alternative $i \in \mathcal{U}$, and let $\boldsymbol{r} = (r_1, \ldots, r_N)$ denote the profit vector. The expected profit from offering assortment $\mathcal{S} \subseteq \mathcal{U}$ is $f(\mathcal{S}) = \sum_{i \in \mathcal{S}} r_i \pi_i(\mathcal{S})$, and the platform's objective is to identify the profit-maximizing assortment:

$$\mathcal{S}^* = \arg\max_{\mathcal{S} \subseteq \mathcal{U}} f(\mathcal{S}).$$

We model choice behavior using the Generalized Random Regret Minimization (G-RRM) framework (Chorus, 2014), which posits that individuals minimize anticipated regret from unfavorable attribute comparisons. Each alternative $i \in \mathcal{U}$ is described by an attribute vector $\boldsymbol{x}_i = (x_{i1}, \ldots, x_{iM}) \in \mathbb{R}^M$.

Given a choice set $\mathcal{S} \subseteq \mathcal{U}$, the total regret of selecting $i \in \mathcal{S}$ is:

$$RR_i(\mathcal{S}) = R_i(\mathcal{S}) + \varepsilon_i,$$

where $\varepsilon_i \sim \text{Gumbel}(0,1)$ is an i.i.d. error term, and $R_i(\mathcal{S})$ captures deterministic regret from pairwise comparisons:

$$R_i(\mathcal{S}) = \sum_{j \in \mathcal{S} \setminus \{i\}} \sum_{m=1}^{M} \left[\ln\left(\gamma + e^{\beta_m (x_{jm} - x_{im})}\right) - \ln(\gamma + 1) \right],$$

with $\beta_m \in \mathbb{R}$ denoting sensitivity to attribute m, and $\gamma > 0$ controlling smoothness and curvature of regret response.

The resulting choice probability is:

$$\pi_i(\mathcal{S}) = \frac{e^{(-R_i(\mathcal{S}))}}{\sum_{j \in \mathcal{S}} e^{(-R_j(\mathcal{S}))}} \tag{4.1}$$

4.3.1 Behavioral Properties of G-RRM

The parameter $\gamma \in [0,1]$ in the Generalized Random Regret Minimization (G-RRM) model modulates the curvature of the regret function and determines the extent of context sensitivity. As γ increases, the model places greater emphasis on pairwise regret comparisons, amplifying context effects in choice behavior.

When $\gamma = 0$, G-RRM becomes behaviorally neutral to the choice context. Under this setting, and assuming linear utility, the regret function simplifies to a linear form. The following proposition establishes that, in this regime, G-RRM yields choice probabilities equivalent to those produced by the Multinomial Logit (MNL) model.

Proposition 1 *Let $\pi_i(\mathcal{S})$ denote the G-RRM choice probability for alternative $i \in \mathcal{S}$, as defined in Equation* (4.1)*. Assume that utility is linear in attributes. Then, in the limit $\gamma \to 0$, G-RRM converges to a scaled version of the Multinomial Logit (MNL) model with transformed weights $\omega_m = |\mathcal{S}| \cdot \beta_m$, such that:*

$$\lim_{\gamma \to 0} \pi_i(\mathcal{S}) = \frac{e^{\left(\sum_{m=1}^{M} \omega_m x_{im}\right)}}{\sum_{j \in \mathcal{S}} e^{\left(\sum_{m=1}^{M} \omega_m x_{jm}\right)}}.$$

A complete derivation is provided in A.1.

When $\gamma > 0$, G-RRM incorporates context-sensitive regret comparisons that induce behavioral asymmetries not captured by MNL. In this regime, the model no longer satisfies the properties of regularity and independence of irrelevant alternatives (IIA), which are central to multinomial logit model. The following propositions formally demonstrate that G-RRM violates these properties through its structural dependence on the composition of the choice set.

Proposition 2 (Violation of Regularity under G-RRM) *Let $\mathcal{S} \subset \mathcal{U}$ be a choice set, and let $\mathcal{T} = \mathcal{S} \cup \{k\}$ for some $k \notin \mathcal{S}$. Under the G-RRM model, there exist attribute vectors $\{x_i\}_{i \in \mathcal{T}}$ and regret parameters β, γ such that for some $i \in \mathcal{S}$,*

$$\pi_i(\mathcal{T}) > \pi_i(\mathcal{S}).$$

Hence, G-RRM violates the regularity property.

The proof is provided in A.2.

Proposition 3 (Violation of IIA under G-RRM) *Let $\mathcal{S} \subset \mathcal{U}$ be a choice set, and let $\mathcal{T} = \mathcal{S} \cup \{k\}$ for some $k \notin \mathcal{S}$. Under the G-RRM model, there exist attribute vectors $\{x_\ell\}_{\ell \in \mathcal{T}}$ and parameters β, γ such that for some pair $i, j \in \mathcal{S}$,*

$$\frac{\pi_i(\mathcal{S})}{\pi_j(\mathcal{S})} \neq \frac{\pi_i(\mathcal{T})}{\pi_j(\mathcal{T})}.$$

Therefore, the model does not satisfy the independence of irrelevant alternatives (IIA) property.

The proof is provided in A.3.

In summary, G-RRM recovers MNL behavior as $\gamma \to 0$, satisfying regularity and IIA in that limit. For $\gamma > 0$, however, context-sensitive comparisons introduce systematic violations of these axioms. Propositions 2 and 3 show that choice probabilities may increase for dominated alternatives or shift disproportionately with added options—patterns inconsistent with classical RUM logic. These departures motivate the need for diagnostic tools to quantify context effects in applied assortment design.

4.4 Context Effects in Assortment Design

To characterize when and how context effects distort choice behavior, this section formalizes attraction and compromise effects using geometric constructs over attribute vectors that remain invariant under monotonic transformations of utility. Each alternative $i \in \mathcal{U}$ is represented by an attribute vector $\boldsymbol{x}_i \in \mathbb{R}^M$, where all attributes are (pre)oriented so that higher values are weakly preferred.

We first define the structural conditions under which attraction effects arise.

Definition 4.1 Asymmetric Domination
Let $\boldsymbol{x}_i, \boldsymbol{x}_d \in \mathbb{R}^M$ be attribute vectors for alternatives i and d. Alternative d is *asymmetrically dominated* by i if: $x_{dm} \leq x_{im}$ $\forall m$, and there exists m such that $x_{dm} < x_{im}$.

Definition 4.2 Attraction Zone
Let $a_1, c \in \mathcal{U}$ be alternatives with attribute vectors $\boldsymbol{x}_{a_1}, \boldsymbol{x}_c \in \mathbb{R}^M$. The *attraction zone* $\mathcal{A}_{a_1,c} \subset$

$\mathbb{R}^M$ is the set of vectors $\boldsymbol{x}_d$ such that: (1) $\boldsymbol{x}_d$ is asymmetrically dominated by $\boldsymbol{x}_{a_1}$; (2) $\boldsymbol{x}_d$ is not asymmetrically dominated by $\boldsymbol{x}_c$; and (3) $\boldsymbol{x}_c$ is not asymmetrically dominated by $\boldsymbol{x}_{a_1}$. This zone characterizes decoy positions that amplify preference for a_1 without inducing dominance over c.

Definition 4.3 Attraction Effect
Let $a_1, c \in \mathcal{S} \subseteq \mathcal{U}$ be two alternatives, and let $d \in \mathcal{S}$ be a third alternative such that $\boldsymbol{x}_d \in \mathcal{A}_{a_1,c}$. An *attraction effect* is present if the inclusion of d increases the absolute or relative choice probability of a_1:

$$\frac{\pi_{a_1}(\{a_1, c, d\})}{\pi_c(\{a_1, c, d\})} > \frac{\pi_{a_1}(\{a_1, c\})}{\pi_c(\{a_1, c\})} \quad \text{or} \quad \pi_{a_1}(\{a_1, c, d\}) > \pi_{a_1}(\{a_1, c\}).$$

The attraction effect violates regularity and IIA by increasing demand for a_1 when a dominated alternative from $\mathcal{A}_{a_1,c}$ is added to the choice set. We next formalize the compromise effect, which emerges from the inclusion of extreme options rather than dominance relations.

Definition 4.4 Compromise Region
Let $i, j \in \mathcal{S} \subseteq \mathcal{U}$ be mutually non-dominating alternatives with attribute vectors $\boldsymbol{x}_i, \boldsymbol{x}_j \in \mathbb{R}^M$. The *compromise region* between i and j is

$$\mathcal{C}_{ij} = \left\{ \boldsymbol{z} \in \mathbb{R}^M \;\middle|\; \min\{x_{im}, x_{jm}\} \le z_m \le \max\{x_{im}, x_{jm}\} \quad \forall m \right\},$$

where z_m denotes the m-th coordinate of $\boldsymbol{z}$.

Definition 4.5 Compromise Effect
Let $i, j \in \mathcal{S}$ be mutually non-dominating alternatives and let $\mathcal{C}_{ij}$ be their compromise region. An alternative $k \in \mathcal{S}$ ($k \neq i, j$) is a *compromise candidate* with respect to i and j if (i) $\boldsymbol{x}_k \in \mathcal{C}_{ij}$; (ii) k is not asymmetrically dominated by i or j; and (iii) $\|\boldsymbol{x}_k - \boldsymbol{x}_i\| > \varepsilon$ and $\|\boldsymbol{x}_k - \boldsymbol{x}_j\| > \varepsilon$ for some $\varepsilon > 0$.

A *compromise effect* is present if

$$\frac{\pi_k(\{i, j, k\})}{\pi_i(\{i, j, k\})} > \frac{\pi_k(\{i, k\})}{\pi_i(\{i, k\})} \quad \text{or} \quad \frac{\pi_k(\{i, j, k\})}{\pi_j(\{i, j, k\})} > \frac{\pi_k(\{j, k\})}{\pi_j(\{j, k\})},$$

and it additionally violates regularity when

$$\pi_k(\{i, j, k\}) > \pi_k(\{i, k\}) \quad \text{and} \quad \pi_k(\{i, j, k\}) > \pi_k(\{j, k\}).$$

These constructs provide a testable basis for detecting violations of regularity and IIA in observed choice behavior. We next use them to construct diagnostic metrics in Section 4.5.

4.4.1 Illustrative Behavior of G-RRM under Context Effects

We use diagrammatic examples to show how G-RRM produces context effects when the position of a decoy changes within the attribute space. The setting builds on two diagnostics defined earlier: the regularity ratio $\dot{\pi}_{a_1}$ and the IIA ratio $\ddot{\pi}_{a_1,c}$, which respectively capture violations of regularity and IIA. By mapping these ratios over alternative placements of the decoy, we can visualize how geometric configuration shapes the strength of these effects.

Figures 4.1a–4.1b consider a three-alternative set composed of focal a_1, competitor c, and decoy d. The outlined rectangle marks the attraction zone $\mathcal{A}_{a_1,c}$, where d is asymmetrically dominated by a_1 but not by c. Contour lines show the values of $\dot{\pi}_{a_1}$ and $\ddot{\pi}_{a_1,c}$ across the attribute space. Within $\mathcal{A}_{a_1,c}$, both ratios are greater than one, and their values increase as d moves deeper into the zone and farther from the equal-attribute frontier between a_1 and c. This pattern reflects the fact that greater dominance of a_1 over d amplifies the shift in choice probability toward a_1.

Figures 4.1c–4.1d add a second focal alternative a_2 while retaining the same decoy placements in $\mathcal{A}_{a_1,c}$. The presence of a_2 changes the pairwise regret terms, which now compare d not only to a_1 and c but also to a_2. As a result, placements that previously yielded large increases in $\dot{\pi}_{a_1}$ and $\ddot{\pi}_{a_1,c}$ now produce ratios close to one, with only mild increases in the IIA measure. The reduction occurs because part of the decoy's disadvantage relative to a_1 is offset by its disadvantage relative to a_2.

These examples show that the impact of a decoy depends jointly on its location in the attribute space and the composition of the choice set.

4.5 Diagnostic Framework

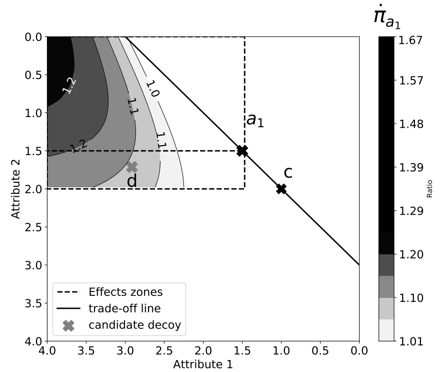
We define four categories of indicators to capture the presence, strength, and strategic consequences of context effects in assortment optimization. Each category plays a distinct role in diagnosing the influence of compromise and attraction dynamics on the structure and outcomes of choice sets.

4.5.1 Structural Preconditions for Context Effects

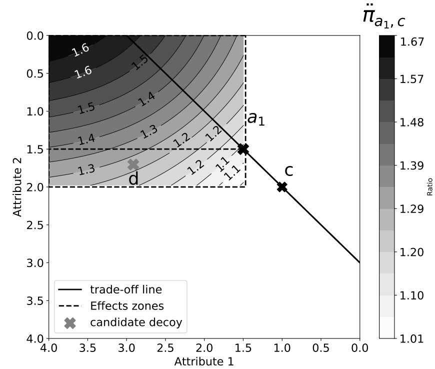
We begin by identifying structural conditions under which context effects can theoretically arise. These indicators assess the spatial configurations that make an assortment geometrically susceptible to compromise or attraction influences, independent of user preferences.

Compromise Configurations $\kappa^{\text{cfg}}(\mathcal{S})$ This indicator quantifies the number of structural configurations within $\mathcal{S}$ that permit a compromise effect. Each valid configuration consists of a pair of mutually non-dominating extremes $\{j,k\} \subseteq \mathcal{S}$ that define a compromise region $\mathcal{C}_{jk}$, containing at least one intermediate alternative $i \in \mathcal{S} \setminus \{j,k\}$ that is not asymmetrically dominated.

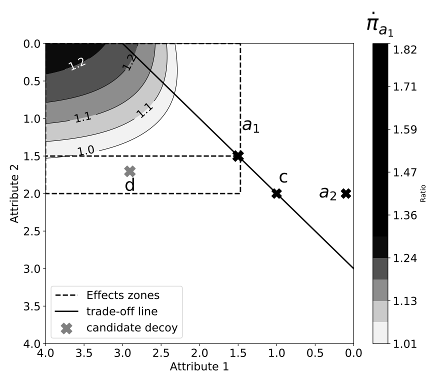
Attraction Configurations $\alpha^{\text{cfg}}(\mathcal{S})$ This indicator quantifies the number of structural configurations within $\mathcal{S}$ that support attraction effects. Each valid configuration consists of a decoy-target pair $(i,d) \in \mathcal{S} \times \mathcal{S}$, where the decoy d is asymmetrically dominated by the focal alternative i, and not by any other competitor $j \in \mathcal{S} \setminus \{i,d\}$.



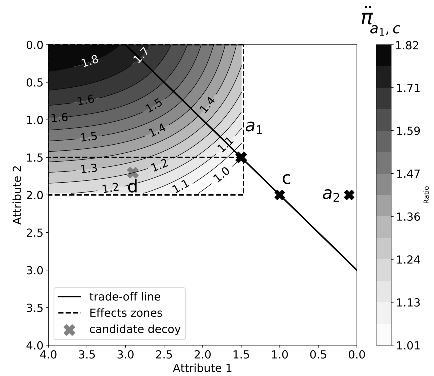
(a) Regularity ratio $\dot{\pi}_{a_1}$ for (a_1, c, d). Contours show values across the attribute space; the outlined region is the attraction zone $\mathcal{A}_{a_1,c}$, where d is dominated by a_1 but not by c. Ratios above one indicate absolute share gains for a_1.



