UNIVERSITY OF TWENTE

MASTER THESIS

Explaining the variation in the demand for roundtrip B2C carsharing in neighbourhoods in the G44 cities in the Netherlands

An explanation based on neighbourhood characteristics and the current distribution of shared cars

J.A. VOSSEBELD
FEBRUARY 2, 2022

A thesis submitted in fulfillment of the requirements for the degree of Master of Science in

Civil Engineering & Management
Faculty of Engineering Technology

Author:
J.A. VOSSEBELD
s1719114
February 2, 2022

Educational Institution:
University of Twente
Drienerlolaan 5
7522 NB Enschede
The Netherlands

Host Institution:
Goudappel Groep
Snipperlingsdijk 4
7417 BJ Deventer
The Netherlands

Internal Supervisors:
dr. T. THOMAS
prof. dr. ing K.T. Geurs

External Supervisors:
R. TER AVEST
T. TIMMERMANS

Preface

This thesis report, titled: "*Explaining the variation in the demand for roundtrip B2C carsharing in neighbourhoods in the G44 cities in the Netherlands*", marks the end of my time as a Civil Engineering student at the University of Twente. I am looking back at a time that has allowed me to develop myself, both professionally and personally, thanks to the many opportunities that crossed my path. I am very grateful for the amazing people I have met along the way, that made this an unforgettable journey and inspired me to take on every challenge I came across.

I would like to thank dr. Tom Thomas for his extensive feedback throughout the whole process of conducting this research. The clear input and valuable discussions during each feedback session aloud me to continuously improve the quality of my work. I would also like to thank prof. dr. ing. Karst Geurs for sharing his knowledge on how to extract the right information from a survey and his relevant feedback on the major milestones in my research.

Moreover, I would like to express my gratitude towards my supervisors at Goudappel. Firstly, I want to thank Richard ter Avest for his endless stream of ideas that could contribute to my research, for providing me with the opportunity to visit the NVC and of course for being available during my green light meeting. I also want to thank Erik Timmermans for his profound feedback and providing me with the tools and guidance in setting up a survey. The drive and enthusiasm of both men guided me towards a successful completion of this thesis. I am grateful for working together with them.

Finally, I would like to thank my girlfriend, friends and family for supporting me during this research.

I hope you will enjoy reading this thesis report!

Aron Vossebeld
Amsterdam, February 2022

Management Summary

The space in the urban areas of the Netherlands is limited, while simultaneously, there exists an increasing demand to live in this urban space. When carsharing would be implemented on a larger scale, fewer cars are needed and cars will be more in use. This could decrease the space needed for parking, which contributes to the increase in space in the urban areas to live in, without decreasing the inhabitants' mobility options. Currently, carsharing providers are doing extensive market research to identify the neighbourhoods, districts or municipalities with the most potential demand for carsharing. However, literature shows that a large number of factors that affect the demand for carsharing are factors that can be expressed by the characteristics of an area. When the areas with the most potential demand for carsharing could be identified based on only the characteristics of an area, the time and money spent on market research would be reduced. This could help increase the use of shared cars and decrease the parking space in urban areas.

To better understand the effect of the characteristics of an area on the demand for carsharing, the goal of this research is to reduce the need for extensive market research for the identification of an area with a high demand for carsharing, by developing a model that can predict the demand for carsharing in a neighbourhood based on its characteristics. To do this, the following research question is established: *How are neighbourhood characteristics explaining the variation in the supply of roundtrip Business to Consumer (B2C) shared cars in neighbourhoods in the G44 cities in the Netherlands and can this explanation of variation be used to explain the variation in the demand for roundtrip B2C carsharing in these neighbourhoods?*. B2C means that a fleet of cars is owned by a business and the cars are rented out to users and roundtrip means that users must return the car at the same place or zone as they started using it. The G44 is a collaboration between the 44 largest cities in the Netherlands.

In order to answer the research question, a conceptual model was drafted in which all factors that have been described in previous literature as factors that affect the demand for and supply of shared cars are framed. This conceptual model consists of individual and neighbourhood factors and is used to answer four sub-questions. To answer the first sub-question, two regression models are developed that explain the variation in the supply of roundtrip B2C shared cars in neighbourhoods in the G44 cities in the Netherlands based on the characteristics of the neighbourhoods that are framed in the conceptual model as neighbourhood factors. The first regression model is based on the neighbourhoods in Almere, Arnhem, Enschede, Nijmegen, Zoetermeer and Zwolle because together, these cities are a good representation of the G44 cities. The second regression model is based on the neighbourhoods in The Hague, Amsterdam and Utrecht because these cities have the highest number of B2C shared cars per 100,000 inhabitants and using their neighbourhoods, therefore, results in data with a higher variation. Both models consist of a binary logistic regression and a negative binomial regression. The binary logistic regression predicts whether there is a roundtrip B2C shared car present in a neighbourhood. Then the neighbourhoods that have a predicted presence of a shared car are the input for the negative binomial regression that is used to explain the variation in the number of shared cars per 100,000 inhabitants in a neighbourhood. This is also referred to as the shared car supply rate in a neighbourhood. Comparing the results of both regression models showed that the regression model based on The Hague, Amsterdam and Utrecht is best able to explain the variation in the supply of roundtrip B2C shared cars in the neighbourhoods of the 6 municipalities that together represent the G44 cities.

To answer the other sub-questions, a survey is distributed over the 6 municipalities that together represent the G44 cities. The survey is used to gather data about the demand for roundtrip B2C carsharing, and all individual and neighbourhood factors that could affect the demand for carsharing or the supply of shared cars that are framed in the conceptual model. The second research question has been answered by comparing the demand for carsharing of the inhabitants of the G44 cities with the observed supply of shared cars in their neighbourhood. This research shows that the inhabitants that live in a neighbourhood with a roundtrip B2C shared car show an almost identical level of demand for carsharing as individuals without a roundtrip B2C shared car in their neighbourhood. Therefore, there is no indication of an effect of the observed supply of roundtrip B2C shared cars in a neighbourhood on the demand for roundtrip B2C carsharing. This means that the supply of shared cars does not have to be taken into account when explaining the variation in the demand for roundtrip B2C carsharing.

The answer to the third sub-question is used to find out whether the explanation of the variation in the supply of roundtrip B2C shared cars in neighbourhoods in the G44 cities in the Netherlands can be used to explain the variation in the demand for roundtrip B2C carsharing in these neighbourhoods as well. This is done by using three different methods that were used to find the relationship between the neighbourhood characteristics that affect the demand for roundtrip B2C carsharing and the neighbourhood characteristics affecting the supply of roundtrip B2C shared cars. The stronger this relationship, the better the explanation of the variation in the supply of shared cars can be used to explain the variation in the demand for carsharing. The results of the three methods show that there is a significant relationship between the neighbourhood characteristics that affect the demand for roundtrip B2C carsharing and the neighbourhood characteristics that affect the supply of roundtrip B2C shared cars. Also, it is shown that the inhabitants of the neighbourhoods with a high predicted shared car supply rate by the explanation of the variation in the supply of roundtrip B2C shared cars have a relatively high demand for roundtrip B2C carsharing as well. However, the explanation of the variation in the supply of shared cars can not be used to predict the exact demand in a neighbourhood in the G44 cities.

This research also provides insight into the reason why a model based on only neighbourhood characteristics cannot exactly explain the variation in the demand for carsharing in neighbourhoods in the G44 cities by analysing what factors explain the variation in the individual demand for roundtrip B2C carsharing of inhabitants of the G44 cities in the Netherlands. This analysis is the answer to the fourth sub-question and is also using the individual factors that affect the demand for carsharing or the supply of shared cars framed in the conceptual model. These individual factors consist of factors based on experience with shared mobility, attitudes, and motives and purposes for carsharing. The results of the analysis show that the experience with a shared car and experience with other shared mobility appear to be the most important characteristics of an individual to show demand for roundtrip B2C carsharing. Also, individuals that travel more often with a non-car form of transport are more likely to show demand for roundtrip B2C carsharing. Therefore, to obtain a better explanation of the variation in the demand for roundtrip B2C carsharing in neighbourhoods in the G44 cities, these individual factors are needed. Besides these individual factors, the analysis also confirms the importance of a large number of neighbourhood characteristics in the explanation of the variation in the demand for roundtrip B2C carsharing.

A limitation of the explanation of the variation in the demand for roundtrip B2C carsharing developed in this research is that it can only be used to identify neighbourhoods with a high demand for roundtrip B2C carsharing and is not able to distinguish between neighbourhoods with a comparable predicted shared car supply rate. Hence, it is recommended to only use the model when it is desired to identify a neighbourhood, district or municipality with the most demand for roundtrip B2C carsharing without doing extensive research or to use this research to gain knowledge on the factors that affect the supply of and demand for roundtrip B2C carsharing. For further research, it is recommended to find a better indicator for the modal split in a neighbourhood to be able to improve the model in this research, and to do further research on how to increase the awareness of the presence of shared cars in a neighbourhood. Because when inhabitants do not know about the presence of a shared car in their neighbourhood, they will never consider using it.

Dutch Management Summary

De ruimte in het stedelijke gebied van Nederland is beperkt, terwijl er tegelijkertijd een toenemende vraag is om in deze stedelijke ruimte te wonen. Als autodelen op grotere schaal zou worden ingevoerd, zijn er minder auto's nodig en zullen auto's meer in gebruik zijn. Hierdoor kan de benodigde ruimte voor parkeren afnemen, wat bijdraagt aan een toename van de ruimte voor huizen in het stedelijke gebied, zonder dat dit ten koste gaat van de mobiliteitsopties van de inwoners. Momenteel doen deelautoaanbieders uitgebreid marktonderzoek om de buurten, wijken of gemeenten met de meeste potentiële vraag naar autodelen te identificeren. Uit de literatuur blijkt echter dat een groot aantal factoren die van invloed zijn op de vraag naar autodelen, factoren zijn die kunnen worden uitgedrukt in de karakteristieken van een gebied. Wanneer de gebieden met de meeste potentiële vraag naar autodelen kunnen worden geïdentificeerd op basis van alleen de karakteristieken van een gebied, zou de tijd en het geld dat aan marktonderzoek wordt besteed kunnen worden verminderd. Dit kan helpen om het gebruik van deelauto's te vergroten, waardoor er minder parkeerruimte nodig is in het stedelijke gebied.

Om het effect van de karakteristieken van een gebied op de vraag naar autodelen beter te begrijpen, is het doel van dit onderzoek om de behoefte aan uitgebreid marktonderzoek voor het identificeren van een gebied met een grote vraag naar autodelen te verkleinen, door een model te ontwikkelen die de vraag naar autodelen in een buurt kan voorspellen op basis van de buurtkarakteristieken. Om dit te doen is de volgende onderzoeksvraag opgesteld: *Hoe verklaren buurtkarakteristieken de variatie in het aanbod van roundtrip Business to Consumer (B2C) deelauto's in buurten in de G44 steden in Nederland en kan deze verklaring van de variatie worden gebruikt om de variatie in de vraag naar roundtrip B2C autodelen in deze buurten te verklaren?*. B2C betekent dat een wagenpark met deelauto's eigendom is van een bedrijf en de auto's worden verhuurd aan gebruikers en roundtrip betekent dat gebruikers de auto moeten inleveren op dezelfde plaats of zone als dat ze hem zijn gaan gebruiken. De G44 is een samenwerking tussen de 44 grootste steden van Nederland.

Om de onderzoeksvraag te beantwoorden is een conceptueel model opgesteld waarin alle factoren zijn vermeld die in bestaande literatuur zijn beschreven als factoren die de vraag naar en het aanbod van deelauto's beïnvloeden. Dit conceptueel model bestaat uit individuele factoren en buurtfactoren en wordt gebruikt om vier deelvragen te beantwoorden. Om de eerste deelvraag te beantwoorden zijn twee regressiemodellen ontwikkeld die de variatie in het aanbod van roundtrip B2C deelauto's in buurten in de G44 steden in Nederland verklaren op basis van de karakteristieken van de buurten die in het conceptueel model als buurtfactoren zijn gegroepeerd. Het eerste regressiemodel is gebaseerd op de buurten in Almere, Arnhem, Enschede, Nijmegen, Zoetermeer en Zwolle omdat deze steden samen een goede afspiegeling zijn van de G44 steden. Het tweede regressiemodel is gebaseerd op de buurten in Den Haag, Amsterdam en Utrecht omdat deze steden het hoogste aantal B2C deelauto's per 100.000 inwoners hebben en het gebruik van deze buurten dus resulteert in data met een grotere variatie. Beide modellen bestaan uit een binaire logistische regressie (binary logistic regression) en een negatieve binominale regressie (negative binomial regression). De binaire logistische regressie voorspelt of er een roundtrip B2C deelauto aanwezig is in een buurt. De buurten met een voorspelde aanwezigheid van een roundtrip B2C deelauto zijn de input voor de negatieve binominale regressie die wordt gebruikt om de variatie in het aantal deelauto's per 100.000 inwoners in een buurt te verklaren. Dit wordt ook wel de 'shared car supply rate' in een buurt genoemd. Vergelijking van de resultaten van beide regressiemodellen laat zien dat het regressiemodel op basis van Den Haag, Amsterdam en Utrecht het beste de variatie in het aanbod van roundtrip B2C deelauto's kan verklaren in de buurten van de 6 gemeenten die samen de G44 steden vertegenwoordigen.

Om de overige deelvragen te beantwoorden is een enquête verspreid over de 6 gemeenten die samen de G44 steden vertegenwoordigen. De enquête is gebruikt om data te verzamelen over de vraag naar roundtrip B2C autodelen en over alle individuele en buurtfactoren die van invloed kunnen zijn op de vraag naar autodelen of op het aanbod van deelauto's, die zijn vermeld in het conceptueel model. De tweede onderzoeksvraag is beantwoord door de vraag van de inwoners van de G44 steden te vergelijken met het waargenomen aanbod van deelauto's in hun buurt. Uit dit onderzoek blijkt dat de inwoners die in een buurt wonen met een roundtrip B2C deelauto een bijna identieke vraag hebben als individuen zonder een roundtrip B2C deelauto in hun buurt. Er zijn dan ook geen aanwijzingen voor een effect van het aanbod van roundtrip B2C deelauto's in een buurt op de vraag naar roundtrip B2C autodelen. Dit betekent dat bij het verklaren van de variatie in de vraag naar roundtrip B2C autodelen geen rekening hoeft te worden gehouden met het aanbod van deelauto's.

Het antwoord op de derde deelvraag wordt gebruikt om na te gaan of de verklaring van de variatie in het aanbod van roundtrip B2C deelauto's in buurten in de G44 steden in Nederland kan worden gebruikt om ook de variatie in de vraag naar roundtrip B2C autodelen in deze buurten te verklaren. Dit is gedaan door met behulp van drie verschillende methoden de relatie te vinden tussen de buurtkarakteristieken die van invloed zijn op de vraag naar roundtrip B2C autodelen, en de buurtkarakteristieken die van invloed zijn op het aanbod van roundtrip B2C deelauto's. Hoe sterker deze relatie, hoe beter de verklaring van de variatie in het aanbod kan worden gebruikt om de variatie in de vraag te verklaren. De resultaten van de drie methoden laten zien dat er een significante relatie is tussen de buurtkarakteristieken die van invloed zijn op de vraag naar roundtrip B2C autodelen en de buurtkarakteristieken die van invloed zijn op het aanbod van roundtrip B2C deelauto's. Ook blijkt dat de inwoners van de buurten met een hoog voorspeld deelautoaanbod door de verklaring van de variatie in het aanbod van roundtrip B2C deelauto's een relatief grote vraag naar roundtrip B2C deelauto's hebben. De verklaring van de variatie in het aanbod van deelauto's kan echter niet worden gebruikt om de exacte vraag in een buurt in de G44 steden te voorspellen.

Dit onderzoek geeft ook inzicht in de reden waarom een model gebaseerd op alleen buurtkarakteristieken de variatie in de vraag naar autodelen in buurten in de G44 steden niet precies kan verklaren door te analyseren welke factoren de variatie in de individuele vraag naar roundtrip B2C autodelen van inwoners van de G44 steden in Nederland verklaren. Deze analyse is het antwoord op de vierde deelvraag en maakt ook gebruik van de individuele factoren die de vraag naar autodelen of het aanbod van deelauto's beïnvloeden uit het conceptueel model. Deze individuele factoren bestaan uit factoren die gebaseerd zijn op ervaring met deelmobiliteit, attitudes, en motieven en doelen voor autodelen. De resultaten van de analyse laten zien dat de ervaring met een deelauto en ervaring met andere vormen van deelmobiliteit de belangrijkste karakteristieken van een individu met vraag naar roundtrip B2C autodelen blijken te zijn. Ook hebben personen die vaker met een ander vervoermiddel dan de auto reizen, vaker vraag naar roundtrip B2C autodelen. Daarom zijn deze individuele factoren nodig om een betere verklaring te geven voor de variatie in de vraag naar roundtrip B2C autodelen in buurten in de G44-steden. Naast deze individuele factoren bevestigt de analyse ook het belang van een groot aantal buurtkarakteristieken in de verklaring van de variatie in de vraag naar roundtrip B2C autodelen.

Een beperking van de verklaring van de variatie in de vraag naar roundtrip B2C autodelen uit dit onderzoek, is dat deze alleen gebruikt kan worden om buurten te identificeren met een hoge vraag naar roundtrip B2C autodelen en geen onderscheid kan maken tussen buurten met een vergelijkbaar voorspeld aanbod van deelauto's. Het is daarom aan te raden om het model alleen te gebruiken wanneer het gewenst is om te indentificeren in welke buurt, wijk of gemeente de meeste vraag naar roundtrip B2C autodelen is, zonder uitgebreid onderzoek te doen, of om dit onderzoek te gebruiken om kennis te vergaren over de factoren die het aanbod van deelauto's en de vraag naar roundtrip B2C autodelen verklaren. Voor verder onderzoek is het aan te raden om een betere indicator voor de modal split in een buurt te vinden om het model in dit onderzoek te kunnen verbeteren, en om nader onderzoek te doen naar hoe de bekendheid van de aanwezigheid van deelauto's in een buurt kan worden vergroot. Want als inwoners niet op de hoogte zijn van de aanwezigheid van een deelauto, zullen ze er nooit over nadenken om deze te gebruiken.

Contents

List of Figures

List of Tables

List of Abbreviations

AIC	**A**kaike **I**nformation **C**riterion
Avg.	**A**verage
B2C	**B**usiness to **(2) C**onsumer
BIC	**B**ayesian **I**nformation **C**riterion
P2P	**P**eer to **(2) P**eer
MCAR	**M**issing **C**ompletely **A**t **R**andom
N	**N**umber
VIF	**V**ariable **I**nflation **Factors**

Chapter 1

Introduction

1.1 Problem statement

The Netherlands wants to build one million houses before 2030 (Aedes, 2021). Meanwhile, the country encounters urbanization. From 2009 to 2019, the percentages of people that live in cities increased from 86% to 92% (O'Neill, 2020). This means that a large part of the demand for houses will be in urban areas and that a large part of the one million extra houses, must be built in these urban areas. However, the space in urban areas is limited.

A potential terrain to gain space for the construction of houses is the parking space. One parking place accounts for at least 12.5 m^2 (TU Delft, 2016). Therefore, decreasing the number of parking places in urban areas would result in a significant amount of extra space. However, the number of cars per inhabitant in the Netherlands is growing and has increased from 528 cars per 1,000 inhabitants in 2015, to 543 cars per 1,000 inhabitants in 2020 (CBS, 2020a). This means that an increasing number of inhabitants own a car and are parking this car near their house. Hence, it would be beneficial to implement carsharing on a larger scale. The idea behind carsharing is that the usage of a car is shared by members for trip making on a per-trip basis (Ferrero, Perboli, Rosano, & Vesco, 2018). When implemented on a large scale, carsharing has the potential to cause a decrease in private car ownership since the supply and demand for cars could be adjusted to people's needs. This ensures that the vehicles will be more in use and therefore, will be parked less. In this way, a situation is created where less parking space will be needed without decreasing mobility options.

It is known that the municipalities with the highest number of shared cars per inhabitant and the strongest carsharing growth are all highly urbanized (CROW, 2021b). However, the potential shared car use in a smaller area like a neighbourhood is not only based on the urban density. Also, other factors like the average income and the percentage of highly educated individuals are important when identifying the most suitable neighbourhood of a city to place a shared car. Currently, carsharing providers are doing extensive market research to identify the neighbourhoods, districts or municipalities with the most potential demand for carsharing (Van der Molen, 2021). However, literature shows that a large number of factors that affect the demand for carsharing are factors that can be expressed by the characteristics of a neighbourhood. When neighbourhoods with a high potential demand for carsharing could be identified based on only these neighbourhood characteristics, an interesting opportunity for municipalities and carsharing providers arises since the time and money spent on market research would be reduced. This could help increase the use of shared cars and decrease the parking space in urban areas. However, it is unknown whether a good representation of the demand for carsharing in a neighbourhood could be predicted by only neighbourhood characteristics and it is unknown which neighbourhood characteristics are important in this prediction. To better understand the effect of neighbourhood characteristics on the demand for carsharing, the goal of this research is to reduce the need for extensive market research for the identification of a neighbourhood with a high demand for carsharing, by developing a model that is able to predict the demand for carsharing in a neighbourhood based on its characteristics.

Although there exists a lot of literature on which factors are affecting the demand for carsharing, most existing literature does not distinguish between different business models and organizational structures of carsharing and the different carsharing adopters these business models and organizational structures might attract. The different business models and organizational structures have important differences in use and impact and may well prosper in different neighbourhoods (Münzel, Boon, Frenken, Blomme, & van der Linden, 2020). Therefore, the prediction of the demand for carsharing in a neighbourhood based on neighbourhood characteristics depends on which business model and which organizational structure of carsharing is focused on. This research focuses on roundtrip Business-to-Consumer (B2C) shared cars. B2C means that a fleet of cars is owned by a business and the cars are rented out to users and roundtrip means that users must return the car at the same place or zone as they started using it (CROW, 2021b).

It would be preferable to use the characteristics of the neighbourhoods with the highest number of booked shared cars to predict the extent to which each neighbourhood characteristic plays a role in the demand for shared car usage in these neighbourhoods. This is because the number of bookings could give the most precise measure of shared car use in each neighbourhood since it would correct for individuals who are not regularly making use of carsharing. Unfortunately, a complete database with active member numbers or bookings of each carsharing provider is not publicly available and the will of carsharing providers to share their data is limited.

Data on the supply of shared cars per neighbourhood, however, is available on the website of shared car providers. But it is unknown whether the distribution of the supply of shared cars can be used to be able to draw conclusions on the demand for carsharing in a neighbourhood because the size of the relation between demand and supply and the way the demand and supply are affecting each other is unknown. Therefore, this knowledge gap must be eliminated to be able to use the distribution of the supply of shared cars, to predict the demand for carsharing.

In summary, an interesting opportunity for municipalities and carsharing providers arises when neighbourhoods with a high potential demand for carsharing could be identified based on only neighbourhood characteristics. However, it is unknown whether a good representation of the demand for carsharing in a neighbourhood could be predicted by only neighbourhood characteristics and could be based on the distribution of the shared car supply. It is also unknown which neighbourhood characteristics are important in this prediction. Since different business models and organizational structures of carsharing will prosper in different neighbourhoods and therefore need distinct studies, the prediction of the demand for carsharing, in general, is not specific enough. Hence, this research focuses on the demand for roundtrip B2C carsharing and answers the following research question:

How are neighbourhood characteristics explaining the variation in the supply of roundtrip B2C shared cars in neighbourhoods in the G44 cities in the Netherlands and can this explanation of variation be used to explain the variation in the demand for roundtrip B2C carsharing in these neighbourhoods?

The market for carsharing is more developed in larger cities in the Netherlands and therefore, more data bout carsharing is available in the (medium) large cities. The focus of this research is on neighbourhoods in the cities that are part of the G44. The G44 is a collaboration between the 44 largest Dutch cities (CBS, 2022).

To be able to answer the research question, a regression model is developed that explains the variation in the supply of roundtrip B2C shared cars in neighbourhoods in the G44 cities in the Netherlands. Then it is analysed whether this regression model can be used to explain the variation in the demand for roundtrip B2C carsharing in neighbourhoods in the G44 cities as well by making use of data from a survey.

The study is executed in collaboration with Goudappel. Goudappel is a label of Goudappel Groep, a cooperation of companies that solves mobility issues by improving the living environment and contributing to a sustainable, vital and attractive society. Multiple municipalities have indicated to Goudappel that they are interested in exploring the potential for carsharing in their municipality.

1.2 Thesis outline

In Chapter 2, the theoretical framework is discussed in which all factors that have been described in previous literature as factors that affect the demand for and supply of shared cars are framed in a conceptual model. Also, the different business models and organizational structures of carsharing are elaborated on and the current demand for and supply of carsharing in the Netherlands are described and visualized in this chapter. Based on the theoretical framework, research questions and corresponding hypotheses are derived in Chapter 3. In this chapter, also the scope of this research is elaborated on. Chapter 4 describes the methodology of this study. The research approach is discussed and also the methods used to explain the variation in the supply of roundtrip B2C shared cars in neighbourhoods in the G44 cities is elaborated on. Then also the structure of the survey and the analysis methods used to find the relation between the demand for and supply of shared cars are discussed. In Chapter 5, the distribution of the shared car supply over the neighbourhoods in The Hague, Amsterdam and Utrecht is analysed and the representativeness of the survey is elaborated on. Chapter 6 describes the results of the explanation of the variation in the supply of roundtrip B2C shared cars in neighbourhoods in the G44 cities and Chapter 7 shows the effect of the observed supply of roundtrip B2C shared cars on the demand for roundtrip B2C carsharing. The results of analysing the relationship between the neighbourhood characteristics affecting the demand for carsharing and the neighbourhood characteristics affecting the supply of roundtrip B2C shared cars are presented in Chapter 8. Then it is known to what extent the explanation of the variation in the supply of roundtrip B2C shared cars based on only neighbourhood characteristics can be used to explain the variation in the demand for roundtrip B2C carsharing. Subsequently, in Chapter 9 the explanation of the variation in the individual demand for roundtrip B2C shared cars based on individual factors and neighbourhood characteristics is discussed. This results in the individual factors that could improve the model. This thesis then draws conclusions on the results in Chapter 10 and finalizes this report by discussing the results and limitations of the work and providing recommendations for further research in Chapter 11.

Chapter 2

Theoretical framework

To find whether the demand for roundtrip B2C carsharing can be predicted based on neighbourhood characteristics and the current supply of shared cars, all factors that affect the demand for carsharing and the supply of shared cars are important. Therefore, the main aim of this theoretical framework is to frame all factors that have been described in previous literature as factors that affect the demand for and supply of shared cars in a conceptual model. The conceptual model is shown in Figure 2.1 and works as a point of departure for this research. The demand for and supply of carsharing are the dependent variables and the different factors that affect the demand and supply are the independent variables in the model. All independent variables will be discussed independently in this chapter.

There is not much literature on factors affecting only the supply of shared cars. This could be because of a correlation between factors affecting the demand for carsharing and factors affecting the supply of shared cars. A lead to evidence for this correlation can be found in the work of Münzel, Piscicelli, Boon, and Frenken (2019) in which the authors discuss a large number of factors that are elaborated on in other literature as factors affecting the demand for carsharing, as factors affecting the supply for carsharing. This could mean that the factors affecting the demand for carsharing are also somehow affecting the supply of shared cars. However, there is no previous literature that explains the size of this correlation and the way the demand and supply are affecting each other. That there is a presumable coherence between the carsharing demand and supply is presented in the conceptual model in Figure 2.1 by the connection between the two. Because of this coherence, the independent variables that directly affect the carsharing supply are also independent variables that affect the carsharing demand and the other way around.

In the theoretical framework, a distinguishment has been made between neighbourhood factors and individual factors. The difference is that a neighbourhood factor is a factor that explains what neighbourhood characteristics affect the demand for and supply of carsharing and an individual factor is based on the individual choice for carsharing. The neighbourhood factors have been categorized as spatial factors, sociodemographic factors and socioeconomic factors and the individual factors as factors based on the experience with shared mobility, factors based on attitudes and factors based on motives and purposes for carsharing. All categories will be discussed independently for the demand for and supply of carsharing in the same order as they are shown in the conceptual model in Figure 2.1.

After all factors that affect the demand for and supply of carsharing have been discussed, the different business models and organizational structures of carsharing are elaborated on and the current demand for and supply of carsharing in the Netherlands are presented and visualized in this chapter.

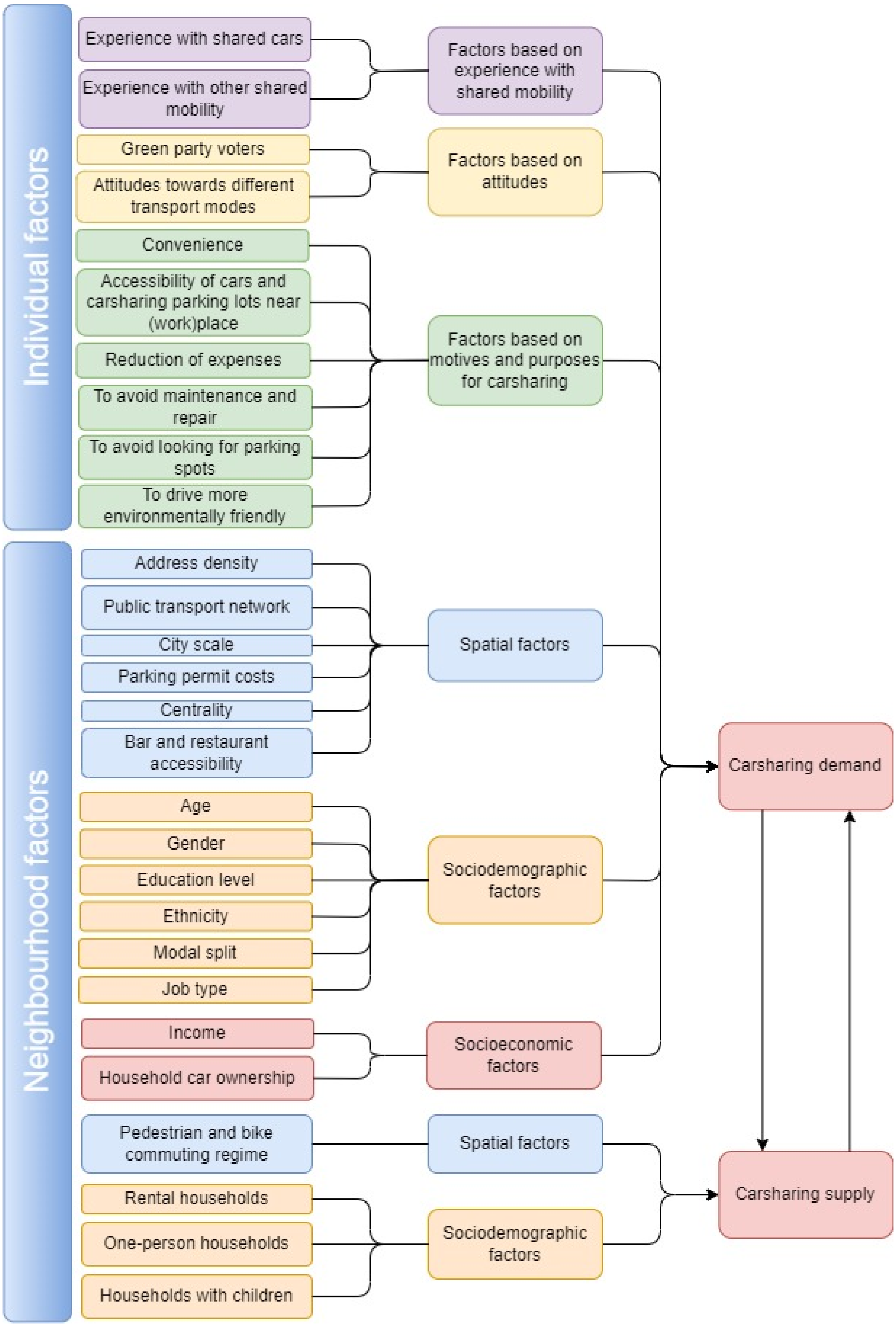

FIGURE 2.1: Conceptual model

2.1 Factors affecting the demand for carsharing

As has been visualized in the conceptual model in Figure 2.1, a large number of factors are affecting the demand for carsharing. These factors are independently discussed in this section, where a distinguishment is made between the individual factors and the neighbourhood factors that affect the demand for carsharing.

2.1.1 Individual factors

First, the individual factors affecting the demand for carsharing are described. These factors are grouped as factors based on the experience with shared mobility, factors based on attitudes and factors based on motives and purposes for carsharing.

2.1.1.1 Factors based on the experience with shared mobility

The demand for carsharing can be influenced by the presence of other types of shared mobility. The presence of bike sharing systems may be beneficial for the demand for carsharing. The first reason for this is that individuals can get familiar with the use of shared transportation and will therefore show more demand for other shared transportation as well. This is known as the spillover effect among innovations (Jaffe, Trajtenberg, & Fogarty, 2000). The second reason is that individuals get less dependent on owning a car if a good bike sharing system is in place already (Münzel et al., 2019). But not only experience with bike sharing systems will increase the individual demand for carsharing, but also experience with other shared mobility as shared scooters or the shared car itself are affecting the demand for carsharing in the same way (Münzel et al., 2019).

2.1.1.2 Factors based on attitudes

Truffer (2003) identified members of environmental groups or others with a strong ideological background as early adopters of carsharing. Münzel et al. (2019) are clarifying this as that people with strong environmental ideologies identify carsharing as a more environmentally friendly transportation mode and act on these ideologies by showing more demand for carsharing. The variable that is included to explain this attitude in their research is the number of green voters in a city. The number of green voters has been applied in the model in Figure 2.1 as an independent variable as well. Lee, Aultman-Hall, Coogan, and Adler (2016) also proved that attitudes related to driving preference are relevant for the choice of sharing or not sharing a ride. An individual's attitude towards more sustainable transport modes such as public transport, cycling or walking increases the likelihood of making use of shared transportation and are therefore also included in the conceptual model.

2.1.1.3 Factors based on motives and purposes for carsharing

Previous studies suggested that there are also several motivations an individual can have to show demand for carsharing. According to Lane (2005), the convenience of having a car only when necessary is the most important reason for joining carsharing. Ramos, Bergstad, Chicco, and Diana (2020) measured the motives of current users of carsharing and distinguished between different motives. One of them is convenience as well. Others are the reduction of mobility expenses, to drive more environmentally friendly, to have more comfort when travelling, to avoid responsibilities with maintenance and repair of a private car, to avoid looking for parking spots and the accessibility of cars and carsharing parking lots near the respondent's place or workplace. The convenience of having a car only when needed and avoiding responsibilities with maintenance and repair turned out to be the most important motives for using a shared car. The accessibility of cars and carsharing pick-up locations near an individual's place or workplace was considered to be important as well. Wang, Dane, and Timmermans (2020) are confirming this by stating that carsharing users prefer a shared-car accessibility distance of fewer than 5 minutes for convenience. According to the CROW (2021d), shared cars are mainly used for occasional trips that are often relatively long. Carsharing is less attractive for short journeys and also for commuting, the shared car is rarely used. Nevertheless, it is used for occasional business trips. The most common purposes to make use of a shared car are to visit a friend or a relative and for a day out (CROW, 2021d).

2.1.2 Neighbourhood factors

Most factors in Figure 2.1 that affect the demand for carsharing are neighbourhood factors. These factors originate from spatial, sociodemographic and socioeconomic factors and are discussed in this section.

2.1.2.1 Spatial factors

Factors that regard the location of a collective are referred to as spatial factors. The municipalities with the highest number of shared cars per inhabitant and the strongest growth in carsharing are all highly urbanized (CROW, 2021b). Wang et al. (2020) investigated the preference of residents in these densely populated urban areas in the Netherlands for carsharing-facilitating neighbourhoods and found some interesting factors that affect the demand for carsharing in these neighbourhoods. Wang et al. (2020) are characterizing the urbanization of a city as city scale. They state that residents in neighbourhoods in large cities are more likely to face parking pressure problems and that therefore these neighbourhoods are more likely to show carsharing potential. Besides, the advanced public transport network in large cities make these residents less dependent on private cars, which creates a positive effect on the demand for carsharing (Shaheen & Cohen, 2007). The distance to public transport facilities is, therefore, an important factor affecting the demand for carsharing. Opposing to this, carsharing could also be useful in neighbourhoods in cities with a weaker public transit system, since a car is needed more frequently to get to a location (Münzel et al., 2020). Celsor and Millard-Ball (2007) found several additional spatial neighbourhood factors necessary for carsharing to succeed. One of them is the population density or address density of a neighbourhood. A high population density brings a large customer base within walking distance of each carsharing location. Doubling the population density doubles the potential customers and the residents in these neighbourhoods have lower rates of vehicle ownership and trips as well, causing even more demand for shared car usage. Wang et al. (2020) indicate that also the high costs of a parking permit in large cities is a reason for residents to show a higher demand for carsharing. Müller, Correia, and Bogenberger (2017) state that other important determinants that result in a high number of carsharing bookings are the area's centrality and the bar and restaurant accessibility. The centrality has been defined as the average distance of multiple points in an area to the centre in this area and the bar and restaurant accessibility as the number of bars and restaurants that can be reached within one kilometre.

2.1.2.2 Sociodemographic factors

Sociodemographic factors are the key drivers of mobility patterns and travel modes and determine the diffusion of carsharing services in the urban population (Prieto, Baltas, & Stan, 2017). In the study of Wang et al. (2020) sociodemographic factors in terms of gender, age, education, job type and income are factors that affect the carsharing preference of a neighbourhood. Wang et al. (2020) conclude that with increasing age, the utility of carsharing-neighbourhoods decreases monotonically. This is in line with research by Becker, Ciari, and Axhausen (2017), in which the researchers conclude that young individuals are usually the adopters of new products and innovations, such as carsharing. Wang et al. (2020) state that respondents with a part-time job are more likely to make use of carsharing than respondents with a full-time job because full-time workers have to travel to work daily, and private cars are more convenient for daily use. Part-time workers are commuting on a more occasional basis, suiting more to the use of a shared car. A different study by Wang, Dane, and Timmermans (2021), shows the role of gender to estimate the preference to live in carsharing-facilitating neighbourhoods. Their research shows that the group that is interested in carsharing has a male to female ratio of 41:59, whereas this ratio in the group that is not interested is 53:47. This means that a neighbourhood consisting of a relatively high number of females is more likely to be interested in carsharing. Other research showed that highly-educated individuals are more likely to make use of carsharing (Ciari, Weis, & Balac, 2016). This is in line with the research of Wang et al. (2021), where the group of individuals with higher education levels that are interested in living in a carsharing-facilitating neighbourhood is significantly larger (70,4%) than the group of lower educated individuals (60,4%). Also, ethnicity could affect the demand for carsharing. A Northern American study showed that some minorities, typically Hispanics and African Americans, carpool more than other racial and ethnic groups (Shaheen, 2018). Another factor affecting the demand for carsharing is the modal split in the neighbourhood. Residents in carsharing neighbourhoods are far more likely than their regional counterparts to walk or take public transport rather than drive, to work. Research by Celsor and Millard-Ball (2007) shows that 55% of residents in a neighbourhood with a low level of carsharing service are driving alone to work, where this percentage is only 35-40% in a neighbourhood with a high level of carsharing service. Also, a more recent study by Le Vine, Zolfaghari, and Polak (2014) confirms that carsharing users are relatively heavy users of non-car forms of urban transport such as public transport, walking and cycling, showing the impact of the modal split of a neighbourhood on the carsharing demand in the neighbourhood.

2.1.2.3 Socioeconomic factors

Also, socioeconomic factors affect the demand for carsharing. People with lower incomes are more likely to have a higher demand for carsharing, lending support to the hypothesis that carsharing is an attractive alternative for poorer people who find car ownership too expensive (Abraham, 1999). That carsharing is a good option for households without a car is confirmed by the research of Münzel et al. (2019) and Le Vine et al. (2014) in which it is shown that carsharing users often live in carless or single-car households.

2.2 Factors affecting the supply of shared cars

In this section, only factors that have not been discussed as factors affecting the carsharing demand are discussed. These are all factors of which literature shows a direct relationship with the supply of shared cars. Since there are no individual factors, all discussed factors in this section are neighbourhood factors.

2.2.1 Spatial factors

Münzel et al. (2020) found that a high supply of carsharing is correlated with strong pedestrian and bike commuting regimes. Research by Celsor and Millard-Ball (2007) is confirming this by showing that only 5% of the residents that live in a neighbourhood with a low level of carsharing service is walking to work, where this percentage is 15-20% in a neighbourhood with a high level of carsharing service.

2.2.2 Sociodemographic factors

As the carsharing supply increases, so does the proportion of rental households and one-person households. At the same time, the proportion of households with children decreases (Celsor & Millard-Ball, 2007). The study of Celsor and Millard-Ball (2007) also confirmed their hypothesis by showing that the correlation between the carsharing supply and the proportion of rental households and one-person households is 0,301 and 0,478 respectively. The correlation between the carsharing supply and the proportion of households with children is -0,412.

2.3 Business models and organizational structures

Carsharing can be organized in multiple business models and organizational structures. The two most common business models are Business-to-Consumer (B2C) and Peer-to-Peer (P2P). B2C means that a fleet of cars is owned by a business and the cars are rented out to users. In a P2P business model, individuals rent out their own cars to other consumers through a two-sided platform operated by a coordinating carsharing organisation (Shaheen, Cohen, & Zohdy, 2016). Within these business models, multiple organizational structures exist. The most common structures in the Netherlands have been defined in the dashboard on carsharing of the CROW (CROW, 2021b). These organizational structures are roundtrip/station-based, free-floating, local communities, business and P2P. This means that the business model P2P has its own organizational structure in the dashboard (CROW, 2021b). B2C carsharing can be organized in three ways: roundtrip, free-floating and for business use. The organizational structure 'local communities' is not covered by the B2C or P2P business models. An overview of the relations between all business models and organizational structures is shown in Figure 2.2. All organizational structures will be discussed more extensively in the remainder of this section. An overview of all carsharing providers and platforms per organizational structure in the Netherlands can be found in Appendix A.

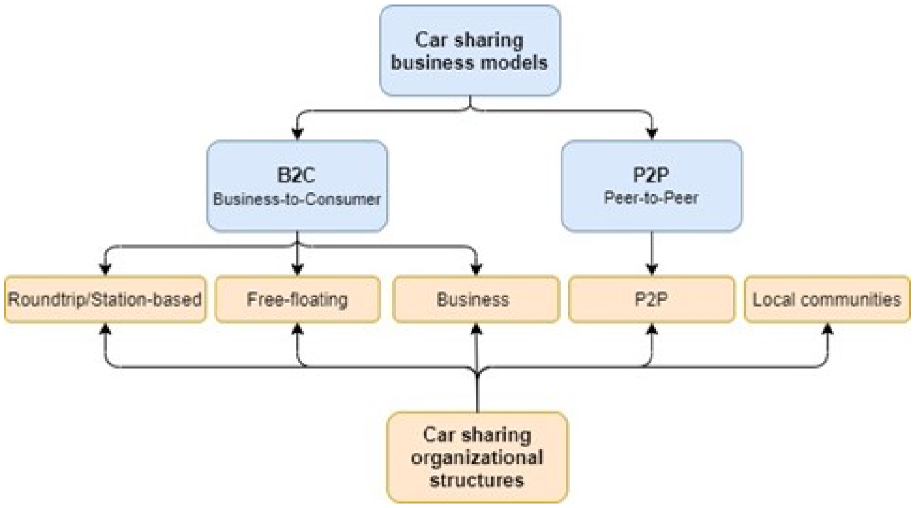

FIGURE 2.2: Relations between all carsharing business models and organizational structures

The demand for carsharing of an individual is dependent on the business model that is used. The individual choice between B2C and P2P shared cars could be affected by two factors: the car type and costs of the shared car. B2C shared cars are often part of a large fleet of cars that are likely to be newer than the national average car. Since newer vehicles have better fuel efficiency and emit fewer pollutants than comparable older models, this contributes to better urban air quality and less consumed fuel (FleetCarma, 2018). Therefore, an individual may consider B2C carsharing as more environmental friendly than P2P carsharing and shows more demand for it. However, P2P could be considered as a cheaper alternative because someone else's vehicle is mostly older and therefore cheaper (Oots, 2018).

2.3.1 Roundtrip/station-based carsharing

Roundtrip carsharing, also called station-based carsharing, is an organizational structure of carsharing in which users must return the car at the same place or zone as they started using it (CROW, 2021b). Initially, all B2C carsharing was based on a roundtrip organizational structure (Münzel et al., 2019). Within roundtrip carsharing, a difference is made between fixed parking places and zonal parking spaces. When the parking place is fixed, the user has to return the car to the same parking place from which he collected it. Cars with a zonal parking space do not have their own parking place, but can be returned in a marked zone (MyWheels, 2021). The fixed parking place in a roundtrip organized structure gives the user the guarantee that a parking place is always available when returning to the starting location to end the use. The demand for roundtrip carsharing is increasing because of this guarantee. However, the obligation of using the fixed parking spot limits the destination location flexibility and could therefore also lower the individual demand for roundtrip carsharing (Van den Berg, 2017).

2.3.2 Free-floating carsharing

In 2008, free-floating carsharing was introduced as a new organizational structure of B2C carsharing (Münzel, Boon, Frenken, & Vaskelainen, 2018). In a free-floating organizational structure, cars do not have a fixed parking place. The car can be returned to a different location than where it has been picked up. There are multiple fixed stations where a shared car can be parked and the user can pick up a car at one station and return it at another station (CROW, 2021b). Free-floating carsharing is mainly used for short inner-city trips (Martin & Shaheen, 2016). When carsharing is organized as free-floating, there is a greater level of location flexibility than when it is organized roundtrip. This is because carsharing members are allowed to use a car for only as little as they need the car for. Free-floating carsharing users could also save money on fuel because roundtrip members must return the vehicle to the allocated parking place and may have to drive more (Steiger, 2011). A disadvantage of a free-floating structure can be that sometimes a user may not find a car when they need one because a car is not always returned at the same parking spot (Krztoń, 2018).

2.3.3 Peer-to-Peer (P2P)

P2P is a business model and organizational carsharing structure that appeared around 2010 (Münzel et al., 2019). In this structure, individuals rent out their own cars to others. Open platforms are being used where people can put their car for rent and get a payment for each time their car is being rented (CROW, 2021b). The main providers of P2P carsharing in the Netherlands are SnappCar and MyWheels. To increase the user-friendliness of peer-to-peer carsharing, SnappCar developed the Keyless-formula. With this development users can access the cars with their smartphones, making P2P shared cars accessible 24 hours per day as well (SnappCar, 2021).

2.3.4 Business carsharing

In the organizational structure 'business carsharing' the main target group are businesses. Typical business carsharing providers are car lease companies, but also providers that focus on private users are open for business users as well. Because the business and private markets have different peak hours, these markets are easy to combine (CROW, 2021b). Business carsharing can also be organized with a 'pool car'. A pool car can be described as a lease car, but where a lease car can be used by only one employee, a pool car is accessible for multiple employees. In this way, the car is used more efficiently (MoveMove, 2021). Some providers that offer B2C carsharing are offering their service for business carsharing.

2.3.5 Local communities

The organizational structure 'local communities' is not covered by the business models B2C and P2P, but is an organizational structure that also occurs within the Netherlands. A fixed and closed group of users are sharing one or more cars. In most cases, the shared car has a fixed parking place or street where it must be returned to. Because the users are often living close to each other or are neighbours, there is more social control over the use and maintenance of the car (CROW, 2021b). Although local communities can organize their shared car themselves, there are also platforms such as Stapp.in and Amigo that make it easier for the communities. This organizational structure is different from P2P carsharing because the car is not owned by one individual and can only be used by members of the community.

2.4 Carsharing in the Netherlands

Now all factors affecting the demand for and supply of carsharing have been covered and the different business models and organizational structures have been described, the current state of carsharing in the Netherlands is discussed. First, the supply of shared cars is elaborated on and this is followed by how much and where these cars are being used.

2.4.1 Supply of shared cars in the Netherlands

In the spring of 2021, the Netherlands counted 87,000 shared cars. In 2020, 23,000 cars were added to this total. As can be seen in Figure 2.3, the most shared cars are shared P2P. However, shared cars that are not shared P2P are showing the strongest growth in the total number of shared cars. To provide a better overview of the shared cars that are not shared P2P, their growth is shown in Figure 2.4. This figure shows that the supply of business carsharing is growing most significantly, followed by roundtrip carsharing.

