Norwegian University of Life Sciences
School of Economics and Business

Philosophiae Doctor (PhD)
Thesis 2021:19

Freight trip generation and consumer preferences for reducing externalities from last mile deliveries

Godsturgenerering og konsumenters preferanser for å redusere eksternaliteter fra sisteledds-distribusjon

Elise Caspersen

Freight trip generation and consumer preferences for reducing externalities from last mile deliveries

Godsturgenerering og konsumenters preferanser for å redusere eksternaliteter fra sisteledds-distribusjon

Philosophiae Doctor (PhD) Thesis

Elise Caspersen

Norwegian University of Life Sciences
School of Economics and Business

Ås (2021)

Thesis number 2021:19
ISSN 1894-6402
ISBN 978-82-575-1791-5

Acknowledgements

I am grateful for the opportunity to pursue this Ph.D., given to me by the Institute of Transport Economics and Jardar Andersen, the manager of the project financing this research. Thank you Jardar for all the discussions, motivating words and massive support during the last five years, as well as for taking on the role as co-supervisor. I am also grateful to the School of Economics and Business at the Norwegian University of Life Science for taking me in as a Ph.D. student, and for providing my main supervisor Ståle Navrud and co-supervisor Jens Bengtsson. Thank you Ståle and Jens for interesting discussions, good advice and feedback on my work.

I would also like to thank Rensselaer Polytechnic Institute (RPI) represented by Professor Jose Holguin-Veras and his research team for including me as an exchange student and inspiring my research for 6 months in 2018. The stay contributed greatly to my motivation and strengthened this thesis. It resulted in a paper co-authored with Mario Arrieta-Prieto and Xiaokun (Cara) Wang, of whom I learned a lot.

Furthermore, I want to thank my colleagues at the Institute of Transport Economics, particularly my research group Logistics and Innovation for giving me the freedom to pursue this degree, the Ph.D. lunch group for being a space for ventilation, and Stefan Flügel for answering all my stupid questions.

Finally, I would like to thank my family and friends for their love, support and motivating talks. To Jørn, toddler Viktor and the baby to come, I am grateful for all the love and joy you provide me every day, keeping me sane and motivated to finish this task.

Writing this thesis has been quite a journey where I have visited places in my mind I did not know existed. It would be a lie to say I have enjoyed it all, but hopefully the benefits of persevering this task will come.

Haslum, December 2020

Elise Caspersen

Contents

List of expressions

Some expressions are frequently repeated throughout this thesis and are important for understanding its content. These are summarized in Table 1:

Table 1: Explanations of expressions commonly used in this thesis. Abbreviations in parenthesis.

Freight trip generation (FTG)	Number of trips generated (attracted and produced) by an economic agent (building, establishment, consumer) during a certain time period
Freight trip attraction (FTA)	Number of trips attracted by an economic agent during a certain time period
Freight trip production (FTP)	Number of trips produced by an economic agent during a certain time period
Establishments	Public or private firms (i.e. warehouses, retail stores, restaurants, hotels, schools, offices etc.)
Last mile delivery services	Goods delivery services provided to consumers (i.e. at home, pick-up point, etc.), often from an online platform.
Carriers	Companies undertaking freight transport activities for a fee
Business-to-business (B2B)	Delivery/transport between two firms (i.e. from warehouse to store)
Business-to-consumer (B2C)	Delivery/transport between firm and consumer (i.e. from warehouse to the end consumer)
E-commerce	The buying and selling of goods and services online
Crowdshipping	Using the crowd (a large group of people not necessarily employed in freight logistics) to perform (parts of) the last mile delivery.
Willingness to pay (WTP)	Maximum price a consumer is willing to pay for a service or a good (here measured in Norwegian kroner; NOK)
Stated preference (SP) methods	Methods asking consumer to state their WTP for changes in (public) goods or services in questionnaire surveys (as opposed to revealed preference methods that elicit consumers' preferences from their observed behavior in markets related to the (public) good in question).
Discrete choice experiment (DCE)	An indirect SP method, where consumers state their preferences through choices between alternatives with different levels of a set of attributes/characteristics of a good or service.

Summary

Freight transport is a key contributor to negative externalities in terms of congestion, noise, emissions of local air pollutants and greenhouse gases, road damage and accident risk. The issue is particularly significant in cities, with lots of activities generating freight transport and a high density of people inflicted by their externalities. At the same time, freight transport is indispensable for a city's economy and urban lifestyle, and achieving sustainable solutions is important. However, existing measures targeting urban freight transport are often aimed at carriers, although research has shown that carriers' ability to influence freight traffic is limited by shippers and receivers, i.e. the generators of freight. This thesis **provides knowledge to municipalities and urban planners working on measures to reduce externalities from urban road freight transport** through research on freight trip generation (FTG) and consumer preferences for last mile deliveries.

The work consists of five chapters: one introductory chapter and four chapters written as scientific papers that are either published or in the process of getting published in peer-reviewed journals. The introductory chapter provides an overview of the overall research gap targeted in this thesis as well as each of the four papers' individual and joint contribution to answering the overall research questions. The main findings within and between each paper are highlighted.

The first paper is presented in chapter 2 and presents FTG models for the Norwegian context. It contributes research on functional form and explanatory variables, fills some of the identified knowledge gap for FTG numbers in Norway, and identifies key industries contributing FTG from establishments.

The second paper (chapter 3), also deals with FTG, but considers shipments to consumers rather than establishments. The background for researching FTG to consumers is the growth in e-commerce, which inflicts changes to urban freight transport. The paper reveals that consumers generate a significant amount of freight trips and should be considered in future urban planning and policy making.

The third and fourth papers (chapter 4 and 5) shift focus from FTG numbers to measures targeting FTG and corresponding externalities. Both papers investigate if and how consumers could be involved in the transformation to environmentally sustainable freight deliveries through changes in last mile delivery services. The main findings from these papers are that consumers value environmentally sustainable deliveries and are willing to take on some of the costs, either by paying for reduced CO_2-emissions or waiting some extra days for the delivery if it implies reduced PM and/or CO_2-emissions.

The main overall contribution of this thesis is that consumers are a key agent in understanding urban freight traffic and should be included in the transition towards environmentally sustainable last mile deliveries, along with establishments.

Sammendrag

Godstransport er en viktig bidragsyter til negative eksternaliteter som kø, støy, utslipp, veislitasje og ulykkesrisiko. Problemet er stort i byer hvor det ofte er mange aktiviteter som etterspør godstransport og høy tetthet av mennesker som blir skadelidende under eksternalitetene. Samtidig er godstransport uunnværlig for byens økonomi og urbane livsstilsaktiviteter, slik at en bærekraftig godstransport er ønskelig. Tiltak for å redusere eksternaliteter fra godstransporten er ofte rettet mot transportører. Dette til tross for at forskning har vist at transportørers evne til å påvirke godstransporten er begrenset av avsendere og mottakere, dvs. aktørene som etterspør godstransport. **Denne avhandlingen gir kunnskap til kommuner og byplanleggere som arbeider med tiltak for å redusere eksternaliteter fra godstransport i by**, og bidrar med forskning på godsturgenerering og konsumenters preferanser for miljøvennlig sisteledds-distribusjon.

Avhandlingen består av fem kapitler: ett innledende kapittel og fire kapitler skrevet som vitenskapelige artikler som enten er publisert eller i en prosess for publisering i fagfellevurderte tidsskrifter. Innledningskapitlet gir en oversikt over kunnskapshullet denne avhandlingen tilstreber å fylle, samt hvert av de fire artiklenes individuelle og felles bidrag til å svare på forskningsspørsmålene. Hovedfunn i og mellom hver artikkel fremheves.

Den første artikkelen (kapittel 2) presenterer turgenereringsmodeller for godstransport til bedrifter i Groruddalen i Oslo og eksperimenterer med modellens form og forklaringsvariabler. Resultatet er turgeneringstall i en norsk sammenheng, som viser hvilke næringer som er viktige for å forstå omfanget av turgenerering i området.

Den andre artikkelen (kapittel 3) omhandler også turgenerering, men fra forbrukere som netthandler. Bakgrunnen er veksten i e-handel som har bidratt til å transformere godstransporten via økt sisteledds-distribusjon. Artikkelen viser at turgenerering fra forbrukere er betydelig, og bør vurderes i fremtidig byplanlegging og politikkutforming sammen med turgenerering fra bedrifter.

Den tredje og fjerde artikkelen (kapittel 4 og 5) flytter oppmerksomheten fra turgenereringsmodeller til mulige tiltak for å redusere antall turer og tilhørende eksternaliteter. Målet har vært å undersøke om og hvordan forbrukere kan bidra til en mer bærekraftig godsnæring ved hjelp av sisteledds-distribusjonsløsningene de velger. Hovedfunnene fra disse to artiklene er at forbrukere verdsetter miljøvennlig levering og er villige til å ta noe av kostnadene, enten ved å betale for reduserte CO_2-utslipp eller vente noen ekstra dager på levering for reduserte PM- og/eller CO_2-utslipp.

Hovedbidraget til avhandlingen er at forbrukere er en sentral aktør for å forstå godstransport i by, som kan og bør inkluderes (sammen med bedrifter og transportaktører) for å oppnå miljøvennlige leveringsløsninger.

Chapter 1:

Introduction

1.1 Background

"Economics is a study of mankind in the ordinary business of life" (Pigou, 1920). This thesis presents research on how this ordinary business of life generates urban road freight transport and how consumers can contribute to reduce its environmental impact. The main motivation has been to provide knowledge for municipalities and urban planners seeking to reduce the externalities from urban road freight transport. The latter is of interest to municipalities and urban planners as externalities may harm the population. One example is emissions of local air pollutants, which are of high concern in both Norway and Europe due to increased health care costs, lower labor productivity and premature deaths (European Environment Agency, 2019).

1.1.1 Urban road freight transport externalities

Externalities can be defined as the costs or benefits that the actions of one agent inflicts on other agents without compensation. The idea was introduced by Pigou (1920), who pointed out the difference between economic welfare (from production or monetary evaluation of a good) and social welfare (including utility or satisfaction from a good). Externalities provide a gap between the economic and social welfare resulting in a sub-optimal amount of the action: too little if the action is beneficial to others and too much if it is costly to others. An illustration of a market with negative externalities is presented in Figure 1. The x-axis represents the quantity of a good produced and the y-axis represents its price. If we assume that negative externalities are ignored and only marginal private costs are presented to the consumer, the market equilibrium provides a quantity that exceeds the social optimum. The difference is caused by externalities.

In the context of urban road freight transportation, Figure 1 can represent a typical market for a private good that needs a road freight transport service. Thus, one can assume that the marginal private cost (MPC) and marginal external cost (MEC) include both costs of production and road freight transport. A part of the marginal social cost (MSC) of producing the good will then be the marginal social cost of road freight transportation, including both marginal private costs and externalities. Many freight operations, in particular e-commerce with free or low-priced transportation costs presented to the consumer, conceal the full MSC, including the externalities, leading to a greater quantity of road freight transport services than what is socially optimal.

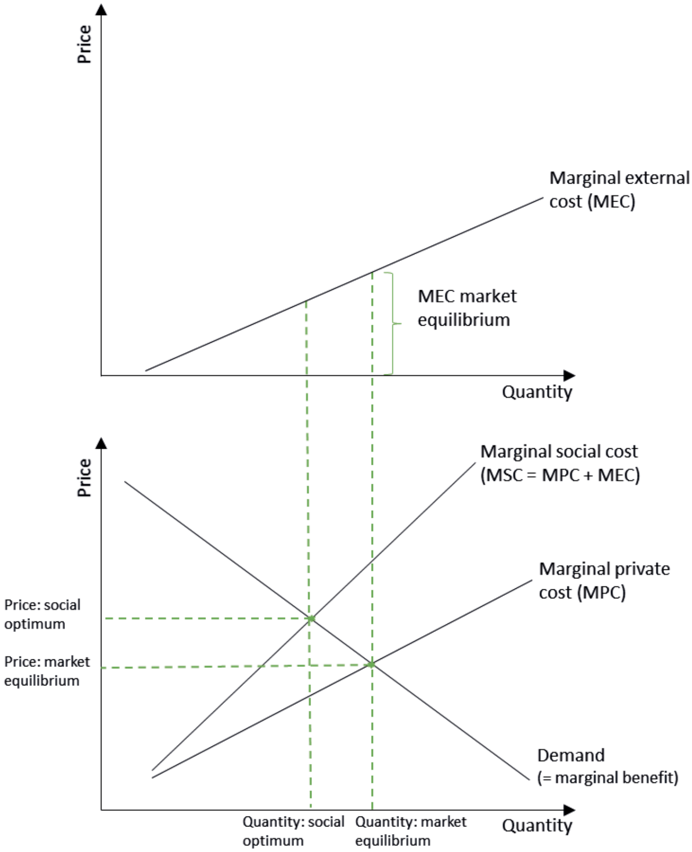

Figure 1: A market with external costs. When externalities are ignored, the market equilibrium provides too high a quantity produced as compared to the social optimum.

Road transport causes negative externalities in terms of congestion, noise, emissions of local air pollutants and greenhouse gases, road damage and accidents risk being imposed on other road users or people located in its proximity (Calthrop and Proost, 1998, Santos et al., 2010), with road freight transport having higher externalities than other types of road transport (Rødseth et al., 2019, Santos, 2017). Several measures can be, and are, taken to internalize them, such as command and control measures (standards, laws and regulation), incentive based measures (taxes, subsidies, charges) or tradeable emission permits (European Commission, 2019). However, not all measures are equally successful. Santos et al. (2010) argued that command and control measures, including fuel and vehicle standards and access restrictions, are inefficient to tackle externalities, although widely used and accepted in public. Fiscal policy instruments (taxes or subsidies) are found to be more efficient (Santos et al., 2010), although De Borger et al. (1997) found that fuel tax or road tolls alone were insufficient to reach the welfare gains from optimal pricing in Belgium, and Santos (2017) found that fuel taxes are not high enough to internalize externalities from road transport in 22 European countries, with the gap being particularly large for heavy duty vehicles.

The efficiency of measures targeting externalities relies on a strong link between the tax and the source of externalities (Santos et al., 2010), which is not always the case with road freight traffic. To reduce externalities from road freight traffic, the right agent must be targeted: carriers (companies undertaking freight transport activities) for changes in vehicle technology or increased productivity, but suppliers and receivers for freight traffic altering measures, like changes in the time of delivery or freight trip reduction (Holguín-Veras et al., 2015). The reasons is that carrier behaviour depends on decisions made by shippers and receivers, leaving them unable to (solely) enforce the necessary changes to influence freight traffic (Holguín-Veras et al., 2016, Holguín-Veras et al., 2015). Still, most suggested measures aimed at reducing road freight externalities affect the carriers. Examples are urban tolls or low-emission zones (LEZs) being suggested as a possible measure to reduce emissions from freight traffic in the city centre of Paris (Coulombel et al., 2018), or the existing road tolls in Oslo, Norway, distinguishing between peak and off-peak travel, vehicle size, and vehicle technology (fossil fuel or zero-emission vehicles) when charging each crossing[1].

Another way of reducing the externalities from urban road freight transport is to target the demand side to inflict changes to how, where and when the transport is made. As senders and receivers are a heterogeneous group of establishments, and in the later years, consumers, this requires extensive knowledge of the demand side. One way to achieve this is through research on freight trip generation (FTG). FTG is a measure of freight trips produced or attracted by economic activities and helps explain the freight traffic observed in different areas. As both the total and marginal costs of road transport externalities depend on traffic volume, as shown for instance by Calthrop and Proost (1998), targeting FTG is expected to contribute to a general reduction in road transport externalities. FTG has already been used to model freight traffic and externalities in a given area, as done by Coulombel et al. (2018) in their model for pollutant emissions in Paris, and for understanding each industry sectors' contribution to congestion (Holguín-Veras et al., 2016). How the demand side would respond to policies is also an important part of identifying the correct measures to reduce freight traffic and externalities (Holguín-Veras et al., 2016). Thus, this thesis contributes the magnitude of freight trips in Norway and how these, and the corresponding negative externalities, can be reduced by targeting consumers. As illustrated in Figure 2, the first two papers are related to FTG and investigate freight trip attraction (FTA) by establishments and consumers, while the second two papers analyse consumer preferences for different aspects of environmentally sustainable last mile deliveries.

[1] For more information about the Oslo road tolls visit https://www.fjellinjen.no/private/

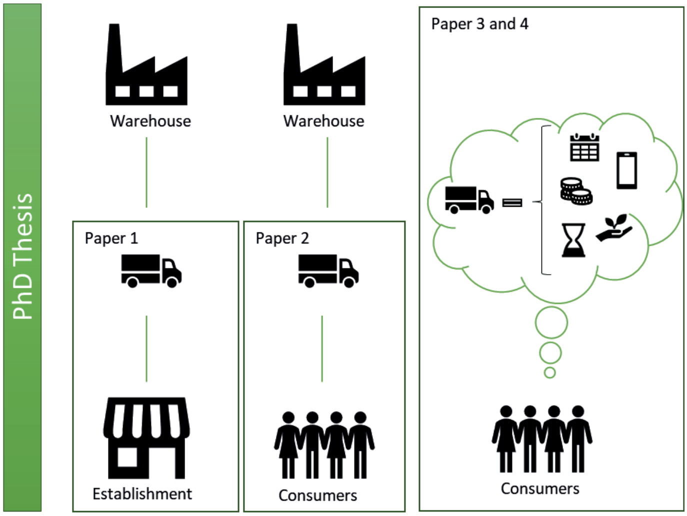

Figure 2: Illustration of supply and demand side of urban freight transport and the positioning of the four papers. Icons are borrowed from Microsoft Office.

The rest of the introduction chapter (Chapter 1) highlights the research gaps in the literature, presents the data and methodology used, and summarizes the four papers before finishing with key findings and conclusion. Finally, the four papers constituting this thesis are presented in Chapter 2 to 5.

1.1.2 Overall research question

The research in this thesis can be summarized by the following overall research questions:

E-commerce alters purchase behaviour and increases freight trip generation to consumers: i) What is the magnitude of freight trip generation by consumers compared to freight trip generation by establishments? ii) Can externalities of urban freight traffic be reduced by targeting consumer last mile delivery alternatives?

Urban freight is a key contributor to externalities, but is, at the same time, indispensable for a city's economy and urban lifestyle activities. As a result, conflicts arise as to how to secure freight transport admission to cities while reducing its negative burden on city users. To solve this issue Sustainable Urban Logistics Plans (SULPs), bringing together local actors and improving planning and actions for sustainable urban logistics, are suggested. One such project, *Sustainable urban logistics plans in Norway (NORSULP),* received funding from the Research Council of Norway in the 4-year period 01.01.2016-31.12.2019. The aim was to develop user-validated guidelines for SULPs and to contribute to capacity building and knowledge generation within the field of urban freight transport in Norwegian cities. The project financed two Ph.D. positions, one of which is documented in this thesis.

The original objective of this Ph.D. was to shed light on establishments as generators of urban freight transport and their impact on freight traffic in different cities using quantitative measures. These measures should contribute a knowledge base to reduce the negative externalities of urban freight transport in the NORSULP member cities. However, it was soon discovered that the cities did not have much data about freight transport or freight traffic (as documented in Fossheim et al. (2019)), and that the data required, or the relevant substitutions, were not covered by public data collection (for instance by Statistics Norway). Thus, this thesis' research objective could not be met without excessive city-wise data collection efforts. At the time of these discoveries, e-commerce and business-to-consumer (B2C) deliveries had increased in size and altered freight traffic as we know it. Hence, the objective was changed to shed light on establishments and consumers as generators of urban freight transport and the potential measures to reduce its impact, as stated in the overall research questions above.

1.2 Freight trip generation and consumer change

1.2.1 Freight trips from establishments

Cities depend on freight activities to support the urban lifestyle, but as it is a source of several externalities its magnitude should be as low as possible. Thus understanding the relationship between urban economic activities and freight activities is key to adequately address the challenges faced by cities and policy makers on a daily basis (Holguin-Veras et al., 2018). As traffic counts or other relevant data may be scarce for cities, as found in the NORSULP project, FTG indicators may be useful proxies for freight traffic in cities. A study of economic activity and the generation of freight and service activity in metropolitan areas reveals high numbers of FTG, even for a relatively small city such as Oslo: the average number of freight trips generated in the larger Oslo area is 1.53 trips per establishment per work day or 70.4 trips per 1000 residents (Holguin-Veras et al., 2018)[2]. Through its connection to economic activities, FTG numbers provide more information than traffic counts and are useful for analysis on different aggregation levels. Campbell et al. (2018) use FTG models to estimate parking needs at two small commercial areas in New York State, Gonzalez-Feliu and Peris-Pla (2017) uses freight trip attraction (FTA) numbers to analyze the relation between freight trips and shopping trips for different urban space categories while Holguin-Veras et al. (2018) contribute overall understanding of freight and service activities in metropolitan areas.

Despite its usefulness, FTG numbers are often lacking. When starting on Paper 1 in this thesis, the recommended guidelines for FTG numbers for Norwegian urban planners dated back to 1989 (Reinsborg et al., 1989). Further, a review of local development projects and plans fostering dense areas and land use combining workplaces, shopping, leisure and living (DARK + ADEPT (2013), COWI AS et al. (2014)), regional strategies seeking to reduce urban sprawl (Akershus fylkeskommune, 2018) and local public development guidelines (Plan- og bygningsetaten, 2019a, Plan- og bygningsetaten, 2019b), revealed that freight transport is often neglected both in initial proposals and in concept studies. The lack of emphasis of freight traffic in Norwegian public planning is strengthened by a review of economic analysis within urban densification (Magnussen et al., 2017), confirming that freight traffic is mostly neglected. During the later years, data collections of FTG have been conducted, although through local initiatives that are not publicly available. Thus, Paper 1 targets a gap in the literature by contributing FTG numbers from the Norwegian context. The paper also contributes knowledge on functional form (as called for by Holguin-Veras et al. (2011)) and potential relevant explanatory variables of the FTA-models.

1.2.2 From establishments to consumers

The lack of freight transport in urban planning is troublesome due to its supporting role for economic activities, but also due to its relation to societal trends. Although

[2] The numbers are based on the same dataset as used in Paper 1 of this thesis.

freight transport is still dominated by business-to-business (B2B) shipments, new trends are emerging. One of the most influential changes is the increase in online shopping or e-commerce seen over the past decade: the European E-commerce turnover was €373 billion in 2015, and predicted to be €717 billion in 2020 (Ecommerce News, 2020). Suddenly consumers can access everything from home, using their computer, smart board or phone, requesting goods and services to be sent to them, rather than the opposite. This increase in B2C deliveries puts pressure on urban freight transport systems.

Compared to B2B, B2C provides smaller deliveries to more recipients, reducing the efficiency of freight transport (Visser et al., 2014, Allen et al., 2018, Dablanc et al., 2017), and has led to increased use of light goods vehicles for parcel deliveries and last mile movements in many urban areas, for instance in the UK (Allen et al., 2018). This emerging trend and change in freight transport renders FTG models focusing on establishments insufficient when providing an overview of the freight traffic in an area: when adding online deliveries to the FTG numbers for Oslo presented above (using an internet delivery ratio of 0.12 estimated by the authors for New York City using the US National Household Transportation Survey), the FTG numbers per 1000 residents in Oslo more than doubles, from 70.4 to 190.4 daily freight trips (Holguin-Veras et al., 2018). These numbers are somewhat overrated as the internet delivery ratio used is for New York City. Still, they illustrate that knowledge about freight transport to households is vital. Paper 2 of this thesis analyzes shipments to consumers from online shopping, contributing knowledge on FTG to consumers. By splitting shipments across five aggregated commodity groups, Paper 2 also contributes a detailed discussion of consolidation as a potential freight trip reducing measures.

1.2.3 The current and future role of consumers

Although increasing FTG by consumers, e-commerce might be beneficial for society if it is performed efficiently and reduces overall vehicle kilometers. Studies have shown that last mile delivery services using a carrier generate the lowest amount of CO_2 per customer/item due to its influence on total vehicle miles traveled (Wygonik and Goodchild, 2018, Carling et al., 2015). However, trends, like instant deliveries (within the same day or a couple of hours) or sharing economy (goods or services that are shared through online platforms, including renting, swapping, and trading (Hamari et al., 2016)), might distort these assumptions, as they generate new freight patterns with the potential to increase emissions (Dablanc et al., 2017, Lin et al., 2018, Carbone et al., 2018). At the center of such trends we find consumers.

By including consumers in the last mile, for instance through self-collection (Wang et al., 2019) or crowdsourcing (Buldeo Rai et al., 2021), delivery efficiency can be improved. Additionally, incorporating knowledge of consumers into service design, including last-mile delivery services, helps online retailers and transport suppliers to offer the right bundle of alternatives and encourages new innovative solutions (Vakulenko et al., 2019). Paper 3 and Paper 4 investigate consumer preferences for environmentally friendly last mile delivery solutions using a hypothetical situation of clothing rentals online.

In Paper 3, the tradeoff between delivery time (ranging from 1 to 20 days) and emission (CO_2 and particulate matter (PM)) is investigated. The paper contributes a discussion of the importance of delivery time as compared to airborne emissions and whether female consumers would accept an increased waiting time for an environmentally sustainable delivery. The objective has been to counterweight some of the perceptions that consumers demand ever decreasing delivery times. In Paper 3, the cost of the last mile delivery is assumed to be zero. Free delivery is expected by many consumers and offered by many online retailers (Bring Research, 2019). Monetary valuation of nature is also being questioned as nature cannot be bought in a market or have a market for close substitutes (Victor, 2020), and current behavior influences future climate damage (van den Bergh and Botzen, 2015). However, according to economic theory consumers make choices subject to both money and time budgets (McFadden and Train, 2000) and willingness to pay is the commonly used term for evaluating and comparing consumer preferences for a good. Thus, a similar model as in Paper 3 is presented in Paper 4, with the main difference being that free delivery is no longer assumed. The analysis in Paper 4 contributes findings on consumers' willingness to pay for CO_2-emission reduction, and how this amount differs between female consumers.

As identified research on consumer engagement in last mile delivery services is concerned with value generation (Rouquet et al., 2017, Wang et al., 2019, Carbone et al., 2018), the papers contribute a new angle. Paper 3 and 4 also contribute to the research field of environmental sustainability, which is scarce, but increasing in magnitude (Ellram and Murfield, 2017). Knowledge of how consumers think about their FTG activities is also relevant for urban planners working to achieve sustainability and establishments working for corporate social responsibility, and both are research gaps in sustainable transport literature (Zhao et al., 2020).

1.3 Data and methodology

The data used in this thesis stems from two surveys developed and completed as part of this thesis and a dataset shared by a large Norwegian logistics company. Papers 1 and 2 are based on different datasets and methodology, while papers 3 and 4 are based on the same survey dataset (but different subsamples) and same methodology. Although being a collector of national data and statistic for freight flows and freight transport with large trucks, small trucks and vans, data from Statistics Norway were found to be unsuitable for the purpose of this thesis. The data are suited to analyze differences between cities at a macro-level, like volumes of freight transport, each commodity groups' share of the transport, main modes for transport etc., but does not include information about freight generating activities from business and consumers, nor consumer preferences for last mile deliveries. Thus, data for the thesis was collected separately from Statistics Norway's data.

1.3.1 Paper 1

Paper 1 builds on data from an establishment survey inquiring about freight and service traffic in Groruddalen in Oslo, Norway. The survey was inspired by the *Prototype Freight and Service Trip Generation Survey* in Holguín-Veras et al. (2012, Appendix H). The objective was to provide data input for a FTG model (as part of this thesis) as well as an analysis to clarify the necessary future road transport system in Groruddalen (documented in Caspersen and Pinchasik (2017)). The survey was distributed as a self-reporting web-survey and ran from August 30th – September 23rd 2016. The net response rate was 20 %. This is as expected from establishment surveys (Allen and Browne, 2008). Although this is perceived as a suitable way to obtain knowledge of freight and service activity from establishments, Browne et al. (2010) point to one obvious weakness with establishment surveys: the responses rely on the respondents' knowledge and ability to provide correct information about the vehicles entering and exiting the firm over a given time period. The data in this survey were compared with other data for Groruddalen, and found to be fairly representative in terms of establishment size, location and industry classification as well as suitable for freight activity analysis of the area (Caspersen and Pinchasik, 2017). However, some industry groups had very few respondents, which is a weakness for industry specific analysis.

In this thesis, the data were used to estimate a FTG model for freight trips attracted to establishments in Groruddalen. The chosen methodology followed the research field on FTG modeling at that time: linear or non-linear regression models estimated using ordinary least squares with employment size as the main explanatory factor (Holguín-Veras et al., 2017). As summarized by Holguín-Veras et al. (2017), the recommended models for FTG estimation was linear $f_i = \alpha + \beta E_i$ or non-linear $f_i = \alpha E_i^{\beta}$, with f_i being the freight trip generation metric for establishment i, α a constant term and E_i employment at establishment i. It is common that the significance of the constant and the employment variable vary between industries. Heterogeneity between industry groups was explored as well as additional explanatory variables in accordance with the literature on the field (Alho and Silva, 2014, Holguín-Veras et al., 2013, Jaller et al., 2015, Sánchez-Díaz, 2016, Holguín-Veras et al., 2017). Also in later years linear models have been estimated using OLS with employment being the main measure of business size, non-linearity is introduced through logarithmic transformations of employment and the FTG variable, and different industries are accounted for (De Bakshi et al., 2020, Gonzalez-Feliu and Sánchez-Díaz, 2019, Gonzalez-Feliu et al., 2020).

1.3.2 Paper 2

The dataset used in Paper 2 stems from a Norwegian consumer survey about online shopping behavior and was shared with the researchers by one of the largest logistics companies in Norway. The survey was conducted by a third party, using national representative samples of the population between 18 and 79 years of age with internet access. It is one of a series of regularly conducted surveys, which as of 2016, have been conducted monthly. The dataset in question was collected in January 2017 and includes online shopping behavior for December 2016. It consists of 1515 respondents and found to be close to representative in terms of gender, age, and geography by the data collector. Information in the sample includes frequency and commodity type purchased online last month (December 2016), online shopping preferences, demographics, socioeconomic factors, and household characteristics.

The data was used to investigate the number of shipments received from online shopping in total and per aggregated commodity group by utilizing information on numbers of shipments received and indicator variables for commodity type purchased. Commodity type is of interest as it influences freight transport through probability of first time failure, time sensitivity and freight handling requirements (Allen et al., 2018), but collecting the number of shipments per commodity group from consumers would be a tedious and time-consuming task. As no papers splitting the number of shipments across subgroups were identified within urban freight research, the methodology used in this paper sought inspiration from accident research. More precisely, the research on latent marginal counts proposed by Afghari et al. (2016) and Afghari et al. (2018) where the total count $\lambda^{(i)}$ is found by summarizing counts from each subgroup, j, which are modeled in terms of group specific-covariates, X_{ij}, and parameters, β_j: $\lambda^{(i)} = \sum_{j=1}^{J} \lambda_j^{(i)} = \sum_{j=1}^{J} \exp\,(\beta_j^T X_{ij})$. The methodology allows for zero counts through an indicator variable $I_j^{(i)}$, taking on the value 1 if the consumer bought from the j-th commodity type or 0 otherwise. Count data models have been used to estimate the number of shipments per person from online shopping (Wang and Zhou, 2015, Saphores and Xu, 2020), but without the additional methodology allowing us to distinguish by commodity group. Other research also presents average shipments to individuals or households based on existing data sources shared by other researchers (Dablanc, 2019) or own data collection designed for this purpose (Gardrat et al., 2016).

1.3.3 Paper 3 and Paper 4

The second survey completed as part of this thesis provided data for both Paper 3 and Paper 4. The main motivation was to capture consumer preferences for environmentally sustainable last-mile delivery solutions for online shopping: who and when, if any, chose the most environmentally friendly home delivery solutions. Prior to data collection, some online retailers in the starting pit of providing their customers with environmentally sustainable last mile delivery alternatives were contacted and inquired about their data inventory (to collect data about the consumers' actual choices (revealed preferences)). Unfortunately, these companies' knowledge was limited to the chosen delivery alternative, at most; no information about the offered but rejected alternatives was available. Thus, little could be said about consumer preferences for environmentally sustainable last mile delivery solutions. When no revealed preference data is available, stated preferences (SP) enable the investigation of consumer preferences. SP methods put consumers in a hypothetical situation in which they are asked to state their valuation of the presented alternative(s). The hypothetical component allows for the construction of an unreal situation, which is both the strength and weakness of the methodology (Johnston et al., 2017). To capture SP either discrete choice experiments (DCE), capturing preferences for multi-attribute alternatives, or contingent valuation (CV)), capturing preferences for a proposed change, are suitable (Johnston et al., 2017). For the purpose outlined here, DCE was preferred as multiple last mile delivery attributes were of interest.

The design of the DCE was highly influenced by FJONG (a Norwegian platform for clothing rentals; www.fjong.com), through an existing project between The Norwegian University of Life Science (NMBU) and FJONG, shifting the scope of the survey towards clothing rentals in addition to online shopping. As free delivery is of importance to many consumers and offered by many online retailers (Bring Research, 2019), including FJONG, it was of interest to investigate both a hypothetical situation which included delivery at a cost for the consumer and a hypothetical situation with delivery free of charge. Thus, two DCEs were developed, with attributes and attribute levels as presented in Table 2. Each respondent was randomly assigned to one of the two versions with attribute A5.1 or A5.2.

The levels of CO_2-emissions were calculated based on the distances for last mile deliveries (Statistics Norway[3]) and emission factors (Handbook Emission Factors for Road Transport[4]) for light duty vehicles (weighing 3,5 tons or less) reported in Caspersen and Ørving (2018). The middle value is based on average numbers, while the low value assumes zero emission vehicles and/or zero net distance of delivery, and the high value assumes instant or express deliveries with low consolidation and/or large distance of delivery. As PM is an unfamiliar concept for many respondents, the qualitative terms (Low-Medium-High) were found adequate for this purpose. Price (measured in Norwegian kroner (NOK)) takes on values found for online clothing retailers. The DCE was combined with other questions about habits, preferences, and attitudes towards online shopping, clothing rentals, and the environment, as well as consumer socio-economic data, into a survey. The survey was distributed to an online response panel consisting of females between 18 and 70 years of age through the survey firm NORSTAT during summer 2020. In total,

[3] https://www.ssb.no/transpsg
[4] https://www.hbefa.net/e/index.html

4,602 panelists received the survey link, of which 1,200 responded (response rate = 26 %). Of these 1,200, 595 respondents received the DCE with CO_2 and PM (Paper 3), and 605 received the DCE with CO_2 and Price (Paper 4).

Table 2: Characteristics of the discrete choice experiment: attribute description and levels.

Attribute (A)	*Description and levels*
A1: Delivery time	Number of days the respondent accepts to wait for the parcel: *1-5-10-20 days*
A2: Delays (dummy)	Uncertainty with respect to delivery time: *"No", "Yes, 1-2 days"*
A3: Information (dummy)	Notifications by SMS or e-mail when 1) the good is approved for shipping and 2) the parcel is shipped to the consumer: *"No", "Yes"*
A4: CO_2 -emission	CO_2-emission resulting from transport of the parcel. Differ with transport mode, time, degree of consolidation etc.: *0, 0,28kg, 1,40kg*
A5.1: Particulate matters (PM)	PM resulting from transport of the parcel. Differs with transport mode, time, degree of consolidation etc.: *"Low", "Medium", "High"*
A5.2: Price (for the delivery):	Price: *0 NOK, 49 NOK, 99 NOKr*

The estimations in Paper 3 and Paper 4 were performed using the choice modeling framework originating from McFadden (1974) and Ben-Akiva and Lerman (1985), of which the utility (U) is given by a deterministic (V) and a stochastic (unobserved) component (ε). The choice maker is assumed to choose the alternative with highest utility, and thus the probability of observing alternative i over alternative j is $P_i = P\left(U_i \geq U_j\right) = P(V_i - V_j \geq \varepsilon_j - \varepsilon_i)$. The solution depends on the distribution of ε, which, as is most common, was assumed independent and identically distributed following a standard type I extreme (Gumbel) distribution. This provides the conditional or multinomial logit model (MNL). When analyzing environmental preferences, past experiences, beliefs and attitudes are found to be important (as argued in both Paper 3 and 4) but difficult to measure. Such unobserved heterogeneity is accounted for using the latent class model (LC) in Paper 3 and the mixed logit model (MMNL) in Paper 4. The LC model has been used on several occasions to reveal consumer preferences within transport research, for instance in combination with MNL to model socioeconomic and attitudinal influence on preferences for environmental policy drivers (Valeri et al., 2016) and consumer preferences for last mile deliveries of e-groceries (Gatta et al., 2020), or for dissecting preference heterogeneity in consumer airport choice (Marcucci and Gatta, 2012). In Paper 4, the modeling choice was inspired by existing research on consumer preferences for reduced CO_2-emissions by Achtnicht (2012) and Alberini et al. (2018), both combining the MNL and the MMNL models to analyze preferences. The disadvantages that that researchers must decide the distribution of the parameters (Shen, 2009, Daziano and Achtnicht, 2013) and the undefined moments of willingness to pay in the MMNL-model (Sillano and Ortúzar, 2005, Daly et al., 2012, Carson and Czajkowski, 2019), led to the choice of keeping price fixed (the other attributes vary according to a normal or log-normal distribution). Although being in a reasonable range compared to market prices for last mile delivery services, the estimated WTP to reduce 1 kg CO_2 from Paper 4 provides a WTP per tCO_2 that exceeds those found by Achtnicht (2012) and Alberini et al. (2018). However, the preferences from Paper 4 might not be constant for all levels of CO_2, and scalable up to 1 ton.

1.4 Paper summaries

1.4.1 Paper 1: An Explorative Approach to Freight Trip Attraction in an Industrial Urban Area.

Caspersen, E. (2018). An Explorative Approach to Freight Trip Attraction in an Industrial Urban Area. In E. Taniguchi & R. G. Thompson (Eds.), City Logistics 3: Towards Sustainable and Liveable Cities (pp. 249-266). London, UK: Wiley.

This paper contributes to the field of FTG through the estimation of establishment FTA models in a Norwegian context. Two model specifications are presented: one controlling for business size (using number of employees), industry group and whether the establishment is intermediary or not; another extending the first model to include firm and shipment specific explanatory variables. The models are estimated on data from 271 establishments in Groruddalen, an urban area in the Norwegian capital Oslo. Due to the small dataset, some industry groups suffered from low response, while others were represented by more than 100 respondents. Both model specifications are estimated on all industries and observations, and for each of the three industry groups with more than 30 observations: manufacturing, construction and sanitation, and retail.

The paper supports existing literature suggesting that employment is a key variable in the FTG modelling and that different industries require different models. Other variables in addition to employment, like degree of consolidation and decision making, seem to be relevant for explaining FTA, at least in an industrial area with a large mix of industries. The paper also reveals that retail, being the largest industry in the population, net sample and dataset, is an important sector for further research.