(b) IIA ratio $\ddot{\pi}_{a_1,c}$ for (a_1, c, d). Contours show values across the attribute space; the outlined region is $\mathcal{A}_{a_1,c}$. Ratios above one indicate relative odds gains for a_1 against c.



(c) Regularity ratio $\dot{\pi}_{a_1}$ after adding a_2. Decoy positions match those in panels (a)–(b). The presence of a_2 changes regret terms, reducing the uplifts previously observed.



(d) IIA ratio $\ddot{\pi}_{a_1,c}$ after adding a_2. Decoy positions match those in panels (a)–(b). Most placements yield ratios near one, indicating little advantage for a_1 over c.

Figure 4.1: Impact of decoy placement on regularity and IIA under G-RRM. In the three-alternative set (a_1, c, d), uplifts in both diagnostics occur exclusively within the attraction zone $\mathcal{A}_{a_1,c}$ and intensify as d moves deeper into the zone and farther from the equal-attribute frontier. Adding a second focal alternative a_2 changes the structure of pairwise regret comparisons, as d is now evaluated relative to both a_1 and a_2 in addition to c. This additional competition neutralizes most of the uplifts, yielding ratios close to one in much of the attribute space. The contrast between panels (a)–(b) and (c)–(d) illustrates how decoy effects depend jointly on geometric dominance and the composition of the choice set.

4.5.2 Context-Induced Shifts in Choice Probability

This section evaluates how structurally admissible context effects translate into observable distortions in user choice behavior under the G-RRM model. These indicators measure the behavioral realization of compromise and attraction dynamics for individual alternatives and across the entire assortment.

Compromise Gain $\kappa_k^{(\mathcal{S})}$ Measures the behavioral uplift in choice probability for an alternative $k \in \mathcal{S}$ due to its role in structurally valid compromise configurations. For each pair of mutually non-dominating extremes $\{i, j\} \subseteq \mathcal{S} \setminus \{k\}$, if $k \in \mathcal{C}_{ij}$, we evaluate how the inclusion of both extremes increases k's probability relative to either extreme alone:

$$\kappa_k^{(\mathcal{S})} = \sum_{\substack{\{i,j\} \subseteq \mathcal{S} \setminus \{k\} \\ k \in \mathcal{C}_{ij}}} \left[\pi_k(\{i, j, k\}) - \pi_k(\{i, k\})\right]_+ + \left[\pi_k(\{i, j, k\}) - \pi_k(\{j, k\})\right]_+ .$$

A positive value signals that k's appeal is behaviorally amplified by its placement between extremes. Summing across all alternatives yields the total compromise gain: $\kappa^{\text{tot}}(\mathcal{S}) = \sum_{k \in \mathcal{S}} \kappa_k^{(\mathcal{S})}$.

Attraction Gain $\alpha_i^{(\mathcal{S})}$ Captures the behavioral gain in choice probability for a focal alternative $i \in \mathcal{S}$ driven by attraction-inducing decoys. For each competitor $j \in \mathcal{S} \setminus \{i\}$ and decoy $d \in \mathcal{A}_{i,j} \cap \mathcal{S}$, we evaluate how the presence of d elevates the appeal of i relative to the baseline pair $\{i, j\}$:

$$\alpha_i^{(\mathcal{S})} = \sum_{\substack{j \in \mathcal{S} \setminus \{i\} \\ d \in \mathcal{A}_{i,j} \cap \mathcal{S}}} \left[\pi_i(\{i, j, d\}) - \pi_i(\{i, j\})\right]_+ .$$

This indicator isolates decoy effects that reinforce i's position without also benefiting competitors. The total attraction gain across the assortment is: $\alpha^{\text{tot}}(\mathcal{S}) = \sum_{i \in \mathcal{S}} \alpha_i^{(\mathcal{S})}$.

4.5.3 Competitor Suppression

We evaluate how context effects diminish the relative standing of a focal competitor c. Two indicators quantify suppression via compromise and attraction dynamics.

Compromise-Based Suppression $\bar{\kappa}_c(\mathcal{S})$ Measures the extent to which intermediate alternatives $i \in \mathcal{C}_{cj}$ weaken c's position when introduced alongside another extreme j. For each valid triplet $\{i, c, j\}$, we compute:

$$\bar{\kappa}_c(\mathcal{S}) = \sum_{\substack{\{i,j\} \subseteq \mathcal{S} \setminus \{c\} \\ i \in \mathcal{C}_{cj}}} \left[\log \frac{\pi_i(\{i, c, j\})}{\pi_c(\{i, c, j\})} - \log \frac{\pi_i(\{i, c\})}{\pi_c(\{i, c\})}\right]_+ .$$

A positive value indicates a regularity violation, where the presence of a compromise draws preference away from c.

Attraction-Based Suppression $\bar{\alpha}_c(\mathcal{S})$ Assesses whether decoys asymmetrically dominated by a target i distort its comparison with c. For each $d \in \mathcal{A}_{i,c}$, we evaluate:

$$\bar{\alpha}_c(\mathcal{S}) = \sum_{\substack{i \in \mathcal{S} \setminus \{c\} \\ d \in \mathcal{A}_{i,c} \cap \mathcal{S}}} \left[\log \frac{\pi_c(\mathcal{S} \setminus \{d\})}{\pi_i(\mathcal{S} \setminus \{d\})} - \log \frac{\pi_c(\mathcal{S})}{\pi_i(\mathcal{S})} \right]_+ .$$

A positive value signals a violation of IIA, as the decoy alters relative odds without directly competing with c.

4.5.4 Indicators for Spatial and Behavioral Centrality

This section evaluates how assortments align with geometric and preference-based centers. We define centrality through two constructs: the geometric center, which reflects spatial symmetry, and the behavioral center, which captures the concentration of user preference. These indicators assess whether optimization shifts assortments toward central profiles, and whether such shifts favor high-probability or high-revenue alternatives.

Geometric Center We define spatial centrality to assess how assortment optimization repositions alternatives relative to the geometric center of the attribute space. The geometric center of a set $\mathcal{S}$ is given by the geometric median:

$$\bar{\boldsymbol{x}}^{\mathrm{geo}}(\mathcal{S}) = \arg\min_{\boldsymbol{z} \in \mathbb{R}^M} \sum_{i \in \mathcal{S}} \|\boldsymbol{x}_i - \boldsymbol{z}\|.$$

This point minimizes total distance to all alternatives and serves as a spatial anchor for evaluating symmetry.

To measure how optimization affects this anchor, we compare its distance to a fixed competitor $\boldsymbol{x}_c$ before and after optimization:

$$\Delta^{\mathrm{geo}} = \|\bar{\boldsymbol{x}}^{\mathrm{geo}}(\mathcal{S}^*) - \boldsymbol{x}_c\| - \|\bar{\boldsymbol{x}}^{\mathrm{geo}}(\mathcal{U}) - \boldsymbol{x}_c\|.$$

A negative value indicates that the optimized set $\mathcal{S}^*$ moves the center closer to c, suggesting greater alignment.

We also evaluate the centrality shift for each item $i \in \mathcal{S}^*$ relative to the geometric center:

$$\Delta_i^{\mathrm{geo}} = \|\boldsymbol{x}_i - \bar{\boldsymbol{x}}^{\mathrm{geo}}(\mathcal{U})\| - \|\boldsymbol{x}_i - \bar{\boldsymbol{x}}^{\mathrm{geo}}(\mathcal{S}^*)\|.$$

Positive values suggest that optimization draws item i closer to the spatial center, reinforcing its geometric centrality.

Behavioral Center To capture the focal point of user attention within a choice set, we define the behavioral center as the preference-weighted average of alternative attributes $\bar{\boldsymbol{x}}^{\mathrm{beh}}(\mathcal{S}) = \sum_{i \in \mathcal{S}} \pi_i(\mathcal{S}) \cdot \boldsymbol{x}_i$, where $\pi_i(\mathcal{S})$ is the choice probability of alternative i under the G-RRM model.

This point reflects where user preference is concentrated. To quantify how far each item deviates from this center, we define its behavioral distance:$d_i^{\text{beh}} = \|\boldsymbol{x}_i - \bar{\boldsymbol{x}}^{\text{beh}}(\mathcal{S})\|$. Lower values indicate stronger alignment with collective preference, highlighting candidates that act as behavioral anchors.

Alignment Indicator To evaluate whether preference concentrates near the geometric center, we compute the Spearman correlation:

$$\rho^{\text{geo}}(\mathcal{S}) = \operatorname{corr}\left(\{\|\boldsymbol{x}_i - \bar{\boldsymbol{x}}^{\text{geo}}(\mathcal{S})\|\}_{i\in\mathcal{S}},\ \{\pi_i(\mathcal{S})\}_{i\in\mathcal{S}}\right).$$

A negative value supports compromise behavior by indicating that central options attract higher probability.

4.6 Resolution Approach

This section develops solution methods for G-RRM under a common scale normalization applied uniformly for all γ. The main results are: (i) at $\gamma = 0$, the model reduces to MNL and the optimal assortment is revenue-ordered, computable in $O(N \log N)$ (Proposition 1, Theorem 4.1); and (ii) for any fixed $\gamma > 0$, the cardinality-k assortment problem is NP-hard (Theorem 4.2), hence the general problem is also NP-hard. We then introduce the Context-Aware Policy (CAP), a one-exchange local search that evaluates full G-RRM probabilities and coincides with the MNL optimum as $\gamma \to 0$. CAP further achieves near-optimal revenue in the low-regret regime (Corollary 4.3).

4.6.1 Behavior and Solvability at $\gamma = 0$

As $\gamma \to 0$, the G-RRM probabilities reduce to a logit form with scale proportional to the set size $|\mathcal{S}|$ (Proposition 1). Because this scale varies across sets, the odds between two fixed items can shift when new options are added, violating IIA and regularity.

To eliminate this artifact, we normalize by dividing each deterministic regret by the number of displayed rivals,

$$\pi_i(\mathcal{S};\gamma) \ \propto\ \exp(-R_i(\mathcal{S};\gamma)/T(\mathcal{S})), \qquad T(\mathcal{S}) := \max\{|\mathcal{S}| - 1, 1\}.$$

This convention fixes the comparison scale across sets (equivalently, it rescales the Gumbel errors) and yields at $\gamma = 0$ the standard Multinomial Logit model with utilities

$$V_i = \sum_{m=1}^{M} \beta_m x_{im}.$$

The normalized $\gamma = 0$ case therefore inherits all structural properties of MNL (IIA, regularity, monotonicity), enabling polynomial-time solvability and the revenue-ordered optimal assortment results that follow.

Proposition 4 (Structural properties at $\gamma = 0$) *In the normalized $\gamma = 0$ case, the G-RRM* *choice rule coincides with the Multinomial Logit model with utilities $\{V_i\}$. It therefore inherits the standard structural properties of MNL: independence of irrelevant alternatives (IIA), regularity, and monotonicity in utilities.*

Proof: From the softmax form $\pi_i(\mathcal{S}) = e^{V_i} / \sum_{j\in\mathcal{S}} e^{V_j}$: (i) the odds ratio π_i/π_k depends only on $V_i - V_k$, establishing IIA; (ii) enlarging the choice set increases the denominator while leaving the numerator unchanged, so π_i cannot increase, proving regularity; and (iii) differentiating with respect to V_i and V_k yields

$$\frac{\partial \pi_i}{\partial V_i} = \pi_i(1-\pi_i) > 0, \qquad \frac{\partial \pi_i}{\partial V_k} = -\,\pi_i\pi_k < 0,$$

so choice probabilities are strictly monotone in utilities. □

Proposition 4 establishes that the normalized $\gamma = 0$ regime behaves exactly as an MNL model, inheriting its structural axioms. This equivalence allows us to leverage classical assortment theory. In particular, expected revenue admits the standard logit form, which implies that the optimal assortment follows a simple threshold rule in item revenues. The next theorem formalizes this revenue-ordered structure.

Theorem 4.1 (Revenue-ordered optimal assortment at $\gamma = 0$) *At $\gamma = 0$, the optimal assortment is characterized by a revenue threshold: there exists $\tau^\star$ such that*

$$\mathcal{S}^\star \;=\; \{\, i \in \mathcal{U} : r_i \geq \tau^\star \,\}, \qquad \tau^\star = f(\mathcal{S}^\star).$$

Equivalently, ordering items by revenue yields a nested family of assortments $\{A_k\}_{k=0}^{|\mathcal{U}|}$, and $\mathcal{S}^\star = A_{K^\star}$ for some $K^\star$. The optimum can therefore be found by sorting revenues and scanning prefixes, an $O(N \log N)$ procedure.

A complete proof is given in Appendix A.4.

4.6.2 Stability Near $\gamma = 0$

A natural question is whether the revenue-ordered assortment that is optimal under MNL ($\gamma = 0$) continues to be optimal when regret effects are small but positive. The next proposition shows that this is indeed the case, provided the MNL optimum is unique and we adopt a consistent normalization across all γ.

Proposition 5 (Stability near $\gamma = 0$) *Let $\mathcal{S}^{\mathrm{NR}}$ denote the revenue-ordered assortment that is optimal at $\gamma = 0$. Suppose this assortment is unique and that the revenue gap*

$$\Delta \;=\; f(\mathcal{S}^{\mathrm{NR}}; 0) \;-\; \max_{\mathcal{S} \neq \mathcal{S}^{\mathrm{NR}}} f(\mathcal{S}; 0)$$

is strictly positive. Then there exists $\gamma_0 > 0$ such that $\mathcal{S}^{\mathrm{NR}}$ remains optimal for all $\gamma \in [0, \gamma_0]$.

Proof: For any fixed $\mathcal{S}$, the revenue $f(\mathcal{S};\gamma)$ is continuous in γ. Since the number of assortments is finite, the difference

$$g(\gamma) \;=\; \max_{\mathcal{S}\neq\mathcal{S}^{\mathrm{NR}}} f(\mathcal{S};\gamma) \;-\; f(\mathcal{S}^{\mathrm{NR}};\gamma)$$

is also continuous. At $\gamma = 0$, we have $g(0) = -\Delta < 0$. By continuity, there exists $\gamma_0 > 0$ such that $g(\gamma) < 0$ for all $\gamma \in [0, \gamma_0]$, proving the claim. □

Thus the revenue-ordered structure established at $\gamma = 0$ is robust under small positive regret. Beyond this regime, however, assortment optimization becomes computationally intractable.

4.6.3 Computational Limits Beyond $\gamma = 0$

At $\gamma = 0$, assortment optimization reduces to MNL and admits a simple revenue-ordered solution in polynomial time. This tractability is fragile: as soon as regret effects are introduced, the problem becomes intractable.

Theorem 4.2 (NP-hardness for $\gamma > 0$) *For any fixed $\gamma > 0$ and $\beta > 0$, the decision version of the cardinality-k assortment optimization problem under G-RRM is NP-hard.*

Thus, the revenue-ordered solvability of the $\gamma = 0$ regime is a knife-edge phenomenon: any positive regret sensitivity makes assortment optimization computationally intractable in general. This rules out efficient exact optimization and motivates the approximation strategies developed in the next section. The proof is by reduction from the *Densest k-Subgraph* problem and is deferred to A.5.