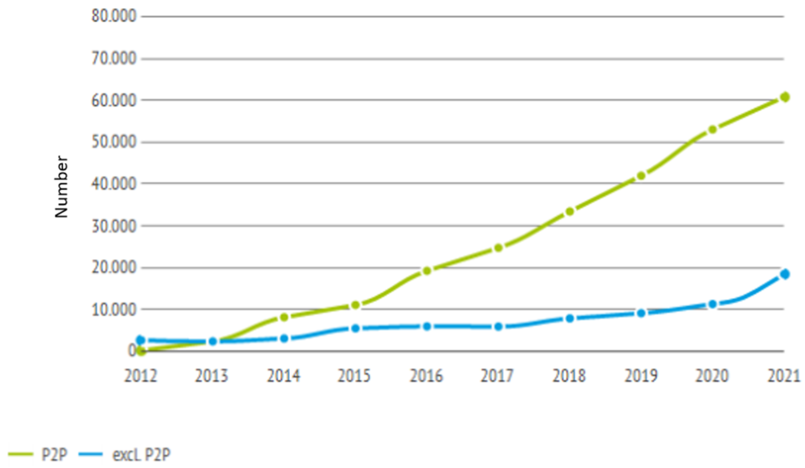

FIGURE 2.3: Number of shared cars in the Netherlands, 2012–2020 (CROW, 2021b)

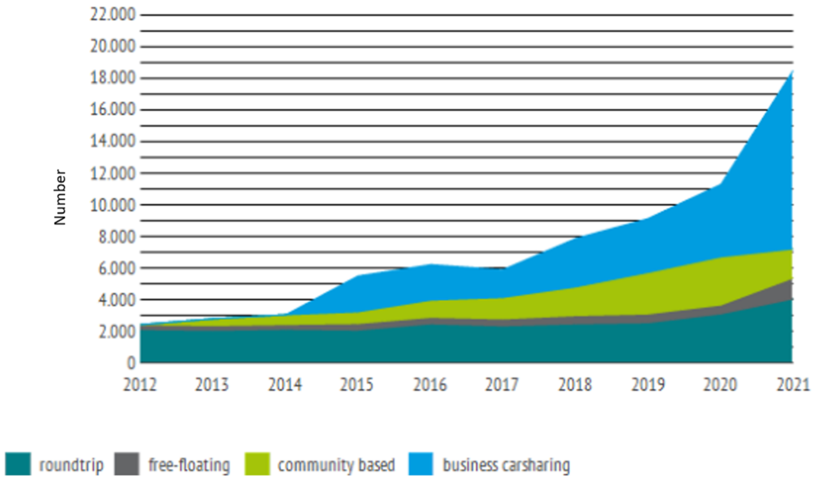

FIGURE 2.4: Number of shared cars in the Netherlands excl. P2P, 2012-2020 (CROW, 2021b)

The largest number and the highest increase of shared cars can be found in Amsterdam. Utrecht is the municipality that has the most shared cars per 100,000 inhabitants with more than 1,400 shared cars per 100,000 inhabitants (De Boer, 2020). In 2021, the Netherlands has 502 shared cars per 100,000 inhabitants. P2P shared cars are mostly responsible for this number. When only roundtrip, free-floating and community based carsharing are considered, it is only 42 shared cars per 100,000 inhabitants (CROW, 2021b).

The top 10 carsharing municipalities are all highly urbanized. Also, the strongest growth of carsharing is in the more urban municipalities, with the strongest growth in the largest four cities: Amsterdam, Rotterdam, The Hague and Utrecht. In highly urbanized areas, 2% of the private car fleet consists of shared cars (CROW, 2021a). Figure 2.5 shows the top 10 cities with the highest number of shared cars per 100,000 inhabitants in the Netherlands.

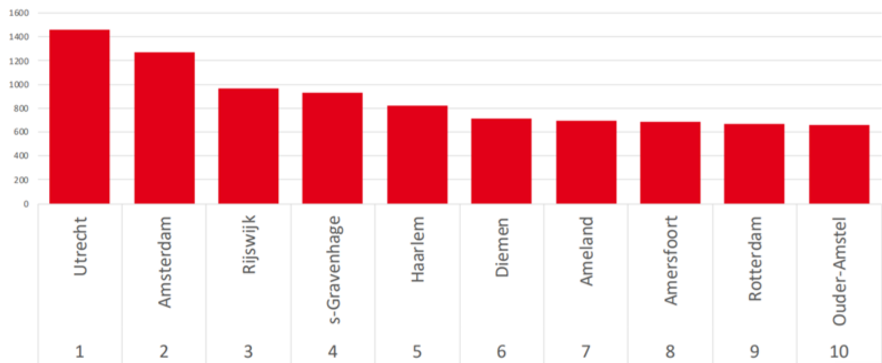

FIGURE 2.5: Top 10 cities in the number of shared cars per 100,000 inhabitants (De Boer, 2020)

2.4.2 Use of shared cars in the Netherlands

The Netherlands counted 970,000 users of shared cars in 2021 (CROW, 2021b). Unfortunately, there is no open data or literature available on the use of shared cars per municipality or lower spatial level as a district or neighbourhood. The number of users per shared car fluctuates strongly per business model and organizational structure. A free-floating organizational structure results in the most users per shared car. An average free-floating shared car has almost 200 users. A roundtrip shared car has an average of 32 users per shared car, which still is a relatively large number. An average P2P shared car has only 8 users per car. Since free-floating and roundtrip carsharing are both organizational structures that are used in B2C carsharing, the number of users per B2C shared car is much higher than per P2P shared car. The highest growth in the number of users per car is seen in free-floating carsharing as well. This growth is shown in Figure 2.6. Figure 2.7 shows the growth in the number of users per car for the other organizational structures. The figure shows that the number of users per car, when not considering free-floating carsharing, is mainly increasing for roundtrip carsharing and business carsharing. Because the highest growth is in free-floating, roundtrip and business carsharing, it can be concluded that the highest growth in the number of users per shared car is in the business model B2C carsharing as well (CROW, 2021b).

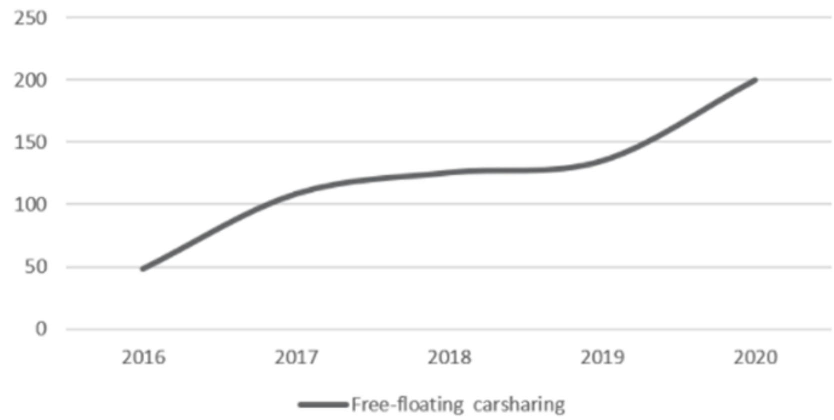

FIGURE 2.6: Number of users per shared car for free-floating carsharing (CROW, 2021b)

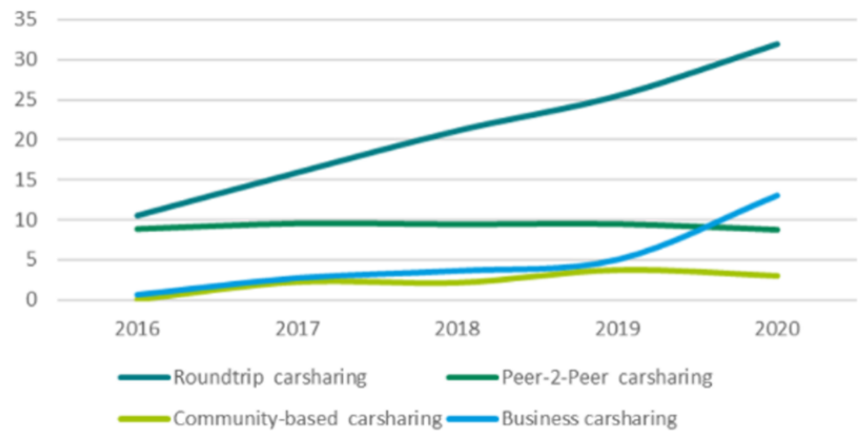

FIGURE 2.7: Number of users per shared car for roundtrip, peer-2-peer, community based and business carsharing (CROW, 2021b)

Although some users of carsharing are driving more kilometres than they did before making use of carsharing, the average user of carsharing is driving fewer kilometres. Before making use of carsharing, the average car sharer owned one car per household. Since making use of shared cars, the average number of cars per household that makes use of carsharing has decreased to 0.7. The average car sharer is driving less as well. Before making use of carsharing, the average number of kilometres per year was 9,100. Since making use of shared cars, this average has dropped to 7,500 kilometres per year, a decline of 17.5% (Autodelen, 2021).

Chapter 3

Scope, hypotheses and research questions

This chapter starts with elaborating on the scope of the research. Then the sub-questions, which are the building blocks of the answer to the main research question, are discussed. Finally, hypotheses are drafted for further guidance of the research.

3.1 Scope

Before the research questions and hypotheses of this research are formulated, the scope is elaborated on by discussing the study area and focus group of the research and the business model that is focused on more elaborately.

3.1.1 Study area and focus group

To be able to predict a detailed location of the areas with the most potential users for carsharing, the focus in this research is on a neighbourhood level. Municipalities in the Netherlands are subdivided into districts and neighbourhoods whereof neighbourhoods make up the lowest regional level (CBS, 2020b). The CBS (2020b) defines a neighbourhood as a part of a municipality that is homogeneously demarcated based on historical or urban features in which homogeneous means that one function is dominant. This definition is also used in this research. The focus of this research is on neighbourhoods in the cities that are part of the G44. The G44 is a collaboration between the 44 largest Dutch cities (CBS, 2022). The focus is on (medium) large cities since the market for carsharing is more developed in these larger cities and therefore, more data about carsharing is available.

3.1.2 Focus on roundtrip B2C carsharing

The most common business models are B2C and P2P carsharing. Since research shows that these different carsharing business models attract different groups, the predictions of the demand in a neighbourhood for both business models desire two distinct studies. The focus of this study is on roundtrip B2C carsharing.

The choice to focus on B2C carsharing originates from multiple reasons. Although previous literature shows that the majority of shared cars in the Netherlands are shared P2P, an average P2P shared car is not used much. An average B2C shared car has a much higher number of users than P2P shared cars. More users mean that there is a greater variation in when the car is used, resulting in a higher total use of one B2C shared car. This means that a B2C shared car has to be parked less and can therefore be considered a better solution to decrease the parking space in the urban areas of the Netherlands than a P2P shared car. Additionally, the supply of P2P shared cars is dependent on residents that would like to share their private car and the supply of B2C shared cars is more dependent on the decisions of municipalities and carsharing providers to invest in the placing of a shared car. Therefore, the prediction of the demand for carsharing to identify potential neighbourhoods for the future placing of shared cars is more relevant for B2C shared cars.

As explained in Section 2.3, the B2C business model consists of three organizational structures: 'roundtrip', 'free-floating' and 'business'. These organizational structures all need distinct studies as well since the intended use is different for all three of them. Business carsharing is mainly dependent on the location of businesses and the policy of a business to make use of carsharing or not. Since these cars are not available for private use and it is clear that a business area has the largest demand for business shared cars, this organizational structure is excluded from this research. The prediction of the demand for roundtrip and free-floating shared cars is different since a 'free-floating' shared car can be used for one-way trips and a 'roundtrip' shared car cannot. This results in a different demand for both organizational structures. Section 2.4.1 shows that a free-floating organizational structure is very uncommon in the Netherlands. Since this means that there is a lower availability of data on this organizational structure, the focus in this research is only on roundtrip B2C carsharing. Section 2.3.1 explained that within roundtrip carsharing, a difference is made between fixed parking places and zonal parking spaces. In this research, both ways of return are included in the term 'roundtrip B2C carsharing'.

3.1.3 Focus on supply of roundtrip B2C shared cars

Unfortunately, a complete database with active member numbers or bookings of each carsharing provider is not publicly available and the will of carsharing providers to share their data is limited. Therefore, to obtain data on the demand for carsharing that is representative for a complete neighbourhood, data about the demand of a large number of inhabitants of a neighbourhood is needed. Thereby, to distinguish between the demand for carsharing in different neighbourhoods, this data needs to be gathered for a large number of neighbourhoods as well. Hence, the analysis in this research starts with a focus on the supply of roundtrip B2C shared cars to obtain representative data on a large number of neighbourhoods in the G44 cities. Then an attempt is made to establish the link between the supply of shared cars and the demand for carsharing to be able to explain the variation in the demand for roundtrip B2C carsharing.

3.2 Main research question

This research aims at contributing knowledge to the explanation of the variation in the demand for roundtrip B2C carsharing in a neighbourhood in the G44 cities based on the characteristics of a neighbourhood. Consequently, this research answers the following research question:

How are neighbourhood characteristics explaining the variation in the supply of roundtrip B2C shared cars in neighbourhoods in the G44 cities in the Netherlands and can this explanation of variation be used to explain the variation in the demand for roundtrip B2C carsharing in these neighbourhoods?

3.3 Sub-questions

To provide an answer to the main research question, sub-questions are established. The first part is about explaining the variation in the supply of roundtrip B2C shared cars in neighbourhoods in the G44 cities in the Netherlands. Then, the effect of the observed supply of roundtrip B2C shared cars on the demand for roundtrip B2C carsharing is analysed and the relationship between the neighbourhood characteristics that affect the demand for carsharing is compared to the neighbourhood characteristics that affect the supply of roundtrip B2C shared cars. Finally, it is analysed what factors other than neighbourhood characteristics explain the variation in the demand for roundtrip B2C carsharing in neighbourhoods in the G44 cities in the Netherlands. The sub-questions are formulated as followed:

1. How are neighbourhood characteristics explaining the variation in the supply of roundtrip B2C shared cars in neighbourhoods in the G44 cities in the Netherlands?
2. What is the effect of the observed supply of roundtrip B2C shared cars on the demand for roundtrip B2C carsharing?
3. What is the relationship between the neighbourhood characteristics affecting the demand for and the neighbourhood characteristics affecting the supply of roundtrip B2C shared cars?
4. What factors are explaining the variation in the demand for roundtrip B2C carsharing of inhabitants of the G44 cities in the Netherlands?

3.4 Hypotheses

Based on the conceptual model that is presented in Figure 2.1 and the research questions that are discussed in the previous section, this section drafts hypotheses for further guidance to this research. The hypotheses are grouped by their corresponding research question.

3.4.1 Hypothesis reseach question 1

The existing literature presented in the theoretical framework in the previous chapter discussed a large number of factors that affect the demand for or supply of shared cars in a neighbourhood. It is expected that all factors that have been framed as neighbourhood factors, will affect the supply of roundtrip B2C shared cars in a neighbourhood in the G44 cities in the way that is presented in the hypothesis below.

1. There is more supply of roundtrip B2C shared cars in neighbourhoods within the Dutch G44 cities that...

(a) are highly urbanized

(b) are located within large cities

(c) have a high parking permit cost

(d) have high centrality

(e) have a high bar and restaurant accessibility

(f) have a low average age

(g) have a high number of females

(h) have a large number of highly educated individuals

(i) have a large number of residents with a migration background

(j) have a relatively high use of non-car forms of urban transport

(k) have a large number of part-time workers

(l) have a low average income

(m) have a low number of cars per household

(n) have a high share of walking in their modal split

(o) have a large percentage of rental houses

(p) have a large percentage of one-person households.

3.4.2 Hypothesis research question 2

The size of the correlation between the demand and supply of shared cars and whether the supply of shared cars creates more demand for them is unknown. However, from the literature discussed in the previous section, there is a lead to believe that there is a correlation between the demand for roundtrip B2C carsharing and the observed supply of roundtrip B2C shared cars. Therefore, it is expected that more observed supply of roundtrip B2C shared cars in a neighbourhood is causing more demand for roundtrip B2C carsharing in the same neighbourhood.

2. More observed supply of roundtrip B2C shared cars in a neighbourhood is causing more demand for roundtrip B2C carsharing in the same neighbourhood

3.4.3 Hypothesis research question 3

Furthermore, since there is a presumable coherence between the factors affecting the demand for carsharing and the factors affecting the supply of shared cars, it is expected that the neighbourhood characteristics of the inhabitants of the G44 cities that show the most demand for roundtrip B2C carsharing are similar to the characteristics of the neighbourhoods with the highest supply of roundtrip B2C shared cars. This means that it is expected that there is more demand for a roundtrip B2C shared car by individuals that live in neighbourhoods with the characteristics that comply with the neighbourhood characteristics listed in hypothesis 1.

3. The characteristics of the neighbourhoods of individuals that show the most demand for roundtrip B2C carsharing are similar to the characteristics of the neighbourhoods with the most supply of roundtrip B2C shared cars.

3.4.4 Hypotheses research question 4

The literature described in the theoretical framework in the previous chapter showed that the most important motives for using a shared car are to have the convenience of having a car only when needed and to avoid the responsibilities with maintenance and repair of a private car. Another strong motive is to start using a shared car when shared cars and their carsharing pick-up locations are accessible near an individual's place or workplace. Since the use of a roundtrip B2C shared car is related to all of these motives in the same way as a shared car in general, it is expected that these are the strongest motives for the use of a roundtrip B2C shared car in the G44 cities in the Netherlands.

4. Roundtrip B2C shared cars are mostly used because of the convenience of having a car only when needed, to avoid responsibilities with maintenance and repair, and when these cars and their carsharing pick-up locations are accessible near an individual's place or workplace

The literature also showed that shared cars are used the most to visit a friend or a relative and for a day out. It is expected that a roundtrip B2C shared car is used the most for these purposes as well.

5. Roundtrip B2C shared cars are mostly used to visit a friend or a relative and for a day out.

Literature also shows that individuals that already have experience with other shared mobility than shared cars are showing more demand for a shared car as well. This is because they are already familiar with the use of shared transportation, which makes the threshold to start using shared cars lower. Since the shared car is considered a relatively sustainable transportation mode, individuals with strong environmental ideologies tend to show more demand for carsharing. It is expected that the demand for roundtrip B2C carsharing is affected by these factors in the same way.

6. There is more demand for a roundtrip B2C shared car by individuals that have experience with other shared mobility and have strong environmental ideologies

3.4.5 Hypothesis main research question

When all previous hypotheses are accepted it is expected that the variation in the demand for roundtrip B2C carsharing in the neighbourhoods in the G44 cities in the Netherlands can be explained by the explanation of the variation in the supply of roundtrip B2C shared cars in the neighbourhoods in the G44 cities.

7. The variation in the demand for roundtrip B2C carsharing in a neighbourhood in the G44 cities in the Netherlands can be explained by the neighbourhood characteristics that explain the variation in the supply of roundtrip B2C shared cars in the neighbourhoods in the G44 cities

Chapter 4

Methodology

This chapter describes the methods that are applied in this study. First, the approach that is used to answer the main research question is discussed. Then the data collection methods are described and the final part elaborates on the methods for analysis used in this study.

4.1 Research approach

The main objective of this research is to find out whether the variation in the demand for roundtrip B2C carsharing in a neighbourhood in the G44 cities in the Netherlands can be explained based on neighbourhood characteristics.

The research is approached by first using the characteristics of a neighbourhood to explain the variation in the supply of roundtrip B2C shared cars in neighbourhoods that represent the neighbourhoods in the G44 cities. This is done by making use of a regression model with the number of roundtrip B2C shared cars per 100,000 inhabitants in a neighbourhood as the dependent variable and all neighbourhood factors described in the conceptual model in Figure 2.1 as independent variables. The results of this regression are answering the first research question and are presented in Chapter 6.

Subsequently, it has been studied whether this explanation of the variation in the supply could be used to explain the variation in the demand for roundtrip B2C carsharing in neighbourhoods in the G44 cities as well. To do this, first, the effect of the observed supply of roundtrip B2C shared cars on the demand for roundtrip B2C carsharing is analysed to find out whether the supply must be taken into account when explaining the variation in the demand for roundtrip B2C carsharing. The results of this analysis are answering the second research question and are discussed in Chapter 7. Then, the relationship between the neighbourhood characteristics that affect the demand for carsharing and the neighbourhood characteristics that affect the supply of roundtrip B2C shared cars is analysed by making use of three different methods. The results from this analysis provide an answer to the third research question and are presented in Chapter 8. The input for the relationship between the supply and demand results from a survey that is distributed over 6 municipalities that together represent the G44 cities in the Netherlands. The representativeness of this survey is discussed in Chapter 5.

In addition to the data on the demand and neighbourhood characteristics, also data on all individual factors framed in Figure 2.1 is obtained from the survey. This is used to find out what factors other than neighbourhood characteristics explain the variation in the demand for roundtrip B2C carsharing of inhabitants of the G44 cities in the Netherlands. The results of this analysis are presented in Chapter 9 and are answering the fourth research question.

To create a manageable overview of the steps taken in this research, the research approach is visualized in Figure 4.1. The figure shows the contribution of each sub-question to the answer to the main research question. The different colours show exactly what external input is used and what information is generated to obtain the output of the sub-questions.

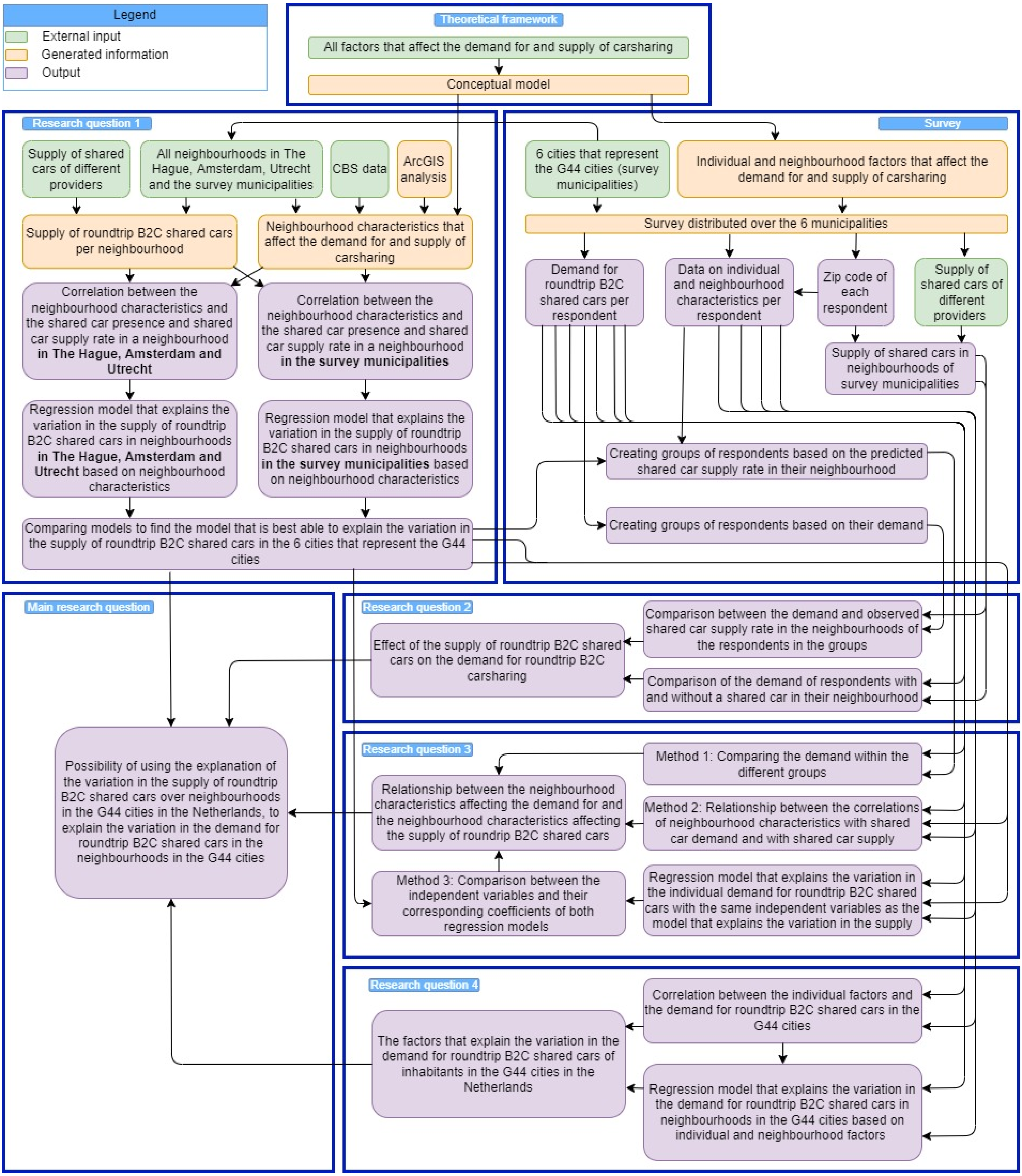

FIGURE 4.1: Research approach

4.2 Data collection

This study makes use of data regarding the current distribution of the supply of B2C shared cars in neighbourhoods in the G44 cities and the neighbourhood characteristics of each of these neighbourhoods. Both are the input for the regression model that is used to answer the first research question. This study also encompasses a survey with a focus on obtaining data on the demand for roundtrip B2C carsharing in different municipalities and the individual factors that affect the demand for carsharing. The exact way of collecting this data is elaborated on in this section.

4.2.1 Roundtrip B2C shared car supply

The number of roundtrip B2C shared cars per 100,000 inhabitants in a neighbourhood is the dependent variable in the regression model that explains the variation in the supply of roundtrip B2C shared cars in a neighbourhood. This dependent variable is used to provide a fair comparison between the supply of shared cars in different neighbourhoods. Without dividing it by the number of inhabitants in a neighbourhood, the number of shared cars in a neighbourhood is not saying much about the supply of shared cars in this neighbourhood since the size of a neighbourhood can be very different. The number of shared cars per square meter is not a fair comparison either, since some large neighbourhoods have a very low number of inhabitants. Thereby, the number of shared cars per 100,000 inhabitants is also used by the CROW (2021a) in their data set about shared cars in different municipalities and can therefore be considered a good measure of the supply in a neighbourhood that corrects for the neighbourhood size. In the remainder of this report, the number of roundtrip B2C shared cars per 100,000 inhabitants is also referred to as the shared car supply rate. 2 regression models that explain the variation in the shared car supply rate in neighbourhoods in the G44 cities are developed. One is based on the neighbourhoods in 6 municipalities that together represent the G44 cities. These municipalities are Almere, Arnhem, Enschede, Nijmegen, Zoetermeer and Zwolle. Since the survey in this research is also distributed over these municipalities, they are also referred to as the survey municipalities. The way in which the survey municipalities represent the G44 cities in the Netherlands is discussed in Chapter 5. The other regression model is based on cities with a large variation in neighbourhood types and carsharing providers. Therefore, neighbourhoods in the largest cities with the highest B2C shared car supply rate are used to obtain data on the carsharing supply. Data on the shared car supply rate in municipalities is retrieved from the database of the CROW that keeps track of the sustainability score of each municipality (CROW, 2021c). The 5 Dutch cities with the highest B2C shared cars per 100,000 inhabitants are Rijswijk, Tubbergen, The Hague, Amsterdam and Utrecht. Because Rijswijk and Tubbergen are not part of the G44 cities, this second regression model is based on the neighbourhoods of The Hague, Amsterdam and Utrecht.

To find out whether the number of neighbourhoods in these municipalities is large enough to obtain a representative sample, it must be taken into account that the regression models must also be validated on a data set that has not been used to create the model. The data used to create the model is called the training set and data used to validate the model is referred to as the validation set. The validation set provides an unbiased evaluation of a model fit on the training data set. Applying the Pareto principle indicates that 20% of the neighbourhoods must be used as the validation set and 80% of the neighbourhoods as the training set (Tardi, 2020). There are a total of 704 neighbourhoods in The Hague, Amsterdam and Utrecht. When applying the Pareto principle to all neighbourhoods in these three municipalities, the sample size of the training set is 563 neighbourhoods. Because there are 11,726 neighbourhoods in the G44 cities, 563 neighbourhoods are enough to obtain a data set that represents the neighbourhoods in the G44 cities with a 95% confidence level and a margin of error of 4% (SurveyMonkey, 2021).

There are a total of 368 neighbourhoods in the survey municipalities that together represent the G44 cities. When leaving 20% of these neighbourhoods out for the validation set, the training set consist of 294 neighbourhoods. This is enough to obtain a data set that represents the neighbourhoods in the G44 cities with a 95% confidence level and a margin of error of 6% (SurveyMonkey, 2021).

The supply of shared cars in each neighbourhood in The Hague, Amsterdam, Utrecht and in each neighbourhood in the survey municipalities is measured by counting the number of roundtrip B2C shared cars that are offered by all providers that are listed in Appendix A as having a roundtrip organizational structure. The number of shared cars in each neighbourhood are taken from the carsharing providers' websites during weekdays between 25 and 28 May 2021. The total number of roundtrip B2C shared cars in each neighbourhood is divided by the neighbourhood population size and multiplied by 100,000 to obtain the shared car supply rate in each neighbourhood. The neighbourhood population size has been retrieved from the CBS database on districts and neighbourhoods (CBS, 2020b).

4.2.2 Neighbourhood characteristics

All factors that have been framed in the conceptual model in Figure 2.1 as neighbourhood factors are used as independent variables in the regression models that explain the variation in the shared car supply rate. Data on these variables have been collected for each neighbourhood in The Hague, Amsterdam and Utrecht and for each neighbourhood in the survey municipalities by making use of indicators that express the characteristics of a neighbourhood. Most of these indicators and the data on these indicators are retrieved from the CBS database on districts and neighbourhoods (CBS, 2020b). However, also some data and indicators are based on simple analyses in ArcGIS or other open data sources than the CBS database. All neighbourhood characteristics and the corresponding indicators are shown in Table 4.1. An overview of the sources or techniques used to find the indicators for each neighbourhood characteristic is presented in Appendix B. All neighbourhood factors that have been framed in the conceptual model in Figure 2.1 are covered except for the average job type. This is because there is no data available on the average job type per neighbourhood.

Almost all neighbourhood characteristics are derived from CBS data and therefore most indicators in Table 4.1 are based on the indicators that are used by the CBS (CBS, 2020b). However, to obtain an indicator to find the modal split in a neighbourhood, the urbanization level of the neighbourhood is used. This is because to obtain more detailed modal split data on a neighbourhood level, extensive analysis is required. This extensive analysis is out of the scope of this research since it outweighs the benefits of the application of the analysis. The CBS (2017) provides insight into the average modal split per urbanization level in the Netherlands. Therefore, with the urbanization level of a neighbourhood, a rough estimate of the modal split in a neighbourhood can be included in the regression model.

Other exceptions in the collection of data on the independent variables are made to obtain an indication of the centrality and parking costs in a neighbourhood. ArcGIS is used to calculate the centrality of each neighbourhood by calculating the average Euclidean distance from a large number of points in a neighbourhood to the centroid of the corresponding neighbourhood. To obtain data on the parking costs, the costs of a first resident parking permit in each neighbourhood have been derived from the website of each municipality. All neighbourhood characteristics and their corresponding indicators are presented in Table 4.1. An overview of the sources or techniques used to find the indicators for each neighbourhood characteristic is presented in Appendix B.

TABLE 4.1: Neighbourhood factors and their indicators

Neighbourhood characteristic	**Indicator**
Address density	Number of addresses per km^2 per 1,000 addresses
Public transport network	Average distance to train station in km
City scale	Number of inhabitants in the complete municipality per 100,000 inhabitants
Parking costs	Costs for a parking permit for the first resident in a household in euros
Centrality	Average distance to the centre of the neighbourhood in km
Bar and restaurant density	Number of restaurants within 1 km Number of bars within 1 km
Age	Percentage of individuals between 0-14 years old Percentage of individuals between 15-24 years old Percentage of individuals between 25-44 years old Percentage of individuals between 45-64 years old Percentage of individuals 65 years old or older
Gender	Percentage of male individuals Percentage of female individuals
Education level	Percentage of low educated individuals [1] Percentage of moderately educated individuals[2] Percentage of highly educated individuals [3]
Ethnicity	Percentage of individuals with a western migration background[4] Percentage of individuals with a non-western migration background[5]
Modal split & pedestrian and bike commuting regime	Percentage of individuals in urbanization class 1 ($\geq$ 2,500 addresses per km^2) Percentage of individuals in urbanization class 2 (1,500-2,500 addresses per km^2) Percentage of individuals in urbanization class 3 (1,000-1,500 addresses per km^2) Percentage of individuals in urbanization class 4 (500-1,000 addresses per km^2) Percentage of individuals in urbanization class 5 ($<$500 addresses per km^2)
Income	The percentage of individuals in the lowest 40% of the national average income per individual The percentage of individuals in the highest 20% of the national average income per individual
Car ownership	Average number of private cars per household
Rental households	Percentage of individuals living in rental houses
Household composition	Percentage of individuals living in single-person households Percentage of individuals living in multiple-person households with children Percentage of individuals living in multiple-person households without children

[1] *Individuals of which the highest attained level of education is elementary school, VMBO, the first three years of HAVO/VWO and the entrance education, the former MBO1.*
[2] *Individuals of which the highest attained level of education is HAVO/VWO, MBO2, MBO3 and MBO4.*
[3] *Individuals of which the highest attained level of education is HBO or University.*
[4] *Individuals with a migration background with as origin group one of the countries in the continents Europe (excl. Turkey), North-America and Oceania or Indonesia or Japan.*
[5] *Individuals with a migration background with as origin group one of the countries in the continents Africa, Latin America and Asia (excl. Indonesia and Japan) or Turkey.*

4.2.3 Survey

The most important data source in this research is a descriptive survey that is used to find out whether the explanation of the variation in the supply of roundtrip B2C shared cars can be used to explain the variation in the demand for roundtrip B2C carsharing. To do this, the survey has the aim to collect data on the demand for roundtrip B2C carsharing and the individual and neighbourhood factors in the conceptual model that could affect this demand. Also, data on the perceived supply of shared cars is gathered to find out whether individuals know about the presence of a shared car in their neighbourhood. The survey is distributed in Almere, Arnhem, Enschede, Nijmegen, Zoetermeer and Zwolle and is in Dutch since this is the native language of the respondents. The representativeness of the sample is discussed in Chapter 5.

In summary, the survey aims to gain information on the demand for roundtrip B2C carsharing, on all individual and neighbourhood factors that are framed in the conceptual model in Figure 2.1 and on the supply of roundtrip B2C shared cars in the neighbourhood of the respondent. The survey consists of multiple parts that are further elaborated on in this section. The complete survey can be found in Appendix C and the survey is structured as followed:

1. Introduction
2. Quota questions
3. Questions about the respondent's demand for carsharing, the individual factors and the respondent's perceived supply of shared cars in their neighbourhood
4. Questions about the neighbourhood characteristics of the respondents

4.2.3.1 Introduction and quota questions

In the introduction, the respondents are introduced to the topic of the survey and it is explicitly mentioned that the respondents could also participate when the respondent has never made use of a shared car before. Then it is explained what a shared car is and that the survey is focusing on roundtrip B2C shared cars without mentioning the terms 'roundtrip' and 'B2C' to prevent the survey from getting too difficult. Also, the structure and expected duration of the survey have been explained.

To ensure that only respondents that fit the requirements of the target group are part of the survey, quota questions are used. These quota questions consist of a maximum of respondents that meet a certain requirement. Quota questions are used to ensure a maximum number of respondents per municipality, to reach an equal number of respondents per municipality. The survey is also restricted for respondents without a driver's license or respondents that are aged younger than 18 years old.

4.2.3.2 Demand for roundtrip B2C carsharing

To measure the demand for roundtrip B2C carsharing of each respondent, they are asked how likely it is that they would make use of a shared car at least once a year when a shared car is always available at walking distance from their house. A 5-point Likert scale is used to indicate the demand with the following options:

1. I would definitely not make use of the shared car
2. It is not likely that I would make use of the shared car
3. Neutral/I do not know whether I would make use of the shared car
4. I would probably make use of the shared car
5. I would definitely make use of the shared car

4.2.3.3 Individual factors

The survey is also used to gather data on the individual factors that are framed in the conceptual model in Figure 2.1. The individual factors are grouped as factors based on experience with shared mobility, factors based on attitudes and factors based on motives and purposes for carsharing.

Factors based on experience with shared mobility
The experience with shared mobility is measured by questions regarding the experience with a shared car and with other forms of shared mobility. The respondents were asked about whether they ever made use of a shared car in particular, and if they did, what the use frequency was in the past year. The possible answers for the frequency range from 1 to 2 times a year, to 4 times a week or more. Then the respondents were asked whether they ever made use of another form of shared mobility. This could be a shared bicycle, shared scooter, shared step or other form of shared mobility. When the respondent ever used another form of shared mobility, they were asked about their use frequency in the past year in the same way as has been done for the shared car experience.

Factors based on attitudes
The factors based on attitudes in the conceptual model in Figure 2.1 consist of green party voters and the attitude towards different transport modes. Just as in the literature described in Chapter 2, a green party voter is considered an indication of acting on having an environmental ideology. To find out whether an individual is a green party voter the political spectrum of the Netherlands is used (Kieskompas, 2021). In this research, it is assumed that green party voters can be categorized as individuals with a left-oriented political view. The question is stated as followed: 'political preferences are often referred to as 'left' or 'right' orientated. In general, where would you place your political views on a scale from 1 (left) to 5 (right)?' The possible answers are based on the Dutch political spectrum and are presented with their corresponding parties (Kieskompas, 2021). The possible answers are presented in the fourth section of Appendix C.

To find the attitude and perception of the respondents towards different transport modes, the respondents are asked about the importance of the infrastructure and flow of different transport modes. The level of importance is indicated on a 5-point Likert scale. This means that for each transport mode, they indicate their level of agreement towards a statement that describes their attitude towards the transport mode from 1 (totally disagree) to 5 (totally agree) with point 3 being neutral. The statements that are used to find the level of importance for different transport modes are derived from research of Te Morsche, La Paix Puello, and Geurs (2019). In that research, the attitudes are more extensively studied than is the objective in this study. To avoid the survey getting too extensive on the topic of attitudes, only one statement per transport mode has been used. The statements used in the survey are listed in the third section of Appendix C. These statements have been chosen because they represent the attitude towards these different transport modes in the most general way.

Factors based on motives and purposes for carsharing
The respondents had to indicate their level of agreement to 6 possible motives for making use of a shared car on a 5-point Likert scale as well. The possible motives have been derived from the theoretical framework in Chapter 2 and have been stated in two possible ways. For the respondents that indicated to have experience with a shared car they are stated as 'I am using a shared car to ...' and for the respondents that never used a shared car they are stated as 'I am considering using a shared car to ...'. Subsequently, the respondents must indicate their level of agreement to 5 possible purposes to make use of a shared car on the same 5-point Likert scale. These possible purposes have also been derived from the literature in Chapter 2. The statements about the possible motives and purposes can be found in the third section of Appendix C as well.

4.2.3.4 Neighbourhood characteristics and roundtrip B2C shared car supply in neighbourhoods of respondents

The survey is also used to find all characteristics of the respondents or neighbourhoods of the respondents that are framed as neighbourhood characteristics in the conceptual model. Data on the number of private cars in the household of the respondents, their year of birth and net income is obtained by making use of open questions. Thereafter, closed questions about the costs of a first resident parking permit, education level, migration background, household composition and whether their house is rented or owner-occupied are asked. The possible answers are based on the indicators used in Table 4.1. However, for the costs of a first resident parking permit, a choice can be made between multiple price ranges and the options that there is no parking permit necessary or that there is a permit necessary but the respondent does not know the costs.

To be able to find what the effect of the working situation of an individual is on the demand for carsharing, the respondents are asked how many hours of paid work they do per week and which of the situations below applies the most to them. Although the CBS (2020b) does not feature data about these situations on a neighbourhood level, the categories for the situations are derived from their database. The different situations are the following:

1. I am working (full time or part-time)
2. I am (partly) incapacitated or unemployed
3. I am retired
4. I am studying or doing an internship

5. I am a housewife/houseman
6. I am doing voluntary work
7. I do not know
8. A different situation applies to me

Since information on some neighbourhood characteristics is not generally known by everyone, the zip code of the respondents is obtained to be able to find the remaining neighbourhood characteristics of each respondent by making an analysis in ArcMap or using CBS data on the indicators discussed in Section 4.2.2. The characteristics that are found with the zip code are all based on the neighbourhood of the respondents and are the following: the average distance to the city centre, the address density, the average distance to a train station, the city scale, the bar and restaurant accessibility and the urbanization class of the neighbourhood. The zip code is also used to find the number of roundtrip B2C shared cars in the neighbourhood of the respondent.

Because data on the modal split of a neighbourhood can only be retrieved from the zip code of the respondents by using the urbanization level of a neighbourhood as rough estimate of the modal split, data on the modal split is retrieved from the respondents themselves. The respondents are asked how many times they made use of different transport modes in the past year. The possible answers for the frequency are the same possible answers that have been used for the use frequency of shared mobility and range from 1 to 2 times a year, to 4 times a week or more. The respondents must indicate their use frequency for a private car (also as a passenger), their own bicycle, walking as a transportation mode, taxi, motorcycle/moped/scooter and public transport.

4.2.3.5 Perceived supply of shared cars

To find out whether the respondents know about the presence of a shared car in their neighbourhood they are asked whether there is a shared car placed in their neighbourhood or not. Since it was expected that not all respondents know whether there is a shared car located in their neighbourhood, they are able to choose between 'yes', 'no' and 'I don't know'.

4.3 Analysis methods

This section is discussing the analysis methods that are applied in this study. The first part elaborates on the explanation of the variation in the supply of shared cars in a neighbourhood by making use of a regression model based on The Hague, Amsterdam and Utrecht and a regression model based on the survey municipalities. Then the analysis methods used to find out whether the explanation of the variation in the supply could be used to explain the variation in the demand for roundtrip B2C carsharing in neighbourhoods in the G44 cities are discussed. Finally, the methods used to find out what factors other than neighbourhood characteristics explain the variation in the demand for carsharing in the G44 cities is elaborated on.

4.3.1 Variation in the supply of roundtrip B2C shared cars in a neighbourhood

The indicators that are listed in Table 4.1 are used as independent variables to develop two regression models that predict the number of roundtrip B2C shared cars per 100,000 inhabitants in a neighbourhood. The dependent variable is the shared car supply rate (the number of roundtrip B2C shared cars per 100,000 residents) in each neighbourhood. The software that is used to build the model is SPSS, a program that can be used for statistical analysis.

25% of all neighbourhoods in The Hague, Amsterdam and Utrecht and 78.5% of the neighbourhoods in the survey municipalities have an observed shared car supply rate of 0. Because of this relatively large number of zeros in the data on the dependent variable, fitting a regression model that predicts the shared car supply rate based on the complete data set would result in a bias in the predictions of the models towards 0. Additionally, there may be other dynamics behind the characteristics that explain the presence of a shared car and the characteristics that explain how many shared cars per inhabitant there are in a neighbourhood. Therefore, first, a regression is used to explain the variation in the presence of a shared car in a neighbourhood. Then the neighbourhoods in which the presence of a shared car is predicted by this first regression, are used as input for another regression that explains the variation in the number of roundtrip B2C shared cars per 100,000 inhabitants in a neighbourhood. Since the second regression is based on fewer neighbourhoods with an observed shared car supply rate of 0, there is less bias towards a prediction of a shared car supply rate of 0. Also, by making use of two different models, the possible other dynamics behind the characteristics that explain the shared car presence and the shared car supply rate are better taken care of. This section starts by explaining how the data is prepared and is discussing the regression techniques that are eventually used to predict the shared car supply rate more extensively.

4.3.1.1 Data preparation

This section describes the steps that are taken to prepare the data to be able to develop the regression models. First, it is discussed what corrections have been made to obtain data that is more representative for the neighbourhoods and then the technique that is used to cope with missing data is elaborated on.

Correcting data
The data on the independent variables that describe the age in a neighbourhood should together sum up to 100%. Since this was not always the case, all sums in the range between 95% and 105%, were extrapolated to sum up to 100%. The same has been done for the education level, gender and household composition in a neighbourhood. When the sum was not in this range, all data on the particular group of variables has been indicated as missing data.

Coping with missing data

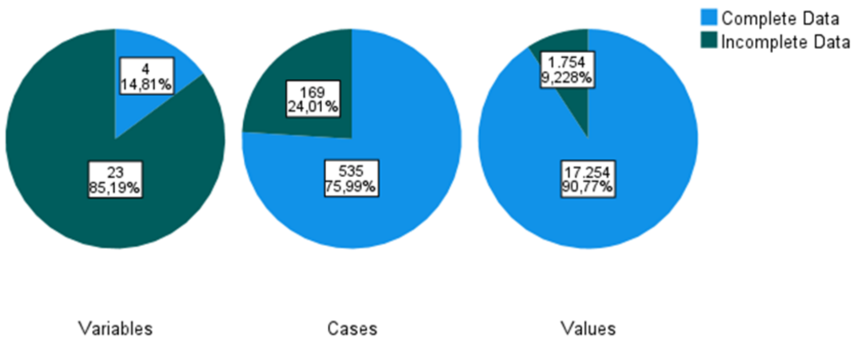

FIGURE 4.2: Overall summary of missing values in the neighbourhood in The Hague, Amsterdam and Utrecht

Figure 4.2 shows that for almost all variables at least one neighbourhood has missing data in The Hague, Amsterdam and Utrecht. When only neighbourhoods with complete data are used, only 76% of the neighbourhoods can be used and this results in the reduction of statistical power. Also, the analysis could be biased if the missing data is not Missing Completely At Random (MCAR) (Statistics Solutions, 2021b). Because a relatively large number of the neighbourhoods with missing data are neighbourhoods with no shared cars, the missing data is not MCAR. To complete the data set, multiple imputation is used (Zhang, 2016).

Imputation is the process of replacing missing values with imputed values and can be done with different methods (Zhang, 2016). Since the percentage of missing data is between 5% and 40% and the data is not only missing on the dependent variable, multiple imputation is used. Multiple imputation is a simulation-based statistical technique for handling missing data (Jakobsen, Gluud, Wetterslev, & Winkel, 2017).

Since only the independent variables consist of missing values, single variable imputation can be used as multiple imputation method (Jakobsen et al., 2017). Humphreys (2015) discussed multiple techniques to carry out single variable imputation of which the regression method is used in this study. This is done by identifying and replacing missing values by making use of the predicted scores of a linear regression equation for each independent variable with missing data. The regression equations are automatically generated by SPSS and are only used to complete the data set and are therefore different from the regressions that are used to explain the variation in the supply of shared cars. SPSS creates data for the missing values multiple times, resulting in a different data set each time. These data sets are analysed and combined into a single multiple-imputation result (Jakobsen et al., 2017).

The independent variables with the most missing values are education level, household type, age and income. Because the linear regressions are based on the variables with the strongest correlation with the variable with missing data, data is needed on variables with a strong correlation with the missing variables. The correlations between the independent variables are presented in Appendix D. The variables describing the education level, household type, age and income all have multiple strong correlations with independent variables that only have missing data in a few neighbourhoods. Therefore, the imputation is based on significant data and the data set is suitable for multiple imputation. Hence, this technique is used to cope with the missing data. The disadvantages of applying regression imputation to cope with missing data are that the model fit could be overestimated and that the multicollinearity of each independent variable increases. The term 'multicollinearity' is further discussed in Section 4.3.1.2. Multiple imputation is also used to cope with missing data in the neighbourhoods in the survey municipalities.

Creating training and validation set
For the validation of the model, the neighbourhoods in The Hague, Amsterdam and Utrecht and the neighbourhoods in the survey municipalities must be split into a training set and a validation set. All neighbourhoods are randomly grouped into both sets with a training set/validation set-ratio of 80:20. This means that the training set of the regression model that explains the shared car presence based on the neighbourhoods in The Hague, Amsterdam and Utrecht consists of 563 randomly selected neighbourhoods and that the validation set consists of the other 141 neighbourhoods. For the regression model that explains the shared car presence based on the survey municipalities, the training set and validation set consist of respectively 294 and 74 randomly selected neighbourhoods. On all neighbourhoods of which the regression predicts the presence of a shared car, also a training set/validation set-ratio of 80:20 is used to explain the variation in the shared car supply rate.

4.3.1.2 Regression techniques

For both the regression model that is based on The Hague, Amsterdam and Utrecht and the regression model that is based on the survey municipalities, first, a regression model is used to explain the variation in the presence of a roundtrip B2C shared car. As discussed before, the neighbourhoods with a predicted presence can then be used in the explanation of the variation in the number of roundtrip B2C shared cars per 100,000 inhabitants in a neighbourhood. The techniques used for both regressions are discussed in this section.