The novelties in this research are i) the estimation of an FTA model on an industrial urban area working as a freight hub, but not a city center, ii) contribution to the limited research on FTG in Europe, iii) investigation of model functional form (as called for by Holguin-Veras et al. (2011)), providing reasons to believe in a non-linear relationship between freight trips and employment (through a logarithmic transformation of both variables), and iv) revealing that information other than establishment size and industry type might be relevant when explaining FTA.

1.4.2 Paper 2: Latent split of aggregate counts: revealing home deliveries per commodity types and potential freight trip implications

Caspersen, E., Arrieta-Prieto, M. E. & Wang, X.

This paper also deals with FTG but considers shipments to consumers rather than to establishments. The main motivation for studying shipments to consumers is the increase in online shopping and B2C deliveries.

The objective of this paper was to estimate the propensity to shop online and the corresponding number of shipments received in total and per aggregated commodity group for online shoppers. As such, a negative binomial hurdle model, taking into account the latent split of the aggregated counts, was chosen for estimation purposes. The model builds on the work by Afghari et al. (2016) and Afghari et al. (2018), and has been extended to allow for zero observations in the commodity groups. A validation procedure for the proposed splitting of the total counts is discussed; followed by a remedial mechanism when the data suggest minor deviations from the assumptions. The model was estimated on 1,440 respondents (from a dataset collected by a third party in January 2017).

The estimation results confirm existing research on online shopping behavior in the literature: elderly people are less likely to buy online; high income, education and having kids motivate online shopping. The analysis of the number of shipments per commodity group shows that an average online shopper receives 2.4 shipments per month, or 0.077 shipments per day. Clothes, beauty and interior products have the highest relative frequency of shipments, followed by children's and leisure goods. The number of shipments depend on the consumer profile: consumers with kids in the household have the highest average number of shipments in total, and for all but one commodity group, females receive more than twice as many shipments of clothes, beauty and interior products than men, while men purchase electronics to a much higher extent than females. Correlations between different commodity groups are explored, revealing great consolidation opportunities and potential for reducing the number of shipments up to 30 %.

The novelties in this research are i) the estimation of the number of shipments per commodity and hence revealing some of the heterogeneity in FTA to consumers, and ii) suggesting how knowledge of commodity groups and purchase behavior can help policy makers and transport companies achieve more sustainable freight transport in urban areas through consolidation.

1.4.3 Paper 3: Consumer preferences for reducing environmental impacts of last mile deliveries. Case: female clothing rentals

Caspersen, E. & Navrud, S.

In Paper 3, the research has shifted from FTG activities to freight trip reducing measures. Consumers are again the targeted group, of which their preferences for environmentally friendly last mile deliveries are investigated. More specifically, the tradeoff between delivery time (ranging from 1 to 20 days) and emission (CO_2 and particulate matter (PM)) are of interest. The tradeoff is investigated using a discrete choice experiment (DCE) where respondents are asked to choose their preferred last mile delivery alternative when renting clothes online. The different last mile delivery alternatives differ in terms of five attributes: delivery time, delays, information services (notification of product quality and departure), PM emission, and CO_2 emission. The data from the DCE were analyzed using both multinomial logit (MNL) and latent class (LC) models to reveal (observed and unobserved) heterogeneity among consumer preferences. An internet panel of female respondents between 18- and 70-years were used to collect the data, resulting in a sample of 513 females.

The results revealed that female consumers renting clothes online have a negative utility of both increased emissions and delivery time. Low age and high frequency of online shopping are key indicators of time sensitivity, while environmental awareness is the key indicator of emission mitigation. The LC model reveals four consumer groups, of which two groups are found likely to accept a delivery time of up to 10 days to reduce CO_2- and PM-emissions from last mile transport. The most important characteristic explaining this likelihood is environmental awareness: those who totally agree that consumers must change their attitudes and behavior to solve the environmental challenges of today are more likely to accept a high delivery time to reduce CO_2- and PM-emissions from last mile transport than others. The overall take-home result is that people can accept a delivery time of 5 or even 10 days, but must be compensated in terms of reduced air emissions from the transport. The findings are viewed in relation to how they can contribute to more sustainable last mile delivery transport.

This paper is novel in many ways as it: i) reports consumers' valuation of important last mile delivery attributes: delivery time, delays, information services, PM emission and CO_2 emission, ii) documents a revealed general disutility of CO_2- and PM-emission, iii) adds to the scarce literature on last mile delivery from the sharing economy, represented by clothing rentals online, and iv) provides knowledge on how measures other than price can incentivize consumers to choose environmentally friendly deliveries.

1.4.4 Paper 4: Act locally? Are female online shoppers willing to pay to reduce the carbon footprint of last mile deliveries?

Caspersen, E., Navrud, S. & Bengtsson, J.

Paper 4 also analyzes consumer preferences for last mile deliveries, but uses a different sample than Paper 3 (although the data is collected in the same survey), with price replacing PM in the discrete choice experiment (DCE). The objective was to investigate if consumers are willing to pay for climate friendly deliveries (measured using CO_2-emission), and if so, how much. The sample for analysis was collected using an internet panel of female respondents between 18- and 70-years and cleaned to represent general online shopping, not clothing rentals in particular. This resulted in 460 respondents for analysis. The data from the DCE were analyzed using both multinomial logit (MNL) and mixed multinomial logit (MMNL) models to identify (observed and unobserved) heterogeneous consumer preferences.

The results show that females are willing to pay for CO_2-mitigation, and that the willingness to pay (WTP) increases with consumer income, employment, willingness to change habits to solve the environmental challenges of today, and a preference for sustainable online shopping and delivery alternatives, but falls with the frequency of online shopping. The WTP for 1kg CO_2 exceeds the WTP for any other aspects of the last mile delivery; i.e. delivery time, delays and information services (notification of departure and arrival). The results indicate that freight operators (carriers) and online retailers can transfer (some of) the cost of climate-friendly last-mile delivery to their customers. This knowledge is important for urban planners as it provides back up for CO_2-mitigating measures aimed at last mile delivery services to achieve more environmentally sustainable urban freight transport.

This paper's novel contribution is i) the investigation of female consumers' WTP for climate-friendly home deliveries in terms of reduced or no CO_2-emissions of last mile deliveries, and ii) the result that carriers can transfer some the costs from implemented measures to (parts of) the demand side. The limited ability for carriers to do so has been pointed out as a reason why measures targeting freight traffic have had a low effect on externalities (Holguín-Veras et al., 2015, Holguín-Veras et al., 2016).

1.5 Key findings and future research

1.5.2 Key findings

The following part summarizes the key findings of the papers and relates it to existing literature.

Explanatory variables should be explored in freight trip generation modeling: consolidation and decision-making may reduce freight trip attraction to establishments.

Paper 1 contributes to the, back then, limited literature of FTG models in a European context, and to FTG models for small industrial urban areas that are not the city center but still an important generator of freight activity. By investigating model functional form and explanatory variables not previously included in the literature, an exploratory approach to FTA model was taken. Models were estimated both across all industries and for separate industries with at least 30 observations: construction and sanitation, manufacturing and retail. When estimating across all industries, establishments using carriers for all their shipments were found to have fewer shipments than those not using carriers for all their shipment. This indicated that consolidation in fact helps to reduce freight trips to establishments. The impact of a high degree of decision-making is also positive, indicating that the firms themselves might not provide the most efficient transport solutions. The industry specific models confirmed the findings in the literature: different industries have different FTG models.

Similar results can be found in newer research, strengthening the findings in Paper 1. Gonzalez-Feliu and Sánchez-Díaz (2019) investigated aggregation levels and functional forms and found that the majority of the models display a functional form with a logarithmic transformation. Other studies also showed interested in exploring new explanatory variables, such as demographic, socio-economic and accessibility variables (Gonzalez-Feliu et al., 2020) or population density, employment density and the level of mixed land (De Bakshi et al., 2020).

Freight trip attraction numbers in a Norwegian industrial area.

Paper 1 also contributes updated numbers for FTA in Norway and how they differ between industries. The estimation results presented in Table 3 in Paper 1 are transformed (only significant variables are included) to reveal the relationship between FTA and employment rather than ln(FTA) and ln(employment). The models are presented in Table 3 below, and illustrated by calculating the FTA per firm using employment of 1, 10 and 20. These establishment numbers are chosen to represent the median number for employment from the data, ranging from 9 to 23 employees per establishments (the average employment number is higher due to some large businesses in the data).

Table 3: Relationship between employment and freight trip attraction during a typical week from Table 3 in Paper 1.

Industry	Model		Employment (empl) = 1	= 10	= 20
All industries	$FTA = e^{\alpha} \cdot e^{\beta_1 TW} \cdot empl^{\beta_2}$	$FTA_{TW} = 6.939 \cdot empl^{0.291}$	6.9	13.6	16.6
		$FTA_{Other} = 1.091 \cdot empl^{0.291}$	1.1	2.1	2.6
Construction and sanitation (inc. waste management)	$FTA = e^{\beta_1 Inter} \cdot empl^{\beta_2}$	$FTA_{Inter} = 11.450 \cdot empl^{0.877}$	11.5	86.3	158.4
		$FTA_{NoInter} = empl^{0.877}$	1.0	7.5	13.8
Manu-facturing	$FTA = e^{\alpha} \cdot empl^{\beta_1}$	$FTA = 4.702 \cdot empl^{0.638}$	4.7	20.4	31.8
Retail	$FTA = e^{\alpha} \cdot empl^{\beta_1}$	$FTA = 2.807 \cdot empl^{0.693}$	2.8	13.8	22.4

α=constant, TW = Transportation and warehousing industry, Inter = Intermediary firm that both produces and attracts freight trips

Sánchez-Díaz (2016) estimated a linear model for FTA in Gothenburg and showed that a typical establishment attracts a base number of trips ranging from 1.23 to 10.24 trips per week, plus an extra number of trips per employee, that can be as high as 1.42 (depending on the industry). Assuming 10 employees, this provides an FTA range of 15.43 to 24.44 trips per establishment per week. A separate model was estimated for non-perishable retail revealing an FTA per establishment per week of 14.2 (with 10 employees). Holguín-Veras et al. (2017) presented recommended non-linear FTA models for deliveries per day for different industries in New York City and New York State Capital Region (Table 10 p. 41 in Holguín-Veras et al., 2017). Using their models, an average of 5.5 delivery days per week (Holguin-Veras et al., 2018) and 10 employees per establishment, provides 16.21 deliveries/week for construction, 8.09 for manufacturing, 23.23 for retail and 20.47 for all industries. Thus, the numbers presented in Table 3 are in the range of both studies despite the different industry classifications and cities making the transferability uncertain. The transferability of estimates has not been analyzed in this thesis and is a matter for future research.

Paper 1 reveals that retail contributes greatly to FTA, both through a high number of establishments but also through a high number of FTA. On average, a retail firm with 10 employees received 2.52 shipments per day (assuming 5.5 delivery days per week).

Online shopping increases deliveries to consumers: what is the magnitude of freight trip generation by consumers compared to freight trip generation by firms?

Paper 2 contributes to the limited knowledge on FTG from consumers by developing a model for individual probability to shop online as well as the number of shipments purchased in total and per aggregated commodity group. The model is estimated on demographic variables. The results showed that higher income, education, number of kids and the distance (km) to an urban area motivate online shopping while age has the opposite effect. On average, an individual that shops online attracts 2.4

shipments per month, which equals 0.077 shipments per day if assuming delivery all days, or 0.096 if assuming 5.5 delivery days per week (as done for establishments above). This number places itself in the upper part of existing estimates from the US and Europe. In New York State, Wang and Zhou (2015) found 0.43 home deliveries per person per month, while it has reached 0.12 operations per person per day in the New York metro area (Dablanc, 2019 (provided by José Holguín Veras), Holguin-Veras et al., 2018). For European cities, Dablanc (2019) finds that, based on results from Gardrat et al. (2016) and Allen et al. (2017), the deferred purchase and reception in Lyon is 0.03 per person per day, and the number of operations from e-commerce in London is approximately 0.05 operations per person per day, including take-outs and other home delivered meals. In Paris, instant deliveries are calculated to represent 12 % of B2C related deliveries and pick-ups, with 0.2 instant deliveries per home per week (in 2016) (Dablanc et al., 2017).

Although the number of shipments per individual from online shopping (B2C) is smaller than to retail businesses (B2B), 0.096 shipments per person per day vs 2.52 shipments per retail firm per day, the number of individuals and their location in residential or mixed land use areas should put B2C on the agenda for municipalities and public planners together with B2B. For illustration purposes, only 26 individuals who shop online are expected to attract the same amount of freight trips per day as one retail firm employing 10 persons. Note that returns (freight trip production) are not included in these estimates, and that one shipment is assumed to imply one trip (consolidation is not considered). How these factors would influence the results is left for future research.

Online shopping increases deliveries to consumers: are there unexplored consolidation opportunities?

Paper 2 also provides estimates of the number of shipments received per commodity group, and the percentage split is presented in Figure 3.

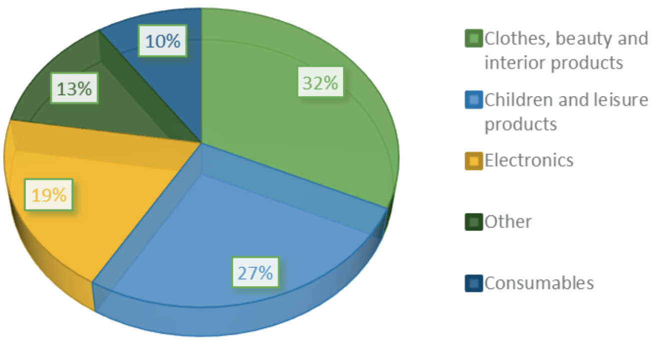

Figure 3: the percentage split across commodity groups found in Paper 2. The total average number of shipments per person per month is estimated to 2.4.

The results resemble those found by Buldeo Rai et al. (2021): clothing, shoes and accessories are the most popular products bought online, followed by electronics, beauty and healthcare, consumables and interior.

The correlation between the commodity groups was investigated for different consumer demographics (having kids, age and income). This revealed that several commodity groups were commonly bought together. Shipments of consumables and clothes showed a positive correlation for almost all combinations of age, income and family composition (kids), with as much as 56 % of some consumer types receiving shipments of both commodities in the same month. Similar results are found for consumables and children and leisure products when looking at consumers with kids: as much as 42 % of consumers received shipments of both commodity types in the same month. This tells us that considerable consolidation opportunities exist and that freight trips can be saved by offering consumers the opportunity to consolidate their shipments of different commodity groups. Consolidation opportunities were also found for homogenous neighborhoods, in particular suburban areas with high-income families.

The results correspond with the business strategy of a Norwegian online retailer: Kolonial.no. Kolonial.no started out as a provider of groceries online but has now expanded to sell and deliver goods from other established brands, such as Clas Ohlson (product and leisure store) or Iceland (frozen groceries). Additionally, they have partnered with Sprell (toy store), Milrab.no and Anton Sport (both sports shops) to distribute their deliveries[5]. As such, Kolonial.no increases their consolidation of goods that in Paper 2 are found to be attracted by the same customers and neighbourhoods. Through a matrix of alternative delivery times, the customers of Kolonial.no can choose delivery at a more environmentally sustainable slot, contributing even more to consolidation.

One relevant question is whether consumers are interested in waiting to receive their deliveries so that freight operators can increase consolidation. Paper 3 shows that some consumer types, though not all, might be willing to wait as much as 10 days for a delivery if it benefits the environment (measured in reduced airborne emission of CO_2 and PM from last mile delivery of clothing rentals online). The key identifying variable is related to environmental behavior: consumers who are willing to change their attitudes and behavior to solve the environmental challenges of today. This attitude is found among average consumers, but also among high income, frequent online shoppers. Young individuals or people feeling less responsible for the environmental challenges are less likely to trade delivery time for emission reduction from last mile transport.

Can the provision of environmentally sustainable last mile delivery alternatives reduce the negative externalities of urban freight traffic?

[5] https://no.ehandel.com/kolonial-no-inngar-samarbeid-med-milrab-anton-sport-og-sprell?utm_campaign=cmp_2307123&utm_source=getanewsletter&utm_medium=email

Delivery services are of concern to both the suppliers and demanders of e-commerce, have been for a while (Morganti et al., 2014), and experience increased interest from researchers. Both Paper 3 and Paper 4 investigate consumer preferences for environmentally sustainable last mile deliveries and reveal that consumers have a negative utility of airborne emissions from last mile delivery transport, and that measures to reduce emissions can take different approaches. Using different approaches to reduce the negative impacts of urban freight traffic is not a new idea. Allen et al. (2018) suggest several measures to increase the efficiency of parcel deliveries, including measures that might benefit from more patient consumers, such as last-mile collaboration, logistics hotels, crowdshipping and pick-up points, the latter being related to benefits such as consolidating parcel drop-offs and optimizing delivery rounds (de Oliveira et al., 2017, Morganti et al., 2014). However, adapting delivery solutions to different consumer preferences might improve the success rate of the measures.

The main finding in Paper 3, that (some) consumers are willing to increase delivery time for reduced emissions from last mile delivery, can support consolidation as shown above, but also other measures. Examples are pick-up points, where delivery time increases as consumers themselves must pick up the goods at a given location, rather than at home, or crowdshipping, which involves an extra transfer of the shipment to the crowd, and, if delivered to pick-up points, a pickup by the consumer. If the environmental benefits of such measures are clearly stated, consumers might prefer them even if they imply increased delivery time. The results from Paper 4 also support environmentally sustainable last mile deliveries as they reveal that consumers are willing to pay to reduce CO_2-emissions. Thus, cities or carriers can expect to be able to share some of their investment costs or incremented costs from environmentally sustainable delivery technology with the consumers. The low ability to share costs with the demand side has been pointed out as one of the reasons why measures targeting freight traffic externalities through carriers have a small effect and why carriers oppose command and control measures enforcing new investments (Holguín-Veras et al., 2015, Holguín-Veras et al., 2016).

Is last mile delivery service less important than we think?

In Paper 3, a group of consumers with what seems to be inconclusive preferences for last mile delivery attributes was identified. The same finding was made by Buldeo Rai et al. (2021) in their study of consumer preferences for innovative last mile deliveries. This begs the question: can the sophisticated consumer be a result of competition in the market rather than actual consumer preferences? This calls for further research on whether last mile delivery service is less important to customers than what researchers, online retailers and transport operators think.

Can this research help municipalities and public planners reduce externalities from urban freight traffic?

This thesis shows that FTG by consumers is non-negligible and should be emphasized in public planning along with FTG to establishments. This is a key finding for urban development, in particular for projects planning for mixed land use including businesses, residents and recreational areas like parks, playgrounds etc., but also for political measures targeting freight traffic externalities. For the latter, knowledge about the demand side is useful to develop a strong link between the measure and the target. Several potential measures to reduce urban road freight externalities have prevailed from this research.

Consolidation opportunities for shipments from online shopping exist, both at individual and neighborhood level, and can reduce freight trips to consumers. Allen et al (2018) also state that operating costs might be reduced by increasing the vehicle load factor as well as the drop density per round trip. Delivering more goods to the same consumer at a time, would target these aims. However, if not solved by the market, public planners might advocate measures consolidating the transportation of goods such as groceries and clothing, interior, and beauty products, or groceries and children's goods and leisure products.

Several consumers accept longer delivery times to reduce emissions of PM and CO_2. As suggested above, this supports innovative delivery solutions with lower environmental footprint than home deliveries, such as pick-up points and crowdshipping. The research also shows that consumers are interested in and willing to pay for environmentally sustainable last mile transport, although the degree of interest differs with their attitudes and habits. This suggests that measures aimed at improving the sustainability of freight transport might be feasible without directly reducing firms' profit margins. However, a clever selection of last mile delivery alternatives must be provided to the consumers to secure their contribution.

Implementation of measures to improve the environmental sustainability of last mile delivery and reduce its externalities is key. Past experience of innovative logistics solutions may discourage future participation intention, as found for crowdshipping (Punel and Stathopoulos, 2017), self-collection services (Wang et al., 2019) and parcel lockers (Vakulenko et al., 2019). Hence, when enforcing new solutions for the consumers, professional initial customer interaction is important and defines the future of the solution (Vakulenko et al., 2019). Thus, municipalities and urban planners should contribute to the smooth implementation of innovative solutions when necessary and possible.

1.5.2 Future research

Starting with the words of Pigou, this thesis also finishes with them: "Economic science must always speak with an uncertain voice" (Pigou, 1920). As with all other economic studies, there are uncertainties in the research presented in this thesis. The results from Papers 1, 3 and 4 rely on samples that are limited in terms of geography (Paper 1) and gender and purchase (Paper 3 and Paper 4). As such, its generalization potential is limited. More knowledge on FTG to establishments in Norway could be gained by replicating the analysis in Paper 1 for different areas in Oslo or different cities in Norway. Reproducing the DCE reported in Papers 3 and 4 including males, consumer goods other than clothing rentals, or even countries, would shed light on the generalizability of the results from these two papers. The study in Paper 2 uses data from online shopping in December only, and a test using data from other months could strengthen the results. Returns are not investigated, despite being present for both establishments (return of products, packaging, waste) and consumers (return of products). Further, attribute levels in the DCE can possibly influence the results. Although attributes and attribute levels are chosen under careful consideration, and consumer knowledge and understanding of emission levels might be limited, the impression is that knowledge of emission, in particular CO_2-emissions, has increased in the public over the last decade, with an increasing emphasis on the carbon footprint of activities. However, the results might have been different if the levels were presented in other ways, i.e. as a percentage of the average consumer's annual emissions of CO_2 instead of in kg. These questions are not addressed in this thesis but of interest for future research.

1.6 Concluding remarks

The overall research questions motivating this thesis are:

What is the magnitude of freight trip generation by consumers compared to freight trip generation by establishments? Can externalities of urban freight traffic be reduced by targeting consumer last mile delivery alternatives?

This has been sought to be answered through four papers focusing on freight trip attraction (FTA) from establishments, FTA from consumers, as well as consumer preferences for environmentally sustainable last mile delivery transport. The overall finding is that freight trip generation (FTG) per consumer is still smaller than FTG per establishment (measured using FTA models for retail activities), but that the large numbers of consumers compared to retail stores makes their contribution non-negligible. Thus, FTG from both establishments and consumers should be considered in urban planning. With the (increasing) role of consumers as generators of freight traffic comes new opportunities to reduce its negative impacts. Research in this thesis shows that consumers care for environmentally sustainable last mile deliveries. Hence, by offering consumers the right bundle of choices (in terms of consumer group), they can contribute to more sustainable urban freight transport and reduced externalities just by choosing their preferred delivery alternative. Consumers are even willing to take on some of the costs, either by paying for reduced CO_2-emission or waiting some extra days for the delivery to reduce PM- and CO_2-emissions. Thus, the findings support the idea of freight demand management being a useful approach for targeting freight traffic externalities as consumers, an important source of freight traffic, can be influenced to choose more environmentally sustainable freight transport. Another contribution of this thesis is the use of an all-female sample in two of the four papers. Many papers on environmental preferences within transport, at least on consumer preference for emission reduction in terms of vehicle choice, use male dominant samples (Costa et al., 2019, Achtnicht, 2012).

The objective of this thesis has been to increase the knowledge and understanding of urban freight trip generation and possible measures to reduce its magnitude and externalities imposed on urban residents. The humble opinion of the author is that the research presented has succeeded in this goal, but that a huge amount of research remains to be done to achieve a full understanding of the topic.

1.7 Bibliography

ACHTNICHT, M. 2012. German car buyers' willingness to pay to reduce CO2 emissions. *Climatic Change,* 113**,** 679-697.

AFGHARI, A. P., HAQUE, M. M., WASHINGTON, S. & SMYTH, T. 2016. Bayesian Latent Class Safety Performance Function for Identifying Motor Vehicle Crash Black Spots. *Transportation Research Record,* 2601**,** 90-98.

AFGHARI, A. P., WASHINGTON, S., HAQUE, M. M. & LI, Z. 2018. A comprehensive joint econometric model of motor vehicle crashes arising from multiple sources of risk. *Analytic Methods in Accident Research,* 18**,** 1-14.

AKERSHUS FYLKESKOMMUNE 2018. Regional plan for handel, service og senterstruktur i Akershus.

ALBERINI, A., BIGANO, A., ŠČASNÝ, M. & ZVĚŘINOVÁ, I. 2018. Preferences for Energy Efficiency vs. Renewables: What Is the Willingness to Pay to Reduce CO2 Emissions? *Ecological Economics,* 144**,** 171-185.

ALHO, A. & SILVA, J. D. A. E. 2014. Freight-Trip Generation Model: Predicting Urban Freight Weekly Parking Demand from Retail Establishment Characteristics. *Transportation Research Record: Journal of the Transportation Research Board,* 2411**,** 45-54.

ALLEN, J. & BROWNE, M. 2008. Review of survey techniques used in urban freight studies. Green Logistics Report, University of Westminster.

ALLEN, J., PIECYK, M. & PIOTROWSKA, M. 2017. An analysis of online shopping and home delivery in the UK. *FTC 2050.* Westminster, UK: University of Westminister.

ALLEN, J., PIECYK, M., PIOTROWSKA, M., MCLEOD, F., CHERRETT, T., GHALI, K., NGUYEN, T., BEKTAS, T., BATES, O., FRIDAY, A., WISE, S. & AUSTWICK, M. 2018. Understanding the impact of e-commerce on last-mile light goods vehicle activity in urban areas: The case of London. *Transportation Research Part D: Transport and Environment,* 61**,** 325-338.

BEN-AKIVA, M. & LERMAN, S. R. 1985. *Discrete Choice Analysis: Theory and Application to Travel Demand,* USA, The MIT Press.

BRING RESEARCH 2019. Slik velger kundene deg. *In:* NORGE, P. (ed.).

BROWNE, M., ALLEN, J., STEELE, S., CHERRETT, T. & MCLEOD, F. 2010. Analysing the results of UK urban freight studies. *Procedia - Social and Behavioral Sciences,* 2**,** 5956-5966.

BULDEO RAI, H., VERLINDE, S. & MACHARIS, C. 2021. Who is interested in a crowdsourced last mile? A segmentation of attitudinal profiles. *Travel Behaviour and Society,* 22**,** 22-31.

CALTHROP, E. & PROOST, S. 1998. Road Transport Externalities. *Environmental and Resource Economics,* 11**,** 335.

CAMPBELL, S., HOLGUÍN-VERAS, J., RAMIREZ-RIOS, D. G., GONZÁLEZ-CALDERÓN, C., KALAHASTHI, L. & WOJTOWICZ, J. 2018. Freight and service parking needs and the role of demand management. *European Transport Research Review,* 10**,** 47.

CARBONE, V., ROUQUET, A. & ROUSSAT, C. 2018. A typology of logistics at work in collaborative consumption. *International Journal of Physical Distribution & Logistics Management,* 48**,** 570-585.

CARLING, K., HAN, M., HÅKANSSON, J., MENG, X. & RUDHOLM, N. 2015. Measuring transport related CO2 emissions induced by online and brick-and-mortar

retailing. *Transportation Research Part D: Transport and Environment,* 40, 28-42.
CARSON, R. T. & CZAJKOWSKI, M. 2019. A new baseline model for estimating willingness to pay from discrete choice models. *Journal of Environmental Economics and Management,* 95, 57-61.
CASPERSEN, E. & ØRVING, T. 2018. Kunnskapsgrunnlag for mer klimavennlig næringstrafikk i Oslo. *TØI rapport 1622/2018.* Oslo: Transportøkonomisk institutt.
CASPERSEN, E. & PINCHASIK, D. R. 2017. Innsamling og bruk av virksomhetsdata for informasjon om næringstrafikk i et byområde. Eksempel fra Groruddalen i Oslo. Oslo, Norway: Transportøkonomisk institutt.
COSTA, E., MONTEMURRO, D. & GIULIANI, D. 2019. Consumers' willingness to pay for green cars: a discrete choice analysis in Italy. *Environment, Development and Sustainability,* 21, 2425-2442.
COULOMBEL, N., DABLANC, L., GARDRAT, M. & KONING, M. 2018. The environmental social cost of urban road freight: Evidence from the Paris region. *Transportation Research Part D: Transport and Environment,* 63, 514-532.
COWI AS, MAD AS & GRO AS 2014. Forslagsstillers planbeskrivelse Økern Torgvei 30 Planforslag til politisk behandling
DABLANC, L. 2019. *E-commerce trends and implications for urban logistics. In, Browne, M., Behrends, S., Woxenius, J., Giuliano, G., Holguin-Veras, J. Urban logistics. Management, policy and innovation in a rapidly changing environment*.
DABLANC, L., MORGANTI, E., ARVIDSSON, N., WOXENIUS, J., BROWNE, M. & SAIDI, N. 2017. The rise of on-demand 'Instant Deliveries' in European cities. *Supply Chain Forum: An International Journal,* 18, 203-217.
DALY, A., HESS, S. & TRAIN, K. 2012. Assuring finite moments for willingness to pay in random coefficient models. *Transportation,* 39, 19-31.
DARK + ADEPT 2013. Helhetsplanen Føyka/Elvely.
DAZIANO, R. & ACHTNICHT, M. 2013. Accounting for Uncertainty in Willingness to Pay for Environmental Benefits *ZEW - Centre for European Economic Research Discussion,* Paper No. 13-059.
DE BAKSHI, N., TIWARI, G. & BOLIA, N. B. 2020. Influence of urban form on urban freight trip generation. *Case Studies on Transport Policy,* 8, 229-235.
DE BORGER, B., OCHELEN, S., PROOST, S. & SWYSEN, D. 1997. Alternative transport pricing and regulation policies: A welfare analysis for Belgium in 2005. *Transportation Research Part D: Transport and Environment,* 2, 177-198.
DE OLIVEIRA, L. K., MORGANTI, E., DABLANC, L. & DE OLIVEIRA, R. L. M. 2017. Analysis of the potential demand of automated delivery stations for e-commerce deliveries in Belo Horizonte, Brazil. *Research in Transportation Economics,* 65, 34-43.
ECOMMERCE NEWS. 2020. *Ecommerce in Europe: €717 billion in 2020* [Online]. https://ecommercenews.eu/ecommerce-in-europe-e717-billion-in-2020/. [Accessed].
ELLRAM, L. M. & MURFIELD, M. L. U. 2017. Environmental Sustainability in Freight Transportation: A Systematic Literature Review and Agenda for Future Research. *Transportation Journal,* 56, 263-298.

EUROPEAN COMMISSION 2019. Handbook on the external costs of transport. *In:* TRANSPORT, D.-G. F. M. A. (ed.). Delft: CE Delft.

EUROPEAN ENVIRONMENT AGENCY 2019. Cutting air pollution in Europe would prevent early deaths, improve productivity and curb climate change. *In:* EUROPEAN ENVIRONMENT AGENCY (ed.). https://www.eea.europa.eu/highlights/cutting-air-pollution-in-europe.

FOSSHEIM, K., CASPERSEN, E., BJØRGEN, A., KARLSSON, H. & EIDHAMMER, O. 2019. What do Norwegian citites need to plan for urban logistic? Oslo, Norway: Transportøkonomisk institutt.

GARDRAT, M., TOILIER, F., PATIER, D. & ROUTHIER, J.-L. 2016. The impact of new practices for supplying households in urban goods movements: method and first results. An application for Lyon, France. *VREF conference on Urban Freight 2016.* Göteborg, Sweden.

GATTA, V., MARCUCCI, E., PIRA, M. L., INTURRI, G., IGNACCOLO, M. & PLUCHINO, A. 2020. E-groceries and urban freight: Investigating purchasing habits, peer influence and behaviour change via a discrete choice/agent-based modelling approach. *Transportation Research Procedia,* 46**,** 133-140.

GONZALEZ-FELIU, J., PALACIOS-ARGÜELLO, L. & SUAREZ-NANEZ, C. 2020. Links between freight trip generation rates, accessibility and socio-demographic variables in urban zones. *Archives of Transport,* 53.

GONZALEZ-FELIU, J. & PERIS-PLA, C. 2017. Impacts of retailing attractiveness on freight and shopping trip attraction rates. *Research in Transportation Business & Management,* 24**,** 49-58.

GONZALEZ-FELIU, J. & SÁNCHEZ-DÍAZ, I. 2019. The influence of aggregation level and category construction on estimation quality for freight trip generation models. *Transportation Research Part E: Logistics and Transportation Review,* 121**,** 134-148.

HAMARI, J., SJÖKLINT, M. & UKKONEN, A. 2016. The sharing economy: Why people participate in collaborative consumption. *Journal of the Association for Information Science and Technology,* 67**,** 2047-2059.

HOLGUÍN-VERAS, J., AROS-VERA, F. & BROWNE, M. 2015. Agent interactions and the response of supply chains to pricing and incentives. *Economics of Transportation,* 4**,** 147-155.

HOLGUIN-VERAS, J., JALLER, M., DESTRO, L., XUEGANG, B. J., LAWSON, C. & LEVINSON, H. S. 2011. Freight Generation, Freight Trip Generation, and the Perils of using Constant Trip Rates. *Transportation Research Record,* 2224**,** 68-81.

HOLGUÍN-VERAS, J., JALLER, M., SÁNCHEZ-DÍAZ, I., WOJTOWICZ, J., CAMPBELL, S., LEVINSON, H., LAWSON, C., POWERS, E. & TAVASSZY, L. 2012. NCHRP Report 739/NCFRP Report 19: freight trip generation and land use. *Washington DC: Transportation Research Board of the National Academies*.

HOLGUÍN-VERAS, J., LAWSON, C., WANG, C., JALLER, M., GONZÁLEZ-CALDERÓN, C., CAMPBELL, S., KALAHASHTI, L., WOJTOWICZ, J. & RAMIREZ-RÍOS, D. 2017. Using Commodity Flow, Survey Microdata, and Other Establishment Data to Estimate the Generation of Freight, Freight Trips, and Service Trips: Guidebook. National Academies of Science, Engineering, and Medicine.

HOLGUIN-VERAS, J., RAMIREZ-RÍOS, D. G., ENCARNACIÓN, T., GONZÁLEZ-FELIU, J., CASPERSEN, E., RIVERA-GONZÁLEZ, C., GONZÁLEZ-CALDERÓN, C. A. & DA SILVA LIMA, R. 2018. Metropolitan economies and the generation of freight and service activity: an international perspective. *In:* BROWNE, M.,

BEHRENDS, S., WOXENIUS, J., GIULIANO, G. & HOLGUIN-VERAS, J. (eds.) *Urban Logistics: management, policy and innovation in a rapdily changing environment.* London (UK) and New York (US): Kogan Page Limited.
HOLGUÍN-VERAS, J., SÁNCHEZ-DÍAZ, I. & BROWNE, M. 2016. Sustainable Urban Freight Systems and Freight Demand Management. *Transportation Research Procedia,* 12**,** 40-52.
HOLGUÍN-VERAS, J., SÁNCHEZ-DÍAZ, I., LAWSON, C. T., JALLER, M., CAMPBELL, S., LEVINSON, H. S. & SHIN, H.-S. 2013. Transferability of freight trip generation models. *Transport Research Record,* 2379**,** 1-8.
JALLER, M., SÁNCHEZ-DÍAZ, I. & HOLGUÍN-VERAS, J. 2015. Identifying Freight Intermediaries. *Transportation Research Record: Journal of the Transportation Research Board,* 2478**,** 48-56.
JOHNSTON, R. J., BOYLE, K. J., ADAMOWICZ, W. V., BENNETT, J., BROUWER, R., CAMERON, T. A., HANEMANN, W. M., HANLEY, N., RYAN, M., SCARPA, R., TOURANGEAU, R. & VOSSLER, C. A. 2017. Contemporary Guidance for Stated Preference Studies. *Journal of the Association of Environmental and Resource Economists,* 4**,** 319-405.
LIN, J., ZHOU, W. & DU, L. 2018. Is on-demand same day package delivery service green? *Transportation Research Part D: Transport and Environment,* 61**,** 118-139.
MAGNUSSEN, K., GIERLOFF, C. W., SEEBERG, A. R. & NAVRUD, S. 2017. Den tette byens verdi. *MENON-PUBLIKASJON.* MENON.
MARCUCCI, E. & GATTA, V. 2012. Dissecting preference heterogeneity in consumer stated choices. *Transportation Research Part E: Logistics and Transportation Review,* 48**,** 331-339.
MCFADDEN, D. 1974. Conditional logit analysis of qualitative choice behavior. *In:* P.ZAREMBKA (ed.) *Frontiers in econometrics.* New York, 1973: Academic Press.
MCFADDEN, D. & TRAIN, K. 2000. Mixed MNL models for discrete response. *Journal of Applied Econometrics,* 15**,** 447-470.
MORGANTI, E., SEIDEL, S., BLANQUART, C., DABLANC, L. & LENZ, B. 2014. The Impact of E-commerce on Final Deliveries: Alternative Parcel Delivery Services in France and Germany. *Transportation Research Procedia,* 4**,** 178-190.
PIGOU, A. C. 1920. *The Economics of Welfare,* London, Macmillan.
PLAN- OG BYGNINGSETATEN 2019a. Østre Aker vei 29, Økern: PBEs område- og prosessavklaring til oppstartmøte for planer med krav om konsekvensutredning. Oslo Kommune.
PLAN- OG BYGNINGSETATEN 2019b. Planprogram med veiledende plan for offentlige rom for Rødtvet Forslag til politisk behandling. Oslo: Oslo kommune.
PUNEL, A. & STATHOPOULOS, A. 2017. Modeling the acceptability of crowdsourced goods deliveries: Role of context and experience effects. *Transportation Research Part E: Logistics and Transportation Review,* 105**,** 18-38.
REINSBORG, J., KAURIN, T., SANDELIEN, S., NYQUIST, B., MARTINSEN, J. A. & FOSS, T. 1989. *Trafikkberegninger*, Vegdirektoratet.
RØDSETH, K. L., WANGSNESS, P. B., VEISTEN, K., HØYE, A. K., ELVIK, R., KLÆBOE, R., THUNE-LARSEN, H., FRIDSTRØM, L., LINDSTAD, E., RIALLAND, A., ODOLINSKI, K. & NILSSON, J.-E. 2019. The external costs of transport – Marginal damage cost estimates for passenger and freight transport in

Norway. *In:* INSTITUTT, T. (ed.) *TØI-report.* Oslo, Norway: Transportøkonomisk institutt.
ROUQUET, A., GOUDARZI, K. & HENRIQUEZ, T. 2017. The company-customer transfer of logistics activities. *International Journal of Operations & Production Management,* 37**,** 321-342.
SÁNCHEZ-DÍAZ, I. 2016. Modeling urban freight generation: A study of commercial establishments' freight needs. *Transportation Research Part A: Policy and Practice,* 102**,** 3-17.
SANTOS, G. 2017. Road fuel taxes in Europe: Do they internalize road transport externalities? *Transport Policy,* 53**,** 120-134.
SANTOS, G., BEHRENDT, H., MACONI, L., SHIRVANI, T. & TEYTELBOYM, A. 2010. Part I: Externalities and economic policies in road transport. *Research in Transportation Economics,* 28**,** 2-45.
SAPHORES, J.-D. & XU, L. 2020. E-shopping changes and the state of E-grocery shopping in the US - Evidence from national travel and time use surveys. *Research in Transportation Economics***,** 100864.
SHEN, J. 2009. Latent class model or mixed logit model? A comparison by transport mode choice data. *Applied Economics,* 41**,** 2915-2924.
SILLANO, M. & ORTÚZAR, J. D. D. 2005. Willingness-to-pay estimation with mixed logit models: Some new evidence. *Environment and Planning A,* 37**,** 525-550.
VAKULENKO, Y., SHAMS, P., HELLSTRÖM, D. & HJORT, K. 2019. Service innovation in e-commerce last mile delivery: Mapping the e-customer journey. *Journal of Business Research,* 101**,** 461-468.
VALERI, E., GATTA, V., TEOBALDELLI, D., POLIDORI, P., BARRATT, B., FUZZI, S., KAZEPOV, Y., SERGI, V., WILLIAMS, M. & MAIONE, M. 2016. Modelling individual preferences for environmental policy drivers: Empirical evidence of Italian lifestyle changes using a latent class approach. *Environmental Science & Policy,* 65**,** 65-74.
VAN DEN BERGH, J. C. J. M. & BOTZEN, W. J. W. 2015. Monetary valuation of the social cost of CO2 emissions: A critical survey. *Ecological Economics,* 114**,** 33-46.
VICTOR, P. A. 2020. Cents and nonsense: A critical appraisal of the monetary valuation of nature. *Ecosystem Services,* 42**,** 101076.
VISSER, J., NEMOTO, T. & BROWNE, M. 2014. Home Delivery and the Impacts on Urban Freight Transport: A Review. *Procedia - Social and Behavioral Sciences,* 125**,** 15-27.
WANG, X., YUEN KUM, F., WONG YIIK, D. & TEO, C.-C. 2019. Consumer participation in last-mile logistics service: an investigation on cognitions and affects. *International Journal of Physical Distribution & Logistics Management,* 49**,** 217-238.
WANG, X. & ZHOU, Y. 2015. Deliveries to residential units: A rising form of freight transportation in the U.S. *Transportation Research Part C: Emerging Technologies,* 58**,** 46-55.
WYGONIK, E. & GOODCHILD, A. V. 2018. Urban form and last-mile goods movement: Factors affecting vehicle miles travelled and emissions. *Transportation Research Part D: Transport and Environment,* 61**,** 217-229.
ZHAO, X., KE, Y., ZUO, J., XIONG, W. & WU, P. 2020. Evaluation of sustainable transport research in 2000–2019. *Journal of Cleaner Production,* 256**,** 120404.