4.6.4 Context-Aware Policy (CAP)

At $\gamma = 0$, assortment optimization admits a revenue-ordered solution (Theorem 4.1). Once regret interactions are introduced ($\gamma > 0$), this structure disappears: the marginal effect of an item depends on the full incumbent set through nonlinear pairwise terms, and the problem becomes NP-hard (Theorem 4.2). To cope with this complexity, we develop the *Context-Aware Policy (CAP)*, a local-search method that evaluates exact G-RRM revenues and delivers behaviorally consistent, computationally efficient solutions.

We formalize the search procedure as follows. The outside option c is always available with $r_c = 0$. For an assortment $\mathcal{S} \subseteq \mathcal{U}$, define

$$V_i(\mathcal{S};\gamma) = \exp\{-R_i(\mathcal{S};\gamma)\}, \qquad \pi_i(\mathcal{S};\gamma) = \frac{V_i(\mathcal{S};\gamma)}{V_c(\mathcal{S};\gamma) + \sum_{j\in\mathcal{S}} V_j(\mathcal{S};\gamma)},$$

and expected revenue

$$Z_{\mathcal{S}}(\gamma) = \sum_{i\in\mathcal{S}} r_i\,\pi_i(\mathcal{S};\gamma).$$

Given $\mathcal{S}$, its 1-exchange neighborhood is

$$\mathcal{N}(\mathcal{S}) = \{\,\mathcal{S}\cup\{i\} : i \notin \mathcal{S}\,\} \cup \{\,\mathcal{S}\setminus\{i\} : i \in \mathcal{S}\,\},$$

optionally restricted to enforce $|\mathcal{S}| \leq k$ under a cardinality limit. The *Context-Aware Policy (CAP)* iteratively applies the best-improvement rule: from $\mathcal{S}$, it moves to the neighbor $\mathcal{S}' \in \mathcal{N}(\mathcal{S})$ with maximum $Z_{\mathcal{S}'}(\gamma)$ whenever $Z_{\mathcal{S}'}(\gamma) > Z_{\mathcal{S}}(\gamma)$; otherwise it halts. To reduce sensitivity to initialization, CAP is run from each singleton $\{i\}$ and returns the best final revenue. The detail of the algorithm is presented in A.6.

Connection to the Nested-by-Revenue Policy

At $\gamma = 0$, G-RRM reduces to MNL and the optimal assortment follows the Nested-by-Revenue (NR) rule (Theorem 4.1). CAP is consistent with this tractable regime: starting from any singleton, every improving move under MNL admits the highest-revenue remaining item until the fixed-point condition

$$\mathcal{S}^\star = \{\, i \in \mathcal{U} : r_i \geq Z(\mathcal{S}^\star)\,\}$$

is satisfied. Thus CAP collapses exactly to the NR solution when regret effects vanish.

More generally, if regret comparisons remain aligned with revenue rankings (e.g., higher revenue implies lower regret, rankings are invariant under expansions, and zero-profit items are dominated), then CAP selects items monotonically in revenue order. These conditions are restrictive and mainly serve as a *consistency check*: they show that CAP degenerates to the familiar greedy revenue-nesting policy whenever regret interactions behave monotonically.

Robustness of CAP in the Low-Regret Regime

At $\gamma = 0$, CAP coincides with the revenue-ordered optimum of the MNL model (Theorem 4.1). A natural question is whether CAP continues to perform well when γ is small but positive. To address this, we derive a first-order expansion of G-RRM probabilities around their MNL limit and use it to bound the revenue gap of CAP.

Proposition 6 (First-order approximation of G-RRM by MNL) *As $\gamma \to 0$, the G-RRM choice probabilities converge smoothly to their MNL counterparts. For any fixed assortment $\mathcal{S}$,*

$$\pi_i^{\text{G-RRM}}(\mathcal{S};\gamma) = \pi_i^{\text{MNL}}(\mathcal{S}) + O(\gamma), \qquad i \in \mathcal{S},$$

with deviation bounded linearly in γ under uniformly bounded attribute differences. More precisely,

$$\pi_i^{\text{G-RRM}}(\mathcal{S};\gamma) = \pi_i^{\text{MNL}}(\mathcal{S}) + \gamma\,\pi_i^{\text{MNL}}(\mathcal{S}) \left(\sum_{j\in\mathcal{S}} \pi_j^{\text{MNL}}(\mathcal{S})\,\Delta_j - \Delta_i \right) + O(\gamma^2),$$

where $\Delta_i = \sum_{j\in\mathcal{S}\setminus\{i\}} \sum_m (e^{-\beta_m(x_{jm}-x_{im})} - 1)$. In particular, there exists $K(\mathcal{S}) > 0$ such that

$$\max_{i\in\mathcal{S}} \left| \pi_i^{\text{G-RRM}}(\mathcal{S};\gamma) - \pi_i^{\text{MNL}}(\mathcal{S}) \right| \;\leq\; K(\mathcal{S})\,\gamma$$

for all sufficiently small γ.

The result follows by a first-order expansion of the regret terms in γ.

Corollary 4.3 (Value guarantee for CAP in the low-regret regime) *Suppose revenues are bounded ($r_i \le \bar{r}$) and attribute differences satisfy the condition of Proposition 6. Then there exists $C > 0$ such that, for all sufficiently small γ,*

$$\sup_{\mathcal{S}\subseteq\mathcal{U}} \left| f^{\text{G-RRM}}(\mathcal{S};\gamma) - f^{\text{MNL}}(\mathcal{S}) \right| \;\le\; C\gamma.$$

Consequently, if $\mathcal{S}_{\text{CAP}}(\gamma)$ is the CAP solution and $\mathcal{S}^*(\gamma)$ the G-RRM optimum, then

$$0 \;\le\; f^{\text{G-RRM}}(\mathcal{S}^*(\gamma);\gamma) - f^{\text{G-RRM}}(\mathcal{S}_{\text{CAP}}(\gamma);\gamma) \;\le\; 2C\gamma,$$

so CAP attains $O(\gamma)$ near-optimality as $\gamma \to 0$.

The uniform $O(\gamma)$ convergence of G-RRM to MNL (Proposition 6), together with bounded revenues, implies the first bound. The CAP guarantee then follows immediately: both the G-RRM optimum and the CAP solution lie within $C\gamma$ of the same MNL benchmark, so their difference is at most $2C\gamma$ by triangle inequality.

4.7 Performance Analysis and Assortment Insights

This section quantifies when context effects matter for assortment design and how they alter optimal sets under G-RRM. Section 4.7.1 specifies the experimental design and parameterization. Section 4.7.2 establishes that the CAP algorithm attains near-optimal solutions with modest runtime. Section 4.7.3 contrasts MNL and G-RRM outcomes across competitor profiles and dimensions, reporting agreement rates, change patterns, and revenue/share impacts. Section 4.7.4 explains the mechanisms: it measures within-set structures (compromise and attraction) and shows how optimization repositions assortments relative to competitors in attribute and behavioral space.

4.7.1 Experimental Setup

We conduct experiments across three attribute dimensions ($M \in \{2, 3, 4\}$), three universal set sizes ($|\mathcal{U}| \in \{8, 12, 16\}$), and five competitor profiles, generating 100 independent instances for each combination, for a total of $4{,}500$ instances. Competitor performance is summarized by a relative attribute rank (RAR) $\sigma_m \in [0, 1]$, defined as the percentile position of the competitor on attribute m. We consider five cases: (i) underperforming ($\sigma_m = 0$), (ii) price-strong ($\sigma_{\text{price}} = 0.75, \sigma_{m'} = 0.25$), (iii) balanced ($\sigma_m = 0.5$), (iv) attribute-strong ($\sigma_{\text{price}} = 0.25, \sigma_{m'} = 0.75$), and (v) high-performing ($\sigma_m = 1$).

Behavioral variation is introduced through the G-RRM coefficients. For each attribute dimension $M \in \{2, 3, 4\}$, we consider five scenarios: (i) equal sensitivities ($\beta_m = -1\ \forall m$); (ii) one attribute stronger ($\beta = -2$ vs. -1); (iii) the reverse; (iv) uniformly weaker ($\beta_m = -0.5$); and (v) very weak ($\beta_m = -0.25$). These cases span environments with symmetric versus asymmetric sensitivities and with strong versus weak overall preferences.

For each attribute dimension, an MNL surrogate is obtained by projecting G-RRM choice probabilities onto the MNL model. The projection is implemented via Kullback–Leibler divergence minimization on training menus, with evaluation on holdout sets. The complete set of estimated parameters is reported in A.7, Table 1.

4.7.2 Computational Performance

We evaluate the computational efficiency of the Context-Aware Policy (CAP) algorithm by varying the number of alternatives ($|\mathcal{U}| \in \{6, 10, 16\}$) and the number of attributes ($M \in \{2, 3, 4\}$). For these instance sizes, the global optimum can be obtained by exhaustive enumeration, allowing direct measurement of optimality gaps.

Table 4.2 reports the average optimality gap and computation time. Across all cases, the average gap is 0.02%. Runtime increases with $|\mathcal{U}|$ and M, but remains below one second even for the largest configuration ($|\mathcal{U}| = 16, M = 4$). These results show that CAP computes solutions that are effectively optimal with modest computation time.

Table 4.2: Computational performance of CAP algorithm

	Gap (%)				**CAP Time (sec)**			
$\lvert\mathcal{U}\rvert$	M=2	M=3	M=4	**Ave.**	M=2	M=3	M=4	**Ave.**
6	0.04	0.03	0.01	0.03	0.01	0.01	0.01	0.01
10	0.03	0	0	0.01	0.10	0.12	0.12	0.11
16	0.00	0.04	0.06	0.03	0.83	1.00	1.00	0.94
Ave.	0.02	0.02	0.02	0.02	0.31	0.38	0.38	0.36

4.7.3 Assortment Outcomes under MNL and G-RRM

We compare the optimal assortments under the benchmark MNL model and the context-sensitive G-RRM model across five competitor profiles and three attribute dimensions ($M \in \{2, 3, 4\}$), averaging over 100 replications. Results are reported as proportions with Wilson 95% confidence intervals and as means with bootstrap 95% intervals; change decompositions and exclusion rates are conditional on non-identical cases, and performance deltas are measured as relative percentage changes under the G-RRM data-generating process. The main results are summarized in Table 4.3, with additional tables provided in Appendix A.8.

Table 4.3 summarizes differences between MNL- and G-RRM-optimal assortments across competitor profiles and attribute dimensions ($M = 2, 3, 4$), with preference coefficients set to -1 for all attributes. The **Agreement** columns report the share of identical assortments (Id.), the Jaccard overlap of products between the two sets (Overlap), and the mean number of differing items ($|\Delta S|$). The **Change** column records the share of non-identical cases in which the G-RRM set strictly expands the MNL set (Expand). The **Excl.** columns measure the frequency with which G-RRM omits the MNL top-revenue item (Top-Rev Excl.) or the MNL top-choice item (Top-Share Excl.). Finally, the **Impact** columns report the relative change in expected revenue (ΔRev) and market share (ΔShare) when both assortments are evaluated under the G-RRM model.

Agreement between MNL- and G-RRM-optimal assortments is high only when competition is weak. With underperforming rivals, assortments coincide in up to 35% of cases in two dimensions and exceed 60% with more attributes, with Jaccard overlaps of 74–95%. Disagreements

*Table 4.3: MNL vs. G-RRM optimal assortments: compact summary by dimension (M) and competitor profile (*Prof*).*

		Agreement			**Change**	**Excl.**		**Impact**	
M	Prof (tag)	Id. (%)	Overlap (Jac,%)	$\|\Delta S\|$ (#)	Expand (%)	Top-Rev Excl. (%)	Top-Share Excl. (%)	ΔRev (%)	ΔShare (%)
2	AS	4	60.5	2.95	13.5	7.3	8.3	-3.2	-4.6
	BAL	1	56.0	3.20	0.0	14.1	22.2	-0.5	-4.3
	HP	0	16.0	10.2	0.0	55.0	79.0	11.9	-26.8
	PS	11	63.7	3.04	0.0	36.0	44.9	15.4	7.3
	UP	35	77.2	1.31	98.5	1.5	1.5	-0.5	-0.1
3	AS	0	52.2	4.45	36.0	13.0	18.0	-2.5	-0.8
	BAL	2	51.6	4.05	0.0	29.6	34.7	3.0	-6.5
	HP	0	10.8	12.6	0.0	76.0	92.0	40.4	-39.5
	PS	9	63.6	3.08	0.0	44.0	47.3	19.4	7.8
	UP	31	74.4	1.36	100	0.0	0.0	-0.8	-0.4
4	AS	0	42.8	5.61	26.0	41.0	44.0	1.6	-11.0
	BAL	3	53.5	3.99	2.1	32.0	38.1	7.3	-3.7
	HP	0	9.6	13.8	0.0	87.0	98.0	94.0	-40.3
	PS	21	70.0	2.84	16.5	46.8	53.2	18.2	6.6
	UP	12	67.6	1.23	100	0.0	0.0	-0.8	-0.6

Legend. Prof tags: UP=*underperforming*, PS=*price-strong*, BAL=*balanced*, AS=*attribute-strong*, HP=*high-performing*. *Agreement*: Id.=share of identical sets; Overlap=Jaccard overlap; $|\Delta S|$=mean # differing items (non-identical cases). *Change*: Expand=share of non-identical cases with $S^{\mathrm{G}} \supset S^{\mathrm{M}}$. *Excl.*: Top-Rev Excl.=share of non-identical cases where the MNL top revenue-contribution item ($r_i\pi_i(S^{\mathrm{M}})$) is absent from S^{G}; Top-Share Excl.=analogous for the highest MNL probability item in S^{M}. *Impact*: ΔRev, ΔShare are relative changes (G-RRM optimal minus MNL optimal), both evaluated under G-RRM.

are minor ($|\Delta S| < 1.4$) and almost always expansions, while exclusions of top-revenue or top-choice items are essentially absent. Revenue and share differences remain within $\pm 1\%$. In such settings, context effects have little impact: the standard MNL model captures assortments well when competitors are clearly dominated.

As competition strengthens, divergence becomes systematic. High-performing rivals produce assortments that never coincide with MNL (0% identical across all M), with overlaps collapsing below 20% and disagreement sizes exceeding 10 items in 2D and 14 in 4D. G-RRM frequently excludes the very products that MNL highlights: the MNL top-revenue item is dropped in 55–87% of cases and the top-choice item in 79–98%. These shifts yield substantial revenue gains under G-RRM evaluation (up to +40% in 3D and +94% in 4D) but at the cost of steep share losses (–27 to –40 percentage points). Price-strong competitors also drive consistent divergence: assortments coincide in only 9–26% of cases, top-revenue items are excluded in 36–53%, and revenues rise by 15–19% with modest share gains. In these markets, regret sensitivity reshapes assortments fundamentally, reversing MNL's emphasis on "star" products and altering platform strategy.

Dimensionality amplifies these differences. With balanced competitors, identical assortments fall from 1–5% in two dimensions to 0–3% in four, while Jaccard overlaps decline to just above 50%. Attribute-strong rivals show a similar trajectory: overlaps drop from 60% in 2D to 43% in 4D, disagreement sizes increase from 3 to over 5, and top-revenue exclusions climb from 7–13% to 26–41%. These changes magnify performance swings: while revenue effects are modest in low dimensions ($\pm 3\%$), in 4D they reach +7% for balanced profiles and –11% for attribute-strong profiles. As the attribute space expands, behavioral effects interact with competition to alter assortments more forcefully and with clearer economic consequences.