Predicting the shared car presence
A roundtrip B2C shared car in a neighbourhood can either be present or not present and hence, the presence of a shared car in a neighbourhood is a binary dependent variable. A binary dependent variable can be explained with multiple regression techniques of which the most commonly used techniques are the linear probability model, the probit model and the logit, or logistic, model. The linear probability model has the disadvantage that it is not able to capture the nonlinear nature of the binary dependent variable and therefore may predict probabilities that lie outside the interval between 0 and 1. Probit and logit models are able to capture this nonlinearity and therefore only predict probabilities in the interval between 0 and 1 (Hanck, Arnold, Gerber, & Schmelzer, 2021). Although the predictions of all three models are often close to each other, it is preferable to use a regression technique that only predicts a probability in the interval between 0 and 1. The book of Hanck et al. (2021) suggests that the choice between a probit or logit model is based on which regression technique is easiest to use in the statistical software of choice. Since a logit model is easier to apply in SPSS, the presence of a shared car

in a neighbourhood is predicted with a logit model. The logit model can also be referred to as logistic model and this term is used in this research.

To predict the presence of a shared car in a neighbourhood, a (binary) logistic regression model has been developed. This is a regression technique to predict a binary (or dichotomous) dependent variable based on one or more continuous or nominal independent variables. That the dependent variable is binary means that the outcome of the binary logistic regression can only be 0 or 1. In the context of this study, 0 means that the model predicts that a neighbourhood does not have a roundtrip B2C shared car and a 1 means that the neighbourhood has at least 1 roundtrip B2C shared car. This means that the presence of a roundtrip B2C shared car is the dependent variable in the binary logistic regression. In many ways, a binary logistic regression can be considered as a multiple linear regression, but for a dichotomous rather than a continuous dependent variable (Laerd, 2018). The logistic regression equation expresses the multiple linear regression equation in logarithmic terms and is structured as followed (Schüppert, 2009):

$$P(Y) = \frac{1}{1 + e^{-Y}} \tag{4.1}$$

In which

$$Y = b_0 + b_1 X_1 + b_2 X_2 + ... + b_p X_p \tag{4.2}$$

In Equation 4.1, P is the probability that the predicted value of the dependent variable, Y, is occurring. In Equation 4.2, X_1 through X_p are p distinct independent variables that describe the dependent variable, b_0 is the value of Y when all of the independent variables are equal to zero and b_1 through b_p are the estimated regression coefficients. These regression coefficients represent the change in Y relative to a one-unit change in the respective independent variable (Schüppert, 2009).

In the binary logistic regression, the probability of the presence of a roundtrip B2C shared car in a neighbourhood is estimated. This probability has a value between 0 and 1. The classification cutoff is the value of the boundary between an event that is classified as occurring and an event that is classified as not occurring. The value for the classification cutoff determines whether the model classifies the event as occurring, based on the predicted probability (Laerd, 2018). In the context of this study, a classification cutoff of 0.5 means that if the predicted probability of the presence of a shared car is greater than 0.5, the particular neighbourhood is classified as a neighbourhood with the presence of a shared car. Otherwise, the particular neighbourhood is classified as a neighbourhood without the presence of a shared car.

To determine the classification cutoff value, the classification table is used. The structure of a classification table is shown in Table 4.2.

TABLE 4.2: Structure of a classification table (Zaiontz, 2015)

	Predicted neighbourhoods without shared car	**Predicted neighbourhoods with shared car**	***Percentage correct***
Observed neighbourhoods without shared car	True negative	False positive	*Predicted negative*
Observed neighbourhoods with shared car	False negative	True positive	*Predicted positive*
			Overall percentage correct

The classification cutoff value is determined based on the false positives and false negatives in the classification table. The object with determining a cutoff value is to minimize the cost of the false positives and false negatives and depends on which one is considered more important (Griffin, 2020). In the context of this research, the goal is to have a smaller number of neighbourhoods without a shared car as input for the regression that eventually predicts the shared car supply rate in a neighbourhood. The number of neighbourhoods without a shared car in the data set is brought down by using the neighbourhoods with a predicted shared car presence as input for this next model. If there are a large number of false positives in this data set, then the model is still based on a large number of neighbourhoods without a shared car and the model still has a bias towards 0. False negatives are neighbourhoods that are unduly not included in the data set. Therefore, too many false negatives lead to a decrease in the number of neighbourhoods in the data set and with that a decrease in statistical power. Therefore, the number of neighbourhoods that are incorrectly classified as neighbourhoods with the presence of a shared car must be the same as the number of neighbourhoods that are incorrectly classified as neighbourhoods without the presence of a shared car in this research. This is the case when the false positives are equal to the false negatives.

Prediction of the shared car supply rate

All neighbourhoods of which the binary logistic regression model predicts the presence of a shared car are used as input to explain the variation in the number of roundtrip B2C shared cars per 100,000 inhabitants in a neighbourhood with a new regression. To obtain a data set that is as large as possible, also the neighbourhoods with a predicted shared car presence that were used in the validation set of the binary logistic regression are used as input for the negative binomial regression. Since the dependent variable is a count that is divided by the population size of a neighbourhood, the dependent variable can be characterized as a count per unit of exposure. According to O'Hara and Kotze (2010), a regression model based on Poisson or negative binomial distributions should be used for this type of dependent variable. A unit of exposure may be time, space, distance, area, volume, or population size in a Poisson or negative binomial regression (NCSS, 2021)

The choice to use a Poisson or negative binomial regression model depends on whether the observed data on the independent variable is overdispersed or not. When the variation in the observations is equal to its mean, a Poisson regression can be used. However, when the variation exceeds the mean, the data is overdispersed. When overdispersion occurs, a negative binomial regression model can be used since this type of regression corrects for overdispersion by using an additional parameter that is used to model the variation. This variable is referred to as the dispersion parameter and is estimated automatically by SPSS (Gardner, Mulvey, &

Shaw, 1995). Since the variation in the observations on the number of shared cars per 100,000 inhabitants is greater than the mean, the data is overdispersed (Date, 2017). Therefore, a negative binomial regression model is used to predict the number of roundtrip B2C shared cars per 100,000 inhabitants in a neighbourhood.

A negative binomial regression is similar to a regular multiple regression except that the dependent variable is an observed count or count per unit of exposure that follows the negative binomial distribution. This means that instead of a regular regression equation, the exponent of the regression equation is used and that all possible values of the dependent variable are nonnegative integers. Frequently, a negative binomial regression is used when the dependent variable is a count without a unit of exposure. However, it can also be used when the dependent variable is a count per unit of exposure by applying an offset variable. This is necessary to control for the differences in the unit of exposure. Since the dependent variable in this study is the number of roundtrip B2C shared cars per 100,000 inhabitants, the number of inhabitants in a neighbourhood per 100,000 is the unit of exposure. The offset variable is created by including the natural logarithm of the variable that reflects the unit of exposure in the regression equation. The variable is incorporated into the model as an independent variable, with its regression coefficient fixed at 1 (Coxe, West, & Aiken, 2008).

The negative binomial regression is structured as followed (Federal Highway Administration, 2016):

$$\mu_i = exp(ln(t_i) + \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + ... + \beta_p X_{pi}) \tag{4.3}$$

In Equation 4.3, μ_i is the mean incidence rate of y per unit of exposure t for a particular observation i. In the context of this research, the mean incidence rate is the number of roundtrip B2C shared cars in a neighbourhood and the unit of exposure is the number of inhabitants in a neighbourhood per 100,000. X_1 through X_p are p distinct independent variables and β_1 through β_p are the estimated regression coefficients (Federal Highway Administration, 2016)(NCSS, 2021).

4.3.1.3 Process of finding the most effective model

The process of finding the most effective model for a binary logistic regression and for a negative binomial regression is similar. It is a process of finding the independent variables that describe the dependent variable most effectively and simultaneously show a significant contribution to the prediction of the dependent variable and do not show multicollinearity.

The correlations between the dependent variable and each independent variable are used to find the independent variables with the strongest correlation with the dependent variable since these independent variables describe the variation of the dependent variable most effectively. The most effective regression model can be found by starting with the independent variable with the strongest correlation with the dependent variable and adding independent variables that also have a strong correlation with the dependent variable. In an iterative process, new independent variables are added. When one of the variables is insignificant or shows multicollinearity, it is deleted and another variable is added.

The significance of the contribution of all independent variables to the prediction of the dependent variable is represented by the *p*-value. A *p*-value, or probability value, is a number describing how likely it is that data, or the relationship between an independent variable and the dependent variable, would have occurred by random chance. An independent variable with a *p*-value of 0.05 or less is statistically significant. It indicates strong evidence that the relation between the variables is not random since there is less than a 5% probability the relation is random (McLeod, 2019).

Multicollinearity occurs when two or more independent variables are highly correlated with one another and are both used to predict the dependent variable in a regression model. This means that an independent variable can be predicted from another independent variable in the model. Multicollinearity can be a problem in a regression model because it is not possible to distinguish between the individual effects of the independent variables on the dependent variable (Bhandari, 2020). This multicollinearity can lead to the redundancy of variables. This is because two independent variables might be providing the same information about the dependent variable and are therefore leading to unreliable coefficients of the independent variables. To check whether the independent variables do not show multicollinearity, Variable Inflation Factors (VIF) are used. The VIF shows for each independent variable in the regression how well the variable is explained by other independent variables in the regression. The higher the value of VIF for an independent variable, the higher the multicollinearity with another independent variable. The lowest value for the VIF is 1, which means that there is no correlation between the independent variable and the other variables in the regression. VIF exceeding 5 indicates a correlation between the particular independent variable and the other independent variables that is too high, and the variable should be dropped (Bhandari, 2020). By making use of the VIF, the problem of a higher multicollinearity that is caused by using multiple imputation is also taken care of.

To find the combinations of independent variables that fit the observations the best, the goodness of fit is used. In a regular multiple regression, the R^2 measure of goodness of fit is used to measure the proportion of the variation in the dependent variable that is predictable from the independent variables. However, in a binary logistic regression and negative binomial regression, only a so-called pseudo R^2 can be used. Pseudo R^2's are generally lower than the R^2 in multiple regression but are interpreted in the same manner (Laerd, 2018). However, since there is no consensus on what can be considered a good value for a pseudo R^2, a pseudo R^2 can only be used to compare different fits (Sieben & Linssen, 2009). The pseudo R^2 that is used to compare the different binary logistic regression models that predict the dependent variable, is the Nagelkerke R^2. The fit with the highest Nagelkerke R^2, in which all variables are significant and do not show multicollinearity is chosen as the binary logistic regression model that describes the variation in the presence of roundtrip B2C shared cars in the best way.

To measure the goodness of fit of the negative binomial regression model the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) are the most commonly used methods (Hardin & Hilbe, 2014). The quantity calculated for the BIC is different from the AIC, although it is proportional to the AIC. The AIC is derived from the most frequent probability and the BIC is derived from a Bayesian probability. A lower score for both the AIC and BIC represents a better-fitted model. Unlike the AIC, the BIC penalizes the model more for its complexity, meaning that more complex models will have a worse score and will be less likely to be selected (Brownlee, 2019). Both the AIC and BIC are used to compare the different fits and obtain the most effective model. When the AIC and BIC are not providing a clear distinction between the fit of different models, also the log-likelihood is used as a method to measure the goodness of fit.

Besides the different methods to measure the goodness for fit, another difference in the process of finding the most effective model for a binary logistic and a negative binomial regression, is the use of an offset variable in the negative binomial regression. The natural logarithm of the number of inhabitants of a neighbourhood is included by default in the negative binomial regression that explains the variation in the shared car supply rate. Additionally, the outcome of the negative binomial regression equation is the predicted number of roundtrip B2C shared cars per 100,000 inhabitants and there is no need to classify the prediction into different classifications, as is necessary for the binary logistic regression.

4.3.1.4 Regression validation

In this section, the process of deciding whether the results obtained from the regression models are acceptable as descriptions of the data is discussed. For both the binary logistic and the negative binomial regression, the process of analyzing the goodness of fit of the regression, the residual analysis and the validation sample are elaborated on.

The explanation of the variation in the shared car presence by the binary logistic regression can be evaluated based on the classification table discussed in Section 4.3.1.2. In the column on the right in Table 4.2, it is shown how well the binary logistic regression model predicts the presence of a roundtrip B2C shared car in a neighbourhood. The column shows how many neighbourhoods with zero shared cars have correctly been predicted, how many neighbourhoods with the presence of a shared car have correctly been predicted and the overall percentage of correctly predicted neighbourhoods.

To validate the goodness of fit of both the binary logistic and the negative binomial regressions, the log-likelihood test is used. In a regression analysis, the model is always compared to a model that is not based on independent variables but is random. This model is called the null model. The log-likelihood test is used to test whether the full model is a significant improvement of fit over the null model, which has a chi-squared distribution. When the significance of the log-likelihood test is below 0.05, then the model is considered to be an improvement of fit over the null model (McClave, Benson, Sincich, & Knypstra, 2011).

The next part of the model validation is the residual analysis. McClave et al. (2011) define a residual as the difference between an observation and the corresponding predicted value. Residual analysis is used to search for outliers in the residuals of the prediction and in most regressions, to control for certain regression assumptions (McClave et al., 2011).

The design of binary logistic regressions is made to overcome many of the restrictive assumptions of regressions. For example, linearity, normality and equal variations are not assumed, nor is it assumed that the error term variation is normally distributed (Statistics Solutions, 2021a). Therefore, the residual analysis of a binary logistic regression only consists of searching for outliers in the residuals of the prediction. The residuals in the binary logistic regression are the difference between the observations, that are 1 or 0, and the predicted probabilities between 0 and 1. The largest residuals are the predictions with a large predicted probability of shared car presence and an observation of 0 shared cars, and the predictions with a small probability of shared car presence and an observation that a shared car is present. The largest residuals must be checked for mistakes or unusual data in the variables that are the base of the prediction.

The residual analysis of a negative binomial regression is used to search for outliers as well, but also to control whether the model is not misspecified. To control whether the model is not misspecified, the residuals are plotted against each independent variable in the regression equation. When a curvilinear relation is present, it might be the case that a quadratic term must be included in the model (McClave et al., 2011). To search for outliers in the residuals of the negative binomial regression the predicted values are often plotted against the residuals or the standard deviation of the predicted values. One limitation to these plots is that the residuals and standard deviations can be larger for higher predicted values. Therefore, it is difficult to use the residuals or standard deviations to determine whether an observation is an outlier. Therefore, studentized residuals are used. Studentized residuals are more effective in detecting outliers and are calculated by dividing the residual by an estimate of its standard deviation (JMP, 2021). The studentized residuals are plotted against the predictions of the negative binomial regression. When the studentized residual of a prediction exceeds an absolute value of 3, it is generally classified as an outlier. However, the cutoffs of either 2 or 3 must not be taken too literally (Pennsylvania State University, 2021). Therefore, studentized residuals with an absolute value higher than 2 are checked for mistakes or unusual data.

To assess the performance of both the binary logistic and negative binomial regression, the regression equations are also applied to the validation sets. The validation sets are used as more objective measures of the performance of the model (Statistics.com, 2021). The binary logistic regression is validated by comparing the classification table of the validation set to the classification table of the training set. The coefficients in the tables must be relatively equal. The assessment of the performance of the negative binomial regression is done by plotting the studentized residuals against the predicted values of the validation set and comparing it with the plot of the studentized residuals against the predicted values of the training set.

4.3.1.5 Comparison of both regression models

Because the aim is to explain the variation in the supply of shared cars in the neighbourhoods in the G44 cities, the regression equation of the model that is based on The Hague, Amsterdam and Utrecht is also applied to the 6 municipalities that represent the G44 cities. The predictions of the model based on The Hague, Amsterdam and Utrecht and the model based on the survey municipalities are compared to assess which model performs the best. The model that is best able to explain the variation in the supply of roundtrip B2C shared cars in the neighbourhoods in the G44 cities is used to find out whether the explanation can be used to explain the variation in the demand for roundtrip B2C carsharing in these neighbourhoods as well.

4.3.2 Effect of the observed supply of roundtrip B2C shared cars on the demand for roundtrip B2C carsharing

To find out whether the supply of shared cars must be taken into account when explaining the variation in the demand, the effect of the observed supply of roundtrip B2C shared cars on the demand for roundtrip B2C carsharing is analysed in two ways. First, the observed shared car supply rate in neighbourhoods of respondents that indicated a different level of demand for carsharing is compared to each other. Then, a comparison is made between the distribution of the demand of the respondents with and without a shared car in their neighbourhood. Finally, also the awareness of the respondents about the presence of a shared car is analysed. When the supply affects the demand, the relation between the variation in the demand and the variation in the supply of roundtrip B2C shared cars becomes complicated and a model that explains the variation in the demand must also take the supply in a neighbourhood into account.

4.3.2.1 Observed shared car supply rate in neighbourhoods of respondents with different demand for carsharing

To find the effect of the observed supply of roundtrip B2C shared cars on the demand for roundtrip B2C carsharing, first, all respondents that showed the same demand for roundtrip B2C carsharing are binned based on their level of demand. The respondents that answered 1 on the 5-point Likert scale are grouped in the first bin, the respondents that answered 2 are grouped in a second bin, etcetera. Then the average observed shared car supply in the neighbourhoods of all inhabitants in a certain group is calculated by three different measures of the supply of roundtrip B2C shared cars:

1. The average shared car presence: the total number of respondents that live in a neighbourhood with a shared car divided by the number of respondents in the bin, e.g. an average shared car presence of 0.44.
2. The average number of shared cars: the total number of shared cars in the neighbourhoods of the respondents in the bin divided by the number of respondents in the bin, e.g. an average of 2 shared cars.
3. The average shared car supply rate: the number of shared cars per 100,000 inhabitants in the neighbourhoods of the respondents in the bin divided by the number of respondents in the bin, e.g. an average shared car supply rate of 150 roundtrip B2C shared cars per 100,000 inhabitants.

The average shared car presence, the number of shared cars and the shared car supply rate are measured for each bin. To show the significance of these averages, also the standard error is calculated by dividing the standard deviation of the data in the bin by the square root of the number of neighbourhoods in the bin. By comparing the averages, it is clear to see whether the second hypothesis, which states that more supply of roundtrip B2C shared cars results in more demand for roundtrip B2C carsharing, can be accepted or rejected.

4.3.2.2 Demand of respondent with and without a shared car in their neighbourhood

The second method of finding the effect of the observed supply of roundtrip B2C shared cars on the demand for roundtrip B2C carsharing is to compare the demand of all respondents with and all respondents without a roundtrip B2C shared car in their neighbourhood. To be able to make a good comparison between the demand for carsharing in both these groups, the average answer and distribution of the answers on the 5-point Likert scale are presented in a table. To also show the significance of the average demand, also the standard error is calculated. Because this standard error is not used to show the significance of the supply of shared cars in a neighbourhood but to show the significance of the demand for an individual, this standard error is calculated by dividing the standard deviation of the answers on the 5-point Likert scale by the square root of the number of individuals. When the averages and distributions are comparable, it means that the supply of shared cars is not affecting the demand.

4.3.2.3 Awareness of shared car presence

The observed supply of roundtrip B2C shared cars can only affect the demand for roundtrip B2C carsharing when the respondents are aware of the presence of a shared car in their neighbourhood. Therefore, the answer to the question of whether the respondents know if there is a shared car located in their neighbourhood is analysed and their perceived shared car presence is compared to the observed shared car presence in their neighbourhood.

4.3.3 Relationship between the neighbourhood characteristics affecting the demand for and the neighbourhood characteristics affecting the supply of roundtrip B2C shared cars

To find out whether the explanation of the variation in the supply of shared cars can be used to explain the variation in the demand, the relationship between the neighbourhood characteristics that affect the supply of shared cars and the neighbourhood characteristics that affect the demand for carsharing must be comparable. To find this relationship, three methods are applied. In the first method, the relation between the individual demand for carsharing and the predicted supply of shared cars in the neighbourhood of the individual is analysed. Then a comparison is made between the correlations of the neighbourhood characteristics with the shared car supply and the demand for carsharing. The final method consists of comparing the regression coefficients of the regression model that explains the variation in the supply with the regression coefficients of a new model that predicts the demand for carsharing with the same independent variables that are used in the model that predicts the supply of shared cars in a neighbourhood.

4.3.3.1 Method 1: Relation between the individual demand for carsharing and the predicted supply of shared cars in the neighbourhood of the individual

To find out what the relationship between the neighbourhood characteristics that affect the demand for carsharing and the neighbourhood characteristics that affect the supply of shared cars is, it is important that the distribution of the individual demand for carsharing is comparable to the distribution of the supply in the neighbourhoods of the respondents predicted by the model. Therefore, the relation between the demand for roundtrip B2C carsharing and the predicted supply of shared cars is analysed. To find this relation, the respondents are binned in 4 new groups that are based on the predicted shared car supply rate in their neighbourhoods. The shared car supply rate is predicted by the regression model that explains the variation in the supply of roundtrip B2C shared cars. Group 4 consists of all respondents that have a predicted shared car supply rate of 0. All other respondents are equally distributed over the other groups. Group 1 consists of the respondents with the highest predicted shared car supply rates in their neighbourhood and group 3 consists of the respondents with the lowest predicted shared car supply rates. The other respondents together fill group 2. Subsequently, the relation between the average demand and distribution of the demand in these groups is compared to each other to find the relation between the demand and the predicted supply.

4.3.3.2 Method 2: Comparison between the correlations of the neighbourhood characteristics with the shared car supply and demand for carsharing

The second method to find the relationship between the neighbourhood characteristics that affect the demand for carsharing and the neighbourhood characteristics that affect the supply of shared cars, is to compare the correlation between the neighbourhood characteristics and the shared car supply with the correlation between the neighbourhood characteristics and the demand for carsharing. When the signs and magnitudes of the correlation of a neighbourhood characteristic with the demand is comparable to the correlation of the same neighbourhood characteristic with the supply of shared cars, it means that the characteristic has the same effect on the demand as it has on the supply of shared cars. This then signifies that there is a relation between the demand and the supply of roundtrip B2C shared cars. To compare these correlations a chart is created in which for each neighbourhood characteristic, the correlation with the demand, shared car presence and shared car supply rate is presented. With this chart, a clear overview is provided from which the relation between the correlations can be derived. When the correlations of a neighbourhood characteristic with the demand, the shared car presence and the shared car supply rate have the same sign and magnitude, it means that this neighbourhood characteristic has a similar effect on the demand as it has on the supply of shared cars. By taking the significance of the correlations into account, the meaning of the correlations can be discussed.

4.3.3.3 Method 3: Comparing regression coefficients

The third method to find the relationship between the neighbourhood characteristics that affect the demand for carsharing and the neighbourhood characteristics that affect the supply of shared cars, is to create a regression model that explains the variation in the demand for roundtrip B2C carsharing by making use of the same independent variables as have been used in the regression that explains the variation in the supply of shared cars. The regression coefficients of this new model are then compared to the regression coefficients of the regression model that explains the variation in the supply of shared cars to find whether the explanation of the variation in the demand and supply can be based on the same neighbourhood characteristics.

The dependent variable that is used to explain the variation in the individual demand is the respondents' level of interest in the use of a roundtrip B2C shared car that has been indicated on the 5-point Likert scale. Because there is a lot of discussion on how to interpret Likert scale data and what statistical tools to use when dealing with Likert scale data (Leppink, 2021), the data on the respondents' level of interest has been binned. A binary dependent variable is created in which 2 bins are possible. The first bin indicates that a respondent shows demand for roundtrip B2C carsharing and consists of the respondents that answered '4. I would probably make use of the shared car' and '5. I would definitely make use of the shared car' on the Likert scale. The other bin is indicating that a respondent does not show demand for roundtrip B2C carsharing and consists of the other 3 answers on the Likert scale. Since the dependent variable is binary, the regression technique that is used to explain the dependent variable is the same as the technique used to explain the variation in the presence of roundtrip B2C shared car in a neighbourhood, a binary logistic regression.

A binary logistic regression model is developed that explains the variation in the demand for roundtrip B2C carsharing of an individual by using the same independent variables that are used in the regression model that explains the variation in the supply of roundtrip B2C shared cars in a neighbourhood. Since this regression consists of a binary logistic regression that explains the variation in the shared car presence and a negative binomial regression that explains the variation in the shared car supply rate, three different models are developed. One model is created with the same independent variables as have been used in the binary logistic regression, one model has the same independent variables as the negative binomial regression and one model uses the independent variables used in both models. After the regressions have been developed, the signs and magnitudes of their regression coefficients are compared to the signs and magnitudes of the coefficients in the regression that explains the variation in the supply of roundtrip B2C shared cars. When these coefficients are similar, it means that the variation of the demand for roundtrip B2C carsharing can be explained by the same independent variables that explain the variation of the supply for roundtrip B2C shared cars.

4.3.3.4 Comparison between methods

After the results of the three distinct methods to find the relationship between the neighbourhood characteristics that affect the demand for carsharing and the neighbourhood characteristics that affect the supply of shared cars are discussed, the results are compared to each other. This comparison is used to draw a conclusion on what kind of relationship exists between the demand and the supply for roundtrip B2C shared cars.

4.3.4 Explanation of the variation in the demand for roundtrip B2C carsharing

To find out what factors other than neighbourhood characteristics explain the variation in the demand for roundtrip B2C carsharing, the variation in the demand is directly explained by all individual factors and neighbourhood characteristics that are framed in the conceptual model in Figure 2.1 of the inhabitants of the G44 cities in the Netherlands. The dependent variable used for this explanation is again a binary variable in which the same 2 bins that are discussed in the previous section are possible. This means that an individual is classified as showing demand or not. Because the data about the motives and purposes for using a shared car cannot be used in a regression model, the motives and purposes are analysed separately. Then the correlations between each variable and the demand for roundtrip B2C carsharing is discussed. Finally, the variables are used in a binary logistic regression model that explains the variation in the demand for roundtrip B2C carsharing.

4.3.4.1 Motives and purposes

The motives and purposes to use a shared car are obtained by asking the respondents to indicate their level of agreement about statements like 'I am using / I am considering using a shared car to avoid maintenance and repair'. Therefore, the correlation between the level of agreement on each statement and the demand for carsharing does not explain anything about the variation in the demand for roundtrip B2C carsharing. Therefore, the motives and purposes are discussed separately.

The respondents have indicated their level of interest in using a shared car for 6 different motives and 5 different purposes on a 5-point Likert scale. The distributions of their answers are analysed by comparing them to each other. Also, a distinction is made between respondents that have experience with a shared car and respondents that do not have experience. The 6 motives and 5 purposes that the analysis is based on are listed in the third section of Appendix C.

4.3.4.2 Correlations of the individual factors and neighbourhood characteristics with the demand for roundtrip B2C carsharing

To be able to explain the variation in the demand for roundtrip B2C carsharing, first, the correlation between the demand and all individual and neighbourhood characteristics framed in Figure 2.1 are analysed. There are a total of 65 variables that represent these factors and characteristics. Data on these variables all result from the survey and the variables are listed in Appendix E. The correlation of the variables with the strongest correlation with the demand are plotted in a graph and based on these variables, it is known what factors other than neighbourhood characteristics are important when providing an explanation of the variation in the demand in different neighbourhoods in the G44 cities.

4.3.4.3 Explanation of the variation in the demand for roundtrip B2C carsharing

The 65 independent variables and their correlations with the demand are used to explain the variation in the demand for roundtrip B2C carsharing. The independent variable is the binary variable that represents the demand of an individual. A binary logistic regression is used to explain the variation in the demand for the same reason as discussed in Section 4.3.1.2. Therefore, also the process of finding the most effective fit is the same as elaborated on in that section. Likewise, a training set/validation set ratio of 80/20 is used. This means that 496 of the 620 respondents are used to base the model on. The independent variables that are in the most effective model, also show what characteristics other than neighbourhood characteristics are important when providing an explanation of the variation in the demand in different neighbourhoods in the G44 cities.

Chapter 5

Descriptive analysis

This chapter presents a descriptive analysis of the shared car supply rate and the survey that is the main data source in this research. First, the distribution of the shared car supply over the neighbourhoods in The Hague, Amsterdam and Utrecht is analysed. Then the representativeness of the survey is described.

5.1 Analysis of the distribution of the shared car supply

The dependent variable in the explanation of the variation in the supply, the number of roundtrip B2C shared cars per 100,000 inhabitants, has been visualized to analyse the distribution of the shared car supply rate over the different neighbourhoods in The Hague, Amsterdam and Utrecht. This is done to identify possible extra spatial independent variables that affect the shared car supply rate in a neighbourhood but have not been discussed by literature. The distribution of the shared car supply rate in Amsterdam is shown in Figure 5.1 and has been visualized by making use of ArcGIS. The distribution of the observed shared car supply rate for The Hague and Utrecht can be found in Appendix F.

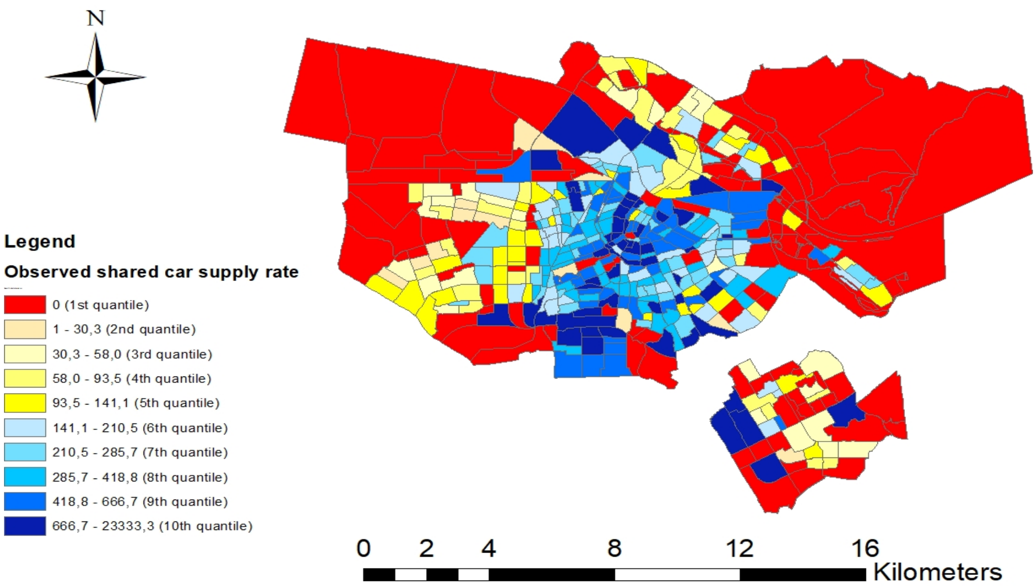

FIGURE 5.1: The distribution of the observed shared car supply rate in Amsterdam

In Figure 5.1 and Appendix F it is clear to see that the closer a neighbourhood is located to the city centre, the higher the roundtrip B2C shared car supply rate in that particular neighbourhood. Despite that the distance to the city centre might be a proxy for the address density or for another variable, it might be of added value to the explanation of the variation in the supply of shared cars. Therefore, in addition to the independent variables in Table 4.1, also the distance to the city centre is used as an independent variable to base the regression models on in the next chapter. Data on this additional independent variable has been collected by calculating the Euclidean distance from the centroid of each neighbourhood to the city centre of the corresponding city and dividing this value by the average distance to the city centre of the corresponding city in ArcMap.

5.2 Representativeness of the survey

To obtain a sample that represents the inhabitants of the G44 cities, the survey is distributed among 6 of the 44 municipalities. In Table 5.1, spatial, sociodemographic and socioeconomic characteristics of the G44 cities and the 6 cities in which the survey is distributed are presented. These municipalities are Almere, Arnhem, Enschede, Nijmegen, Zoetermeer and Zwolle. The table shows that these 6 municipalities together have similar characteristics and are therefore a good representation of the cities in the G44 cities. To make sure that enough data about neighbourhoods with a high supply and with a low supply of shared cars is obtained, 3 cities with a relatively high and 3 cities with a relatively low number of B2C shared cars per 100,000 inhabitants are chosen. The cities with a high number of B2C shared cars per inhabitant are Arnhem, Nijmegen and Zoetermeer. The cities with a low number of B2C shared cars per inhabitant are Almere, Enschede and Zwolle. The locations of these municipalities are shown in Figure 5.2.

The survey is distributed by making use of PanelClix. PanelClix is an ISO certified online panel that can deliver respondents that fit into any desired profile. Data of 631 respondents have been collected. Because the zip code of 11 respondents turned out invalid and the zip code is used to collect a large amount of data, the sample in this research consists of 620 respondents. These 620 respondents are almost equally divided over the 6 municipalities by closing the survey for respondents from a municipality of which already enough data was gathered. Since the survey focuses on the demand for carsharing, it is desired that all respondents are able to make use of a shared car. Therefore, all respondents fit the profile of an individual who is at least 18 years old and owns a car driver's license. The spatial, sociodemographic and socioeconomic characteristics of the sample that consists of the 620 respondents are presented in Table 5.1 next to the CBS data about the characteristics of the G44 cities and the 6 municipalities to be able to find the representativeness of the sample.

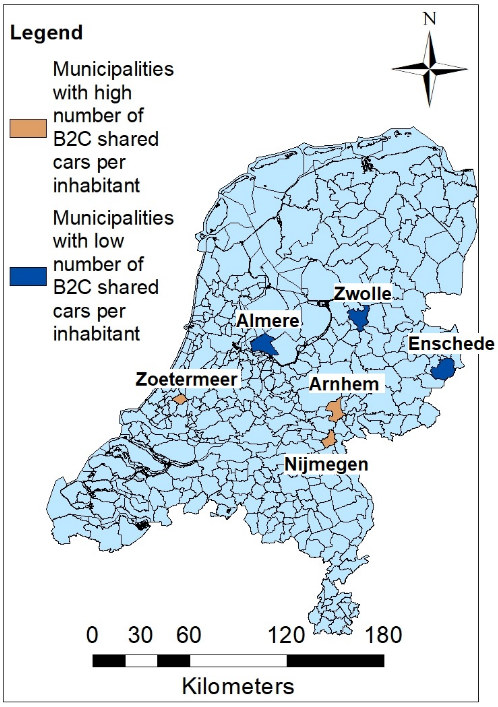

FIGURE 5.2: The municipalities in which the survey is distributed

TABLE 5.1: Representativeness of samples

Characteristics	**CBS G44 (CBS, 2020b)**	**CBS survey municipalities (CBS, 2020b)**	**Sample**
Spatial characteristics			
Number of inhabitants in municipality	170,808	160,778	160,830
Number of addresses per km^2	2,379	2,157	2,091
Sociodemographic characteristics			
Percentage of individuals between 0-14 years old	15%	16%	N/A
Percentage of individuals between 15-24 years old[1]	13%	14%	5%
Percentage of individuals between 25-44 years old	26%	28%	30%
Percentage of individuals between 45-64 years old	27%	26%	39%
Percentage of individuals 65 years old or older	18%	16%	25%
Percentage of male individuals	50%	50%	50%
Percentage of female individuals	50%	50%	50%
Percentage of low educated individuals	21%	20%	15%
Percentage of moderately educated individuals	31%	31%	44%
Percentage of highly educated individuals	25%	26%	38%
Total percentage of individuals with an education level	*78%*	*77%*	*97%*
Percentage of individuals with a western migration background	12%	11%	8%
Percentage of individuals with a non-western migration background	17%	20%	5%
Percentage of individuals living in rental houses	47%	48%	39%
Percentage of individuals living in single-person households	42%	43%	30%
Percentage of individuals living in multiple-person households with children	31%	32%	26%
Percentage of individuals living in multiple-person households without children	27%	25%	40%
Socioeconomic characteristic			
Average number of private cars per household	0.99[2]	0.87[2]	1.18

[1]For the sample this variable only consists of individuals aged between 18 and 24 years old.
[2]There is no data available on the average number of private cars per household in Almere.

Because according to the CBS (2020b), approximately 7,5 million people are living in the G44 cities, a sample size of 620 respondents is enough to obtain a sample that represents the G44 cities with a confidence interval of 95% and a margin of error of 4% (SurveyMonkey, 2021). However, Table 5.1 shows that some groups are over- or underrepresented in the sample.

Only 5% of the individuals in the sample have a non-western migration background, whereas 17% of the inhabitants of the G44 cities are individuals with a non-western migration background. Therefore, this group is underrepresented in the sample. Individuals with a western migration background are also a little underrepresented, but this deviation is relatively small. Also, the individuals living in a single-person household are underrepresented. Furthermore, the individuals living in a multiple-person household without children are overrepresented. The male/female ratio in the sample is similar to the male/female ratio in the G44 cities. Also, the deviations in the representation of individuals living in a rental house and in the representation of individuals living in a multiple-person household with children in the sample are relatively small.

Since the sample consists of only individuals who are at least 18 years old and own a car driver's license, the percentages for the different age groups are hard to compare. Because the age group of individuals between 0 and 14 years old is empty and also part of the individuals in the age group between 15 and 24 years old were not allowed to complete the survey, the percentages of the other age groups in the sample are automatically higher when compared to the percentages for the G44 cities. The percentage of individuals between 0 and 14 years old in the G44 cities is 15%. When correcting for this 15% in the number of individuals between 25 and 44 years old, 45 and 64 years old and 65 years old or older in the sample, the percentages for these age groups become 25%, 33% and 21% respectively. Then only the individuals between 45 and 64 years old show a significant overrepresentation when compared to the characteristics of the G44 cities. Furthermore, only 5% of the individuals in the sample are between 15 and 24 years old. Although this could partly be caused by the fact that individuals between 15 and 18 years old were not allowed to complete the survey, this age group still seems underrepresented. Therefore, it seems that individuals between the age of 45 and 64 years old are overrepresented and individuals between the age of 18 and 24 years old are underrepresented in the sample.

Because all individuals in the sample are 18 years old or older, the total percentage of individuals with an education level is higher in the sample. This difference can be corrected by dividing the percentages about the education levels in the sample by the total percentage of individuals with an education in the sample and multiplying it by the total percentage in the G44 cities. The values for the percentages of low educated, moderately educated and highly educated individuals in the sample then become 12%, 35% and 30% respectively. After this correction, it can be argued that low educated individuals are underrepresented in the sample when compared to the values for the G44 cities.

Also, the average number of private cars per household is overrepresented in the sample. This could be caused by the constraint that each respondent must have a car driver's license to complete the survey because individuals without a car driver's license are less likely to own a car. However, this overrepresentation can also result from the missing data on the average number of private cars per household in Almere. The CBS (2020b) does include information on the number of private cars and the number of households and therefore the average number of private cars per household in Almere can be calculated manually. However, this results in an average of more than 3 private cars per household, which seems incorrect. That could be the reason why the CBS did not include this number in their database. Nevertheless, the high number of private cars is an indication of a relatively high number of private cars per household in Almere, which would have led to a higher average in the G44 cities and a smaller overrepresentation in the sample.

In conclusion, the sample size is large enough to represent the G44 cities, however, some characteristics are over- or underrepresented in the sample.

Chapter 6

Explanation of the variation in the supply of roundtrip B2C shared cars in neighbourhoods in the G44 cities in the Netherlands

This chapter provides an answer to the first research question by discussing the results of the regression models that explain the variation in the supply of roundtrip B2C shared cars in the neighbourhoods in the G44 cities in the Netherlands. By comparing the model based on the neighbourhoods in The Hague, Amsterdam and Utrecht with the model based on the municipalities that represent the G44 cities, the most effective explanation of the variation is obtained. The first section is about the correlation of the neighbourhood characteristics with the observed shared car presence and the observed shared car supply rate. Then the results of both models are discussed.

6.1 Correlation of the neighbourhood characteristics with the shared car presence and shared car supply rate

This section presents the correlation of the neighbourhood characteristics with the shared car presence and shared car supply rate in The Hague, Amsterdam and Utrecht and in the survey municipalities. Besides that these correlations are the input for the regression models, the correlations also provide an answer to the first hypothesis, in which is expected that there is more supply of roundtrip B2C shared cars in neighbourhoods that consist of a large number of characteristics. These characteristics can be found in Section 3.4.

6.1.1 Comparison between the correlations in The Hague, Amsterdam and Utrecht and the survey municipalities

In this section, the correlations between the roundtrip B2C shared car presence and each independent variable, and the roundtrip B2C shared car supply rate (number of roundtrip B2C shared cars per 100,000 inhabitants) and each independent variable is presented. The correlations of the neighbourhood characteristics with the shared car presence and the shared car supply rate are distinctively plotted in Figure 6.1 and Figure 6.2 respectively. In these figures, the correlations are presented for The Hague, Amsterdam and Utrecht and for the survey municipalities. The neighbourhood characteristics are sorted by the strength of the correlation in the survey municipalities. The significance of each correlation is presented in Appendix G.

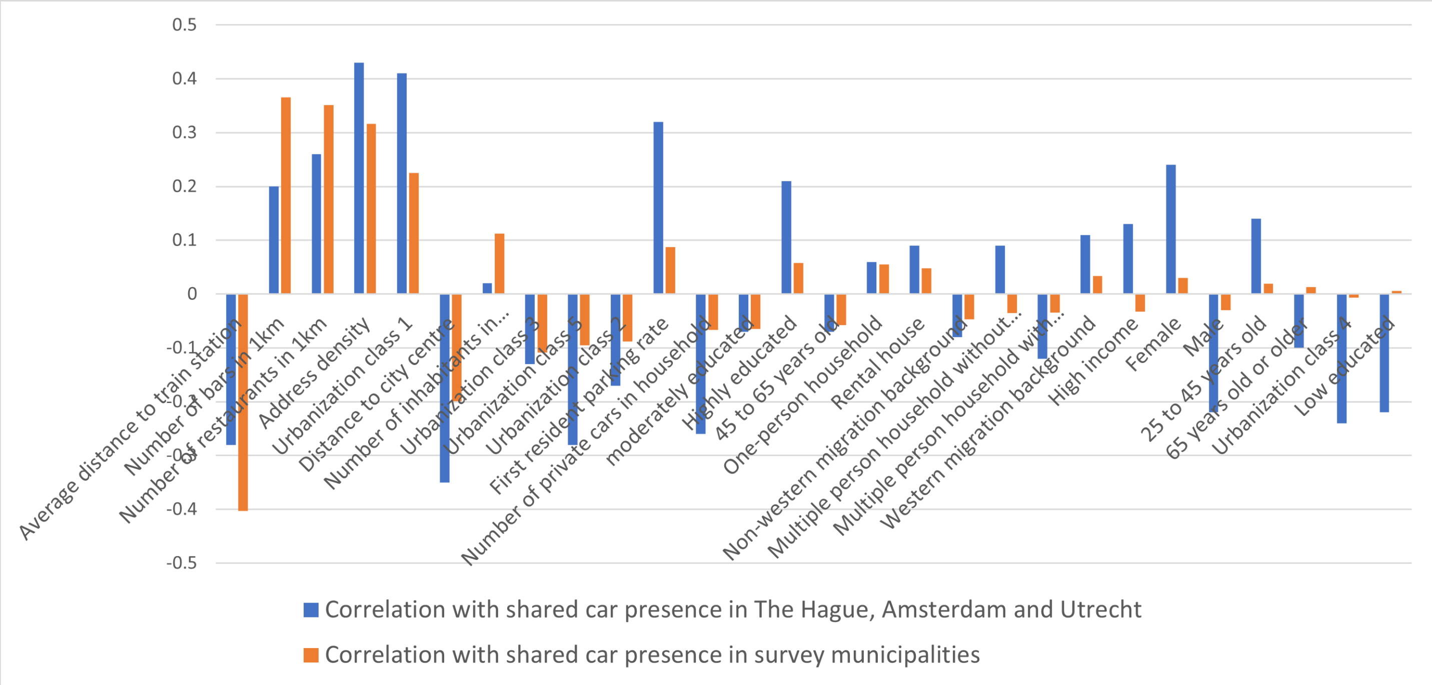

FIGURE 6.1: Comparison between the correlation of the neighbourhood characteristics with the shared car presence in the survey municipalities and in The Hague, Amsterdam and Utrecht

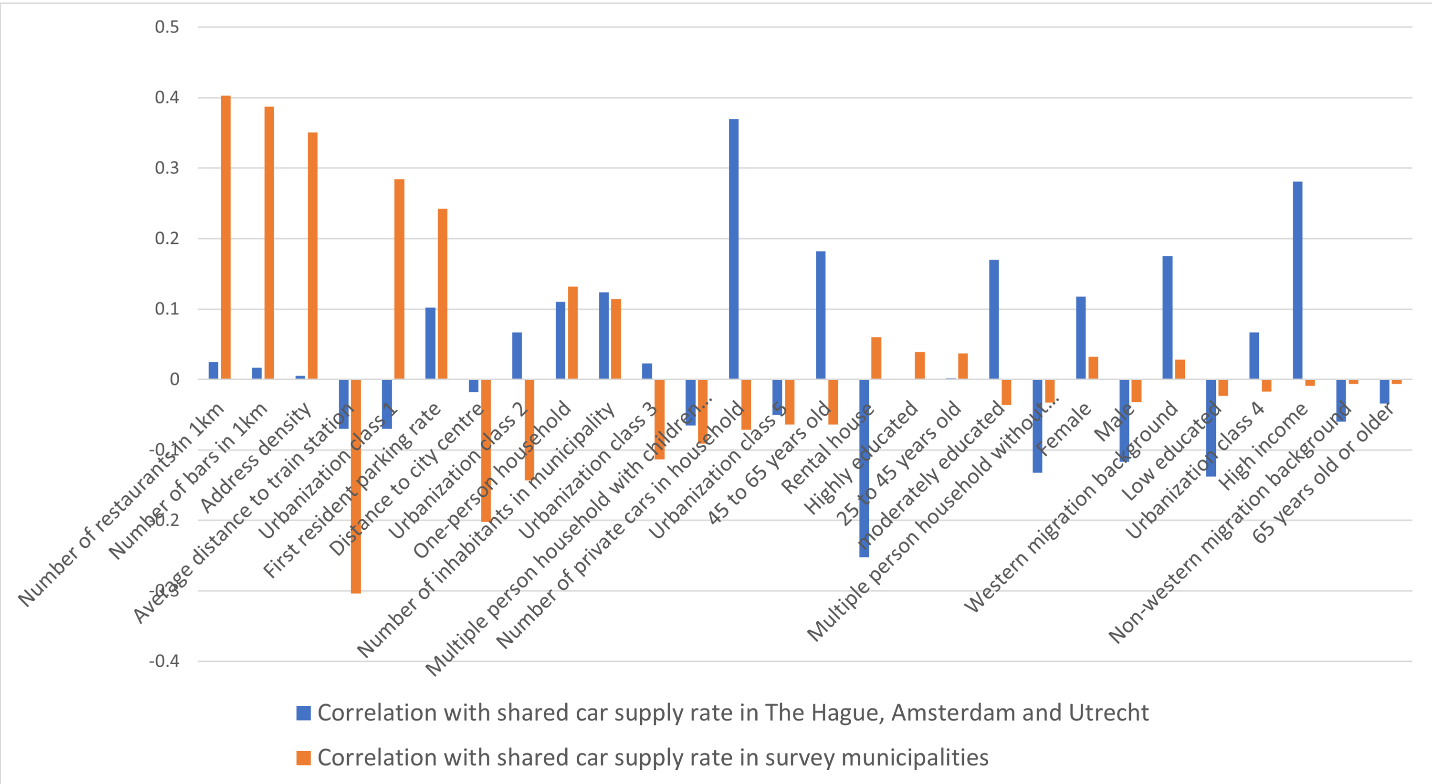

FIGURE 6.2: Comparison between the correlation of the neighbourhood characteristics with the shared car supply rate in the survey municipalities and in The Hague, Amsterdam and Utrecht

The correlation of the neighbourhood characteristics with the shared car presence in both groups of municipalities is shown in Figure 6.1. The figure shows that almost all signs are the same for each neighbourhood characteristic and also the magnitude of the correlations are comparable in most cases. However, the more the neighbourhoods factors are located on the right side of the figure, the smaller the correlation with the shared car presence in the survey municipalities. Appendix G shows that most of the time, these smaller correlations also result in a smaller significance. Therefore, the comparison of the signs of these correlations is not as significant as the comparison of the signs of the correlation on the left side of the figure. Nevertheless, the correlations of the neighbourhood characteristics with the shared car presence in both groups of municipalities seem comparable.

The correlations of the neighbourhood characteristics with the shared car supply rate are shown in Figure 6.2. This figure shows that the magnitude of the correlations of each neighbourhood characteristic with the shared car supply rate is different in both groups of municipalities. This could mean that the effect of most neighbourhood characteristics on the shared car supply rate is different in the survey municipalities than it is in The Hague, Amsterdam and Utrecht. However, Appendix G shows that the significance of most correlations between the neighbourhood characteristics and the shared car supply rate in the survey municipalities is low. Therefore, the difference in the correlations in Figure 6.2 could also be explained by the fact that there are fewer neighbourhoods with a shared car and there is a lower number of shared cars per inhabitant in the survey municipalities. This could result in a lack of variation in the shared car supply rate and therefore unreliable correlations between the neighbourhood characteristics and the shared car supply rate in the survey municipalities.