Chapter 2:

An explorative approach to freight trip attraction in an industrial urban area.

Elise Caspersen

Published in City Logistics 3: Towards Sustainable and Liveable Cities

14

An Explorative Approach to Freight Trip Attraction in an Industrial Urban Area

This chapter presents a freight trip attraction (FTA) model estimated in a Norwegian context. The novelty of this research is the explanatory approach taken in a European context by estimating the FTA model on an industrial urban area, which functions as a freight hub but not as a city center. The model's functional form is investigated, providing reasons to believe in a nonlinear relationship between freight trips and employment. An extended version of the model, including establishment and shipment characteristics, shows that information other than establishment size and industry type might be relevant when explaining freight trip attraction. The model is estimated using data from 271 establishments in Groruddalen (an urban area in Oslo).

14.1. Introduction

Many cities suffer from inefficient management of urban freight transport. Consequently, freight pickup and delivery becomes tedious and time-consuming, leaving urban areas more congested than necessary. One reason for inefficient management is a general lack of understanding of freight activities in cities. Compared with personal transport, there is little knowledge about the supply and demand of freight transport and how heavy vehicle drivers' route choices affect road occupancy [BAS 09, HOL 11, BEN 16]. However, increased knowledge can be gained to solve these issues. For instance, a broader understanding of the underlying causes for demand and supply of freight transport can support estimates of freight traffic to and from establishments. Such information, also referred to as freight trip generation, is key to planning transportation systems [SÁN 16b]. Knowing the cause of transport (freight trip generation) and hence urban congestion can assist decision makers and transport planners when identifying and evaluating urban policy

Chapter written by Elise CASPERSEN.

City Logistics 3: Towards Sustainable and Liveable Cities, First Edition.
Edited by Eiichi Taniguchi and Russell G. Thompson.

[HOL 16]. The relevance of freight trip generation is also found with respect to transport models. González-Feliu *et al*. [GON 14] categorized transport models into four groups according to their applied methodological framework. Common to these groups is that all include calculation of freight trip generation in one way or another.

Despite the usefulness of freight trip generation in public planning and policy-making, there is, to the author's knowledge, limited knowledge about freight trip generation both in a Norwegian, and to some extent, a European context. The majority of freight trip generation models are estimated on data from the US [HOL 11, LAW 12, HOL 13, JAL 15, SÁN 16b, HOL 17]. These models are mostly estimated on relatively small areas or city centers in large cities, taking on an explorative approach. The applicability of US models to European cities is however uncertain [ADJ 16]. European freight trip generation models are less exploratory in terms of explicative variables and estimated for either a large city employer [ADJ 16], city center or city [ALH 14] (Portugal), [GON 14] (France), [GON 16] (France), [SÁN 16a] (Gothenburg), or national samples of establishments [IDI 02] (Netherlands), [REI 89] (Norway). Hence, there seems to be a research gap of models taking an in-depth approach to explicative variables affecting freight trip generation in a European context as well as freight trip generation models for non-negligible but small urban areas that are not city centers.

Based on the identified gaps in the literature, the contribution of this chapter is threefold. The first is an explanatory modeling approach for a European context. The existing literature shows that most models treat freight trips either as a constant per establishment or a linear function of explanatory variables, particularly business size. However, although freight production might be approximately linear to business size, there are several reasons why the number of freight trips to and from the establishment is not. For instance, increased business size and freight demand might result in larger shipment size or a change in transport mode [HOL 11]. Hence, the model functional form is explored.

In addition, the impact of extending the model with establishment and shipment characteristics are investigated. For freight trip generation models to be relevant for public policy planning, publicly available or easily retrieved policy-oriented explanatory variables are often used. However, this approach excludes how freight trip generation is affected by internal decisions. For instance, pre-arranged transport agreements or the flexibility to make daily changes in deliveries might affect freight trip generation. The second contribution is an analysis of FTG in a non-negligible industrial urban area, which functions as a freight hub but not as a city center. Third, as the analysis is done on the establishment data conducted in Groruddalen, an urban area in the Norwegian capital Oslo, a contribution to freight trip generation modeling in Norway is provided.

The rest of this chapter is organized as follows: the second section provides background literature and places this contribution among existing FTG models. The third section presents the data collection and corresponding data set. The estimation results, including a discussion of freight trip generation models' functional form, are presented in the fourth section, and followed by a suggested extension of the model. The final section sums up the chapter and concludes.

14.2. Background

The most frequently used FTG model in the reviewed literature seems to be a linear regression model estimated by ordinary least squares (OLS) with business size as the only explanatory variable. This approach often results in a FTG that is either constant per establishment or a linear function of business size, with or without a constant term. Business size is mostly represented by the establishment's number of employees [LAW 12, HOL 11, GON 14], building area [TAD 94] or a combination [REI 89, IDI 02] of both. Separate models are generally estimated for different industries or land use groups, resulting in different model specifications for different categories. Although such models are common in the field, Holguín-Veras *et al.* [HOL 13] pointed out that FTG models presented in the literature often have too low explanatory power (in general). In addition, a few studies have tested the significance of the independent variable but not the validity of its functional form [HOL 11].

Perhaps as a response, recent literature on the field shows a development within FTG modeling, including the introduction of new explanatory variables, nonlinear transformation of variables and nonlinear models. For instance, there has been a change in the way that land use and industry classification is included in the models: instead of estimating one FTG model per category, binary variables are included to control for the impact of different industries, land use classification and geographical areas [HOL 13, ALH 14, JAL 15, SÁN 16a].

Other variables, like supply chain variables [ALH 14], interaction terms [JAL 15] and location variables [SÁN 16b], are introduced. Alho and Silva [ALH 14] found non-normality between FTG and the independent variables for the number of deliveries per week, number of employees, number of suppliers, establishment area and warehouse area resulting in a transformation of variables. Sánchez-Díaz *et al.* [SÁN 16b] also found nonlinearity in both the dependent and independent variables, and that for the extreme skewness case, logarithmic variable transformation is most suitable.

Holguín-Veras *et al.* [HOL 17] used a power function to control for nonlinearity which, when taking its logarithm, equals a model where both dependent and independent variables are logarithmic. Nonlinearity is also introduced in the models itself through binary logic models and discrete-continuous models [JAL 15, SÁN 16a]. Other recent contributions to the research field are the introduction of freight intermediaries (firms that both attract and produce freight trips) [JAL 15], challenges related to aggregation levels, including how constant FTG models are not necessarily suitable when estimating FTG models on aggregated categories [GON 16] and how to control for spatial interaction between establishments [SÁN 16b].

With respect to Norwegian models of freight trip generation, only one study has been located. The work by Reinsborg *et al*. [REI 89] is a part of the Norwegian national guidelines on traffic calculations and it presented three ways to calculate trip generation. The first method relies on a constant, experience-based relationship between vehicle trips and activity per employee or per 100m^2 floor area. A distinction is made between the manufacturing, retail and office industry groups. The second method is a linear regression model similar to those in the literature. The third method is like the first, but parameter size differs with the explanatory variables. The authors present recommended values for commercial areas using the first method and data from a national survey conducted by the National Public Road Administration (NPRA) prior to 1989. However, the numbers include both commuting and commercial trips by establishment employees and visiting customers as well as freight trips. They are also generic, providing no clear distinction between freight trip generation in urban and non-urban areas. Hence, the numbers are not comparable with estimates in the literature. Using the second and third method, recommended values are provided for residential areas only. Although these methods are from 1989, they are still included in the public guidelines by the NPRA. This also underlines the need for updated estimates.

14.3. Data from establishments in Groruddalen

The data used in this analysis is collected by an establishment survey in Groruddalen. Groruddalen is a broad valley in the eastern part of Oslo, Norway, and it covers four of Oslo’s boroughs and a wide area of commercial and private activity including industries, offices, retails, private housing, schools, commercial premises and warehouses. Approximately 140,000 people reside in the valley [TVE 16]. In addition, Groruddalen houses the national railway terminal, Alnabru, and it is served by several motorways and highways. Groruddalen is an area of interest because it suffers from severe road congestion resulting from high commercial and private activity. This has advocated a need for the Norwegian Public Roads Administration (NPRA) to clarify the future highway system in the area. Hence, the data collection was conducted through collaboration with the NPRA.

The data collection was carried out through a self-reporting Internet survey. The survey was distributed by e-mail to establishments in Groruddalen. The survey was distributed on 30th of August 2016 following an informative e-mail distributed on 24th of August. The data collection period lasted until 23rd of September 2016. Two reminders were sent to establishments that had not completed the survey.

The questions in the Internet survey inquired about freight and service traffic to and from the establishment. Questions about freight transport included industrial activity taking place at the establishment, whether the establishment belongs to a business chain, storage area (in m^2), importance of real-time delivery, requirements with respect to collection and distribution of freight, the number of freight transporting vehicles attracted (FTA) and/or produced (FTP) in a typical week, the vehicles' origin/destination, type of vehicles, use of carriers and the establishment's possibility to impact the freight transport. Questions regarding service traffic inquired about the type of service taking place at the establishment, number of service vehicles attracted (STA) and produced (STP) in a typical week, geographical area of origin/destination and type of vehicles used in service transport. Filter functions were added to avoid irrelevant questions and reduce the number of questions per respondent.

The data of registered establishments in Groruddalen was retrieved from the Registry of Establishments in Norway [STA 17]. The available registry was updated in March 2015 and included 2,531 establishments with visiting addresses in Groruddalen. Contact information was successfully retrieved for 2,184 establishments who were the recipients of the survey. In total, 221 establishments provided feedback that resulted in a categorization of invalid observations. Typical feedback was change of location, i.e. the establishment ceased to exist or e-mail delivery failure. This resulted in a final sample of 1,963 establishments from whom it could be expected to receive an answer. In total, 385 establishments replied. This is summarized in Table 14.1. The respondents are found representative given aggregated measures of geography, size and industry.

Group	**Observations**
Population	2,531
Gross sample	2,184
Net sample	1,963
Number of respondents	385
Gross response rate	*18 %*
Net response rate	*20 %*

Table 14.1. *An overview of the number of establishments in the population, gross and net sample, retrieved answers and response rates*

Because of privacy concerns, establishments with only one employee had to be removed from the data set. In total, this applied to 14 of 385 respondents. Of these 14, only seven answered that they had either freight or service transport, hence removing one-man firms had a minor impact on the data. Of the 371 observations left, 363 answered at least one question. Not all respondents answered every question, hence there is some missing information in the data.

14.3.1. *Industry classification*

The Registry of Establishments in Norway includes information about each establishment's industry sector at five-digit SN2007 level. This classification system is similar to the North American Industry Classification System (NAICS) and the statistical classification of economic activities in the European Community (NACE)[1]. Using the five-digit disaggregation level directly provides too many industry groups and hence very few or no observations in each group. However, it is used as a starting point for aggregation of the establishments to get as homogenous groups as possible. To get a satisfactory number of establishments in each group, the one-digit NACE classes (highest level of aggregation) were aggregated into 10 groups. This is presented in Table 14.2. Column two presents the one-digit NACE main industry classification in the data, while column one shows the aggregation, resulting in the 10 chosen industry groups. Despite the comprehensive aggregation of industries, some groups still have too few observations to be used in estimation: "Accommodation and eating places", "Information" and "Public services and education" include less than 20 establishments, and "Health care services", "Offices and commercial services" and "Transportation and Warehousing" less than 30 establishments. For this reason, the establishments are grouped into freight-intensive sector (FIS) and non-freight-intensive sector (Non-FIS), inspired by the assumption that sectors where production or consumption of freight is the primary activity are freight-intensive [HOL 16]. "Manufacturing" and "Offices and commercial services" are split between freight-intensive and non-freight-intensive industries.

1 The SN2007 (Norwegian Standard Industrial Classification) builds on rev. 2 of the NACE-classification, except for the lowest disaggregation level, where SN2007 is adapted to fit Norwegian conditions [STA 08, EUR 08].

Aggregation of NACE main industries	NACE main industries	Population	Net sample	Answers	Freight-intensive industry?
Accommodation and eating places	Accommodation and eating places	77	46	4	FIS
Construction and sanitation	Construction	341	265	47	FIS
	Energy supply	1	1		Non-FIS
	Water supply, sewage and waste management	17	15	7	FIS
Health care services	Health care services	163	117	25	Non-FIS
Information	Information	106	80	14	Non-FIS/FIS
Manufacturing	Manufacturing	189	156	39	Non-FIS/FIS
Offices and commercial services	Commercial services	148	115	12	Non-FIS/FIS
	Finance and insurance	17	9	2	Non-FIS
	Managing real estate	75	44	10	Non-FIS/FIS
Other services	Other services	44	31	9	Non-FIS
	Professional, scientific and technical services	194	157	25	Non-FIS
Public services and education	Art, entertainment and recreation	31	24	4	Non-FIS
	Education	26	23	6	Non-FIS
	Public administration, defense and social insurance	16	14	1	Non-FIS
Wholesale and retail trade	Wholesale and retail trade	955	759	144	FIS
Transportation and warehousing	Transportation and warehousing	129	106	22	FIS
Unknown	Unknown	2			
Total		**2 531**	**1 962**	**371**	

Table 14.2. *Industry sectors and frequency distribution of firms*

14.4. Estimating freight trip generation models

Freight trip generation (FTG) often refers to both freight trip production (FTP) and freight trip attraction (FTA) [JAL 15, SÁN 16b, GON 16]. However, it is not clear whether FTA and FTP should be estimated together or separately. Combining freight production and attraction is beneficial as it increases the number of observations in the data set if the data includes firms that either generate or attract freight trips, not both. Moreover, a combined FTG might be useful for planning purposes [GON 16]. Sánchez-Díaz *et al*. [SÁN 16b] argued, on the contrary, that FTA and FTP are driven by different factors, and hence should be seen separately and summarized to get FTG. In addition, Jaller *et al*. [JAL 15] showed that freight trip generation from intermediary firms might differ from the sum of estimated freight trips produced and freight trips attracted, which is often ignored in combined models. Although combining FTA and FTP in one model would let us utilize more of the observations from Groruddalen, running the risk of ignoring that FTA and FTP are driven by different factors advocates separation.

Freight trip generation is mostly estimated as a linear relation between FTG and explanatory variables like employment, industry sector or land use. However, there might be reasons to introduce nonlinearity in freight trip generation models. The first and main reason is that while freight generation (FG) is proportional to business size, freight trip generation (FTG) might also depend on variables like shipment size, vehicle size and transportation costs. Hence, an increase in FG might not result in increased traffic [HOL 11, SÁN 16b]. Such a framework, where freight volume impacts vehicle size and even mode choice, is for instance used in the Norwegian logistics model [MAD 15].

The second reason is that variation in the FTG might differ between business size, introducing heteroskedasticity in the estimation, and third, extreme values in the data set might lead to skewed estimators if not attended to. By taking a logarithmic transformation of the variables, these challenges might be controlled for. Using the data collected in Groruddalen, the relationship between FTG and employment is investigated, showing a positive and nonlinear relationship for all industry groups except "Accommodation and eating places". It is worth noting that the relationship for accommodation and eating places is highly uncertain, as it consists of only four observations, where only three observations generated freight trips. In general, it seems that employment has a higher impact on freight trip generation for freight-intensive industries than non-freight-intensive industries. Similar results are found when looking at FTA and FTP in isolation.

Based on the above argument, there are reasons to estimate separate FTA and FTP models, and investigate the model functional form, which for Groruddalen is expected to be nonlinear. In addition, as most establishments in Groruddalen attract freight trips, the focus is on freight trip attraction models. It is worth emphasizing that the number of attracted freight trips in this analysis is the number of vehicles that visits an establishment per week with the purpose of delivering goods. Reinsborg *et al*. [REI 89] claimed that on average, the vehicle movements between two areas are symmetrical. Hence, a simplified calculation of the total number of vehicle movements generated can be found by multiplying the FTA measure by two.

14.4.1. *FTA model functional form*

Based on FTA and FTG models in the literature, which to a large extent include publicly available and public policy relevant variables, a model for FTA is suggested as follows:

$$\begin{aligned}\mathrm{FTA} = {} & \beta_0 + \beta_1 \mathrm{employment} + \sum \tau_{k-1} \mathrm{industry}_{k-1} + \delta_1 \mathrm{intermediary} \\ & + \delta_2 \mathrm{intermediary} * \mathrm{employment} + \mathrm{u}\end{aligned} \qquad [14.1]$$

Where employment is the establishments' number of employees as provided in the questionnaire. If no answer is provided, the number of employees is retrieved from the Registry of Establishments in Norway. The notation k denotes the 10 different industries presented in Table 14.2, and the variable intermediary is a binary variable equal to 1 if the establishment is intermediary and otherwise. An intermediary establishment both produces and attracts freight trips, and it is expected to generate a different amount of freight trips than establishments that only attract or produce freight trips [JAL 15]. In addition, the model controls for the possibility that intermediary establishments have a different return on employment than non-intermediary establishments. Finally, β, τ and δ are parameters to be estimated. Based on equation (1), a model is estimated for all industry groups and each of the industry groups including more than 30 freight trip attracting establishments. The industry specific models do not include binary variables for industry sector.

The results from estimating equation [14.1] on all industries show that only the transportation and warehousing industry has a significant impact on FTA. However, the validity of this result is uncertain, as several of the industry groups have very few observations. One way to still control for differences between industries is to include a binary variable controlling for whether an establishment is freight-intensive (FIS) or not. Hence, a binary variable denoting freight-intensive industry, an interaction term with employment and a binary variable for transportation and warehousing industry is included instead. The revised model for all industry groups is presented in equation [14.2]. The industry specific models are still given by equation [14.1] without binary variables for industry sector.

$$\mathrm{FTA} = \beta_0 + \beta_1 \mathrm{employment} + \delta_1 \mathrm{intermediary} + \delta_2 \mathrm{intermediary} * \mathrm{employment} + \delta_3 \mathrm{FIS} + \delta_4 \mathrm{FIS} * \mathrm{employment} + \delta_5\, \mathrm{industry}_{\mathrm{TW}} + u \quad [14.2]$$

All variables included in model [14.1] are expected to have a positive impact on FTA: the establishment's number of employees is thought to reflect firm size and hence increases FTA, both existing literature [JAL 15] and the observations in the data show that intermediary firms attract more freight than other firms, and by definition, freight-intensive sectors are expected to have a higher FTA than non-freight-intensive sectors. At last, an employee in intermediary and freight intensive firms is expected to attract more freight than others, all else being equal.

14.4.1.1. *Estimation results*

In total, 271 of 363 establishments in the data from Groruddalen stated that their business attracts freight transport. These are included in the analysis. As argued above, special interest is given to the functional form of the relationship between FTA and employment. At the simple level, there are four different functional forms given by whether the independent and the dependent variables are linear or logarithmic. When plotted against each other, the relationship between FTA and employment in Groruddalen seems to be nonlinear and the presence of extreme values advocates a model where both FTA and employment are logarithmic (log–log). However, models treating both variables as linear (lin–lin) are most common in the literature and should be investigated. A natural starting point is therefore to compare the model fit of a linear (lin–lin) model with a non-linear (log–log) model.

The lin–lin and log–log models were estimated using ordinary least square (OLS), and compared and tested against the necessary assumptions for the OLS estimators to be the best unbiased linear estimator [WOO 06]. This shows that the lin–lin model suffers from heteroskedastic error terms, while the log–log model for all industries does not. This is confirmed by a Breusch–Pagan test for heteroskedasticity [WOO 06]. For the industry specific models, both the lin–lin and the log–log model specifications have heteroskedastic error terms.

Heteroskedasticity is a problem in model estimation, as it biases confidence intervals and invalidates hypothesis testing with standard methods [WOO 06]. However, it can easily be corrected for using robust standard errors. A larger problem might occur if the zero-conditional mean assumption is violated, as this might lead to estimation bias. An analysis of the residuals from both model specifications shows that the log–log model also outperforms the lin–lin model in terms of model fit. Hence, there are reasons to believe that estimators from the log–log model are less biased than those from the lin–lin model.

The results hold for all models, although they are more evident for the model including all industries. For this model, an interesting finding is that the lin–lin model provides a constant FTA for establishments that are neither freight intensive nor intermediary. Such results are often seen in the literature. The models are also tested for a lin–log and log–lin functional form. A lin–log model (regression of FTA on ln(employment)) shows a good fit to the data, however worse than the log–log model. The reason seems to be the extreme values, which are not well accommodated in the lin–log model. Since Groruddalen includes a mixture of firms of different sizes, the chosen model should be able to explain some of this variation. Hence, log–log models are preferred above lin–log models. The log–lin functional form (regression of ln(FTA) on employment) did not fit the data and was abandoned.

Estimation results are presented in Table 14.3. Model (I) includes all establishments that attract freight trips, irrespective of industry. Models (II)–(IV) are industry specific and include establishments that attract freight trips within the industries "Construction and Sanitation", "Manufacturing" and "Retail" respectively. All the models are estimated using a logarithmic transformation of both freight trip attraction (dependent variable) and employment (independent variable) using OLS with robust standard errors. For the industry group "Manufacturing", the data shows that 29 of 32 manufacturing firms are intermediate. Hence, this variable is excluded from Model (III).

The results in Table 14.3 show that employment has a positive and significant impact on freight trip attraction in all models. This is as expected. The impact is largest for construction and sanitation industries, where the elasticity[2] is almost 0.9. Manufacturing and retail has an elasticity between 0.6 and 0.7, while the general model including all industries has the lowest value of almost 0.3. All models except Model (III) "Manufacturing" include variables denoting whether the establishment is intermediary or not, along with an interaction term between this variable and employment. The variable for the intermediary firm is only significant for Model (II) "Construction and sanitation". Model (I) "All industries" includes binary variables for transportation and warehousing industries and freight-intensive sectors. The estimation results show that the variable for transport and warehousing industries is significant and positive.

---

2 When both the dependent and the independent variables are in logarithmic form, the parameter value is interpreted as elasticity, meaning that an increase in the independent variable leads to a β-% change in the dependent variable [WOO 06].

	(I)	(II)	(III)	(IV)
	ln(FTA) All industries	ln(FTA) Construction and sanitation	ln(FTA) Manufacturing	ln(FTA) Retail
Ln(employment)	0.291**	0.877**	0.638***	0.693*
	(0.110)	(0.248)	(0.148)	(0.286)
Intermediary firm	0.223	2.438*		0.806
	(0.322)	(0.992)		(0.535)
Intermediary firm × ln(employment)	0.170	–0.589		–0.255
	(0.123)	(0.319)		(0.295)
Freight-intensive sector	0.466			
	(0.362)			
Freight-intensive sector × ln(employment)	0.0220			
	(0.118)			
Transportation and warehousing industry	1.850***			
	(0.335)			
Constant	1.091**	–0.285	1.548***	1.032*
	(0.342)	(0.662)	(0.347)	(0.486)
R^2	0.393	0.350	0.529	0.278
adj. R^2	0.379	0.293	0.514	0.259
F	22.67	7.693	18.59	21.86
N	271	38	32	120

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 14.3. *Result from estimating freight trip attraction using a logarithmic transformation of freight trip attraction and employment*

The low significance of the other variables might be explained by correlation between the explanatory variables and the interaction terms with employment. For Model (I), an F-test shows that a model consisting of the binary variable for the transportation and warehousing industry as well as the interaction terms between intermediary firms and employment and freight-intensive sector and employment is an acceptable restriction. Such a restriction could not be done for the other models. The explanatory power of the models (measured by R^2 and adjusted R^2) is relatively low, ranging from 0.250 to 0.530. However, this is common for freight trip models (in general).

14.4.2. *Model extension with establishment and shipment characteristics*

The second part of this chapter analyzes an extension of the freight trip attraction models presented above. The extension is based on the hypothesis that there are factors other than establishment size and industry that explain freight trip generation. This hypothesis is partly motivated by the low explanatory power of general freight trip generation models, like the one presented above, and partly by the data collected from establishments in Groruddalen.

The data collected in Groruddalen includes information about the establishments that are not necessarily publicly available. Examples are the establishment's use of logistics provider or degree of time sensitive shipments. The complete list of information included in the data collected from establishments in Groruddalen is presented in the third section of this chapter. Given that the relationship between FTA and employment appears in logarithmic form, as shown above, the extended model is suggested as follows (for the model including all industries, the binary variables for the industry group are replaced by a binary variable denoting freight-intensive industry, an interaction term with employment and a binary variable for the transportation and warehousing industry):

$$\begin{aligned}
&\ln(FTA) = \\
&\beta_0 + \beta_1 \ln(employment) + \delta_1 intermediary + \delta_2 intermediary * \\
&\ln(employment) + \delta_3 FIS + \delta_4 FIS * \ln(employment) + \\
&\delta_5 Industry_{TW} + \delta_6 Carrier_{Yes} + \delta_7 Carrier_{No} + \delta_8 Time_{Yes} + \\
&\delta_9 Time_{No} + \delta_{10} Decision_{High} + \delta_{11} Decision_{Low} + u
\end{aligned} \qquad [14.3]$$

The first six variables are the same as presented in equation [14.2]. Variables seven and eight denote whether the firm uses a carrier for all its shipments or none at all. Variable nine denotes whether more than 75% of the firm's shipments are time sensitive, while variable ten says the opposite, i.e. less than 25% of the firm's shipments are time sensitive. Variable eleven and twelve represent the firm's degree of decision making. If the degree of decision making is high, the firm can choose when and how the transport arrives at the establishment. On the contrary, a low degree of decision making means that another party makes the decisions. All variables except employment are discrete. The six discrete variables included in the model extension can take on more than two values (including the answers “don’t know” and “no answer”).

14.4.2.1. *Estimation results*

The model presented in equation [14.3] is estimated for all industries and each of the three industries “Construction and sanitation”, “Manufacturing” and “Retail”. All models are estimated using ordinary least squares (OLS) with robust standard errors. For the model including all industries, a restricted version with the binary variable for the transportation and warehousing industry and interaction terms is used.

As before, the variable for the industry and freight-intensive sector is not included in the industry specific models. Because of few observations, the variable $\text{Carrier}_{\text{No}}$ is excluded from all industry specific models, Time_{Yes} is excluded from the construction and sanitation model, and $\text{Decision}_{\text{High}}$ and variables for intermediary firms are excluded from the model for manufacturing industries. The estimation results are presented in Table 14.4. Model (V) includes all industries, while Models (VI)–(VIII) are industry specific models.

	(V)	(VI)	(VII)	(VIII)
	ln(FTA) All industries	ln(FTA) Construction and sanitation	ln(FTA) Manufacturing	ln(FTA) Retail
Ln(employment)	0.172*	0.897**	0.537**	0.713**
	(0.0728)	(0.262)	(0.188)	(0.263)
Intermediary firm		2.333*		0.876
		(1.043)		(0.525)
Intermediary firm × ln(employment)	0.195**	–0.616		–0.294
	(0.0639)	(0.334)		(0.270)
Freight-intensive sector × ln(employment)	0.120*			
	(0.0549)			
Firm use carrier for all shipments	–0.508***	–0.644	–0.547	–0.317
	(0.151)	(0.522)	(0.280)	(0.199)

Firm do not use carrier	–0.413			
	(0.513)			
Above 75% time-sensitive shipments	0.483		0.292	0.460
	(0.256)		(0.435)	(0.311)
Below 25% time-sensitive shipments	–0.280	0.0920	–0.00267	–0.265
	(0.164)	(0.462)	(0.343)	(0.215)
High degree of decision-making	0.523*	0.922		0.721*
	(0.236)	(0.545)		(0.306)
Low degree of decision-making	0.231	0.440	–0.464	0.481*
	(0.152)	(0.490)	(0.276)	(0.200)
Transportation and warehousing industry	1.746***			
	(0.335)			
Constant	1.708***	–0.641	2.075***	0.820
	(0.202)	(0.740)	(0.426)	(0.535)
R^2	0.436	0.408	0.577	0.365
adj. R^2	0.415	0.270	0.496	0.319
F	18.49	3.621	15.59	9.257
N	271	38	32	120

Standard errors in parentheses
$^{*}\ p < 0.05$, $^{**}\ p < 0.01$, $^{***}\ p < 0.001$

Table 14.4. *Estimation result from estimating FTA on establishment and shipment characteristics in addition to public policy oriented attributes*

A look at the estimation results in Table 14.4 shows that, where comparable, the results resemble the findings presented in Table 14.3: there is a positive relationship between the number of employees and FTA and transportation and warehousing firms, intermediary firms and freight-intensive industries tend to have a higher FTA than other establishments, all else being equal. A look at the establishment and shipment specific variables shows that use of a carrier for all shipments is negative in all models, and significant in Model (V). Hence, establishments that use a carrier for all their shipments tend to have fewer shipments than those who do not use a carrier for all their shipments. However, the variable denoting no use of carrier is also included in Model (V) and has a negative impact on ln(FTA), although not significant.

This might reflect that a mix of own- and third-party transport increases freight trip attraction more than choosing only one of the transport solutions. More research on this topic would be interesting. The binary variable representing a high degree of decision making is positive in Models (V), (VI) and (VIII), but only significant in Models (V) and (VIII). The variable denoting a low degree of decision making is also positive, although with a smaller impact than a high degree of decision making and only significant in Model (VIII). The variables representing time sensitivity are all insignificant; however, the parameter signs show that a high degree of time sensitivity increases ln(FTA), while a low degree of time sensitivity indicates a (close to) zero or negative impact.

The explanatory power of the models, represented by the ordinary R^2, lies between 0.400 and 0.600, and the adjusted R^2, controlling for inclusion of irrelevant variables, lies between 0.270 and 0.500. If these results are compared with the model fit from the simpler model, presented in Table 14.3, we can see that all models have an increase in the R^2, but only two out of four got an increase in adjusted R^2. Hence, the model estimated for all industries and for retail benefits from more variables, while the model for construction and sanitation and manufacturing does not. This confirms that different industries require different models to explain freight trip generation, and hence the need for industry specific freight trip models. The results also show that some industries might require additional information as well as business size to explain their freight trip generation.

14.5. Conclusion

This chapter presented a freight trip generation model estimated using Norwegian establishment data. The analysis has taken on an explorative approach in a European context, estimated freight trip attraction (FTA) models for a non-negligible urban area that is not a city center and explored the relevance of different explicative variables in the model. The data used was collected through a self-completion Internet survey distributed to establishments in Groruddalen, an urban area in the Norwegian capital Oslo. Two model specifications have been tested: one controlling for business size through the establishment's number of employees, industry group and whether the establishment is intermediary or not, as well as a model extension including establishment and shipment specific explanatory variables. Both model specifications were estimated for all establishments attracting freight trips, as well as manufacturing establishments, construction and sanitation establishments and retail establishments, respectively.

One of the main results from the model estimation is that the relationship between freight trip attraction and employment in Groruddalen is nonlinear, where a model including a logarithmic transformation of both freight trip attraction and employment (log–log) provides the best fit. A linear relationship between FTA and employment provides a constant FTA for certain establishments, as frequently found in the literature, but this model suffers from heteroskedastic error terms and violates the zero-conditional mean assumption. Another contribution is that establishment and shipment characteristics are found relevant to explain freight trip attraction in a model including all industries and for retail establishments only.

Despite interesting findings, there are some weaknesses that should be highlighted. First, the models presented in this chapter are estimated on data retrieved from a sample of establishments that is fairly representative of its population, measured in terms of business size (number of employees), geographical area and industry sector. However, some industry classification groups suffer from very few respondents. This is mainly a problem for accommodation and eating places and public services and education with 4 and 11 respondents, respectively. In addition, industry groups including information, health care service, offices and commercial services as well as transportation and warehousing have less than 30 respondents.

The lack of establishments in certain industry classification groups might be a reason for their low significance. In addition, the industry group consisting of retail establishments includes more than 100 establishments, and hence dominates the sample. Such problems might be a consequence of data collected through a self-completion Internet survey. Even though a self-completion Internet survey allows the researcher to reach many respondents at low cost and little time use, it is at the expense of low response rate and control over the answers provided [ALL 12]. More targeted sampling, for instance stratified sampling by industry groups, might be preferred to allow more research on industry specific freight trip attraction models.

Another weakness is related to the grouping of industries. A random sample, as with the one retrieved from Groruddalen, limits the numbers of industry groups available for estimation. This might lead to groups that are in fact too heterogeneous for a FTA model. Moreover, relying on industry classification when grouping the establishments requires assumptions about industry homogeneity within freight trip attraction. This is not necessarily the case.

An alternative way of grouping establishments for freight trip attraction models is by commodity type, as in the Norwegian logistics model [DEJ 07, DEJ 13]. This grouping relies on commodity characteristic similarity instead of industry, differentiating between transportation requirements like shipment size, frequency of delivery and available mode choices. The Norwegian logistics model currently uses 39 commodity groups [HOV 15], which is too disaggregated for the dataset in question. However, if more data are collected this might be a subject for further research. Despite some challenges, the research presented in this chapter provides some interesting findings which illustrate the usefulness of taking a more explorative approach when estimating freight trip generation in European cities, as well as areas for further research.

14.6. Bibliography

[ADJ 16] ADITJANDRA P.T., GALATIOTO F., BELL M.C. *et al.*, "Evaluating the impacts of urban freight traffic: application of micro-simulation at a large establishment", *European Journal of Transport & Infrastructure Research*, vol. 16, no. 1, pp. 4–22, 2016.

[ALH 14] ALHO A., SILVA J.D.A.E., "Freight-trip generation model: predicting Urban Freight weekly parking demand from retail establishment characteristics", *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2411, pp. 45–54, 2014.

[ALL 12] ALLEN J., BROWNE M., CHERRETT T., "Survey techniques in Urban Freight transport studies", *Transport Reviews*, vol. 32, no. 3, pp. 287–311, 2012.

[BAS 09] BASTIDA C., HOLGUÍN-VERAS J., "Freight generation models", *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2097, pp. 51–61, 2009.

[BEN 16] BEN-AKIVA M.E., TOLEDO T., SANTOS J. *et al.*, "Freight data collection using GPS and web-based surveys: Insights from US truck drivers' survey and perspectives for urban freight", *Case Studies on Transport Policy*, vol. 4, no. 1, pp. 38–44, 2016.

[DEJ 07] DE JONG G., BEN-AKIVA M., "A micro-simulation model of shipment size and transport chain choice", *Transportation Research Part B: Methodological*, vol. 41, no. 9, pp. 950–965, 2007.

[DEJ 13] DE JONG G., BEN-AKIVA M., JAAP B. *et al.*, Method Report - Logistics Model in the Norwegian National Freight Model System (Version 3), Significance, Project 12028, The Netherlands, 2013.

[EUR 08] EUROEPAN COMMISSION, Nace Rev. 2 - Statistical classification of economic activities in the European Community [Online], accessed 12 January 2017, available at: http://ec.europa.eu/eurostat/documents/3859598/5902521/KS-RA-07-015-EN.PDF/dd5443f5-b886-40e4-920d-9df03590ff91?version=1.0, 2008.

[GON 14] GONZÁLEZ-FELIU J., CEDILLO-CAMPO M.G., GARCÍA-ALCARAZ J. L., "An emission model as an alternative to O-D matrix in urban goods transport modelling", *Dyna*, vol. 81, no. 187, pp. 249–256, 2014.

[GON 16] GONZÁLEZ-FELIU J., SÁNCHEZ-DÍAZ I., AMBROSINI C., "Aggregation level, variability and linear hypotheses for urban delivery generation models", *Transportation Research Board 95th Annual Meeting Compendium of Proceedings*, TRB, Washington, 10–14 January 2016.

[HOL 11] HOLGUÍN-VERAS J., JALLER M., DESTRO, L. *et al.*, "Freight generation, freight trip generation, and the Perils of using Constant Trip Rates", *Transportation Research Record*, vol. 2224, pp. 68–81, 2011.

[HOL 13] HOLGUÍN-VERAS J., SÁNCHEZ-DÍAZ I., LAWSON C.T. *et al.*, "Transferability of freight trip generation models", *Transport Research Record*, vol. 2379, pp. 1–8, 2013.

[HOL 16] HOLGUÍN-VERAS J., SÁNCHEZ-DÍAZ I., BROWNE M., "Sustainable Urban Freight systems and freight demand management", *Transportation Research Procedia*, vol. 12, pp. 40–52, 2016.