Overall, three insights emerge. First, MNL suffices in simple, weakly competitive markets, where ignoring context effects rarely changes outcomes. Second, in stronger markets G-RRM systematically overturns MNL prescriptions, often displacing the very items MNL deems indispensable. Third, higher dimensionality magnifies these differences, producing larger shifts in composition and performance. For practice, these findings imply that logit-based models are adequate for low-stakes design problems, but in competitive and high-dimensional environments they risk systematically misguiding decision makers. In such cases, incorporating regret-based choice is not only behaviorally realistic but economically decisive, as it reshapes platform menus, revenue potential, and market share in ways that the standard logit framework cannot anticipate.

4.7.4 Context Effects in Optimal Assortments and Competitive Repositioning

Compromise and attraction effects are central to behavioral choice theory, but their implications for assortment optimization remain ambiguous. They can in principle act as *assets*, amplifying the demand for targeted items, or as *liabilities*, distorting substitution patterns and weakening overall performance.

This section evaluates their role along two dimensions. *Internally*, we examine how optimal assortments handle context-sensitive structures—whether they preserve them, eliminate them, or selectively retain them. *Externally*, we investigate how optimization shapes the position of the entire assortment relative to the competitor, through shifts in geometric and behavioral centers and through suppression mechanisms.

Table 4.4 quantifies the *internal composition* of assortments by tracking two types of context-sensitive structures: compromise (κ) and attraction/decoy (α). For each, we report both their frequency ($\kappa_{\text{cfg}}, \alpha_{\text{cfg}}$) and their aggregate behavioral gain ($\kappa_{\text{tot}}, \alpha_{\text{tot}}$), defined as the increase in the focal item's choice probability due to the presence of contextual alternatives. The table presents these diagnostics for the universal set, the G-RRM optimum, and the MNL optimum, averaged across 100 replications by competitor profile and preference vector. These measures reveal whether optimization preserves, eliminates, or selectively retains context effects when forming optimal assortments.

Whereas Table 4.4 focuses on structures within assortments, Table 4.5 evaluates their *external positioning* relative to competitors. Here, distances (d^{geo}, d^{beh}) capture how far assortment centers lie from the competitor, with U, S_G, and S_M giving levels for the universal set, the G-RRM optimum, and the MNL optimum. The ΔS_G and ΔS_M columns record changes relative to the universal baseline. Alignment metrics assess whether optimization reduces distortions between geometry and behavior, through the change in misalignment ($\Delta\|\mathbf{x}^{beh} - \mathbf{x}^{geo}\|$) and the correlation ρ^{geo} between centrality and choice probability. Finally, $\bar{\kappa}_c$ and $\bar{\alpha}_c$ quantify competitor suppression, capturing how compromise and attraction effects reduce the rival's share. Together, these diagnostics establish how assortment optimization shapes both the *internal structure* of menus and their *external stance* in the market.

Table 4.4: Within-set context effects: compromise (κ) and attraction (α) structures in the universal set and under optimal assortments. Values are averaged over 100 replications.

		Universal Set				G - RRM Opt.				MNL Opt.			
β	Prof (tag)	κ_{cfg} (#)	α_{cfg} (#)	κ_{tot} (gain)	α_{tot} (gain)	κ_{cfg} (#)	α_{cfg} (#)	κ_{tot} (gain)	α_{tot} (gain)	κ_{cfg} (#)	α_{cfg} (#)	κ_{tot} (gain)	α_{tot} (gain)
(-1,-1)	AS	9.31	60.48	0.24	0.28	2.63	3.32	0.01	0.00	3.73	5.84	0.01	0.01
	BAL	9.31	60.48	0.24	0.28	0.37	1.93	0.00	0.00	3.86	7.04	0.01	0.00
	HP	9.31	60.48	0.24	0.28	0.00	0.00	0.00	0.00	24.43	19.24	0.18	0.04
	PS	9.31	60.48	0.24	0.28	0.42	1.54	0.00	0.00	3.80	5.38	0.01	0.01
	UP	9.31	60.48	0.24	0.28	0.52	1.59	0.00	0.00	0.04	0.74	0.00	0.00
(-1,-2)	AS	9.31	60.48	0.59	0.86	7.00	5.38	0.32	0.08	3.66	6.63	0.04	0.04
	BAL	9.31	60.48	0.59	0.86	0.32	1.22	0.00	0.00	1.64	5.88	0.03	0.02
	HP	9.31	60.48	0.59	0.86	0.00	0.00	0.00	0.00	27.55	27.83	1.41	0.26
	PS	9.31	60.48	0.59	0.86	0.30	1.12	0.00	0.00	0.62	2.66	0.00	0.01
	UP	9.31	60.48	0.59	0.86	0.36	1.07	0.00	0.00	0.00	0.13	0.00	0.00
(-2,-1)	AS	9.31	60.48	0.63	0.71	7.50	13.98	0.40	0.25	3.51	8.63	0.03	0.06
	BAL	9.31	60.48	0.63	0.71	5.66	9.40	0.23	0.11	7.07	13.32	0.12	0.12
	HP	9.31	60.48	0.63	0.71	0.00	0.07	0.00	0.00	10.50	55.32	0.63	0.58
	PS	9.31	60.48	0.63	0.71	2.03	6.34	0.00	0.02	7.55	18.33	0.15	0.18
	UP	9.31	60.48	0.63	0.71	0.43	1.47	0.00	0.00	0.05	0.54	0.00	0.00
(-0.5,-0.5)	AS	9.31	60.48	0.00	0.00	0.50	1.31	0.00	0.00	2.86	4.07	0.00	0.00
	BAL	9.31	60.48	0.00	0.00	0.36	1.00	0.00	0.00	2.31	5.12	0.00	0.00
	HP	9.31	60.48	0.00	0.00	0.13	0.19	0.00	0.00	8.44	17.06	0.00	0.00
	PS	9.31	60.48	0.00	0.00	0.46	0.95	0.00	0.00	2.17	4.33	0.00	0.00
	UP	9.31	60.48	0.00	0.00	0.77	1.91	0.00	0.00	0.26	1.63	0.00	0.00
(-0.25,-0.25)	AS	9.31	60.48	0.00	0.00	0.71	1.90	0.00	0.00	1.71	2.31	0.00	0.00
	BAL	9.31	60.48	0.00	0.00	0.62	1.93	0.00	0.00	1.81	3.00	0.00	0.00
	HP	9.31	60.48	0.00	0.00	0.14	0.24	0.00	0.00	5.36	7.28	0.00	0.00
	PS	9.31	60.48	0.00	0.00	0.61	1.69	0.00	0.00	1.61	2.64	0.00	0.00
	UP	9.31	60.48	0.00	0.00	0.94	1.87	0.00	0.00	0.51	1.87	0.00	0.00

Table 4.4 shows that context-sensitive structures are ubiquitous in universal product sets. On average, menus contain about nine compromise and sixty attraction opportunities, yet their aggregate behavioral pull is modest ($\kappa_{\text{tot}} \approx 0.2$, $\alpha_{\text{tot}} \approx 0.3$). Most structures are therefore weak in isolation, but their sheer abundance creates ample potential for distortions if they are carried over into optimized menus.

Optimization reveals a stark contrast between the two models. G-RRM assortments aggressively prune context effects, typically retaining fewer than one configuration and often none at all. MNL assortments, by contrast, systematically preserve them, with counts ranging from single digits to more than fifty depending on the competitive setting. The divergence is sharpest under high-performing rivals: G-RRM eliminates all compromise and attraction opportunities ($\kappa_{\text{cfg}} = \alpha_{\text{cfg}} = 0$), whereas MNL preserves dozens, with aggregate attraction pull reaching $\alpha_{\text{tot}} = 0.58$ at $\beta = (-2, -1)$.

Competitive intensity and preference strength amplify these differences. When rivals are weak, both models converge to nearly structure-free menus, consistent with dominance making contextual distortions irrelevant. When preferences are strong, MNL retains large totals, exposing assortments to skew and cannibalization, while G-RRM suppresses them entirely. When preferences are weak, both models again converge, as context effects dissipate when attribute trade-offs are muted. Taken together, these patterns clarify the mechanism behind the divergences observed in Table 4.3: universal sets contain plentiful but weak context structures, MNL passes them through into optimal assortments, while G-RRM eliminates them unless strategically useful. The methodological value of regret-based optimization therefore lies not in exploiting context effects, but in preventing their harmful persistence in competitive assortments.

Table 4.5: Distances, alignment, and suppression metrics under MNL- and G-RRM-optimal assortments compared to the universal set. Distances d^{geo} and d^{beh} are to competitor c; deltas are relative to the universal set U.

Config.		U		S_G		S_M		ΔS_G		ΔS_M		Align.		$\bar{\kappa}_c, \bar{\alpha}_c$	
Root	Prof.	d^{geo}	d^{beh}	d^{geo}	d^{beh}	d^{geo}	d^{beh}	Δd^{geo}	Δd^{beh}	Δd^{geo}	Δd^{beh}	ρ^{geo}	$\Delta\lVert\mathbf{x}^{\text{beh}} - \mathbf{x}^{\text{geo}}\rVert$	$\bar{\kappa}_c$	$\bar{\alpha}_c$
(-1,-1)	AS	0.81	1.36	0.35	0.23	0.41	0.17	-0.46	-1.13	-0.40	-1.19	-0.07	-0.35	0.06	0.09
	BAL	0.17	0.99	0.94	0.88	0.78	0.72	0.77	-0.11	0.61	-0.26	-0.05	-0.33	0.13	0.07
	HP	1.16	0.06	1.38	1.32	1.33	0.69	0.22	1.26	0.16	0.63	—	-0.31	0.00	0.00
	PS	0.94	1.30	1.88	1.86	1.69	1.69	0.93	0.57	0.74	0.39	-0.16	-0.35	0.68	0.15
	UP	1.36	2.43	1.27	1.54	1.43	1.57	-0.08	-0.89	0.07	-0.85	-0.28	-0.42	0.86	0.98
(-1,-2)	AS	0.81	1.36	0.39	0.27	0.42	0.21	-0.42	-1.09	-0.39	-1.15	-0.08	-0.30	0.16	0.16
	BAL	0.17	0.99	0.97	0.98	0.89	0.86	0.80	-0.01	0.72	-0.13	-0.04	-0.32	0.53	0.12
	HP	1.16	0.06	1.46	1.40	1.29	0.59	0.30	1.34	0.12	0.53	—	-0.28	0.00	0.00
	PS	0.94	1.31	1.87	2.00	1.85	1.94	0.93	0.69	0.90	0.63	-0.13	-0.28	1.47	0.81
	UP	1.36	2.43	1.28	1.64	1.51	1.64	-0.07	-0.79	0.15	-0.79	-0.30	-0.31	1.96	2.21
(-2,-1)	AS	0.81	1.37	0.46	0.28	0.41	0.24	-0.35	-1.09	-0.40	-1.13	-0.26	-0.37	0.35	0.59
	BAL	0.17	0.99	0.67	0.53	0.59	0.45	0.50	-0.46	0.42	-0.55	-0.07	-0.34	0.27	0.29
	HP	1.16	0.06	0.71	0.68	1.21	0.13	-0.46	0.63	0.05	0.08	-0.25	-0.19	0.00	0.00
	PS	0.94	1.30	1.73	1.53	1.39	1.31	0.79	0.23	0.45	0.01	-0.13	-0.36	1.48	0.43
	UP	1.36	2.43	1.28	1.53	1.45	1.58	-0.07	-0.90	0.09	-0.86	-0.25	-0.41	0.86	0.99
(-0.5,-0.5)	AS	0.81	1.31	0.36	0.30	0.42	0.25	-0.45	-1.00	-0.39	-1.06	-0.22	-0.38	0.00	0.02
	BAL	0.17	0.92	0.97	0.97	0.87	0.81	0.80	0.06	0.70	-0.11	-0.15	-0.32	0.02	0.01
	HP	1.16	0.13	1.66	1.64	1.46	1.12	0.50	1.51	0.30	0.99	-0.01	-0.31	0.00	0.00
	PS	0.94	1.25	1.88	1.88	1.76	1.75	0.94	0.64	0.82	0.50	-0.21	-0.34	0.14	0.02
	UP	1.36	2.35	1.30	1.45	1.32	1.43	-0.05	-0.90	-0.04	-0.92	-0.24	-0.23	0.66	0.32
(-0.25,-0.25)	AS	0.81	1.11	0.48	0.42	0.45	0.39	-0.34	-0.69	-0.36	-0.72	-0.28	-0.38	0.00	0.00
	BAL	0.17	0.65	0.99	0.95	0.94	0.88	0.82	0.30	0.78	0.23	-0.28	-0.34	0.01	0.00
	HP	1.16	0.41	1.70	1.70	1.63	1.51	0.54	1.30	0.47	1.10	-0.25	-0.22	0.00	0.00
	PS	0.94	1.07	1.87	1.83	1.82	1.77	0.93	0.76	0.88	0.70	-0.28	-0.36	0.03	0.00
	UP	1.36	2.08	1.33	1.34	1.33	1.36	-0.03	-0.74	-0.02	-0.73	-0.29	-0.26	0.34	0.15

Note: Spearman's ρ^{geo} is only reported for assortments with at least 3 items.

Having shown that optimization prunes most context-sensitive structures from universal menus (Table 4.4), we now examine how assortments reposition relative to the competitor. The diagnostics in Table 4.5 show consistent adjustments across profiles that reveal how regret-based optimization reshapes market interactions.

A striking feature is that assortments do not remain neutral in attribute space but systematically shift with the strength of the rival. Against underperforming or attribute-strong competitors, assortments move inward, drawing closer in both geometric and behavioral distance (e.g., AS at $(-1,-1)$: $\Delta d^{\text{geo}}_{S_G} = -0.46$, $\Delta d^{\text{beh}}_{S_G} = -1.13$). In contrast, balanced and price-strong rivals trigger outward shifts (e.g., PS at $(-2,-1)$: $\Delta d^{\text{geo}}_{S_G} = +0.79$), a defensive repositioning that increases separation. High-performing rivals generate yet another pattern: geometry changes little, but behavioral centers are pulled closer (e.g., HP at $(-1,-1)$: $\Delta d^{\text{beh}}_{S_G} = +1.26$), indicating that demand remains anchored by dominant competitors even when menus move outward.

These positional changes are accompanied by deeper structural adjustments. Across nearly all cases, behavioral responses exceed geometric ones, with shifts in d^{beh} outpacing those in d^{geo} (e.g., BAL at $(-2,-1)$: $\Delta d^{\text{geo}}_{S_G} = +0.50$ versus $\Delta d^{\text{beh}}_{S_G} = -0.46$). Optimization also reduces the misalignment between behavioral and geometric centers, producing consistently negative

$\Delta\|\mathbf{x}^{\text{beh}} - \mathbf{x}^{\text{geo}}\|$. The effect is most pronounced under strong rivals (PS, HP), where universal sets display large distortions (e.g., HP at $(-1, -1)$: $\Delta\|\cdot\| = -1.05$).

Finally, competitor suppression emerges as profile-dependent. For underperforming and price-strong rivals, assortments retain just enough context structures to reduce the rival's share (e.g., PS at $(-2, -1)$: $\bar{\kappa}_c = 1.48$, $\bar{\alpha}_c = 0.43$). Balanced and attribute-strong competitors experience weaker suppression, and in the case of high-performing rivals both measures collapse to zero, underscoring that contextual mechanisms cannot erode a dominant competitor.