6.1.2 Testing hypothesis 1

The correlations between the neighbourhood characteristics and the shared car presence and shared car supply rate in the survey municipalities are less significant than the correlations in The Hague, Amsterdam and Utrecht. Therefore, hypothesis 1, in which is expected that there is more supply of roundtrip B2C shared cars in neighbourhoods that consist of certain characteristics, can best be tested based on the correlations between the neighbourhood characteristics and the supply of shared cars in The Hague, Amsterdam and Utrecht. Because the figures in the previous section are sorted by the strength of the correlation in the survey municipalities, new figures are created to provide a clear overview of the neighbourhood characteristics with the strongest correlation with the shared car presence and shared car supply rate in The Hague, Amsterdam and Utrecht. The correlations of the neighbourhood characteristics with the shared car presence and shared car supply rate are presented in Figure 6.3 and Figure 6.4. The figures are sorted by the strength of the correlation in The Hague, Amsterdam and Utrecht. The significance of each correlation is presented in Appendix G.

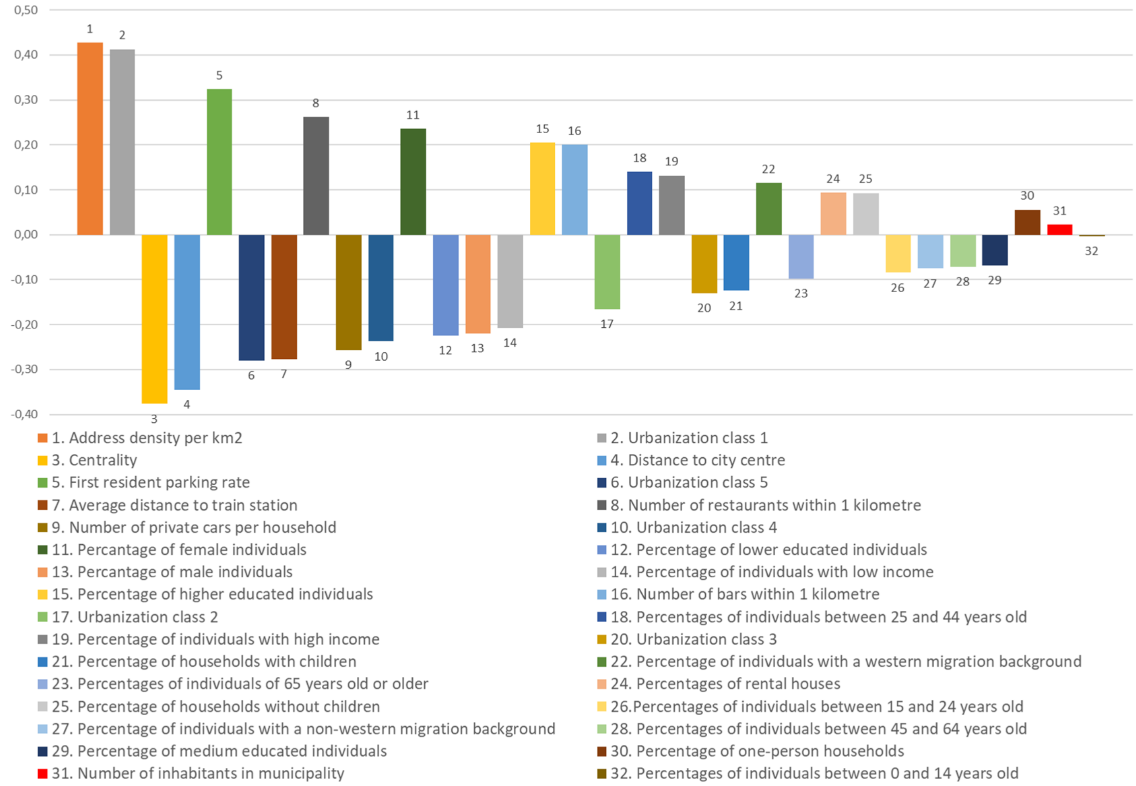

FIGURE 6.3: Correlation of the independent variables with the presence of a shared car in a neighbourhood

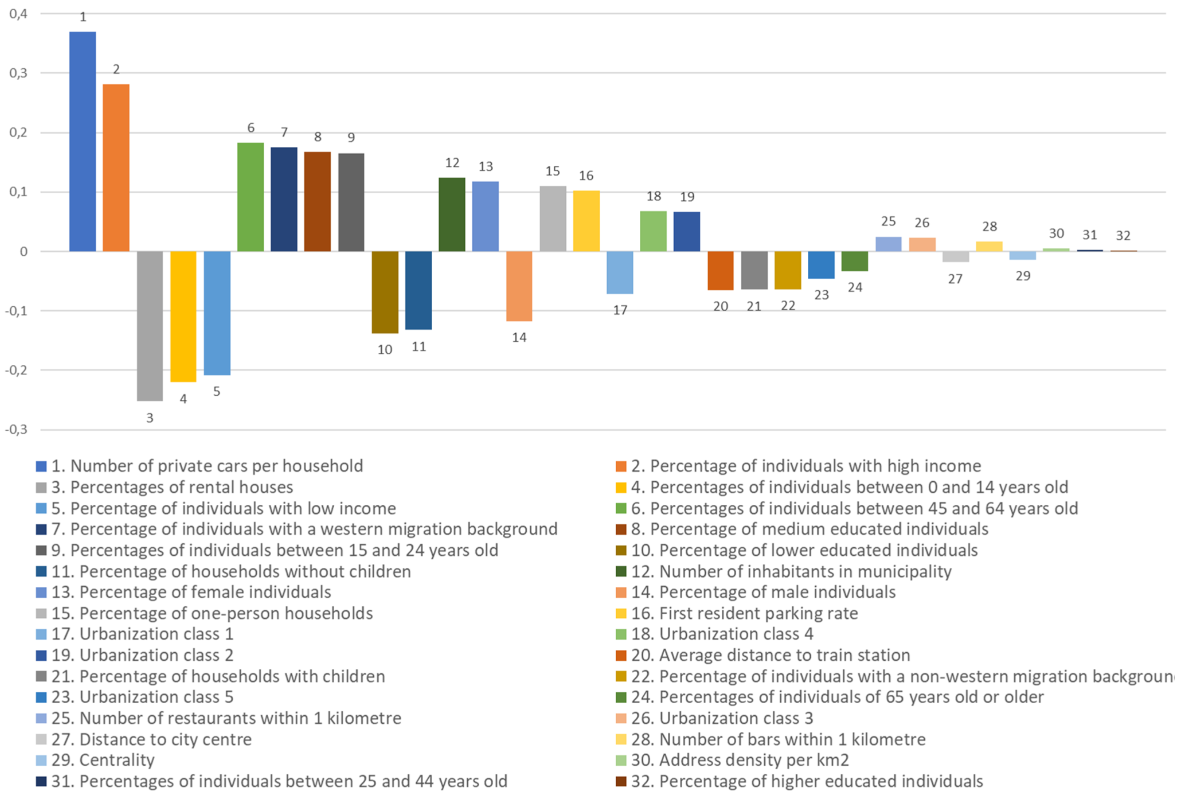

FIGURE 6.4: Correlation of the independent variables with the shared car supply rate in a neighbourhood

Figure 6.3 shows that the address density per km^2 and neighbourhoods in urbanization class 1 show the strongest correlation with the presence of a roundtrip B2C shared car in a neighbourhood. Individuals in neighbourhoods with urbanization class 1 make relatively more use of non-car forms of urban transport and have a relatively high share of walking in their modal split (CBS, 2017). Also, the distance to the city centre, which has been added as a variable based on the current shared car distribution, shows a strong negative correlation. This means that it is more likely that there is a shared car placed in a neighbourhood that is closer to the city centre. All of the above is in line with the expectations in hypothesis 1. The average age in a neighbourhood and the number of inhabitants in a municipality, however, were expected to be more important for the presence of a shared car in a neighbourhood.

Figure 6.4 shows that the number of private cars per household and the percentage of individuals with a high income are important factors for the number of shared cars per inhabitant in a neighbourhood. However, the number of private cars per household has a positive correlation, where a negative correlation with the supply of shared cars was expected according to hypothesis 1. Other independent variables with a strong correlation are the percentage of rental houses and the percentage of individuals between 0 and 14, and 45 and 64 years old. The percentage of rental houses has an unexpected effect on the supply of shared cars as well. It is expected that a large percentage of rental houses leads to a high supply, however, this is not the case for its correlation with the shared car supply rate. Since the hypothesis about the age was based on the effect of the average age in a neighbourhood, not only the correlation with the percentage of individuals between 0 and 14, and 45 and 64 years old must be compared, but the correlation with all age groups. Figure 6.4 shows that the percentage of individuals between 0 and 14 years old has a negative effect on the shared car supply rate in a neighbourhood. It could be that the expected positive effect of a young age on the supply of shared cars is not visible in this age group since these individuals are not allowed to drive a car yet. The figure shows a comparable positive correlation with the shared car supply rate and the percentages of individuals between 15 and 24, and 45 and 64 years old. There is a small negative correlation with the percentage of individuals that are 65 years old or older and an almost negligible correlation with the percentages of individuals between 25 and 44 years old. The small negative correlation with the percentage of individuals that are 65 years old or older and a positive correlation with the younger age groups indicate a small positive effect of a young age on the supply of shared cars. Other interesting particularities arise between the correlations about the individuals with a migration background. Figure 6.4 shows that the percentage of individuals with a western migration background in a neighbourhood has a strong positive correlation with the shared car supply rate. This is interesting because the percentage of individuals with a non-western migration background has a negative correlation with the number of roundtrip B2C shared cars per 100,000 inhabitants.

In conclusion, most of the expected correlations in hypothesis 1 are accepted. However, the expectation of a positive correlation between the number of residents with a migration background and the supply only holds for the number of residents with a western migration background and is rejected for the number of residents with a non-western migration background. Also, the hypothesis that having a low number of private cars per household and a large percentage of rental houses would result in a higher supply of shared cars is only accepted for the shared car presence and is rejected for correlation with the shared car supply rate in a neighbourhood.

6.2 Results explanation of the variation in the shared car supply rate

In this section, the results of the explanation of the variation in the shared car supply rate by the regressions based on The Hague, Amsterdam and Utrecht and based on the survey municipalities are discussed. First, the process of finding the combination of independent variables with the best fit is described, then the regression equations are presented. Subsequently, the regressions are validated and finally, the two explanations of the variation in the shared car supply rate are compared to each other.

6.2.1 Process of finding the combination of variables with the best fit

To find the independent variables that describe the variation in the presence of roundtrip B2C shared cars and the variation the shared car supply rate the best, the process described in Section 4.3.1.3 has been followed for neighbourhoods in The Hague, Amsterdam and Utrecht and for the survey municipalities.

In the previous section, the independent variables with the strongest correlation with the presence of a roundtrip B2C shared car and with the shared car supply rate are presented for both groups of municipalities. The variables with the strongest correlation in the figures for the shared car presence are the first variables that are inserted in the regressions. Then the next variables with a strong correlation have been added one after the other and each time a new regression is fitted with the new combination of variables. The binary logistic regressions are assessed on their Nagelkerke R^2 and the significance of its variables. The negative binomial regressions are assessed on their AIC and BIC and also on the significance of the independent variables. The best models are optimized by adding new variables that could improve the models. The steps that have been taken to find the most effective binary logistic and negative binomial regression models based on The Hague, Amsterdam and Utrecht and on the survey municipalities are shown in Appendix H. In the tables in the appendix, the combination of independent variables is shown with their corresponding p-value to show the significance of each variable. Also, the Nagelkerke R^2 or AIC and BIC of each combination is presented. By making use of superscripts, the independent variables that are insignificant and the regressions with only significant variables are indicated. The most effective model of the binary logistic regressions and negative binomial regressions for The Hague, Amsterdam and Utrecht and for the survey municipalities is made bold in Appendix H. Also, independent variables with a smaller correlation than the independent variables in the tables were included but this did not result in a better fit. To be certain that the independent variables in the models do not show multicollinearity, also the VIF of the independent variables is calculated. Since the VIF for all independent variables in the regressions are below 5, there is no multicollinearity in the models.

To develop the negative binomial regression models, all neighbourhoods in which the binary logistic regression predicts the presence of a roundtrip B2C shared car have been used as input. This means that the data set for the negative binomial regression in The Hague, Amsterdam and Utrecht consists of 534 neighbourhoods. The data set for the negative binomial regression in the survey municipalities consists of 77 neighbourhoods.

Since the AIC and BIC did not provide a clear distinction between the fits of different models to explain the shared car supply rate in the survey municipalities, also the log-likelihood is used as a method to measure the goodness of fit. Model 17 in Table H.4 in Appendix H turned out to be the model with the best fit because this model has a relatively low AIC and BIC, and the lowest log-likelihood.

6.2.2 Regression equations

After the process of finding the best combination, the regression equations that correspond to the best combination of variables for each model is presented in this section. The regression coefficients of the models based on The Hague, Amsterdam and Utrecht and the models based on the survey municipalities are combined in one table to be able to compare the equations. First, the regression equations of the binary logistic regressions are discussed, and then the regression equations of the negative binomial regressions are presented and compared.

6.2.2.1 Regression equations of the binary logistic regressions

The binary logistic regressions that are most effectively explaining the variation in the shared car presence consist of the independent variables and their corresponding regression coefficients that are presented in Table 6.1. The regression equations have the in Equation 4.1 and Equation 4.2 presented structure. The table shows independent variables and their corresponding regression coefficients for the model based on The Hague, Amsterdam and Utrecht and for the model based on the survey municipalities.

TABLE 6.1: Independent variables in the binary logistic regression equations

Independent variable $[X_p]$	**Regression coefficient in the binary logistic regression based on The Hague, Amsterdam and Utrecht** $[b_p]$	**Regression coefficient in the binary logistic regression based on the survey municipalities** $[b_p]$
Constant	-7.039	-7.118
Address density	0.381	
Percentage of female individuals	9.295	14.684
Percentage of individuals with high income	3.347	
Urbanization class 1	1.627	
Urbanization class 2	1.194	
Urbanization class 3	1.040	
Urbanization class 4	-1.018	
Average distance to city centre		-1.492
Average distance to train station		-0.446
Percentage of highly educated individuals		2.788

Table 6.1 shows that only the percentage of female individuals in a neighbourhood come forward in the most effective model to explain the variation in the presence of a roundtrip B2C shared car based on both groups of municipalities. The regression coefficient of the percentage of female individuals is very similar in both models.

The equation for the regression based on The Hague, Amsterdam and Utrecht shows a regression coefficient for urbanization class 1 to 4 and no coefficient for urbanization class 5. Since the urbanization class of a neighbourhood is a categorical variable it needs a reference variable. This means that the regression coefficient of urbanization class 1 to 4 represents the presence of a shared car compared to the reference variable, urbanization class 5. Table 6.1 shows that the higher the urbanization class of a neighbourhood, the higher the predicted classification value. However, urbanization class 4 results in a lower predicted classification value than urbanization class 5. The equation also shows that the address density, percentage of female individuals and the percentage of individuals with a high income in a neighbourhood have a positive effect on the predicted classification value to classify a neighbourhood as being a neighbourhood with the presence of a roundtrip B2C shared car.

The equation for the regression based on the survey municipalities shows that the percentage of highly educated individuals and the previously discussed percentage of female individuals have a positive effect on the predicted classification value. The average distance to the city centre and train station have a negative effect. This means that the closer a neighbourhood is located to the city centre and the smaller the average distance of a neighbourhood to a train station, the higher the predicted classification value to classify a neighbourhood as being a neighbourhood with the presence of a roundtrip B2C shared car.

As discussed in Section 4.3.1.2, the classification cutoff value is used to classify a neighbourhood as a neighbourhood with or without the presence of a shared car. A neighbourhood with a predicted classification value above the cutoff value is classified as a neighbourhood with the presence of a roundtrip B2C shared cars and a neighbourhood with a predicted value below the cutoff value is classified as a neighbourhood without the presence of a shared car. The cutoff value is determined by choosing a value that ensures that there is an equal number of false positives and false negatives. For the model based on The Hague, Amsterdam and Utrecht, this is the case with a cutoff value of 0.655. For the model based on the survey municipalities, the false positives and false negatives are equal when a cutoff value of 0.4 is used.

6.2.2.2 Regression equations of the negative binomial regressions

The negative binomial regressions that are most effectively explaining the variation in the shared car supply consist of the independent variables and their corresponding regression coefficients that are presented in Table 6.1. The regression equations have the in Equation 4.3 presented structure. In the equations, t_i is the number of inhabitants per neighbourhood with a fixed coefficient of one. The table shows independent variables and their corresponding regression coefficients for the model based on The Hague, Amsterdam and Utrecht and for the model based on the survey municipalities.

TABLE 6.2: Independent variables in the negative binomial regression equations

Independent variable $[X_p]$	**Regression coefficient in the negative binomial regression based on The Hague, Amsterdam and Utrecht** $[b_p]$	**Regression coefficient in the negative binomial regression based on the survey municipalities** $[b_p]$
Dispersion parameter	1.220	13.338
Constant	-2.352	-52.147
Number of inhabitants in neighbourhood $[t_i]$	1.000	1.000
Percentage of individuals with high income	3.956	
Percentage of individuals between 0 and 14 years old	-7.674	-11.293
Percentage of households without children	-5.308	
Number of inhabitants in municipality	0.310	
Average distance to train station	-0.238	
Percentage of individuals with a non-western migration background	-1.753	
Distance to city centre	-0.459	-1.760
Urbanization class 1		29.641
Urbanization class 2		30.975
Urbanization class 3		30.586
Urbanization class 4		30.781
Number of restaurants in 1 km		0.036
Address density		0.161
Percentage of highly educated individuals		5.574
Percentage of individuals between 25 and 44 years old		-16.458
Percentage of individuals with a low income		5.124
Percentage of individuals between 15 and 24 years old		-1.416
Percentage of individuals between 45 and 64 years old		10.736
Percentage of female individuals		32.048

The first thing that points out when comparing both regressions in Table 6.2, is the different dispersion parameter. The dispersion parameter in the model based on the survey municipalities is much higher than it is in the model based on The Hague, Amsterdam and Utrecht. The higher the dispersion parameter, the more the model is correcting for overdispersion. This means that the variation in the shared car supply rate in the survey municipalities is much higher than the average shared car supply rate in these municipalities. Also, there are not many independent variables that occur in both regressions. Besides the number of inhabitants in a neighbourhood that is applied in the regression as the offset variable and is fixed at 1, only the percentage of individuals between 0 and 14 years old and the distance to the city centre occur in both regressions. The regression coefficient for these independent variables in both regressions is comparable. However, in the model based on the survey municipalities, the magnitude of the coefficients are larger. When comparing all magnitudes, it points out that the magnitudes of the regression coefficient in the model based on the survey municipalities are generally larger as well.

The equation for the regression based on The Hague, Amsterdam and Utrecht shows that neighbourhoods with more individuals with a high income and more inhabitants in the complete municipality lead to a higher number of roundtrip B2C shared cars per 100,000 inhabitants. It also shows the negative effect of a large percentage of individuals between 0 and 14 years old and a high percentage of households without children on the shared car supply rate. Also, a higher share of individuals with a non-western migration background leads to a lower shared car supply rate. Finally, a smaller average distance to a train station and a smaller distance to the city centre is leading to more roundtrip B2C shared cars per 100,000 inhabitants.

Table 6.2 shows that besides the regression coefficients in the model that is based on the survey municipalities being much larger than the regression coefficients in the model based on The Hague, Amsterdam and Utrecht, there are also a lot more independent variables in this model. The regression coefficients of the different urbanization classes do not deviate much from each other. This means that according to the model based on the survey municipalities, the urbanization class of a neighbourhood affects the shared car supply rate minimally. Furthermore, the percentage of highly educated individuals, the percentage of individuals with a low income, the percentage of individuals between 45 and 64 years old and the percentage of female individuals are positively affecting the shared car supply rate. The percentage of individuals between 0 and 14, between 15 and 24 and between 25 and 44 years old, however, have a negative effect on the shared car supply rate. This means that the model that explains the variation in the shared car supply rate in the survey municipalities indicates that a younger age has a negative effect on the shared car supply rate. Finally, the model based on the survey municipalities shows that the smaller the distance of a neighbourhood to the city centre, the higher the shared car supply rate.

6.2.3 Regression validation

The developed regression models are validated in the way described in Section 4.3.1.4. Because there are some differences in the validation of the binary logistic regression and the negative binomial regression, both validations are distinctively discussed.

6.2.3.1 Validation of the binary logistic regressions

To evaluate the results of the binary logistic regression, the classification table is used. The classification tables of the regression based on The Hague, Amsterdam and Utrecht and based on the survey municipalities are presented in Table 6.3 and Table 6.4 respectively. This table shows the performance of the models in the municipalities that they are based on. The classification cutoff values that the tables are based on are 0.655 and 0.4 respectively. After these classification tables are discussed, the classification table of the predictions of the model that is based on The Hague, Amsterdam and Utrecht, applied in the survey municipalities is presented.

TABLE 6.3: Classification table of binary logistic regression model based on The Hague, Amsterdam and Utrecht

	Predicted neighbourhoods without shared car	**Predicted neighbourhoods with shared car**	*Percentage correct*
Observed neighbourhoods without shared car	81	55	*59.6%*
Observed neighbourhoods with shared car	55	369	*87.0%*
			80.4%

TABLE 6.4: Classification table of the binary logistic regression based on the survey municipalities

	Predicted neighbourhoods without shared car	**Predicted neighbourhoods with shared car**	*Percentage correct*
Observed neighbourhoods without shared car	191	32	*85.7%*
Observed neighbourhoods with shared car	32	39	*54.9%*
			78.2%

As shown in Table 6.3, the model based on The Hague, Amsterdam and Utrecht is correctly predicting the variation in the presence of a roundtrip B2C shared car in a neighbourhood for 80.4% of all predicted neighbourhoods in these municipalities. This is done by making a correct prediction for 87% of all neighbourhoods with an observed presence of a shared car and a correct prediction for 59.6% of all neighbourhoods without an observed presence of a shared car.

The chi-squared value for the binary regression model based on The Hague, Amsterdam and Utrecht is 182.9. Because there are 7 degrees of freedom in the model, the significance value is smaller than 0.001. This means that the chi-square is highly significant and the model is significantly better than a random model.

Table 6.4 shows that the model based on the survey municipalities is correctly predicting the variation in the presence of a roundtrip B2C shared car in a neighbourhood for 78.2% of all predicted neighbourhoods in the survey municipalities. This is done by making a correct prediction for 54.9% of all neighbourhoods with an observed presence of a shared car and a correct prediction for 85.7% of all neighbourhoods without an observed presence of a shared car.

The chi-squared value for the binary regression model based on the survey municipalities is 73.8. Because there are 4 degrees of freedom in the model, the significance value is smaller than 0.001. This means that the chi-square is highly significant and this model is also significantly better than a random model.

The performance of the binary logistic regression has also been assessed by applying the regression equations on the validation sets and using the same classification cutoff value to validate the model. The classification tables that result from this are shown in the first section of Appendix I. The percentages in these classification tables are similar to the percentages in the initial classification tables. This means that the overfitting of the models is minimal.

The model that explains the variation in the presence of a roundtrip B2C shared car based on The Hague, Amsterdam and Utrecht has also been applied on the survey municipalities, that together represent the G44 cities. Except for the number of inhabitants in the municipality, the data on each neighbourhood characteristic in the survey municipalities is within the same range as the data on the neighbourhood characteristics in The Hague, Amsterdam and Utrecht. Therefore, using the model on the survey municipalities should not result in very deviant predictions. The classification table is shown in Table 6.5.

TABLE 6.5: Classification table of the prediction of the presence of a roundtrip B2C shared car in the survey municipalities

	Predicted neighbourhoods without shared car	**Predicted neighbourhoods with shared car**	***Percentage correct***
Observed neighbourhoods without shared car	192	97	*66.4%*
Observed neighbourhoods with shared car	42	37	*47.0%*
			62.2%

Table 6.5 shows that 62.2% of the predictions are correct. 47.0% of the predictions in the neighbourhoods with an observed presence of a roundtrip B2C shared car and 66.4% of the predictions in the neighbourhoods without an observed presence of a shared car are correct. Although these percentages are not very high, the model is performing significantly better than a null model. Also, the model is predicting no shared car presence in 63.5% of the neighbourhoods. This percentage comes close to the observed neighbourhoods without a shared car in the survey municipalities since this percentage is 78%. The 63.5% does not come close at all to the 25% of neighbourhoods without a shared car in The Hague, Amsterdam and Utrecht. This means that the model can be applied to different municipalities than it is based on.

To search for outliers in the residuals of the predictions, the predicted probability of having the presence of a shared car for each neighbourhood in both models are plotted against the observed shared car presence. These plots are presented in the second section of Appendix I with their corresponding classification cutoff value. The predicted probabilities of the neighbourhoods with a shared car in the model based on The Hague, Amsterdam and Utrecht are not below 0.35. Because the cutoff value for this model is 0.655, this means that the false negatives are not that bad. Also, there are way more predicted probabilities above the cutoff value in the neighbourhoods with an observed shared car than in the neighbourhoods without. Therefore, the model performs well. The predicted probabilities of the neighbourhoods without a shared car in the model based on the survey municipalities do not exceed 0.62. This is also close to the cutoff value of 0.4 and means that there are no large outliers in the false positives of this model. Also, there are way more predicted probabilities underneath the cutoff value in the neighbourhoods without an observed shared car than in the neighbourhoods with an observed shared car. Therefore, this model is performs good as well. The largest residuals are the predictions with a large predicted probability and an observation of 0 shared cars and the predictions with a small predicted probability and an observation of a neighbourhood with a shared car. These residuals have been analysed and no mistakes have been found in the variables that are the base of the predictions. Therefore, this analysis confirms that no neighbourhoods in The Hague, Amsterdam and Utrecht and in the survey municipalities are considered an outlier. To be sure, the binary logistic regressions have also been fitted on a data set that excludes the neighbourhoods with the highest residuals and did not show an improved result. Therefore, it can be concluded that there are no significant outliers in the models.

The predicted presence of roundtrip B2C shared cars per neighbourhood and the residuals of these predictions are also shown on the map for The Hague, Amsterdam and Utrecht in Appendix J. The figures show that most neighbourhoods where a shared car is predicted are around the city centre of each city. The figures that show the correctness of the predicted presence of a roundtrip B2C shared car show maps that are largely blue for each city. This means that the prediction of the major part of the neighbourhoods is correct. From the approximately 20% wrongly predicted neighbourhoods, there are relatively more in Amsterdam-Zuidoost and in the west part of Utrecht. The model predicts no shared car presence because these neighbourhoods are relatively far from the city centre and do not have a relatively high score on the other variables in the model. Except for these two areas, there is no other spatial pattern in the wrongly predicted neighbourhoods. Therefore, there is no reason to believe that the model results in relatively more wrongly predicted neighbourhoods when the neighbourhood is located relatively far from the city centre.

6.2.3.2 Validation of negative binomial regressions

The likelihood chi-squared value for the negative binomial regression model based on The Hague, Amsterdam and Utrecht is 214.7. Because there are 7 degrees of freedom in the model, the significance value is smaller than 0.001. This means that the chi-square is highly significant and the model is considered to be an improvement of fit over the null model. The likelihood chi-squared value for the negative binomial regression model based on the survey municipalities is only 20.3. Because there are 13 degrees of freedom in the model, the significance value is greater than 0.05. This means that the chi-square is not significant and the model based on the survey municipalities is not considered to be an improvement of fit over a null model.

Because the model based on The Hague, Amsterdam and Utrecht is an improvement of fit over a null model, also the residuals are analysed to control whether the model is not misspecified and to search for outliers in the residuals of the prediction. To control whether the model is not misspecified, the residuals have been plotted against each independent variable. These plots can be found in Appendix K. Since none of the plots shows a curvilinear relation between the residuals and the particular independent variable, the model is not misspecified.

To search for outliers in the residuals of the prediction, the studentized residuals are plotted against the predicted values. This plot is presented in Figure 6.5.

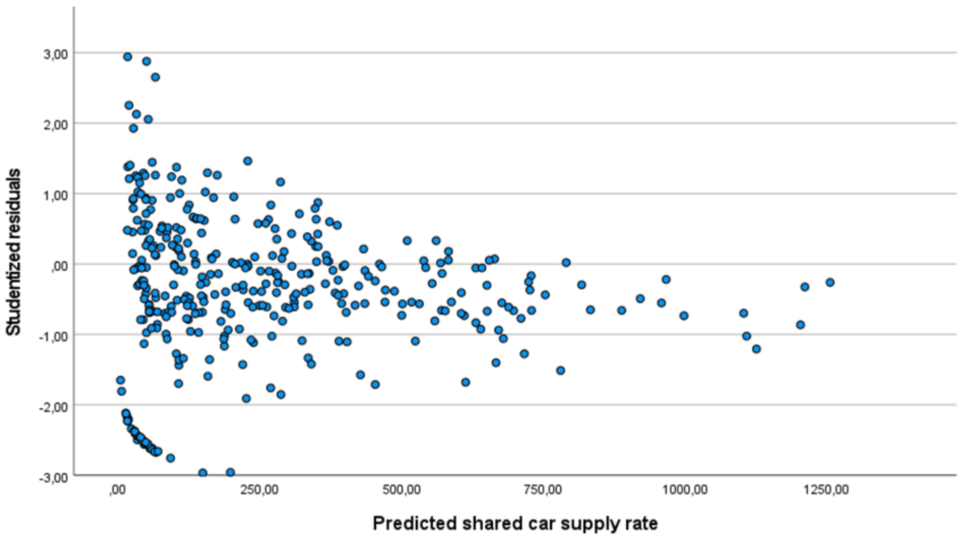

FIGURE 6.5: Plot of the studentized residuals against the predictions of the negative binomial regression

Figure 6.5 shows that all studentized residuals of the predictions of the shared car supply rate in each neighbourhood are within the interval of three but there are some predictions with studentized residuals with a higher absolute value than 2. Since the cutoff of 3 must not be taken too literally, the predictions with studentized residuals with a higher absolute value than 2 are checked for mistakes or unusual data.

To be able to discuss the studentized residuals that are higher than 2 or lower than -2, it is important to note that a studentized residual below 0 results from a higher predicted shared car supply rate than the observed shared car supply rate in a neighbourhood. When looking at the studentized residuals below -2, a pattern of exponential decay can be detected. Although the exponential decay stands out, it can be explained by the observed shared car supply rate of the neighbourhoods in this pattern. The pattern consists of all neighbourhoods in which no observed shared cars are present. Therefore, a higher prediction results in a studentized residual that deviates more from 0 and approaches -3. Because there are no mistakes in the data on these neighbourhoods and the data is not unusual, the observations in the neighbourhoods with studentized residuals lower than -2 are not classified as outliers. The observations with studentized residuals higher than 2 are caused by a high observed shared car supply rate and a low predicted shared car supply rate. Since there are also no mistakes or unusual data in these predictions, the predictions with studentized residuals higher than 2 are also not classified as outliers.

Subsequently, it points out in Figure 6.5 that the studentized residuals of the higher predicted shared car supply rates are all below zero. This means that the higher predicted values for the shared car supply rate are all higher than the observed shared car supply rate in these neighbourhoods. These higher predictions are caused by the fact that there are neighbourhoods in the data set with a shared car supply rate that is around the values of these high predictions. Hence, the model makes predictions in the same range of these observed shared car supply rates. Because the predicted shared car supply rates of these high observed shared car supply rates are not high, all studentized residuals of the higher predicted shared car supply rates are below zero. Because there are no mistakes or unusual data in the neighbourhoods with a high observed shared car supply rate but a low predicted shared car supply rate, these neighbourhoods are not classified as an outlier either.

The performance of the prediction by the negative binomial regression has also been assessed by applying the regression equation on the validation set to validate the model. The plot of the studentized residuals against the predicted values that results from this is shown in Figure 6.6.

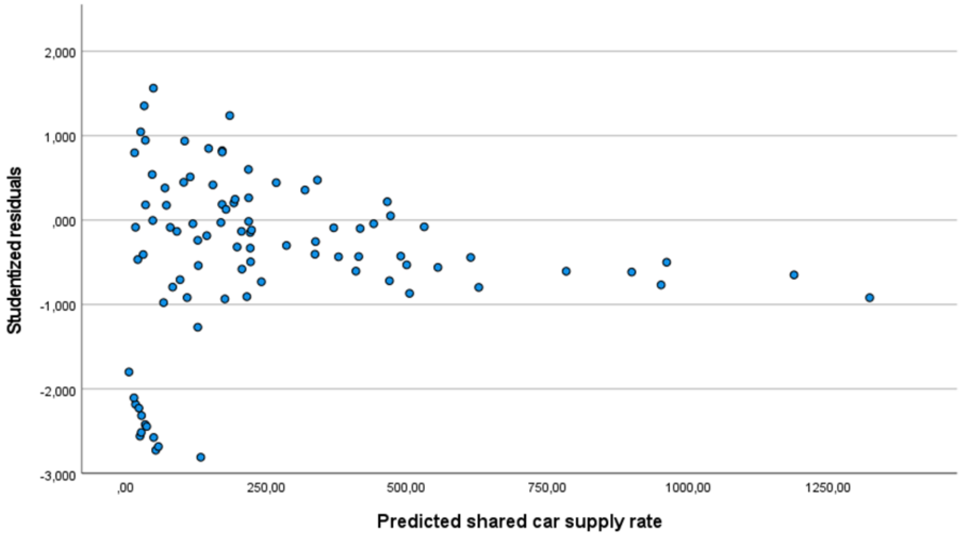

FIGURE 6.6: Plot of the studentized residuals against the predictions of the negative binomial regression validation set

Figure 6.6 shows that the plot of the studentized residuals against the prediction of the number of shared cars for the neighbourhoods in the validation set is similar to the plot of the studentized residuals and predictions in the training set of the model. The predicted values are in the same range and also the distribution of the studentized residuals is similar. This means that the overfitting of the model is minimal.

The predicted number of roundtrip B2C shared cars per 100,000 inhabitants for each neighbourhood in The Hague, Amsterdam and Utrecht are also visualized on the maps of these cities in Appendix J. The figures also show the residuals of the predictions, where a positive value means that the prediction is too low and a negative value means that the prediction is too high. It is clear to see that the model predicts a higher shared car supply rate in Amsterdam than it does in The Hague and Utrecht. This is in line with the observed shared car supply rate. Also, most of the distribution of the predicted shared car supply rate over the neighbourhoods in the municipalities is similar to the distribution of the observed shared car supply rate that is presented in Appendix F. Especially within the ring of Amsterdam a high shared car supply rate is predicted. Also in Utrecht, neighbourhoods near the city centre have a relatively higher predicted shared car supply rate than the neighbourhoods in the suburbs. The distribution of the predicted shared car supply rate over the neighbourhoods in The Hague is more spread out across the complete municipality.

Because of the high predictions in the neighbourhoods near the city centres, also the highest residuals are observed in these neighbourhoods. This could be explained by the larger prediction and larger observed shared car supply rate around the city centre. Larger observations and predictions are more likely to cause larger residuals. What also points out is that the residuals in The Hague are mostly negative, meaning that the predictions in The Hague are often too high. In Amsterdam and Utrecht, the residuals are more diverse.

6.3 Most effective explanation of the variation in the supply of roundtrip B2C shared cars in neighbourhoods in the G44 cities in the Netherlands

The models are explaining the variation in the shared car supply rate by first using a binary logistic regression to predict the presence of a roundtrip B2C shared car and then using a negative binomial regression to predict the number of roundtrip B2C shared cars per 100,000 inhabitants in a neighbourhood where the first regression predicts the presence of a shared car. In this section, both the model based on The Hague, Amsterdam and Utrecht and the model based on the survey municipalities are used to explain the variation in the supply of roundtrip B2C shared cars in the 6 municipalities that together represent the G44 cities in the Netherlands. For each neighbourhood in the 6 municipalities of which the models predict the presence of a shared car, the predicted shared car supply rate has been plotted against the observed predicted shared car supply rate in Figure 6.7 and Figure 6.9. In this way, the residuals of each prediction can be retrieved from the difference between the observed and predicted shared car supply rate. To provide a clear overview of differences in the lower shared car supply rates, Figure 6.8 and Figure 6.10 present a variant of the plots with a smaller scale. Also, exponential trend lines of the predicted shared car supply rates are added to the figures.

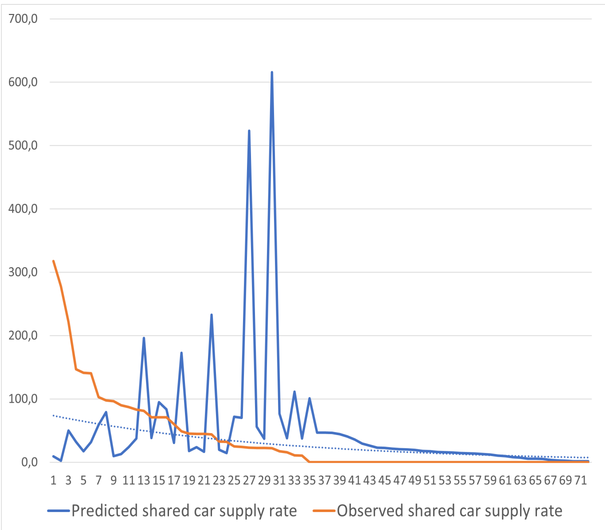

FIGURE 6.7: Observed shared car supply rate plotted against the predicted shared car supply rate in the survey municipalities for the model based on The Hague, Amsterdam and Utrecht

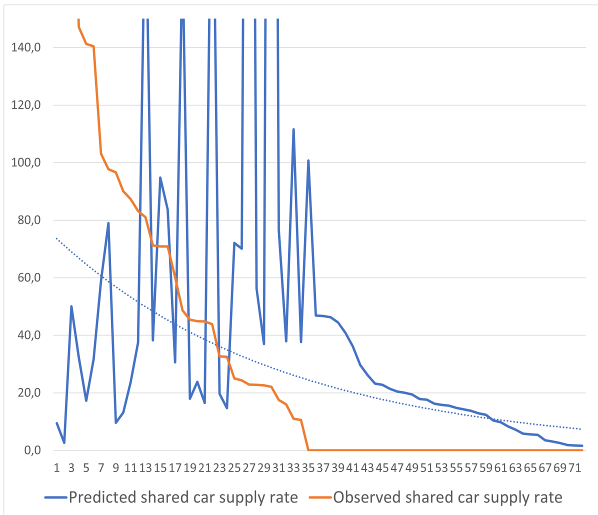

FIGURE 6.8: Observed shared car supply rate plotted against the predicted shared car supply rate in the survey municipalities for the model based on The Hague, Amsterdam and Utrecht (smaller scale)

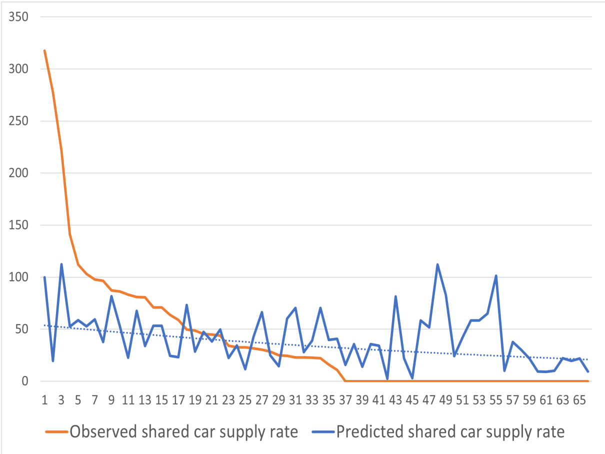

FIGURE 6.9: Observed shared car supply rate plotted against the predicted shared car supply rate for the model based on the survey municipalities

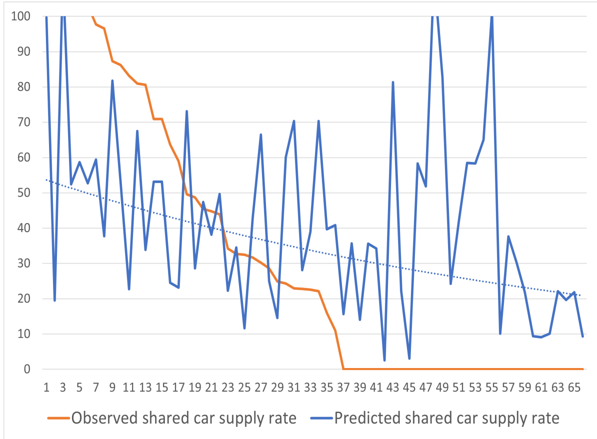

FIGURE 6.10: Observed shared car supply rate plotted against the predicted shared car supply rate for the model based on the survey municipalities (smaller scale

When comparing the predicted shared car supply rate by the regression model based on the survey municipalities in Figure 6.9 with the shared car supply rate predicted by the regression model based on The Hague, Amsterdam and Utrecht in Figure 6.7, some differences can be detected. Although the peaks in the predicted shared car supply rate in the model based on the survey municipalities are smaller and therefore deviate less from the observed shared car supply rate, the trend line deviates more from the observed shared car supply rate. The trend line of the predicted shared car supply based on the survey municipalities starts at a shared car supply rate of 50 and goes down to approximately 25. The trend line of the predicted shared car supply based on The Hague, Amsterdam and Utrecht starts at a shared car supply rate of 80, has a more similar pattern to the observed shared car supply rate and ends, just as the observed shared car supply rate at a shared car supply rate of approximately 0. This means that despite the larger peaks, the model based on The Hague, Amsterdam and Utrecht is better able to predict the shared car supply rate in the 6 municipalities that represent the G44 cities.

Table 6.4 in Section 6.2.3.1 shows that the prediction of the presence of a roundtrip B2C shared car in the survey municipalities is better with the binary logistic regression based on the survey municipalities than it is with the binary logistic regression based on The Hague, Amsterdam and Utrecht. However, the predictions of the number of roundtrip B2C shared cars per 100,000 inhabitants in a neighbourhood is better with a negative binomial regression based on The Hague, Amsterdam and Utrecht. This is probably because the negative binomial regression based on the survey municipalities is not significantly better than a null model. This is presumably caused by the lack of variation in the shared car supply rate in the survey municipalities. Since the model based on The Hague, Amsterdam and Utrecht is based on more neighbourhoods with a high shared car supply rate, the variation in the data that is the input for the model is higher. This results in a better fit to explain the variation in the shared car supply rate, also when applying the model on the municipalities that represent the G44 cities.

The trend line in Figure 6.7 shows that the explanation of the variation in the supply of roundtrip B2C shared cars based on The Hague, Amsterdam and Utrecht has a predicted shared car supply rate in the neighbourhoods of the survey municipalities with a similar pattern as the observed shared car supply rate. However, because of the large peaks, the model is not able to make an exact prediction of the shared car supply rate in other municipalities than The Hague, Amsterdam and Utrecht. However, since the trend line has a similar pattern, an indication of the shared car supply rate in neighbourhoods in a different municipality can be provided.

The model based on the survey municipalities does not provide a good fit of the shared car supply rate in the municipalities that represent the G44 cities and the model based on The Hague, Amsterdam and Utrecht can provide a good indication of the shared car supply rate in the survey municipalities. Therefore, the model based on The Hague, Amsterdam and Utrecht is used to find out whether the explanation of the variation in the supply of roundtrip B2C shared cars in the G44 cities can be used to explain the variation in the demand for roundtrip B2C carsharing in the G44 cities in the next chapters.

Chapter 7

The effect of the observed supply of roundtrip B2C shared cars on the demand for roundtrip B2C carsharing

In this chapter, the effect of the observed supply of roundtrip B2C shared cars on the demand for roundtrip B2C carsharing is elaborated on and with that, the chapter is providing an answer to the second research question. As explained in Chapter 4, the effect of the observed supply of roundtrip B2C shared cars on the demand for roundtrip B2C carsharing is analysed in two ways. First, the respondents that indicated to have the same level of demand are binned and the observed shared car supply rate in their neighbourhoods is compared. Then, a comparison is made between the distribution of the demand of the respondents with and without a shared car in their neighbourhood. Finally, also the awareness of the presence of a shared car in the neighbourhoods of the respondents is analysed.

7.1 Observed shared car supply rate in neighbourhoods of respondents with a different demand for carsharing

To find the observed shared car supply rate in the neighbourhoods of respondents with a different demand for carsharing, all respondents that showed the same demand for roundtrip B2C carsharing have been binned. Three measures of supply have been calculated for each bin and these are presented in Table 7.1. These measures are the average shared car presence, the average number of shared cars and the average shared car supply rate in each group. Table 7.1 also shows the number of respondents and number of distinct neighbourhoods in each group.

TABLE 7.1: The supply of roundtrip B2C shared cars for respondents grouped by their demand for roundtrip B2C carsharing

Group	Number of respondents	Number of distinct neighbourhoods	Average shared car presence	Standard error	Average number of shared cars in neighbourhood	Standard error	Average shared car supply rate	Standard error
1	175	109	0.29	0.04	0.42	0.08	12.79	2.98
2	213	131	0.31	0.04	0.47	0.07	10.29	1.79
3	151	108	0.27	0.04	0.41	0.08	9.87	2.42
4	62	55	0.31	0.06	0.42	0.10	8.73	2.25
5	19	18	0.37	0.12	0.74	0.28	18.52	8.83
1 and 2 combined	388	240	0.30	0.03	0.45	0.05	11.42	1.67
4 and 5 combined	81	73	0.32	0.05	0.50	0.10	11.02	2.73

As shown in Table 7.1, the highest average shared car presence, the highest average number of cars and the highest average shared car supply rate can all be found in group 5. This is the group that answered 5 on the 5-point Likert scale and therefore showed the most demand. This means that the highest supply of shared cars is observed in the group that shows the most demand. However, the standard error of the averages in group 5 is also relatively high. Because of this relatively high standard error, the averages in group 5 are not significantly different from the averages in the other groups.

The average shared car presence is higher in group 2 than it is in groups 3 and 4. Group 3 even comes forward as the group with the lowest average shared car presence. Table 7.1 also shows that when groups 1 and 2 are combined and groups 4 and 5 are combined, the average shared car presence in these combined groups is more or less the same with an average shared car presence of 0.30 and 0.32 and a standard error of 0.03 and 0.05 respectively. This indicates no clear effect of the observed shared car presence in a neighbourhood on the demand of the respondent.

Table 7.1 is also showing no clear pattern between group 1 to group 4 in the average number of shared cars in a neighbourhood. However, the combination of groups 1 and 2 has a lower average number of shared cars in a neighbourhood than the combination of groups 4 and 5 and therefore indicates that a higher number of shared cars in a neighbourhood results in more demand for carsharing. Nevertheless, when taking the standard error into account, the ranges of the averages of both groups are overlapping. Hence, the average number of shared cars in the neighbourhoods of respondents that show no demand for carsharing is not significantly different from the average number of shared cars in the neighbourhoods of the respondents that show demand for carsharing.

The distribution of the average shared car supply rate over the different groups shows an unexpected pattern. When group 5 is excluded, the average shared car supply rate is higher in the groups that show a lower demand. Also, there is no clear difference in the average shared car supply rate in the combination of groups 1 and 2 and the average shared car supply rate in the combination of groups 4 and 5.

In conclusion, the supply of roundtrip B2C shared cars in a neighbourhood is not showing a clear effect on the demand for roundtrip B2C carsharing of the respondents that live in that neighbourhood. This means that hypothesis 2, which states that more supply of roundtrip B2C shared cars in a neighbourhood is causing more demand for roundtrip B2C carsharing in the same neighbourhood is rejected by this first method.

7.2 Demand of respondents with and without a shared car in their neighbourhood

The other method to find the effect of the observed supply of shared cars on the demand for carsharing is to compare the demand of the respondents that live in a neighbourhood with a shared car, with the demand of the respondents that live in a neighbourhood without a shared car. The distribution of the answers on the 5-point Likert scale and the average demand for both groups is presented in Table 7.2. The answers range from no demand (1) to a high demand (5).

TABLE 7.2: The supply of roundtrip B2C shared cars for respondents grouped by their demand for roundtrip B2C carsharing

Shared car presence	Yes	No	Total
Number of respondents	183	437	620
Average demand	2.27	2.24	2.25
Standard error	0.006	0.002	0.002
Demand on Likert scale			
1	27%	29%	28%
2	36%	34%	34%
3	22%	25%	24%
4	10%	10%	10%
5	4%	3%	3%
1 and 2 combined	63%	62%	63%
4 and 5 combined	14%	13%	13%

Table 7.2 shows that most of the respondents with a roundtrip B2C shared car in their neighbourhood show no demand for carsharing by indicating their demand with 1 or 2 on the 5-point Likert scale. The total percentage of respondents with these answers is 63% against a total percentage of respondents that indicated their demand with 4 or 5 of only 14%. The distribution of the demand of respondents that live in a neighbourhood without a shared car is almost identical to the distribution of the demand of respondents with a shared car in their neighbourhood. Also, the average demand is almost the same with an average of 2.27 and 2.24 respectively. The equal distribution and comparable average demand are also indicating no clear effect of the observed supply of shared cars on the demand for carsharing. This means that hypothesis 2 is also rejected by this second method.