[HOL 17] HOLGUÍN-VERAS J., LAWSON C., WANG C. *et al.*, *Using Commodity Flow, Survey Microdata, and Other Establishment Data to Estimate the Generation of Freight, Freight Trips, and Service Trips: Guidebook*, National Academies of Science, Engineering, and Medicine, 2017.

[HOV 15] HOVI I.B., CASPERSEN E., GRUE B., Commodity flow matrices in Norway, TØI report 1399, Institute of Transport Economics, Oslo, Norway, 2015.

[IDI 02] IDING M.H.E., MEESTER W.J., TAVASSZY L.A., "Freight trip generation by firms", *42nd European Congress of the Regional Science Association*, Dortmund, Germany, 2002.

[JAL 15] JALLER M., SÁNCHEZ-DÍAZ I., HOLGUÍN-VERAS J., "Identifying freight intermediaries". *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2478, pp. 48–56, 2015.

[LAW 12] LAWSON C., HOLGUÍN-VERAS J., SÁNCHEZ-DÍAZ I. *et al.*, "Estimated generation of freight trips based on land use", *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2269, pp. 65–72, 2012.

[MAD 15] MADSLIEN A., STEINSLAND C., GRØNLAND S.E., The National Norwegian freight transport model: How to use the model, TØI report 1429, Institute of Transport Economics, Oslo, Norway, 2015.

[REI 89] REINSBORG J., KAURIN T., SANDELIEN S. *et al.*, Trafikkberegninger, Vegdirektoratet, 1989.

[SÁN 16a] SÁNCHEZ-DÍAZ I., "Modeling urban freight generation: A study of commercial establishments' freight needs", *Transportation Research Part A: Policy and Practice*, vol. 102, pp. 3–17, 2016.

[SÁN 16b] SÁNCHEZ-DÍAZ I., HOLGUÍN-VERAS J., WANG X., "An exploratory analysis of spatial effects on freight trip attraction", *Transportation*, vol. 43, no. 1, pp. 177–196, 2016.

[STA 08] STATISTICS NORWAY, Næringsstandard og næringskoder, available at: http://www.ssb.no/virksomheter-foretak-og-regnskap/naeringsstandard-og-naeringskoder, accessed 12 January 2017, 2008.

[STA 17] STATISTICS NORWAY, About the statistics, available at: https://www.ssb.no/en/virksomheter-foretak-og-regnskap/statistikker/bedrifter/aar/2016-01-22#content, accessed 11 January 2017.

[TAD 94] TADI R.R., BALBACH P., "Truck trip generation characteristics of nonresidential land uses", *ITE Journal*, vol. 64, no. 7, pp. 43–47, 1994.

[TVE 16] TVEDT K.A., ASKHEIM S., "Groruddalen", *Store norske leksikon*, available at: https://snl.no/Groruddalen, accessed 9 January 2017, 2016.

[WOO 06] WOOLDRIDGE J.M., *Introductory Econometrics: A Modern Approach*, Thomson South-Western, Mason, 2006.

Chapter 3:

Latent split of aggregate counts: revealing home deliveries per commodity types and potential freight trip implications

Elise Caspersen

Mario Enrique Arrieta-Prieto

Xiaokun (Cara) Wang

Latent split of aggregate counts: revealing home deliveries per commodity types and potential freight trip implications

Elise Caspersen[1,4,*], Mario Arrieta-Prieto[2], Xiaokun (Cara) Wang [3]

[1]Institute of Transport Economics – Norwegian Centre for Transport Research, Gaustadalléen 21, 0349 Oslo, Norway
[2] Department of Industrial and Systems Engineering. Rensselaer Polytechnic Institute, 110 8th St., Troy, NY 12180, USA
[3] Department of Civil and Environmental Engineering. Rensselaer Polytechnic Institute, 110 8th St., Troy, NY 12180, USA
[4] School of Economics and Business; Norwegian University of Life Sciences, Ås, Norway.
[*]Corresponding author at: Institute of Transport Economics – Norwegian Centre for Transport Research, Gaustadalléen 21, 0349 Oslo, Norway. Tel.: + 922 444 52, E-mail address: elc@toi.no

Abstract: This paper suggests a joint econometric model that allows estimating latent marginal counts when only total counts and types of commodities purchased are available. The basis for this model is the Negative binomial Hurdle model. A validation procedure for the proposed splitting is discussed; followed by remedial mechanisms to adjust the estimated marginal counts when necessary.

The proposed methodology was used to estimate and validate a model for the propensity to shop online and the corresponding number of shipments per commodity group. The results confirm existing research on online shopping behavior: elderly is less likely to buy online, while high income, education and having kids motivate online shopping. Per-commodity-group analysis of shipments shows that clothes, beauty and interior products have the highest relative frequency followed by children and leisure goods. The average online shopper receives 2.4 shipments per month, or 0.077 shipments per day, with variations depending on the consumer profile. The correlation between the orders of different commodity groups is explored, revealing great consolidation opportunities among the groups, and potential reductions of up to 30% of shipments.

Keywords: e-commerce, freight trip generation, hurdle model, latent count estimation, hypothesis testing, consolidation opportunities.

1 Introduction

Online shopping is growing extensively, with European e-commerce forecasted to hit €717 billion in 2020 (Ecommerce News, 2020). An early assumption was that e-commerce would crowd out shopping trips and as such, reduce total traffic and its burden on society; but empirical evidence contradicts this (Zhou and Wang, 2014, Cao et al., 2012, Pettersson et al., 2018). Instead, online shoppers are increasingly sophisticated consumers, pushing for faster and cheaper deliveries, including new delivery services like "instant" deliveries (Dablanc et al., 2017). As such, duplication of delivery services with low vehicle utilization in terms of payload has been observed (Allen et al., 2018, Ni et al., 2016), and consolidation is challenged as competitors fight for market shares (Allen et al., 2018, Pettersson et al., 2018). The contrast to traditional freight and retail transportation – where large volumes are transported between warehouses or retail stores (Visser et al., 2014) – is huge, defying existing infrastructure as well as political goals to redesign urban areas in favor of walking, cycling and public transportation (Allen et al., 2018).

A combination of high-level delivery services and receivers located in residential units makes online shopping one of the main contributors to growing freight traffic in residential areas (Visser et al., 2014, Allen et al., 2018). Although a vast research on drivers of online shopping exists (Farag et al., 2007, Cao et al., 2012, Zhou and Wang, 2014), there is relatively little knowledge on the relationship between consumers' online shopping activity, number of shipments and freight traffic, despite the magnitude of freight trips generated by residential units being no longer trivial (Wang and Zhou, 2015, Holguín-Veras et al., 2018, Saphores and Xu, 2020). In later years, online shopping has expanded from commodities with high searchability and ease of making a quality judgement prior to experience, identified as attractive commodities for online shopping (de Figueiredo, 2000; Girard, Silverblatt, & Korgaonkar, 2002; Lowengart & Tractinsky, 2001); to clothes or even groceries, with particularly strict delivery requirements and difficulties for quality assessments online (Lagorio and Pinto, 2020, Saphores and Xu, 2020). This results in heterogeneous online purchases with different delivery requirements, lower degree of consolidation and a higher degree of returns, resulting in increased freight traffic.

Hence, more knowledge on shipments of different types of commodities from online shopping is necessary, motivating the following research questions addressed in this paper: i) *which factors explain the probability to shop online and the corresponding number of shipments received, separated on commodity groups*; and ii) *what are the freight trip generation implications of this new knowledge*? To accommodate for the potential tedious and time-consuming task it is to collect individual data about number of shipments per commodity group, this paper builds on the methodology by Afghari et al. (2016), Afghari et al. (2018) to estimate latent marginal counts when only total counts and types of commodities purchased are available. The estimation results reveal shopping patterns that suggest potential for increased consolidation between commodities to the same consumer or neighborhood. The latter is further explored, revealing potential of reducing number of shipments up to 30%.

To the best of our knowledge, this is the first paper that develops a complete statistical framework to infer latent counts and correct them based on the evidence provided by the data. It is also the first paper that uses information about total number of shipments, indicators for commodity type purchase and joint econometric modeling to estimate number of shipments per aggregated commodity type from online shopping. This work is also pioneer in identifying the impact of consolidation strategies based on correlation analysis of the different consumption profiles of interest.

The rest of the paper is organized as follows: Section 2 introduces the proposed methodological development, which is tested and validated in Section 3. The model is applied to the total reported number of shipments from online shopping and indicator variables for commodity groups purchased to estimate the number of shipments per commodity group from online shopping. The case study is described in Section 4, followed by estimation results and implications in Section 5 and 6. A brief conclusion is provided in Section 7.

2 Hurdle model and latent marginal count estimation

The proposed model attempts to describe the driving factors related with the decision of buying online or not and, in the case of online shoppers, which characteristics influence the amount of shipments received. To deal with excess of zeros from no shopping, zero-inflated or hurdle models are commonly used alternatives, with the main difference being how they treat the process of obtaining the zeros (Mullahy, 1986). The dataset at hand considers a priori only one source of zeros: no online shopping. However, this might result from both individuals who never shop online and individuals who usually shop online, just not that particular time period. As the month in question is related to Christmas shopping (see Chapter 4 for a description of the data), it is likely that regular online shoppers did shop, justifying that zeros occurred from only one source. Additionally, the hurdle model, unlike the zero-inflated model, estimates the splitting between zero and non-zero counts and the positive count data model in two stages. This allows for inclusion of variables that are only available for shoppers, of which are in abundance in the dataset. Thus, the authors chose to proceed with the hurdle model, despite that both Wang and Zhou (2015) and Saphores and Xu (2020) used the Zero-inflated model for similar purposes.

The Hurdle model firstly characterizes the probability of buying online by means of a binary-outcome model (usually logit), and then, incorporates a count data model to explain the nonzero counts observed for buyers (Washington, Karlaftis, & Mannering, 2003). In the context of online shopping, it is of interest to identify the count of shipments received per type of commodity and which socio-economic variables influence that marginal count. Often, as it is the case with the data considered in this paper, the amount of shipments per commodity type is not collected in the surveys, because it might induce fatigue due to the broad number of commodity types considered. It is also questionable whether respondents can remember their purchases in such a detailed level. Hence, latent marginal-count estimation procedures are necessary to elicit the split for different commodity types given only the information about total counts and, in the case of this research, records of the types of commodities purchased.

Let Y_i represent the total number of shipments received by individual i provided that he/she is an online shopper ($Y_i > 0$). Let Y_{ij} represent the (latent) number of shipments received from commodity type j. It is clear, then, that if m commodity types are considered

$$Y_{\mathrm{i}} = \sum_{j=1}^{m} Y_{\mathrm{ij}}, \ \forall i. \tag{1}$$

If $\lambda^{(\mathrm{i})}$ represents the mean of total shipments for individual i and $\lambda_{\mathrm{j}}^{(\mathrm{i})}$ the mean of shipments from commodity type j for that same individual, it holds that

$$\lambda^{(\mathrm{i})} = \sum_{j=1}^{m} \lambda_{\mathrm{j}}^{(\mathrm{i})}. \tag{2}$$

Y_{i} is usually modeled as a Negative Binomial distribution of parameters $\lambda^{(\mathrm{i})}$ and *α,* with *α* accounting for heterogeneity and overdispersion. With a similar framework, Afghari et al propose a methodology for estimation of latent marginal counts in the context of road safety (Afghari et al., 2016, Afghari et al., 2018). In their papers, each latent count mean, $\lambda_{\mathrm{j}}^{(\mathrm{i})}$, is modeled in terms of a vector of group-specific covariates, X_{ij}, and a vector of group-specific parameters, β_{j}, such that

$$\lambda_{\mathrm{j}}^{(\mathrm{i})} = \exp\left(\beta_{\mathrm{j}}^{T} X_{\mathrm{ij}}\right). \tag{3}$$

After substituting **(3)** in **(2)**, the resulting expression for $\lambda^{(\mathrm{i})}$ is replaced for each individual in the likelihood function for parameter estimation.

Afghari et al's methodology also assumes that all j subgroups have a non-negative count (Afghari et al., 2016, Afghari et al., 2018). The dataset in this work contains information of commodity types purchased, and, although a maximum number of commodity types *m* is defined, this model allows null latent counts for a given buyer as a result of the reasonable statement that not all commodity types are always purchased by the same individual. The proposed model's goal is to identify the major drivers that affect the total mean count, $\lambda^{(\mathrm{i})}$, while allowing the impact of the covariates to vary with the commodities the user reported to have purchased.

Following the principles of a count data model, the total mean count is characterized as

$$\ln\left(\lambda^{(\mathrm{i})}\right) = I_{1}^{(\mathrm{i})} \cdot \left(\sum_{k=1}^{K_1} \beta_{\mathrm{k1}} x_{\mathrm{k1}}^{(\mathrm{i})}\right) + I_{2}^{(\mathrm{i})} \cdot \left(\sum_{k=1}^{K_2} \beta_{\mathrm{k2}} x_{\mathrm{k2}}^{(\mathrm{i})}\right) + \ldots + I_{\mathrm{J}}^{(\mathrm{i})} \cdot \left(\sum_{k=1}^{K_J} \beta_{\mathrm{kJ}} x_{\mathrm{kJ}}^{(\mathrm{i})}\right). \tag{4}$$

The binary variable $I_{\mathrm{j}}^{(\mathrm{i})} \in \{0,1\}$ indicates if the *i*-th individual buys from the *j*-th commodity type or not. The terms in parenthesis are the linear contributions of covariates associated with the latent count of commodity type $j \in [J] := \{1,2,\ldots,J\}$, with $x_{\mathrm{kj}}^{(\mathrm{i})}$ being the *k*-th covariate considered to explain the count of the *j*-th commodity type for the *i*-th individual and β_{kj} its corresponding parameter. As such, the model consists of interactions between the explanatory and the indicator variables for each commodity type.

There are several reasons to favor this model structure over a more conventional one with only first-order general terms when the main purpose is to characterize the latent counts:

1. Based on the objective of identifying the driving factors of the number of purchases of J commodity types, this model structure allows to isolate and directly obtain the marginal effects of the covariates for a specific commodity *j*. For instance, if "*age*" is included in the covariates explaining the counts for commodity $u \in [J]$ and $t \in [J] \setminus \{u\}$, the corresponding parameter associated with each type will explicitly show the marginal effect of "*age*" in the purchase of that specific group.
2. The candidate variables to explain the count for a commodity type can be determined independently from those considered for a different type. Continuing with the example above, this implies that the marginal effect of "*age*" in type *u* is independent of age being present in any of the other types, i.e, *t*.
3. Another approach to capture the purchase behavior of the individuals in the sample could be based on developing an explanatory model with different covariates for each one of the different purchase profiles included in the sample. If J commodity types are included in the analysis, there are 2^J purchase profiles depending on the types of commodities bought by the user (a purchase profile is intended as a record of J "Yes/No" answers to the question if *j*-th commodity type was purchased or not). However, even for a moderate-size number of commodities (e.g., $J = 5$), the number of purchase profiles grows exponentially (e.g., $2^5 = 32$), adding an unnecessary computational burden to the estimation routines.

The parameters of the model are estimated via maximum likelihood under the assumption of negative-binomial-distributed outcomes for each one of the buyers. Once the estimation routine is finalized, the following estimate for the latent marginal counts is proposed

$$\hat{\lambda}_{\mathrm{j}}^{(\mathrm{i})} = \begin{cases} \exp\left(\sum_{k=1}^{K_j} \hat{\beta}_{\mathrm{kj}} x_{\mathrm{kj}}^{(\mathrm{i})} \right), \text{ if } I_{\mathrm{j}}^{(\mathrm{i})} = 1. \\ 0, \text{ if } I_{\mathrm{j}}^{(\mathrm{i})} = 0. \end{cases} \tag{5}$$

A methodology to test the validity of these latent count estimates is presented in Section 3. Via a statistical test of hypothesis, it is possible to verify if the data provide evidence towards favoring the proposed count splitting procedure. The next section also presents a way to adjust the latent-count estimators for each commodity type if the statistical evidence suggests that some assumptions are not met; but at the same time, the deviations are not significantly large.

3 Hypothesis testing for validation of the latent-count estimation

3.1 Procedure for hypothesis testing

It is expected that the mean latent counts satisfy the assumption that their summation equals the mean total count, i.e., $\lambda^{(i)} = \lambda_1^{(i)} + \lambda_2^{(i)} + \cdots + \lambda_J^{(i)}$, as introduced in **(2)**. However, the proposed construction of the latent-count estimators provides no guarantee of that. To test this assumption, an indicator of the magnitude of discrepancy from the assumption can be defined as

$$\Delta^{(i)} := \ln\left(\frac{\sum_{j \in J} \lambda_j^{(i)}}{\lambda^{(i)}}\right). \tag{6}$$

The natural logarithm is conveniently introduced so that $\Delta^{(i)}$ can take any value on the real line. A negative value would imply that the summation of the mean latent counts underestimates the mean total count, while a positive value would be an indicator of the opposite.

If the latent-count estimators proposed are reasonable, one would expect that $\Delta^{(i)} = 0$. Therefore, the following hypothesis system is of interest

$$\begin{cases} H_0 : \Delta^{(i)} = 0 \\ H_1 : \Delta^{(i)} \neq 0 \end{cases}. \tag{7}$$

Using the estimates for $\hat{\lambda}_j^{(i)}$, it is possible to compute an estimate of $\Delta^{(i)}$ ($\hat{\Delta}^{(i)}$) for each e-online shopper in the sample. Due to sample variability, measurement errors and unobserved effects; these estimates will not be exactly equal to zero even if all the assumptions are met, motivating the use of statistical tools to assess the hypothesis system presented. It must also be noted that the information of online shoppers who bought only from one commodity type should not be considered in the analysis since no splitting was necessary. These individuals have a $\hat{\Delta}$ value equal to zero by definition, but those zeroes do not provide evidence in favor or against the splitting procedure proposed and must be discarded.

Assuming that $\tilde{N}$ individuals purchased shipments from more than one commodity type, the vector of Δ-estimates, $\vec{\Lambda} := \left(\hat{\Delta}^{(1)}, \hat{\Delta}^{(2)}, \ldots, \hat{\Delta}^{(i)}, \ldots, \hat{\Delta}^{(\tilde{N})}\right)$, can be used to assess the hypothesis system in **(7)**. However, it cannot be assumed that these realizations are independent from each other; although they come from different individuals, because they are all obtained using the same set of estimated parameters, which induces a co-dependence structure between them. After evaluating expression **(6)** with the proposed estimators, it is obtained that

$$\hat{\Delta}^{(i)} := k\left(\hat{\beta}, X_i\right) := \ln\left(\frac{\hat{\lambda}_1^{(i)}\left(\hat{\beta}, X_i\right) + \hat{\lambda}_2^{(i)}\left(\hat{\beta}, X_i\right) + \ldots + \hat{\lambda}_J^{(i)}\left(\hat{\beta}, X_i\right)}{\hat{\lambda}^{(i)}\left(\hat{\beta}, X_i\right)}\right), \tag{8}$$

where $k(\cdot,\cdot)$ is the function that relates the parameter estimates , $\hat{\beta}$, and the vector of covariates, X_{i}, for individual i; with the delta estimate $\hat{\Delta}^{(\mathrm{i})}$. The $\hat{\beta}$ estimators are present in the expression for each delta value, making them dependent, up to some extent, on each other. This prevents direct consideration of the delta estimates for any standard procedure of statistical hypothesis testing of their mean.

To overcome this difficulty, a three-step procedure can be put in place.

1. Identify a matrix B such that the vector $\vec{\delta} := \mathrm{B}\vec{\Lambda}$ has all components with unit variance and uncorrelated from each other. This first step allows to break the co-dependence structure that the delta estimates might exhibit, by decorrelating its components. Appendix B presents a discussion on how to find the matrix B.
2. Run a normality test over the elements in the vector $\vec{\delta} := \mathrm{B}\vec{\Lambda}$ using any of the classical goodness-of-fit tests for normality. This step is crucial in the successful implementation of this methodology, because evidence of normality in addition to no-correlation (achieved in step 1) implies independence of the elements of $\vec{\delta}$. If normality is rejected, independence cannot be claimed, preventing the use of this methodology, since absence of correlation is not enough to build a statistical test supported in random sample theory. Under this scenario, a new methodology should be developed based, for example, on bootstrapping techniques.
3. Since the vector $\vec{\delta} := \mathrm{B}\vec{\Lambda}$ consists of a sequence of independent, unit-variance random variables, a simple t-test to check if the mean of these components is equal to zero suffices to provide a recommendation regarding the hypothesis presented in **(7)**, since a zero mean in the original $\Delta^{(\mathrm{i})}$ variables is equivalent to a zero mean in the elements of the vector $\vec{\delta}$.

3.2 Approximate solutions when the null hypothesis gets rejected

By means of the approach previously described, if the hypothesis $\Delta^{(\mathrm{i})} = 0,\ \forall i$ cannot be rejected, there is evidence that $\hat{\lambda}_{\mathrm{j}}^{(\mathrm{i})}, \forall j$ are reasonable estimates for the counts of each commodity type the *i*-th online shopper purchases. On the other hand, if there is evidence to reject the hypothesis, a new proposal for the estimation of the latent counts should be proposed.

The correction method presented in this section attempts to mediate between the methodological developments done in this work and an eventual rejection of the null hypothesis. This correction is based on the idea that even if $\Delta^{(\mathrm{i})}$ cannot be claimed to be statistically equal to zero, its value is reasonably close to zero, hence there is not a substantial deviation from the assumptions already made. Nevertheless, the nonzero value for $\Delta^{(\mathrm{i})}$ is incorporated in a redefinition of the latent mean count estimates. Recall from **(6)** that

$$e^{\Delta^{(\mathrm{i})}} = \frac{\sum_{j \in J} \lambda_{\mathrm{j}}^{(\mathrm{i})}}{\lambda^{(\mathrm{i})}} \Leftrightarrow \lambda^{(\mathrm{i})} e^{\Delta^{(\mathrm{i})}} = \sum_{j \in J^{(\mathrm{i})}} \lambda_{\mathrm{j}}^{(\mathrm{i})} \ . \tag{9}$$

Since $\Delta^{(i)} \neq 0 \therefore \exp\left(\Delta^{(i)}\right) \neq 1$, the sum of the actual latent count parameters is not equal to the mean count, violating the first assumption in **(2)**. In order to respect that expression, it is necessary to find an adjustment for the latent count estimates, so that **(2)** is observed. One way to adjust the parameters $\hat{\lambda}_j^{(i)}, j \in J$ is to define another latent marginal count

$$\bar{\lambda}_j^{(i)} := h\left(\Delta^{(i)}\right) \cdot \lambda_j^{(i)}, \tag{10}$$

for some function $h(\cdot)$ that depends on $\Delta^{(i)}$. A multiplicative adjustment is reasonable as it increases the counts proportionally based on their previous value. The multiplicative constant can be derived as follows

$$\begin{aligned} \sum_{j=1}^{J} \bar{\lambda}_j^{(i)} &= \lambda^{(i)}, \\ \sum_{j=1}^{J} \left[\lambda_j^{(i)} \cdot h\left(\Delta^{(i)}\right) \right] &= \lambda^{(i)}, \\ h\left(\Delta^{(i)}\right) \cdot \sum_{i=1}^{J} \lambda_j^{(i)} &= \lambda^{(i)}. \end{aligned} \tag{11}$$

After replacing **(9)** in the last equality, it gets that

$$\begin{aligned} h\left(\Delta^{(i)}\right) \cdot e^{\Delta^{(i)}} \cdot \lambda^{(i)} &= \lambda^{(i)}, \\ h\left(\Delta^{(i)}\right) \cdot e^{\Delta^{(i)}} &= 1, \\ h\left(\Delta^{(i)}\right) &= e^{-\Delta^{(i)}}. \end{aligned} \tag{12}$$

In conclusion, $\bar{\lambda}_j^{(i)} = e^{-\hat{\Delta}^{(i)}} \hat{\lambda}_j^{(i)}$ is an adjustment that respects that the summation of the latent count estimates equals the total mean count, while maintaining the reasoning behind the entire modeling effort. The new estimates, $\bar{\lambda}_j^{(i)}$, can then be used as proxies of the mean counts for each of the commodity types an individual purchases.

4 Data description and variable selection

The modeling methodology proposed in Chapters 2 and 3 was applied to online shopping data to estimate number of shipments from online shopping.

4.1 Sample description

The data was shared with the researchers by a Norwegian logistic company. It was collected in January 2017 from a panel of Norwegian respondents between 18 and 79 years of age with internet access (which in 2015 was 97 % of the age segment (PostNord, 2017)), and contains information about online shopping in December 2016. In this context, online shopping is defined as any goods purchased online and delivered to consumers at home, to a pick-up point, or collected by consumers themselves in a store, warehouse, etc. All purchases in store are excluded, even though the purchase is reserved online in advance. Purchase of services not providing any shipments and online purchases between businesses (B2B) or consumers (C2C) are also excluded. The total sample consists of 1,515 respondents, of which 1,019 respondents had shopped online at least once in December 2016 and got follow-up questions about their purchase(s). At most, 27 questions were asked, including total number of shipments from online shopping and commodity purchased, separating between 41 commodity types. For the 496 respondents who did not shop online in December 2016, only information about individual characteristics (demographics, socioeconomics factors and household characteristics) are available.

The dataset holds missing pieces of information due to non-response. Out of the variables included in the analysis, 75 respondents (5%) have missing data in at least one of the following variables: own income, number of shipments received, main benefit of online shopping, and preferred payment options when shopping online. Depending on the mechanism behind missing data, observations can be deleted, or values can be imputed. Deleting missing observations is accepted in the frequentist approach if both the missing and the observed data are at least random (Graham, 2009, Rubin, 1976). There is no guarantee imputation will produce unbiased estimates of the missing variables, and when the missing data points are in the dependent variable and these observations are missing at least at random, several imputation methods produce the same results as deletion (Allison, 2000). Additionally, when the total number of missing data points is low, the imputation mechanism is of negligible relevance (Schafer, 1999). Based on this evidence, deletion was chosen leaving 1440 respondents for analysis.

4.2 Variable selection based on multiple correspondence analysis and literature review

To select variables for model estimation, existing literature on online shopping were consulted. It showed that income, internet use and experience, e-shopping attitude, potential for lower price, higher education, working full-time jobs or long hours, kids in the household and residency in urban areas have a positive impact on online shopping; while risk of security breach, household size, ethnicity and age have a negative impact (Farag et al., 2007, Cao et al., 2012, Zhou and Wang, 2014, Limayem et al., 2000, Lohse et al., 2000). Regarding freight trip generation by households, Wang and Zhou (2015) and Saphores and Xu (2020) estimated number of shipments from online shopping per household from the US 2009 and 2014 National Household Travel Survey. They showed that web use, education, income, age, race and lifestyle, including household composition impact both the probability to purchase online as well as the number of deliveries received at home. Gardrat et al. (2016) found tendencies towards an increase in both practice and number of deferred purchase and reception (DPR) with the social profile of the head of the household, covering both deliveries from online shopping and shopping trips.

Gardrat et al. (2016) also distinguished between commodity groups and found that groceries (including catering) were the most frequent commodity purchased and delivered to consumers, followed by clothing; high-tech items and culture; household appliances, furniture and other products, and last; healthcare and cosmetics. This paper, together with the degree of quality judgement online, risk of erroneous purchase and consumer characteristics (inspired by a multiple correspondence analysis (see Appendix B)) influenced the aggregation of the 41 commodity types into 5 main commodity groups:

Children and leisure goods include products about which the same quality judgement can be made online as in-store. No pre-knowledge of the product is necessary, and the combination of high quality judgement and low price makes it a low risk commodity for online purchases (Lowengart and Tractinsky, 2001). The MCA showed that buyers of these items suffer from "time starvation" with kids under 15 years in the household and some distance to their closest urban area. Their main benefit of online shopping is either better selection or lower price.

Electronics covers commodities about which consumers can make quality judgements based on information online, given that the customer has or gains some pre-knowledge about the product, and to a high degree meet customer's expectations at arrival. Hence online shopping induces relatively low risk of erroneous purchase or mismatch of expectations. The MCA showed that this consumer group has a predominance of male consumers, searching for the lowest price and potentially a better selection. The same tendencies are found for electronics and computer products in Levin et al. (2005).

Clothes, beauty and interior products include commodities that need experience prior to a quality judgement, and as such, induce some risk of erroneous purchase. The MCA shows that the consumers are typically women, and assess time saving or flexibility to shop at a preferred time (for instance outside of regular opening hours) as main benefit of online shopping. The latter can be related to shopping interests, which is an attribute that is relatively more important for clothes and products of its like than other commodities (Levin et al., 2005). Consumers who want to enjoy shopping, tend to prefer offline to online shopping or add traditional shopping to online shopping (Levin et al., 2005, Zhou and Wang, 2014).

Consumables are goods of which the quality cannot be known prior to consumption. Hence, the risk of buying a good that does not meet expectations is relatively high but can be reduced by multiple purchases of the same brand or from the same seller. The MCA indicates that these customers are usually women, have kids, assess comfort or time savings as the main benefit of online shopping, and tend to live in urban areas.

The last group, ***other*** *(miscellaneous)* commodities, includes commodities that do not fit in the previous four groups. These are i) antiques and nutritional supplements, of which the quality cannot be assessed without careful inspection or long time experience (de Figueiredo, 2000, Girard et al., 2002); ii) commodities that fall out of the 41 given commodity types and are included in the open alternative "Other…", typically tickets and trips; and iii) purchases where the consumer did not know or did not report the commodity type. This group has high variability and diversity. No joint assumptions can be made about quality judgement, risk of erroneous purchase or consumer characteristics.

A summary of the five aggregated commodity groups is presented in Table 1, followed by descriptive statistics of variables sought relevant for analysis in Table 2. The latter presents statistics in total, for buyers and non-buyers, and per commodity group. For the variables *age*, *male*, *kids*, *education* and *income,* average numbers for the Norwegian population were collected from Statistics Norway (2017b), Statistics

Norway (2017a), Statistics Norway (2016a), Statistics Norway (2016b). Note that the population statistics have some discrepancies from the sample. The main difference is that population numbers include both individuals with and without internet access. Other discrepancies are highlighted in the table below.

Table 1: Summary and examples of the five aggregated commodity groups based on MCA and literature review.

Group	*Quality judgement*	*Risk of erroneous purchase*	*Consumer characteristics*	*Example*
Children and leisure goods	High	Low	Have kids, motivated by better selection and price online	CDs, board games, toys, magazines
Electronics	Medium/high	Low	Male, motivated by lower price and better selection online	Phones, digital games, kitchen and beauty appliances, TVs, PCs
Clothing, beauty and interior	Low/medium	Medium	Female, motivated by time savings and flexibility, have interest in shopping	Sweaters, Makeup, Shoes
Consumables	Low	Medium/high	Female, have kids, “time-starved”, live in urban areas	Food and other groceries, lenses and glasses
Other goods	-	-	-	Antiques, nutritional supplements, tickets and trips

Table 2: Descriptive statistics (min/mean/max) for the variables included in the estimation as a total and for each model. For the count data portion, descriptive statistics are presented per commodity group covering only individuals who purchased from that group. Variable names are partly given by authors. For binary variables (B) only the proportion of "ones"/"yes" outcomes is presented.

		Total[1]	Binary-outcome model		Count data model				
Var.	**Descriptions**	**Sample (Popul. '16)**	**Do not shop**	**Shop**	**Children and leisure**	**Electronics**	**Interior, clothing and beauty**	**Consumables**	**Other**
Ship	Number of shipments (dependent variable)	0/1.97/21	Na	1/2.95/21	1/3.76/21	1/3.69/21	1/3.62/21	1/4.56/21	1/2.96/18
Age	Age	18/50/79 (18/46/79)	18/57/79	18/47/78	18/47/78	19/46/75	18/44/78	20/46/77	18/48/78
Inc	Personal annual gross income (in 1000 NOK)	100/442/1,000 (100/432[4]/1,000)	100/426/1,000	100/450/1,000	100/453/1,000	100/467/1,000	100/427/1,000	100/505/1,000	100/448/1,000
Km	Number of kilometers from nearest urban area	0.0/5.1/63.5	0.04/5.43/46.82	0.01/4.94/63.5	0.09/5.23/56.77	0.09/5.37/56.77	0.04/5.17/63.5	0.13/4.26/42.03	0.01/5.02/40.4
Male	The respondent is a man (B)	0.50 (0.50)	0.51	0.50	0.49	0.71	0.35	0.45	0.52
Kids	Have kids under 15 years in the household (B)	0.20 (0.50[2])	0.101	0.26	0.31	0.27	0.30	0.33	0.24
Heduc	Have university or college education (B)	0.60 (0.33[3])	0.50	0.59	0.64	0.53	0.60	0.64	0.58
Empl	Working fulltime or as self-employed (B)	0.50	0.46	0.60	0.61	0.61	0.60	0.70	0.58
Price	Lower price is most important benefit of online shopping (B)	0.14	Na	0.21	0.22	0.29	0.19	0.22	0.19
Select	Better selection is most important … (B)	0.13	Na	0.19	0.18	0.18	0.17	0.15	0.22
Time	Time savings is most important … (B)	0.06	Na	0.09	0.09	0.08	0.11	0.09	0.08
Flex	Shop at preferred time is most important … (B)	0.23	Na	0.35	0.36	0.31	0.38	0.34	0.36
Comf	Comfort is most important … (B)	0.07	Na	0.10	0.11	0.11	0.11	0.18	0.08
N	**Number of observations**	**1440**	**479**	**961**	**455**	**327**	**437**	**132**	**245**

[1]Population equivalents are presented under parenthesis, when available; [2]Includes children at 15 to 17 years, and people younger than 18 years and older than 79 years with kids in the household; [3]Includes children at 16 and 17 years; [4]Total income includes social welfare transfers

The variable *ship,* ranging from 0 (non-shoppers) to 21, is a combination of the binary variable for online shopping and the count variable for number of shipments from online shopping. Number of shipments after the value of 10 was originally reported through intervals, with values being greater than 21 in the last interval. These were recoded using the interval midpoint values. Hence, the variable exhibits censoring from 10 on (reported as intervals) and right censoring at 21. This affects few observations (for example, only two observations lie above 21), and no correction mechanism was applied. Income was also reported using intervals and recoded using midpoint values. The variable *Km* was generated based on consumers' zip code and a distance matrix between the zip code and its nearest urban area using QGIS (QGIS Development TEAM, 2018). In this context, an urban area is defined as a center zone consisting of one or more center kernels and a 100-metre zone surrounding them (Statistics Norway). A center kernel is an area with at least 3 different main types of economic activities within 50 meters distance, where retail trade, government administration, health and social services or social and personal services must be present (Statistics Norway).

5 Estimation results and implications

5.1 Estimation results

Table 3 presents estimation results of the model parameters given by the methodology proposed in Chapter 2 and the variables in Table 2. The results for the hurdle splitting (buy/no buy decision) and the parameters for the actual buyers are presented separately and per commodity group.

Table 3: Results from estimation of a hurdle model with five commodity groups.

Zero hurdle splitting model coefficients (binomial with logit link):					
		Estimate	Std, Error	z value	Pr(>|z|)
	(Intercept)	0.135	1.179	0.114	0.909
	age	-0.041	0.004	-9.432	< 2e-16 ***
	Log(inc)	0.191	0.096	1.980	0.048 *
	Male	-0.074	0.122	-0.606	0.545
	Kids	0.512	0.184	2.788	0.005 **
	Higheduc	0.324	0.122	2.648	0.008 **
	Km	1.21e-04	0.008	0.015	0.988
Count model coefficients (truncated negbin with log link):					
		Estimate	Std. Error	z value	Pr(>|z|)
Other	Constant	0.334	0.068	4.920	8.63e-07 ***
Electronics	Male	0.444	0.090	4.942	7.73e-07 ***
	Price	0.155	0.130	1.186	0.236
	Selection	0.352	0.147	2.395	0.017 *
Clothes, beauty and interior products	Male	-0.131	0.103	-1.268	0.205
	Log(inc)	0.047	0.008	5.964	2.47e-09 ***
	Selection	0.068	0.152	0.446	0.655
	Time	0.108	0.165	0.655	0.513
	Flexibility	0.138	0.116	1.190	0.234
Consumable	Km	-0.024	0.012	-2.053	0.040 *
	Male	-0.315	0.168	-1.879	0.060 .
	Comfort	0.374	0.208	1.799	0.072 .
	Time	0.131	0.284	0.461	0.645
	Kids	0.194	0.170	1.142	0.254
	Log(inc)	0.050	0.011	4.455	8.39e-06 ***
Children and leisure goods	Km	0.016	0.005	3.086	0.002 **
	Kids	0.404	0.093	4.358	1.32e-05 ***
	Price	0.330	0.120	2.749	0.006 **
	Selection	0.305	0.132	2.308	0.021 *
Negative binomial	Log(alpha)	0.965	0.139	6.935	4.06e-12 ***
Alpha: count = 2.6239					
Number of iterations in BFGS optimization: 31					
Log-likelihood: -2525 on 27 Df					

Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1

Starting with the zero-hurdle splitting model, the results show that higher income, education, number of kids and distance (km) to nearest urban area are positively correlated with, and thus, motivate online shopping. One reason might be "time starvation", as all the factors above imply less time for traditional shopping. On the other hand, age are negatively related to the decision of buying online. Gender and number of kilometers are insignificant. The results are in line with existing research on the subject, as presented in Chapter 4.

The first thing to notice about the count model is that the negative binomial parameter for the variance (referred to as *Log(alpha)*) is significant, exhibiting the importance of allowing heterogeneity in the count data model. Among the covariates, *gender* indicates that males are more willing to buy electronic items than females, while females purchase more clothes, beauty and interior products, and consumables. This might be related to interests, and that women (still) tend to be more occupied with the household than men. *Income* has a positive influence for clothes, beauty and interior products, and consumable purchases; likely representing purchasing power and time starvation. Having kids under the age of 15 in the household also increases purchases of consumables and children and leisure goods; probably reflecting demand as well as time-starvation. Distance to urban areas favors the acquisition of children and leisure goods via internet, but for consumables the effect is the opposite. Both might be explained by lower service in rural areas than in cities, or that people in rural areas have a higher propensity to travel by car and stop by the local supermarket to buy groceries and pick up parcels on their way. In terms of perceived benefits, the possibility of having lower prices and a greater selection motivates consumers to buy electronics and children and leisure items online. Flexibility to shop at any given time influences positively the online purchase of clothes, beauty and interior products. Time savings are conceived as an advantage for both consumables and clothes, beauty and interior products. Comfort is associated with online purchase of consumables.

5.2 Validation of the methodology for count splitting

The significant variables (with p-value ≤ 0.1) in Table 3 are used to calculate average number of shipments in total and per commodity group. To validate the latent count estimates for each commodity group, the delta values must be computed and a hypothesis system to determine if they are statistically equal to zero or not must be tested.