Two mechanisms connect the internal pruning of context effects (Table 4.4) with the external repositioning of assortments (Table 4.5).

The first is the reduction of behavior–geometry misalignment. In universal sets, abundant compromise and attraction structures distort predicted choice probabilities, shifting behavioral centers away from their geometric counterparts. G-RRM eliminates these distortions: by pruning context effects, it produces negative values of $\Delta\|\mathbf{x}^{\text{beh}} - \mathbf{x}^{\text{geo}}\|$, bringing behavioral centers back into line with assortment geometry. MNL, which preserves weak but numerous context structures, carries these distortions into optimal menus. Thus, regret-based optimization is effective not because it creates new demand leverage, but because it removes the misalignments that otherwise dilute geometric positioning.

The second mechanism is strategic suppression. While G-RRM usually eliminates context effects, it sometimes retains a small number when they weaken moderate rivals. This targeted use is visible in profiles such as underperforming (UP) and price-strong (PS), where $\bar{\kappa}_c$ and $\bar{\alpha}_c$ take positive values (e.g., PS at $\beta = (-2, -1)$: $\bar{\kappa}_c = 1.48$, $\bar{\alpha}_c = 0.43$). Against dominant high-performing (HP) rivals, however, suppression collapses to zero: no local distortion can dislodge a competitor that already anchors demand. Pruning under G-RRM is therefore selective: harmful structures are eliminated, but those offering tactical leverage are retained.

These contrasts are amplified by preference sensitivity. When preferences are strong (large $|\beta|$), MNL preserves dozens of context structures and often overexposes weak items (e.g., HP at $\beta = (-2, -1)$: $\alpha_{\text{tot}} = 0.58$ with more than 55 attraction structures). G-RRM, in contrast, eliminates them entirely ($\kappa_{\text{cfg}} = \alpha_{\text{cfg}} = 0$). When preferences are weak (e.g., $\beta = (-0.25, -0.25)$), both models converge toward negligible totals, since shallow trade-offs make context effects largely irrelevant. The advantage of regret-based design therefore emerges most clearly when preferences are strong, where unpruned context structures would otherwise amplify misalignment and competitor leverage.

Conclusion The evidence across Tables 4.4 and 4.5 shows that G-RRM's advantage lies in *elimination rather than exploitation*. It prunes context structures that dilute demand or misalign choice probabilities with geometry, and retains them only when they provide tactical suppression against moderate rivals (UP, PS). As a result, assortments are more internally aligned, more effectively positioned against competitors, and more robust across preference regimes.

MNL, by contrast, preserves context indiscriminately: it sustains compromise and attraction configurations even when they amplify distortions or cannibalization. Regret-based optimization is therefore best understood as a *discipline of selective pruning*, not a strategy of amplification.

4.8 Conclusion

This study examined the role of context effects in assortment optimization by embedding the General Random Regret Model (G-RRM) into a competitive choice environment. Our results show that the value of context-aware optimization does not lie in amplifying compromise or attraction effects, but in *selectively eliminating them when they dilute demand and redeploying them only when they suppress vulnerable rivals*. In weakly competitive settings assortments under G-RRM and the benchmark Multinomial Logit (MNL) model converge, but in stronger markets the differences are systematic: G-RRM prunes harmful structures, realigns behavioral and geometric centers, and strategically repositions assortments relative to competitors.

Methodologically, the paper contributes a unified framework for studying context effects in assortment design. We introduce diagnostics to quantify compromise and attraction structures, show how they drive divergences between MNL and G-RRM, and propose an efficient optimization algorithm. Importantly, the model nests MNL as a special case when regret sensitivity vanishes, ensuring consistency between context-free and context-sensitive regimes.

The experimental evaluation demonstrates that regret-based optimization yields assortments that are more robust to competition. In revenue-oriented environments, G-RRM systematically suppresses competitor appeal and reallocates demand to profitable items. In share-oriented environments, it trims distortions and prevents overexposure to weak alternatives. These results highlight how behavioral assortment optimization can inform the design of platform-based mobility and logistics services, where competition and attribute trade-offs are central.

Naturally, the analysis rests on synthetic data, which allows systematic exploration but limits immediate external validity. Future research should complement these findings with empirical studies, testing whether the pruning and suppression mechanisms observed here also arise in real-world assortment settings.

Bibliography

Atasoy, B., T. Ixkeda, X. Song, M. E. Ben-Akiva (2015) The concept and impact analysis of a flexible mobility on demand system, *Transportation Research Part C: Emerging Technologies*, 56, pp. 373–392.

Ausseil, R., J. A. Pazour, M. W. Ulmer (2022) Supplier menus for dynamic matching in peer-to-peer transportation platforms, *Transportation Science*, 56(5), pp. 1304–1326.

Azadeh, S. S., B. Atasoy, M. E. Ben-Akiva, M. Bierlaire, M. Maknoon (2022) Choice-driven dial-a-ride problem for demand responsive mobility service, *Transportation Research Part B: Methodological*, 161, pp. 128–149.

Ben-Akiva, M. E., S. R. Lerman (1985) *Discrete choice analysis: theory and application to travel demand*, vol. 9, MIT press.

Berbeglia, G., A. Garassino, G. Vulcano (2022) A comparative empirical study of discrete choice models in retail operations, *Management Science*, 68(6), pp. 4005–4023.

Chorus, C. G. (2014) A generalized random regret minimization model, *Transportation research part B: Methodological*, 68, pp. 224–238.

Davis, J. M., G. Gallego, H. Topaloglu (2014) Assortment Optimization Under Variants of the Nested Logit Model, *Operations Research*, 62(2), pp. 250–273.

Guevara, C. A., M. Fukushi (2016) Modeling the decoy effect with context-rum models: Diagrammatic analysis and empirical evidence from route choice sp and mode choice rp case studies, *Transportation Research Part B: Methodological*, 93, pp. 318–337.

Heger, J., R. Klein (2024) Assortment optimization: a systematic literature review, *OR Spectrum*, 46(4), pp. 1099–1161.

Huber, J., J. W. Payne, C. Puto (1982) Adding asymmetrically dominated alternatives: Violations of regularity and the similarity hypothesis, *Journal of Consumer Research*, 9(1), pp. 90–98.

Huber, J., C. Puto (1983) Market boundaries and product choice: Illustrating attraction and substitution effects, *Journal of Consumer Research*, 10(1), pp. 31–44.

Kivetz, R., O. Netzer, V. Srinivasan (2004) Alternative models for capturing the compromise effect, *Journal of Marketing Research*, 41(3), pp. 237–257.

Loomes, G., R. Sugden (1982) Regret theory: An alternative theory of rational choice under uncertainty, *The Economic Journal*, 92(368), pp. 805–824.

Luce, R. D. (1977) The choice axiom after twenty years, *Journal of Mathematical Psychology*, 15(3), pp. 215–233.

Rooderkerk, R. P., H. J. Van Heerde, T. H. Bijmolt (2011) Incorporating context effects into a choice model, *Journal of Marketing Research*, 48(4), pp. 767–780.

Simonson, I. (1989) Choice based on reasons: The case of attraction and compromise effects, *Journal of Consumer Research*, 16(2), pp. 158–174.

Talluri, K., G. Van Ryzin (2004) Revenue management under a general discrete choice model of consumer behavior, *Management Science*, 50(1), pp. 15–33.

Ulmer, M. W. (2020) Dynamic pricing and routing for same-day delivery, *Transportation Science*, 54(4), pp. 1016–1033.

Wedell, D. H. (1991) Distinguishing among models of contextually induced preference reversals., *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 17(4), p. 767.

Yang, X., A. K. Strauss, C. S. Currie, R. Eglese (2016) Choice-based demand management and vehicle routing in e-fulfillment, *Transportation Science*, 50(2), pp. 473–488.

Chapter 5

Conclusions

This dissertation examined demand management strategies for shared mobility services from a behavioral perspective. The motivation arose from the critical role that both users and service providers play in realizing the environmental and societal benefits of shared mobility, such as reducing emissions, improving urban mobility, and enhancing accessibility. Ultimately, users' decisions determine whether and how these services are utilized, while providers' decisions shape the characteristics of the services offered. Although these decisions are inherently interdependent, existing studies often treat user choices as external, independent of providers' actions, or model them under assumptions of perfect rationality. In reality, user behavior is influenced not only by service attributes, such as quality, which is determined by providers' operational strategies, but also by inter- and intrapersonal behavioral factors that cannot be fully explained by rational choice theory. These behavioral factors can be predicted and incorporated as demand management levers for shared mobility services.

The research questions outlined in Chapter 1 are addressed in Section 5.1 with a summary of key findings, the key contributions are discussed in Section 5.2, and practical insights for providers and policymakers are presented in Section 5.3. Finally, directions for future research are outlined in Section 5.4.

5.1 Summary of the Key Findings

This dissertation attempted to provide an answer to the following main research question:

How can shared mobility service providers incorporate users' behavioral factors in their demand management strategies?

This dissertation addresses this research question by identifying key behavioral factors that have a predictable impact on user choices with potential implications for shared mobility operations. To this end, we first conducted a literature review to detect relevant behavioral factors in user decision-making that are largely unexplored in operational models of shared mobility services. To integrate these factors and analyze their implications for operations, we adopted an analytical modeling approach grounded in behavioral theory, embedding the identified factors into optimization models and evaluating their effects through scenario-based sensitivity analysis. The specific studies, findings, and implications are presented through the following sub-research questions, formulated to address the main research question.

SQ1. What users' behavioral factors should be integrated into the operations of shared mobility services?
Chapter 2 addresses this question through a literature review and the development of a general conceptual framework that connects user behavior with provider operations. The review identified the behavioral factors most relevant yet underrepresented in previous studies that model providers' operations. It shows that accounting for user decisions is essential to tackle the persistent challenges of adoption, profitability, and environmental impact in shared mobility. Incorporating behavioral effects in modeling users' decision-making holds significant potential to make operational models more realistic and impactful, bridging the gap between behavioral insights and mobility practice. Nevertheless, the existing literature overwhelmingly relies on rational choice theory, overlooking empirically observed behavioral factors that generate systematic and predictable patterns of behavior.

This study highlights social influence and context effects as highly relevant for the service providers' demand management, which are largely overlooked. Social influence, as an interpersonal behavioral factor, shapes users' decisions to adopt and use shared mobility services. Accordingly, potential users who are willing to use a service often imitate the choices of others and rely on their experiences before transitioning into active users. This adoption behavior results in a gradual market penetration of the service. Service deployment should align with the pace of market penetration and gradual adoption in order to effectively support the realization of market potential and profitability. In addition, context effects, as intrapersonal factors, systematically influence user preferences depending on the set of available alternatives. These effects are particularly relevant in operations where providers determine which options to present from the set of feasible service alternatives in response to users' requests for the service.

SQ2. How can service design be adapted to the gradual demand realization of new services as an operational demand management strategy?
Chapter 3 addresses this question by jointly incorporating the impact of service quality and social influence into the modeling of users' adoption decisions, whether adopting a new service provider or a new mobility vehicle. This approach enables the synchronization of service design decisions with adoption growth. In the proposed service design model, user adoption decisions are modeled at an aggregated level and depend explicitly on both the number of current adopters (social influence) and service characteristics, such as coverage and vehicle availability (i.e., level of service). We formulated the service design problem as a mixed-integer linear model for phased deployment, where fleet size and coverage expand over time as adoption grows, with the objective of maximizing total profit over a given planning horizon divided into multiple periods (phases).

We conducted scenario-based analysis by defining different adoption behavior scenarios, with trip patterns from an open-source bike-sharing dataset. The results show that selective coverage with a high level of service (i.e., vehicle availability) outperforms broad coverage with a low level of service. Phased deployment aligned with adoption trajectories improves profitability by matching supply with evolving demand. This increases profits by 4% compared to deployment strategies based on stabilized demand, which assume that all potential users adopt immediately at service introduction, a common but unrealistic assumption in the literature that ignores social influence. Our experiments show that both service quality and social influence must be considered in modeling adoption and service design; ignoring service quality while considering the social influence can reduce profits by up to 54%. We also examined cases where adoption predictions deviate from observations and explored the benefits of phased deployment, which allows predictions and decisions to be updated based on observed outcomes. In our simulations, adjusting predictions and decisions at each deployment phase reduced profit losses caused by inaccurate adoption estimates by up to 47%.

SQ3. How can context effects be leveraged as a behavioral demand management strategy for shared mobility services?
Chapter 4 addresses this question by focusing on the problem of selecting which options to offer a user in response to their mobility request. To this end, we propose an assortment optimization framework that incorporates context effects when selecting the best subset from the given set of all feasible options, for the user and provider, in the presence of a competing option. The competing option represents an external mobility alternative, reflecting the possibility that a user may reject the entire menu and select the competitor.

The assortment optimization model is formulated using the Generalized Random Regret Minimization (G-RRM) model, which predicts context effects, specifically compromise and attraction effects, in user choices. In this Chapter, we also proposed diagnostic metrics to quantify the presence of context effects in any choice set, grounded in empirical and theoretical behavioral insights from the literature. These metrics enable comparison of different choice sets based on the presence and strength of context effects and show how they drive divergences between choice sets selected under context-sensitive and context-ignorant choice models.

We conducted structured, scenario-based numerical experiments to compare choice sets designed using the G-RRM model against those designed with the multinomial logit (MNL) model, which is widely used in assortment optimization due to its tractability but does not account for context effects. The experimental evaluation demonstrates that context-sensitive menu selection yields assortments that are more robust to competition. The context-sensitive G-RRM model improves menu outcomes through two mechanisms: it either systematically suppresses competitor appeal to steer choices toward profitable items, or it selectively removes items to prevent preference shifts that weaken market share against competitors. In weakly competitive settings, assortments generated by G-RRM and the benchmark MNL model converge. In more competitive markets, however, systematic differences emerge: G-RRM removes harmful structures and strategically manages the composition of alternatives. These results highlight how behavioral assortment optimization can inform the design of platform-based mobility and logistics services, where competition and attribute trade-offs are central.

Together, our studies highlight that user behavior is as critical for effective demand management in shared mobility services as operational factors such as fleet size, resource allocation, and pricing. The literature shows that user behavior systematically deviates from rationality assumptions in predictable ways, which can be embedded into operational decision-making.

Our findings suggest that user behavior should therefore be treated as a central component of demand management, rather than as an unpredictable or negligible element separate from operational models. This dissertation develops two complementary modeling frameworks for integrating user behavior into operations as demand management strategies: (i) using adoption growth, shaped by interpersonal social influence, to synchronize service design decisions such as service location, fleet sizing, and repositioning as an operational demand management lever; and (ii) incorporating intrapersonal context effects into service menu design as a behavioral demand management lever. These frameworks enable providers to better guide their demand management strategies, ensuring alignment with user behavior.

5.2 Contributions

This study contributes to behavioral operations management in shared mobility by integrating a behavioral aspect, user adoption behavior, into an operations management aspect, long-term service design models for free-floating services. The main contributions of this dissertation are discussed below.

In this dissertation, we assert that focusing solely on service characteristics is insufficient for promoting shared mobility adoption and managing demand. A shift is needed from treating users as passive or fully rational decision-makers to recognizing them as actors who interact with providers' choices and respond in systematic and predictable ways. This perspective establishes directions for embedding behavioral realism into shared mobility operations and reframes behavioral effects as opportunities to enhance service design and demand management strategies.