7.3 Awareness of shared car presence

The awareness of the shared car presence of all respondents is measured by their answer to the question of whether the respondent knows if there is a shared car located in their neighbourhood and comparing this answer to the observed shared car presence in their neighbourhood. The distribution of answers to the question 'Is there a shared car present in your neighbourhood?' is shown in Figure 7.1. The comparison of the answers to the observed shared car presence in the neighbourhood of the respondents is presented in Table 7.3.

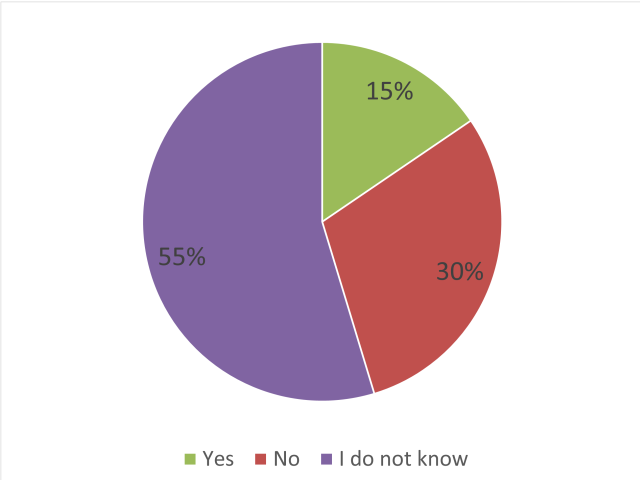

FIGURE 7.1: Distribution of answers to the question 'Is there a shared car present in your neighbourhood?'

TABLE 7.3: The observed shared car presence compared to the perceived shared car presence

		Observed shared car presence		**Total**	**% correct**
		No	*Yes*		
Perceived shared car presence	*No*	138	47	185	75%
	Yes	43	53	96	55%
Total		181	100	281	
% correct		76%	53%		

Figure 7.1 shows that 55% of the respondents does not know whether a shared car is placed in their neighbourhood. From Table 7.3 can be concluded that from the 96 respondents that indicated to live in a neighbourhood with a shared car, also only 55% of the respondents was correct about it. Also, from the 100 respondents that actually live in a neighbourhood with a shared car, only 53% knows about this. Considering these percentages, it can be concluded that the respondents have a low awareness of shared car presence in their neighbourhood.

Chapter 8

The relationship between the neighbourhood characteristics affecting the demand for carsharing and the neighbourhood characteristics affecting the supply of roundtrip B2C shared cars

This chapter provides an answer to the third research question by elaborating on the relation between the neighbourhood characteristics that affect the demand for roundtrip B2C carsharing and the neighbourhood characteristics that affect the supply of roundtrip B2C shared cars. The difference between this chapter and the previous chapter is that in the previous chapter it was analysed whether the observed supply of shared cars affects the demand for roundtrip B2C carsharing and in this chapter, it is about whether the same characteristics have the same effect on the demand for carsharing as on the supply of shared cars. The analysis in this chapter is making use of three distinct methods. First, the demand of respondents with a different predicted shared car supply rate is compared. Then the correlations of the neighbourhood characteristics with the demand are compared to the correlations of the neighbourhood characteristics with the supply of shared cars. Finally, the regression coefficients of the explanation of the variation in the supply of roundtrip B2C shared cars are compared to the regression coefficients of three new models that explain the variation in the demand for carsharing with the same independent variables as are used in the model that explains the variation in the supply.

8.1 Method 1: Comparing the demand of respondents with a different predicted shared car supply

The relation between the individual demand for carsharing and the predicted supply of shared cars in the neighbourhood of the individual is analysed to find out whether it is possible to use the regression model that explains the variation in the supply of roundtrip B2C shared cars, to explain the variation in the demand for roundtrip B2C carsharing in a neighbourhood. To find this relation, the respondents are binned in new groups that are based on the predicted shared car supply rate in their neighbourhoods. The model based on The Hague, Amsterdam and Utrecht that is developed in Chapter 6 is predicting a shared car supply rate that is higher than 0 in the neighbourhoods of 159 of the 620 respondents. 4 groups are created with the predicted shared car supply rate in the neighbourhoods of the respondents. The 159 respondents with a predicted shared car supply rate higher than 0 are split into 3 groups of 53 respondents per group. Group 1 consists of the 53 respondents with the highest predicted shared car supply rate and group 3 with the lowest. The remaining 455 respondents that have a predicted shared car supply rate of 0 are forming group 4. The average predicted shared car supply rate and average demand per group is shown in Table 8.1. The distribution of the demand in all groups is presented in Figure 8.1 to Figure 8.4.

TABLE 8.1: The predicted shared car supply rate in each group

Group	**Number of respondents in group**	**Average predicted shared car supply rate [Number of roundtrip B2C shared cars per 100,000 inhabitants]**	**Average demand**	**Standard error of average demand**
1	53	254.3	2.57	0.14
2	53	33.7	2.34	0.16
3	53	16.4	2.15	0.13
4	455	0.0	2.22	0.05

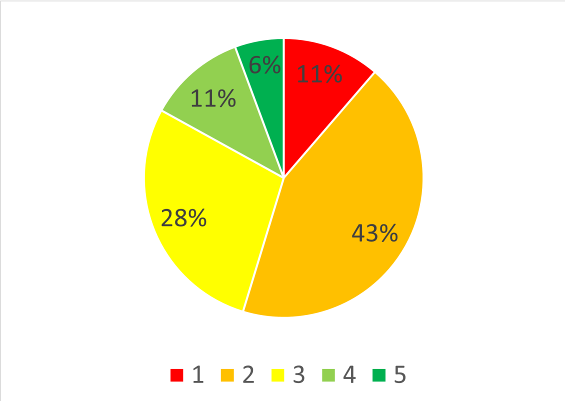

FIGURE 8.1: Distribution of the demand of respondents in group 1

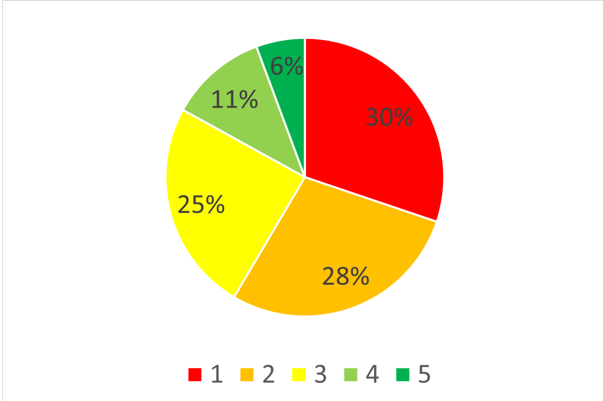

FIGURE 8.2: Distribution of the demand of respondents in group 2

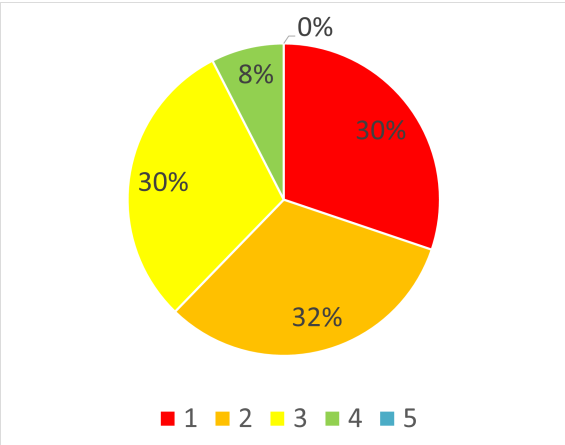

FIGURE 8.3: Distribution of the demand of respondents in group 3

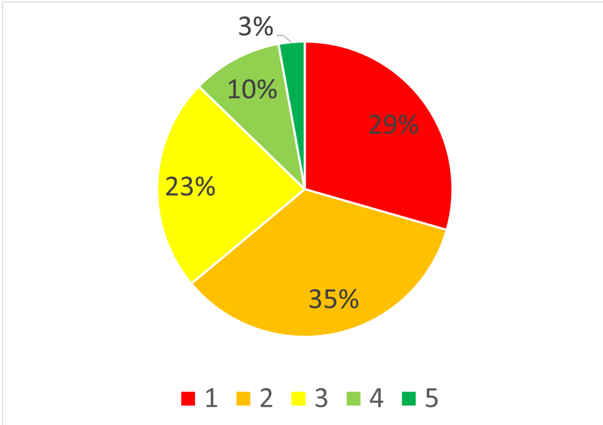

FIGURE 8.4: Distribution of the demand of respondents in group 4

Table 8.1 shows that the average demand is the highest in the group with the highest average predicted shared car supply rate, group 1. Also, a pattern can be detected that indicates that a group with a higher predicted shared car supply, shows more demand. However, the average demand in group 4 is slightly higher than in group 3.

The figures show that group 1 and group 2 both have the same percentages of respondents that indicated their demand with 4 or 5 on the 5-point Likert scale. In groups 3 and 4 these percentages are lower. When looking at the combination of percentages of respondents that indicated their demand with 1 or 2, a pattern can be detected as well. The combination of respondents that show a demand of 1 or 2 is 54% in group 1, 58% in group 2, 62% in group 3 and 64% in group 4. This indicates that the respondents in a group with a higher average predicted shared car supply rate are indicating their level of demand less frequently with 1 or 2. However, the difference between the percentages of adjacent groups is not significantly large. In conclusion, Table 8.1 and Figure 8.1 to Figure 8.4 indicate that the demand for roundtrip B2C carsharing is higher in the group with the highest predicted shared car supply rate. However, because there is no clear difference in the average demand and distribution of the demand in groups 3 and 4, no clear distinction can be made in the demand between adjacent groups.

8.2 Method 2: Comparison between the correlations of the neighbourhood characteristics with the shared car supply and with the demand for carsharing

This section presents the results of the comparison of the correlation between the neighbourhood characteristics and the demand for roundtrip B2C carsharing and the correlation between the neighbourhood characteristics and the supply of roundtrip B2C shared cars.

8.2.1 Comparison of characteristics correlations

To compare the correlations of the neighbourhood characteristics with the demand and the with the supply, the correlations of the neighbourhood characteristics with the demand, shared car presence and the shared car supply rate have been plotted next to each other in Figure 8.5. Figure 8.5 shows the correlation of the neighbourhood characteristics with the demand in the G44 cities and with the shared car presence and shared car supply rate in The Hague, Amsterdam and Utrecht. Descriptive information on these dependent variables is presented in Appendix L. The correlation of some neighbourhood characteristics with the shared car supply rate is relatively high compared to the correlation of the same neighbourhood characteristic with the demand and shared car presence. Therefore, the square roots of all correlations have been used in the graphs to make it easier to compare the correlations. Thereby, to maintain well-organised figures, only the variables that have a correlation with the demand that has a higher absolute value of the square root of the correlation than 0.2 are shown in the figures. This is because the correlation gets less significant when the correlation is smaller. After all, the difference between a positive and a negative sign is then smaller as well. The exact correlation and corresponding significance of the correlation are presented in the second section of Appendix L. The correlation with the shared car presence and shared car supply rate is based on all 704 neighbourhoods in The Hague, Amsterdam and Utrecht and the correlation with the demand is based on the complete sample of the G44 cities.

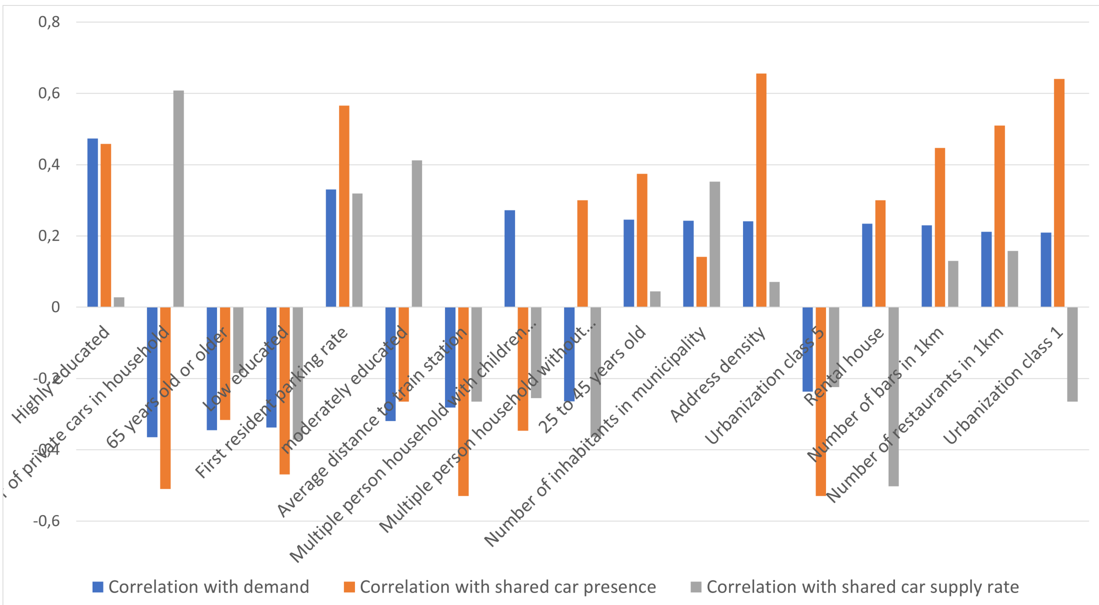

FIGURE 8.5: Comparison between the correlation of the neighbourhood characteristics with the demand, and the shared car presence and shared car supply rate in The Hague, Amsterdam and Utrecht

The second section in Appendix L shows that only the first 7 variables have a significant correlation with the demand for carsharing, with a p-value that is not higher than 0.05. Therefore, the signs and magnitudes of the correlations of these 7 variables are the most important to compare. Because the correlations of the other variables are less significant, the magnitudes of these correlations do not say much. Nevertheless, the signs of the correlations can still be used to provide an indication of a similar correlation between a neighbourhood characteristics and the demand, and the same neighbourhood characteristic and the supply of shared cars.

Figure 8.5 shows that in most cases, the correlation of a neighbourhood characteristic with the demand has the same sign as the correlation with the shared car presence. From the 17 variables in the figure, there are 15 variables of which the correlation with the shared car presence has the same sign as the correlation with the demand. Also, the magnitudes of the correlations of the first 7 variables with the demand are very similar to the magnitudes of the correlations of the first 7 variables with the shared car presence.

The sign of the correlation between each variable and the demand, and the sign of the correlation between each variable and the shared car supply rate is the same for 12 of the 17 variables. From the 5 variables of which the correlation with the demand and shared car supply rate have a non-matching sign, 2 variables are part of the 7 variables that have a significant correlation with the demand. The variables with non-matching signs are the number of private cars per household and moderately educated individuals.

In summary, the signs and magnitudes of the correlations of a neighbourhood characteristic with the demand and with the shared car presence are almost similar. The correlations of the neighbourhood characteristics with the demand and with the shared car supply rate are less similar, but still show to be very comparable. Therefore, in general, the correlations between the neighbourhood characteristics and the demand are similar to the correlations between the same neighbourhood characteristics and the supply of shared cars. This indicates that the supply in The Hague, Amsterdam and Utrecht is based on the same neighbourhood characteristics as the demand in the G44 cities. Therefore, hypothesis 3, which states that the characteristics of the neighbourhoods of individuals that show the most demand for roundtrip B2C carsharing are similar to the characteristics of the neighbourhoods with the highest supply of roundtrip B2C shared cars, can gently be accepted.

8.3 Method 3: Comparing regression coefficients

This section discusses the results of the comparison between the regression coefficients of the regression models that explain the variation in the individual demand for roundtrip B2C carsharing by making use of the same independent variables as have been used in the regression that explains the variation in the supply of shared cars. This is done by creating three new regression models. After the results of these models are presented, the regression coefficients are compared to the regression coefficients in the models that explain the variation in the supply of roundtrip B2C shared cars.

8.3.1 Explanation of the variation in the demand with the independent variables used in the explanation of the variation in the supply of roundtrip B2C shared cars

The explanation of the variation in the individual demand for roundtrip B2C carsharing with the independent variables used in the explanation of the variation in the supply of roundtrip B2C shared cars consists of three distinct regressions models. First, the variation in the demand is explained by using the same independent variables that are used in the binary logistic regression that explains the variation the shared car presence, then with the same independent variables as used in the negative binomial regression that explains the variation in the shared car supply rate, and finally, the independent variables of both regressions are used simultaneously.

8.3.1.1 Regression 1: Explanation of the variation in the demand with the independent variables used in the explanation of the variation in the shared car presence

The explanation of the variation in the shared car presence in a neighbourhood has been discussed in Section 6.2 and consists of a binary logistic regression that uses multiple neighbourhood characteristics as independent variables. The independent variables in the model are the address density, percentage of female individuals, percentage of individuals with a high income and the urbanization class of a neighbourhood. In this first regression, the independent variables that describe the demand in the most similar way as the variables in the regression that explains the variation in the shared car presence are used. Therefore, instead of using the percentage of female individuals and the percentage of individuals with a high income, a dichotomous independent variable is used that is either 1 or 0. This means that an individual is either female or not female, and has a high income or does not have a high income. A respondent has been identified as an individual with a high income when the respondents indicated to have a net income of more than 5000 euros because this is the highest category in the survey. Furthermore, the address density and urbanization class of the neighbourhood of an individual has been used in the regression. The independent variables in the binary logistic regression that explains the variation in the individual demand for roundtrip B2C carsharing are presented in Appendix M together with their corresponding regression coefficients and their p-value to show the significance of each variable. Since it is a binary logistic regression, the equation has the in Equation 4.1 and Equation 4.2 presented structure.

As shown in Appendix M, all independent variables in the equation are insignificant since their p-value is larger than 0.05. Also, the Nagelkerke R^2 of the regression is only 0.013 and with a chi-squared value of 0.838 and 7 degrees of freedom, the chi-square is not significant and the model is not significantly better than a random model.

8.3.1.2 Regression 2: Explanation of the variation in the demand with the independent variables used in the explanation of the variation in the shared car supply rate

The explanation of the variation in the roundtrip B2C shared car supply rate in a neighbourhood is based on a negative binomial regression model that also uses a large number of independent variables. The independent variables in the model are the number of inhabitants in the neighbourhood, percentage of individuals with high income, percentage of individuals between 0 and 14 years old, the percentage of households without children, the number of inhabitants in the municipality, the average distance to a train station, the percentage of individuals with a non-western migration background and the average distance to the city centre. As in the previously developed model, the independent variables that explain the variation in the demand for carsharing that are most similar to these variables are used in the regression model that explains the variation in the individual demand for carsharing. This means that instead of the percentages in a neighbourhood, dichotomous variables are used to describe the situation of an individual. Instead of the percentage of individuals between 0 and 14 years old, the age of a respondent is used since all respondents are at least 18 years old and therefore no data about individuals between 0 and 14 years old exist. For the number of inhabitants in the neighbourhood, the number of inhabitants in the municipality, the average distance to a train station and the average distance to the city centre, data about the neighbourhood of the respondent has been used. The independent variables in the binary logistic regression that explains the variation in the individual demand for roundtrip B2C carsharing are presented in Appendix M together with their corresponding regression coefficients and their p-value as well. Again, the equation has the in Equation 4.1 and Equation 4.2 presented structure.

As shown in Appendix M, also all independent variables in this equation are insignificant. The Nagelkerke R^2 of the regression is 0.027 and with a chi-squared value of 0.477 and 8 degrees of freedom, the chi-square is not significant and this model is also not significantly better than a random model.

8.3.1.3 Regression 3: Explanation of the variation in the demand with the independent variables used in the explanation of the variation in the shared car presence and the shared car supply rate

In the third regression, the explanation of the variation in the individual demand for roundtrip B2C carsharing is based on a combination of the independent variables used in the explanation of the variation in the shared car presence and shared car supply rate. The regression coefficients and p-values for all independent variables in the regression are also shown in Appendix M. Again, the equation has the in Equation 4.1 and Equation 4.2 presented structure.

As presented in Appendix M, also all independent variables in this equation are insignificant. The Nagelkerke R^2 of the regression is 0.040 and with a chi-squared value of 0.685 and 14 degrees of freedom, the chi-square is not significant and this model is also not significantly better than a random model.

8.3.2 Comparison of regression coefficients

Although all independent variables in the three regressions that explain the variation in the individual demand are insignificant, the regression models are used to compare the sign and magnitude of the regression coefficient of each independent variable with the sign and magnitude of the regression coefficients of the model that explains the variation in the supply of shared cars that uses the same independent variables. To make a clear comparison, the regression coefficients of the independent variables for each regression have been presented in Table 8.2. Since the independent variable 'high income' appears in both the binary logistic and the negative binomial regression, it has been included in the table twice.

TABLE 8.2: Coefficients of the explanation of the variation in the demand with the independent variables used in the explanation of the variation in the shared car presence and supply rate

Independent variable $[X_p]$	**Regression coefficient in shared car presence model**	**Regression coefficient in demand model with variables of shared car presence model**	**Regression coefficient in shared car supply rate model**	**Regression coefficient in demand model with variables of shared car supply rate model**	**Regression coefficient in demand model with combined variables**
Address density	0.381	-0.195			-0.188
Female	9.295	0.160			0.046
High income	3.347	0.256			0.265
Urbanization class 1	1.627	0.918			1.340
Urbanization class 2	1.194	0.931			1.161
Urbanization class 3	1.040	0.719			0.909
Urbanization class 4	-1.018	-0.599			-0.518
Number of inhabitants in neighbourhood			1.000	4.977	3.405
High income			3.956	0.211	*0.265*
Individuals between 0-14 / Age			-7.674	-0.015	-0.017
Living in a household without children			-5.308	-0.001	0.002
Number of inhabitants in municipality			0.310	0.716	0.793
Average distance to train station			-0.238	0.093	0.120
Having a non-western migration background			-1.753	-0.558	-0.600
Distance to city centre			-0.459	-0.017	-0.008

Table 8.2 shows that the magnitude of the coefficients in the explanation of the variation in the demand with the same independent variables as used in the model that explains the variation in the shared car presence is smaller for each independent variable. The coefficients for the urbanization classes, however, are somewhat similar. An individual that lives in urbanization class 4 has a clear negative effect on the demand, just as urbanization class 4 has a negative effect on the presence of a shared car. Also, the more urban the neighbourhood of an individual is, the more demand the individual seems to show. However, it is clear that for the demand of an individual, it is not important whether the individual lives in a neighbourhood in urbanization class 1 or 2. This is different from the effect of the urbanization class on the supply, where a neighbourhood in urbanization class 1 shows to have a significantly higher chance of having the presence of a shared car than a neighbourhood in urbanization class 2. Furthermore, a female individual and an individual with a high income are both more likely to show demand for roundtrip B2C carsharing, just as a roundtrip B2C shared car is more likely to be present in a neighbourhood with more females and individuals with a high income. The address density of an individual's neighbourhood, however, seems to have a negative effect on the demand, where it has a positive effect on the supply of roundtrip B2C shared cars.

The regression coefficients of the explanation of the variation in the demand with the same variables as the explanation of the variation in the shared car supply rate, also have mostly smaller magnitudes than the model that explains the variation in the supply rate. However, the number of inhabitants in the neighbourhood and also the number of inhabitants in the municipality of an individual have coefficients with a higher magnitude than in the regression of the supply of shared cars. The magnitudes of the age, the distance to the city centre and whether an individual lives in a household without children are very low and therefore have a small contribution to the predicted demand. The table also shows that the average distance to a train station has a negative effect on the demand, where it has a positive effect on the shared car supply rate. Having a non-western migration background has a relatively large negative effect on the demand for carsharing. This is similar to the effect of a neighbourhood with a large percentage of individuals with a non-western migration background on the shared car supply rate.

The regression coefficients in the model that explains the variation in the demand based on all independent variables that are used in the regressions that explain the variation in the supply of shared cars are comparable to the previously discussed regression coefficients. However, with this combination of variables, a difference between the demand of an individual living in urbanization class 1 or 2 occurs. Living in a neighbourhood in urbanization class 1 has a stronger positive effect on the demand for roundtrip B2C carsharing than living in a neighbourhood in urbanization class 2. This is more in line with the effect of the urbanization class on the supply of shared cars. Also, whether an individual is female, is having less effect on the predicted demand than it had in the first regression.

In conclusion, the signs of almost all independent variables in the models that explain the demand for carsharing are equal to the signs of the same independent variable in the model that explains the supply of shared cars. However, the magnitude of most variables is deviating largely from the magnitude in the supply model. This could be caused by the high p-value of almost all independent variables in the models that explain the variation in the demand. Because of this high p-value, the contribution of each independent variable to the regressions is insignificant and its regression coefficient may therefore be small. Nevertheless, since the signs of almost all independent variables are equal, method 3 indicates that the neighbourhood characteristics used in the model have a comparable relationship with the supply of shared cars as with the demand for carsharing.

8.4 Comparing the results of the three different methods

To find the relationship between the neighbourhood characteristics that affect the demand for roundtrip B2C carsharing and the neighbourhood characteristics that affect the supply of roundtrip B2C shared cars 3 methods have been applied. The results of these methods are compared to each other to identify whether they produced the same outcomes.

Method 1 indicates that a high predicted shared car supply rate is an indication for more demand for roundtrip B2C carsharing. This indicates that the neighbourhood characteristics used in the model that explains the variation in the supply of shared cars have the same relation with the demand as they have with the supply of shared cars. However, the difference between the demand in adjacent groups with a different predicted shared car supply rate is not clearly visible.

The results of method 2 show that the signs and magnitudes of the correlations of a neighbourhood characteristic with the demand and with the shared car presence are almost similar. The correlations of the neighbourhood characteristics with the demand and with the shared car supply rate are less similar, but still show to be very comparable. This indicates that the supply in The Hague, Amsterdam and Utrecht is based on similar neighbourhood characteristics as the demand in the sample that represents the G44 cities.

Method 3 showed that most of the neighbourhood characteristics used as independent variables in the model that explains the variation in the supply have a comparable relationship with the supply of shared cars as with the demand for carsharing. However, because all regression coefficients in the models that explain the variation in the demand are insignificant, the magnitude of most variables is deviating largely from the magnitude in the model that explains the variation in the supply.

All three methods indicate that the neighbourhood characteristics that affect the demand for roundtrip B2C carsharing and the neighbourhood characteristics that affect the supply of roundtrip B2C shared cars are comparable. Method 2 and method 3 show that most neighbourhood characteristics have the same relation with the supply as with the demand when looking at this relation per neighbourhood characteristic and when combining neighbourhood characteristics in a regression. Method 1 is the only method that is significantly showing the size of the relation by indicating that in a group with the highest average predicted shared car supply rate, the demand is significantly larger than in the other groups. Therefore, the combined results of the three methods show that the explanation of the variation in the supply of roundtrip B2C shared cars over neighbourhoods in the G44 cities is not able to indicate the exact demand for roundtrip B2C shared cars in a neighbourhood but can identify the neighbourhoods with a relatively high expected demand for roundtrip B2C carsharing.

Chapter 9

Explanation of the variation in the demand for roundtrip B2C shared cars

This chapter focuses on all factors that explain the variation in the demand for roundtrip B2C shared cars of inhabitants of the G44 cities in the Netherlands and thus provides an answer to the fourth research question. To address each aspect of the answer to the main question, the analysis in this chapter provides insight into the reason why the model that is based on only neighbourhood characteristics, cannot exactly explain the variation in the demand in different neighbourhoods in the G44 cities. This is done by analysing the different motives and purposes for using a shared car and developing a regression model that also includes the remaining individual factors framed in Figure 2.1 in the explanation of the variation in the demand for roundtrip B2C shared cars.

The chapter starts with the analysis of the motives and purposes of using a roundtrip B2C shared car. Then the correlations between all remaining individual factors, neighbourhood characteristics and the demand for roundtrip B2C shared cars are discussed. Finally, a regression model is developed in which the variation in the demand for roundtrip B2C shared cars in the G44 cities is explained based on these individual factors and neighbourhood characteristics.

9.1 Motives and purposes of using a roundtrip B2C shared car

The respondents in the 6 municipalities that represent the G44 cities have indicated their level of interest in using a shared car for different motives and purposes on a 5-point Likert scale. The distributions of their answers are analysed by comparing them to each other. First, the motives are discussed and then the purposes are elaborated on.

9.1.1 Motives to use a roundtrip B2C shared car

Within the survey, the respondents had to indicate their level of agreement on 6 possible motives to use a roundtrip B2C shared car. As discussed in the methodology in Section 4.3.4, the motives are about lowering travel costs, parking places near the respondent's house, avoiding owning a car, avoiding maintenance and repair, avoiding searching for parking places and driving more environmentally friendly. The respondent's level of agreement about the motives has been indicated on a 5-point Likert scale and a distinction has been made between respondents with experience with a shared car and respondents without experience with a shared car. The average answer and the distribution of answers on the level of agreement per motive are shown in Figure 9.1 to Figure 9.12. The distribution of answers of the respondents with experience with a shared car is based on 80 respondents and the distribution of answers of individuals without experience is based on 540 respondents.

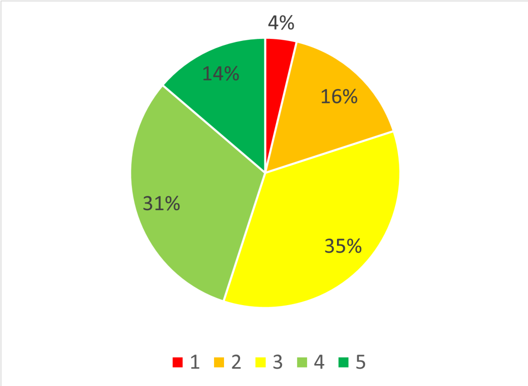

FIGURE 9.1: Shared car user that is using a shared car because of lower travel costs (avg. 3.35)

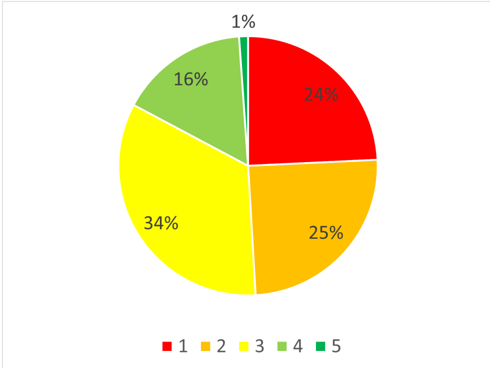

FIGURE 9.2: Individual without experience with a shared car that is considering to use a shared car to have lower travel costs (avg. 2.45)

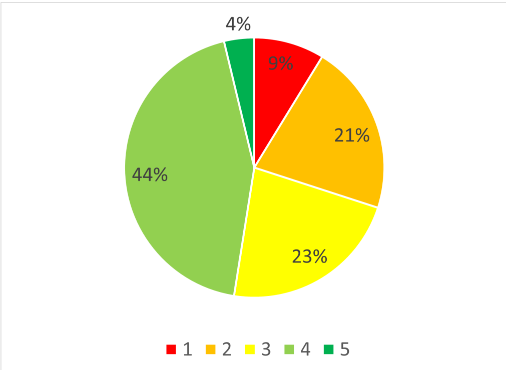

FIGURE 9.3: Shared car user that is using a shared car because there are shared car parking places near his/her house (avg. 3.13)

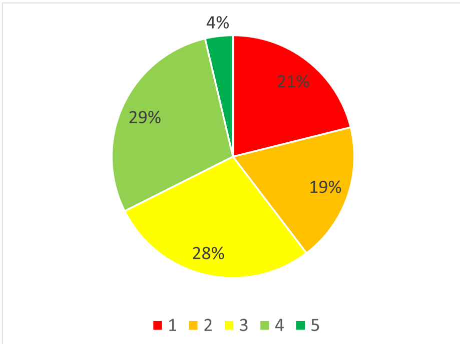

FIGURE 9.4: Individual without experience with a shared car that is considering to use a shared car when there are shared car parking places near his/her house (avg. 2.75)

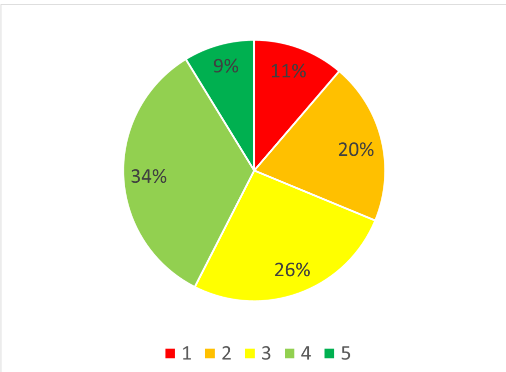

FIGURE 9.5: Shared car user that is using a shared car to avoid owning a car himself/herself (avg. 3.09)

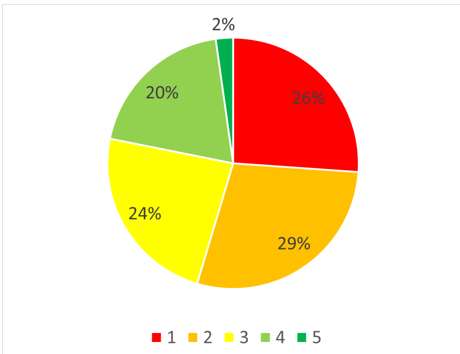

FIGURE 9.6: Individual without experience with a shared car that is considering to use a shared car to avoid owning a car himself/herself (avg. 2.43)

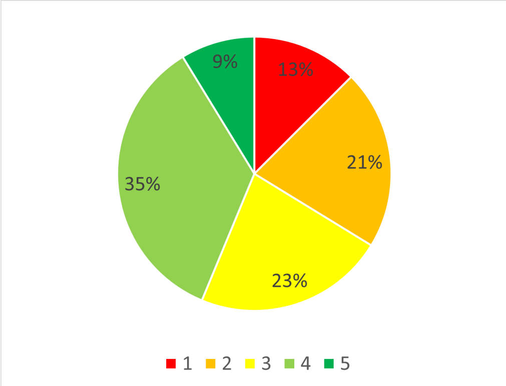

FIGURE 9.7: Shared car user that is using a shared car to avoid maintenance and repair of own car (avg. 3.06)

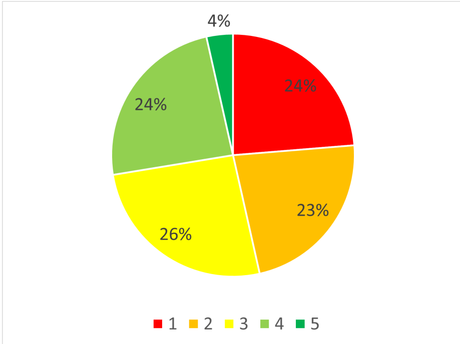

FIGURE 9.8: Individual without experience with a shared car that is considering to use a shared car to avoid maintenance and repair of own car (avg. 2.61)

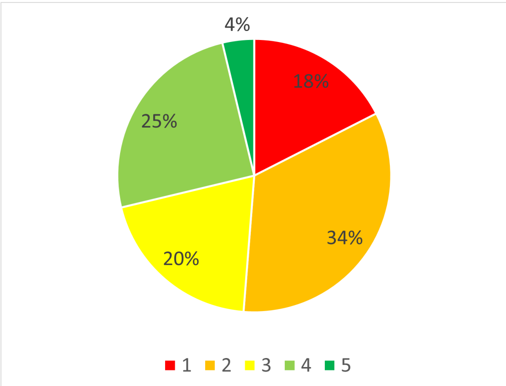

FIGURE 9.9: Shared car user that is using a shared car to avoid searching for parking places (avg. 2.64)

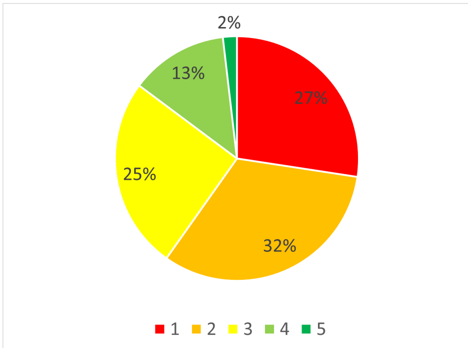

FIGURE 9.10: Individual without experience with a shared car that is considering to use a shared car to avoid searching for parking places (avg. 2.29)

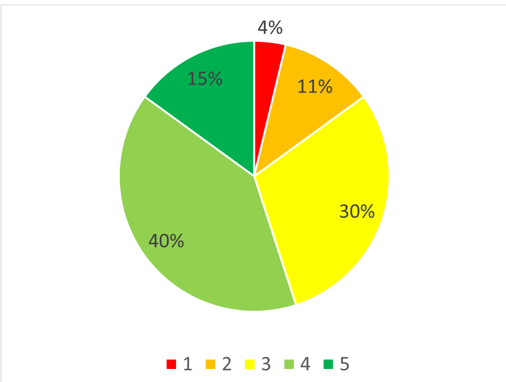

FIGURE 9.11: Shared car user that is using a shared car to drive more environmentally friendly (avg. 3.51)

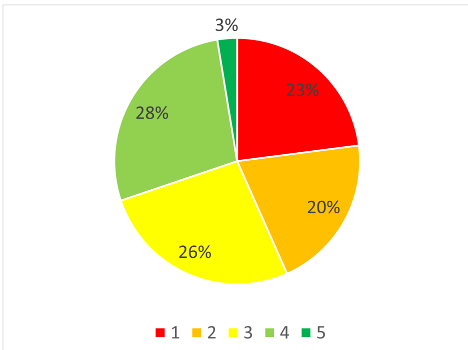

FIGURE 9.12: Individual without experience with a shared car that is considering to use a shared car to drive more environmentally friendly (avg. 2.66)

From the figures can be concluded that for the respondents that have experience with a shared car, the main motive to use a shared car is to drive more environmentally friendly. This motive has the highest average answer and 55% of all answers about the motive to drive more environmentally friendly is 4 or 5 and only 15% of all answers is 1 or 2 on the 5-point Likert scale. The figures also show that there is a relatively large number of shared car users that is using a shared car because of low travel costs. There are also a lot of shared car users that indicate a level of agreement of 4 or 5 on the motive to use a shared car because there are shared car parking places near their house. Furthermore, using a shared car to avoid searching for parking places is the motive with the lowest average answer, having the least answers being 4 or 5 and most answers being 1 or 2.

The respondents without experience with a shared car are indicating a generally lower level of agreement per motive. Nevertheless, some of the same conclusions can be drawn from the comparison between the motives. The respondents without experience with a shared car are also indicating the motive to avoid searching for parking places as the motive with the lowest average answer and also having the least answers being 4 or 5 and the most answers being 1 or 2. The most important motive for an individual without experience, however, is to consider making use of a shared car when parking places are available near the respondent's house. This motive is closely followed by driving more environmentally friendly and also the avoidance of maintenance and repair is a relatively popular motive.

Concluding, the main motive for shared car users to make use of a shared car is to drive more environmentally friendly and also non-users of shared cars find driving more environmentally friendly an important motive to consider using a shared car. However, the most important motive for non-users is to consider the use of a shared car when there are shared car parking places near their house. Another difference between users and non-users is that users find the motive to use a shared car because of lower travel cost relatively more important than non-users do. Non-users find the avoidance of maintenance and repair more important. Finally, the search for parking places is the least important motive for both users and non-users of a shared car.

Hypothesis 4 stated that roundtrip B2C shared cars are used to have the convenience of having a car only when needed and to avoid responsibilities with maintenance and repair, and when these cars and their carsharing pick-up locations are accessible near an individual's place or workplace. To use the shared car to have the convenience of having a car only when needed is presumed to be equal to using a shared car to avoid owning a car. Since the results show that this is not a popular motive, this part of the hypothesis is rejected. The other part of the hypothesis is accepted since non-users find the avoidance of maintenance and repair relatively important and both users and non-users consider the accessibility of carsharing parking places near their house an important motive.

9.1.2 Purposes for using a roundtrip B2C shared car

The respondents also had to indicate their level of agreement on 5 possible purposes for the use of a roundtrip B2C shared car. These purposes are to use a shared car to go to work or study with, to do the groceries with, to use to visit a friend or relative, to go to an activity and to go to another city or village. The respondent's level of agreement has also been indicated on a 5-point Likert scale. Since the experience with a shared car did not affect the way the question was asked, the distribution is shown for all 620 respondents. To be able to find the difference in the purposes of individuals with experience with the shared car, these results are also presented. The average answer and the distribution of answers on the level of agreement per purpose are shown in Figure 9.13 to Figure 9.22.

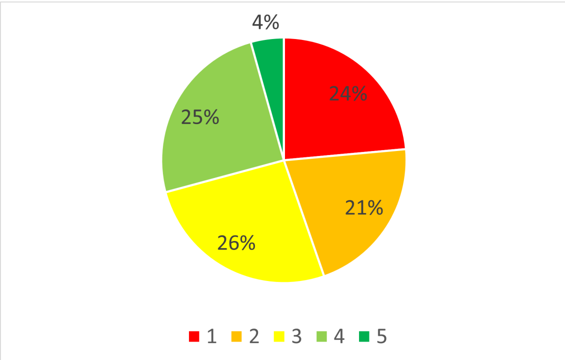

FIGURE 9.13: Respondents that find a shared car an attractive option to go to work or study with (avg. 2.65)

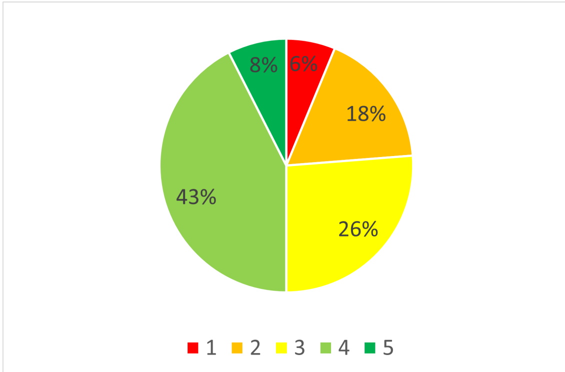

FIGURE 9.14: Shared car user that finds a shared car an attractive option to go to work or study with (avg. 3.28)

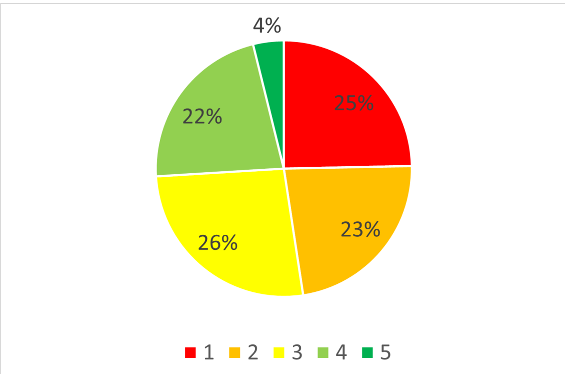

FIGURE 9.15: Respondents that find a shared car an attractive option to do the groceries with (avg. 2.58)

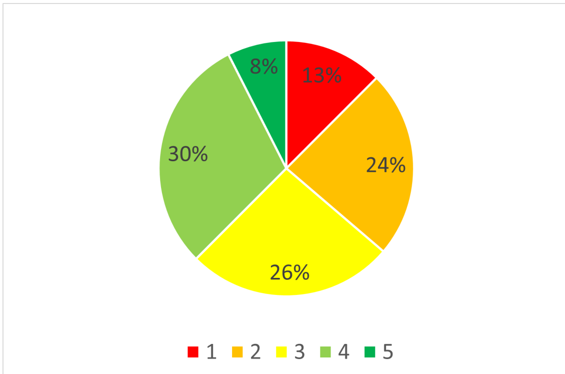

FIGURE 9.16: Shared car user that finds a shared car an attractive option to do the groceries with (avg. 2.96)

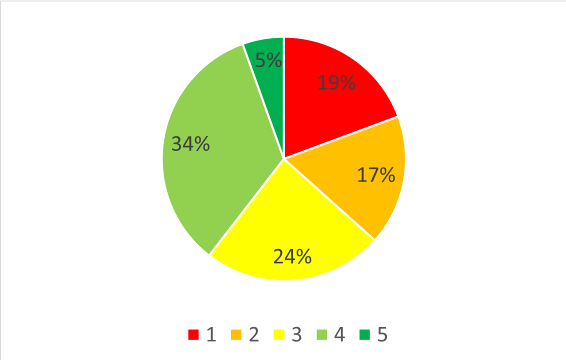

FIGURE 9.17: Respondents that find a shared car an attractive option to use to visit a friend or relative (avg. 2.89)

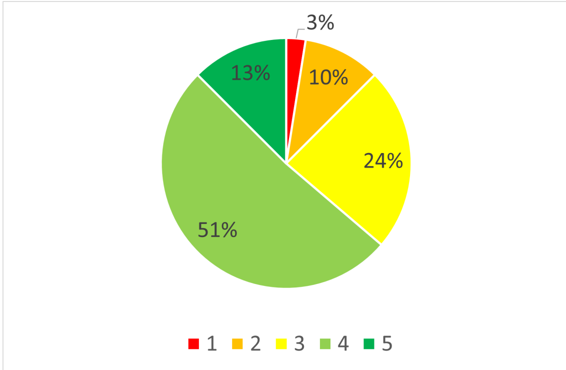

FIGURE 9.18: Shared car user that finds a shared car an attractive option to use to visit a friend or relative (avg. 3.61)

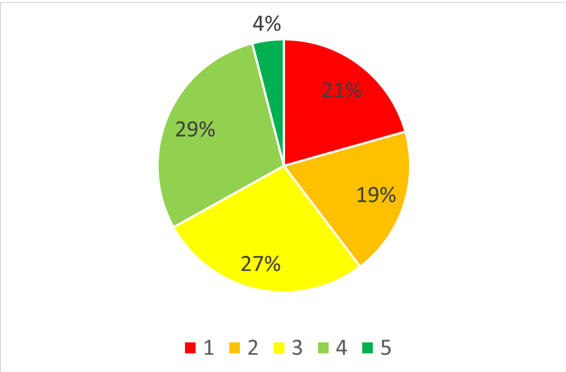

FIGURE 9.19: Respondents that find a shared car an attractive option to use to go to an activity (avg. 2.77)

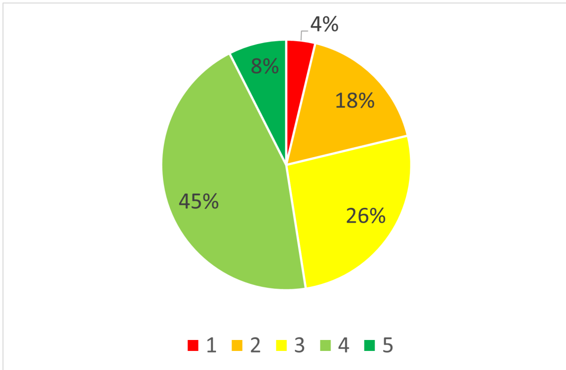

FIGURE 9.20: Shared car user that finds a shared car an attractive option to use to go to an activity (avg. 3.35)

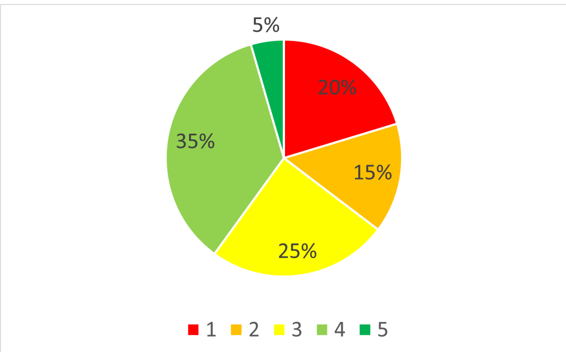

FIGURE 9.21: Respondents that find a shared car an attractive option to use to go to another city or village (avg. 2.89)

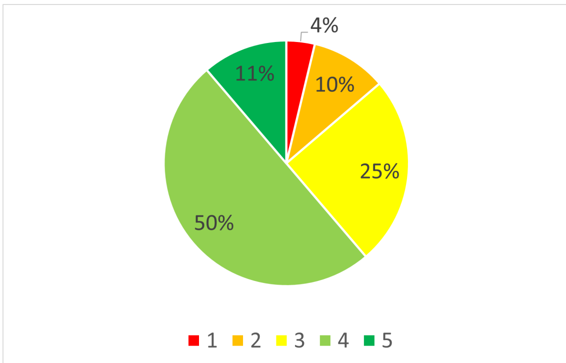

FIGURE 9.22: Shared car user that finds a shared car an attractive option to use to go to another city or village (avg. 3.55)

The figures show that the most attractive purposes are to use a shared car to visit a friend or relative and to go to another city or village. The average attractiveness of both purposes is 2.89 and also the percentage of respondents that chose 1 or 2, and 4 or 5 is more or less the same. To use a shared car to go grocery shopping and to go to work or study are the least popular purposes to use a shared car. This means that a roundtrip B2C shared car is most likely to be used for occasional trips that are often relatively long.