The corresponding delta values computed from the estimation are shown using a histogram in Figure 1. As it can be seen from Figure 1, the values for delta are relatively small, ranging from -0.6 to 0.9 with an average value of 0.11, which means that, in average, the summation of the latent marginal counts corresponds to $e^{0.11} \approx 1.12$ times the estimated total count (12% larger). In terms of median performance, the median value of the deltas is equal to zero. This suggests that the discrepancies between the observed counts and the proposed split are minor.

The histogram for the standardized delta values (after using a first-order approximation to compute the variance-covariance structure and using it to decorrelate and standardize them according to Appendix A) is presented in Figure 2. A normality test using the Jarque-Bera method was conducted with a p-value of 0.734, allowing to conclude that the normality assumption is valid (at a 5% significance level). Figure 2 also includes the overlaid curve of a normal density function with parameters provided by the standardized delta values to see the fit graphically. However, the p-value for the t-test establishes that there is evidence to reject the zero-mean hypothesis (p-value=0.008).

Given the outcome of the hypothesis testing and the low value of the original delta values (presented in Figure 1), the adjustment presented in Subsection 3.2 was considered to be a reasonable approach to account for non-null, yet small, delta values in the estimate of the latent marginal counts.

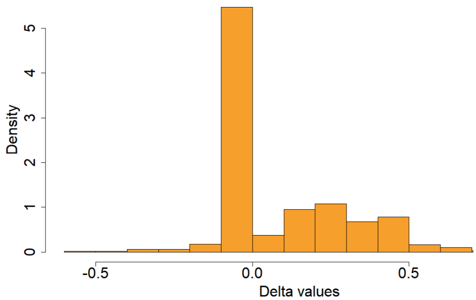

Figure 1: Histogram of computed delta values

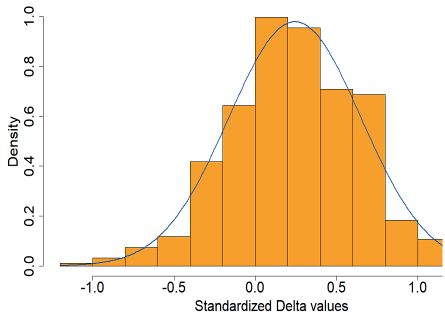

Figure 2: Histogram of standardized deltas

6 Application: number of shipments per commodity group

The estimation results above showed that, on average, an individual attracts 2.4 shipments per month (0.077 shipments per day (considering that the data collecting process was conducted in December, which has 31 days)). *Clothes, beauty and interior products* and *Children and leisure products* have the highest number of shipments with 0.77 shipments/person-month (0.025 shipments/person-day) and 0.64 shipments/person-month (0.02 shipments/person-day) respectively. They are followed by *Electronics* (0.46 shipments/person-month, or 0.015 shipments/person-day), *Other* (0.31 shipments/person-month, or 0.01 shipments/person-day), and lastly, *Consumables* (0.23 shipments/person-month, or 0.0075 shipments/person-day). Data collection in December might explain the relatively small contribution of consumables. However, purchases of groceries online were probably less common in 2016 than at present.

Table 4 presents the average number of shipments per person-month for each commodity group according to population segments for *age*, *income*, *gender* and *kids*. Some key findings are that consumers with kids have the highest average number of total shipments, followed by consumers in the second highest income interval and consumers in the age group 38-47 years. Females receive almost twice as many shipments of consumables, clothes, beauty and interior products than men, while men purchase electronics to a much higher extent than females (almost five times). Measured in total number of shipments, the gender difference is small.

Table 4: Average number of shipments per person per month for each commodity group according to different population segmentations (age, income, gender, presence of kids). Max (and min) in each column are highlighted.

Age	Electronics	Interior, clothing and beauty products	Consumables	Children and leisure products	Other goods	**Row total**
18-27	0.46	0.78	0.09	0.48	0.31	2.12
28-37	0.48	0.97	0.33	0.79	0.27	2.85
38-47	0.52	0.91	0.34	0.89	0.30	2.96
48-57	0.54	0.72	0.25	0.62	0.29	2.42
58-67	0.35	0.61	0.18	0.42	0.36	1.92
68-78	0.36	0.50	0.09	0.52	0.34	1.81
Group average	0.46	0.77	0.23	0.64	0.31	2.40

Income (NOK)*	Electronics	Interior, clothing and beauty products	Consumables	Children and leisure products	Other goods	**Row total**
100,000	0.42	0.73	0.11	0.53	0.32	2.11
250,000	0.39	0.85	0.26	0.65	0.35	2.49
349,999	0.43	0.75	0.19	0.55	0.30	2.22
450,000	0.45	0.81	0.26	0.68	0.26	2.46
549,999	0.40	0.79	0.28	0.65	0.37	2.48
650,000	0.60	0.73	0.19	0.69	0.28	2.49
749,999	0.41	0.92	0.19	0.69	0.24	2.46
899,999	0.79	0.86	0.30	0.76	0.33	3.04
1000000	0.64	0.32	0.43	0.70	0.36	2.45
Group average	0.46	0.77	0.23	0.64	0.31	2.40

Gender	Electronics	Interior, clothing and beauty products	Consumables	Children and leisure products	Other goods	**Row total**
Female	0.16	0.98	0.30	0.62	0.30	2.36
Male	0.75	0.56	0.16	0.65	0.32	2.44
Group average	0.46	0.77	0.23	0.64	0.31	2.40

Kids	Electronics	Interior, clothing and beauty products	Consumables	Children and leisure products	Other goods	**Row total**
No	0.44	0.68	0.19	0.46	0.31	2.09
Yes	0.51	1.00	0.34	1.13	0.30	3.28
Group average	0.46	0.77	0.23	0.64	0.31	2.40

*Midpoint for each income interval, as reported in the original dataset.

6.1 Freight trip implications of home deliveries per commodity type

Allen et al (2018) states that operating costs might be reduced by increasing the vehicle load factor as well as the drop density per round trip. Consolidating more goods to the same consumer or neighborhood would target these aims. Another issue related to home delivery is failed deliveries. As much as 1 in every 20 orders are not delivered on their first attempt (Loqate, 2018), generating both economic losses and unnecessary freight trips. For establishments, industry grouping is a popular indicator of heterogeneity in freight trip generation estimates. Similarly, splitting home deliveries in commodity groups can explain some of the heterogeneity in freight transport to consumers. The next section provides an example of freight trip reduction opportunities by exploring correlation between commodity groups purchased and consolidation of shipments.

Figure 3 and Figure 4 present positive correlation (red-colored cells, with a stronger color denoting a stronger correlation) and the slope of a linear regression model for purchase behavior between two commodity groups (numbers) for respondents with and without kids. The estimated latent counts are used as input for this calculation based on the expression:

$$\hat{\beta}_{\mathrm{y|x}} = \hat{\rho}(x, y)\frac{\hat{\sigma}_{\mathrm{y}}}{\hat{\sigma}_{\mathrm{x}}} \ , \qquad (12)$$

where $\hat{\beta}_{\mathrm{y|x}}$ is the slope of the regression of shipments for commodity type y taking as input the shipments of commodity type *x*, $\hat{\rho}(x, y)$ is the estimate of Pearson's correlation coefficient for the two commodity groups and $\hat{\sigma}_{\mathrm{y}}, \hat{\sigma}_{\mathrm{x}}$ are their corresponding (estimated) standard deviations. When interpreting the numbers, the commodity group in the *column should be read as input* and the commodity group in the *row as output*. Note that the coefficients are not symmetric, i.e., $\hat{\beta}_{\mathrm{y|x}} \neq \hat{\beta}_{\mathrm{x|y}}$. Still, only half of the matrix of coefficients is presented to illustrate their use due to space limitations. Zero values should be interpreted as either zero or negative correlation between commodities, i.e. correlations that do not translate into consolidation efforts. On the other hand, NA values correspond to cases in which not enough information was available to provide reliable estimates.

There are several implications derived from the result in the figures. First, shipments of consumables and clothes seem to be positively related in almost all combinations. The size varies, with the largest value corresponding to people 37 years or younger, with kids, and in the lowest income group. For people with kids, a general positive correlation for consumables and children and leisure products also exists, except for individuals with both low age and high income. This result suggests consolidation opportunities between consumables (like groceries) and clothing products, beauty and interior products; and/or children and leisure goods. Other, less general correlations are also present.

The linear regression between consumables and clothes, interior and beauty products shows that for several consumer types, as many as 50 % of consumers who received shipments of consumables, also received shipments of clothes, interior and beauty products. Similarly, up to 42 % of consumers receiving shipments of children and leisure goods also received shipments of consumables. Hence, considerable consolidation opportunities exist, and they are found between commodities with different transport and delivery requirements: the transportation operation and characteristics for children and leisure goods, and clothes, interior and beauty products is flexible compared to consumables, with strict requirements regarding transport, handling and delivery (Lagorio and Pinto, 2020, Saphores and Xu, 2020). By consolidating shipments from these commodity groups, multiple freight trips to a given household may be avoided and the share of failed deliveries reduced, the latter as the large cost of failed deliveries motivates e-grocers to experiment with alternative deliveries (Saphores and Xu, 2020). Finally, consolidating consumables with other commodity groups to a household or neighborhood might not suffer from high number of destination points and strict time constraints, as identified when traditional groceries are consolidated with e-groceries (Lagorio and Pinto, 2020).

For consolidation of different types of commodities to work in practice, there must be either a sufficiently large or homogenous demand for the commodity groups in question. Statistics for boroughs in Oslo (the capital of Norway) shows clear patterns regarding demographics (see Appendix C for a table of descriptive statistics and a map of the boroughs). Suburban residents with medium to large households, high income, medium age and low share of people living in apartments (like the boroughs *Nordstrand, Vestre Aker* and *Ullern* in Appendix C) are areas that the model identifies with a potentially large demand of the commodity groups in question. The low share of people living in apartments indicate a relatively low density, and hence the need to consolidate shipments to achieve small routes with high vehicle load factor. Table 5 shows the potential impact a best-case consolidation effort applied to the three boroughs in Oslo might have (in terms of total shipments), after estimating total counts per commodity group using an average individual profile derived from the information presented in Appendix C, and utilizing the slope coefficients for the population segment in which each borough is classified to determine the extent of consolidation strategies. The *Full consolidation* scenario assumes that all shipments can be consolidated according to the indication of the linear regression coefficients shown above (it corresponds to a best-case scenario). It was assumed that the three boroughs were classified as being predominantly composed of households with kids (Figure 4).

As a result of this application, it can be seen that the number of separate shipments can be reduced significantly if cooperative efforts to develop multi-sector and multi-company partnerships are supported.

Households with kids		Income segments (NOK)															
		Less than 350,000				350,000-550,000				550,000-750,000				more than 750,000			
Age segments		cloth	cons	child	oth	cloth	cons	child	oth	cloth	cons	child	oth	cloth	cons	child	oth
less than or equal to 37	elect	0	0	0	0	0	0	0	0	0.07	0.06	0	0	NA	NA	NA	NA
	cloth		0.56	0.22	0		0.37	0.43	0		0.49	0.04	0		0.35	0	NA
	cons			0.31	0.11			0.42	0.11			0.26	0			0	NA
	child				0				0				0				NA
37-57	elect	0	0.13	0.22	0	0	0	0.18	0	0	0	0.14	0	0.14	0.42	0.42	0
	cloth		0.26	0	0.08		0.33	0.2	0		0.12	0.22	0		0.52	0.05	0
	cons			0.08	0.45			0.07	0.12			0.09	0.27			0.15	0
	child				0				0				0				0
Older than 57	elect	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
	cloth		NA	NA	NA		NA	NA	NA		NA	NA	NA		NA	NA	NA
	cons			NA	NA			NA	NA			NA	NA			NA	NA
	child				NA				NA				NA				NA

Figure 3: Beta values of linear regression model between commodity groups. Consumers with kids.

Househ. without kids		Income segments (NOK)															
		Less than 350,000				350,000-550,000				550,000-750,000				more than 750,000			
Age segments		cloth	cons	child	oth	cloth	cons	child	oth	cloth	cons	child	oth	cloth	cons	child	oth
less than or equal to 37	elect	0	0.09	0	0	0	0.09	0	0	0	0	0.11	0	NA	0	0	0
	cloth		0.29	0	0		0.13	0	0		0	0.23	0		NA	NA	NA
	cons			0.04	0			0.05	0			0.03	0			0.03	0
	child				0				0				0				0
37-57	elect	0	0	0.35	0.03	0	0	0	0	0	0	0	0	0	0	0	0
	cloth		0.04	0	0.06		0	0	0		0.18	0.15	0		0.13	0	0
	cons			0	0			0	0.01			0	0			0	0.18
	child				0				0				0.01				0
Older than 57	elect	0	0	0	0	0	0	0	0	0	0	0	0	0	0.34	0	0
	cloth		0.11	0	0		0.16	0	0		0.03	0	0		0	0	0
	cons			0	0			0	0			0	0			0.07	0
	child				0				0				0				0

Figure 4: Beta values of linear regression model between commodity groups. Consumers with NO kids

Table 5: Analysis of the impact of consolidation for three boroughs in Oslo, based on the linear regression coefficients previously presented.

Borough	Income sector	Age sector	Number of shipments per commodity					Total shipments		
			Electronics	Clothing	Consumables	Children	Others	No consolidation	Full consolidation	Relative gap
Nordstrand	550,000-750,000	37-57	12,522	24,043	6,418	18,340	9,751	71,075	59,419	16.40%
Vestre Aker	More than 750,000	37-57	12,105	23,672	6,365	17,972	9,470	69,584	50,575	27.32%
Ullern	More than 750,000	37-57	8,365	16,299	4,379	12,199	6,544	47,787	34,705	27.38%

A drawback with the values (number of shipments) presented in this analysis is its reliance on information of online shopping in December. December might not be representative of online shopping as it includes Christmas shopping and holidays. However, a consumer survey reveals that as much as 67 % of the population between 18-79 years old (with internet access) shop online at least once a month (Postnord, 2020). Although MCA and literature reviews are conducted to classify commodity groups, the result might be influenced by both errors in the grouping and total number of shipments. However, we believe the methodology and potential application are relevant and of further research interest.

7 Conclusions

This paper builds on the econometric model to estimate latent marginal counts based on an observable total count by Afghari et al. (2016), Afghari et al. (2018), and complements it with a statistical test to assess its validity. The model is applied to an e-commerce dataset to estimate online shopping behavior given by the propensity to shop online and corresponding number of shipments in total and applies the splitting methodology to infer the marginal counts of five different commodity groups of interest. The results are assessed in terms of concordance of the purchase profiles obtained and the literature consulted, and the potential freight trip implications of the results, exploiting the co-dependence structure between the demands for different types of commodities.

Accordingly, the novelty in this research are i) proposing a statistical test to measure the validity of the assumptions made in the modeling process and a correction mechanism, based on the evidence provided by the data, in case the deviations from the assumptions are minor, ii) estimating number of shipments per commodity and hence revealing some of the heterogeneity in the freight trip attraction to consumers, and iii) suggesting how knowledge of commodity groups and purchase behavior can help policy makers and transport companies achieve more sustainable freight transport in urban areas through consolidation.

The outcome of the methodology proposed, and the results obtained for the case study of online shopping, shows a promising development in inferring non-observable features, and the implications these results might have for different stakeholders and decision makers involved. This research effort is particularly timely, especially during this unprecedented time of pandemic situation, in which online shopping and household deliveries have become a new standard, in order to avoid life-endangering situations. However, more studies should be undertaken to strengthen the findings in this paper in this context.

Acknowledgements: The authors would like to thank the Norwegian logistics company for sharing their data.

Funding: This work is undertaken as part of the research project 250432 NORSULP (Sustainable Urban Logistics Plans in Norway), financed by the Research Council of Norway and the Norwegian Public Roads Administration. Further information is available from www.norsulp.no (in Norwegian only).

Conflicts of interest: None

Bibliography

AFGHARI, A. P., HAQUE, M. M., WASHINGTON, S. & SMYTH, T. 2016. Bayesian Latent Class Safety Performance Function for Identifying Motor Vehicle Crash Black Spots. *Transportation Research Record,* 2601**,** 90-98.

AFGHARI, A. P., WASHINGTON, S., HAQUE, M. M. & LI, Z. 2018. A comprehensive joint econometric model of motor vehicle crashes arising from multiple sources of risk. *Analytic Methods in Accident Research,* 18**,** 1-14.

ALLEN, J., PIECYK, M., PIOTROWSKA, M., MCLEOD, F., CHERRETT, T., GHALI, K., NGUYEN, T., BEKTAS, T., BATES, O., FRIDAY, A., WISE, S. & AUSTWICK, M. 2018. Understanding the impact of e-commerce on last-mile light goods vehicle activity in urban areas: The case of London. *Transportation Research Part D: Transport and Environment,* 61**,** 325-338.

ALLISON, P. D. 2000. Multiple Imputation for Missing Data: A Cautionary Tale. *Sociological Methods & Research,* 28**,** 301-309.

CAO, X. J., XU, Z. & DOUMA, F. 2012. The interactions between e-shopping and traditional in-store shopping: an application of structural equations model. *Transportation,* 39**,** 957-974.

DABLANC, L., MORGANTI, E., ARVIDSSON, N., WOXENIUS, J., BROWNE, M. & SAIDI, N. 2017. The rise of on-demand 'Instant Deliveries' in European cities. *Supply Chain Forum: An International Journal,* 18**,** 203-217.

DE FIGUEIREDO, J. M. 2000. *Finding Sustainable Profitability in the E-commerce Continuum,* MIT Sloan Management Review.

ECOMMERCE NEWS. 2020. *Ecommerce in Europe: €717 billion in 2020* [Online]. https://ecommercenews.eu/ecommerce-in-europe-e717-billion-in-2020/. [Accessed].

FARAG, S., SCHWANEN, T., DIJST, M. & FABER, J. 2007. Shopping online and/or in-store? A structural equation model of the relationships between e-shopping and in-store shopping. *Transportation Research Part A: Policy and Practice,* 41**,** 125-141.

GARDRAT, M., TOILIER, F., PATIER, D. & ROUTHIER, J.-L. 2016. The impact of new practices for supplying households in urban goods movements: method and first results. An application for Lyon, France. *VREF conference on Urban Freight 2016.* Göteborg, Sweden.

GIRARD, T., SILVERBLATT, R. & KORGAONKAR, P. 2002. Influence of Product Class on Preference for Shopping on the Internet. *Journal of Computer-Mediated Communication,* 8**,** 0-0.

GRAHAM, J. W. 2009. Missing Data Analysis: Making It Work in the Real World. *Annual Review of Psychology,* 60**,** 549-576.

HOLGUÍN-VERAS, J., HODGE, S., WOJTOWICZ, J., SINGH, C., WANG, C., JALLER, M., AROS-VERA, F., OZBAY, K., WEEKS, A., REPLOGLE, M., UKEGBU, C., BAN, J., BROM, M., CAMPBELL, S., SANCHEZ-DÍAZ, I., GONZÁLEZ-CALDERÓN, C., KORNHAUSER, A., SIMON, M., MCSHERRY, S., RAHMAN, A., ENCARNACIÓN, T., YANG, X., RAMÍREZ-RÍOS, D., KALAHASHTI, L., AMAYA, J., SILAS, M., ALLEN, B. & CRUZ, B. 2018. The New York City Off-Hour Delivery Program: A Business and Community-Friendly Sustainability Program. *Interfaces,* 48**,** 70-86.

KESSY, A., LEWIN, A. & STRIMMER, K. 2018. Optimal Whitening and Decorrelation. *The American Statistician,* 72**,** 309-314.

LAGORIO, A. & PINTO, R. 2020. Food and grocery retail logistics issues: A systematic literature review. *Research in Transportation Economics*, 100841.
LEVIN, A. M., LEVIN, I. & WELLER, J. 2005. *A multi-attribute analysis of preferences for online and offline shopping: Differences across products, consumers, and shopping stages*, Journal of Electronic Commerce Research.
LIMAYEM, M., KHALIFA, M. & FRINI, A. 2000. What makes consumers buy from Internet? A longitudinal study of online shopping. *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans,* 30, 421-432.
LOHSE, G. L., BELLMAN, S. & JOHNSON, E. J. 2000. Consumer buying behavior on the Internet: Findings from panel data. *Journal of Interactive Marketing,* 14, 15-29.
LOQATE 2018. Fixing Failed Deliveries. *Improving Data Quality in Retail.* https://www.loqate.com/resources/thank-you/data-quality-report/?submissionGuid=e5417777-de99-4620-b51c-f1627e4d8027.
LOWENGART, O. & TRACTINSKY, N. 2001. Differential Effects of Product Category on Shoppers' Selection of Web-based Stores: A Probabilistic Modeling Approach. *J. Electron. Commerce Res.,* 2, 142-156.
MULLAHY, J. 1986. Specification and testing of some modified count data models. *Journal of Econometrics,* 33, 341-365.
NI, L., WANG, X. & ZHANG, D. 2016. Impacts of information technology and urbanization on less-than-truckload freight flows in China: An analysis considering spatial effects. *Transportation Research Part A: Policy and Practice,* 92, 12-25.
PETTERSSON, F., WINSLOTT HISELIUS, L. & KOGLIN, T. 2018. E-commerce and urban planning – comparing knowledge claims in research and planning practice. *Urban, Planning and Transport Research,* 6, 1-21.
POSTNORD. 2017. Netthandel i Norden 2017. *Norden - en digitalisert region: Slik ser nettkjøpsatferden ut i Norden i 2017.* [Online]. Available: http://www.postnord.no/nb/info/om-postnord/nyheter-og-presse/netthandelsrapporter/netthandel-i-norden-2017.
POSTNORD 2020. Netthandel i Norden - Oppsummering 2019. *In:* POSTNORD (ed.) *Netthandel i Norden.* PostNord.
QGIS DEVELOPMENT TEAM 2018. QGIS Geographic Information System. Open Source Geospatial Foundation Project.
RUBIN, D. B. 1976. Inference and missing data. *Biometrika,* 63, 581-592.
SAPHORES, J.-D. & XU, L. 2020. E-shopping changes and the state of E-grocery shopping in the US - Evidence from national travel and time use surveys. *Research in Transportation Economics*, 100864.
SCHAFER, J. L. 1999. Multiple imputation: a primer. *Statistical Methods in Medical Research,* 8, 3-15.
SEN, P. K. & SINGER, J. M. 2017. *Large sample methods in statistics (1994): An introduction with applications*.
STATISTICS NORWAY 2016a. 07778: Registered incomes for residents persons (mill.NOK) 2006-2017. *In:* STATISTICS NORWAY (ed.) *Income and wealth statistics for households.*
STATISTICS NORWAY 2016b. 08921: Educational attainment, by county, age and sex (C) 1980-2018. *In:* STATISTICS NORWAY (ed.) *Educational attainment of the population.*
STATISTICS NORWAY 2017a. 06071: Persons, by type of household, contents and year. *In:* STATISTICS NORWAY (ed.) *Families and households.*

STATISTICS NORWAY 2017b. 07459: Population, by sex and one-year age groups (M) 1986-2019. *In:* STATISTICS NORWAY (ed.).
VISSER, J., NEMOTO, T. & BROWNE, M. 2014. Home Delivery and the Impacts on Urban Freight Transport: A Review. *Procedia - Social and Behavioral Sciences,* 125**,** 15-27.
WANG, X. & ZHOU, Y. 2015. Deliveries to residential units: A rising form of freight transportation in the U.S. *Transportation Research Part C: Emerging Technologies,* 58**,** 46-55.
ZHOU, Y. & WANG, X. 2014. Explore the relationship between online shopping and shopping trips: An analysis with the 2009 NHTS data. *Transportation Research Part A: Policy and Practice,* 70**,** 1-9.

Appendix A: First-order approximation of the covariance structure

In order to characterize the association structure of the delta estimators, an approximation of the second-order moments for the vector $\vec{\Lambda}$, through a stochastic first-order Taylor expansion of its components, was utilized. The first-order Taylor expansion for $\hat{\Delta}^{(\mathrm{i})}$ around β corresponds to

$$\hat{\Delta}^{(\mathrm{i})} = k\left(\hat{\beta}, X_{\mathrm{i}}\right) = k\left(\beta, X_{\mathrm{i}}\right) + \nabla_{\beta} k\left(\beta, X_{\mathrm{i}}\right)^{T}\left(\hat{\beta} - \beta\right) + o_{\mathrm{p}}\left(\left\|\hat{\beta} - \beta\right\|\right), \tag{13}$$

where $\underline{\nabla_{\beta}}$ is the gradient operator, $o_{\mathrm{p}}(\cdot)$ is the probabilistic equivalent of the little o-notation for random objects and $\|\cdot\|$ the Euclidean norm (Sen and Singer, 2017). The variance of $\hat{\Delta}^{(\mathrm{i})}$, based on this first-order approximation, can be computed as

$$\mathrm{var}\left(\hat{\Delta}^{(\mathrm{i})}\right) \simeq \nabla_{\beta} k\left(\beta, X_{\mathrm{i}}\right)^{T} \cdot \Sigma_{\hat{\beta}} \cdot \nabla_{\beta} k\left(\beta, X_{\mathrm{i}}\right), \tag{14}$$

with $\Sigma_{\hat{\beta}}$ being the variance-covariance matrix of the estimators $\hat{\beta}$. Since the true vector of parameters is not observed, this variance has to be estimated by replacing all the parameters involved in **(14)** with their corresponding estimates

$$\tau_{\mathrm{ii}} := \mathrm{var}\left(\hat{\Delta}^{(\mathrm{i})}\right) \simeq \nabla_{\beta} k\left(\hat{\beta}, X_{\mathrm{i}}\right)^{T} \cdot \hat{\Sigma}_{\hat{\beta}} \cdot \nabla_{\beta} k\left(\hat{\beta}, X_{\mathrm{i}}\right). \tag{15}$$

The estimated variance and covariance matrix, $\hat{\Sigma}_{\hat{\beta}}$, can be easily obtained from the maximum likelihood estimation of the count data model as the inverse of the observed Fisher's information matrix. The elements of the gradient can be computed as

$$\frac{\partial k}{\partial \beta_{\mathrm{l},\mathrm{K}_{\mathrm{l}}}}\left(\hat{\beta}, X_{\mathrm{i}}\right) = -\left[\frac{\sum_{j \in J \setminus \{l\}} \hat{\lambda}_{\mathrm{j}}^{(\mathrm{i})}}{\sum_{j \in J} \hat{\lambda}_{\mathrm{j}}^{(\mathrm{i})}}\right] x_{\mathrm{l},\mathrm{K}_{\mathrm{l}}}^{(\mathrm{i})} I_{\mathrm{l}}^{(\mathrm{i})}, \tag{16}$$

where $x_{\mathrm{l},\mathrm{K}_{\mathrm{l}}}^{(\mathrm{i})}$ is the is the K_{l}-th covariate associated with the *l*-th commodity type for individual *i*, $I_{\mathrm{l}}^{(\mathrm{i})}$ is the indicator variable for the *l*-th commodity type for individual *i*, and $\beta_{\mathrm{l},\mathrm{K}_{\mathrm{l}}}$ is their corresponding parameter. The covariance between any two components can also be estimated using the approximation in **(13)** as

$$\tau_{\mathrm{ij}} := \mathrm{cov}\left(\hat{\Delta}^{(\mathrm{i})}, \hat{\Delta}^{(\mathrm{j})}\right) \simeq \nabla_{\beta} k\left(\hat{\beta}, X_{\mathrm{i}}\right)^{T} \cdot \hat{\Sigma}_{\hat{\beta}} \cdot \nabla_{\beta} k\left(\hat{\beta}, X_{\mathrm{j}}\right). \tag{17}$$

Then, the matrix $\mathrm{T} := \left(\tau_{\mathrm{ij}}\right)_{\mathrm{i},\mathrm{j}}$ is an estimate of the variance-covariance matrix of the vector $\vec{\Lambda}$ that can be used to break the correlation between its components by means of the transformation

$$\vec{\delta} = \mathrm{T}^{-1/2} \vec{\Lambda}, \tag{18}$$

because

$$\operatorname{cov}\left(\vec{\delta}\right) = \mathrm{T}^{-1/2} \operatorname{cov}\left(\vec{\Lambda}\right) \mathrm{T}^{-1/2} \simeq \mathrm{T}^{-1/2} \mathrm{T} \mathrm{T}^{-1/2} = I_{\tilde{\mathrm{N}} \times \tilde{\mathrm{N}}}; \tag{19}$$

being $I_{\tilde{\mathrm{N}} \times \tilde{\mathrm{N}}}$ the identity matrix of size $\tilde{N}$. In the case that T is not a positive definite matrix (it is non-invertible), a similar result using the Moore-Penrose generalized inverse of T, T^{+}, can be implemented

$$\vec{\delta} = \left(\mathrm{T}^{+}\right)^{1/2} \vec{\Lambda}, \tag{20}$$

as long as $\left(\mathrm{T}^{+}\right)^{1/2} \mathrm{T} \left(\mathrm{T}^{+}\right)^{1/2} \approx I_{\tilde{\mathrm{N}} \times \tilde{\mathrm{N}}}$, which means that the variance-covariance matrix of $\vec{\delta}$ is close to the identity. There are other decorrelation matrices different from the proposed ones. For further details, please consult Kessy et al. (2018).

Appendix B: Multiple Correspondence Analysis

Figure B1 summarizes the results from the multiple correspondence analysis (MCA) on the 41 commodity groups. The axes are interpreted in terms of attributes of the online shopping experience inspired by Lowengart and Tractinsky (2001), Levin et al. (2005), Zhou and Wang (2014), de Figueiredo (2000), Girard et al. (2002). Dimension 1 could be associated with the risk (or the propensity) of unmet expectations or erroneous purchase. A high (positive) value suggests a high risk of making purchases that do not meet the buyers' expectations. For instance, some groceries have a high risk of not meeting the expectations because of freshness, transporting and handling conditions; while electronical products do not present that risk due to standard manufacturing procedures and vast marketing campaigns to inform the consumers. Dimension 2 can be interpreted as the amount of knowledge a buyer possesses about a product prior to online purchase. A high value means that the buyer is informed about the characteristics of the product he/she is about to buy. Once again, groceries have a high value in this dimension since the majority of individuals is familiar with the standard expected quality and appearance of the food they are about to purchase. Electronics is also a well-informed purchase, as opposed to clothes, beauty and interior products, of which qualities like texture and color are important aspects and difficult to describe with accuracy on the internet.

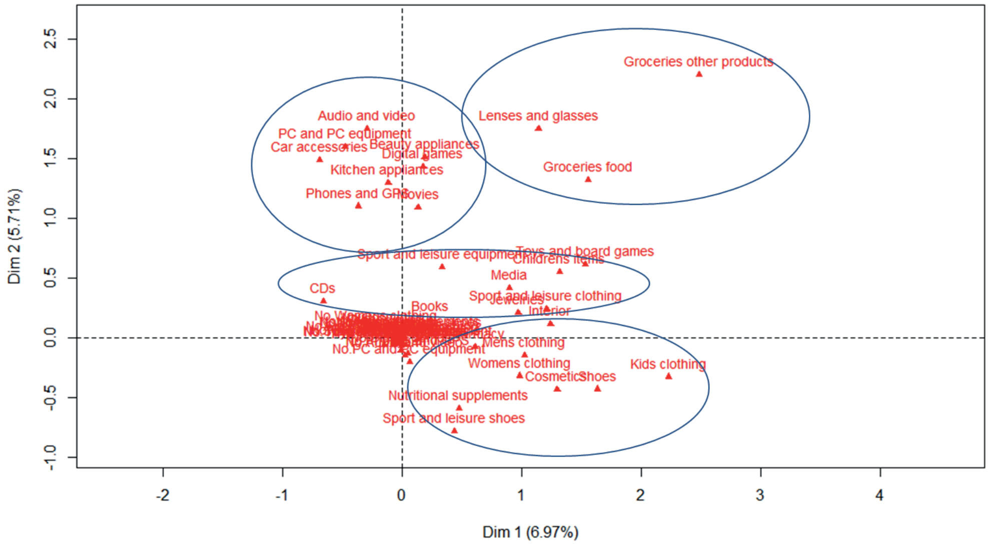

Figure B1: MCA results for the commodity clustering

Even though the MCA itself did not allow to reduce the dimensionality of the dataset due to low percentages of explained variability in the first components, it was of great help in motivating the number and composition of the groups.

Appendix C: The boroughs in Oslo

Table C1: Demographics for the boroughs in Oslo, Norway for year 2018 due to delayed income data for 2019. All statistics are collected from Statistics Norway in August 2020.

	*Adult population**	*Adults* per household*	*Live in apartment building (share)*	*Age (average)*	*Net person income per year***	*Female share*	*Kids*** per adult***
Alna	36,738	1.65	78.4	37.7	394,000	0.50	0.30
Bjerke	23,317	1.60	72.2	35.7	440,500	0.49	0.33
Frogner	49,535	1.41	91.6	39.0	670,300	0.50	0.14
Gamle Oslo	44,594	1.47	93.2	34.4	449,800	0.48	0.21
Grorud	20,502	1.61	75.8	38.4	384,400	0.50	0.29
Grünerløkka	49,735	1.43	93.4	33.5	456,900	0.49	0.17
Nordre Aker	37,495	1.58	43.3	37.3	623,200	0.50	0.30
Nordstrand	37,109	1.64	38.9	39.2	625,300	0.51	0.31
Østensjø	35,877	1.57	61.1	38.7	490,500	0.52	0.32
Sagene	36,341	1.38	95.0	34.3	482,300	0.51	0.17
Sentrum	1,059	1.30	86.6	34.1	352,200	0.40	0.04
Søndre Nordstrand	28,091	1.87	47.1	35.5	394,700	0.50	0.36
St. Hanshaugen	32,691	1.39	94.7	35.1	493,500	0.50	0.14
Stovner	23,801	1.79	63.8	38.0	375,300	0.49	0.33
Ullern	24,680	1.60	57.0	40.9	767,100	0.52	0.29
Vestre Aker	34,974	1.66	41.5	39.0	826,400	0.52	0.33

*18-79years, **All adults 17years and older,***0-17 years

Chapter 4:

Consumer preferences for reducing environmental impacts of last mile deliveries. Case: female clothing rentals

Elise Caspersen

Ståle Navrud

Consumer preferences for reducing environmental impacts of last mile deliveries. Case: female clothing rentals

Elise Caspersen[1,2,*], Ståle Navrud[2]

[1]Institute of Transport Economics – Norwegian Centre for Transport Research, Gaustadalléen 21, 0349 Oslo, Norway
[2] School of Economics and Business; Norwegian University of Life Sciences, Ås, Norway.
[*]Corresponding author at: Institute of Transport Economics – Norwegian Centre for Transport Research, Gaustadalléen 21, 0349 Oslo, Norway. Tel.: + 922 444 52, E-mail address: elc@toi.no

Abstract: *The sharing economy reduces production of goods as ownership is transferred permanently or temporarily between peers or between businesses and consumers, and its participation is, among other things, motivated by environmental concerns. Still, the sharing economy may increase freight transportation if its services require both shipments to consumers and returns to sender, as is the case for clothing rentals. This paper looks at whether consumers' environmental attitudes and behavior are reflected in their stated preferences for last mile delivery options for clothing rentals, and whether preferences are heterogenous across groups of respondents in terms of socioeconomic characteristics, income, and environmental attitudes.*

A discrete choice experiment (DCE) was conducted among females between 18 and 70 years of age, recruited from a nationwide internet panel in Norway, to reveal their preferences for delivery time, delays, information service (notification of product quality and departure), emission of local air pollutants (Particulate Matter - PM) and greenhouse gas emissions (CO_2) from last mile deliveries. Multinomial logit (MNL) and latent class (LC) models were used to analyze the DCE data with regards to respondent groups' preferences. The results show a general negative utility of delivery time, delays and emissions, while a positive utility was observed from information services.

The main results are: i) all respondent groups have negative utility of emission and three out of four groups from the LC model have a negative utility of both local air pollutants (PM) and greenhouse gas emissions (CO_2), ii) some respondent groups are willing to wait 5-10 days for their delivery if it implies reduced emissions, iii) respondents fully agreeing that consumers must change their attitudes and behavior to solve the current environmental challenges are more likely to accept longer delivery time to reduce emissions, and iv) as price is not included as an attribute in the DCE, the results provide knowledge on how other measures than price can incentivize consumers to choose environmentally sustainable deliveries. The findings are relevant for both urban planners, online retailers, and transport operators as they show that consumers should be presented with environmentally sustainable last mile delivery options. Tailoring measures to different consumer groups based on the results from this study could accelerate the transition to environmentally sustainable last mile deliveries.

Key words: sharing economy, clothing rentals, e-commerce, sustainable last mile delivery, discrete choice modeling, nudge

1 Introduction

The sharing economy, where goods or services are rented, swapped or traded through online platforms (Hamari et al., 2016), is expected to rise in future, and predicted to reach USD 335 billion by 2020, up from USD 15 billion in 2014 (Mazareanu, 2019). Only a decade ago, the sharing economy was mainly found within accommodation and transportation, with Couchsurfing, Airbnb and Uber being pioneering companies. Now, it can be found in almost every sector and activity in the world (Rinne, 2019), benefitting the consumer, the environment, and the community, as well as providing new business opportunities for firms that anticipate, welcome and adapt to it (Belk, 2014). One of many sectors adapting to the sharing economy is clothing rentals, covering both specialized online platforms (like Rent The Runway in the US, HURR in the UK and FJONG in Norway) and established fashion retailers (like H&M (Wilen, 2019)), providing consumer access to trendy outfits while benefiting both the wallet and the environment. In many ways, the sharing economy resembles e-commerce, where goods and services are bought and sold online, with the main difference being how ownership is transferred, and the fact that environmental aspects are a motivating factor for both consumers supporting the sharing economy (Hamari et al., 2016, Standing et al., 2019, Jia et al., 2020) and producers altering their production (Jia et al., 2020). Despite the emphasis on the environmental aspects of the sharing economy, little focus has been put on the freight aspect of it (Lim et al., 2018), although it is expected to alter freight traffic both through opportunities for new innovative transport solutions and changes in transfer of ownership (Carbone et al., 2018). The latter is particularly relevant for large-scale renting of goods between peers or between consumers and online platforms, i.e. clothing rentals, where temporary transfer of ownership generates both a delivery to consumer and a return to sender.