Despite considerable literature demonstrating the primary role of service characteristics (e.g., quality and price) (Zhang and Kamargianni, 2023; Bachand-Marleau et al., 2012; Ma et al., 2023; Albiński et al., 2018) in service adoption decisions and the role of social influence (Baddeley, 2010; Deutsch and Gerard, 1955; Hsu and Lu, 2004; Cialdini and Goldstein, 2004) in shaping adoption dynamics, the joint consideration of service quality and user behavior remains largely absent in operational models. Chapter 3 contributes conceptually by challenging the prevailing assumption that users adopt a service once quality expectations are met, and instead introduces an alternative approach that incorporates gradual market penetration driven jointly by service quality and social influence. The proposed service-adoption planning framework leverages gradual adoption behavior and service design to guide demand uptake and optimize profit.

Moreover, this dissertation adopts a different perspective on context effects. While their empirical presence has been widely demonstrated, including the mobility context (Chorus and Bierlaire, 2013; Guevara and Fukushi, 2016; Ding et al., 2025), their implications for choice set design remain underexplored. Chapter 4 conceptualizes context effects as a behavioral demand management lever, demonstrating how behavioral factors (i.e., attraction and compromise effects) can be leveraged to shape preferences and guide choices, thereby expanding the provider's toolkit beyond monetary mechanisms such as pricing and incentives. This research demonstrates that the strength of context-aware optimization lies not in magnifying compromise or attraction effects, but in managing them strategically—removing those that reduce the choice share of the set of expected outcomes while retaining those that weaken competitors. It also highlights that context effects can enhance decision outcomes when harnessed effectively, but undermine the results if not anticipated and controlled.

This dissertation makes methodological contributions that extend existing modeling approaches and practical tools for incorporating behavioral aspects into operations. Chapter 3 advances adoption modeling by extending the Bass diffusion model to incorporate the joint impact of service quality and social influence. This formulation captures the dynamic interplay between service availability, accessibility, and peer influence, providing a behaviorally grounded account of service adoption. This chapter also develops a phased service-adoption model that links deployment decisions, involving fleet sizing and coverage expansion, to adoption growth. The proposed model captures the full adoption trajectory and enables providers to scale operations over time rather than committing to static designs based on mature demand assumptions.

Chapter 4 introduces metrics for quantifying attraction and compromise effects in choice sets of arbitrary composition, moving beyond the stylized three-option settings with fixed relative positions of alternatives (see Huber et al. (1982); Simonson (1989); Guevara and Fukushi (2016)) that are typically examined in the literature. These diagnostic tools translate attraction and compromise effects into measurable constructs, enabling analysis of how context effects shape menu composition and serving as validation tools for assessing their impact on the divergence and convergence of choice sets when benchmarking behaviorally informed designs against rationality-driven models. Moreover, we show that the G-RRM model nests MNL as a special case when regret sensitivity vanishes, ensuring consistency between context-free and context-sensitive models. Additionally, this chapter proposes a tractable algorithm for menu design under the G-RRM model, which captures context effects. The algorithm supports real-time menu design by efficiently selecting subsets of options, making it suitable for integration into platform operations such as matching and routing.

5.3 Practical Implications

This dissertation offers actionable guidance for both managers and policymakers.

For managers engaged in the strategic design and deployment of new services or expansions, we propose gradual service deployment, aligned with market penetration, as a demand management tool that leverages adoption behavior and service design to guide demand uptake and optimize profitability. Our findings indicate that profitability and adoption are best achieved by retaining users through consistently high service quality rather than focusing solely on attracting new users. Depending on the mobility pattern (trip distribution) and users' sensitivity to social influence and service characteristics, launching services in a smaller, high-potential area may be more efficient. In all cases, maintaining high service quality is essential: it secures user retention, supports adoption growth, sustains long-term usage, and underpins profitability.

This research highlights menu design as a competitive tool in providers' interactions with users. Because users are not captive and can easily switch to competing services, providers can strategically present a menu of options rather than only their most desirable offer. By including additional alternatives, providers can increase the choice probability from the set (i.e., probability of purchase) compared to an external competitor. These effects can also be used to curate choice sets that increase the attractiveness of a specific targeted option (e.g., a shared ride, off-peak service, or a highly profitable option) within the choice set and encourage its selection. The findings suggest that the context-ignorant MNL model may be sufficient in situations where the competitor is weak. However, in competitive and high-dimensional environments, where alternatives are evaluated across multiple attributes, ignoring the context effects risks misleading the design of choice sets. In such cases, the use of the context-sensitive G-RRM model is not only behaviorally realistic but also economically critical, as it reshapes platform menus, revenue potential, and market share in ways that the standard logit framework cannot capture. These effects reshape the optimal menu composition to increase the likelihood of purchase against competitors or enhance the appeal of options that are aligned with the provider's objectives (e.g., promoting higher-priced items to maximize revenue). Importantly, context effects can serve as a substitute for, or a complement to, monetary incentives, which are the prevalent demand management strategy.

This work provides policymakers with both the user and provider perspectives on shared mobility adoption. It shows that adoption follows gradual trajectories that must be accounted for when setting targets, allocating subsidies, or evaluating societal impacts. By integrating these perspectives, the research enables the design of regulations and support schemes that are aligned with user behavior, ensure financial sustainability for providers, and motivate their continued engagement.

5.4 Limitations and Future Work

This dissertation advances the integration of behavioral effects as levers in service design and demand management, shifting the focus beyond service quality and attributes toward user behavior. However, much remains to be explored. We outline avenues for further investigation as follows:

- **Empirical validation of behavioral effects.** Although the proposed models in this dissertation are analytically grounded, they rely on synthetic data, limiting real-world generalizability. The challenge is the scarcity of high-resolution, purpose-built datasets that capture both operational attributes and specific behavioral effects. Therefore, future research can prioritize empirical validation of behavioral effects to bridge theory and practice. This requires different data collection strategies. Longitudinal studies using Observational Field Data are needed to track adoption patterns in new services, which requires granular data on user registration and usage patterns to empirically fit the S-shaped adoption curves driven by social influence (Chapter 3) (Jackson and Yariv, 2011). Experimental data, such as A/B Testing and Discrete Choice Experiments, can validate context effects (Chapter 4); providers can integrate controlled A/B tests into digital platforms (e.g., presenting decoy options) to measure shifts in choice probability (Larsen et al., 2024). Similarly, stated preference or discrete choice experiments can quantify user sensitivity to the context effects (Train, 2009).
- **Integrating broader psychological and behavioral determinants of choice.** The proposed framework centers on social influence and context effects that strategically capture two distinct, yet operationally critical, stages of user interaction with shared-mobility ser-

vices. Social Influence is treated as a key interpersonal behavioral effect that determines the aggregated demand and user uptake trajectory. Conversely, Context Effects (Attraction and Compromise effects) are modeled as intrapersonal behavioral factors, which become highly relevant once users have adopted the service and are actively making choices among service options. This dual approach ensures that the models account for both long-term planning based on market penetration (Chapter 3) and short-term decisions (Chapter 4). We acknowledge, however, that the full spectrum of user choice is influenced by broader factors, including a vast range of psychological and individual characteristics. Specifically, user choice is affected by individual socio-demographics and psychological factors (Faghih-Imani et al., 2017; Najmi et al., 2023) These psychological factors include attitudes—an individual's positive or negative evaluative response towards a behavior (e.g., adopting an environmentally friendly service), which is influenced by their beliefs, values, emotions, and perceptions (the subjective interpretation of stimuli, e.g., quality perception, resulting from past experience and the service image like reputation and trust) (Najmi et al., 2023). These are compounded by behavioral factors encompassing cognitive and motivational biases, which include a wide range of documented phenomena, such as reference dependency and loss aversion, risk aversion, the anchoring effect, and confirmation bias. Incorporating these factors impacts the models, primarily by influencing the parameters and/or variables (to incorporate the way users perceive attributes, assign importance to attributes, or evaluate attributes relative to a reference). This may also necessitate changing the structure of the demand functions (i.e., adoption and choice models in chapters 3 and 4) to capture some behavioral effects like risk aversion or reference dependency. While the G-RRM Model captures a form of reference dependency and loss aversion by considering each alternative as a reference point for others, people may also have external reference levels for specific attributes like price. Moreover, people may put more value on the options they are currently using, like their own cars, known as Status Quo Bias Kahneman et al. (1991). This bias represents a strong tendency to maintain the current state. These broader factors could also influence the definition of the choice task itself, as people may ignore certain attributes or alternatives due to cognitive screening.

- **Service design under uncertainty.** In the proposed service design model (Chapter 3), user service quality expectations are assumed to follow known distributions with given parameters. Moreover, adoption parameters, potential market size, and trip distributions are treated as fixed inputs. Future research could relax this assumption by incorporating uncertainty, through stochastic or robust optimization, to improve decision robustness under incomplete data or uncertain adoption and trip patterns.

- **Adaptive service design.** The proposed model for service design under gradual adoption does not incorporate learning mechanisms for updating initial estimates of the adoption based on realized adoption observed during phased implementation. Future research could embed learning mechanisms, such as Bayesian inference or reinforcement learning, directly into the optimization process, enabling real-time adaptation to observed user behavior.

- **Incorporating competition.** In the Service Design model (Chapter 3), competition is implicitly represented by setting the total available market potential (i.e., the maximum number of users) after accounting for the static presence of competing transport options (e.g., transit, private cars). In the Menu Design model (Chapter 4), competition is introduced explicitly as a single competing alternative within the user's choice set. However, competition among shared mobility providers (e.g., two ride-hailing services) introduces complex, dynamic interactions that significantly affect operational outcomes. The interaction between competing providers can be modeled as a non-cooperative dynamic game. Providers' operational decisions—such as pricing, rebalancing, and fleet deployment—are strategic variables that directly influence a competitor's optimal response. Furthermore, competition directly affects user behavior. The existence of multiple competing options shapes user expectations and defines reference points for service attributes (like price, formed from other platforms). Furthermore, users who are already utilizing a competitor may lead to habitual choices (Gärling and Axhausen, 2003). This established usage can also lead to biases, such as the status quo bias (Kahneman et al., 1991), where the current platform is overvalued simply because it is the existing state. Both the formation of habits and the status quo bias significantly impact the user's choice behavior between different platforms or modes.

- **Leveraging digital nudging and serious games for behavioral demand management.** The proposed behavioral demand management strategy explored through menu design (Chapter 4) can be expanded by integrating principles of digital nudging and gamification. Digital nudging, defined as the use of interface design elements to influence behavior (Weinmann et al., 2016), represents a largely untapped opportunity in shared mobility. Techniques such as social norm framing (e.g., "most chosen"), environmental labeling (e.g., CO_2 impact), or default and framing effects (e.g., highlighting a "green route" or "less congested option") can steer user choices toward sustainable outcomes or better system utilization, often without altering underlying service attributes (Thaler and Sunstein, 2021). Furthermore, incorporating serious games, defined as games primarily designed for purposes other than pure entertainment, offers a powerful means to shift long-term user behavior (Ritterfeld et al., 2009). These gamified elements, such as point systems and badges, reward users for choices that align with provider objectives (e.g., parking in undersupplied areas or sharing rides). For instance, platforms can use progress bars showing accumulated CO_2 savings or employ social comparison mechanisms to accelerate adoption through interpersonal influence. Serious games also enable experimental data collection. Both nudges and serious games can be incorporated into the proposed models as exogenous behavioral interventions that modify existing behavioral parameters, such as price sensitivity in the G-RRM model or the imitation coefficient (q) in the Bass Diffusion model.

- **Joint design of alternatives and choice sets under behavioral effects.** Our findings in Chapter 4 indicate that context effects can serve as a demand management lever by influencing user decisions without requiring changes to the configuration of alternatives, like giving price incentives. However, external incentives may still be necessary to achieve desired outcomes, given the limitations imposed by exogenous factors such as the performance of competing options, users' attribute sensitivities, and the composition of the available choice set.

  Future work could go beyond the isolated design of either pricing/incentives or menu composition and instead explore their joint optimization. Moreover, users exhibit spatial or temporal flexibility (Boyacı and Zografos, 2019), but their willingness to accept delays, detours, or pricing trade-offs can be influenced by the choice set composition. Joint design enables platforms to co-optimize the service attributes (e.g., price, pickup window, walking distance) and the service menu, in order to maximize both acceptance and operational efficiency. Context effect helps avoid unnecessary incentives (over-incentivizing).

- **Incorporating user heterogeneity.** The developed models rely on the assumption of homogeneous user behavior. However, shared mobility users may exhibit a degree of heterogeneity. This heterogeneity stems from the fact that individuals vary in their sociodemographics, travel needs, cognitive abilities, attitudes, and perceptions (Liao and Correia, 2022; Aguilera-García et al., 2020; Faghih-Imani et al., 2017; Najmi et al., 2023; Simonson, 1989), which results in variation in their sensitivity to behavioral effects (Social Influence and Context Effects). This variation across users implies that the Bass Diffusion coefficients (p, q) and acceptable service level used in the adoption model (Chapter 3) and the regret coefficient γ used in the G-RRM choice model (Chapter 4) should be estimated as segment-specific. Future research can integrate this heterogeneity using segment-specific parameters into the optimization frameworks of Chapters 3 and 4, allowing providers to deploy personalized and targeted demand management strategies.
- **Multi-objective formulation.** The current operational models prioritize a unidimensional objective function (i.e., profit maximization). However, shared mobility services operate within a broader context that often requires balancing competing objectives, such as profitability and equity (e.g., ensuring service availability in low-income or underserved areas) or sustainability (e.g., reducing carbon emissions). To address this, the analytical models can be extended to Multi-Objective Optimization problems by formulating multiple objectives. Although computationally demanding, multi-objective optimization offers strategic flexibility. It enables service providers to balance operational efficiency and profitability with key policy and user priorities. Rather than producing a single optimal solution, the model generates a set of Pareto-optimal solutions that represent trade-offs among competing objectives. This Pareto front allows providers to incorporate their strategic goals, while users benefit from services that better reflect their preferences and broader societal objectives.

5.5 Concluding Remark

This dissertation demonstrates that demand management in shared mobility is not only about managing operational aspects such as fleet, pricing, and service quality, but also about accounting for user behavior in operational decisions. From a theoretical angle, this research connects empirically observed behavioral effects, such as social influence and context effects, to operational models of shared mobility, showing how predictable behavioral factors can be systematically incorporated into service operations. From a methodological angle, it develops models that extend existing approaches to long-term service design and short-term menu selection by incorporating behavioral effects and analyzing their implications through scenario-based experiments. From a practical and managerial angle, it provides insights for providers and policymakers on how service design can be aligned with gradual adoption dynamics and how context effects can be used in menu design to guide user choices.

Bibliography

Aguilera-García, Á., Gomez, J., Sobrino, N., 2020. Exploring the adoption of moped scooter-sharing systems in spanish urban areas. Cities 96, 102424.

Albiński, S., Fontaine, P., Minner, S., 2018. Performance analysis of a hybrid bike sharing system: A service-level-based approach under censored demand observations. Transportation Research Part E: Logistics and Transportation Review 116, 59–69.

Bachand-Marleau, J., Lee, B.H., El-Geneidy, A.M., 2012. Better understanding of factors influencing likelihood of using shared bicycle systems and frequency of use. Transportation Research Record 2314, 66–71.