The users of a shared car are more positive about all purposes to use a shared car. Nevertheless, also the purposes to visit a friend or relative with, and to go to another city or village come forward as the most popular purposes. Also, the use of a shared car to go grocery shopping is the purpose that is the least popular. The only difference is that the users of a shared car find the purpose of using a shared car to go to work or study relatively more popular than all respondents think about this purpose. Both groups are also relatively positive about using a shared car to go to an activity or day out. However, hypothesis 5, in which it is stated that roundtrip B2C shared cars are mostly used to visit a friend or a relative and for a day out, is rejected. This is because using a shared car to go to another city or village is more popular than using it for a day out. Nevertheless, the expectation that a roundtrip B2C shared car is most likely to be used for occasional trips that are often relatively long is confirmed.

9.2 Correlations of the individual factors and neighbourhood characteristics with the demand for roundtrip B2C shared cars

The correlation of the demand with the different individual factors and neighbourhood characteristics that have the strongest effect on the demand for roundtrip B2C shared cars are presented in Figure 9.23 and are sorted from strongest correlation to weakest correlation. Showing demand is defined as an answer of 4 or 5 on the 5-point Likert scale about the demand for roundtrip B2C shared cars. The correlation with the demand has been calculated for 65 variables that are either individual factors or neighbourhood characteristics. Since the absolute correlation of a large number of these variables is smaller than 0.1, only the 22 variables with the strongest correlation with the demand are shown in Figure 9.23. The correlation between all 65 variables and the demand and the corresponding significance of the correlations are presented in Appendix N.

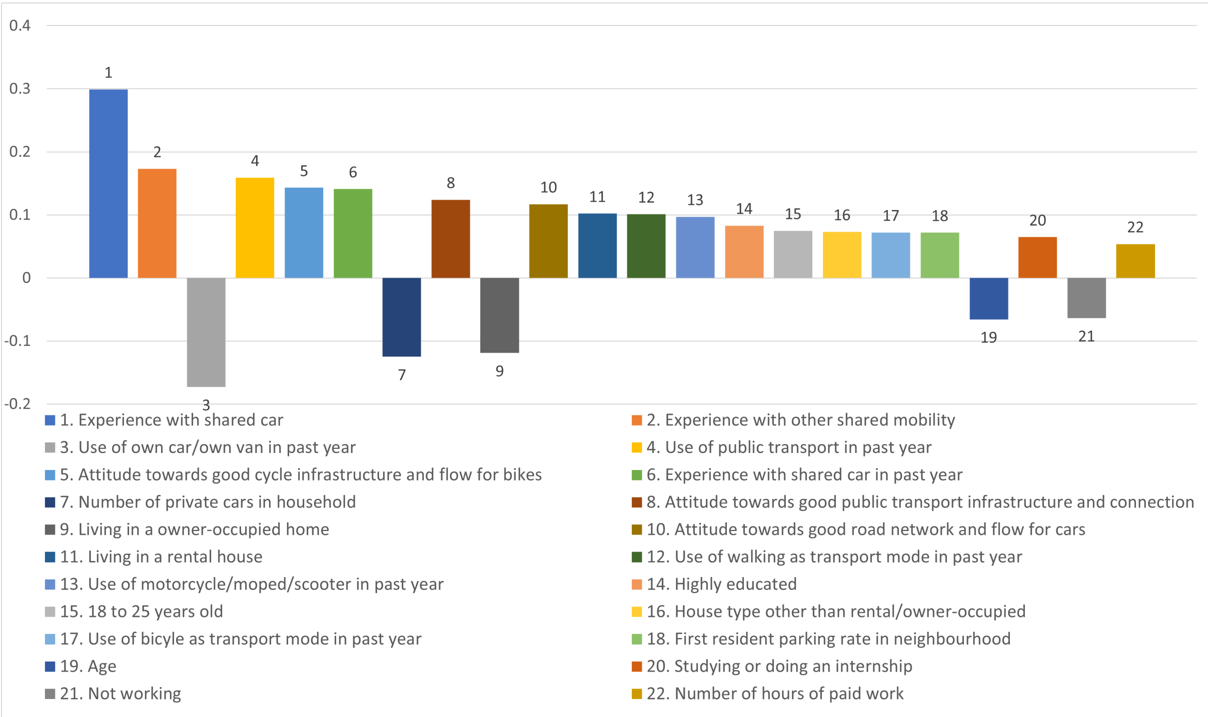

FIGURE 9.23: Characteristics and attitudes with the strongest correlation with the demand for roundtrip B2C shared cars

Figure 9.23 shows that the demand has the strongest relationship with respondents that have experience with a shared car. Also, the experience with other shared mobility turns out to be an important characteristic of an individual that shows demand for roundtrip B2C carsharing. This is in line with the work of (Jaffe et al., 2000) and (Münzel et al., 2019), and part of hypothesis 6 in which is expected that individuals that have experience with other shared mobility and have strong environmental ideologies show more demand for a roundtrip B2C shared car. The environmental ideology of the respondents is measured by their political preference for green parties. This is done with the assumption that Dutch parties on the left side of the political spectrum are often parties that can be identified as greener parties. The correlation of the demand with individuals that are left or central-left oriented on the political scale is 0.028 and with individuals that are right or central-right oriented on the political scale is -0.002. Unfortunately, the significance of both correlations is very low and therefore no statement can be made about the effect of a strong environmental ideology on the demand for carsharing.

Figure 9.23 also shows that the individual demand for carsharing has a strong negative correlation with the use of the own car in the past year and with the number of private cars in a household. In combination with the strong positive correlation between the demand for carsharing and the use of public transport and walking as a transport mode, it can be concluded that individuals that travel more often with a non-car form of transport, are more likely to show demand for roundtrip B2C shared cars. Also interesting is the correlation of the demand with the respondents' attitudes towards the infrastructure and flow for bicycles, public transport and the road network for cars. It was expected that the correlation with the respondents' attitude towards a good infrastructure and network for cars would be significantly lower than the correlation of the demand for carsharing with the attitude towards the infrastructure and flow for bicycles and public transport. Despite that the correlation between the demand and the attitude towards infrastructure and flow for bicycles and public transport is slightly higher than the importance of a good road network for cars, there is no clear distinction between these attitudes. What also points out is that the correlations show that individuals that live in a rental house are more likely to show demand and individuals that live in an owner-occupied house are less likely to show demand for carsharing. Additionally, young and highly educated individuals are also more likely to show demand for roundtrip B2C shared cars.

The correlations between the demand for roundtrip B2C shared cars and the individual factors and neighbourhood characteristics indicate that multiple characteristics other than neighbourhood characteristics are important when explaining the variation in the demand in neighbourhoods in the G44 cities. Figure 9.23 shows that the experience with a shared car and experience with other shared mobility are most important. Thereby, the correlations show that not only the possession but also the use of a private car plays an important part in the explanation of the variation in the demand as well. The figure also confirms the importance of a large number of neighbourhood characteristics in the explanation of the variation in the demand for roundtrip B2C shared cars. This is reflected in the strong correlation between the demand for carsharing and the number of private cars in a household, living in a rental house, being highly educated and having a young age. Also, the importance of the modal split of a neighbourhood to the demand for roundtrip B2C shared cars follows from the correlations.

9.3 Explanation of the variation in the demand for roundtrip B2C shared cars

The explanation of the variation in the demand for roundtrip B2C shared cars is based on a binary logistic regression. Therefore, it is structured in the same way as the binary logistic regressions that are used to explain the variation in the supply of shared cars. First, the process of obtaining the regression equation with the best fit is discussed. Then, the model is validated based on its classification table, residuals and goodness of fit.

9.3.1 Process and regression equation

The variation in the demand for roundtrip B2C shared cars is explained by using a binary logistic regression model with the 65 independent variables discussed in the previous section as a point of departure. The dependent variable in the model is a binary value that represents whether an individual shows demand for roundtrip B2C shared cars or not. The same process as has been used to explain the variation in the supply of shared cars has been followed to obtain the explanation of the variation in the demand as well. The variable with the strongest correlation with the demand in Figure 9.23, is the experience with shared cars. This variable is the first variable that is inserted in the binary logistic regression. Then the next variables with a strong correlation have been added one after the other and each time a new regression is fitted with the new combination of variables. The binary logistic regression is assessed on its Nagelkerke R^2 and the significance of its variables. The steps that have been taken to find the most effective binary logistic regression are presented in Appendix O. Also, independent variables with a smaller correlation than the independent variables in the table were included but this did not result in a better fit. To be certain that the independent variables in the model do not show multicollinearity, also the VIF of the independent variables is calculated. Since the VIF for all independent variables in the regression is below 5, there is no multicollinearity in the models. The regression equation of the most effective model that explains the variation in the demand for roundtrip B2C shared cars is presented in Table 9.1. The equation has the in Equation 4.1 and Equation 4.2 presented structure.

TABLE 9.1: Independent variables in the binary logistic regression equation that explains the variation in the demand for roundtrip B2C shared cars

Independent variable $[X_p]$	**Regression coefficient**$[b_p]$
Constant	-3.762
Experience with shared cars	1.804
Attitude towards a good infrastructure and flow for cycling	0.544
Number of private cars in household	-0.518

Table 9.1 shows that the regression equation of the binary logistic regression equation that explains the variation in the demand for roundtrip B2C shared cars consists of 3 independent variables. The experience with shared cars and the attitude towards a good infrastructure and flow for cycling show a positive relationship with the demand for carsharing. The number of private cars in the household of an individual has a negative effect on the demand. The classification cutoff value in this regression is used to classify an individual as an individual that shows demand or does not show demand for roundtrip B2C shared cars. It is again determined by choosing a value that ensures that there is an equal number of false positives and false negatives in the classification table. This is the case with a cutoff value of 0.21.

9.3.2 Validation of the regression

To evaluate the model that explains the variation in the demand for roundtrip B2C shared cars, the classification table is used. The classification table that is obtained by using a classification cutoff value of 0.21 is presented in Table 9.2.

TABLE 9.2: Classification table of binary logistic regression model that explains the variation in the demand for roundtrip B2C shared cars

	No predicted demand	**Predicted demand**	*Percentage correct*
No observed demand	382	41	*90.3%*
Observed demand	42	25	*37.3%*
			83.1%

Table 9.2 shows that the model is making a correct prediction for 83.1% of all respondents in the training set. However, this high percentage is mainly caused by a correct prediction of 90.3% for the individuals with no observed demand for carsharing. The predictions for the individuals that indicated to show demand for carsharing is only correct in 37.3% of the cases. This is an indication that the model is not performing well despite the high percentage for the overall correctness of the predictions.

When validating the model based on the chi-squared value, the indication of a bad performance is not confirmed. The chi-squared value for the model is 56.2. Because there are 3 degrees of freedom in the model, the significance value is smaller than 0.001. This means that the chi-square is highly significant and the model is significantly better than a random model.

The performance of the regression is also assessed by applying the regression equation on the validation set and using the same classification cutoff value to validate the model. The classification table that results from this is presented in the second section of Appendix O. The percentages in these classification tables are similar to the percentages in the initial classification tables. This means that the overfitting of the models is minimal.

To search for outliers in the residuals of the predictions, the predicted probability of showing demand for roundtrip B2C shared cars are plotted against the observed demand for roundtrip B2C shared cars. This plot is presented in the third section of Appendix O. The largest residuals are the residuals that result from a large predicted probability of showing demand for carsharing and no observed demand. These residuals have been analysed and no mistakes have been found in the variables that are the base of the prediction. Therefore, these residuals are not considered as outliers. The plot also shows that no clear distinction can be made between the predicted probabilities for the individuals that showed demand and for the individuals that did not show demand for roundtrip B2C shared cars. This indicates that the model is not performing well.

Although the model that explains the variation in the demand for roundtrip B2C shared cars most effectively is significantly better than a random model, the classification table and plot of the predicted probabilities against the observed demand showed that the model is not performing well. Therefore, the model can not be used to gain knowledge about what characteristics other than neighbourhood characteristics are important when explaining the variation in the demand in different neighbourhoods in the G44 cities. Nevertheless, the correlations between the demand and the individual factors and neighbourhood characteristics that are discussed in the previous section did indicate what characteristics other than neighbourhood characteristics are important.

Chapter 10

Conclusions

This research aims to contribute knowledge to the identification of the demand for roundtrip B2C carsharing in neighbourhoods in the G44 cities in the Netherlands by adding knowledge on the current characteristics that explain the distribution of the supply of and demand for roundtrip B2C carsharing over these neighbourhoods by answering the following research question: 'How are neighbourhood characteristics explaining the variation in the supply of roundtrip B2C shared cars in neighbourhoods in the G44 cities in the Netherlands and can this explanation of variation be used to explain the variation in the demand for roundtrip B2C carsharing in these neighbourhoods?'. To answer this research question, an explanation of the variation in the supply of roundtrip B2C shared cars in neighbourhoods in the G44 cities by the characteristics of a neighbourhood is provided. Then it is analysed whether this explanation could be used to explain the variation in the demand for roundtrip B2C carsharing in neighbourhoods in the G44 cities in the Netherlands.

By combining a model that predicts the presence of a roundtrip B2C shared car in a neighbourhood in the G44 cities with a model that predicts the number of roundtrip B2C shared cars per 100,000 inhabitants in the neighbourhoods with a predicted presence, a significant explanation of the variation in the supply of roundtrip B2C shared cars in neighbourhoods in the G44 cities in the Netherlands is provided.

To find out whether the explanation of the variation in the supply of roundtrip B2C shared cars can be used to explain the variation in the demand for roundtrip B2C carsharing as well, the effect of the observed supply on the demand for roundtrip B2C carsharing is analysed. Individuals that live in a neighbourhood with a roundtrip B2C shared car show an almost identical level of demand as individuals without a roundtrip B2C shared car in their neighbourhood. Therefore, there is no indication of an effect of the observed supply of roundtrip B2C shared cars in a neighbourhood on the demand for roundtrip B2C carsharing and the observed supply of shared cars does not have to be included in the explanation of the variation in the demand for roundtrip B2C carsharing.

There is no indication of an effect of the observed supply of roundtrip B2C shared cars on the demand for roundtrip B2C carsharing. However, a relation between the neighbourhood characteristics that affect the demand for roundtrip B2C carsharing and the neighbourhood characteristics affecting the supply of roundtrip B2C shared cars is detected. The correlation of each neighbourhood characteristic with the demand is directly related to almost every correlation of the same neighbourhood characteristic with the presence of a roundtrip B2C shared car in a neighbourhood. The relation between the correlations of the neighbourhood characteristics with the demand and the correlations of the neighbourhood characteristics with the number of roundtrip B2C shared cars per 100,000 inhabitants is a little smaller but is still comparable. The neighbourhood characteristics have also been applied in a regression model that explains the variation in the demand for roundtrip B2C carsharing based on the same variables as the regression model that explains the variation in the number of shared cars per 100,000 inhabitants for the neighbourhoods in the G44 cities. Most of the regression coefficients of the neighbourhood characteristics used as independent variables in the model that explains the variation in the supply have a comparable sign as the regression coefficients of the neighbourhood characteristics in the model that explains the variation in the demand. And also, individuals that live in a neighbourhood in which the model predicts a high number of shared cars per 100,000 inhabitants, show a high demand for roundtrip B2C carsharing. This indicates that the neighbourhood characteristics that are used in the model that explains the variation in the supply of roundtrip B2C shared cars in neighbourhoods in the G44 cities in the Netherlands have a similar effect on the supply as they have on the demand of roundtrip B2C shared cars and that the explanation of the variation in the supply of shared cars can identify neighbourhoods with a high demand for roundtrip B2C carsharing. However, the explanation can not be used to predict the exact demand in a neighbourhood in the G44 cities.

This research also provides insight into the reason why a model based on only neighbourhood characteristics cannot exactly explain the variation in the demand for carsharing in neighbourhoods in the G44 cities by analysing the effect of individual factors and neighbourhood characteristics on the individual demand for roundtrip B2C carsharing. Experience with a shared car and experience with other shared mobility turned out the most important characteristics of an individual to show demand for roundtrip B2C carsharing. Also, individuals that travel more often with a non-car form of transport are more likely to show demand for roundtrip B2C carsharing. The analysis also confirms the importance of a large number of neighbourhood characteristics in the explanation of the variation in the demand for roundtrip B2C carsharing. This is reflected in the strong correlation between the demand for carsharing and the number of private cars in a household, living in a rental house, being highly educated and having a young age.

The main motives and purposes for using a roundtrip B2C shared car are analysed separately. This analysis shows that the main motive for shared car users to make use of a shared car is to drive more environmentally friendly and the main motive for non-users of shared cars is to consider the use of a shared car when there are shared car parking places near their house. To visit a friend or relative, and to go to another city or village come forward as the most popular purposes for using a roundtrip B2C shared car. This means that a roundtrip B2C shared car is most likely to be used for occasional trips that are often relatively long.

In conclusion, a significant explanation of the variation in the supply of roundtrip B2C shared cars in neighbourhoods in the G44 cities in the Netherlands is provided and it is shown that the neighbourhood characteristics that are used in the model indicate a similar relationship with the supply as with the demand. Additionally, the findings of this research showed that there is no indication of an effect of the observed supply of roundtrip B2C shared cars in a neighbourhood on the demand for roundtrip B2C carsharing and that the model can identify a high demand in a neighbourhood by predicting the number of roundtrip B2C shared cars per 100,000 inhabitants. Therefore, this research has shown that the explanation of the variation in the supply of roundtrip B2C shared cars in neighbourhoods in the G44 cities in the Netherlands can be used to identify neighbourhoods in these cities that should have a relatively high demand for roundtrip B2C carsharing.

Chapter 11

Discussion

This chapter delves into the meaning, importance and relevance of the results of this study. It discusses the results and how the results match the expectations that were established in the hypotheses. Also, the limitations of this research are elaborated on and finally, some recommendations for further research are presented.

For this research, two explanations of the variation in the supply of roundtrip B2C shared cars in neighbourhoods in the G44 cities in the Netherlands have been developed. One is based on the number of roundtrip B2C shared cars and the characteristics of neighbourhoods in The Hague, Amsterdam and Utrecht and the other is based on the number of roundtrip B2C shared cars and the characteristics of neighbourhoods in 6 municipalities that together represent the G44 cities. The explanation of the variation in the supply of roundtrip B2C shared cars in the 6 municipalities has a better fit when based on the neighbourhoods in The Hague, Amsterdam and Utrecht than when it is based on the neighbourhoods in the 6 municipalities that represent the G44 cities themselves. A possible explanation of this result is that the variation in the neighbourhood types and number of shared cars per neighbourhood in The Hague, Amsterdam and Utrecht is higher than it is in the 6 municipalities.

Then it is analysed whether the explanation of the variation in the supply of roundtrip B2C shared cars can be used to explain the variation in the demand for roundtrip B2C carsharing as well. It was expected that there is an effect of the observed supply of roundtrip B2C shared cars on the demand for roundtrip B2C carsharing and that this effect would make it more difficult to use the explanation of the variation in the supply, to explain the variation in the demand for roundtrip B2C carsharing. However, the answers from the survey showed that there is no indication of an effect of the observed supply of roundtrip B2C shared cars on the demand for roundtrip B2C carsharing. This result can be explained by the fact that this research also showed that in general, individuals have a low awareness of whether a shared car is present in their neighbourhood or not.

To prove that the explanation of the variation in the supply of roundtrip B2C shared cars can be used to explain the variation in the demand for roundtrip B2C carsharing, it was also expected that the characteristics of the neighbourhoods of individuals that show the most demand for roundtrip B2C carsharing are similar to the characteristics of the neighbourhoods with the highest supply of roundtrip B2C shared cars. This research shows that this is the case when the neighbourhood characteristics are combined in a model that is explaining the variation in the demand for carsharing by using the same neighbourhood characteristics as are used to explain the variation in the supply of shared cars. Also, when using the model that explains the variation in the supply of shared cars to predict the shared car supply rate, it turns out that individuals that live in a neighbourhood with a high predicted shared car supply show a high demand for roundtrip B2C carsharing. This also indicates that there is a similar relation between the characteristics of the neighbourhoods of individuals that show the most demand for roundtrip B2C carsharing and the characteristics of the neighbourhoods with the highest supply of roundtrip B2C shared cars. This does mean that the explanation of the variation in the supply of roundtrip B2C shared cars over neighbourhoods in the G44 cities can only be used to identify neighbourhoods in the G44 cities that should have a relatively high demand for roundtrip B2C carsharing and can not be used to identify the difference in demand between neighbourhoods with predicted shared car supply rates that are close to each other. This could be caused by the fact that the correlations between the neighbourhood characteristics and the shared car supply rate are not as similar to the correlations between the same neighbourhood characteristics and the demand, as is the case for the correlations between the neighbourhood characteristics and the shared car presence.

Because the correlation between the demand for roundtrip B2C carsharing and the political preference was not significant, no statement could be made about the effect of a strong environmental ideology on the demand for carsharing. Therefore, only the first part of the hypothesis in which was expected that there is more demand for a roundtrip B2C shared car by individuals that have experience with other shared mobility and have strong environmental ideologies could be tested. The results of this research show that experience with a shared car and experience with other shared mobility turned out the most important characteristics of an individual to show demand for roundtrip B2C carsharing and therefore the first part of the hypothesis is accepted.

It was also expected that roundtrip B2C shared cars are used to have the convenience of having a car only when needed. Since the results show that this is not a popular motive, this part of the hypothesis is rejected. This could mean that the inhabitants of the G44 cities are not negative towards owning a private car and the responsibilities that come with it and believe that the advantages of owning a car outweigh the disadvantages. It was also expected that roundtrip B2C shared cars are mostly used to visit a friend or a relative and for a day out. This research did not completely meet this expectation since using a shared car to go to another city or village turned out to be more popular than using it for a day out. This could be caused by the coherence of both purposes since going to another city or village can be the same as going for a day out.

This research is an addition to the existing literature on the factors that explain the demand for and supply of roundtrip B2C shared cars because previous studies did not conclude on the relation between the factors that explain the demand for carsharing and the factors that explain the supply for shared cars. Also, most previous studies discuss the relation of the supply of shared cars in different cities or countries, but elaborating on this relation within the narrow scope of a neighbourhood is not very common. Making a distinction between factors of which data can be obtained without extensive research and factors of which data can only be obtained with for example a survey, is also a practical contribution to existing literature. Based on this research, policymakers and carsharing providers can identify the neighbourhood, district or municipality with the most demand for roundtrip B2C carsharing without doing extensive market research or can use this research to gain knowledge on the factors that affect the supply and demand for roundtrip B2C carsharing.

It must be taken into account, however, that this research shows a relation between a high predicted shared car supply rate and a high demand for roundtrip B2C carsharing, but no clear pattern between groups with a different predicted shared car supply rate. Therefore, the model can only be used to identify neighbourhoods with a high demand for roundtrip B2C carsharing and is not able to distinguish between neighbourhoods with a comparable predicted shared car supply rate. Also, to find the correlation between the supply of and the demand for carsharing, a survey was distributed that used closed questions. The respondents that completed the survey, were rewarded with points that they can exchange for a discount in particular webshops. The reliability of the findings in this survey may be less accurate because closed questions are easy to answer without reading the question and the reward after filling in the survey may be an incentive to not truthfully answer the question in the survey. Another limitation to the model is that since it predicts the number of roundtrip B2C shared cars per 100,000 inhabitants, the predictions are not reliable for a neighbourhood with only a few inhabitants since the observed number of shared cars per 100,000 inhabitants is probably underestimated by the model. In addition, it must be taken into account that this research is only focusing on the supply of and demand for roundtrip B2C carsharing. Once the focus is on a different business model or organizational structure of carsharing or carsharing in general, the results could be different.

Additionally, when identifying the characteristics of an individual with a high demand for roundtrip B2C carsharing, some characteristics could only be proxied through indicators that give room to flaws and missed particularities. For example, it was assumed that an individual's attitude towards different transport modes can be translated directly to the individual's opinion on the importance of a good flow or connection of a certain transport mode. Also, the political preference of an individual was assumed to indicate how environmentally minded the individual is. Further research into attitudes might have deepened the understanding of the characteristics of an individual with a high demand for roundtrip B2C carsharing.

For further research, it is recommended to get more insight into the relationship between the demand for carsharing and the supply of shared cars by an enlarged data set with respondents that are equally distributed over all cities that are part of the G44 cities in the Netherlands. With this data set, the results could be better generalised to fit the variation in the demand for roundtrip B2C carsharing in all of the G44 cities. Also, the analysis of the factors that affect the individual demand for carsharing, showed that the modal split of an individual turned out a very important factor. The explanation of the variation in the supply of shared cars is only using the urbanization level of a neighbourhood as an indicator for the modal split. Therefore, it is recommended to do more research into finding a better indicator for the modal split in a neighbourhood to be able to improve the model developed in this research. Nevertheless, the most important recommendation is to do further research on how to increase the awareness of the presence of shared cars in a neighbourhood to boost the demand for carsharing. Because when inhabitants do not know about the presence of a shared car in their neighbourhood, they will never consider using it.

References

Abraham, J. E. (1999). A survey of carsharing preferences. *World Transport Policy and Practice*, 1–12. Retrieved from https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.186.3784&rep=rep1&type=pdf

Aedes. (2021). *Brede wooncoalitie gaat woningtekort aanpakken en woonklimaat verbeteren.* Retrieved 20-06-2021, from https://www.aedes.nl/artikelen/woningmarkt/verkiezingen/tweede-kamer/brede-wooncoalitie-gaat-woningtekort-aanpakken-en-woonklimaat-verbeteren.html

Autodelen. (2021, May 20). *Cijfers en onderzoek.* Retrieved 25-07-2021, from https://autodelen.info/cijfers-autodelen

Becker, H., Ciari, F., & Axhausen, K. W. (2017). Comparing car-sharing schemes in switzerland: User groups and usage patterns. *Transportation Research Part A: Policy and Practice*, *97*, 17–29. doi: https://doi.org/10.1016/j.tra.2017.01.004

Bhandari, A. (2020). What is multicollinearity? here's everything you need to know. *Analytics Vidhya.* Retrieved from https://www.analyticsvidhya.com/blog/2020/03/what-is-multicollinearity/

Brownlee, J. (2019). *Probabilistic Model Selection with AIC, BIC, and MDL.* Retrieved 27-10-2021, from https://machinelearningmastery.com/probabilistic-model-selection-measures/

CBS. (2017). *Ritten van inwoners per provincie naar vervoerswijze, 2014/2016.* Retrieved from https://www.cbs.nl/-/media/_excel/2017/45/ovin-2014-2016-modal-split-provincies.xlsx

CBS. (2018). *Kerncijfers wijken en buurten 2018 [Data set].* Retrieved from https://www.cbs.nl/nl-nl/maatwerk/2018/30/kerncijfers-wijken-en-buurten-2018

CBS. (2019). *Kerncijfers wijken en buurten 2019 [Data set].* Retrieved from https://www.cbs.nl/nl-nl/maatwerk/2019/31/kerncijfers-wijken-en-buurten-2019

CBS. (2020a). *Autopark groeit sterker dan bevolking.* Retrieved 21-06-2021, from https://www.cbs.nl/nl-nl/nieuws/2020/10/autopark-groeit-sterker-dan-bevolking#:%7E:text=Nederland%20telde%20begin%202020%20bijna,op%20naam%20van%20een%20bedrijf.

CBS. (2020b). *Kerncijfers wijken en buurten 2020 [Data set].* Retrieved from https://www.cbs.nl/-/media/_excel/2021/12/kwb-2020.xls

CBS. (2022). *Gemeentegrootte en stedelijkheid.* Retrieved from https://www.cbs.nl/nl-nl/onze-diensten/methoden/classificaties/overig/gemeentegrootte-en-stedelijkheid

CBS Statline. (2020). *Centraal bureau voor de statistiek statline nabijheid voorzieningen [data set].* Retrieved from https://opendata.cbs.nl/statline/#/CBS/nl/dataset/84953NED/table?ts=1627397554756

Celsor, C., & Millard-Ball, A. (2007). Where does carsharing work? *Transportation Research Record: Journal of the Transportation Research Board*, *1992*(1), 61–69. doi: https://doi.org/10.3141/1992-08

Ciari, F., Weis, C., & Balac, M. (2016). Evaluating the influence of carsharing stations' location on potential membership: a swiss case study. *EURO Journal on Transportation and Logistics*, *5*(3), 345–369. doi: https://doi.org/10.1007/s13676-015-0076-6

Coxe, S., West, G., S, & Aiken, S., L. (2008). The analysis of count data: A gentle introduction to poisson regression and its alternatives. *Journal of Pesronality Assessment*, *91*, 121–136. doi: https://doi.org/10.1080/00223890802634175

CROW. (2021a, April). *CROW Nieuws - CROW.* Retrieved 24-07-2021, from https://www.crow.nl/over-crow/nieuws/2021/april/terugblik-op-tweede-webinar-autodelen-en-beleid

CROW. (2021b). *Dashboard autodelen - crow.* Retrieved 21-06-2021, from https://www.crow.nl/dashboard-autodelen/

CROW. (2021c). *Jive [Data set].* Retrieved 20-07-2021, from https://crow.databank.nl/jive?cat_open_code=cgdghlOghVLIx&var=duurz_ind

CROW. (2021d). *Wie is de autodeler?* Retrieved 10-12-2021, from https://www.crow.nl/downloads/pdf/verkeer-en-vervoer/crow-kpvv/factsheet-autodelen-3-wie-is-de-autodeler.aspx#:~:text=Deelauto's%20worden%20met%20name%20gebruikt,Dit%20zijn%20vaak%20langere%20ritten.&text=De%20auto's%20die%20autodelers%20wegdeden,dan%20het%20gemiddelde%20Nederlandse%20wagenpark.

Date, S. (2017). *The Negative Binomial Regression Model.* Retrieved 22-10-2021, from https://timeseriesreasoning.com/contents/negative-binomial-regression-model/

De Boer, B. (2020). *Autodelen in 2020 waar staan we nu?* Presentation slides. Retrieved from https://advier.nl/wp-content/uploads/2020/10/Presentatie-autodeelcijfers-NVC-definitief.pdf

Federal Highway Administration. (2016). *Accident Models for Two-Lane Rural Roads: Segment and Intersections.* Retrieved 2-11-2021, from https://www.fhwa.dot.gov/publications/research/safety/98133/ch05_01.cfm

Ferrero, F., Perboli, G., Rosano, M., & Vesco, A. (2018). Car-sharing services: An annotated review. *Sustainable Cities and Society*, *37*, 501–518. doi: https://doi.org/10.1016/j.scs.2017.09.020

FleetCarma. (2018, October 1). *The environmental benefits of carsharing.* Retrieved 28-06-2021, from https://www.fleetcarma.com/environmental-benefits-car-sharing/

Gardner, W., Mulvey, E. P., & Shaw, E. C. (1995). Regression analyses of counts and rates: Poisson, overdispersed poisson, and negative binomial models. *Psychological Bulletin*, *118*(3), 392–414. doi: https://doi.org/10.1037/0033-2909.118.3.392

Gemeente Amsterdam. (2021). *Wat kost een parkeervergunning voor bewoners in amsterdam?* Retrieved 21-09-2021, from https://www.amsterdam.nl/veelgevraagd/?caseid=%7B49FE150C-D91F-4022-8CFF-4FE610B7E616%7D

Gemeente Den Haag. (2021). *Parkeervergunning voor bewoners aanvragen.* Retrieved 21-09-2021, from https://www.denhaag.nl/nl/parkeren/parkeervergunning-bewoners-en-hun-bezoek-aanvragen/parkeervergunning-voor-bewoners-aanvragen.htm

Gemeente Utrecht. (2021). *Parkeerkosten.* Retrieved 21-09-2021, from https://www.utrecht.nl/wonen-en-leven/parkeren/parkeerkosten/#c316781

Griffin, S. (2020). *Determining a Cut-Off or Threshold When Working With a Binary Dependent (Target) Variable.* Retrieved 25-10-2021, from https://medium.com/swlh/determining-a-cut-off-or-threshold-when-working-with-a-binary-dependent-target-variable-7c2342cf2a7c

Hanck, C., Arnold, M., Gerber, A., & Schmelzer, M. (2021). *Introduction to ecnometrics with r.* Essen, Germany: Department of Business Administration and Economics, University of Duisburg-Essen. Retrieved from https://www.econometrics-with-r.org/11.2-palr.html

Hardin, W., J, & Hilbe, J. (2014). Regression models for count data based on the negative binomial(p) distribution. *The Stata Journal*, *14*(2), 280–291. doi: https://journals.sagepub.com/doi/pdf/10.1177/1536867X1401400203

Humphreys, M. (2015). *Missing data how to deal: An overview of missing data.* Presentation slides. Retrieved from https://liberalarts.utexas.edu/prc/_files/cs/Missing-Data.pdf

Jaffe, A. B., Trajtenberg, M., & Fogarty, M. S. (2000). Knowledge spillovers and patent citations: Evidence from a survey of inventors. *American Economic Review*, *90*(2), 215–218. doi: https://doi.org/10.1257/aer.90.2.215

Jakobsen, J., Gluud, C., Wetterslev, J., & Winkel, P. (2017). When and how should multiple imputation be used for handling missing data in randomised clinical trials – a practical guide with flowcharts. *BMC Medical Research Methodology*, *17*(162). doi: https://doi.org/10.1186/s12874-017-0442-1

JMP. (2021). *Multiple Regression Residual Analysis and Outliers.* Retrieved 22-11-2021, from https://www.jmp.com/en_nl/statistics-knowledge-portal/what-is-multiple-regression/mlr-residual-analysis-and-outliers.html

Kieskompas. (2021). *Kieskompas 2021.* Retrieved 12-12-2021, from https://pbs.twimg.com/media/EwoW_QPXIAIHZlq?format=jpg&name=4096x4096

Krztoń, M. (2018, May 15). *Car sharing — everything you should know.* Retrieved 04-07-2021, from https://medium.com/@michakrzton/car-sharing-future-of-mobility-ffb6b363eb6a

Laerd. (2018). Binomial logistic regression analysis using stata. *Lund Research Ltd*. Retrieved from https://statistics.laerd.com/spss-tutorials/binomial-logistic-regression-using-spss-statistics.php

Lane, C. (2005). Phillycarshare: First-year social and mobility impacts of carsharing in philadelphia, pennsylvania. *Transportation Research Record: Journal of the Transportation Research Board*, *1927*(1), 158–166. doi: https://doi.org/10.1177\%2F0361198105192700118

Lee, B.-Y., Aultman-Hall, L., Coogan, M., & Adler, T. (2016). Rideshare mode potential in non-metropolitan areas of the northeastern united states. *Journal of Transport and Land Use*, *9*(3), 111–126. doi: https://doi.org/10.5198/jtlu.2015.669

Leppink, J. (2021). *Do Likert scales like being averaged? Let's ask them!* Retrieved 15-12-2021, from https://www.maastrichtuniversity.nl/nl/news-events/newsletters/article/HDpIb2dvqSEckXtCFerwIw

Le Vine, S., Zolfaghari, A., & Polak, J. (2014). Carsharing: Evolution, challenges and opportunities. *Centre for Transport Studies, Imperial College London*. Retrieved from https://www.acea.auto/uploads/publications/SAG_Report_-_Car_Sharing.pdf

Martin, E., & Shaheen, S. (2016). Impacts of car2go on vehicle ownership, modal shift, vehicle miles traveled, and greenhouse gas emissions: An analysis of five north american cities. *Innovative Mobility*. Retrieved from http://innovativemobility.org/wp-content/uploads/2016/07/Impactsofcar2go_FiveCities_2016.pdf

McClave, T., J, Benson, G., P, Sincich, T., & Knypstra, S. (2011). *Statistiek: Een inleiding* (11th ed.). Amsterdam: Pearson Benelux BV.

McLeod, S. A. (2019). What a p-value tells you about statistical significance. *Simply Psychology*. Retrieved from https://www.simplypsychology.org/p-value.html

MoveMove. (2021). *De poolauto slimme oplossing voor mkb'ers.* Retrieved 21-07-2021, from https://movemove.com/blog/de-poolauto-slimme-oplossing-voor-mkbers

MyWheels. (2021). *Where do i return the car to?* Retrieved 30-06-2021, from https://mywheels.nl/en/help/rijden/ophalen-en-parkeren/waar-breng-ik-de-auto-terug

Müller, J., Correia, G., & Bogenberger, K. (2017). An explanatory model approach for the spatial distribution of free-floating carsharing bookings: A case-study of german cities. *Sustainability*, *9*(7), 1290. doi: https://doi.org/10.3390/su9071290

Münzel, K., Boon, W., Frenken, K., Blomme, J., & van der Linden, D. (2020). Explaining carsharing supply across western european cities. *International Journal of Sustainable Transportation*, *14*(4), 243–254. doi: https://doi.org/10.1080/15568318.2018.1542756

Münzel, K., Boon, W., Frenken, K., & Vaskelainen, T. (2018). Carsharing business models in germany: characteristics, success and future prospects. *Information Systems and e-Business Management*, *16*(2), 271–291. doi: https://doi.org/10.1007/s10257-017-0355-x

Münzel, K., Piscicelli, L., Boon, W., & Frenken, K. (2019). Different business models – different users? uncovering the motives and characteristics of business-to-consumer and peer-to-peer carsharing adopters in the netherlands. *Transportation Research Part D: Transport and Environment*, *73*, 276–306. doi: https://doi.org/10.1016/j.trd.2019.07.001

NCSS. (2021). *Negative Binomial Regression.* Retrieved 25-10-2021, from https://ncss-wpengine.netdna-ssl.com/wp-content/themes/ncss/pdf/Procedures/NCSS/

Negative_Binomial_Regression.pdf

O'Hara, R., & Kotze, J. (2010). Do not log-transform count data. *British Ecological Society*, *1*(2). doi: https://doi.org/10.1111/j.2041-210X.2010.00021.x

O'Neill, A. (2020). *Urbanization in the netherlands 2019.* Retrieved 20-06-2021, from https://www.statista.com/statistics/276724/urbanization-in-the-netherlands/

Oots, S. (2018). The next step for p2p carsharing: How to reach early majority? *University of Utrecht*. Retrieved from https://dspace.library.uu.nl/bitstream/handle/1874/368054/Sander%20Oots%20thesis.pdf?sequence=2&isAllowed=y

Pennsylvania State University. (2021). *Identifying Outliers (Unusual y Values).* Retrieved 15-12-2021, from https://online.stat.psu.edu/stat501/lesson/11/11.3

Prieto, M., Baltas, G., & Stan, V. (2017). Carsharing adoption intention in urban areas: What are the key sociodemographic drivers? *Transportation Research Part A: Policy and Practice*, *101*, 218–227. doi: https://doi.org/10.1016/j.tra.2017.05.012

Ramos, R. M. S., Bergstad, C. J., Chicco, A., & Diana, M. (2020). Mobility styles and carsharing use in europe: attitudes, behaviours, motives and sustainability. *European Transport Research Review*, *12*(1). doi: https://doi.org/10.1186/s12544-020-0402-4

Schüppert, A. (2009). *Binomial (or binary) logistic regression.* Presentation slides. Retrieved from http://www.let.rug.nl/nerbonne/teach/rema-stats-meth-seminar/presentations/Binary-Logistic-Regression-Schueppert-2009.pdf

Shaheen, S. (2018). Shared mobility: The potential of ridehailing and pooling. *Three Revolutions*, 55–76. doi: https://doi.org/10.5822/978-1-61091-906-7_3

Shaheen, S., Cohen, A., & Zohdy, I. (2016). Shared mobility: current practices and guiding principles. *Unites States Department of Transportation: Federal Highway Administration*. Retrieved from https://ops.fhwa.dot.gov/publications/fhwahop16022/fhwahop16022.pdf

Shaheen, S., & Cohen, A. P. (2007). Growth in worldwide carsharing. *Transportation Research Record: Journal of the Transportation Research Board*, *1992*(1), 81–89. doi: https://doi.org/10.3141/1992-10

Sieben, I., & Linssen, L. (2009). *Logistische regressie analyse: een handleiding .* Retrieved 9-11-2021, from https://www.ru.nl/publish/pages/865809/logistischeregressie.pdf

SnappCar. (2021). *Snappcar keyless: Vaker verhuren, meer verdienen, geen gedoe.* Retrieved 21-07-2021, from https://keyless.snappcar.nl/

Statistics Solutions. (2021a). *Binary Logistic Regressions.* Retrieved 29-10-2021, from https://www.statisticssolutions.com/free-resources/directory-of-statistical-analyses/logistic-regression-assumptions/

Statistics Solutions. (2021b). *Missing Data: Listwise vs. Pairwise.* Retrieved 29-10-2021, from https://www.statisticssolutions.com/missing-data-listwise-vs-pairwise/

Statistics.com. (2021). *Validation Set.* Retrieved 29-10-2021, from https://www.statistics.com/glossary/validation-set/

Steiger, M. (2011). *The pros and cons of one-way carsharing.* Retrieved 02-07-2021, from https://earthgarage.com/2011/10/06/eco-friendly-car-sharing-gas-savings/

SurveyMonkey. (2021). *Sample Size Calculator: Understanding Sample Sizes.* Retrieved 10-07-2021, from https://www.surveymonkey.com/mp/sample-size-calculator/

Tardi, C. (2020). *What is the 80-20 rule?* Retrieved 2-11-2021, from https://www.investopedia.com/terms/1/80-20-rule.asp

Te Morsche, W., La Paix Puello, L., & Geurs, K. (2019). Potential uptake of adaptive transport services: An exploration of service attributes and attitudes. *Transport policy*, *84*, 1–11. doi: https://doi.org/10.1016/j.tranpol.2019.09.001

Truffer, B. (2003). User-led innovation processes: The development of professional carsharing by environmentally concerned citizens. *Innovation: The European Journal of Social Science Research*, *16*(2), 139–154. doi: https://doi.org/10.1080/13511610304517

TU Delft, O. (2016). *Parkeren.* Retrieved 20-06-2021, from https://ocw.tudelft.nl/course-readings/parkeren/#:%7E:text=Voor%20haaksparkeren%20zijn%20parkeervakken%20van,de%20straatbreedte%206m%20breed%20zijn.

Van den Berg, L. (2017). The road to the future of cars: Opportunities and key features for car sharing services. *Radboud University*. Retrieved from https://www.ru.nl/publish/pages/769526/z_lennard_van_den_berg.pdf

Van der Molen, F. (2021). *2000ste deelauto in nederland geplaatst door greenwheels in 2020.* Retrieved 30-06-2021, from https://www.duurzaam-ondernemen.nl/2000ste-deelauto-in-nederland-geplaatst-door-greenwheels-in-2020/

Wang, J., Dane, G., & Timmermans, H. (2021). Individuals who have zero-interest in living in carsharing-facilitating neighbourhoods: a case study in the netherlands. *European Planning Studies*, 1–17. doi: https://doi.org/10.1080/09654313.2021.1903840

Wang, J., Dane, G. Z., & Timmermans, H. J. P. (2020). Carsharing-facilitating neighbourhood choice: a mixed logit model. *Journal of Housing and the Built Environment*, *36*, 1033–1054. doi: https://doi.org/10.1007/s10901-020-09791-z

Zaiontz, C. (2015). *Classification Table.* Retrieved 25-10-2021, from https://www.real-statistics.com/descriptive-statistics/roc-curve-classification-table/classification-table/

Zhang, Z. (2016). Missing data imputation: focusing on single imputation. *ATM*, *4*(1), 9. doi: https://dx.doi.org/10.3978\%2Fj.issn.2305-5839.2015.12.38

Appendix A

Carsharing providers and platforms

TABLE A.1: Carsharing providers and platforms

Roundtrip/Station-based	**Free-floating**	**P2P platforms**	**Local communities**	**Business carsharing**
CareCar	Fetch Carsharing	Snappcar	A2B sharing	ALD sharing
ConnectCar	LEV	Vansharing	Axxel	Alphacity
Drive carsharing	Share’n Go	MyWheels	Coöperatieauto	Amber Mobility
E-mobility Park	ShareNow		DEEL	Arval carsharing
Easy Driving			HET Coöperatie	AudiSharedFleet
Elektrip			Huub	Car2Use
GoAbout			Izoof	CommShare
Greenwheels			BiroShare	Driessen Business Carshare
Hely				Free2Move
Juuve				FreeToGo
Kav2Go				Goodmoovs
MobilNoord				Greenwheels
MyWheels				Kyoto Share
StudentCar				Mijn Domein Auto
Tuk				Mobeazy
				Mobiliteitsmeesters

Appendix B

Sources used to collect data on independent variables used in the explanation of the variation in the supply of roundtrip B2C shared cars

TABLE B.1: Sources and approaches used to collect data on each independent variable

Independent variable	**Source**
Address density	'Wijken en buurten 2020' (CBS, 2020b)
Public transport network	'Nabijheid voorzieningen' (CBS Statline, 2020)
City scale	'Wijken en buurten 2020' (CBS, 2020b)
Parking permit costs	Websites municipality of The Hague, Amsterdam and Utrecht[1] (Gemeente Den Haag, 2021) (Gemeente Amsterdam, 2021) (Gemeente Utrecht, 2021)
Centrality	*Analysis in ArcGIS*
Bar and restaurant density	'Nabijheid voorzieningen' (CBS Statline, 2020)
Age	'Wijken en buurten 2020' (CBS, 2020b)
Gender	'Wijken en buurten 2020' (CBS, 2020b)
Education level	'Wijken en buurten 2019' (CBS, 2019)
Ethnicity	'Wijken en buurten 2020' (CBS, 2020b)
Modal split & pedestrian and bike commuting regime	Modal split per urbanization level in 'CBS Statline Mobiliteit' (CBS, 2017). The urbanization level of every neighbourhood in 'Wijken en buurten 2020' (CBS, 2020b)
Job type	*Not available*
Income	'Wijken en buurten 2018' (CBS, 2018)
Household car ownership	'Wijken en buurten 2020' (CBS, 2020b)
Rental households	'Wijken en buurten 2019' (CBS, 2019)
One-person households	'Wijken en buurten 2020' (CBS, 2020b)
Households with children	'Wijken en buurten 2020' (CBS, 2020b)

[1] *When a neighbourhood is located in multiple parking zones the average parking permit costs has been used.*

Appendix C

Complete survey

In this appendix the complete survey is presented. The original survey is in Dutch and has been translated to English.

C.1 Welcome text

Dear participant,

This questionnaire is part of a graduation research at the University of Twente. The research is being conducted in collaboration with mobility consultancy Goudappel. The study looks at the demand for carsharing in different types of neighbourhoods in the Netherlands. You can also complete this questionnaire if you have never used a shared car.

A shared car is a car that is used by multiple people, this is also called carsharing. The main difference with a rental car is that a shared car is often nearby, so you can quickly walk or cycle to reach it instead of having to go to the car rental. You can often pick up and return shared cars 24 hours a day and 7 days a week, because you open them with an app or card. In addition, it is also easier to rent a shared car for a few hours. The questionnaire is based on shared cars from a provider that has several shared cars and therefore not a shared car from a private individual.

The questionnaire consists of 3 parts. In part 1 it is checked whether you belong to the target group. In part 2 you will be asked about your experience with shared mobility and your travel behaviour. In the final part you will be asked about some personal characteristics. Your participation is anonymous, your answers will be treated confidentially and will only be used for this study. Goudappel works according to the guidelines for information security and market research (ISO 27001 and ISO 20252).

Your participation is voluntary and you can stop at any time. By participating in this questionnaire, you agree to the aforementioned conditions and you give permission that the data you have entered may be used for this research.

Filling in the questionnaire takes about 10 minutes and it works best on a computer or tablet.

Thank you for your participation!