Freight transport has already been altered by e-commerce through online shopping, and sharing activities like clothing rentals are assumed to fuel this change. The growth in deliveries directly to consumers increases freight traffic in general and in residential areas (Allen et al., 2018, Visser et al., 2014). At the same time, the competition of market shares fosters a continuous push for shorter lead times and amplifies the poor vehicle utilization and duplication of delivery services (Allen et al., 2018). Even worse, it also increases instant deliveries (deliveries within 2 hours of ordering), resulting in higher total distance traveled as trucks are substituted with smaller modes of transport, as well as higher energy consumption and emissions if vans, private cars and motorcycles are used (Dablanc et al., 2017). On the other hand, increased delivery time flexibility is found to reduce vehicle kilometers and environmental impact (Manerba et al., 2018). Counteracting the continuous reduction in lead time might help reduce emissions from urban freight deliveries of both online shopping and renting. Accepting that parcel deliveries take more than 1 day, or even 5-10 days, might ease the implementation of last-mile collaboration or consolidation, logistics hotels, or crowdshipping (utilizing existing journeys), initiatives suggested by Allen et al. (2018) to improve efficiency in parcel delivery. Local pick-up points have also been found to increase consolidation, optimize delivery rounds, reduce vehicle kilometers and emissions compared to home deliveries (Morganti et al., 2014, de Oliveira et al., 2017, Heshmati et al., 2018). Thus, getting consumers to accept increased delivery time might foster sustainable urban freight transport. This paper contributes to this topic by analyzing consumer last mile delivery preferences when renting an outfit online, and is motivated by the following two main research questions: *i) Are consumers concerned about local air pollution and greenhouse gas emissions from last mile delivery, and if so, can they accept increased delivery time as a measure to reduce these emissions, and ii) do the results vary with consumer type or supplementary information (nudge) about environmental consequences of the fashion industry and last mile deliveries (provided to half the sample)?*

The research questions are answered using an internet panel survey of female respondents between 18 and 70 years of age, inquiring about online shopping and clothing rental habits, preferences and attitudes as well as performing a discrete choice experiment (DCE) to reveal last mile delivery preferences when engaging in clothing rentals. The novelty of this research is how consumers value different last mile delivery services (other than by price), and how the tradeoff between delivery time and emission of local air pollutants (Particulate Matter - PM) and greenhouse gas emissions (CO_2) can be utilized to achieve more sustainable urban freight transport. The authors have not found other papers addressing consumer preference for airborne emissions from last mile delivery nor from the sharing economy.

The rest of the paper is organized as follows: Section 2 provides a literature summary of consumer preferences for the sharing economy and environmentally sustainable deliveries. Section 3 presents the data collection and resulting sample. Section 4 briefly discusses the chosen model specification, followed by estimation results in section 5. Section 6 provides a discussion of the results including implications for urban freight transport. Section 7 concludes the paper.

2 Sharing economy and environmentally sustainable deliveries

This section summarizes existing literature on the field and helps identify relevant attributes for analysis.

2.1 Consumer behavior and motivation in the sharing economy

Hamari et al. (2016) identified positive attitudes towards the sharing economy, enjoyment of the activity, and potential of economic benefit as key motivations for participating in the sharing economy, with sustainability being an important factor through its impact on attitude. Piscicelli et al. (2015) identified environmental, social and financial motivations for joining an online marketplace to lend and borrow between peers. Separate studies have been done for millennials engaging in the sharing economy. Godelnik (2017) found that economic motives were by far the most important for millennials, followed by social and environmental ones, while Hawlitschek et al. (2018) identified financial-, trust-, lifestyle-, effort- and sustainability-related factors as key points. Standing et al. (2019) reviewed transport sharing activities and identified potential for reduced cost, enjoyment, environmental savings and practical aspects as key points for engaging. Hartl et al.

(2018) found that a green image is less important than price, easiness of use and flexibility as motivation for carsharing, and that environmental motives work only as a bonus when a consumer chooses P2P (peer-to-peer) over B2C (business-to-consumer) carsharing services. Cerutti et al. (2019) identified health and environment, social influence and lifestyle as main motivations for bike-sharing. Punel and Stathopoulos (2017) revealed that previous experience impacted preferences for using crowd-shipping services. Tradition and preserving the “status quo” were found to work against sharing economy participation (Piscicelli et al., 2015), as do time use, less freedom of product use and worry of product damage (“hassle cost”) in clothing rentals (Choi and He, 2019).

Demographics seems to play a smaller role in explaining sharing economy motivation. Diamantopoulos et al. (2003) found that socio-demographic variables overall explain a small proportion of the variance, although females hold stronger attitudes towards environmental qualities and green behavior than men, age influences environmental attitudes, high education has a partly positive influence on green behavior, while marriage status and number of kids are insignificant. Diamantopoulos et al. (2003) further argued that the importance of demographic measures is not necessarily transferable and might be outdated if attitudes are changing in society.

2.2 Consumer preferences for environmentally sustainable last mile deliveries

Few research papers were found on consumer preferences for last mile deliveries in the sharing economy. The issue was raised by Carbone et al. (2018), who identified four different freight sharing economy solutions: peer-to-peer and business logistics from freight generated by the sharing economy, and crowd or open logistics as part of the sharing economy for logistics (Carbone et al., 2018). Standing et al. (2019) reviewed literature on sharing economy in transport, and identified online platforms supporting increased vehicle utilization and better use of existing capacity as main activities. The impact of freight trips from the sharing economy was not included. Research on crowdshipping, being a freight-sector sharing activity, includes consumer preferences, but these are related to the adaption of the service (Punel and Stathopoulos, 2017, Buldeo Rai et al., 2021, Punel et al., 2018), not as a solution to reduce freight trips from the sharing economy itself. Hence, an extension beyond the sharing economy was needed to identify consumer preferences for environmentally sustainable last mile deliveries.

Schniederjans and Starkey (2014) investigated intention to buy and willingness to pay (WTP) for a green transportation t-shirt and concluded that a positive attitude towards green freight transport and peer pressure influence purchase intention. Polinori et al. (2018) analyzed students WTP for environmentally labelled last mile delivery when purchasing (also) a green t-shirt, and found that signals or opinions of external parties increased attitudes and WTP for ecofriendly labeling, and that females were more positive than males. Collins (2015) mapped customer preferences for last mile attributes when choosing between home delivery or pick-up point, and the environmental impact through choice of transport to and from the pick-up points. He found that price, quality, location of pick-up points and other delivery alternatives influenced choices and environmental behavior (Collins, 2015). de Oliveira et al. (2017) investigated the potential demand for home delivery and the more environmentally sustainable automated delivery stations (ADS). They found that location, delivery time, information and traceability, and cost of transport had a significant impact on utility. While home delivery was the preferred alternative for most people, ADS had a considerable potential market share if compensated through more flexible delivery time, information services or reduced costs (de Oliveira et al., 2017). Punel et al. (2018) investigated the difference between users and non-users of crowdshipping and revealed that crowdshipping was more prevalent among male, low income respondents who work full-time jobs, are less concerned with safety, trust and privacy issues, but more concerned about the environment, and thus choose crowdshipping to reduce the environmental impact of the delivery, not for monetary gains. Buldeo Rai et al. (2021) investigated consumer preferences towards innovative last mile initiatives, like crowdshipping, and found four consumer segments differing in terms of preference and attitudes, with socio-demographic variations being of less importance. The segment that was most positive to crowdshipping included frequent online shoppers, who preferred parcels delivered at home, were drawn towards innovation and benefitting their local community or the environment, and were willing to wait to avoid additional vehicle kilometers (Buldeo Rai et al., 2021). Valeri et al. (2016) identified environmental awareness and behavior intention to be a key contributor explaining environmental policy preferences.

The role of information in altering consumer preferences for emission reduction was inspired by the work on nudging behavior by Thaler and Sunstein (2009) and is supported in the literature. Godelnik (2017) identified that participants pre-project thought little about their consumption behavior, but post-project had increased their awareness. Polinori et al. (2018) found that information influenced student's WTP for green urban freight transport and should be widely available at low cost. Agatz et al. (2020) found that use of green labels denoting more environmentally sustainable delivery alternatives is an effective tool for steering behavior, with the effect being larger for eco-conscious consumers.

To sum up, positive attitudes towards green consumption or environmentally sustainable behavior influence both participation in the sharing economy and the demand for sustainable last mile delivery. Enjoyment of the activity also motivates sharing, while information or peer pressure motivate environmentally sustainable behavior. Research on demographic factors is inconclusive, but age, income, education, and gender might be relevant.

3 Survey and data collection

The data was collected through a survey composed of four parts: i) preliminary questions about habits and preferences for online shopping in general and clothing rentals in particular, ii) assertions related to clothing rentals and environment aiming to reveal attitudes, including some repeated questions to test consistency (as suggested by Mathews et al. (2007)), iii) stated choice scenarios (presented in detail in subsection 3.1) including debriefing questions and rating of attributes to reveal whether the respondents process all attributes equally, or some more than others (inspired by Hensher (2006)), and iv) socio-economic data including age, educational level, occupation, own and household income, and household members. Additionally, respondents were randomly assigned to get supplementary information regarding environmental aspects of the textile industry and last mile deliveries (described in subsection 3.2).

3.1 Discrete choice experiments

The aim of the survey was to capture consumer preferences for last mile delivery services attributes (including non-market goods time and emission), for which stated choice methods (SC) with discrete choice experiments (DCE) (as opposed to contingent valuation (CV)) are recommended (Johnston et al., 2017). To capture as realistic experiments as possible, attributes and attribute levels were inspired by consumer surveys conducted by Postnord (2020) and Bring Research (2019), World Economic Forum (2020), as well as existing knowledge of the research team. FJONG, a Norwegian online platform for clothing rental, provided valuable insights and comments to the development of the survey. The resulting design included 5 attributes with 2-4 levels each, as presented in Table 1.

Table 1: Characteristics of the experimental design: attribute description and levels.

Attribute	Description and levels
1. Delivery time	Number of days the respondent accepts to wait for the parcel:
	1-5-10-20 days
2. Delays (dummy)	Uncertainty with respect to delivery time:
	"No", "Yes, 1-2 days"
3. Information (dummy)	Notifications by SMS or e-mail when 1) the good is controlled and approved for shipping and 2) the parcel is shipped to the consumer:
	"No", "Yes"
4. CO_2 -emission	CO_2-emission resulting from last mile delivery of the parcel. The emission levels differ with respect to transport mode, time, degree of consolidation etc.:
	0kg, 0.28kg, 1.40kg
5. Particulate matter (PM)	PM resulting from last mile delivery of the parcel. Differs with respect to transport mode, time, degree of consolidation etc.:
	"Low", "Medium", "High"

Table 1 shows that delivery time varies between 1 and 20 days. The latter might be unrealistic, potentially violating the assumption about realistic experiments. On the other hand, the probability that consumers consider all attributes increases and more information is obtained when attribute levels widen (Hensher, 2007, Johnson et al., 2007). The 20 days delivery time was included to investigate maximum wait time for consumers. The levels of CO_2-emission are calculated based on average last mile delivery distance (from Statistics Norway) and emission levels (using the

Handbook Emission Factors for Road Transport (HBEFA)) for light duty vehicles. The levels for PM are qualitative, which should be avoided in stated preference surveys (Johnston et al., 2017, Johnston et al., 2012). However, as PM is an unfamiliar concept for many, the use of qualitative terms was found to be the best way of presenting the levels of this attribute. Price is not included as an attribute in the DCE. Free delivery is a "must-have" to attract customers for many online retailers (Allen et al., 2018), and as many as 21% of Norwegian male customers and 31% of female customers expect free delivery (Bring Research, 2019). Excluding price enables an assessment of how other measures (than price) can incentivize consumers to choose environmentally sustainable deliveries.

To reduce the complexity of the survey, 3 unlabeled alternatives were used. Two alternatives differed in terms of the attributes presented above, while the third was an "opt-out" alternative, to ensure realism for accepting last mile delivery and reliable tradeoffs between the attributes. Each respondent got 9 choice sets, drawn randomly from 16 blocks. The blocks were generated from a full factorial design cleaned from dominant alternatives and grouped into nine groups based on environment and service criteria. One row per group was drawn at random for each block, selecting blocks with low correlation and high balance between attributes (close to orthogonal design), while at the same time avoiding the same row appearing in multiple blocks. A choice set example is given in Figure 1. The survey was written in QuenchTec.

***Did you know** that without any precautionary measures, last mile deliveries are expected to increase with 30 % and congestion with 20 % in the largest cities by 2030? (Source: World Economic Forum).*

Imagine that you are to attend a birthday party, a wedding or a business meeting that is known to you in advance. You need an outfit, and decide to rent this online. After choosing the rental period, you are asked to choose how to get the outfit delivered.

Select your preferred option. You can take for granted that the outfit arrives at your preferred place of delivery, is tracked in the usual way and delivered free of charge.

	Alternative 1	Alternative 2	Alternative 3
Delivery time	20 days	10 days	I would not shop if these were my only delivery options
Delays	No	Yes, 1-2 days	
Information	Yes, 2 notifications	No	
CO2-emission	0,28 kg	1,40 kg	
Particulate Matter (PM)	Medium	High	

Figure 1: Example of choice set asked in the survey. Supplementary information in dotted box on top.

3.2 Nudge

The supplementary information regarding environmental aspects of the textile industry and last mile deliveries was presented in questions related to habits, attitudes and in the DCE (as shown by the dotted box at the top of Figure 1). The aim was to test if this type of nudge influenced attitudes and preferences, as suggested by previous studies presented in section 2. About half of the respondents, drawn at random, received this supplementary information.

3.3 Survey administration

Prior to distribution, both qualitative (general feedback and one-on-one interviews with representatives from both experts and user group) and quantitative (using data from a pilot survey among the online response panel) pretesting were conducted as recommended (Mansfield and Pattanayak, 2007, Champ and Welsh, 2007, Krupnick and Adamowicz, 2007, Harrison, 2007, Mathews et al., 2007, Johnston et al., 2017). Focus group interviews were not an option due to Covid-19 socializing restrictions. Pre-test results helped design the attribute levels in the DCE as well as fine tune questions and information text.

Originally, the plan was to distribute the survey among receivers of FJONGs newsletter to reach females who were familiar with clothing rentals online, either by first- or secondhand experience. But the Covid-19-pandemic led to a change in plans; a pilot-test indicated respondent fatigue (probably) due to an increase in consumer surveys after the outbreak. To secure an adequate sample size, the survey firm NORSTAT[6] was contacted. The NORSTAT panel consists of 81,000 active panelists, 52% female, evenly spread out on age groups starting from 15 years. Respondents are rewarded for their participation using bonus points. When distributing the survey in question, only females between 18 and 70 years of age were targeted. Survey links were sent out continuously starting June 29th and ending August 3rd 2020. The survey resulted in 595 responses[7]. The frequency of choosing the 3 alternatives presented in the choice sets is provided in Appendix A. The frequency indicates that consumers were able to make tradeoffs between the alternatives, as respondents chose the opt out alternative in only 22% of the choices (i.e. Alternative 3 in Figure 3).

3.4 Sample description and summary statistics

After screening for respondents focusing on only one attribute in the DCE, inconsistencies in debriefing questions (i.e. answer honest and random at the same time), very quick respondents (i.e. speeders) or very slow respondents[8], 82 respondents were removed from the sample. Hence, the dataset for analysis includes 513 observations. Summary statistics are presented in Table 2. Statistics are presented for the whole sample, a subsample of respondents who were familiar with FJONG, and the female Norwegian population. The summary statistics for the Norwegian population differs somewhat in terms of the age group included. This is highlighted in the table using notes. Attitude variables are collected asking respondents to disagree with or agree to statements, using a 5-point Likert scale ranging from "totally disagree" to "totally agree". For statements relating to *purchase planning*, *fashion interest* and *clothing rentals* the responses were converted to

[6] https://norstat.no/

[7] The sample was part of a survey collecting data for two DCEs with respondents being randomly allocated to one of the two DCEs. In total, 4602 links were distributed with a response of 1200 (response rate = 26%), of which 605 respondents answered the other DCE and is not included in this paper. The average time used to complete the survey for all 1200 respondents was 11,5 minutes.

[8] Inspired by Hensher (2007), Mathews et al. (2007), and as done in an example by Alberini et al (2007, p.214).

dummy (binary) variables by merging the "totally agree" and "somewhat agree" responses into the dummy variable "agree" option versus the rest. For the two constructed "change" dummy variables (*Society should change* and *Consumers should change*), only "totally agree" defined the "agree" option. A variable (*Top25pop*) denoting whether the respondents live in one of the 25 (out of 356) most populated Norwegian municipalities (ranging from Oslo with 681,000 residents to Lørenskog with 41,000 residents (Statistics Norway, 2018)) was included to capture those who are likely to have shopping opportunities in their proximity and live in municipalities where local air pollution might be an issue. Income was split at 600,000 NOK[9] to capture potential impacts of high income[10].

The sample seems to be representative of the population with respect to demographics. The exception is education: respondents in the sample are more educated than the population as a whole, reflecting also the composition of the NORSTAT internet panel. Approximately 33% of the sample shop online at least once a month. This share is larger than the 12% found by Bjerkan et al. (2020) and lower than the 69% found by Postnord (2020), both surveying online shopping behavior among male and female Norwegian adult respondents. A comparison of the sample with the subsample containing only those who are familiar with FJONG shows that the latter consists of younger, more urban, more highly educated, and more experienced shoppers with fashion interest and positive attitudes to clothing rentals. Based on information provided by FJONG themselves, their active customers are even younger, more urban, have a higher income and higher participation in the work force.

Figure 2 presents how some of the expected key attributes, familiarity with FJONG and their concept of clothing rentals, positive attitude towards clothing rentals, and environmental concerns, vary with generation and location. In general, young adults and those living in one of the most populous municipalities in Norway are most familiar with FJONG and agree that society should pay more attention to the environment. The share with a positive attitude towards clothing rentals is somewhat lower for people living in the 25 most populous municipalities than for those who do not. Overall, the results are according to expectations, showing that the sharing economy is most popular among the younger generations.

[9] The average exchange rate between Euro (€) and NOK (Kr.) at the time of survey (July 2020) was 1€=10.65Kr./1Kr=0.094€ (Norges Bank - Central bank of Norway).

[10] In 2018, only 15% of Norwegian females had an average annual gross personal income of 600,000 NOK or more (Statistics Norway, Table 08411).

Table 2: Descriptive statistics for demographic and attitudinal variables for the sample, and corresponding demographics for Norwegian female population 18-70 years old (from Statistics Norway). All variables are binary, taking on the values 1 = "Yes" or 0 = "No". The values presented are the share of "ones" in the sample.

	Sample *(N=513)*	***Subsample familiar with FJONG*** *(N=52)*	***Female population***
Average age and generation [a]			
Average age (18-70 years); in years	41.9	38.2	43.7
1997-2001 (Generation Z)	9%	10%	9%[b]
1981-1996 (Millennials)	38%	54%	33%
1965-1980 (Generation X)	33%	27%	32%
1949-1964 (Boomers)	20%	10%	25%
Top25pop			
Lived in one of the 25 most populated Norwegian municipalities.	59%	88%	53%[b]
Education			
Primary school	3%	0%	25%[c]
High school	34%	17%	36%
College or university	63%	83%	39%
Employment status			
Employed	65%	69%	65%[d]
Unemployed	4%	2%	3%
Not in work force (incl. students)	27%	25%	32%
Other	4%	4%	
Annual gross personal income in NOK (2019)			
Average income (based on middle value of intervals)	483,000	461,000	382,000[e]
Less than 600,000 NOK	60%	62%	
More than 600,000 NOK	20%	27%	
NA	20%	12%	
Frequent online shopper			
Shopped online at least once a month	33%	54%	
Purchase planning			
Agree that they like to plan their purchases	78%	79%	
Fashion interest			
Agree that they are interested in fashion	40%	71%	
Clothing rentals			
Agree that clothing rentals provide a fashionable wardrobe as well as being more environmentally sustainable than new clothing purchases	45%	62%	
Society should change			
Totally agree that society should pay more attention to the environmental challenges than we do today	52%	60%	
Consumers should change			
Totally agree that consumers must change attitude and behavior to solve the environmental challenges of today	53%	48%	
Information			
Received supplementary information of environmental aspects of clothing purchase and rentals	47%	54%	

Notes: [a] As defined by Pew Research, [b] Females 18-70 years, [c] Females 16 years or older, [d] Females 15 years or older, [e] Average annual gross income for females 17 years and older in 2018 (2019 numbers are postponed until January 2021). The average income of employed females only was 506,000 NOK in 2018 (Statistics Norway, Table 12851).

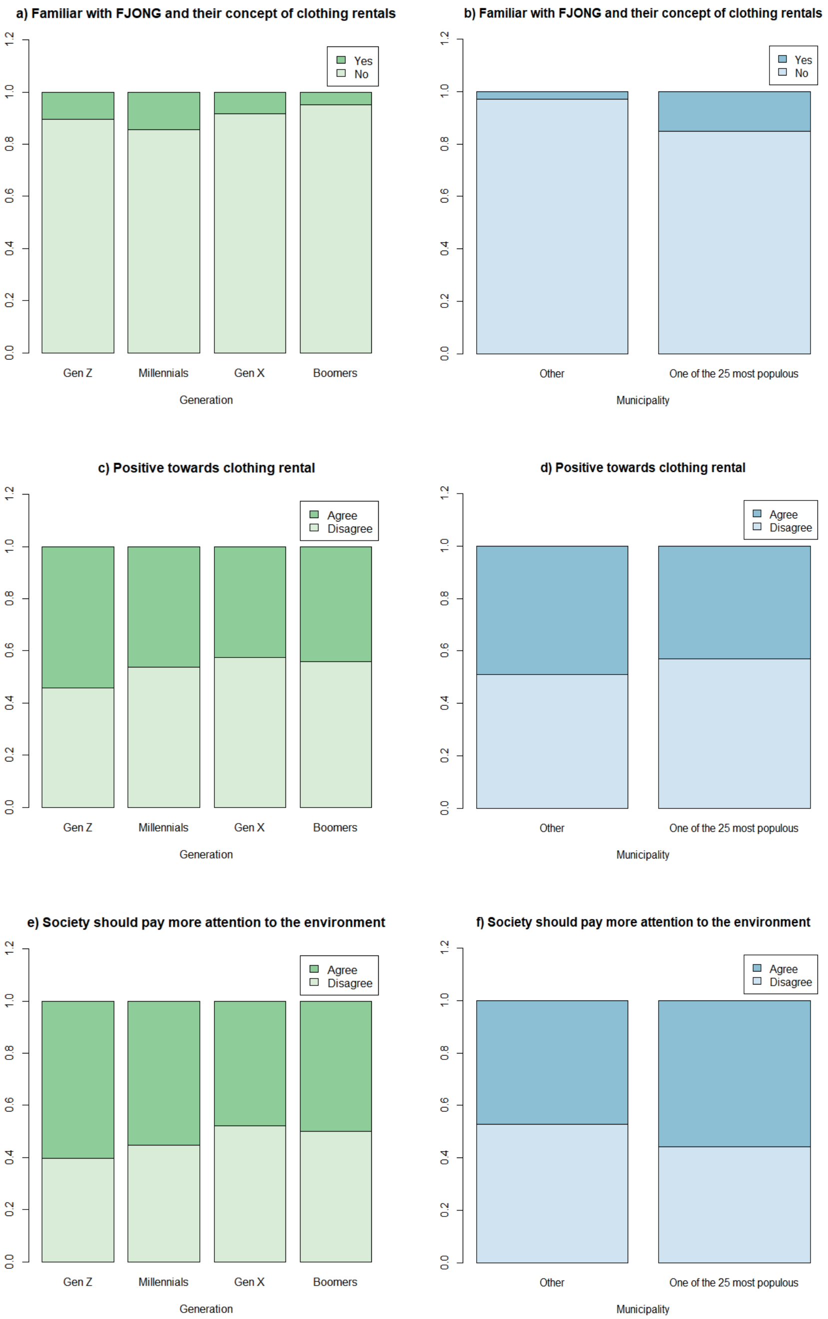

Figure 2 a)-f): visualizations of how knowledge about clothing rentals and attitudes differ with generation and location.

4 Modelling specifications: revealing consumer preferences

Several methods to reveal consumer preferences are available, each with their strengths and weaknesses. Here, a stated preference approach using a DCE was chosen to capture consumer preferences for last mile delivery attributes, and a discrete choice modeling framework was chosen for estimation. The simple Multinomial Logit Model (MNL) provides a good starting point and a useful benchmark for comparison with more complex models (Marcucci and Gatta, 2012, Valeri et al., 2016, Punel and Stathopoulos, 2017, Swait, 2007). The latent class (LC) model allows for a further exploration of individual heterogeneity by splitting the sample in different classes (groups), allowing for group-specific coefficients. MNL in combination with LC is used on several occasions, i.e. to model socioeconomic and attitudinal influence on preferences for environmental policy drivers (Valeri et al., 2016) or consumer preferences for last mile deliveries of e-groceries (Gatta et al., 2020). The next section briefly presents the modeling framework, and is based on the notations by Alberini et al. (2007), Swait (2007) as well as Sarrias and Daziano (2017) for consistency with the gmnl R-package used for estimation.

4.1 Discrete choice modeling framework

Assume N respondents answering T choice tasks, with i alternatives differing in their attribute levels. Respondents are assumed to be rational and utility maximizing individuals choosing the alternative that provides the highest utility, U. Utility consists of a deterministic (observable and estimable) part, $V = V(X, \beta)$, given by a set of observable variables, X (including attributes in the choice tasks and socio-demographic variables), and their corresponding parameters, β, and a stochastic (unobservable) part, called the error term ε. Due to the stochastic part, U is not observable. The probability of observing alternative i over alternative j in choice task t for individual n is written as: $P_{i,t,n} = P\left(U_{i,t,n} \geq U_{j,t,n}\right) = P\left(V_{i,t,n} + \varepsilon_{i,t,n} \geq V_{j,t,n} + \varepsilon_{j,t,n}\right) = P(V_{i,t,n} - V_{j,t,n} \geq \varepsilon_{j,t,n} - \varepsilon_{i,t,n})$. To solve this expression, assumptions must be made about the error term. The most common approach is to assume that $\varepsilon_{i,t,n}$ are independent and identically (IID) distributed following a standard type I extreme (Gumbel) distribution. Under this assumption, the probability of individual n choosing alternative i in choice task t becomes:

$$P_{i,t,n} = \frac{e^{\left(\mu V_{i,t,n}\right)}}{\sum_{j=1}^{J} e^{\left(\mu V_{j,t,n}\right)}}$$

The parameter μ is the scaling parameter, capturing potential heterogeneity in the choices. For identification of the β-parameter in $V_{i,t,n}$ as both μ, β are unknown parameters, μ is often normalized to 1 (at the cost of heterogeneity in the model). This is known as the MNL model.

In the LC model, heterogeneity is included by allowing for group taste variations. Each individual belongs to a group s with w probability, and hence we get coefficients $\beta = \beta_s$ with probability w_{ns}. Group membership also consists of a deterministic and a stochastic part and hence unobservable: $Y_{ns}^{*} = \Gamma_s' Z_n + v_{ns}$.

Where Y^*_{ns} is the membership likelihood function of n being in class s, Z_n is a vector of socio-demographic and attitudinal variables with corresponding parameters Γ'_s, and v_{ns} is the stochastic error term assumed to be IID Gumbel distributed (as ε), and independent of the stochastic conditional utility (V). Hence, if $Y^*_{ns} > Y^*_{ns'}$, $\forall\, s \neq s'\; s, s' = 1, \dots, S$ then n is in class s. The probability of belonging to class s can be estimated, again introducing a scaling parameter λ normalized to one for identification: $W_{ns} = \frac{e^{(\lambda \Gamma'_s Z_n)}}{\sum_{s=1}^{S} e^{\left(\lambda \Gamma'_{s'} Z_n\right)}}$, $w_{ns} > 0, \sum_s w_{ns} = 1$. The probability of individual n choosing alternative i in choice task t is now conditional on class membership (s):

$$P_{i,t,n} = \sum_{s=1}^{S} W_{ns} \cdot P_{i,t,n|s} = \sum_{s=1}^{S} W_{ns} \cdot \frac{e^{\left(V_{i,t,n|s}\right)}}{\sum_{j=1}^{J} e^{\left(V_{j,t,n|s}\right)}}$$

If the replicated choice tasks can be assumed independent within each individual, the joint probability of observing the T choices from the discrete choice experiment are:

$$P_n = \sum_{s=1}^{S} W_{ns} \left(\prod_{t=1}^{T} \frac{e^{\left(V_{i,t,n|s}\right)}}{\sum_{j=1}^{J} e^{\left(V_{j,t,n|s}\right)}} \right)$$

Taking the log-likelihood and summarizing over all individuals provides the log-likelihood function $LL = \sum_{n=1}^{N} \ln\, P_n$ that explicitly recognizes the repeated nature of the data (Swait, 2007), i.e. that the data is panel. Given the right model choice and realistic parameter values, the estimates from this model are known to be asymptotically consistent, unbiased and efficient, and no longer rely on the independence of irrelevant alternatives (IIA) (Swait, 2007). Nested models are compared using the likelihood-ratio (LR) test. Global optimum is found by running several model versions differing in number of underlying classes.

5 Estimation results: consumer preferences for environmentally sustainable last mile deliveries

Two separate models are estimated to understand consumer preferences for environmentally sustainable last mile deliveries: multinomial logit (MNL) and the latent (LC) class models. All models are estimated in R (R Core Team) using the package gmnl (Sarrias, Sarrias and Daziano, 2017). Standard errors are estimated using the sandwich estimator (White's robust alternative). In both models, ten variables are used to control for heterogeneity based on the literature review in section 2. The Pearson correlation between them is investigated, revealing acceptable low levels of correlations (see Appendix B).

5.1 Benchmark models

The MNL model is estimated 21 times; for the full sample and 20 different subsamples, two for each of ten explanatory variables. The only variables in the models are the five attributes from the DCE: delivery time, delays, information service, CO_2 and PM. The idea is to investigate if different groups value attributes differently. This approach also reveals any inconsistencies in the subsamples, for instance due to lack of interest in the topic (e.g. in terms of little e-commerce experience, negative to clothing rentals etc.). The delivery time and PM attributes are found to be non-linear and estimated by piecewise linear approximation. This entails the estimation of different values for different ranges of the selected attribute while maintaining the linear utility function. The base categories are delivery time of 20 days and Low PM. As the models are estimated on different samples, the model fit is not comparable across models. The results are presented in Table 3.

The different subsamples reveal different preferences in the sample. The overall strongest preferences for a reduction in delivery time from 20 days to all other alternatives are found for young adults (Gen Z and Millennials). Young adults also seem to care less for PM-levels than the full sample, but value information services relatively high. Following young adults, frequent online shoppers have the second strongest preference for short delivery time (1 day), followed by respondents with an interest in fashion. However, these consumers have a preference for emission reduction at or slightly above the estimates for the full sample, indicating that emission reduction is also preferred. The lowest time preferences are found for Generation X or Boomers and for respondents who received supplementary information about environmental aspects of the clothing industry. Those receiving supplementary information also accept delays to a higher extent than most groups, suggesting that respondents can be nudged towards more time-flexible deliveries.

The highest disutility of airborne emissions is found for respondents who totally agree that society should pay more attention to the environmental challenges, followed by respondents who totally agree that consumers should change attitude and behavior to solve the environmental challenges of today. These two groups show similar tendencies, which is not surprising due to their high correlations (see Appendix B). The lowest disutility for CO_2-emissions is found for respondents who do not agree that society should pay more attention to the environmental challenges. Residents who are positive towards clothing rentals, have an income above 600,000 NOK and live in one of the top 25 populated municipalities have a strong disutility for particulate matter. Respondents are more negative to high than medium levels of PM.

Debriefing questions inquiring about the importance of the different attributes revealed a somewhat even distribution across the attributes, although delivery time was voted the most important attribute and PM the least (see Appendix D). The Pearson correlation showed, as expected, a positive correlation between the importance of delivery time and delays (0.43), delays and information services (0.40), and CO_2-emission and particulate matter (0.84), confirming the findings from model estimation.

Table 3: Parameter estimates and standard errors (in parenthesis) from estimation of MNL-models on different subsamples with regards to demographic variables (residency, generation, income), whether the respondent received supplementary information in the survey, and frequency of online shopping. The grey cells highlight estimated values at the high and low end (for the given variable) and are commented in the text below. Delivery time: Twenty days and PM Low are base categories for their respective variables.

	Full Sample	*Live in top 25 pop. municipality*		*Gen Z or Millennials*	*Gen X or Boomers*	*Income >600,000 NOK*		*Supplementary information*		*Frequent online shopper*	
		=Yes	= No	= Yes	= Yes	= Yes	*= No*	= Yes	= No	= Yes	= No
Delivery time: One day	1.843*** (0.069)	2.014*** (0.093)	1.616*** (0.105)	2.415*** (0.109)	1.410*** (0.092)	1.792*** (0.159)	1.860*** (0.077)	1.586*** (0.102)	2.073*** (0.097)	2.222*** (0.127)	1.663*** (0.083)
Delivery time: Five days	1.729*** (0.065)	1.794*** (0.085)	1.650*** (0.099)	2.134*** (0.100)	1.425*** (0.086)	1.872*** (0.150)	1.699*** (0.072)	1.588*** (0.091)	1.837*** (0.092)	2.043*** (0.118)	1.588*** (0.078)
Delivery time: Ten days	1.050*** (0.061)	1.114*** (0.080)	0.964*** (0.093)	1.334*** (0.092)	0.846*** (0.082)	1.024*** (0.136)	1.060*** (0.068)	1.069*** (0.091)	1.072*** (0.083)	1.007*** (0.110)	1.076*** (0.073)
Delays	-0.123** (0.040)	-0.114* (0.053)	-0.133* (0.062)	-0.167** (0.060)	-0.091 (0.055)	-0.208* (0.092)	-0.104* (0.044)	-0.038 (0.059)	-0.187*** (0.056)	-0.191** (0.071)	-0.088 (0.049)
Information services	0.218*** (0.037)	0.232*** (0.048)	0.196*** (0.057)	0.246*** (0.055)	0.208*** (0.050)	0.238** (0.083)	0.211*** (0.041)	0.218*** (0.051)	0.185*** (0.053)	0.201** (0.065)	0.226*** (0.044)
CO_2	-0.631*** (0.037)	-0.642*** (0.048)	-0.614*** (0.057)	-0.653*** (0.055)	-0.624*** (0.051)	-0.700*** (0.085)	-0.616*** (0.041)	-0.636*** (0.059)	-0.653*** (0.049)	-0.681*** (0.066)	-0.620*** (0.044)
PM Medium	-0.338*** (0.050)	-0.329*** (0.066)	-0.365*** (0.078)	-0.283*** (0.076)	-0.395*** (0.068)	-0.395*** (0.114)	-0.324*** (0.056)	-0.306*** (0.072)	-0.342*** (0.071)	-0.454*** (0.091)	-0.285*** (0.060)
PM High	-1.068*** (0.057)	-1.185*** (0.076)	-0.922*** (0.086)	-1.063*** (0.085)	-1.096*** (0.077)	-1.217*** (0.129)	-1.034*** (0.063)	-1.033*** (0.084)	-1.054*** (0.079)	-1.072*** (0.100)	-1.077*** (0.069)
Log-likelihood	-4316.17	-2487.627	-1819.033	-1849.56	-2399.866	-861.238	-3450.885	-2107.495	-2198.41	-1377.762	-2920.279
N	4617	2727	1890	2178	2439	918	3699	2187	2430	1539	3078
McFadden pseudo R^2	0.123	0.571	0.552	0.494	0.629	0.623	0.511	0.824	0.298	0.719	0.406

Significance: *** = $p < 0.001$; ** = $p < 0.01$; * = $p < 0.05$

Table 3 cont.: Parameter estimates and standard errors (in parenthesis) from estimation of MNL-models on different subsamples with regards to attitudes towards clothing rentals, environment, purchase planning and fashion. The grey cells highlight estimated values at the high and low end (for the given variable) and are commented in the text below. Delivery time: Twenty days and PM Low are base categories for their respective variables.

	Clothing rentals are beneficial		Society should change		Consumers should change		Like purchase planning		Have a fashion interest	
	= Yes	= No	= Yes	= No	= Yes	= No	= Yes	= No	= Yes	= No
Delivery time: One day	1.853*** (0.103)	1.854*** (0.095)	1.787*** (0.096)	1.920*** (0.101)	1.713*** (0.095)	1.994*** (0.102)	1.776*** (0.078)	2.128*** (0.152)	2.160*** (0.114)	1.643*** (0.088)
Delivery time: Five days	1.695*** (0.096)	1.780*** (0.088)	1.688*** (0.091)	1.805*** (0.092)	1.680*** (0.090)	1.804*** (0.094)	1.671*** (0.073)	1.980*** (0.144)	1.910*** (0.105)	1.620*** (0.082)
Delivery time: Ten days	1.118*** (0.091)	0.995*** (0.082)	1.136*** (0.085)	0.977*** (0.087)	1.100*** (0.084)	1.014*** (0.088)	1.015*** (0.068)	1.193*** (0.133)	1.186*** (0.099)	0.967*** (0.077)
Delays	-0.024 (0.059)	-0.202*** (0.055)	-0.104 (0.056)	-0.133* (0.057)	-0.160** (0.056)	-0.068 (0.057)	-0.152*** (0.046)	-0.021 (0.084)	-0.162* (0.064)	-0.095 (0.051)
Information services	0.184*** (0.055)	0.249*** (0.049)	0.343*** (0.051)	0.088 (0.053)	0.320*** (0.050)	0.106* (0.054)	0.241*** (0.042)	0.131 (0.078)	0.252*** (0.059)	0.197*** (0.047)
CO_2	-0.615*** (0.054)	-0.660*** (0.050)	-0.815*** (0.054)	-0.465*** (0.051)	-0.736*** (0.053)	-0.546*** (0.052)	-0.662*** (0.042)	-0.545*** (0.077)	-0.655*** (0.059)	-0.617*** (0.047)
PM Medium	-0.233** (0.073)	-0.437*** (0.069)	-0.326*** (0.070)	-0.356*** (0.073)	-0.345*** (0.068)	-0.325*** (0.075)	-0.349*** (0.057)	-0.304** (0.107)	-0.325*** (0.080)	-0.352*** (0.064)
PM High	-1.231*** (0.087)	-0.948*** (0.075)	-1.116*** (0.080)	-1.047*** (0.082)	-1.160*** (0.080)	-0.988*** (0.081)	-1.026*** (0.064)	-1.249*** (0.127)	-1.177*** (0.092)	-1.008*** (0.072)
Log-likelihood	-1933.521	-2367.283	-2240.109	-2050.862	-2304.100	-1988.337	-3399.445	-903.180	-1658.117	-2646.567
N	2097	2520	2412	2205	2448	2169	3600	1017	1854	2763
McFadden pseudo ρ^2	0.606	0.518	0.544	0.582	0.531	0.595	0.309	0.815	0.662	0.461

Significance: *** = $p < 0.001$; ** = $p < 0.01$; * = $p < 0.05$

5.2 Heterogenous preference

The latent class (LC) model is estimated using the same attributes and explanatory variables as above. A model with 4 classes was chosen based on test criteria and an assessment of the plausibility of the results (see Appendix C). The results are shown in Table 4. The top part explains group membership, while the lower part shows their preferences for last mile delivery attributes. The LC model can be compared with the full sample model in Table 3. A likelihood ratio-test reveals that the LC model is preferred over the MNL model (λ_{LR}=1272.98), advocating the inclusion of observed and unobserved heterogeneity across groups. In general, the estimation results show that attitudes (*clothing rentals*, *consumer should change* for the environment, *society should change* for the environment, and interest in *purchase planning*) are important and reveal more about the different classes than demographic variables. This is consistent with the literature review presented in section 2.