Baddeley, M., 2010. Herding, social influence and economic decision-making: socio-psychological and neuroscientific analyses. Philosophical Transactions of the Royal Society B: Biological Sciences 365, 281–290.

Boyacı, B., Zografos, K.G., 2019. Investigating the effect of temporal and spatial flexibility on the performance of one-way electric carsharing systems. Transportation Research Part B: Methodological 129, 244–272.

Chorus, C.G., Bierlaire, M., 2013. An empirical comparison of travel choice models that capture preferences for compromise alternatives. Transportation 40, 549–562.

Cialdini, R.B., Goldstein, N.J., 2004. Social influence: Compliance and conformity. Annual Review of Psychology 55, 591–621.

Deutsch, M., Gerard, H.B., 1955. A study of normative and informational social influences upon individual judgment. The Journal of Abnormal and Social Psychology 51, 629.

Ding, J., Kim, E.J., Maksimenko, V., Bansal, P., 2025. Can decoy effects nudge ride-hailing drivers' preferences for electric vehicles? Transportation Research Part A: Policy and Practice , 104470.

Faghih-Imani, A., Hampshire, R., Marla, L., Eluru, N., 2017. An empirical analysis of bike sharing usage and rebalancing: Evidence from barcelona and seville. Transportation Research Part A: Policy and Practice 97, 177–191.

Gärling, T., Axhausen, K.W., 2003. Introduction: Habitual travel choice. Transportation 30, 1–11.

Guevara, C.A., Fukushi, M., 2016. Modeling the decoy effect with context-rum models: Diagrammatic analysis and empirical evidence from route choice sp and mode choice rp case studies. Transportation Research Part B: Methodological 93, 318–337.

Hsu, C.L., Lu, H.P., 2004. Why do people play on-line games? an extended tam with social influences and flow experience. Information & Management 41, 853–868.

Huber, J., Payne, J.W., Puto, C., 1982. Adding asymmetrically dominated alternatives: Violations of regularity and the similarity hypothesis. Journal of Consumer Research 9, 90–98.

Jackson, M.O., Yariv, L., 2011. Diffusion, strategic interaction, and social structure, in: Handbook of social economics. Elsevier. volume 1, pp. 645–678.

Kahneman, D., Knetsch, J.L., Thaler, R.H., 1991. Anomalies: The endowment effect, loss aversion, and status quo bias. Journal of Economic Perspectives 5, 193–206.

Larsen, N., Stallrich, J., Sengupta, S., Deng, A., Kohavi, R., Stevens, N.T., 2024. Statistical challenges in online controlled experiments: A review of a/b testing methodology. The American Statistician 78, 135–149.

Liao, F., Correia, G., 2022. Electric carsharing and micromobility: A literature review on their usage pattern, demand, and potential impacts. International Journal of Sustainable Transportation 16, 269–286.

Ma, L., Zhang, X., Hao, F., Liu, S., 2023. Service quality, perceived price fairness, and users' continuous usage intentions regarding shared bike service. Total Quality Management & Business Excellence , 1–21.

Najmi, A., Waller, T., Memarpour, M., Nair, D., Rashidi, T.H., 2023. A human behaviour model and its implications in the transport context. Transportation Research Interdisciplinary Perspectives 18, 100800.

Ritterfeld, U., Cody, M., Vorderer, P., 2009. Serious games: Mechanisms and effects. Routledge.

Simonson, I., 1989. Choice based on reasons: The case of attraction and compromise effects. Journal of Consumer Research 16, 158–174.

Thaler, R.H., Sunstein, C.R., 2021. Nudge: The final edition. Penguin.

Train, K.E., 2009. Discrete choice methods with simulation. Cambridge university press.

Weinmann, M., Schneider, C., Brocke, J.v., 2016. Digital nudging. Business & Information Systems Engineering 58, 433–436.

Zhang, Y., Kamargianni, M., 2023. A review on the factors influencing the adoption of new mobility technologies and services: autonomous vehicle, drone, micromobility and mobility as a service. Transport Reviews 43, 407–429.

Appendix

A Appendices of Chapter 4

A.1 Proof of Proposition 1

Let the systematic utility of alternative i be

$$U_i^{\boldsymbol{\omega}} = \sum_{m=1}^{M} \omega_m x_{im},$$

with attribute weights $\boldsymbol{\omega}$. The systematic regret of i in $\boldsymbol{\mathcal{S}}$ is

$$R_i^{\boldsymbol{\beta}}(\boldsymbol{\mathcal{S}}) = \sum_{j\in\boldsymbol{\mathcal{S}}, j\neq i} \sum_{m=1}^{M} \Big[\ln\big(\gamma + e^{\beta_m(x_{jm}-x_{im})}\big) - \ln(1+\gamma) \Big],$$

with parameters $\boldsymbol{\beta}$ and γ.

For $\gamma = 0$, the regret simplifies to

$$R_i^{\boldsymbol{\beta}}(\boldsymbol{\mathcal{S}}) = -(|\boldsymbol{\mathcal{S}}|+1) \sum_{m=1}^{M} \beta_m x_{im} + \sum_{j\in\boldsymbol{\mathcal{S}}} \sum_{m=1}^{M} \beta_m x_{jm} + \sum_{m=1}^{M} \beta_m x_{cm},$$

where c is a fixed competitor. The regret of c is

$$R_c^{\boldsymbol{\beta}}(\boldsymbol{\mathcal{S}}) = -(|\boldsymbol{\mathcal{S}}|+1) \sum_{m=1}^{M} \beta_m x_{cm} + \sum_{j\in\boldsymbol{\mathcal{S}}, j\neq c} \sum_{m=1}^{M} \beta_m x_{jm}.$$

The G-RRM choice probability is

$$\pi_i(\boldsymbol{\mathcal{S}}) = \frac{\exp(-R_i^{\boldsymbol{\beta}}(\boldsymbol{\mathcal{S}}))}{\exp(-R_c^{\boldsymbol{\beta}}(\boldsymbol{\mathcal{S}})) + \sum_{k\in\boldsymbol{\mathcal{S}}} \exp(-R_k^{\boldsymbol{\beta}}(\boldsymbol{\mathcal{S}}))},$$

and the MNL probability is

$$\mathcal{P}_i(\boldsymbol{\mathcal{S}}) = \frac{\exp(U_i^{\boldsymbol{\omega}})}{\exp(U_c^{\boldsymbol{\omega}}) + \sum_{k\in\boldsymbol{\mathcal{S}}} \exp(U_k^{\boldsymbol{\omega}})}.$$

Substituting the simplified regret yields

$$\pi_i(\boldsymbol{\mathcal{S}})\big|_{\gamma=0,\boldsymbol{\beta}} = \frac{\exp\Big(U_i^{(|\boldsymbol{\mathcal{S}}|+1)\boldsymbol{\beta}}\Big)}{\exp\Big(U_c^{(|\boldsymbol{\mathcal{S}}|+1)\boldsymbol{\beta}}\Big) + \sum_{k\in\boldsymbol{\mathcal{S}}} \exp\Big(U_k^{(|\boldsymbol{\mathcal{S}}|+1)\boldsymbol{\beta}}\Big)}.$$

The MNL probability under $\boldsymbol{\omega} = (|\boldsymbol{\mathcal{S}}|+1)\boldsymbol{\beta}$ is identical. Hence, the equivalence holds.

A.2 Proof of Proposition 2

In G-RRM, the regret of alternative i in $\mathcal{S}$ is determined by pairwise comparisons with all $j \in \mathcal{S}, j \neq i$. Enlarging the set to $\mathcal{T}$ introduces additional comparisons with k, modifying $R_i(\mathcal{T})$ for all $i \in \mathcal{S}$.

When k is asymmetrically dominated by i but not by a stronger competitor $j \in \mathcal{S}$, the marginal reduction in regret is greater for i than for j. Consequently,

$$\pi_i(\mathcal{T}) > \pi_i(\mathcal{S}),$$

which establishes a violation of regularity.

A.3 Proof of Proposition 3

In G-RRM, the regret of alternative i is determined by pairwise comparisons with all $j \in \mathcal{S}, j \neq i$. Enlarging the set to $\mathcal{T}$ introduces additional comparisons with k, yielding new terms in $R_i(\mathcal{T})$ and $R_j(\mathcal{T})$ that depend on $x_{km} - x_{im}$ and $x_{km} - x_{jm}$ across attributes m.

When these increments differ, the relative change in regret is unequal for i and j, leading to

$$\frac{\pi_i(\mathcal{T})}{\pi_j(\mathcal{T})} \neq \frac{\pi_i(\mathcal{S})}{\pi_j(\mathcal{S})}.$$

This establishes the violation of IIA.

A.4 Proof of Theorem 4.1

Let $f(\mathcal{S})$ denote the revenue of assortment $\mathcal{S}$. If two items i, k satisfy $r_i > r_k$ and $\mathcal{S}$ contains k but not i, replacing k with i weakly increases the numerator of $f(\mathcal{S})$ while leaving the denominator unchanged. Iterating such exchanges implies that any optimal assortment must be defined by a revenue threshold: if i is excluded while k is included, then necessarily $r_i \leq r_k$.

Sorting items by r_i produces the nested family $A_0 \subset A_1 \subset \cdots \subset A_{|\mathcal{U}|}$, where A_k contains the k highest-revenue items. The optimal assortment is one of these prefixes, so it suffices to evaluate them sequentially. The no-purchase option c only shifts the threshold $\tau^\star = f(\mathcal{S}^\star)$ upward but does not alter the revenue-ordered structure. Thus the optimum is obtained by sorting once and scanning candidate prefixes, an $O(N \log N)$ procedure.

A.5 Proof of Theorem 4.2

We prove NP-hardness by reduction from *Densest k-Subgraph* (DkS). Fix $\gamma > 0$ and $\beta > 0$, and normalize

$$\pi_i(\mathcal{S};\gamma) \;\propto\; \exp\Big(- R_i(\mathcal{S};\gamma)/T(\mathcal{S})\Big), \qquad T(\mathcal{S}) = \max\{|\mathcal{S}| - 1, 1\}.$$

Define $\phi(x) = \ln(\gamma + e^{\beta x}) - \ln(\gamma + 1)$. Then ϕ is strictly increasing, convex, and satisfies

$\phi(B) + \phi(-B) > 0$ for all $B > 0$.

For each edge $e = \{u, v\} \in E$, introduce three coordinates (m_e^0, m_e^+, m_e^-) with attributes

$$x_{c,m_e^0} = B, \quad x_{u,m_e^+} = x_{v,m_e^+} = B, \quad x_{u,m_e^-} = B, \; x_{v,m_e^-} = -B,$$

and all other entries set to 0. By direct substitution, the contribution of e to $R_c(\{c\} \cup \mathcal{S}; \gamma)$ is 0 if at most one endpoint lies in $\mathcal{S}$, and 2α if both do, with $\alpha = \frac{1}{2}(\phi(B) + \phi(-B)) > 0$.

To eliminate degree leakage in R_c, for each vertex $i \in V$ add a singleton coordinate $s(i)$ defined by

$$x_{c,s(i)} = T, \quad x_{i,s(i)} = 0, \quad x_{j,s(i)} = 0 \;\; (j \neq i),$$

where $T < 0$ satisfies $\phi(-B) + 2\phi(B) + \phi(-T) = 0$. With this adjustment,

$$R_c(\{c\} \cup \mathcal{S}; \gamma) = C_k + 2\alpha\, |E[\mathcal{S}]|,$$

for a constant C_k depending only on k and fixed parameters.

To enforce symmetry among offered items, for each pair $\{i, j\} \subseteq V$ introduce a coordinate $t\{i, j\}$ with

$$x_{c,t\{i,j\}} = S, \quad x_{i,t\{i,j\}} = x_{j,t\{i,j\}} = S,$$

and 0 otherwise. This yields identical R_i for all $i \in \mathcal{S}$ of fixed size k. Consequently,

$$\pi_c(\{c\} \cup \mathcal{S}; \gamma) = \frac{1}{1 + A_k \exp(R_c(\{c\} \cup \mathcal{S}; \gamma)/T(\mathcal{S}))},$$

with $A_k = k\, \exp(-D_k/T(\mathcal{S}))$, and hence $f(\mathcal{S}; \gamma) = 1 - \pi_c$ is strictly increasing in R_c.

Combining these relations gives

$$f(\mathcal{S}; \gamma) = \varphi(|E[\mathcal{S}]|)$$

for a strictly increasing $\varphi(\cdot)$. Thus maximizing revenue over assortments of size k is equivalent to maximizing $|E[\mathcal{S}]|$. For a DkS instance (G, k, L), construct the above G-RRM instance and set $\tau = \varphi(L)$. Then

$$|E[\mathcal{S}]| \geq L \quad \Longleftrightarrow \quad f(\mathcal{S}; \gamma) \geq \tau,$$

which establishes a polynomial-time reduction.

The construction uses $O(|E| + |V|^2)$ attributes with polynomial bit-length constants, and $f(\mathcal{S}; \gamma)$ is polynomially computable. The reduction is therefore polynomial time, proving NP-hardness of the decision problem and, consequently, of the optimization problem. $\square$

A.6 Context-Aware Policy (CAP)

Algorithm 2 provides the formal specification of CAP, including an optional cardinality constraint. The procedure evaluates each singleton initialization, applies best-improvement local search over the 1-exchange neighborhood, and returns the best final solution.

Algorithm 2 Context-Aware Policy (CAP)

Input: item set $\mathcal{U}$, revenues $\{r_i\}$, regret parameters (β, γ), optional size limit k. **Output:** best assortment $\overline{\mathcal{S}}$ and value $\overline{Z}$.

```
S̄ ← ∅, Z̄ ← −∞. for i ∈ U do
    S ← {i}; compute Z_S. repeat
        // Enumerate all single-item neighbors
        Best ← S; Z_Best ← Z_S. if no size limit or |S| < k then
            for ℓ ∈ U \ S do
                compute Z_{S∪{ℓ}}; if Z_{S∪{ℓ}} > Z_Best then
                    Best ← S ∪ {ℓ}; Z_Best ← Z_{S∪{ℓ}}.
        for ℓ ∈ S do
            compute Z_{S\{ℓ}}; if Z_{S\{ℓ}} > Z_Best then
                Best ← S \ {ℓ}; Z_Best ← Z_{S\{ℓ}}.
        if Z_Best > Z_S then
            S ← Best; Z_S ← Z_Best;
        else
            break;
    until no improving 1-exchange move exists
    if Z_S > Z̄ then
        S̄ ← S; Z̄ ← Z_S.
return S̄, Z̄.
```

Incremental-update variants for evaluating $Z_{\mathcal{S}}(\gamma)$ can be implemented, but do not affect the policy definition.

A.7 Supplementary Estimation Results

Table 1 reports maximum likelihood estimates of MNL parameters when data are generated under the G-RRM model with $\gamma = 1$. Reported estimates are significant at the 10% level.

Table 1: MNL estimates under G-RRM data-generating process ($\gamma = 1$).