For questions please contact: avossebeld@goudappel.nl

C.2 Background information target group

Are you in possession of a car driver's license?
Yes
No

In which municipality do you live?
Almere
Arnhem
Enschede
Nijmegen
Zoetermeer
Zwolle
I live in a different municipality

What age do you have?
0 to 18 years old
18 to 25 years old
25 to 45 years old
45 to 65 years old
65 years old or older

What is your gender?
Male
Female
Other / I prefer not to say

C.3 Experience with shared mobility and travel behaviour

Have you ever made use of a shared car?
Yes
No
I do not know

Have you ever made use of a shared car?
Yes
No
I do not know

How often did you make use of a shared car in the past year?[1]
4 times a week or more
1-3 times a week
1-3 times a month
6-11 times a year
3-5 times a year
1-2 times a year
Never

[1]This question pops up when a respondent indicated to ever made use of a shared car

Have you ever used other forms of shared mobility? (e.g. a shared scooter, shared bicycle (OV-bicycle), shared moppet, etc.)
Yes
No
I do not know

How often did you make use of other forms of shared mobility in the past year? (e.g. a shared scooter, shared bicycle (OV-bicycle), shared moppet, etc.)[1]
4 times a week or more
1-3 times a week
1-3 times a month
6-11 times a year
3-5 times a year
1-2 times a year
Never

[1]This question pops up when a respondent indicated to ever made use of other forms of shared mobility

To what extent do you agree with the statements below?[1]					
	1. Totally disagree	2. Disagree	3. Neutral	4. Agree	5. Totally agree
Motives for using a shared car					
I am using a shared car to reduce mobility expenses					
I am using a shared car because there are carsharing parking places near my house					
I am using a shared car to avoid owning a private car myself					
I am using a shared car to avoid maintenance and repair					
I am using a shared car to avoid searching for parking places					
I am using a shared car to drive more environmentally friendly					
Purposes for using a shared car					
The shared car is an attractive option to go to work or study with when the shared car is within walking distance					
The shared car is an attractive option to do the groceries with when the shared car is within walking distance					
The shared car is an attractive option to use for visiting a friend or relative when the shared car is within walking distance					
The shared car is an attractive option to use to go to an activity/day out (like a concert, sports match etc.) when the shared car is within walking distance					
The shared car is an attractive option to use to go to a different city or village when the shared car is within walking distance					

Attitude towards different transport modes					
A good road network and car traffic flow are important, regardless of the costs for a municipality or government					
Good infrastructure and flow for bicycles are important, regardless of the costs for a municipality or government					
A fast connection and good availability of public transport is important, regardless of the costs for the public transport provider or government					
It is important that a shared car is placed or will be placed in my neighbourhood					

[1]This question pops up when a respondent indicated to ever made use of a shared car

To what extent do you agree with the statements below?[2]					
	1. Totally disagree	2. Disagree	3. Neutral	4. Agree	5. Totally agree
Motives for using a shared car					
I am considering using a shared car to reduce mobility expenses					
I am considering using a shared car when there are carsharing parking places near my house					
I am considering using a shared car to avoid owning a private car myself					
I am considering using a shared car to avoid maintenance and repair					
I am considering using a shared car to avoid searching for parking places					
I am considering using a shared car to drive more environmentally friendly					
Purposes for using a shared car					
The shared car is an attractive option to go to work or study with when the shared car is within walking distance					
The shared car is an attractive option to do the groceries with when the shared car is within walking distance					
The shared car is an attractive option to use for visiting a friend or relative when the shared car is within walking distance					
The shared car is an attractive option to use to go to an activity/day out (like a concert, sports match etc.) when the shared car is within walking distance					

The shared car is an attractive option to use to go to a different city or village when the shared car is within walking distance					
Attitude towards different transport modes					
A good road network and car traffic flow are important, regardless of the costs for a municipality or government					
Good infrastructure and flow for bicycles are important, regardless of the costs for a municipality or government					
A fast connection and good availability of public transport is important, regardless of the costs for the public transport provider or government					
It is important that a shared car is placed or will be placed in my neighbourhood					

[2]This question pops up when a respondent indicated to never made use of a shared car before

Is there a shared car present in your neighbourhood?
Yes
No
I do not know

How many times did you make use of the following transport modes to get to another location in the past year?							
	4 times a week or more	1-3 times a week	1-3 times a month	6-11 times a year	3-5 times a year	1-2 times a year	Never
Private car / private van (also as a passenger)							
Private bicycle							
Walking							
Taxi							
Motorcycle / Moped / Scooter							
Public transport (PT)							

When there is always a shared car within walking distance of your house, how likely is it that you make use of a shared car at least once a year?
Note: To make the assessment easier, the costs of using a shared car are indicated below
Costs of carsharing: *Most providers offer different subscriptions. Greenwheels is used here as an example. At Greenwheels, a shared car costs €6 per hour + €0.34 per km (fuel costs included) if you do not take out a subscription. If you pay €10 per month it is €4 per hour + €0.29 per km and at €25 per month it is €3 per hour and €0.24 per km. The subscriptions can be used by 3 drivers.*
1. I would definitely not make use of the shared car
2. It is not likely that I would make use of the shared car
3. Neutral/I do not know whether I would make use of the shared car
4. I would probably make use of the shared car
5. I would definitely make use of the shared car

When there is always a shared car within walking distance of your house, how likely is it that you would make use of it for the following purposes at least once a year?					
	1. Totally not likely	2. Not likely	3. Neutral	4. Probably	5. Definitely
To go to work or study					
To do the groceries					
To visit a friend or relative					
To go to an activity / To go for a day out (e.g. a concert or sports match)					
To go to a different city or village					

How many private cars are owned within your household?
0
1
2
3
4 or more

When there is always a shared car within walking distance of your house, how likely is it that do not buy a new car when the current car in your household is due for replacement?
Note: In order to make the assessment easier, the costs of the shared car compared to the private car are indicated below.
Cost of shared car compared to own car: The Consumentenbond calculated in March 2010 that the critical mileage limit, at which the costs of a shared car and the cost of a private car are approximately equal, is around 12,000 kilometers per year. So at about 1,000 kilometers per month. They also calculated that buying your own car is cheaper if you need a car for more than 120 days a year.
I have no influence on the replacement of the car in my household
1. I am / we are definitely buying a new car
2. I am / we are probably buying a new car
3. Neutral/I am / we are possibly buying a new car
4. I am / we are probably not buying a new car
5. I am / we are definitely not buying a new car

* This question pops up when the respondent indicates to own exactly one car in their household

When there is always a shared car within walking distance of your house, how likely is it that do not buy a new car at all when the current cars in your household are due for replacement?
Note: In order to make the assessment easier, the costs of the shared car compared to the private car are indicated below.
Cost of shared car compared to own car: The Consumentenbond calculated in March 2010 that the critical mileage limit, at which the costs of a shared car and the cost of a private car are approximately equal, is around 12,000 kilometers per year. So at about 1,000 kilometers per month. They also calculated that buying your own car is cheaper if you need a car for more than 120 days a year.
I have no influence on the replacement of a car in my household
1. I am / we are definitely buying a new car
2. I am / we are probably buying a new car
3. Neutral/I am / we are possibly buying a new car
4. I am / we are probably not buying a new car
5. I am / we are definitely not buying a new car

* This question pops up when the respondent indicates to own 2 cars or more in their household

When there is always a shared car within walking distance of your house, how likely is it that do not buy a new car when the second (or third etc.) car in your household is due for replacement?
Note: In order to make the assessment easier, the costs of the shared car compared to the private car are indicated below.
Cost of shared car compared to own car: *The Consumentenbond calculated in March 2010 that the critical mileage limit, at which the costs of a shared car and the cost of a private car are approximately equal, is around 12,000 kilometers per year. So at about 1,000 kilometers per month. They also calculated that buying your own car is cheaper if you need a car for more than 120 days a year.*
I have no influence on the replacement of a car in my household
I only have influence on the replacement of 1 car in my household
1. I am / we are definitely buying a new second car
2. I am / we are probably buying a new second car
3. Neutral/I am / we are possibly buying a new second car
4. I am / we are probably not buying a new second car
5. I am / we are definitely not buying a new second car

* This question pops up when the respondent indicates to own 2 cars or more in their household

C.4 Background information

What is your zip code?
Note: For example 1234AB
Open question

What is your year of birth?
Open question

What is the bird's-eye view of the distance between your house and the centre of your city or village?
Note: The bird's-eye view of the distance is the the shortest distance between 2 points. This is when a straight line is drawn between your house and the centre of your city or village.
0 - 500 meter (0 - 0.5 km)
500 - 1,000 meter (0.5 - 1.0 km)
1,000 - 2,000 meter (1.0 - 2.0 km)
2,000 - 5,000 meter (2.0 - 5.0 km)
5,000 - 10,000 meter (5.0 - 10.0 km)
10,000 meter or more (10.0 km or more)

What is the quarterly rate for a first residents permit to use a parking place in your street?
Note: In (surrounding neighbourhoods of) the city centres of many cities, a parking permit is required to be allowed to park. Many municipalities use different rates to distinguish between the number of permits per household. The first resident permit is the rate for the first resident in a household that applies for a parking permit.
There is no parking permit necessary to park in my street
€0,- - €10,- per quartile
€10,- - €20,- per quartile
€20,- - €50,- per quartile
€50,- - €80,- per quartile
€80,- - €110,- per quartile
More than €110,- per quartile
A parking permit is necessary to park in my street, but I do not know what the costs are.
I do not know whether a parking permit is necessary to park in my street.

Political preferences are often referred to as 'left' or 'right' orientated. In general, where would you place your political views on a scale from 1 (left) to 5 (right)?
Note: Nowadays, progressive parties are referred to as 'left' and more conservative parties as 'right'. Left-wing parties are in favor of a larger role for government in social life, while right-wing parties want to limit that role.
1. Left (BIJ1, PvdD, SP)
2. Central-left (GroenLinks, DENK, PvdA)
3. Central (D66, Volt, ChristenUnie, 50PLUS, CDA, PVV)
4. Central-right (BBB, SGP, VVD)
5. Right (FVD, JA21)
I prefer not to answer this question

Which of the following situations applies (the most) to you?
I am working (full time or part-time)
I am (partly) incapacitated or unemployed
I am retired
I am studying or doing an internship
I am a housewife/houseman
I am doing voluntary work
I do not know
A different situation applies to me

What is the highest level of education you have completed?
No education, elementary school, VMBO, the first three years of HAVO/VWO and/or MBO1
HAVO, VWO, MBO2, MBO3 or MBO4
HBO, university bachelor, university master or doctor
I do not know
Other

Do you have a migration background?
Note: A person has a migration background when at least one of the parents was born abroad *Western migration background:* *Individuals with a migration background with as origin group one of the countries in the continents Europe (excl. Turkey), North-America and Oceania or Indonesia or Japan.* *Non-western migration background:* *Individuals with a migration background with as origin group one of the countries in the continents Africa, Latin America and Asia (excl. Indonesia and Japan) or Turkey.*
No
Yes, I have a western migration background
Yes, I have a non-western migration background

What is the composition of your household?
A single-person household
Single, with (a) child(ren) living at home
Multiple-person household without children living at home
Multiple-person households with children living at home
Other

How many hours of paid work do you do on average per week?
I do now have paid work
12 hours or less
12 to 29 hours per week
30 hours per week or more

What is your personal monthly net income (the amount of money you receive per month)?
Less than € 1.000,-
€ 1.000 to € 2.500
€ 2.500 to € 5.000
€ 5.000 or more
I prefer not to say

Do you live in an owner-occupied or rental house?
Owner-occupied house
Rental house
Other

Appendix D

Correlation table of independent variables used in the explanation of the variation in the supply for roundtrip B2C shared cars

	Shared car presence	Shared car supply rate	Address density per km2	Average distance to train station	Number of inhabitants in municipality	Centrality	Number of restaurants within 1 kilometre	Number of bars within 1 kilometre	Percentages of individuals between 0 and 14 years old	Percentages of individuals between 15 and 24 years old	Percentages of individuals between 25 and 44 years old	Percentages of individuals between 45 and 64 years old
Shared car presence		0,133	0,427	-0,276	0,022	-0,376	0,262	0,200	-0,004	-0,084	0,139	-0,071
Shared car supply rate	0,133		0,005	-0,065	0,124	-0,014	0,025	0,017	-0,220	0,165	0,002	0,182
Address density per km2	0,427	0,005		-0,388	0,247	-0,568	0,695	0,634	-0,239	-0,133	0,365	-0,112
Average distance to train station	-0,276	-0,065	-0,388		0,127	0,342	-0,279	-0,250	0,202	-0,111	-0,381	0,247
Number of inhabitants in municipality	0,022	0,124	0,247	0,127		-0,319	0,195	0,186	-0,205	-0,047	0,072	0,137
Centrality	-0,376	-0,014	-0,568	0,342	-0,319		-0,376	-0,334	0,124	0,073	-0,264	0,109
Number of restaurants within 1 kilometre	0,262	0,025	0,695	-0,279	0,195	-0,376		0,933	-0,350	-0,048	0,311	-0,059
Number of bars within 1 kilometre	0,200	0,017	0,634	-0,250	0,186	-0,334	0,933		-0,310	-0,027	0,281	-0,047
Percentages of individuals between 0 and 14 years old	-0,004	-0,220	-0,239	0,202	-0,205	0,124	-0,350	-0,310		-0,211	-0,368	-0,039
Percentages of individuals between 15 and 24 years old	-0,084	0,165	-0,133	-0,111	-0,047	0,073	-0,048	-0,027	-0,211		-0,013	-0,435
Percentages of individuals between 25 and 44 years old	0,139	0,002	0,365	-0,381	0,072	-0,264	0,311	0,281	-0,368	-0,013		-0,522
Percentages of individuals between 45 and 64 years old	-0,071	0,182	-0,112	0,247	0,137	0,109	-0,059	-0,047	-0,039	-0,435	-0,522	
Percentages of individuals of 65 years old or older	-0,097	-0,033	-0,092	0,227	0,037	0,102	-0,012	-0,035	-0,204	-0,365	-0,524	0,312
Percantage of male individuals	-0,219	-0,117	-0,049	-0,028	0,099	0,146	0,111	0,143	-0,083	0,049	0,165	-0,052
Percantage of female individuals	0,235	0,118	0,057	0,018	-0,106	-0,151	-0,110	-0,143	0,085	-0,056	-0,155	0,052
Percentage of lower educated individuals	-0,224	-0,138	-0,205	0,165	-0,019	0,106	-0,300	-0,211	0,272	-0,035	-0,231	0,080
Percentage of medium educated individuals	-0,068	0,168	-0,237	0,179	-0,086	0,182	-0,215	-0,156	0,014	0,209	-0,254	0,104
Percentage of higher educated individuals	0,205	0,001	0,294	-0,238	0,067	-0,189	0,358	0,253	-0,211	-0,094	0,334	-0,134
Percentage of individuals with a western immigration background	0,115	0,175	0,339	-0,160	0,270	-0,191	0,452	0,425	-0,347	0,125	0,215	-0,147
Percentage of individuals with a non-western immigration background	-0,075	-0,063	-0,101	-0,088	0,125	-0,141	-0,278	-0,192	0,261	0,119	0,029	-0,128
Percentages of rental houses	0,094	-0,252	0,327	-0,286	0,219	-0,350	0,191	0,181	-0,237	0,151	0,426	-0,397
Percentage of one-person households	0,056	0,110	0,316	-0,279	0,170	-0,227	0,318	0,281	-0,605	0,368	0,378	-0,342
Percentage of households without children	0,093	-0,132	0,082	0,112	-0,051	0,017	0,126	0,094	-0,039	-0,428	0,047	0,138
Percentage of households with children	-0,124	-0,065	-0,415	0,276	-0,153	0,250	-0,441	-0,383	0,706	-0,216	-0,465	0,322
Percentage of individuals with low income	-0,207	-0,208	-0,152	0,016	0,012	0,075	-0,210	-0,134	-0,093	0,452	-0,051	-0,153
Percentage of individuals with high income	0,131	0,281	0,106	-0,050	0,036	0,001	0,221	0,132	0,062	-0,230	-0,030	0,148
Number of private cars per household	-0,256	0,370	-0,430	0,203	-0,063	0,401	-0,251	-0,222	0,079	-0,070	-0,238	0,284
Urbanization_class_1	0,412	-0,071	0,651	-0,358	0,075	-0,522	0,329	0,288	-0,119	-0,168	0,241	-0,128
Urbanization_class_2	-0,166	0,066	-0,370	0,212	-0,046	0,141	-0,218	-0,192	0,170	0,089	-0,135	0,015
Urbanization_class_3	-0,130	0,023	-0,270	0,079	-0,090	0,192	-0,138	-0,124	0,075	0,071	-0,156	0,075
Urbanization_class_4	-0,237	0,068	-0,233	0,087	-0,045	0,269	-0,098	-0,082	-0,045	0,072	-0,062	0,080
Urbanization_class_5	-0,280	-0,046	-0,297	0,264	0,058	0,476	-0,110	-0,091	-0,085	0,077	-0,070	0,108
Distance to city centre	-0,345	-0,018	-0,596	0,277	0,000	0,425	-0,506	-0,479	0,233	0,087	-0,343	0,145
First resident parking rate	0,323	0,102	0,709	-0,358	0,337	-0,468	0,754	0,683	-0,355	-0,069	0,327	-0,019

FIGURE D.1: Correlation table part 1

	Percentages of individuals of 65 years old or older	Percantage of male individuals	Percantage of female individuals	Percentage of lower educated individuals	Percentage of medium educated individuals	Percentage of higher educated individuals	Percentage of individuals with a western immigration background	Percentage of individuals with a non-western immigration background	Percentages of rental houses	Percentage of one-person households	Percentage of households without children
Shared car presence	-0,097	-0,219	0,235	-0,224	-0,068	0,205	0,115	-0,075	0,094	0,056	0,093
Shared car supply rate	-0,033	-0,117	0,118	-0,138	0,168	0,001	0,175	-0,063	-0,252	0,110	-0,132
Address density per km2	-0,092	-0,049	0,057	-0,205	-0,237	0,294	0,339	-0,101	0,327	0,316	0,082
Average distance to train station	0,227	-0,028	0,018	0,165	0,179	-0,238	-0,160	-0,088	-0,286	-0,279	0,112
Number of inhabitants in municipality	0,037	0,099	-0,106	-0,019	-0,086	0,067	0,270	0,125	0,219	0,170	-0,051
Centrality	0,102	0,146	-0,151	0,106	0,182	-0,189	-0,191	-0,141	-0,350	-0,227	0,017
Number of restaurants within 1 kilometre	-0,012	0,111	-0,110	-0,300	-0,215	0,358	0,452	-0,278	0,191	0,318	0,126
Number of bars within 1 kilometre	-0,035	0,143	-0,143	-0,211	-0,156	0,253	0,425	-0,192	0,181	0,281	0,094
Percentages of individuals between 0 and 14 years old	-0,204	-0,083	0,085	0,272	0,014	-0,211	-0,347	0,261	-0,237	-0,605	-0,039
Percentages of individuals between 15 and 24 years old	-0,365	0,049	-0,056	-0,035	0,209	-0,094	0,125	0,119	0,151	0,368	-0,428
Percentages of individuals between 25 and 44 years old	-0,524	0,165	-0,155	-0,231	-0,254	0,334	0,215	0,029	0,426	0,378	0,047
Percentages of individuals between 45 and 64 years old	0,312	-0,052	0,052	0,080	0,104	-0,134	-0,147	-0,128	-0,397	-0,342	0,138
Percentages of individuals of 65 years old or older		-0,123	0,112	0,020	0,041	-0,046	0,049	-0,302	-0,191	-0,020	0,256
Percantage of male individuals	-0,123		-0,997	0,119	-0,074	-0,043	0,267	0,120	0,096	0,152	-0,129
Percantage of female individuals	0,112	-0,997		-0,118	0,079	0,038	-0,264	-0,108	-0,087	-0,158	0,126
Percentage of lower educated individuals	0,020	0,119	-0,118		0,007	-0,745	-0,300	0,558	0,194	-0,178	-0,270
Percentage of medium educated individuals	0,041	-0,074	0,079	0,007		-0,663	-0,081	0,169	-0,078	-0,085	-0,140
Percentage of higher educated individuals	-0,046	-0,043	0,038	-0,745	-0,663		0,284	-0,534	-0,098	0,193	0,283
Percentage of individuals with a western immigration background	0,049	0,267	-0,264	-0,300	-0,081	0,284		-0,331	-0,007	0,333	0,100
Percentage of individuals with a non-western immigration background	-0,302	0,120	-0,108	0,558	0,169	-0,534	-0,331		0,478	-0,008	-0,474
Percentages of rental houses	-0,191	0,096	-0,087	0,194	-0,078	-0,098	-0,007	0,478		0,502	-0,357
Percentage of one-person households	-0,020	0,152	-0,158	-0,178	-0,085	0,193	0,333	-0,008	0,502		-0,482
Percentage of households without children	0,256	-0,129	0,126	-0,270	-0,140	0,283	0,100	-0,474	-0,357	-0,482	
Percentage of households with children	-0,080	-0,104	0,110	0,338	0,167	-0,363	-0,426	0,248	-0,400	-0,906	0,072
Percentage of individuals with low income	-0,167	0,184	-0,183	0,453	0,192	-0,462	-0,325	0,603	0,591	0,342	-0,617
Percentage of individuals with high income	0,101	0,056	-0,057	-0,485	-0,260	0,536	0,483	-0,624	-0,588	-0,232	0,482
Number of private cars per household	0,139	0,172	-0,178	0,097	0,044	-0,090	0,077	-0,181	-0,567	-0,294	0,144
Urbanization_class_1	0,007	-0,093	0,102	-0,088	-0,159	0,148	0,123	0,063	0,345	0,262	0,048
Urbanization_class_2	-0,055	0,011	-0,009	0,062	0,144	-0,123	-0,003	0,050	-0,159	-0,202	-0,052
Urbanization_class_3	0,020	0,077	-0,076	0,003	0,078	-0,042	-0,096	-0,036	-0,168	-0,129	-0,015
Urbanization_class_4	0,019	-0,008	0,002	0,050	0,030	-0,057	-0,068	-0,036	-0,205	-0,068	-0,006
Urbanization_class_5	0,044	0,111	-0,131	0,038	-0,014	-0,014	-0,107	-0,164	-0,123	-0,020	0,009
Distance to city centre	0,069	0,034	-0,035	0,295	0,254	-0,379	-0,321	0,258	-0,215	-0,309	-0,132
First resident parking rate	-0,042	0,030	-0,030	-0,374	-0,323	0,484	0,410	-0,312	0,219	0,336	0,128

FIGURE D.2: Correlation table part 2

	Percentage of households with children	Percentage of individuals with low income	Percentage of individuals with high income	Number of private cars per household	Urbanization_class_1	Urbanization_class_2	Urbanization_class_3	Urbanization_class_4	Urbanization_class_5	Distance to city centre	First resident parking rate
Shared car presence	-0,124	-0,207	0,131	-0,256	0,412	-0,166	-0,130	-0,237	-0,280	-0,345	0,323
Shared car supply rate	-0,065	-0,208	0,281	0,370	-0,071	0,066	0,023	0,068	-0,046	-0,018	0,102
Address density per km2	-0,415	-0,152	0,106	-0,430	0,651	-0,370	-0,270	-0,233	-0,297	-0,596	0,709
Average distance to train station	0,276	0,016	-0,050	0,203	-0,358	0,212	0,079	0,087	0,264	0,277	-0,358
Number of inhabitants in municipality	-0,153	0,012	0,036	-0,063	0,075	-0,046	-0,090	-0,045	0,058	0,000	0,337
Centrality	0,250	0,075	0,001	0,401	-0,522	0,141	0,192	0,269	0,476	0,425	-0,468
Number of restaurants within 1 kilometre	-0,441	-0,210	0,221	-0,251	0,329	-0,218	-0,138	-0,098	-0,110	-0,506	0,754
Number of bars within 1 kilometre	-0,383	-0,134	0,132	-0,222	0,288	-0,192	-0,124	-0,082	-0,091	-0,479	0,683
Percentages of individuals between 0 and 14 years old	0,706	-0,093	0,062	0,079	-0,119	0,170	0,075	-0,045	-0,085	0,233	-0,355
Percentages of individuals between 15 and 24 years old	-0,216	0,452	-0,230	-0,070	-0,168	0,089	0,071	0,072	0,077	0,087	-0,069
Percentages of individuals between 25 and 44 years old	-0,465	-0,051	-0,030	-0,238	0,241	-0,135	-0,156	-0,062	-0,070	-0,343	0,327
Percentages of individuals between 45 and 64 years old	0,322	-0,153	0,148	0,284	-0,128	0,015	0,075	0,080	0,108	0,145	-0,019
Percentages of individuals of 65 years old or older	-0,080	-0,167	0,101	0,139	0,007	-0,055	0,020	0,019	0,044	0,069	-0,042
Percantage of male individuals	-0,104	0,184	0,056	0,172	-0,093	0,011	0,077	-0,008	0,111	0,034	0,030
Percantage of female individuals	0,110	-0,183	-0,057	-0,178	0,102	-0,009	-0,076	0,002	-0,131	-0,035	-0,030
Percentage of lower educated individuals	0,338	0,453	-0,485	0,097	-0,088	0,062	0,003	0,050	0,038	0,295	-0,374
Percentage of medium educated individuals	0,167	0,192	-0,260	0,044	-0,159	0,144	0,078	0,030	-0,014	0,254	-0,323
Percentage of higher educated individuals	-0,363	-0,462	0,536	-0,090	0,148	-0,123	-0,042	-0,057	-0,014	-0,379	0,484
Percentage of individuals with a western immigration background	-0,426	-0,325	0,483	0,077	0,123	-0,003	-0,096	-0,068	-0,107	-0,321	0,410
Percentage of individuals with a non-western immigration background	0,248	0,603	-0,624	-0,181	0,063	0,050	-0,036	-0,036	-0,164	0,258	-0,312
Percentages of rental houses	-0,400	0,591	-0,588	-0,567	0,345	-0,159	-0,168	-0,205	-0,123	-0,215	0,219
Percentage of one-person households	-0,906	0,342	-0,232	-0,294	0,262	-0,202	-0,129	-0,068	-0,020	-0,309	0,336
Percentage of households without children	0,072	-0,617	0,482	0,144	0,048	-0,052	-0,015	-0,006	0,009	-0,132	0,128
Percentage of households with children		-0,087	0,031	0,270	-0,328	0,263	0,164	0,078	0,014	0,428	-0,463
Percentage of individuals with low income	-0,087		-0,857	-0,303	-0,073	0,009	0,024	0,014	0,110	0,165	-0,261
Percentage of individuals with high income	0,031	-0,857		0,391	-0,047	0,021	0,088	0,012	-0,044	-0,177	0,303
Number of private cars per household	0,270	-0,303	0,391		-0,421	0,169	0,179	0,209	0,262	0,266	-0,248
Urbanization_class_1	-0,328	-0,073	-0,047	-0,421		-0,682	-0,390	-0,297	-0,331	-0,483	0,385
Urbanization_class_2	0,263	0,009	0,021	0,169	-0,682		-0,097	-0,074	-0,083	0,271	-0,268
Urbanization_class_3	0,164	0,024	0,088	0,179	-0,390	-0,097		-0,042	-0,047	0,150	-0,085
Urbanization_class_4	0,078	0,014	0,012	0,209	-0,297	-0,074	-0,042		-0,036	0,193	-0,122
Urbanization_class_5	0,014	0,110	-0,044	0,262	-0,331	-0,083	-0,047	-0,036		0,266	-0,187
Distance to city centre	0,428	0,165	-0,177	0,266	-0,483	0,271	0,150	0,193	0,266		-0,579
First resident parking rate	-0,463	-0,261	0,303	-0,248	0,385	-0,268	-0,085	-0,122	-0,187	-0,579	

FIGURE D.3: Correlation table part 3

Appendix E

All individual and neighbourhood characteristics that are used as input for the explanation of the variation in the demand for roundtrip B2C shared cars

TABLE E.1: List of all independent variables

Independent variable
Experience with shared car
Experience with other shared mobility
Use of own car/own van in past year
Use of public transport in past year
Attitude towards good cycle infrastructure and flow for bikes
Experience with shared car in past year
Number of private cars in household
Attitude towards good public transport infrastructure and connection
Living in a owner-occupied home
Attitude towards good road network and flow for cars
Living in a rental house
Use of walking as transport mode in past year
Use of motorcycle/moped/scooter in past year
Highly educated
18 to 25 years old
House type other than rental/owner-occupied
Use of bicyle as transport mode in past year
First resident parking rate in neighbourhood
Age
Studying or doing an internship
Not working
Number of hours of paid work
Number of inhabitants in municipality
Monthly net income
Moderately educated
Retired
Low educated
Experience with other shared mobility in past year
Different household composition than one or mulitple person with or without children
65 years old or older

Independent variable
Working 30 hours per week or more
One-person household without children living at home
Urbanization class 2
Number of inhabitants in neighbourhood
Married, with children living at home
Working 0 to 12 hours per week
Urbanization class 4
Political left oriented
(Partially) incapacitated for work or unemployed
Average distance to train station
Number of bars in 1km
Urbanization class 5
Low income
Distance to city centre
Number of restaurants in 1km
Working (full time/part time)
No migration background
One-person household with children living at home
Western migration background
Male
Female
Urbanization class 3
Multiple-person household, without children living at home
Address density
Non-western migration background
Urbanization class 1
Doing voluntary work
25 to 45 years old
Working 12 to 30 hours per week
I have a different working situation
Political right oriented
Use of a taxi in past year
Different education level than low, moderately or high
High income
45 to 65 years old

Appendix F

The distribution of the observed shared car supply rate in The Hague, Amsterdam and Utrecht

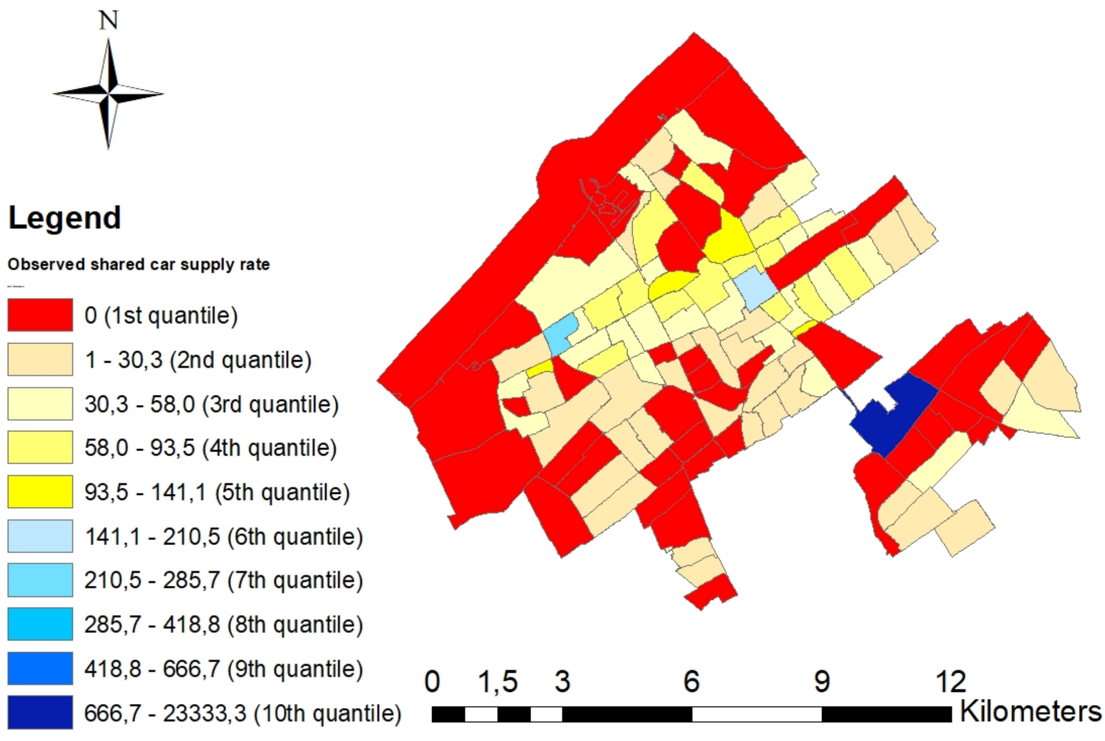

FIGURE F.1: The distribution of the observed shared car supply rate in The Hague

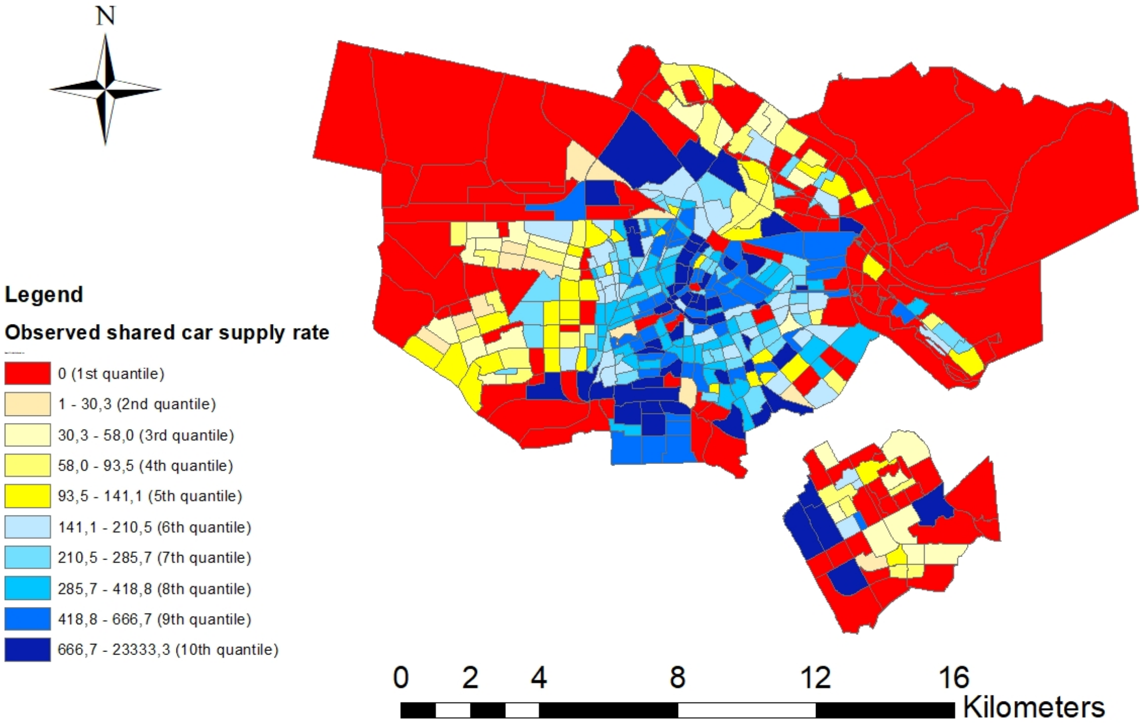

FIGURE F.2: The distribution of the observed shared car supply rate in Amsterdam

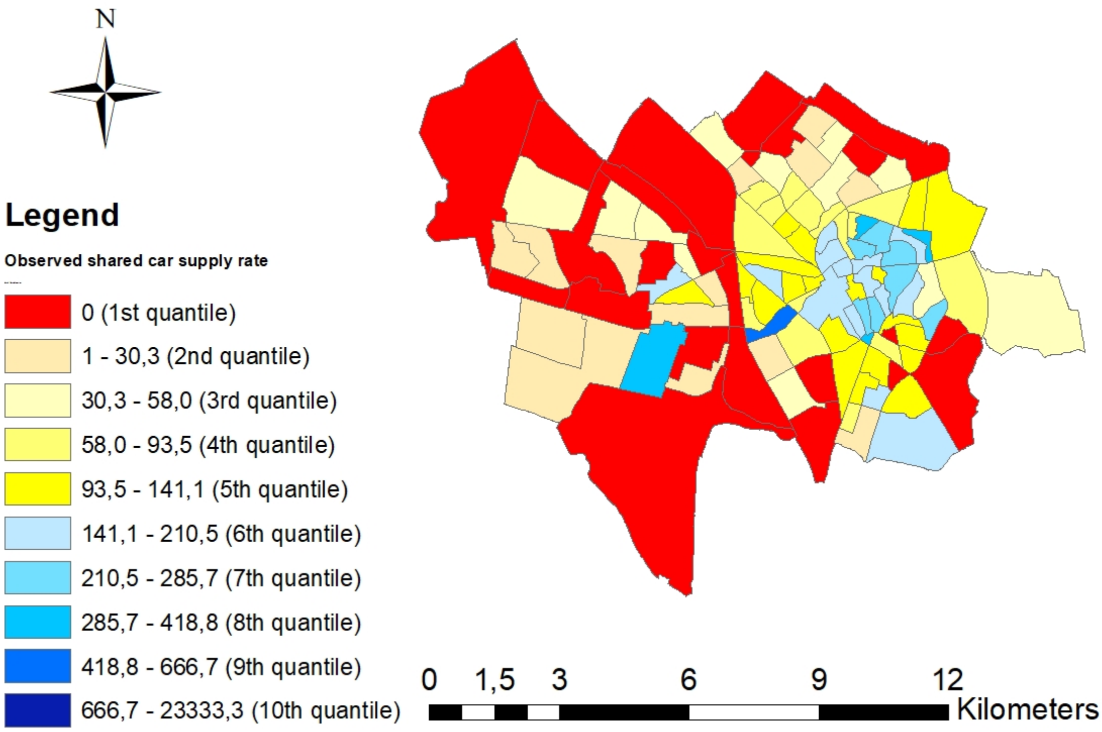

FIGURE F.3: The distribution of the observed shared car supply rate in Utrecht

Appendix G

Significance of the correlation of the neighbourhood characteristics with the shared car presence and shared car supply rate

TABLE G.1: Correlation of the neighbourhood factors with the shared car presence and shared car supply rate

Variable	Correlation with shared car presence in The Hague, Amsterdam and Utrecht	Significance	Correlation with shared car presence in survey municipalities	Significance	Correlation with shared car supply rate in The Hague, Amsterdam and Utrecht	Significance	Correlation with shared car supply rate in survey municipalities	Significance
Address density	0.43	0.000	0.32	0.000	0.01	0.896	0.35	0.000
Urbanization class 1	0.41	0.000	0.23	0.000	-0.07	0.059	0.28	0.000
Distance to city centre	-0.35	0.000	-0.20	0.000	-0.02	0.628	-0.20	0.000
First resident parking rate	0.32	0.000	0.09	0.039	0.10	0.007	0.24	0.000
Average distance to train station	-0.28	0.000	-0.40	0.000	-0.07	0.083	-0.30	0.000
Urbanization class 5	-0.28	0.000	-0.10	0.000	-0.05	0.224	-0.06	0.113
Number of private cars in household	-0.26	0.000	-0.07	0.094	0.37	0.000	-0.07	0.079
Number of restaurants in 1km	0.26	0.000	0.35	0.000	0.03	0.511	0.40	0.000
Urbanization class 4	-0.24	0.000	-0.01	0.000	0.07	0.073	-0.02	0.000
Female	0.24	0.000	0.03	0.460	0.12	0.002	0.03	0.426
Low educated	-0.22	0.000	0.01	0.872	-0.14	0.000	-0.02	0.568
Male	-0.22	0.000	-0.03	0.460	-0.12	0.002	-0.03	0.426
Highly educated	0.21	0.000	0.06	0.150	0.00	0.984	0.04	0.334
Number of bars in 1km	0.2	0.000	0.37	0.000	0.02	0.661	0.39	0.000
Urbanization class 2	-0.17	0.000	-0.09	0.000	0.07	0.079	-0.14	0.000
25 to 45 years old	0.14	0.000	0.02	0.631	0.00	0.955	0.04	0.352
Urbanization class 3	-0.13	0.001	-0.11	0.000	0.02	0.538	-0.11	0.000

Variable	Correlation with shared car presence in The Hague, Amsterdam and Utrecht	Significance	Correlation with shared car presence in survey municipalities	Significance	Correlation with shared car supply rate in The Hague, Amsterdam and Utrecht	Significance	Correlation with shared car supply rate in survey municipalities	Significance
High income	0.13	0.000	-0.03	0.115	0.28	0.000	-0.01	0.068
Multiple person household with children living at home	-0.12	0.001	-0.03	0.396	-0.07	0.087	-0.09	0.023
Western migration background	0.11	0.002	0.03	0.400	0.18	0.000	0.03	0.486
65 years old or older	-0.1	0.010	0.01	0.753	-0.03	0.377	-0.01	0.887
Multiple person household without children living at home	0.09	0.014	-0.04	0.379	-0.13	0.000	-0.03	0.412
Rental house	0.09	0.012	0.05	0.236	-0.25	0.000	0.06	0.134
Non-western migration background	-0.08	0.046	-0.05	0.245	-0.06	0.093	-0.01	0.887
moderately educated	-0.07	0.072	-0.07	0.104	0.17	0.000	-0.04	0.368
45 to 65 years old	-0.07	0.059	-0.06	0.149	0.18	0.000	-0.06	0.113
One-person household	0.06	0.138	0.06	0.172	0.11	0.004	0.13	0.001
Number of inhabitants in municipality	0.02	0.559	0.11	0.005	0.12	0.001	0.11	0.666

Appendix H

Processes of finding the best fit to explain the variation in the supply

H.1 Processes of finding the best fit to explain the variation in the supply based on The Hague, Amsterdam and Utrecht

TABLE H.1: Process of finding the most suitable binary logistic regression model to predict the presence of a shared car

		p-value								
Regression model	**Nagelkerke R^2**	**Address density per km^2**	**Urbanization class**	**First resident parking rate**	**Perc. of female individuals**	**Perc. of low educated individuals**	**Perc. of indidividuals with low income**	**Perc. of indidividuals with high income**	**Perc. of medium educated individuals**	**Number of inhabitants in mun.**
1	*0.353*	0.002	0.047	0.059[2]						
2[1]	*0.391*	0.000	0.038		0.000					
3[1]	*0.401*	0.007	0.013	0.023	0.000					
4	*0.417*	0.004	0.017	0.113[2]	0.000	0.005				
5[1]	*0.412*	0.000	0.035		0.000	0.001				
6	*0.416*	0.005	0.017	0.065[2]	0.000		0.007			
7	*0.423*	0.004	0.020	0.158[2]	0.001	0.052[2]	0.069[2]			
8	*0.420*	0.005	0.011	0.155[2]	0.000			0.003		
9[1]	*0.409*	0.000	0.039		0.000		0.003			
10[1]	***0.416***	**0.000**	**0.022**		**0.000**			**0.001**		
11	*0.416*	0.000	0.021		0.000		0.780[2]	0.059[2]		
12	*0.416*	0.000	0.023		0.000			0.001	0.813[2]	
13	*0.416*	0.000	0.022		0.000			0.001		0.991[2]
14	*0.394*	0.000		0.801[2]	0.000			0.001		

[1] *Binary logistic regression model with independent variables that are all significant*
[2] *Independent variable that is not significant*

TABLE H.2: Process of finding the most suitable negative binomial regression model

			p-value																
Regression model	*AIC*	*BIC*	**Private cars per household**	**High income**	**Rental houses**	**Age between 0 and 14**	**Age between 45 and 64**	**Western migration background**	**Medium educated**	**Age between 15 and 24**	**Lower educated**	**Households without children**	**Inhabitants in municipality**	**Female individuals**	**Parking rate**	**Urbanization class**	**Distance to train station**	**Non-western migration background**	**Distance to city centre**
1[1]	*4693*	*4705*	0.000																
2[1]	*4570*	*4586*	0.000	0.000															
3[1]	*4546*	*4565*	0.000	0.000	0.000														
4[1]	*4517*	*4540*	0.016	0.000	0.000	0.000													
5[1]	*4504*	*4532*	0.003	0.000	0.000	0.000	0.001												
6	*4506*	*4538*	0.003	0.000	0.000	0.000	0.001	0.738[2]											
7	*4506*	*4538*	0.004	0.000	0.000	0.000	0.001		0.748[2]										
8	*4506*	*4538*	0.003	0.000	0.000	0.000	0.001			0.734[2]									
9	*4506*	*4537*	0.006	0.000	0.000	0.000	0.001				0.532[2]								
10[1]	*4498*	*4529*	0.023	0.000	0.000	0.000	0.008					0.008							
11	*4474*	*4509*	0.037	0.000	0.167[2]	0.000	0.499[2]					0.006	0.000						
12	*4473*	*4504*	0.047	0.000	0.209[2]	0.000						0.003	0.000						
13[1]	*4473*	*4500*	0.015	0.000		0.000						0.002	0.000						
14	*4473*	*4504*	0.015	0.000		0.000						0.001	0.000	0.254[2]					
15	*4469*	*4500*	0.231[2]	0.000		0.000						0.021	0.000		0.035				
16[1]	*4469*	*4496*		0.000		0.000						0.016	0.000		0.003				

			p-value																
Regression model	*AIC*	*BIC*	**Private cars per household**	**High income**	**Rental houses**	**Age between 0 and 14**	**Age between 45 and 64**	**Western migration background**	**Medium educated**	**Age between 15 and 24**	**Lower educated**	**Households without children**	**Inhabitants in municipality**	**Female indiv.**	**Parking rate**	**Urbanization class**	**Distance to train station**	**Non-western migration background**	**Distance to city centre**
17	*4470*	*4496*		0.000		0.000						0.021	0.000		0.003	0.358[2]			
18[1]	*4494*	*4505*		0.000		0.000						0.047	0.000		0.015		0.010		
19	*4453*	*4488*		0.002		0.000						0.022	0.000		0.126[2]		0.000	0.002	
20[1]	*4454*	*4485*		0.001		0.000						0.003	0.000				0.000	0.000	
21[1]	***4447***	***4482***		**0.000**		**0.000**						**0.007**	**0.000**				**0.000**	**0.013**	**0.008**

[1] *Negative binomial regression model with independent variables that are all significant*
[2] *Independent variable that is not significant*

H.2 Processes of finding the best fit to explain the variation in the supply based on the survey municipalities

TABLE H.3: Process of finding the most suitable binary logistic regression model to predict the presence of a shared car

		p-value											
Regression model	***Nagelkerke*** R^2	**Address density**	**Distance to city centre**	**Number of bars in 1 km**	**Number of restaurants in 1 km**	**Average distance to train station**	**One person households**	**Number of cars per household**	**Urbanization class**	**Highly educated**	**Male individuals**	**Age between 25 and 44**	**Female individuals**
1	*0.232*	0.000											
2[1]	*0.276*	0.001	0.002										
3	*0.272*	0.061	0.003	0.143[2]									
4	*0.267*	0.033	0.003		0.319[2]								
5	*0.293*	0.079[2]	0.004			0.006							
6[1]	*0.282*		0.000			0.001							
7	*0.294*		0.001	0.090[2]		0.004							
8	*0.288*		0.001		0.199[2]	0.002							
9	*0.284*		0.000			0.002	0.418[2]						
10	*0.282*		0.000			0.003		0.400[2]					
11	*0.301*		0.000			0.003		0.858[2]					
12[1]	*0.307*		0.000			0.002				0.008			
13	*0.313*		0.000			0.003				0.024	0.185[2]		
14	*0.308*		0.000			0.002				0.011		0.648[2]	
15[1]	***0.339***		**0.000**			**0.002**				**0.008**			**0.011**

[1] *Binary logistic regression model with independent variables that are all significant*
[2] *Independent variable that is not significant*

TABLE H.4: Process of finding the most suitable negative binomial regression model based on the survey municipalities

				p-value															
Regression model	*AIC*	*BIC*	*Log Likelihood*	**Number of restaurants in 1 km**	**Number of bars in 1 km**	**Address density**	**Urbanization class**	**Distance to city centre**	**One person households**	**Average distance to train station**	**Highly educated**	**Household with children**	**Number of cars per household**	**Age between 0 and 14**	**Age between 25 and 44**	**Low income**	**Age between 15 and 24**	**Age between 45 and 64**	**Female individuals**
1[1]	*1213*	*1224*	*-603*	0.000															
2	*1215*	*1230*	*-603*	0.200[2]	0.917[2]														
3[1]	*1210*	*1225*	*-601*	0.040		0.020													
4[1]	*1206*	*1237*	*-595*	0.000		0.010	0.010												
5[1]	*1185*	*1220*	*-584*	0.000		0.010	0.010	0.001											
6[1]	*1187*	*1225*	*-583*	0.000		0.017	0.015	0.001	0.000										
7[1]	*1189*	*1231*	*-583*	0.000		0.026	0.000	0.000	0.000	0.000									
8[1]	*1187*	*1225*	*-583*	0.000		0.000	0.014	0.000		0.000									
9[1]	*1184*	*1222*	*-582*	0.000		0.007	0.000	0.016			0.000								

				p-value															
Regression model	***AIC***	***BIC***	***Log Likelihood***	**Number of restaurants in 1 km**	**Number of bars in 1 km**	**Address density**	**Urbanization class**	**Distance to city centre**	**One person households**	**Average distance to train station**	**Highly educated**	**Household with children**	**Number of cars per household**	**Age between 0 and 14**	**Age between 25 and 44**	**Low income**	**Age between 15 and 24**	**Age between 45 and 64**	**Female individuals**
10[1]	*1185*	*1227*	*-581*	0.000		0.000	0.000	0.000			0.000	0.000							
11[1]	*1201*	*1246*	*-589*	0.000		0.000	0.000	0.000				0.000	0.000	0.000					
12[1]	*1183*	*1225*	*-581*	0.000		0.005	0.000	0.000			0.000			0.000					
13[1]	*1181*	*1227*	*-578*	0.000		0.000	0.000	0.000			0.000			0.000	0.000				
14[1]	*1180*	*1230*	*-577*	0.000		0.000	0.008	0.000			0.000			0.000	0.000	0.000			
15[1]	*1182*	*1235*	*-578*	0.000		0.000	0.008	0.000			0.000			0.000	0.000	0.000	0.000		
16[1]	*1183*	*1241*	*-577*	0.000		0.000	0.007	0.000			0.000			0.000	0.000	0.000	0.000	0.000	
17[1]	***1181***	***1242***	***-574***	0.000		**0.000**	**0.000**	**0.000**			**0.000**			**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**

[1] *Negative binomial regression model with independent variables that are all significant*
[2] *Independent variable that is not significant*

Appendix I

Validation of the binary logistic regression in The Hague, Amsterdam and Utrecht and in the survey municipalities

I.1 Classification tables of the binary logistic regressions based on both groups of municipalities

TABLE I.1: Classification table of binary logistic regression model based on the validation set in The Hague, Amsterdam and Utrecht