Table 4: Coefficients from estimation of latent class model with 4 classes (standard errors in parenthesis). Delivery time: Twenty days and PM Low are base categories for their respective variables.

	Class 1	*Class 2*	*Class 3*	*Class 4*
Group membership				
Intercept		-0.720***	0.295*	-0.879***
		(0.183)	(0.124)	(0.170)
Gen Z or Millennials		0.074	-0.073	0.292**
		(0.117)	(0.085)	(0.104)
Top 25 populated municipalities		0.032	0.012	0.341**
		(0.112)	(0.085)	(0.108)
Income over 600,000 NOK		0.438**	0.030	0.271
		(0.141)	(0.109)	(0.145)
Supplementary information		0.047	0.169*	0.078
		(0.114)	(0.083)	(0.104)
Frequent online shopper		0.356**	-0.122	0.057
		(0.116)	(0.093)	(0.110)
Clothing rentals		-0.299**	-0.004	0.512***
		(0.115)	(0.086)	(0.105)
Consumers should change		0.927***	0.515***	0.260
		(0.141)	(0.106)	(0.140)
Society should change		-0.353**	-0.654***	-0.767***
		(0.136)	(0.106)	(0.131)
Purchase planning		-0.305*	-0.481***	0.063
		(0.135)	(0.101)	(0.134)
Fashion interest		-0.293*	0.042	-0.194
		(0.122)	(0.089)	(0.107)
Group shares	*31%*	*15%*	*41%*	*13%*
Last mile delivery attributes				
Delivery time: One day	1.607*** (0.154)	1.310*** (0.349)	5.755*** (0.466)	1.066*** (0.221)
Delivery time: Five days	1.126*** (0.144)	2.177*** (0.305)	4.583*** (0.348)	1.775*** (0.198)
Delivery time: Ten days	0.418** (0.135)	1.678*** (0.254)	2.633*** (0.229)	1.553*** (0.165)
Delays	-0.515***	0.049	-0.315*	0.107
	(0.100)	(0.171)	(0.129)	(0.112)
Information services	-0.026	1.211***	-0.140	0.665***
	(0.091)	(0.255)	(0.116)	(0.117)
CO_2	-1.022***	-2.109***	-0.836***	0.122
	(0.101)	(0.451)	(0.142)	(0.150)
PM Medium	-1.111***	0.482	-0.588**	0.096
	(0.116)	(0.262)	(0.186)	(0.137)
PM High	-2.086***	-0.737**	-1.237***	-0.917***
	(0.157)	(0.271)	(0.187)	(0.191)
Log-likelihood	-3679.68			

AIC	7489.36
BIC	7907.80
McFadden pseudo R^2	0.232
N	4617

Significance: *** = $p < 0.001$; ** = $p < 0.01$; * = $p < 0.05$

The four classes are found to differ. Class 1 (31% of the sample) consists of individuals that do not stand out with respect to demography or online shopping habits. Instead, they tend to totally agree that society, not consumers, should pay more attention to the environmental challenges. They also seem to plan their purchases more than class 2 and 3. Looking at the last mile delivery service attributes, this class prefers shorter delivery time to longer (although the difference in parameter estimates for five- and one-day delivery time is significant only at the 10% level (α=0.087)), and has a negative utility of delays. This group has the overall largest disutility of CO_2- and PM-emissions, thus requesting both a high level of delivery services and environmentally sustainable solutions.

Class 2 (15% of the sample) consists of respondents with higher income and more frequent online shoppers than the other three groups. They are negative to clothing rentals and strongly believe in change for the environment, with consumers being the most important agent. They do not plan their purchases and are not fashion-interested compared to the other groups. These respondents prefer five days' delivery time to one day, and insignificant parameter difference between ten and five days and ten and one day indicates acceptance of ten days' delivery time. This group wants information services and has a strong disutility of CO_2-emission. PM is less important.

Class 3 (41% of the sample) is the largest class and resembles class 1 but believes consumers should change for the environment rather than society, does not plan their purchases and received supplementary information in the survey. This group is also more time sensitive but has a smaller disutility of delays and emissions than class 1 (although still negative and significant). The difference in parameter size for PM Medium and PM High is significant only at α=0.081. The lack of purchase planning might explain the time sensitivity, and supplementary information might have influenced the importance of emission reduction.

The last class, class 4 (13% of the sample), is the smallest class and a bit younger and more urban than the others. They are positive to clothing rentals but not to change for the environment. Five-day delivery time provides the highest utility, although the difference between ten and five days is insignificant (the difference between ten and one day is significant at α=0.085). This group also prefers information services, does not care about CO_2-emission or medium PM-levels, but has a negative disutility of high PM-levels. Overall there seems to be a lack of clear preferences in this class which might indicate that the respondents care less about last mile delivery attributes than in the other classes.

Supplementary information is found to have an influence in class 3. However, to investigate the importance of this variable, the LC model with 4 classes is estimated without the supplementary information variable. A likelihood-ratio test shows that the model without supplementary information is not rejected in favor of the model with supplementary information (λ_{LR}=0.478). Together with only small changes in the parameter estimates, this indicates that supplementary information is not significant in explaining the variation in preferences.

Although the modelling efforts clearly show that there is unobserved heterogeneity in the data, the McFadden pseudo R^2 is small. This might be explained by behavioral intentions. One of the main findings from Hamari et al. (2016) is the discrepancy between reported attitude and actual behavior indicating that consumers tend to overreport their interest in sharing economy and environment. Godelnik (2017) discloses that students are claiming that environmental motivation is more important than social motivation, but their behavior reveals the opposite. One reason might be that consumers are not ready to prioritize the environment over their personal wellbeing and pleasure, and that society is still encouraging non-sustainable lifestyles (ElHaffar et al., 2020). Although efforts have been made in this work to reveal unmotivated consumers and non-rational answers in the DCE, specific questions about behavior intention could increase the models' explanatory power.

6 Discussion

Based on the estimation results above, what can be said about *consumers' concern with airborne emissions from last mile delivery, increased delivery time as an emission mitigating tool, and the role of supplementary information in explaining preferences*, as asked in the research question?

When little consumer information is available or when online retailers can target consumers directly, the MNL might provide enough information about last mile delivery preferences. Some examples are highlighted in Table 5 together with an assessment of the impact on last mile delivery time. Such knowledge might allow online retailers and transport operators to design last mile delivery alternatives that best fit the customer base or each customer, thus achieving environmentally sustainable deliveries at low cost.

Table 5: Consumer type, preferences and acceptance of increased delivery time when little information on consumer segment is available or individual consumers can be targeted directly.

Consumer type	Preferences	Accept increased delivery time as a measure to reduce emissions?
Essentially younger than 40 years	Time sensitive, prefer both quick delivery, avoidance of delay and information about the shipment	Unlikely. Zero-emission vehicles (electric vans, bicycles, drones/droids) with little change in the offered delivery solution might be an option.
Essentially older than 40 years	Time is of less importance, although 20 days is not preferred	Likely. This is probably the least time sensitive group of consumers.
Has income above average	Time is of average importance, with 5 days' delivery being preferred to one day. Emission, particularly PM, should be highlighted.	Likely. Consolidation might be an option, but also crowdshipping, which might take more time but have the potential to reduce emissions from last mile transport.
Shop at least once a month	Consumers are both time sensitive and negative to emissions	Less likely. Investments in technology that increases efficiency and reduced emissions might pay off if consumers return to website.
Consumers believe in change for the environment	Less time sensitive, accept longer delivery time for reduced emissions. Prefer information services	Very likely if the environmental aspect is highlighted. Expected to be flexible towards the delivery solutions: zero-emission vehicles like bicycles, crowdshipping, pick-up points etc.

When more information is available about the consumers, or when group characteristics are of interest, the LC model provides a broader picture. This might be useful for online retailers deciding on their last mile delivery service strategies but also for public planners seeking to achieve more sustainable urban freight traffic. The main findings and their implications are presented in Table 6.

Table 6: Consumer type, preferences and potential measures when more information on consumer segment is available, or group characteristics are of interest.

Consumer type	Preference	Accept increased delivery time as a measure to reduce emissions?
The average consumer (72% of total number of respondents)	Strict delivery requirements in terms of time and emission levels …	
Believe society, not consumers, should act on behalf of the environment (31%)	… with little room for change from the consumer	Unlikely. Prefer quick and environmentally sustainable deliveries. Zero-emission technology with quick delivery (electric vans, bikes) might pay off.
Believe consumers should change for the environment. Do not plan purchases, has the highest share of consumers who received supplementary information in the survey (41%).	… which might be possible to influence through nudges or other carrots that inspire environmentally sustainable behavior	Likely if the consumer is informed about the environmental consequences of their actions
High income, frequent online shoppers who believe consumers must bear the largest burden of changing for the environment. Less interested in fashion, clothing rentals or planning (15%)	Accept longer delivery time if compensated with reduced emissions (particularly CO_2) and information services	Likely. Consumers are expected to choose the more environmentally sustainable service if the reduced emissions and expected day of arrival are clearly communicated. Expected to be flexible in terms of measures.
Young, urban, positive to clothing rentals (13%)	Seem to care less about the last mile delivery solutions than the other groups, accept longer delivery time if information service is provided. Emissions are of less concern.	Uncertain. Can be tested by only offering deliveries that are environmentally sustainable.

The overall result is that (some) consumers accept a delivery time of 5 or even 10 days if compensated in terms of reduced airborne emissions from transport and/or information services. No one seemed to prefer 20 days of delivery time, indicating that this might be a stretch even for the environmentally conscious consumer. Another finding is that one consumer group seems to care less about last mile delivery attributes than the others. Buldeo Rai et al. (2021) also identified a group with neutral positions towards environmental sustainability and last mile innovations. This result is interesting and might indicate that last mile delivery services are of less importance to some consumers than perceived by retailers and freight operators.

The findings above clearly show that consumers prefer reduction in airborne emissions from last mile deliveries. Thus, online retailers and freight operators should offer consumers environmentally sustainable last mile delivery solutions when they shop online. Sustainable urban freight transport should also be on the agenda for governmental authorities and public planners. Governmental authorities might suggest that emissions from last mile deliveries are communicated to online retailers and consumers, either by stating the emissions explicitly or through eco-labels or their like. The latter are found to have the potential to reduce both emissions and costs from last mile deliveries. (Agatz et al., 2020), thus bringing positive effects to both consumers and freight operators. Some consumers accept longer delivery times, and as such increase freight carriers' flexibility to find low cost environmentally sustainable delivery solutions, like consolidation, use of pick-up points or crowdshipping. By tailoring solutions to different consumer preferences, more accurate measures can be provided. However, the bundle of alternatives should be carefully considered and communicated, as people tend to choose short delivery time when this is offered (Allen et al., 2018). Although not significant in the overall model, supplementary information (nudge) is significant for one of the classes, indicating that information might work as a nudge to increase environmental behavior for some consumers.

To get an idea of the transferability of the results, debriefing DCE questions included evaluation of the attributes for subscription rentals in addition to the one-time rentals presented in the DCE. The answers are presented in Appendix D, and reveal a similar distribution as for the one-time rentals, also found in Appendix D. However, delays, information services, and air emissions are rated somewhat more important for subscriptions. The extension from clothing rentals to other parts of the sharing economy is not investigated. Based on the findings by Diamantopoulos et al. (2003), an all-female sample with high education might overestimate the importance of air emissions in last mile delivery if transferred to the overall population, but the effect of demographics is inconclusive. It would be interesting to compare the results with an all-male sample or a mixed-gender sample to test potential differences.

7 Conclusion

"Anticipating new technologies is one of the critical components to improve performance and decrease negative impacts of goods deliveries" (Punel and Stathopoulos (2017) paraphrasing from Macharis and Kin (2017)). The sharing economy is one of many consequences of emerging technologies, bringing us towards new consumption and freight patterns. For the part of the sharing economy where renting of goods is the key concept, freight traffic is highly influenced as both deliveries and returns are generated. This paper has analyzed consumer preferences for environmentally sustainable last mile deliveries from online clothing rentals. A particular focus was put on increased delivery time as an emission mitigating tool. The influence of supplementary information was also investigated to see if consumers can be nudged towards desirable actions. To answer these questions, discrete choice modeling was conducted on data from a discrete choice experiment on an all-female sample in a survey inquiring about online shopping and clothing rental habits, preferences and attitudes.

The estimation results showed that respondents have a negative utility of delivery time, delays, PM- and CO_2-emissions, and a positive utility from information services.

Although varying in size and significance between different respondent types, some were found to be willing to wait 5-10 days to reduce PM and CO_2. Thus, presenting consumers with environmentally sustainable delivery options might contribute to more sustainable urban freight transport. One of the respondent groups responded to supplementary information (nudge), but the variable was irrelevant in the overall model.

This paper provides new evidence on the preferences for last mile delivery attributes in the sharing economy as well as for last mile delivery attributes when free delivery is considered. Free delivery is expected by many customers and offered by many online retailers, and is thus a relevant aspect of online shopping. Suggested future research relates to expanding the sample to cover males, other parts of the sharing economy as well as modeling efforts to understand even more of the consumer preferences.

Acknowledgements: The author would like to thank FJONG for constructive inputs to the survey, Jens Bengtsson for valued comments and feedback during data collection, writing of the first draft and revisions, Jardar Andersen for manuscript revision, and Anna Herzog for proofreading.

Funding: This work is undertaken as part of the research project 250432 NORSULP (Sustainable Urban Logistics Plans in Norway), financed by the Research Council of Norway and the Norwegian Public Roads Administration. Ståle Navrud's work on this paper is funded by the project "Fjong 2025 - an endless, sustainable wardrobe. Powered by research in behavioral economy, environmental impact and artificial intelligence" (Research Council of Norway project no. 309977).

Conflict of interest: None.

References

AGATZ, N., FAN, Y. & STAM, D. 2020. Going green: the effect of green labels on delivery time slot choices. *Available at SSRN 3656982*.

ALBERINI, A., LONGO, A. & VERONESI, M. 2007. Basic Statistical Models For Stated Choice Studies. *Valuing Environmental Amenities Using Stated Choice Studies: A Common Sense Approach to Theory and Practice.* Arlington, Virginia, USA: Springer.

ALLEN, J., PIECYK, M., PIOTROWSKA, M., MCLEOD, F., CHERRETT, T., GHALI, K., NGUYEN, T., BEKTAS, T., BATES, O., FRIDAY, A., WISE, S. & AUSTWICK, M. 2018. Understanding the impact of e-commerce on last-mile light goods vehicle activity in urban areas: The case of London. *Transportation Research Part D: Transport and Environment,* 61**,** 325-338.

BELK, R. 2014. You are what you can access: Sharing and collaborative consumption online. *Journal of Business Research,* 67**,** 1595-1600.

BJERKAN, K. Y., BJØRGEN, A. & HJELKREM, O. A. 2020. E-commerce and prevalence of last mile practices. *Transportation Research Procedia,* 46**,** 293-300.

BRING RESEARCH 2019. Slik velger kundene deg. *In:* NORGE, P. (ed.).

BULDEO RAI, H., VERLINDE, S. & MACHARIS, C. 2021. Who is interested in a crowdsourced last mile? A segmentation of attitudinal profiles. *Travel Behaviour and Society,* 22**,** 22-31.

CARBONE, V., ROUQUET, A. & ROUSSAT, C. 2018. A typology of logistics at work in collaborative consumption. *International Journal of Physical Distribution & Logistics Management,* 48**,** 570-585.

CERUTTI, P. S., MARTINS, R. D., MACKE, J. & SARATE, J. A. R. 2019. "Green, but not as green as that": An analysis of a Brazilian bike-sharing system. *Journal of Cleaner Production,* 217**,** 185-193.

CHAMP, P. & WELSH, M. 2007. Survey Methodologies for Stated-Choice Studies. *Valuing Environmental Amenities Using Stated Choice Studies: A Common Sense Approach to Theory and Practice.* Arlington, Virgina, USA: Springer.

CHOI, T.-M. & HE, Y. 2019. Peer-to-peer collaborative consumption for fashion products in the sharing economy: Platform operations. *Transportation Research Part E: Logistics and Transportation Review,* 126**,** 49-65.

COLLINS, A. 2015. Behavioural Influences on the Environmental Impact of Collection/Delivery Points. *In:* FAHIMNIA B., BELL M., HENSHER D. & J., S. (eds.) *Green Logistics and Transportation.* Springer, Cham.

DABLANC, L., MORGANTI, E., ARVIDSSON, N., WOXENIUS, J., BROWNE, M. & SAIDI, N. 2017. The rise of on-demand 'Instant Deliveries' in European cities. *Supply Chain Forum: An International Journal,* 18**,** 203-217.

DE OLIVEIRA, L. K., MORGANTI, E., DABLANC, L. & DE OLIVEIRA, R. L. M. 2017. Analysis of the potential demand of automated delivery stations for e-commerce deliveries in Belo Horizonte, Brazil. *Research in Transportation Economics,* 65**,** 34-43.

DIAMANTOPOULOS, A., SCHLEGELMILCH, B. B., SINKOVICS, R. R. & BOHLEN, G. M. 2003. Can socio-demographics still play a role in profiling green consumers? A review of the evidence and an empirical investigation. *Journal of Business Research,* 56**,** 465-480.

ELHAFFAR, G., DURIF, F. & DUBÉ, L. 2020. Towards closing the attitude-intention-behavior gap in green consumption: A narrative review of the literature and an overview of future research directions. *Journal of Cleaner Production,* 275**,** 122556.
GATTA, V., MARCUCCI, E., PIRA, M. L., INTURRI, G., IGNACCOLO, M. & PLUCHINO, A. 2020. E-groceries and urban freight: Investigating purchasing habits, peer influence and behaviour change via a discrete choice/agent-based modelling approach. *Transportation Research Procedia,* 46**,** 133-140.
GODELNIK, R. 2017. Millennials and the sharing economy: Lessons from a 'buy nothing new, share everything month' project. *Environmental Innovation and Societal Transitions,* 23**,** 40-52.
HAMARI, J., SJÖKLINT, M. & UKKONEN, A. 2016. The sharing economy: Why people participate in collaborative consumption. *Journal of the Association for Information Science and Technology,* 67**,** 2047-2059.
HARRISON, G. 2007. Making Choice Studies Incentive Compatible. *Valuing Environmental Amenities Using Stated Choice Studies: A Common Sense Approach to Theory and Practice.* Arlington, Virginia, USA: Springer.
HARTL, B., SABITZER, T., HOFMANN, E. & PENZ, E. 2018. "Sustainability is a nice bonus" the role of sustainability in carsharing from a consumer perspective. *Journal of Cleaner Production,* 202**,** 88-100.
HAWLITSCHEK, F., TEUBNER, T. & GIMPEL, H. 2018. Consumer motives for peer-to-peer sharing. *Journal of Cleaner Production,* 204**,** 144-157.
HENSHER, D. 2007. Attribute Processing in Choice Experiments and Implications on Willingness to Pay. *Valuing Environmental Amenities Using Stated Choice Studies: A Common Sense Approach to Theory and Practice.* Arlington, Virginia, USA: Springer.
HESHMATI, S., VERSTICHEL, J., ESPRIT, E. & VANDEN BERGHE, G. 2018. Alternative e-commerce delivery policies. *EURO Journal on Transportation and Logistics*.
JIA, F., YIN, S., CHEN, L. & CHEN, X. 2020. The circular economy in the textile and apparel industry: A systematic literature review. *Journal of Cleaner Production,* 259**,** 120728.
JOHNSON, F., KANNINEN, B. & BINGHAM, M. 2007. Experimental Design For Stated-Choice Studies. *Valuing Environmental Amenities Using Stated Choice Studies: A Common Sense Approach to Theory and Practice.* Arlington, Virginia, USA: Springer.
JOHNSTON, R. J., BOYLE, K. J., ADAMOWICZ, W. V., BENNETT, J., BROUWER, R., CAMERON, T. A., HANEMANN, W. M., HANLEY, N., RYAN, M., SCARPA, R., TOURANGEAU, R. & VOSSLER, C. A. 2017. Contemporary Guidance for Stated Preference Studies. *Journal of the Association of Environmental and Resource Economists,* 4**,** 319-405.
JOHNSTON, R. J., SCHULTZ, E. T., SEGERSON, K., BESEDIN, E. Y. & RAMACHANDRAN, M. 2012. Enhancing the Content Validity of Stated Preference Valuation: The Structure and Function of Ecological Indicators. *Land Economics,* 88**,** 102-120.
KRUPNICK, A. & ADAMOWICZ, W. 2007. Supporting Questions in Stated Choice Studies. *Valuing Environmental Amenities Using Stated Choice Studies: A*

Common Sense Approach to Theory and Practice. Arlington, Virginia, USA: Springer.

LIM, S. F. W. T., JIN, X. & SRAI, J. S. 2018. Consumer-driven e-commerce: A literature review, design framework, and research agenda on last-mile logistics models. *International Journal of Physical Distribution & Logistics Management,* 48**,** 308-332.

MACHARIS, C. & KIN, B. 2017. The 4 A's of Sustainable City Distribution: Innovative Solutions and Challenges ahead. *International Journal of Sustainable Transportation,* 11**,** 59-71.

MANERBA, D., MANSINI, R. & ZANOTTI, R. 2018. Attended Home Delivery: reducing last-mile environmental impact by changing customer habits. *IFAC-PapersOnLine,* 51**,** 55-60.

MANSFIELD, C. & PATTANAYAK, S. 2007. Getting Started. *In:* KANNINEN, B. (ed.) *Valuing Environmental Amenities Using Stated Choice Studies: A Common Sense Approach to Theory and Practice.* Arlington, Virginia, USA: Springer.

MARCUCCI, E. & GATTA, V. 2012. Dissecting preference heterogeneity in consumer stated choices. *Transportation Research Part E: Logistics and Transportation Review,* 48**,** 331-339.

MATHEWS, K., FREEMAN, M. & DESVOUSGES, W. 2007. How and How Much? *Valuing Environmental Amenities Using Stated Choice Studies: A Common Sense Approach to Theory and Practice.* Arlington, Virginia, USA: Springer.

MAZAREANU, E. 2019. Value of the global sharing economy 2014-2025. *In:* STATISTA (ed.). https://www.statista.com/statistics/830986/value-of-the-global-sharing-economy/#statisticContainer: Statista.

MORGANTI, E., SEIDEL, S., BLANQUART, C., DABLANC, L. & LENZ, B. 2014. The Impact of E-commerce on Final Deliveries: Alternative Parcel Delivery Services in France and Germany. *Transportation Research Procedia,* 4**,** 178-190.

PISCICELLI, L., COOPER, T. & FISHER, T. 2015. The role of values in collaborative consumption: insights from a product-service system for lending and borrowing in the UK. *Journal of Cleaner Production,* 97**,** 21-29.

POLINORI, P., MARCUCCI, E., GATTA, V., BIGERNA, S., BOLLINO, C. A. & MICHELI, S. 2018. Eco-labeling and sustainable urban freight transport: How much are people willing to pay for green logistics? *Rivista Internazionale di Economia dei Transporti / International Journal of Transport Economics,* XLV**,** 631-658.

POSTNORD 2020. Netthandel i Norden - Oppsummering 2019. *In:* POSTNORD (ed.) *Netthandel i Norden.* PostNord.

PUNEL, A., ERMAGUN, A. & STATHOPOULOS, A. 2018. Studying determinants of crowd-shipping use. *Travel Behaviour and Society,* 12**,** 30-40.

PUNEL, A. & STATHOPOULOS, A. 2017. Modeling the acceptability of crowdsourced goods deliveries: Role of context and experience effects. *Transportation Research Part E: Logistics and Transportation Review,* 105**,** 18-38.

R CORE TEAM R: A Language and Environment for Statistical Computing and Graphics. Vienna, Austria: R Foundation for Statistical Computing

RINNE, A. 2019. 4 big trends for the sharing economy in 2019. *World Economic Forum Digital Economy and New Value Creation.*

https://www.weforum.org/agenda/2019/01/sharing-economy/: World Economic Forum.

SARRIAS, M. *FAQ: Latent Class Multinomial Logit Model using gmnl Package* [Online]. https://rpubs.com/msarrias1986/335556: RStudio. Available: https://rpubs.com/msarrias1986/335556 [Accessed 03.09 2020].

SARRIAS, M. & DAZIANO, R. 2017. Multinomial Logit Models with Continuous and Discrete Individual Heterogeneity in R: The gmnl Package. *Journal of Statistical Software,* 72**,** 1-46.

SCHNIEDERJANS, D. G. & STARKEY, C. M. 2014. Intention and willingness to pay for green freight transportation: An empirical examination. *Transportation Research Part D: Transport and Environment,* 31**,** 116-125.

STANDING, C., STANDING, S. & BIERMANN, S. 2019. The implications of the sharing economy for transport. *Transport Reviews,* 39**,** 226-242.

STATISTICS NORWAY. 2018. *Norges 100 mest folkerike kommuner* [Online]. https://www.ssb.no/befolkning/artikler-og-publikasjoner/norges-100-mest-folkerike-kommuner: Statistics Norway. [Accessed 12.10 2020].

SWAIT, J. 2007. Advanced Choice Models. *Valuing Environmental Amenities Using Stated Choice Studies: A Common Sense Approach to Theory and Practice.* Arlington, Virginia, USA: Springer.

THALER, R. H. & SUNSTEIN, C. R. 2009. *Nudge: Improving decisions about health, wealth and happiness,* London, Penguin Books.

VALERI, E., GATTA, V., TEOBALDELLI, D., POLIDORI, P., BARRATT, B., FUZZI, S., KAZEPOV, Y., SERGI, V., WILLIAMS, M. & MAIONE, M. 2016. Modelling individual preferences for environmental policy drivers: Empirical evidence of Italian lifestyle changes using a latent class approach. *Environmental Science & Policy,* 65**,** 65-74.

VISSER, J., NEMOTO, T. & BROWNE, M. 2014. Home Delivery and the Impacts on Urban Freight Transport: A Review. *Procedia - Social and Behavioral Sciences,* 125**,** 15-27.

WILEN, A. 2019. H&M Tests Renting Clothes to Address Environment Concern. *Bloomberg.* https://www.bloomberg.com/news/articles/2019-11-29/h-m-tests-renting-clothes-as-fashion-faces-environment-concern: Bloomberg.

WORLD ECONOMIC FORUM 2020. The Future of the Last-Mile Ecosystem. *In:* FORUM, W. E. (ed.) *Transition Roadmaps for Public- and Private-Sector Players.* www.weforum.org.

Appendix A

Frequency of choices from the discrete choice experiment. Each of the 513 respondents received 9 choice sets, resulting in 4,617 observations.

Table A1: Frequency of choices

Alternative	1	2	3
Frequency	1723	1873	1021

The table reveals a good mix of choice alternative 1 and 2, and 22% of the choice sets resulting in choice of alternative 3.

Appendix B

The ten variables used in estimations were checked for correlation together with a variable for time use on the discrete choice experiments and propensity to swap between the answers agree-disagree-do not know on a repeated question. The last two variables were added to rule out that some consumer group answered more at random than others. In general, there are low correlations between variables. The exception is the high, positive correlation between *society should change* and *consumers should change* for the environment, as well as the negative correlation between the aforementioned variables and *Swap (highlighted in grey in Table B1).* This correlation is not surprising and indicates that a fair share of the respondents think all should contribute (both consumers and society) to solve the environmental issues of today, and that respondents who strongly agree change is necessary for the environment are more consistent in their answers.

Table B1: Correlation between variables explaining heterogeneity in the model, plus a variable for time use.

	Frequent online shopper	*Clothing rentals*	*Society should change*	*Consumers should change*	*Purchase planning*	*Fashion interest*	*Top 25 pop*	*Gen Z or Millennials*	*Income over 600,000 NOK*	*Time use DCE*	*Swap*
Information	-0.01	0.00	0.06	0.11	0.02	0.04	0.00	-0.04	0.04	0.08	-0.01
Frequent online shopper		-0.01	-0.08	-0.07	-0.06	0.24	0.14	0.14	0.10	0.05	0.03
Clothing rentals			0.15	0.17	0.09	0.15	-0.06	0.05	-0.02	0.00	-0.12
Society should change				0.60	0.09	0.00	0.08	0.07	0.02	-0.01	-0.30
Consumers should change					0.15	-0.04	-0.01	0.01	-0.02	-0.02	-0.23
Purchase planning						0.08	0.05	0.10	-0.11	-0.14	-0.08
Fashion interest							0.16	0.16	0.02	0.07	-0.05
Top 25 pop								0.12	-0.01	-0.09	-0.03
Gen Z or Millennials									-0.10	0.00	-0.08
Income over 600,000 NOK										0.05	-0.03
Time use DCE											0.03

Appendix C

Models with 4, 5 and 8 classes (highlighted in grey in Table C1) were investigated for inclusion in the paper using test criteria based on the suggestions by Swait (2006) and Sarrias & Dazanio (2017 – gmnl). All inclusion criteria balance the reduction in the log-likelihood function with the increase in the number of parameters. In addition, the models were evaluated in terms of class characteristics, class shares and plausibility of results. Compared to the model with 4 classes, both the models with 5 and 8 classes resulted in difficulties distinguishing between classes as well as some classes that made little sense. With 8 classes, three of five groups had a share below 8%. One class even had a share below 1%.

Table C1: Measures for Selecting Number of Classes in Latent Class Models, based on Swait, 2006.

Number of classes	*K*	*Estimation ended in*	*LogL*	*AIC- 2(LogL-K)*	*AIC3- 2logL+3K*	*BIC LogL+(K*LogN)/2*	*BIC (from R gmnl)*	*McFadden pseudo ρ^2 1-((Log(full)-K)/LogL*
2	27	Successful convergence	-3906.70	7867.41	7894.41	3956.17	8041.22	0.202
3	46	Successful convergence	-3739.85	7571.70	7617.70	3824.13	7867.83	0.232
4	65	Successful convergence	-3679.68	7489.36	7554.36	3798.77	7907.80	0.240
5	84	Successful convergence	-3636.98	7441.97	7525.97	3790.89	7982.72	0.245
6	103	Iteration limits exceeded	-3612.14	7430.28	7533.28	3800.85	8093.34	0.246
7	122	Iteration limits exceeded	-3552.46	7348.93	7470.93	3775.99	8134.30	0.254
8	141	Successful convergence	-3518.34	7318.67	7459.67	3776.67	8226.36	0.258
9	160	Nas produced						
10	179	System is computationally singular						
N	4617							

Appendix D

After the DCE, all respondents were asked to rate the importance of the presented attributes, from very unimportant to very important. The overall frequency distribution for the sample is shown in Figure D1 (one-time rentals) and D2 (subscription). The latter includes an alternative for "don't know".

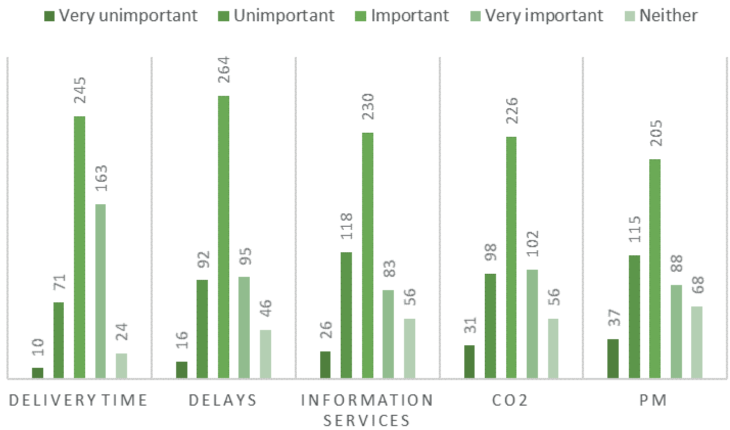

Figure D1: Importance of attributes in debriefing question

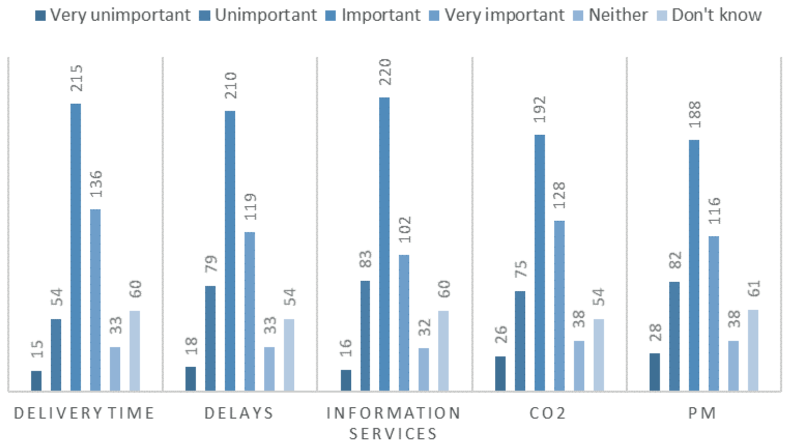

Figure D2: Importance of attributes in debriefing question for a clothing rental subscription

Chapter 5:

Act locally? Are female online shoppers willing to pay to reduce the carbon footprint of last mile deliveries?

Elise Caspersen

Ståle Navrud

Jens Bengtsson

Act locally? Are female online shoppers willing to pay to reduce the carbon footprint of last mile deliveries?

Elise Caspersen[1,2,*], Ståle Navrud[2], Jens Bengtsson[2]

[1]Institute of Transport Economics – Norwegian Centre for Transport Research, Gaustadalléen 21, 0349 Oslo, Norway

[2] School of Economics and Business; Norwegian University of Life Sciences, Ås, Norway.

[*]Corresponding author at: Institute of Transport Economics – Norwegian Centre for Transport Research, Gaustadalléen 21, 0349 Oslo, Norway. Tel.: + 922 444 52, E-mail address: elc@toi.no

Abstract: *E-commerce results in more last mile deliveries, increased freight traffic and potentially also higher CO_2- emissions. This paper is a novel contribution to the literature in terms of investigating consumers' willingness to pay (WTP) for climate-friendly last mile deliveries through reduced or no CO_2-emissions from the delivery. 460 females between 18 and 70 years of age responded to an internet panel survey about their stated preferences for last mile delivery options for online clothing rentals.*

A discrete choice experiment (DCE) was performed and the data analyzed using both multinomial logit (MNL) and mixed multinomial logit (MMNL) models. The results show that females are willing to pay for CO_2-mitigation, and that their WTP increases with consumer income, employment, willingness to change habits to solve the environmental challenges of today, and preferences for sustainable online shopping and delivery alternatives, but falls with the frequency of online shopping. The WTP for 1kg CO_2 exceeds the WTP for any other aspects of the last mile delivery; i.e. delivery time, delays and information services (notification of departure and arrival). The results indicate that freight operators (carriers) and online retailers can transfer (some of) the costs of climate-friendly last mile delivery to their customers. This is important knowledge for urban planners as it provides support for CO_2-mitigating measures aimed at last mile delivery services in order to achieve more environmentally sustainable urban freight transport.

Keywords: climate-friendly, environmentally sustainable, last mile delivery, consumer preferences, discrete choice experiment, females, online shopping.

1 Introduction

1.1 Background and aim

The European Green Deal aims for Europe to become the world's first climate-neutral continent (European Commission). One of its brave goals is zero net emissions of greenhouse gases by 2050 (European Commission, 2019a). In order to achieve this, transport as a large contributor to overall emissions must reduce its emissions by 90 % (European Commission, 2019b). Internalizing climate change externalities to reduce emissions is not enough to reach the goal (as carbon pricing is too weak an instrument and the social cost of carbon (SCC) is too low) and should be combined with other measures like subsidies and regulations (Santos, 2017). Some relevant measures for freight transport in cities are the use of electricity as an alternative power source to diesel (Schulte and Ny, 2018, Teoh et al., 2018), alternative delivery locations like automated delivery stations (de Oliveira et al., 2017), crowdshipping services (Gatta et al., 2018, Gatta et al., 2019), and collaboration and cooperation activities for a better use of freight delivery resources (Ranieri et al., 2018).

The targeted emission reduction coincides with the growth in online shopping and shipments from businesses to consumers (B2C) observed the last decade, and thus several of the above-mentioned measures imply involving end consumers in the last mile delivery solution. However, little is known about consumers' preferences for environmentally friendly last mile deliveries, although freight operators are starting to notice an increased consumer interest in sustainable deliveries. For instance, Postnord (2020) found that sustainable deliveries are preferred by 25% of online shoppers, while 35% are willing to pay for climate-compensated deliveries. Further, the environmental aspect of home deliveries is becoming increasingly important, as online shopping grows at an increased pace during the ongoing Covid-19 pandemic (Postnord, 2020). This paper addresses consumers' willingness-to-pay (WTP) for low or no CO_2-emissions from last mile deliveries. More specifically, it aims to answer the following three questions: *i) How much are female consumers willing to pay for a reduction in CO_2 -emissions from last mile deliveries versus other attributes of the transport, ii) What consumer characteristics determine their WTP to reduce CO_2-emissions from transport, and iii) How can this knowledge support the environmental transformation of last mile transport?*

To address these questions, a sample of Norwegian female internet panelists (18-70 years old) received a discrete choice experiment (DCE) where they were asked to choose between different last mile delivery alternatives for deliveries of clothes rented online. Although inquiring specifically about clothing rentals, the data from the survey is also perceived suitable for analyzing last mile delivery from e-commerce in general for several reasons. First, while the original plan was to sample from the customer base of a Norwegian online platform for clothing rentals (www.fjong.com), a pretest showed very low response rate from this sample. Thus, instead a random sample of female respondents was drawn from an internet panel representing the general Norwegian population. Second, as the sharing economy relies on shared consumption of goods and services using online platforms (Hamari et al., 2016), clothing rentals can be seen as a branch of e-commerce. Third, the survey inquired if the respondents agreed that their answers to the DCE could also work for other purchases than clothing rentals. Only respondents who agreed are included in this analysis.

The contributions made by this paper includes i) New estimates of consumers' WTP for CO_2-mitigation in last mile delivery from an activity resembling online shopping, ii) A topic that is still little researched (cf. the literature review in section 1.2), iii) An emphasis on consumer preference heterogeneity, providing broader knowledge of consumers' CO_2-mitigation preferences for use by urban planners, online retailers and freight operators to pinpoint CO_2-mitigation measures targeting consumers, iv) Knowledge about if and how much consumers are willing to pay for CO_2-mitigation, which may improve public policy makers' confidence in pushing for more sustainable last mile delivery services, and v) An all-female sample to counteract use of male dominant sample in consumer preference for emission reduction (Costa et al., 2019, Achtnicht, 2012).