# Att.	*True G-RRM Parameters*				*Estimated MNL Parameters*			
(M)	β_1	β_2	β_3	β_4	ω_1	ω_2	ω_3	ω_4
2	-1	-1	*	*	-2.13	-2.19	*	*
	-2	-1	*	*	-3.60	-2.32	*	*
	-1	-2	*	*	-2.25	-3.81	*	*
	-0.5	-0.5	*	*	-1.11	-1.11	*	*
	-0.25	-0.25	*	*	-0.56	-0.55	*	*
3	-1	-1	-1	*	-2.39	-2.36	-2.37	*
	-2	-1	-1	*	-4.02	-2.44	-2.43	*
	-1	-2	-1	*	-2.45	-4.08	-2.38	*
	-0.5	-0.5	-0.5	*	-1.22	-1.21	-1.21	*
	-0.25	-0.25	-0.25	*	-0.60	-0.60	-0.60	*
4	-1	-1	-1	-1	-2.69	-2.57	-2.56	-2.58
	-2	-1	-1	-1	-4.51	-2.64	-2.57	-2.67
	-1	-2	-1	-1	-2.72	-4.43	-2.57	-2.56
	-0.5	-0.5	-0.5	-0.5	-1.38	-1.35	-1.35	-1.36
	-0.25	-0.25	-0.25	-0.25	-0.68	-0.68	-0.68	-0.69

A.8 Supplementary Results: MNL vs. G - RRM Assortments

This appendix provides detailed tables comparing the optimal assortments under the Multinomial Logit (MNL) model and the Generalized Random Regret Minimization (G - RRM) model. For each dimensionality of the attribute space ($M = 2, 3, 4$), we report results across different preference parameterizations ($\boldsymbol{\beta}$) and competitor profiles (Prof).

Table 2: Optimal assortments under MNL vs. G - RRM for $M = 2$ attributes, summarized by preference parameters ($\boldsymbol{\beta}$) and competitor profiles (Prof).

		Agreement			**Change**	**Excl.**		**Impact**	
Pref. ($\boldsymbol{\beta}$)	Prof (tag)	Id. (%)	Overlap (Jac,%)	$\lvert\Delta S\rvert$ (#)	Expand (%)	Top-Rev Excl. (%)	Top-Share Excl. (%)	ΔRev (%)	ΔShare (%)
(-1,-2)	AS	1	55.5	3.44	18.2	9.1	10.1	-0.25	-1.28
	BAL	5	60.5	2.65	1.1	21.1	25.3	7.84	2.72
	HP	0	14.0	11.3	0.0	51.0	76.0	52.9	-12.5
	PS	25	72.1	1.97	16.0	24.0	26.7	14.3	9.56
	UP	10	67.2	1.18	100	0.0	0.0	-0.08	-0.01
(-2,-1)	AS	1	74.8	2.07	66.7	0.0	0.0	-0.90	9.92
	BAL	11	68.3	3.24	24.7	3.4	7.9	-5.59	2.74
	HP	0	10.8	13.4	0.0	67.0	75.0	6.20	-49.4
	PS	4	67.3	3.21	0.0	62.5	65.6	29.3	15.1
	UP	35	76.4	1.31	100	0.0	0.0	-0.82	-0.41
(-0.5,-0.5)	AS	10	66.7	2.28	2.2	17.8	21.1	-2.70	-7.56
	BAL	5	61.6	2.60	0.0	24.2	34.7	-2.54	-8.93
	HP	0	25.1	7.28	0.0	63.0	96.0	-10.7	-32.2
	PS	25	70.4	2.64	0.0	26.7	42.7	2.75	-3.06
	UP	37	83.3	1.17	98.4	0.0	0.0	0.80	2.54
(-025,-0.25)	AS	61	90.2	1.41	0.0	0.0	25.6	-1.81	-4.11
	BAL	40	84.8	1.47	0.0	0.0	26.7	-2.22	-5.25
	HP	0	44.1	4.10	0.0	0.0	63.0	-9.00	-20.5
	PS	52	87.4	1.52	0.0	0.0	25.0	-0.79	-3.17
	UP	81	95.3	1.21	100	0.0	0.0	1.35	2.10

Table 3: Optimal assortments under MNL vs. G - RRM for $M = 3$ attributes, summarized by preference parameters ($\boldsymbol{\beta}$) and competitor profiles (Prof).

		Agreement			**Change**	**Excl.**		**Impact**	
Pref. ($\boldsymbol{\beta}$)	Prof (tag)	Id. (%)	Overlap (Jac,%)	$\lvert\Delta S\rvert$ (#)	Expand (%)	Top-Rev Excl. (%)	Top-Share Excl. (%)	ΔRev (%)	ΔShare (%)
(-2,-1,-1)	AS	2	75.3	2.37	69.4	1.0	1.0	5.53	13.2
	BAL	15	72.5	3.15	28.2	9.4	11.8	-0.16	2.85
	HP	0	9.65	14.0	0.0	63.0	67.0	13.0	-46.3
	PS	9	70.3	3.07	0.0	74.7	78.0	39.2	16.3
	UP	29	73.4	1.35	100	0.0	0.0	-1.04	-0.62
(-1,-2,-1)	AS	0	46.5	5.18	33.0	16.0	25.0	2.83	1.22
	BAL	1	56.8	3.06	1.0	13.1	20.2	11.3	6.07
	HP	0	8.92	13.5	0.0	91.0	97.0	72.9	-50.6
	PS	23	69.8	2.13	14.3	28.6	31.2	15.2	8.38
	UP	11	66.0	1.24	100	0.0	0.0	-0.13	-0.04
(-0.5,-0.5,-0.5)	AS	0	52.3	3.49	0.0	34.0	37.0	-4.20	-13.2
	BAL	1	56.0	3.09	0.0	23.2	29.3	-2.60	-9.57
	HP	0	19.0	8.56	0.0	78.0	96.0	-1.73	-32.1
	PS	24	69.7	2.78	0.0	31.6	40.8	3.76	-3.02
	UP	25	79.2	1.21	100	0.0	0.0	-0.19	1.87
(-0.25,-0.25,-0.25)	AS	9	69.3	2.08	0.0	16.5	27.5	-3.52	-9.17
	BAL	19	76.3	1.79	0.0	21.0	29.6	-2.70	-7.52
	HP	0	36.8	5.15	0.0	47.0	87.0	-8.97	-25.4
	PS	51	86.8	1.61	0.0	16.3	26.5	-0.36	-3.02
	UP	78	94.7	1.18	100	0.0	0.0	0.79	1.61

Table 4: Optimal assortments under MNL vs. G-RRM for $M = 4$ attributes, summarized by preference parameters ($\boldsymbol{\beta}$) and competitor profiles (Prof).

		Agreement			**Change**	**Excl.**		**Impact**	
Pref. ($\boldsymbol{\beta}$)	Prof (tag)	Id. (%)	Overlap (Jac,%)	$\lvert\Delta S\rvert$ (#)	Expand (%)	Top-Rev Excl. (%)	Top-Share Excl. (%)	ΔRev (%)	ΔShare (%)
(-1,-2,-1,-1)	AS	0	37.6	6.09	16.0	37.0	45.0	8.03	-7.01
	BAL	1	58.3	3.02	2.0	22.2	26.3	12.6	5.39
	HP	0	8.24	14.3	0.0	89.0	98.0	170	-41.5
	PS	18	67.9	2.05	30.5	22.0	25.6	12.8	6.75
	UP	10	65.3	1.19	100	0.0	0.0	-0.16	-0.06
(-2,-1,-1,-1)	AS	1	69.2	3.13	50.5	10.1	10.1	6.81	11.8
	BAL	5	66.1	3.48	20.0	19.0	19.0	1.76	0.90
	HP	0	8.22	14.7	0.0	70.0	94.0	72.7	-28.9
	PS	2	65.3	3.24	0.0	81.6	82.7	50.1	16.4
	UP	12	67.9	1.19	100	0.0	0.0	-0.90	-0.66
(-0.5,-0.5,-0.5,-0.5)	AS	0	48.4	4.06	0.0	45.0	55.0	-1.39	-13.2
	BAL	1	55.5	3.20	0.0	28.3	34.3	-0.11	-8.09
	HP	0	18.9	9.07	0.0	73.0	96.0	18.8	-24.1
	PS	26	71.4	2.61	17.6	32.4	43.2	4.72	-1.84
	UP	21	75.6	1.28	100	0.0	0.0	-1.68	0.06
(-0.25,-0.25,-0.25,-0.25)	AS	1	61.6	2.54	0.0	26.3	38.4	-1.58	-9.70
	BAL	12	70.2	2.15	0.0	33.0	45.5	-1.00	-8.04
	HP	0	33.5	5.94	0.0	50.0	97.0	-2.15	-23.7
	PS	52	88.1	1.46	8.3	14.6	27.1	0.99	-1.12
	UP	59	90.9	1.07	100	0.0	0.0	-0.43	0.63

About the author

Mahsa Farhani was born in 1992 in Tehran, Iran, and pursued her undergraduate studies in Industrial Engineering at Sharif University of Technology, Iran. After receiving her bachelor's degree in 2014, she was awarded the Exceptional Talents Fellowship to continue with a master's program at the same university. During her master's studies, she focused on game-theoretic mathematical models for optimizing the location of distribution hubs in competitive markets. Alongside her academic pursuits, she acquired professional experience as a research and development specialist and consultant within the logistics industry, enabling her to translate research insights into practical application. Following the completion of her master's degree and professional engagements in Iran, Mahsa joined Delft University of Technology in the Netherlands to pursue doctoral research in Transport and Logistics, a path that culminated in the present dissertation.

During her PhD research, Mahsa focused on shared mobility service design and demand management while explicitly modeling customer responses to supply decisions, as humans play a central role in the success of any system and service. Her work lies at the intersection of behavioral economics and operations management, with a particular focus on understanding how individual choices influence system performance and on applying these insights to improve service operations.

TRAIL Thesis Series

The following list contains the most recent dissertations in the TRAIL Thesis Series, For a complete overview of more than 250 titles see the TRAIL website: www.rsTRAIL.nl.

The TRAIL Thesis Series is a series of the Netherlands TRAIL Research School on transport, infrastructure and logistics.

Farhani, M., *Demand Management Strategies for Operations of Shared Mobility Services*, February 2026, TRAIL Thesis Series, the Netherlands

Yao, X., *Driving Heterogeneity in Traffic Flow Theory: An action-based framework for identification, modelling, and simulation*, January 2026, TRAIL Thesis Series, the Netherlands

Versluis, N.D., *Optimising Railway Traffic Management under Radio-Based Distance-to-Go Signalling*, January 2026, TRAIL Thesis Series, the Netherlands

Jiao, Y., *Proactive Collision Risk Quantification in Multi-directional Traffic Interactions*, January 2026, TRAIL Thesis Series, the Netherlands

Asadi, M., *Accessibility and Road Safety: Integration of road safety in accessibility evaluation*, November 2025, TRAIL Thesis Series, the Netherlands

Akse, R., *Understanding and untangling the uncertainty knot: How to catalyse decision-making in mobility innovations*, November 2025, TRAIL Thesis Series, the Netherlands

Führer, K., *Participatory Decision-making under Deep Uncertainty: Modelling mobility transitions*, November 2025, TRAIL Thesis Series, the Netherlands

Picco, A., *Monitoring and Feedback in Driving*, T2025/17, October 2025, TRAIL Thesis Series, the Netherlands

Cebeci, M.S., *Behaviour of Prosumers in Last-mile Logistics: The case of crowdshipping*, T2025/16, September 2025, TRAIL Thesis Series, the Netherlands

Kuijpers, A., *Enabling Inter-Organizational Collaboration Through Platforms: The role of trust*, T2025/15, September 2025, TRAIL Thesis Series, the Netherlands

Song, R., *Human-MASS Interaction in Decision-Making for Safety and Efficiency in Mixed*

Waterborne Transport Systems, T2025/14, June 2025, TRAIL Thesis Series, the Netherlands

Destyanto, A.R., *A Method for Evaluating Port Resilience in an Archipelago*, T2025/13, June 2025, TRAIL Thesis Series, the Netherlands

Karademir, C., *Synchronized Two-echelon Routing Problems: Exact and approximate methods for multimodal city logistics*, T2025/12, May 2025, TRAIL Thesis Series, the Netherlands

Vial, A., *Eyes in Motion: A new traffic sensing paradigm for pedestrians and cyclists*, T2025/11, May 2025, TRAIL Thesis Series, the Netherlands

Chen, Q., *Towards Mechanical Intelligence in Soft Robotics: Model-based design of mechanically intelligent structures*, T2025/10, April 2025, TRAIL Thesis Series, the Netherlands

Eftekhar, Z., *Exploring the Spatial and Temporal Patterns in Travel Demand: A data-driven approach*, T2025/9, June 2025, TRAIL Thesis Series, the Netherlands

Reddy, N., *Human Driving Behavior when Interacting with Automated Vehicles and the Implications on Traffic Efficiency*, T2025/8, May 2025, TRAIL Thesis Series, the Netherlands

Durand, A., *Lost in Digitalisation? Navigating public transport in the digital era*, T2025/7, May 2025, TRAIL Thesis Series, the Netherlands

Dong, Y., Safe, *Efficient, and Socially Compliant Automated Driving in Mixed Traffic: Sensing, Anomaly Detection, Planning and Control*, T2025/6, May 2025, TRAIL Thesis Series, the Netherlands

Droffelaar, I.S. van, *Simulation-optimization for Fugitive Interception*, T2025/5, May 2025, TRAIL Thesis Series, the Netherlands

Fan, Q., *Fleet Management Optimisation for Ride-hailing Services: from mixed traffic to fully automated environments*, T2025/4, April 2025, TRAIL Thesis Series, the Netherlands

Hagen, L. van der, *Machine Learning for Time Slot Management in Grocery Delivery*, T2025/3, March 2025, TRAIL Thesis Series, the Netherlands

Schilt, I.M. van, *Reconstructing Illicit Supply Chains with Sparse Data: a simulation approach*, T2025/2, January 2025, TRAIL Thesis Series, the Netherlands

Ruijter, A.J.F. de, *Two-Sided Dynamics in Ridesourcing Markets*, T2025/1, January 2025, TRAIL Thesis Series, the Netherlands

Fang, P., *Development of an Effective Modelling Method for the Local Mechanical Analysis of Submarine Power Cables*, T2024/17, December 2024, TRAIL Thesis Series, the Netherlands

Zattoni Scroccaro, P., *Inverse Optimization Theory and Applications to Routing Problems*,

T2024/16, October 2024, TRAIL Thesis Series, the Netherlands

Kapousizis, G., *Smart Connected Bicycles: User acceptance and experience, willingness to pay and road safety implications*, T2024/15, November 2024, TRAIL Thesis Series, the Netherlands

Lyu, X., *Collaboration for Resilient and Decarbonized Maritime and Port Operations*, T2024/14, November 2024, TRAIL Thesis Series, the Netherlands

TRAIL

Summary

This dissertation examines how shared mobility providers can manage demand by aligning operational decisions with user behavior. It highlights behavioral factors that generate systematic and predictable demand responses and embeds them into operational models as demand management levers. By integrating social influence into phased deployment and context effects into choice set design, it shows that high-quality, gradual expansion improves profitability and stabilizes adoption, while behaviorally informed menus strengthen competitive performance.

About the Author

Mahsa Farhani, born in Iran, earned bachelor's and master's degrees in Industrial Engineering from Sharif University of Technology. She completed her PhD in Transport and Logistics at Delft University of Technology, studying behaviorally informed demand management for shared mobility services.

TRAIL Research School ISBN 978-90-5584-379-4

Erasmus University Rotterdam
Radboud University
rijksuniversiteit groningen
TU Delft Delft University of Technology
UNIVERSITY OF TWENTE.
TU/e Technische Universiteit Eindhoven University of Technology