	Predicted neighbourhoods without shared car	**Predicted neighbourhoods with shared car**	***Percentage correct***
Observed neighbourhoods without shared car	19	19	*50.0%*
Observed neighbourhoods with shared car	12	91	*88.3%*
			78.0%

TABLE I.2: Classification table of binary logistic regression model based on the validation set in the survey municipalities

	Predicted neighbourhoods without shared car	**Predicted neighbourhoods with shared car**	***Percentage correct***
Observed neighbourhoods without shared car	52	6	*89.7%*
Observed neighbourhoods with shared car	9	7	*56.3%*
			79.7%

I.2 Plot of the predicted probability against the shared car presence for The Hague, Amsterdam and Utrecht and the survey municipalities

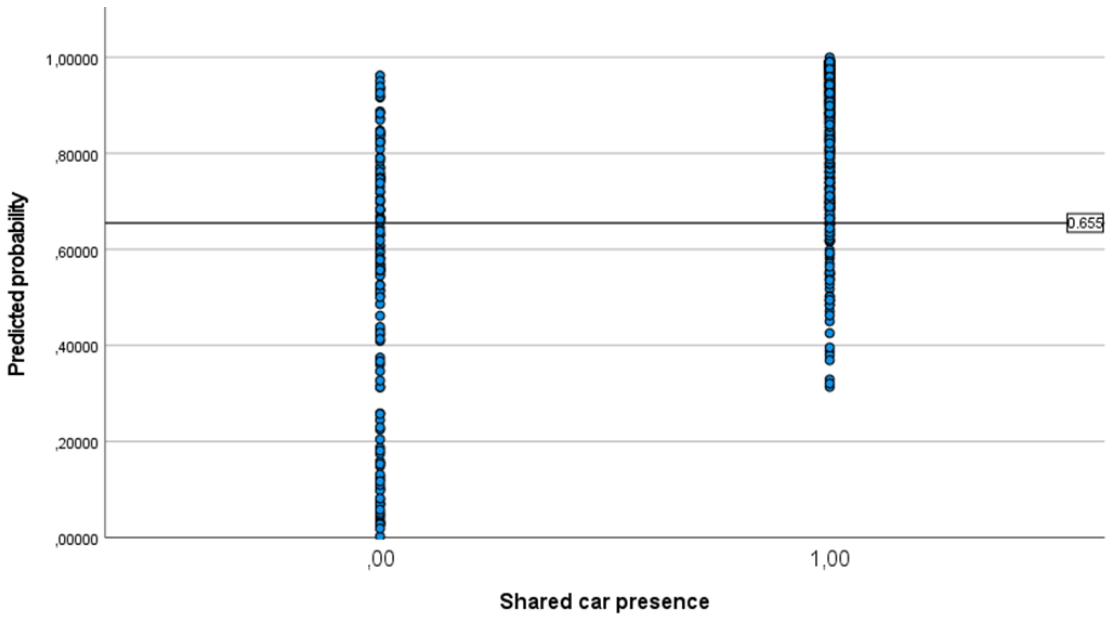

FIGURE I.1: Plot of the predicted probability against the shared car presence for The Hague, Amsterdam and Utrecht

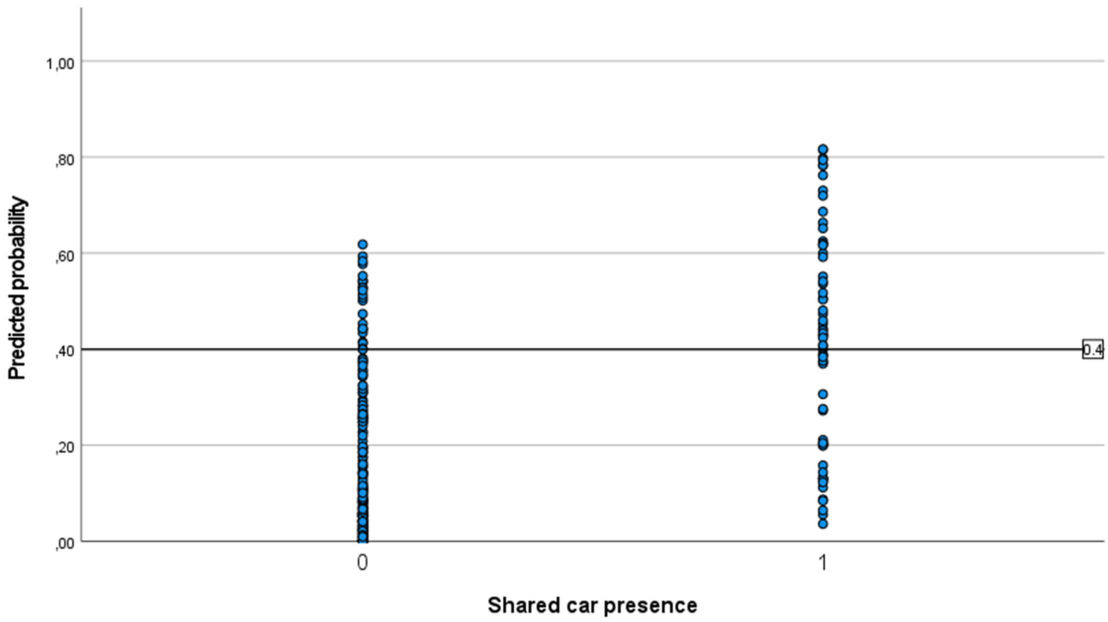

FIGURE I.2: Plot of the predicted probability against the shared car presence for the survey municipalities

Appendix J

Visualized results of regression model in The Hague, Amsterdam and Utrecht

J.1 Visualized results of the binary logistic regression

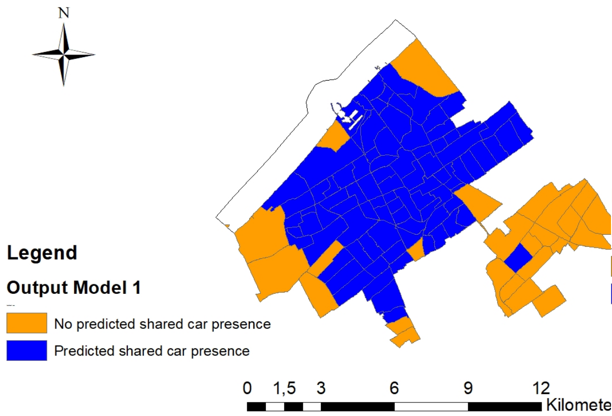

FIGURE J.1: Results of binary logistic regression model in The Hague

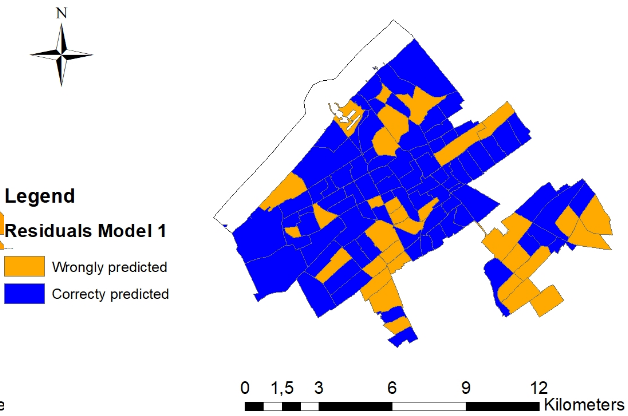

FIGURE J.2: Residuals of binary logistic regression model in The Hague

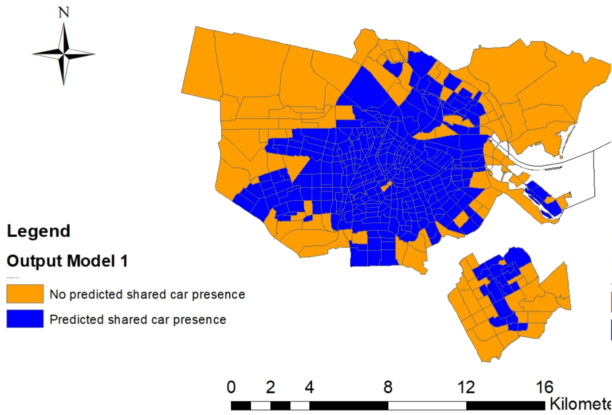

FIGURE J.3: Results of binary logistic regression model in Amsterdam

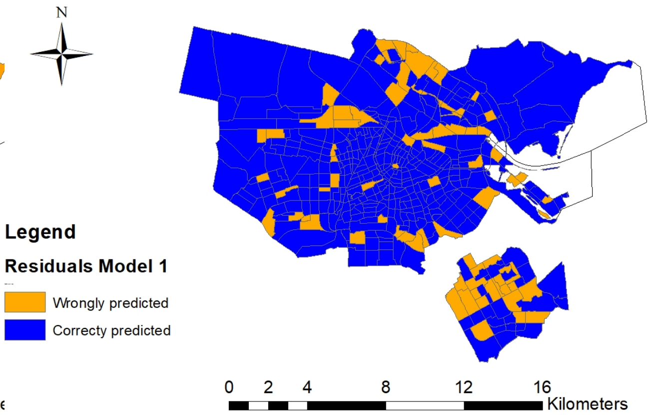

FIGURE J.4: Residuals of binary logistic regression model in Amsterdam

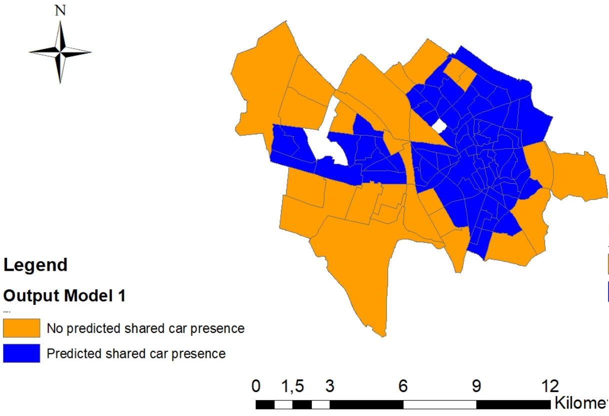

FIGURE J.5: Results of binary logistic regression model in Utrecht

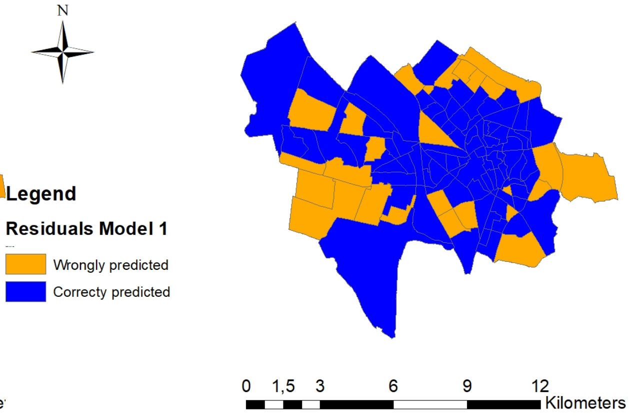

FIGURE J.6: Residuals of binary logistic regression model in Utrecht

J.2 Visualized results of the negative binomial regression

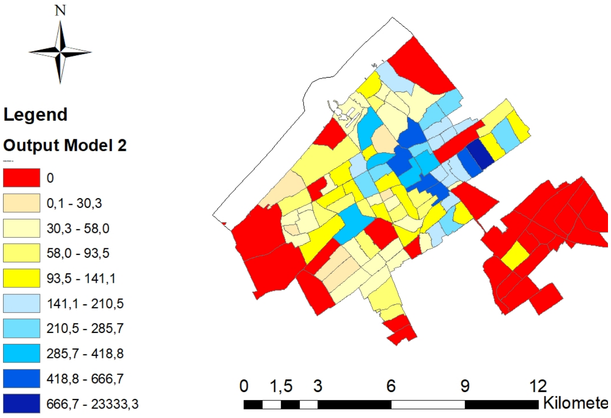

FIGURE J.7: Results of negative binomial regression model in The Hague

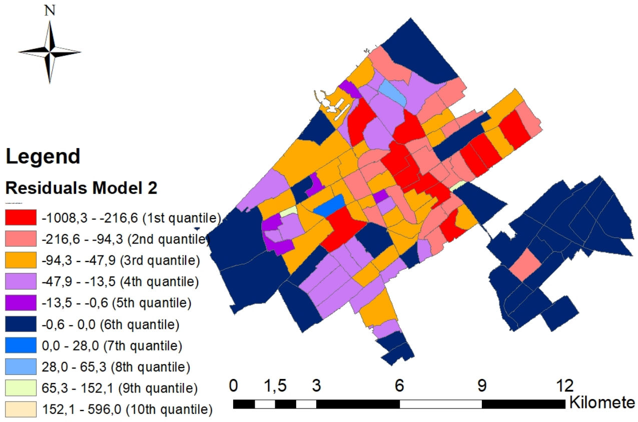

FIGURE J.8: Residuals of negative binomial regression model in The Hague

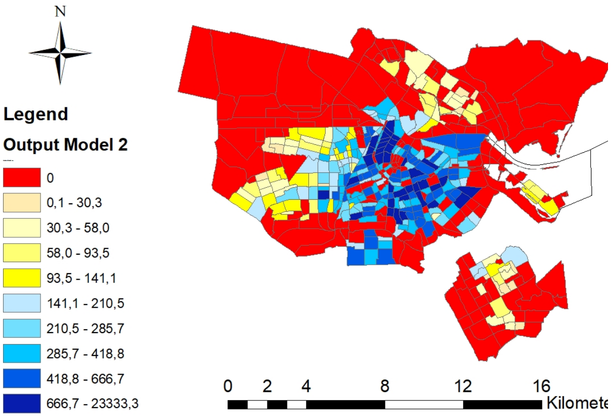

FIGURE J.9: Results of negative binomial regression model in Amsterdam

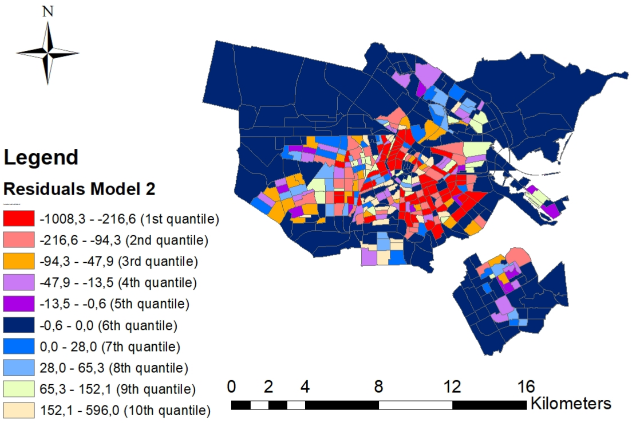

FIGURE J.10: Residuals of negative binomial regression model in Amsterdam

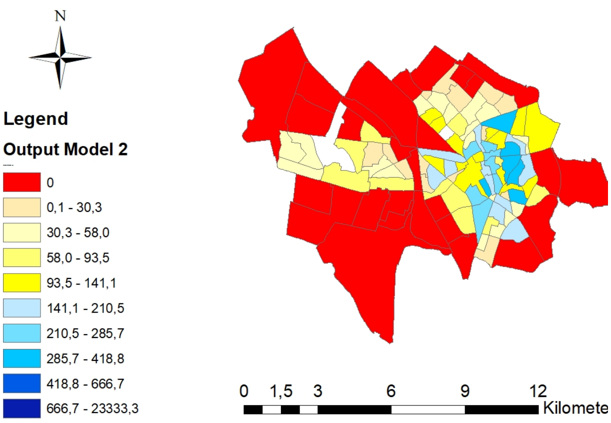

FIGURE J.11: Results of negative binomial regression model in Utrecht

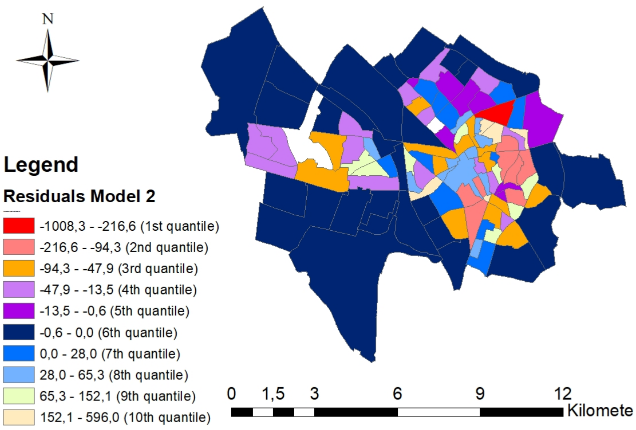

FIGURE J.12: Residuals of negative binomial regression model in Utrecht

Appendix K

Plots of residuals against each independent variable in the negative binomial regression model

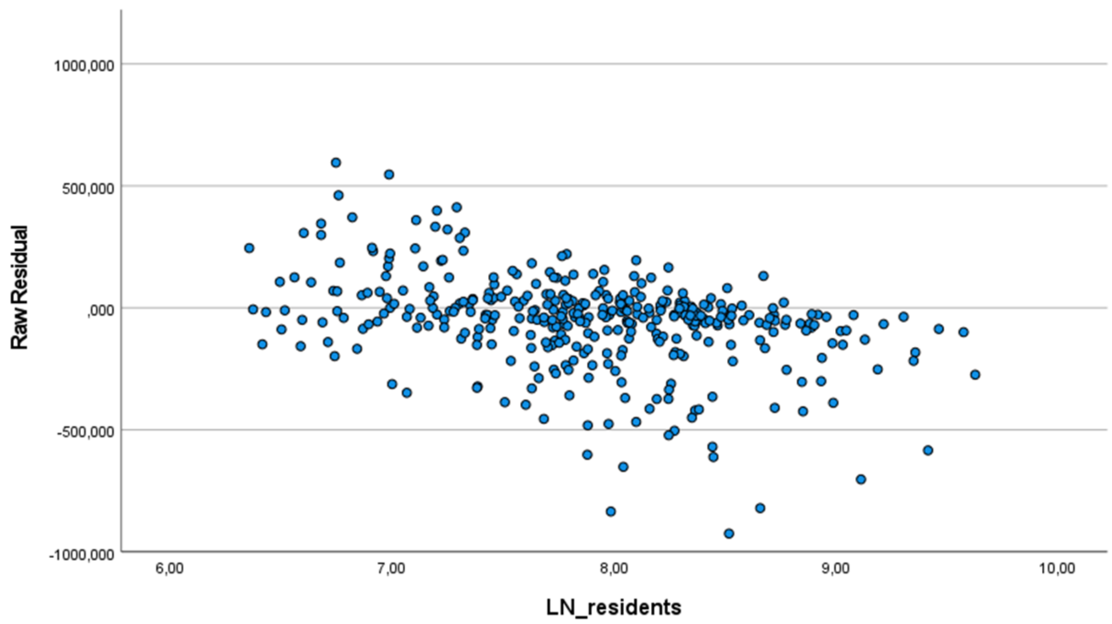

FIGURE K.1: Plot of the residuals against the natural log of the number of inhabitants in a neighbourhood

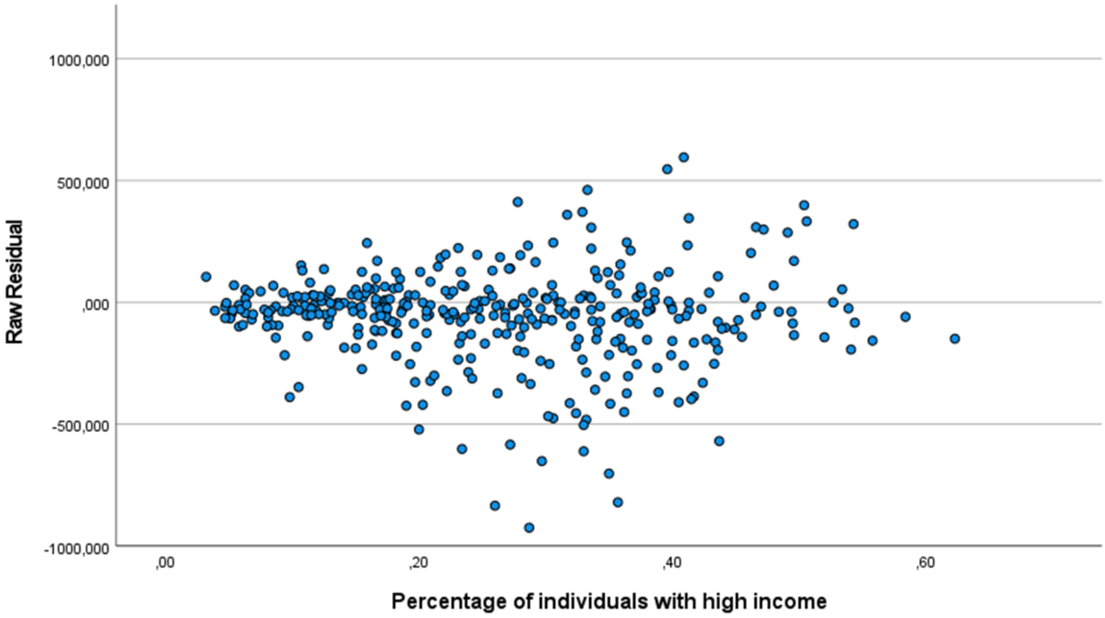

FIGURE K.2: Plot of the residuals against the percentages of individuals with a high income

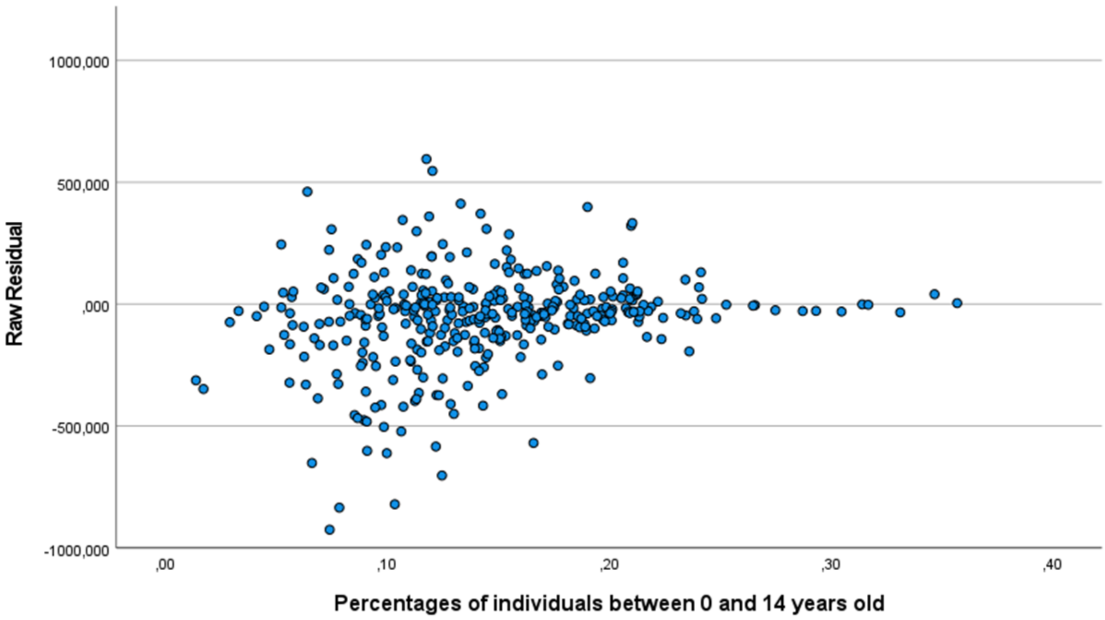

FIGURE K.3: Plot of the residuals against the percentages of individuals aged between 0 and 14 years old

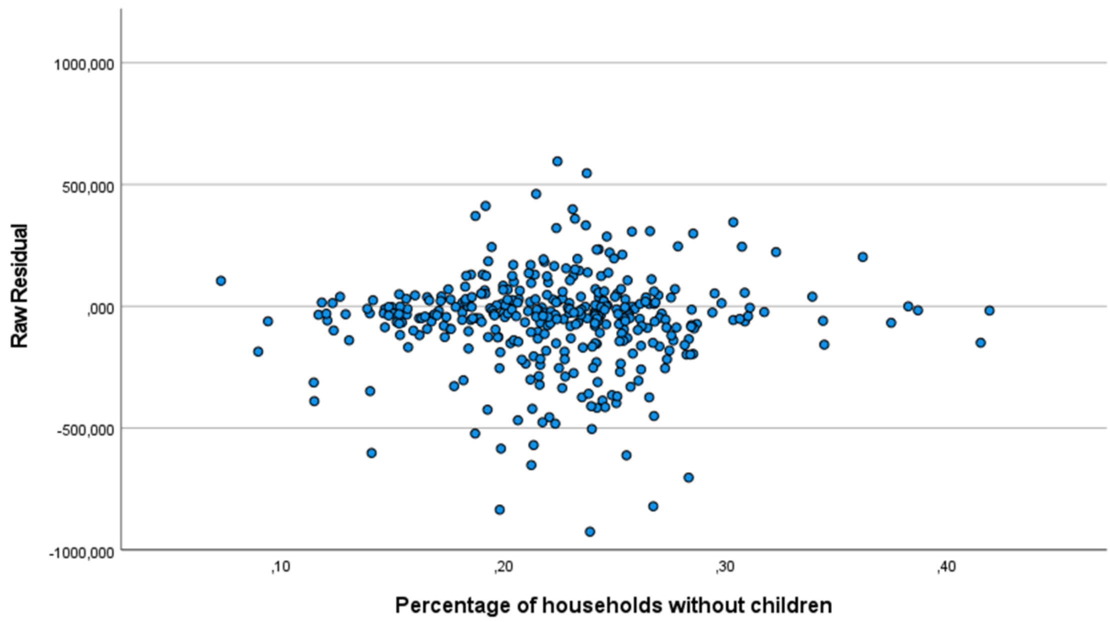

FIGURE K.4: Plot of the residuals against the percentages of households without children

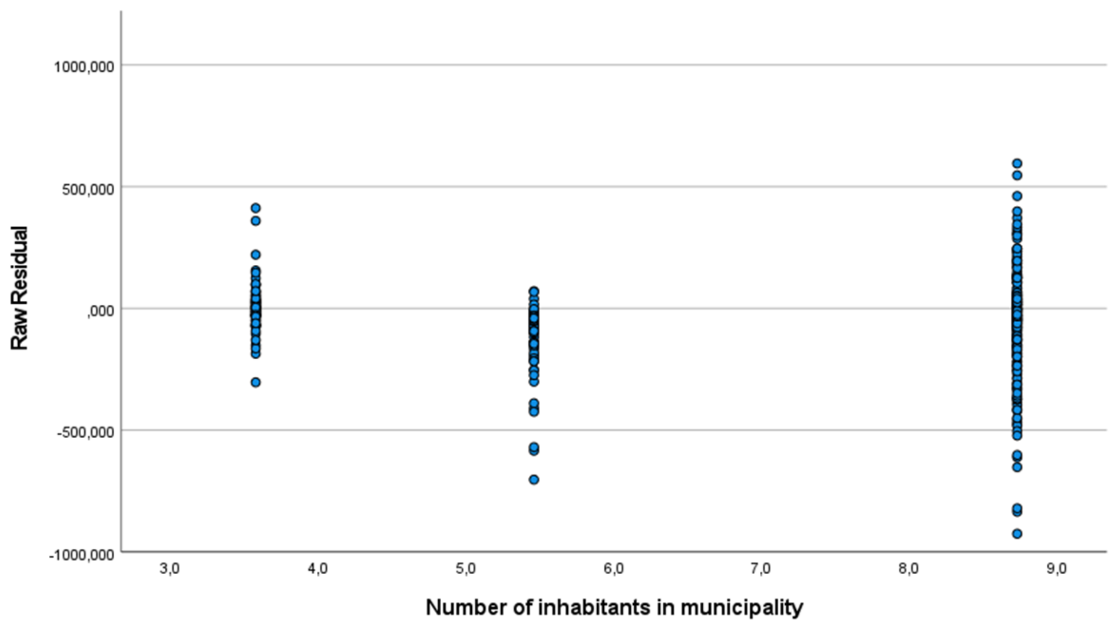

FIGURE K.5: Plot of the residuals against the number of inhabitants in the municipality

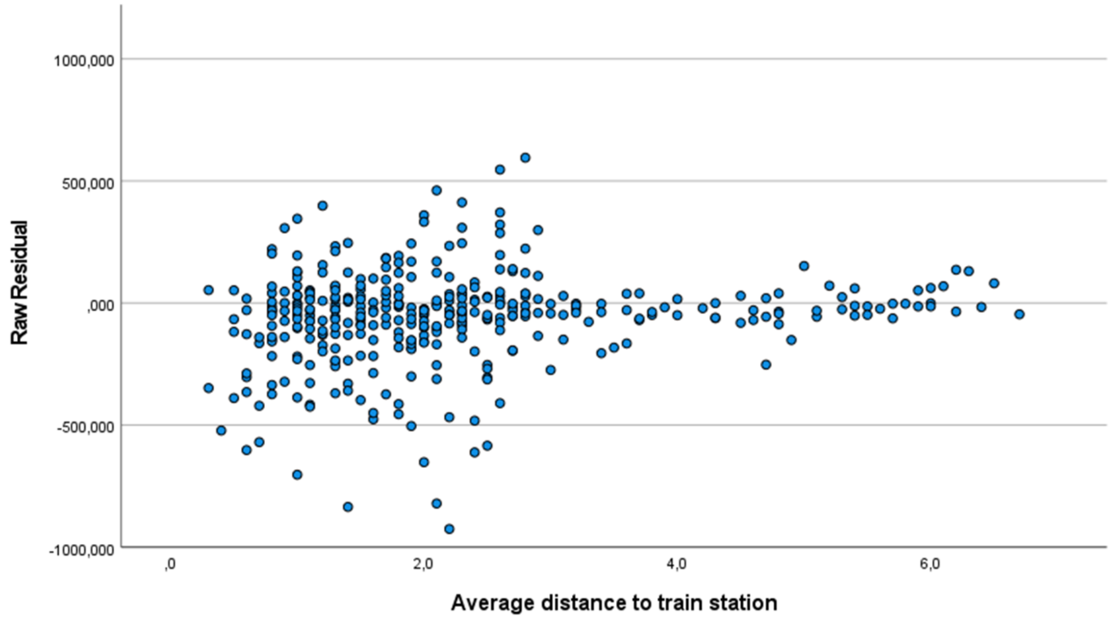

FIGURE K.6: Plot of the residuals against the average distance to a train station

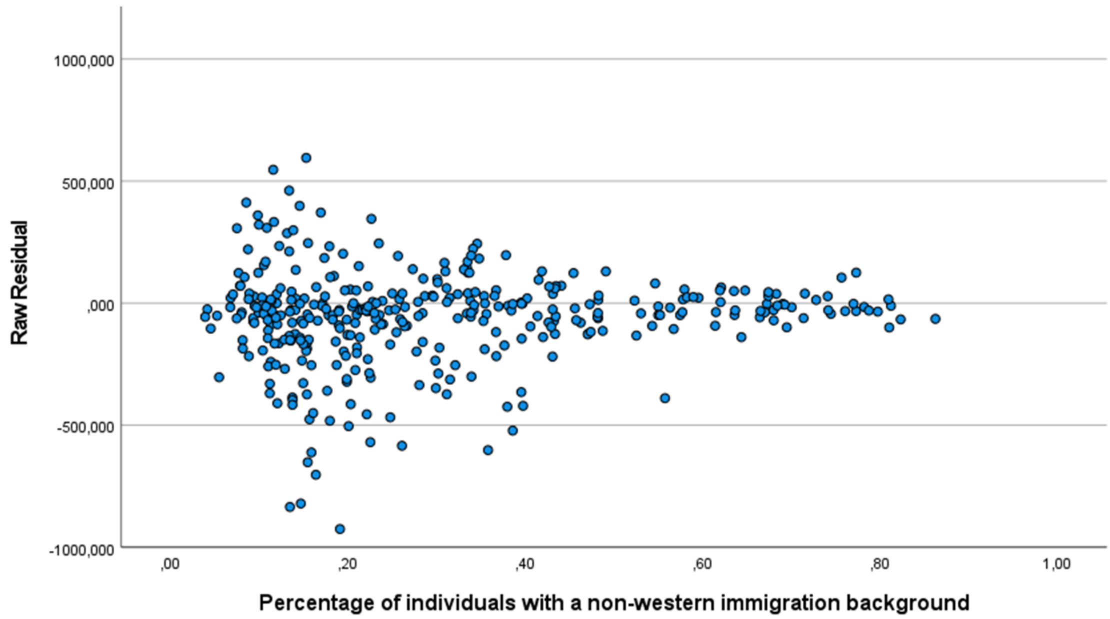

FIGURE K.7: Plot of the residuals against the percentages of individuals with a non-western immigration background

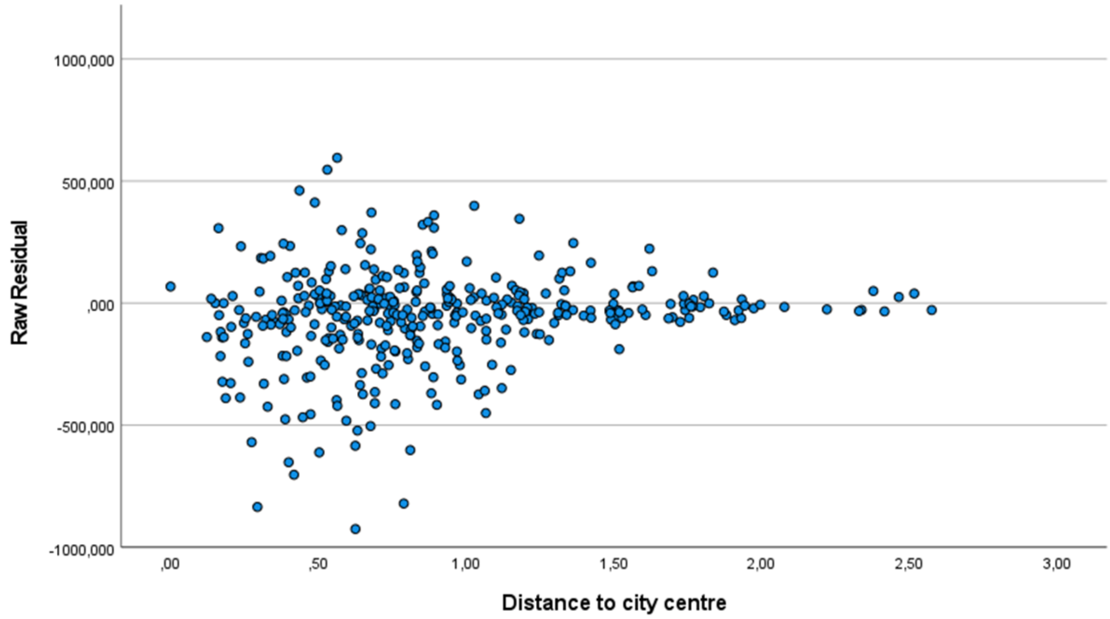

FIGURE K.8: Plot of the residuals against the distance to the city centre

Appendix L

Descriptive information on the dependent variables and the significance of the correlation of the neighbourhood factors with the demand and supply of roundtrip B2C shared cars used in method 2

L.1 Descriptive information of the dependent variables in method 2

TABLE L.1: Descriptive information of the dependent variables in method 2

Dependent variable	Average value	Range of values	N
Demand score on the 5-point Likert scale	2.25	1-5	620 respondents
Shared car presence	0.75	0 - 1	704 neighbourhoods
Shared car supply rate	392.28	0-23,333	704 neighbourhoods

L.2 Significance of the correlation of the neighbourhood factors with the demand and supply of roundtrip B2C shared cars

TABLE L.2: Correlation of the neighbourhood factors with the demand and supply of shared cars for method 2

Variable	Correlation with demand	Significance	Correlation with shared car presence in The Hague, Amsterdam and Utrecht	Significance	Correlation with shared car supply rate in The Hague, Amsterdam and Utrecht	Significance
Highly educated	0.224	0.000	0.205	0.000	0.001	0.984
Number of private cars in household	-0.133	0.001	-0.256	0.000	0.370	0.000
65 years old or older	-0.119	0.003	-0.097	0.010	-0.033	0.377
Low educated	-0.114	0.004	-0.224	0.000	-0.138	0.000
First resident parking rate	0.109	0.010	0.323	0.000	0.102	0.007
moderately educated	-0.102	0.011	-0.068	0.072	0.168	0.000
Average distance to train station	-0.079	0.050	-0.276	0.000	-0.065	0.083
Multiple person household with children living at home	0.074	0.065	-0.124	0.001	-0.065	0.087
Multiple person household without children living at home	-0.070	0.083	0.093	0.014	-0.132	0.000
25 to 45 years old	0.061	0.132	0.139	0.000	0.002	0.955
Number of inhabitants in municipality	0.059	0.142	0.022	0.559	0.124	0.001
Address density	0.058	0.147	0.427	0.000	0.005	0.896
Urbanization class 5	-0.056	0.076	-0.280	0.000	-0.046	0.224
Rental house	0.055	0.172	0.094	0.012	-0.252	0.000
Number of bars in 1km	0.053	0.190	0.200	0.000	0.017	0.661

Variable	Correlation with demand	Significance	Correlation with shared car presence in The Hague, Amsterdam and Utrecht	Significance	Correlation with shared car supply rate in The Hague, Amsterdam and Utrecht	Significance
Number of restaurants in 1km	0.045	0.265	0.262	0.000	0.025	0.511
Urbanization class 1	0.044	0.076	0.412	0.000	-0.071	0.059
Distance to city centre	-0.036	0.366	-0.345	0.000	-0.018	0.628
Non-western migration background	0.036	0.373	-0.075	0.046	-0.063	0.093
Urbanization class 3	-0.035	0.076	-0.130	0.001	0.023	0.538
Urbanization class 4	-0.018	0.076	-0.237	0.000	0.068	0.073
Urbanization class 2	0.012	0.076	-0.166	0.000	0.066	0.079
Male	0.008	0.836	-0.219	0.000	-0.117	0.002
Female	-0.008	0.836	0.235	0.000	0.118	0.002
Western migration background	-0.006	0.874	0.115	0.002	0.175	0.000
45 to 65 years old	0.004	0.926	-0.071	0.059	0.182	0.000
One-person household	-0.004	0.914	0.056	0.138	0.110	0.004
High income	0.000	0.142	0.131	0.000	0.281	0.000

Appendix M

Regression coefficients of the explanation of the variation in the demand with the independent variables used in the explanation of the variation in the supply of shared cars used in method 3

TABLE M.1: Regression coefficients of the explanation of the variation in the demand with the independent variables used in the explanation of the variation in the shared car presence

Independent variable $[X_p]$	**Regression coefficient** $[b_p]$	**p-value**
Constant	-2.340	0.028
Address density	-0.195	0.614
Female	0.160	0.542
High income	0.256	0.747
Urbanization class 1	0.918	0.558
Urbanization class 2	0.931	0.452
Urbanization class 3	0.719	0.534
Urbanization class 4	-0.599	0.686

TABLE M.2: Regression coefficients of the explanation of the variation in demand with the independent variables used in the explanation of the variation in the shared car supply rate

Independent variable $[X_p]$	**Regression coefficient** $[b_p]$	**p-value**
Constant	-2.662	0.007
Number of inhabitants in neighbourhood	4.977	0.229
High income	0.221	0.781
Age	-0.015	0.074
Living in a household without children	-0.001	0.997
Number of inhabitants in municipality	0.716	0.161
Average distance to train station	0.093	0.364
Having a non-western migration background	-0.558	0.376
Distance to city centre	-0.017	0.854

TABLE M.3: Regression coefficients of the explanation of the variation in the demand with the independent variables used in the explanation of the variation in the shared car presence and supply rate

Independent variable $[X_p]$	**Regression coefficient** $[b_p]$	**p-value**
Constant	-3.414	0.059
Address density	-0.188	0.649
Female	0.046	0.869
High income	0.265	0.740
Urbanization class 1	1.340	0.416
Urbanization class 2	1.161	0.370
Urbanization class 3	0.909	0.447
Urbanization class 4	-0.518	0.730
Number of inhabitants in neighbourhood	3.405	0.435
Age	-0.017	0.053
Living in a household without children	0.002	0.993
Number of inhabitants in municipality	0.793	0.152
Average distance to train station	0.120	0.303
Having a non-western migration background	-0.600	0.346
Distance to city centre	-0.008	0.943

Appendix N

Correlations of the demand for carsharing with individual and neighbourhood factors

TABLE N.1: Correlations of the demand for carsharing with individual and neighbourhood factors

Independent variable	**Correlation with the demand**	***p*-value**
Experience with shared car	0.299	0.000
Experience with other shared mobility	0.173	0.000
Use of own car/own van in past year	-0.173	0.000
Use of public transport in past year	0.159	0.000
Attitude towards good cycle infrastructure and flow for bikes	0.143	0.000
Experience with shared car in past year	0.141	0.000
Number of private cars in household	-0.125	0.002
Attitude towards good public transport infrastructure and connection	0.124	0.002
Living in a owner-occupied home	-0.119	0.003
Attitude towards good road network and flow for cars	0.117	0.004
Living in a rental house	0.102	0.011
Use of walking as transport mode in past year	0.101	0.012
Use of motorcycle/moped/scooter in past year	0.097	0.016
Highly educated	0.083	0.038
18 to 25 years old	0.075	0.063
House type other than rental/owner-occupied	0.073	0.069
Use of bicyle as transport mode in past year	0.072	0.073
First resident parking rate in neighbourhood	0.072	0.090
Age	-0.066	0.103
Studying or doing an internship	0.065	0.108
Not working	-0.064	0.113
Number of hours of paid work	0.054	0.181
Number of inhabitants in municipality	0.052	0.192
Monthly net income	0.048	0.285
Moderately educated	-0.047	0.238
Retired	-0.047	0.239
Low educated	-0.047	0.244
Experience with other shared mobility in past year	0.047	0.245
Different household composition than one or mulitple person with or without children	0.046	0.249
65 years old or older	-0.046	0.257

Independent variable	**Correlation with demand**	***p*-value**
Working 30 hours per week or more	0.044	0.271
One-person household without children living at home	-0.036	0.370
Urbanization class 2	0.036	0.375
Number of inhabitants in neighbourhood	0.035	0.385
Married, with children living at home	0.034	0.400
Working 0 to 12 hours per week	0.033	0.416
Urbanization class 4	-0.031	0.444
Political left oriented	0.028	0.479
(Partially) incapacitated for work or unemployed	0.026	0.523
Average distance to train station	-0.025	0.538
Number of bars in 1km	0.025	0.540
Urbanization class 5	-0.023	0.562
Low income	-0.022	0.585
Distance to city centre	0.019	0.634
Number of restaurants in 1km	0.019	0.641
Working (full time/part time)	0.018	0.658
No migration background	0.017	0.671
One-person household with children living at home	-0.014	0.721
Western migration background	-0.014	0.733
Male	-0.013	0.745
Female	0.013	0.745
Urbanization class 3	-0.013	0.754
Multiple-person household, without children living at home	-0.012	0.758
Address density	0.010	0.809
Non-western migration background	-0.009	0.817
Urbanization class 1	-0.008	0.846
Doing voluntary work	-0.007	0.861
25 to 45 years old	0.005	0.910
Working 12 to 30 hours per week	0.004	0.929
I have a different working situation	-0.002	0.962
Political right oriented	-0.002	0.964
Use of a taxi in past year	-0.002	0.970
Different education level than low, moderately or high	0.001	0.975
High income	0.001	0.975
45 to 65 years old	0.001	0.976

Appendix O

Explanation of the variation in the demand for roundtrip B2C carsharing in the G44 cities

O.1 Processes of finding the best fit to explain the variation in the supply based on the survey municipalities

TABLE O.1: Process of finding the most suitable binary logistic regression model to explain the variation in the demand for roundtrip B2C carsharing

Model	**Nagelkerke R^2**	**Experience with shared car**	**Experience with other shared mobility**	**Use of own car/own van in past year**	**Use of public transport in past year**	**Attitude towards cycling**	**Experience with shared car in past year**	**Number of private cars in household**	**Attitude towards public transport**	**Living in a owner-occupied home**	**Attitude towards cars**	**Living in a rental house**	**Walking in past year**	**Motorcycle use in past year**	**Highly educated**	**18 to 25 years old**
1[1]	0.134	0.000														
2	0.148	0.000	0.113[2]													
3[1]	0.152	0.000		0.026												
4	0.158	0.000		0.035	0.180[2]											
5[1]	0.192	0.000		0.048		0.002										
6	0.194	0.000		0.039		0.002	0.439[2]									
7	0.203	0.000		0.200[2]		0.001		0.077[2]								
8[1]	**0.197**	**0.000**				**0.001**		**0.020**								
9	0.192	0.000		0.047		0.015			0.881[2]							
10	0.197	0.000				0.010		0.020	0.794[2]							
11	0.200	0.000				0.001		0.050[2]		0.337[2]						
12	0.198	0.000		0.072[2]		0.001				0.17[2]2						
13	0.193	0.000		0.041		0.027					0.500[2]					
14	0.199	0.000				0.017		0.018			0.507[2]					
15	0.198	0.000				0.001		0.040				0.571[2]				
16	0.195	0.000		0.067[2]		0.001						0.327[2]				

Model	Nagelkerke R^2	Experience with shared car	Experience with other shared mobility	Use of own car/own van in past year	Use of public transport in past year	Attitude towards cycling	Experience with shared car in past year	Number of private cars in household	Attitude towards public transport	Living in a owner-occupied home	Attitude towards cars	Living in a rental house	Walking in past year	Motorcycle use in past year	Highly educated	18 to 25 years old
17	0.194	0.000		0.054[2]		0.003							0.391[2]			
18	0.200	0.000				0.001		0.023					0.384[2]			
19	0.204	0.000				0.001		0.016						0.125[2]		
20	0.197	0.000		0.048		0.001		0.000						0.171[2]		
21	0.192	0.000		0.048		0.002									0.965[2]	
22	0.194	0.000				0.001		0.019							0.714[2]	
23	0.200	0.000				0.001		0.021								0.318[2]
24	0.194	0.000		0.056[2]		0.002										0.384[2]

[1] *Binary logistic regression model with independent variables that are all significant*
[2] *Independent variable that is not significant*

O.2 Classification table of the model applied on the validation set

TABLE O.2: Classification table of the model applied on the validation set

	No predicted demand	**Predicted demand**	***Percentage correct***
No observed demand	100	8	*92.6%*
Observed demand	10	4	*28.6%*
			85.2%

O.3 Residual analysis for the explanation of variation in the demand for roundrip B2C shared cars

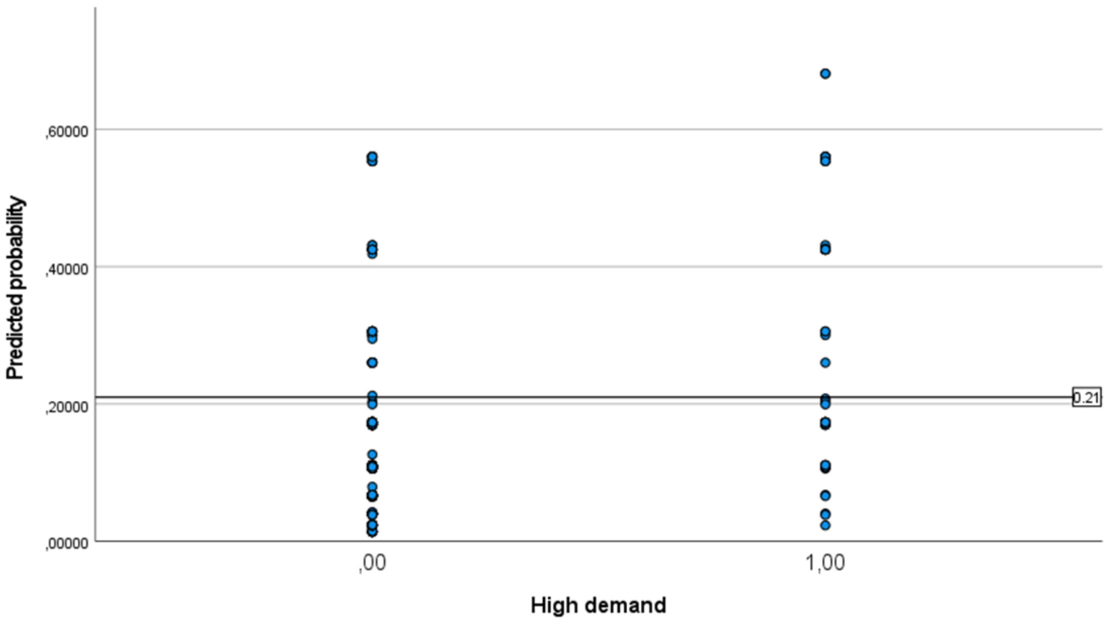

FIGURE O.1: Plot of the predicted probability against the observed demand in the G44 cities