The paper is organized as follows: the introduction (Section 1) ends with a review of stated preference literature on people's WTP to reduce CO_2-emissions from transport. The objective is to identify consumer characteristics expected to influence the WTP for reduced CO_2-emissions, as stated in the research question. Section 2 presents the methodological framework for DCE, while section 3 describes the survey and data collection. Results are presented in Section 4, including descriptions of the collected data, estimates, discussions and potential implications. Section 5 concludes.

1.2 Consumer preferences for climate-friendly transport

While there is a vast amount of studies on consumer environmental preferences and passenger transport, research on consumer environmental preferences and last mile delivery from internet shopping is scarce. Collins (2015) used a DCE to map customers' preferences for last mile attributes when choosing between home delivery or pick-up point and transport mode, with environmental benefits resulting from mode choice. Polinori et al. (2018) analyzed students' stated WTP for environmentally labelled last mile delivery when purchasing a green t-shirt in a contingent valuation (CV) survey. Schniederjans and Starkey (2014) used the theory of planned behavior to assess the impact of attitude, perceived behavioral control and peer pressure on consumers' intention to buy a green transportation t-shirt, and a method resembling CV to capture their WTP for the green transportation. Punel and Stathopoulos (2017) analyzed WTP for delivery times savings when using crowdshipping services. Vakulenko et al. (2019) did not analyze WTP per se but found that customers seek the same benefits from the shopping and the delivery service, implying that consumers seeking low-price products will not be interested in costly transportation, while time sensitive consumers might accept paying for a quick delivery. No measure of consumer WTP for CO_2-reduction was found in the papers referenced above.

Identified studies that report consumers' WTP for CO_2-reductions from transport are not focused on the last mile delivery. Achtnicht (2012) investigated if the stated level of CO_2-emissions influences car purchase decisions, and if it does, how much different consumer types are willing to pay to reduce it. He found that median WTP for the reference group ranges from €90 to €257 per tCO_2., although differing with travel distance and consumer type (Achtnicht, 2012). Achtnicht employs a mixed-logit model based on discrete choice experiment data where CO_2-emissions is one of several attributes, and has greatly inspired this paper. Costa et al. (2019) estimated a (conditional) multinomial logit model using data from a DCE including levels of CO_2-emission and found a WTP of €88 for a CO_2-reduction of 1g/km for Italian consumers. Mabit and Fosgerau (2011), Tanaka et al. (2014), Hidrue et al. (2011), Hackbarth and Madlener (2013), and Hackbarth and Madlener (2016) revealed that the WTP depends on the emission reduction in question, and on consumer segments (Hackbarth and Madlener, 2016). For the abatement of 1% of vehicle CO_2-emissions, the WTP ranges from €5-65 (Tanaka et al., 2014), €20-90 (Hackbarth and Madlener, 2013) and €2-52 (Hackbarth and Madlener, 2016). Alberini et al. (2018) utilized DCE survey data in both multinomial logit and mixed logit models to estimate consumer preferences for different policies to reduce CO_2-emissions in Italy and the Czech Republic. They found that the WTP to reduce CO_2-emissions by one ton was €133 in Italy and €94 in the Czech Republic.

Schniederjans and Starkey (2014) and Achtnicht (2012) both found that females and young adults (up to 44 years) had a higher WTP for green transportation than their counterparts. While Schniederjans and Starkey (2014) did not find any significant impact of education, location and income on WTP, Achtnicht (2012) found a significant, but marginally higher WTP for highly educated individuals. Alberini et al. (2018) showed that people with higher household income are more likely to have environmental knowledge, awareness, and higher marginal utility of emission reductions: the larger the CO_2-emission reduction, the more people are willing to pay. The rural population has a lower willingness to pay for the environment than urban residents (Lera-López et al., 2014, Tianyu and Meng, 2020). WTP for eco-labeled urban freight transport (among students) increases with income, pro-environmental behavior and attitude, knowledge about sustainability issues, interest in and attention to labels (Polinori et al., 2018). Schniederjans and Starkey (2014) also found that a positive attitude towards environmentally friendly consumption increased intentions to choose green transportation. This is consistent with several studies of consumer attitude and behavior, showing that people who are concerned about the environment and the potential damage humans are causing are more likely to be positive towards environmental behavior (Gadenne et al., 2011).

The literature review indicates that characteristics like income, education, age, urban or rural residency and environmental attitude could help explain the variation in people's WTP for CO_2-reductions. Thus, we expect women's WTP for climate-friendly last mile delivery to be higher for i) women with high personal income, ii) highly educated women, iii) younger women (Generation Z or Millennials born 1981 or later), iv) women living in urban areas, and v) women with a pro-environmental attitude.

2 Discrete choice modeling

To identify consumer preferences and WTP for climate-friendly last mile deliveries, a good that is not yet in the market, stated preference methods, more specifically discrete choice experiments (DCE), are used to collect data. As several choice tasks are collected from the same individual (panel data) and consumer attitudes and intentions are important but difficult to capture the random parameter logit/mixed multinomial logit model (MMNL), as suggested by Revelt and Train (1998), McFadden and Train (2000) and others, is chosen for estimation. This is also consistent with other studies on WTP for CO_2-emissions (Achtnicht, 2012, Alberini et al., 2018). The multinomial logit model (MNL) is estimated as a benchmark model. Following the methodology by Hess and Rose (2009), which builds on the framework by Revelt and Train (1998), as well as Sarrias and Daziano (2017) (for consistency with the "gmnl" R software (R Core Team) package used for estimation), the framework is described below.

2.1 The mixed multinomial logit model (MMNL)

An individual n answers t choice tasks, and is assumed to choose the alternative i, out of J options, that provides the highest utility, U. The utility is explained by observable attributes and sociodemographic variables, x, but also by other factors, ε, that are unobservable (random) to the researcher. When assuming a linear relationship between attributes and taste, the utility can be written as follows, denoted the random utility model (RUM): $U_{i,n,t} = V_{i,n,t} + \varepsilon_{i,n,t} = \beta_{n,t} x_{i,n,t} + \varepsilon_{i,n,t}$, where $\beta_{n,t}$ is a vector of taste coefficients that can differ between individuals and choice sets. It is assumed that ε follows an independent and identically type I extreme value (Gumbel) distribution. The variable specification depends on the assumptions made about $\beta_{n,t}$. For the standard MNL, it is assumed that the taste variation is fixed between individuals and choice situations ($\beta_{n,t} = \beta$), and the probability that individual n chooses alternative i becomes:

$$P_n(i|\beta) = \frac{e^{\beta x_{i,n,t}}}{\sum_{j=1}^{J} e^{\beta x_{j,n,t}}}$$

This model relies on the assumption of independence of irrelevant alternatives (IIA), does not allow for unobserved taste heterogeneity, nor correlation between repeated choice tasks by the same respondent (Achtnicht, 2012). The mixed logit model (MMNL) however allows taste to vary between individuals ($\beta_{n,t} = \beta_n$) by following a random (unknown) distribution $\beta_n \sim f(\beta|\Omega)$. Here Ω is parameters explaining the distribution of β (mean and standard deviation, but also individual-specific covariates, for instance $\Omega = \Pi z_n + L\eta$, where z is a set of characteristics that influences the mean of the preference parameter, π is its corresponding vector of parameters, η denotes random parameter distribution, and L is the lower-triangular Cholesky factor of $\sum$ such that $LL^T = VAR(\beta_i) = \sum$ (Sarrias and Daziano, 2017)). In this case the probability function becomes:

$$P_n = \int_{\beta_n} \left(\prod_{t=1}^{T_n} \frac{e^{\beta_n x_{i,n,t}}}{\sum_{j=1}^{J} e^{\beta_n x_{i,n,t}}} \right) f(\beta_n|\Omega) d\beta_n$$

Taking the log-likelihood and summarizing over all individuals provides the log-likelihood function $LL(\Omega) = \sum_{n=1}^{N} \ln P_{n,t}$ that for the MMNL must be simulated. Following the specifications in the gmnl package (Sarrias and Daziano (2017), the Maximum Simulated Likelihood (MSLE) is used, although the Method of Simulated Moments (MSM) is an alternative (McFadden and Train, 2000). Hess and Rose (2009) suggest other model specifications where the taste coefficients vary between choice sets (β_t) or between both choice sets and individuals ($\beta_{n,t}$). The latter is an option also in this model but left for further analysis. Nested models are compared using the likelihood ratio (LR) test.

A drawback with the MMNL is that the random distribution of the taste variation (β) is unknown and must be specified (and thus restricted) by the researcher (Daziano and Achtnicht, 2013). Commonly used distribution in the literature are the normal, lognormal, triangular, uniform and Johnson SD distribution. The normal distribution allows the parameter values to shift sign and take on both negative and positive values, which might induce a problem (Daziano and Achtnicht, 2013), and result in the use of other distributions, like the log-normal, inducing the same coefficient sign for the whole population (Achtnicht, 2012).

2.2 Willingness to pay

Willingness to pay (WTP) is the ratio of the marginal (dis)utility of a quality attribute to the marginal (dis)utility of the cost attribute and measures the amount of money that a consumer is willing to pay for an improvement of a good or service (Masiero and Hensher, 2010). It is commonly estimated as a point estimator, which for a linear specification of parameters is the parameter ratio: $WTP = \frac{\beta_{quality}}{\beta_{cost}}$ (Masiero and Hensher, 2010, Daly et al., 2012). For interpretation of WTP to be meaningful, its distribution must have finite moments (probability, mean, variance), which implies that price (denominator) cannot be zero. When price is randomly distributed across individuals, several distributions provide infinite moments of WTP (Daly et al., 2012), resulting in different correction approaches, each with their drawbacks (see for instance Sillano and Ortúzar (2005) and Carson and Czajkowski (2019)).

One of the suggested correction methods is estimating log-normal distributions keeping the price parameter strictly positive (Carson and Czajkowski, 2019, Daly et al., 2012), restricting variables to not cover zero. However, the wide tail of log-normal distributions tends to give extremely large WTP-values, and is thus not recommended for valuation purposes, including cases with variable restrictions to avoid zero (Sillano and Ortúzar, 2005). A normal distributed cost-parameter implies that the utility of price can be both positive and negative, which is often counterintuitive, and the ratio of two normal distributed parameters is not solvable analytically (Sillano and Ortúzar, 2005). Thus, to get WTP-estimates that are reliable within the scope of this research, the mixed logit model is estimated using fixed price. Although fixed price tends to overestimate the WTP-estimates as random distributed variables tend to have higher mean than fixed variables (see for instance Sillano and Ortúzar (2005)), it solves the issue of identification as pointed out above and is found acceptable in this research.

3 Data collection

Data was collected through a consumer survey composed of four parts: i) questions about habits and preferences for online shopping, ii) statements related to environmental attitudes, including some repeated questions to test if the response changes as respondents move through the questionnaire (as suggested by Mathews et al. (2007)), iii) stated choice scenarios including debriefing questions and rating of attributes to reveal whether the respondents consider all attributes (inspired by Hensher (2007)), and iv) socio-economic characteristics of the respondents including age, educational level, occupation, personal and household gross income.

3.1 Discrete choice experiment design

As the aim was to capture consumer preferences for different last mile delivery attributes, including CO_2-emissions, DCE was chosen above contingent valuation (CV), as suggested by Johnston et al. (2017). In order to design realistic experiments, attributes and attribute levels were inspired by consumer surveys conducted by Postnord (2020), Bring Research (2019), and World Economic Forum (2020), as well as existing knowledge by the research team. The resulting experimental design included 5 attributes with 2-4 levels each, as presented in Table 1. The survey was designed using QuenchTec. A choice set example is given in Figure 1.

Table 1: Characteristics of the experimental design: attribute description and levels.

Attribute	Description and levels
6. Delivery time	Number of days the respondent accepts to wait for the parcel: 1-5-10-20 days
7. Delays (dummy)	Uncertainty with respect to delivery time: "No", "Yes, 1-2 days"
8. Information (dummy)	Notifications by SMS or e-mail when 1) the good is approved for shipping and 2) the parcel is shipped to the consumer: "No", "Yes"
9. CO_2-emission	CO_2-emissions resulting from transport of the parcel. The emission levels differ with respect to transport mode, time, degree of consolidation etc.: 0, 0.28kg, 1.40kg
10. Price (for the delivery)	Price: 0 NOK[a)], 49 NOK, 99 NOK

[a)] The average exchange rate between Euro (€) and Norwegian kroner NOK at the time of survey (July 2020) was 1€=10.65Kr./1NOK=0.094€ (Source: Norges Bank - The Central Bank of Norway).

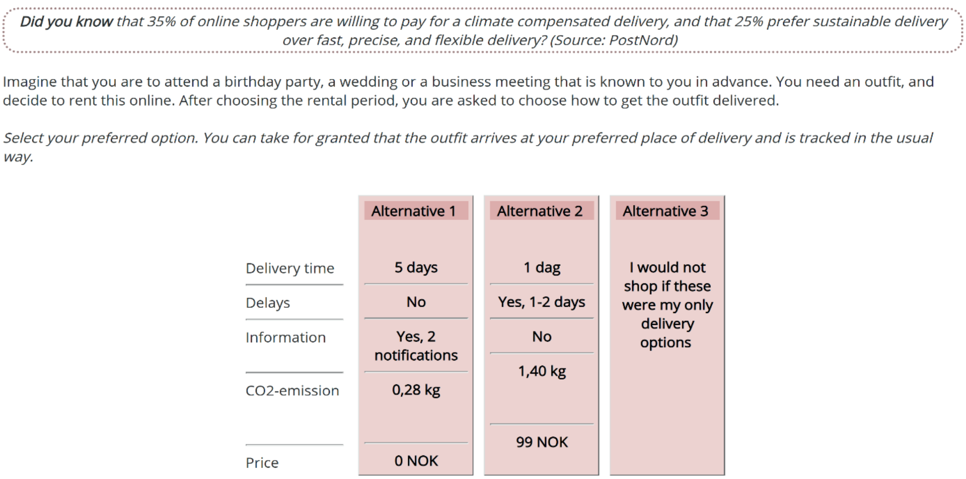

Figure 1: Example of choice question and choice card shown in the survey.

Delivery time of 20 days was included to investigate maximum delivery time for consumers. The levels of CO_2-emission are calculated based on last mile delivery distance (from Statistics Norway) and emission levels (using the Handbook Emission Factors for Road Transport (HBEFA)) for light duty vehicles. The average distance per delivery (1.55 km) and average emission (0.18kg/km) gave the middle value of 0.28kg CO_2-emission per delivery. This is comparable to the average emission of 0.181 CO_2 per delivery found by Edwards et al. (2010)[11]. Delivery price was designed to take on values in the range of those found for online clothing retailers like Zalando, Boozt and H&M. Other commodity types and service levels might have different price levels.

3.2 Data collection

To reduce the complexity of the survey, each choice set consisted of two unlabeled alternatives and an "opt-out" option; see Figure 1. The DCE consisted of 9 choice sets, drawn randomly from 16 blocks from a full factorial design being stripped for dominant alternatives and grouped according to environment and service criteria. Approximately half of the respondents were randomly assigned to get supplementary information about environmental aspects of last mile deliveries. In the DCE this included a sentence about online shopper preferences for climate compensated deliveries (see dotted box in Figure 1). The aim of this split-sample was to test if preferences could be altered, or nudged, by supplementary information; and was inspired by the work of Thaler and Sunstein (2009).

The survey was administered as a web-based survey in Norway using the NORSTAT[12] internet panel. The panel consists of 81 000 active panelists with 52% female and evenly distributed on age groups starting from 15 years. Respondents are rewarded a small incentive for their participation; they receive bonus points which can be exchanged for a gift card (1 minute of response time= 1 point = 1 NOK). When distributing the survey in question, only females between 18 and 70 years of age were targeted. The survey was conducted from June 29th until August 3rd 2020; resulting in a sample of 605 respondents[13]. The frequency of the chosen DCE alternatives is presented in Appendix A.

Both qualitative pretesting (general feedback from testing the survey and one-on-one interviews with representatives from both experts and user group) and quantitative pretesting (using data from a pilot survey of the internet panel) were conducted as recommended by Mansfield and Pattanayak (2007), Champ and Welsh (2007), Krupnick and Adamowicz (2007), Harrison (2007), Mathews et al. (2007), and Johnston et al. (2017). Focus groups were not an option due to Covid-19 socializing restrictions; but the pretesting helped design attribute levels in the DCE as well as to fine tune questions and information text.

4 Results

In the following section all results, from data collection effort, estimation results, WTP calculations to a discussion of potential implications, are presented.

---

11 From Edwards et al. (2010): "A typical 50-mi delivery round by diesel van produces 21,665 g CO_2 in total, and with an average delivery rate of 120 drops per trip, each successful first-time drop would be allocated 181 g CO_2 or its share of the 21,665-g total (this calculation assumes that all drops are delivered successfully; i.e., there are no failed deliveries)."

12 https://norstat.no/

13 The sample was part of a survey collecting data from two DCEs. In total, 4602 persons from the internet panel were invited to take the survey. 1200 responded, yielding an overall response rate of 26%. Respondents were randomly allocated to one of the two DCEs, and of the 1200 respondents, 595 answered the other DCE, and is not included in this paper. The average time used to complete the survey for all 1200 respondents was 11,5 minutes.

4.1 Descriptive statistics

Most respondents carefully consider their survey answers (Mansfield and Pattanayak, 2007). However, some of the responses in the collected sample were flawed. After screening for respondents focusing on only one attribute in the DCE, inconsistencies in debriefing questions (i.e. answer honest and random at the same time), very quick or slow response,[14] 74 respondents were deleted. Additionally, as the interest in this paper is to extend consumer preferences to general online shopping, 52 respondents where the DCE could not extend to purchase of other commodities than clothes and 19 respondents who never shopped online were excluded, leaving 460 observations for analysis. Descriptive statistics for the remaining 460 observations are presented in Table 2 along with statistics for the Norwegian female population. The population is restricted to females aged 18 to 70 years where possible.

[14] Inspired by Hensher (2007), Mathews et al. (2007), and an example by Alberini et al (2007).

Table 2: Descriptive statistics of respondents in the sample, and the Norwegian female population (from Statistics Norway). N=number of observations. All variables are binary, taking on the values 1 = "Yes" or 0 = "No".

	CE1 *(N=460)*	***Female population***
Average age and generation [a]		
Average age (18-70 years)	41.4 years	43.7 years
1997-2001 (Generation Z)	10%	9% [b]
1981-1996 (Millennials)	38%	33%
1965-1980 (Generation X)	33%	32%
1949-1964 (Boomers)	19%	25%
Top25pop		
Lived in one of the 25 most populated Norwegian municipalities.	59%	53% [b]
Education		
Primary school	3%	25% [c]
High school	38%	36%
College or university	59%	39%
Employment status		
Employed	65%	65% [d]
Unemployed	3%	3%
Not in work force (incl. students)	28%	32%
Other	3%	
Annual gross personal income in NOK (2019)		
Average income (based on middle value of intervals)	479,000	382,000 [e]
Less than or equal to 600,000 NOK	61%	
More than 600,000 NOK	18%	
NA	20%	
Frequent online shopper		
Shopped online at least once a month	47%	
Reduced consumption		
Agree that reduced consumption is our most important contribution to solving the environmental challenges	83%	
Change habits		
Agree on being willing to change habits to solve the environmental challenges	79%	
Sustainable shopping		
One of the three most important attributes of online shopping are environmentally friendly shopping and delivery alternatives	7%	
Time savings		
One of the three most important attributes of online shopping is time savings	12%	
Lower price		
One of the three most important attributes of online shopping is that the price is lower than in stores	47%	
Free delivery		
One of the three most important attributes of online shopping is free delivery	37%	
Supplementary information		
Received supplementary information of environmental aspects of home deliveries	43%	

a) As defined by Pew Research b) Females 18-70 years c) Females 16 years or older, d) Females 15 years or older, e) Average annual gross income for females 17 years and older in 2018 (2019 numbers are postponed until January 2021).

Table 2 shows that the sample is more educated than the overall Norwegian female population, which reflects the composition of the internet panel they are drawn from. High education is also reflected in the average annual wage, which is almost 100,000 NOK higher in the sample than in the population. The sample is also somewhat more urban than the general population, which might explain the high average education and income levels. Personal gross income is coded as a binary variable (less than or equal to 600,000 NOK and more than 600,000 NOK) to capture potential income effects on the WTP[15]. Approximately half of the respondents (47 %) are frequent online shoppers and buy online at least once a month, while the rest buys online less than once a month. This number seems to differ between studies of Norwegian online shopping behavior and framing of the question. Bjerkan et al. (2020) found that almost 12 % (56 of 484 respondents) of male and female respondents aged 18 to 87 years living in the Oslo (capital) region are frequent online shoppers (shop at least once a month) based on data collected in the period November 2018 to January 2019 using a survey agency. Postnord (2020) found that approximately 67 % of the male and female Norwegian population aged 18 to 79 years had shopped online last month (at the time of survey), based on data collected in the period January 2019 to December 2019 using a survey agency.

When asked about their preferred delivery location, a higher share of frequent online shoppers preferred their package to be delivered at home rather than at a pick-up point compared to infrequent online shoppers. This is presented in Figure 2 and represents general online purchases (not clothing rentals in particular).

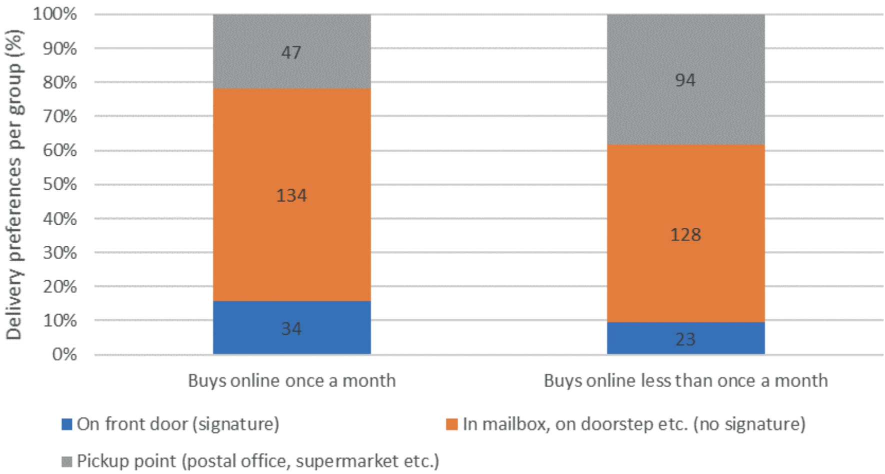

Figure 2: Preferred delivery location by frequency of online shopping. N=460

[15] In 2018, only 15% of Norwegian females had an average annual gross personal income of 600,000 NOK or more (Statistics Norway, Table 08411), indicating that 600,000 NOK is a relatively high wage. Lower numbers could have been used, but as the average annual personal gross income in the sample is high (600,000 NOK) and almost 1/5 of the sample had income above 600,000 NOK, this was found an appropriate level for defining the "high income" group

Figure 3 presents respondents' ranking of the top three most important aspects of online shopping among a list of 14 alternatives relating to delivery, return and shopping experience. It is interesting to observe that infrequent online shoppers value delivery and return service, while frequent online shoppers value a convenient shopping experience. Cost seem to be an important aspect for both groups, with lower product price online than in the store and free delivery being the overall top two aspects. This result coincides with other consumer studies. Postnord (2020) found that free delivery and return are important for the majority (three out of four) of consumers who shopped online the previous month, while Bring Research (2019) found that the price of delivery and return is the main reason for disrupted online purchases by both males and females. It is however interesting to notice that few respondents ranked being able to select the location of the delivery as one of their top three aspects, indicating that there might be room for adjusting the last mile delivery to reduce greenhouse gas emissions.

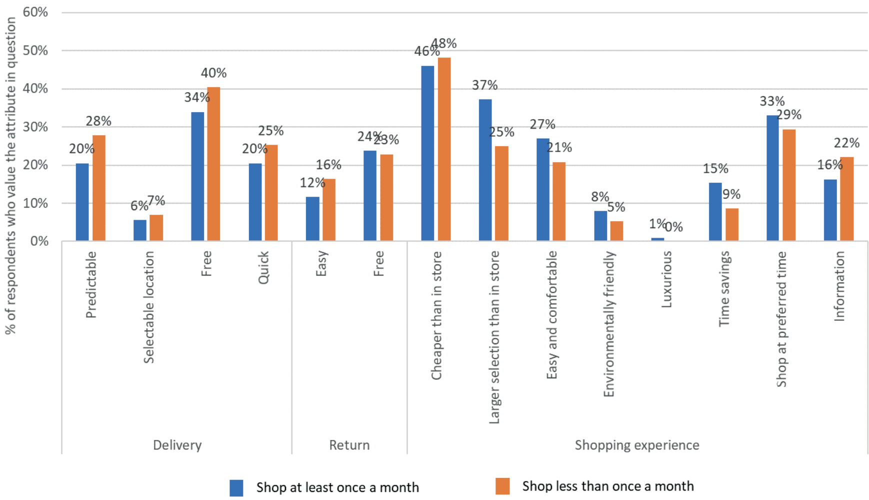

Figure 3: Respondents' votes on the most important aspects when shopping online. Each respondent was asked to vote for their three most important aspects. N=460

4.1 Parameter estimation and discussion

To identify consumer WTP for CO_2-mitigation in last mile deliveries from online shopping, and how this differs between consumers, the following model, inspired by the data and the literature review in section 1, is suggested:

$$U_{n,t,j} = \alpha_j + \beta_{1,n} \cdot Delivery\ Time_{n,t,j} + \beta_{2,n} \cdot Delays_{n,t,j} + \beta_{3,n} \cdot Information\ services_{n,t,j} + (\beta_{4,n} + \delta_{1,n} Sustainable\ shopping + \delta_{2,n} Frequent\ online\ shopper + \delta_{3,n} Change\ habits + \delta_{4,n} Employed + \delta_{5,n} Supplementary\ info) \cdot CO_{2_{n,t,j}} + (\beta_{5,n} + \delta_{6,n} High\ Income + \delta_{7,n} Free\ Delivery + \delta_{8,n} Lower\ Price + \delta_{9,n} Time\ Savings + \delta_{10,n} Reduced\ Consumption) \cdot Price_{n,t,j}$$

where β denotes parameter estimates for attributes and δ parameter estimates for interaction terms. The inclusion of alternative specific constants, α, in the context of an unlabeled choice experiment is motivated to capture inertia (i.e. sticking to one alternative), as well as reading from left to right effects (Hess and Rose, 2009).

Variables for urban residency, education and age were also tested but found insignificant in the introductory models and left out of further analysis. Correlations between variables are found acceptable (see Appendix B).

The model is estimated using both multinomial logit model (MNL) and mixed logit model (MMNL). Although not suitable for discovering unobserved heterogeneity, the MNL provides a useful benchmark for comparison. The MMNL has been estimated using different distributions for price. Although the models with randomly distributed price parameter show a better fit to the data, the chosen model treats price as fixed to secure defined WTP moments, as done by Achtnicht (2012) and Alberini et al. (2018) and explained in subsection 2.2. Models with price following a normal and lognormal distribution are presented in Appendix C for comparison. The remaining four attributes (delivery time, delays, information service and CO_2) follow either a log-normal or normal distribution. As in Achtnicht (2012), individuals are expected to be negative or indifferent towards CO_2, not positive, which is estimated using a log-normal distribution. The same goes for the attributes of delivery time and delays. Although information is a service, some may think notifications from the retailer unnecessary or exhaustive, and the variable is kept normal. The MMNL model is simulated with Halton draws with 1000 replications for the maximum simulated likelihood estimation using the "gmnl-package" (Sarrias and Daziano, 2017) in R software (R Core Team).

The estimation results are presented in Table 3. The first column presents the variables to be estimated, the second presents parameter estimates for MNL, while the third to fifth columns present the results from the MMNL model with price being fixed. For the attribute variables following a log-normal distribution (i.e. delivery time, delays and CO_2,) mean, median and standard deviation are presented. Information service follows a normal distribution and its mean equals its median. The other parameters in the model are fixed. For the interaction terms between CO_2 and the chosen explanatory variables, the shift in CO_2 mean and median for the relevant groups are presented as well as the model estimates (last column).

Table 3: Parameter estimates and standard errors (in parenthesis) from estimation of Model 1 MNL- and Model 2 MMNL models. Model 2 is presented with mean, median and standard deviation for random parameters.

	Model 1: MNL	*Model 2: MMNL - price is fixed* Mean	Median	Standard dev.
α_1	2.883*** (0.091)	4.812*** (0.151)		
α_2	2.794*** (0.088)	4.674*** (0.147)		
Price	-0.020*** (0.002)	-0.028*** (0.002)		
Price x Income>600 000NOK	0.003** (0.001)	0.003* (0.002)		
Price x Free delivery	-0.006*** (0.001)	-0.007*** (0.001)		
Price x Lower Price	-0.003** (0.001)	-0.004** (0.001)		
Price x Time savings	0.003* (0.002)	0.004* (0.002)		
Price x Reduced consumption	0.005*** (0.001)	0.006** (0.002)		
Information services	0.123** (0.042)	0.165** (0.058)		0.542*** (0.096)
Delivery time	-0.095*** (0.004)	-0.251*** (0.026)	-0.133*** (0.01)	0.402*** (0.093)
Delays	-0.140** (0.043)	-0.479*** (0.079)	-0.095* (0.041)	2.369* (1.119)
CO_2	-0.147 (0.099)	-1.733*** (0.291)	-0.542*** (0.083)	5.267** (1.866)
		Shift in CO_2 mean	*Shift in CO_2 median*	*Model estimate*
CO_2 x supplementary information	-0.137 (0.076)	*0.179*	*0.056*	-0.109 (0.328)
CO_2 x Sustainable shopping	-0.734*** (0.172)	*-4.686*	*-1.465*	1.309* (0.510)
CO_2 x Frequent online shopper	0.227** (0.072)	*0.684*	*0.214*	-0.502 (0.346)
CO_2 x Change habits	-0.409*** (0.085)	*-0.879*	*-0.275*	0.410 (0.299)
CO_2 x Employed	-0.301*** (0.074)	*-0.277*	*-0.086*	0.148 (0.320)
AIC	7189.513	6089.101		
BIC	7297.096	6221.999		
Log-likelihood	-3577.756	-3023.551		
N	4140	4140		

Significance: *** = $p < 0.001$; ** = $p < 0.01$; * = $p < 0.05$

A likelihood ratio test rejects the MNL in favor of the MMNL (λ_{LR} =1108.41). Thus, considering the panel structure of the data and unobserved heterogeneity in the attributes significantly improves the model fit[16].

As expected, increments in delivery time, delays, CO_2-emissions and price all have a negative effect on consumer utility, while information service has a positive effect. The attributes are significant in both models, indicating that they influence the utility of last mile delivery alternatives, except for CO_2 in the MNL model, which is insignificant. The significant standard deviations of the random parameters indicate unobserved variation in parameter distribution and heterogenous consumer preference. Price and CO_2 are estimated with interaction terms. The base group for price is shown to the left in Figure 4. As price is fixed in both models, the interaction terms for price have the same size and magnitude: price sensitivity decreases with income, impatience and environmental consciousness, and increases with the importance of getting a lower price online than in the store and with free delivery.

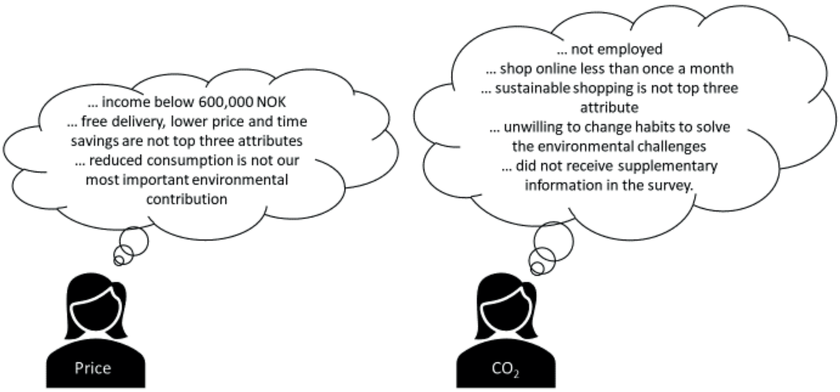

Figure 4: Base groups for Price and CO_2.

With regards to CO_2, the models differ as heterogeneity is introduced in the MMNL. The base group for CO_2 is shown to the right in Figure 4. In the MNL model, the interaction terms explain most of the consumer preferences for CO_2-emissions, as the effect on the base group is non-significant. For the MMNL, the results are opposite: the CO_2-emission variable is negative and highly significant, while only the variable indicating sustainable shopping and delivery options is significant at α=0.05. A comparison with the models allowing price to follow a normal and a log-normal distribution (Appendix C) shows that the price, CO_2 and delivery time variables are in the same range in the three MMNL-models, while the utility of information services is a bit lower in the model with fixed price. The variable for delay is similar in the MMNL-models with fixed price and price following a normal distribution but becomes insignificant when price follows a log-normal distribution. The explanatory power improves when price is allowed to vary, indicating that keeping price fixed is a strict assumption.

[16] A model with uncorrelated random parameters was also tested but could not produce covariance elements for some of the attributes and was rejected.

4.3 Willingness to pay estimates

The MMNL mean of the fixed price (Table 3) is significant at a z-value of -13.19. This is sufficiently large to calculate willingness to pay (WTP) as suggested in subsection 2.2, as well as using the delta method to estimate standard errors (Carson & Czajkowski, 2019). The delta method is found to be sensitive to any departures from normality in the data and small sample size, but appropriate for large datasets with low variation of the cost-parameter where it provides smaller standard errors than several of its counterparts (Gatta et al., 2015). Thus, the delta method is chosen to estimate standard errors for the WTP presented below.

The WTP-calculations distinguish between different consumer groups and price sensitivity using parameter estimates in Table 3. The WTP-calculations from the MNL and MMNL models with price being fixed are presented in Table 4. The top part presents calculations for respondents with income at or below 600,000 NOK, while the bottom part presents calculations for respondents with income above 600,000 NOK. The base group for CO_2 is as shown in Figure 4. As the mean of randomly distributed variables tends to be higher than for fixed variables (see for instance Sillano and Ortúzar (2005) and Achtnicht (2012)), WTP-calculations based on both mean and median values from the MMNL model are presented. The median WTP is the center of the cumulative distribution function and thus a better proxy for the average person's WTP than the median, which is influenced by outliers through the standard errors (Achtnicht, 2012). Hence, the median MMNL WTP is compared with the MNL WTP[17].

[17] WTP for the MMNL models with randomly distributed price (Appendix C) is not calculated as the distribution of the price variables covers zero.

Table 4: Willingness to pay (WTP) in NOK for a change in the attribute level. For CO_2-emission attributes the WTP is presented in NOK/kg CO_2.

Income at or below 600 000 NOK/not reported	***MNL***	***MMNL - price is fixed***	
		Median	*Mean*
Delivery time	-4.9*** (0.4)	-4.8*** (0.5)	-9.1*** (1.1)
Delays	-7.1** (2.3)	-3.4* (1.5)	-17.4*** (3.1)
Information service	6.3** (2.2)	6.0** (2.2)	6.0** (2.2)
CO_2 (base group)	-7.5 (5.1)	-19.6*** (3.3)	-62.8*** (11.3)
CO_2 (received supplementary information)	-14.5* (5.7)	-17.6** (6.2)	-56.3** (19.8)
CO_2 (favor sustainable shopping)	-45*** (10.7)	-72.7 (38.2)	-232.6 (130)
CO_2 (frequent online shopper)	4.1 (5.5)	-11.9** (4.6)	-38** (14.6)
CO_2 (willing to change habits)	-28.4*** (4.3)	-29.6** (9.4)	-94.7** (31)
CO_2 (employed)	-22.9*** (4.9)	-22.8** (8.0)	-72.8** (26.3)
Income above 600 000 NOK	***MNL***	***MMNL - price is fixed***	
		Median	*Mean*
Delivery time	-5.9*** (0.7)	-5.5*** (0.5)	-10.4*** (1.1)
Delays	-8.7** (2.9)	-3.9** (1.5)	-19.9*** (3.1)
Information service	7.7** (2.8)	6.8** (2.2)	6.8** (2.2)
CO_2 (base group)	-9.1 (6.2)	-22.5*** (3.3)	-71.9*** (11.3)
CO_2 (received supplementary information)	-17.6* (7)	-20.1** (6.2)	-64.4** (19.8)
CO_2 (favor sustainable shopping)	-54.7*** (13.8)	-83.2* (38.2)	-266.2* (130)
CO_2 (frequent online shopper)	5 (6.7)	-13.6** (4.6)	-43.5** (14.6)
CO_2 (willing to change habits)	-34.5*** (5.9)	-33.9*** (9.4)	-108.3*** (31)
CO_2 (employed)	-27.8*** (6.4)	-26.1** (8.0)	-83.4** (26.3)

Significance: *** = $p < 0.001$; ** = $p < 0.01$; * = $p < 0.05$

Table 4 shows that the WTP from the MNL-model and the MMNL-model are quite similar for delivery time and information services, suggesting that people are willing to pay around 5 NOK/day in increased transportation cost for decreased delivery time, and around 6 NOK for information services. The WTP to avoid delay is more than twice as large in the MNL-model as (the median) in the MMNL-model. The WTP for reducing CO_2-emissions for the base group differs between the models, being insignificant in the MNL-model. In the MMNL-model, the average respondent in the base group is willing to pay around 20 NOK in increased transportation costs for reducing transport emissions of CO_2 by 1 kg. The WTP is higher for consumers who are employed or willing to change habits for the environment, and lower for frequent online shoppers or those who received supplementary information. It does not differ much with income above or below 600,000 NOK. The exception is for consumers favoring sustainable online shopping: when income increases to 600,000 NOK or more, the WTP becomes significant.

4.4 Discussion of the results and implications for last mile transport

The results above reveal a positive WTP for CO_2-mitigation among female consumers; this also being the attribute with the highest WTP (given the levels used in this survey). Preferences and WTP differ with both observed and unobserved heterogeneity. Among the observed heterogeneity increasing the WTP are pro-environmental attitudes (willingness to change habits for the environment and preferences for sustainable online shopping and delivery alternatives), employment and income. This is as expected and reflects findings in the literature presented in subsection 1.2. Urban or rural residency, education and age were insignificant in the model. The WTP for all attributes is within the range of the most common delivery prices for Norwegian last mile delivery service, but large compared to the WTP estimates in the literature. With a WTP of 20 NOK/kg CO_2 (the base group), the WTP per tCO_2 is 20,000 NOK (approximately €2000), which greatly exceeds the estimates by Achtnicht (2012) (ranges from €90 to €257 per tCO_2) and Alberini et al. (2018) (€133 per tCO_2 in Italy and €94 per tCO_2 in the Czech Republic). However, the levels under study are small (ranging from 0 – 1.40kg) and preferences might not be constant for all levels of CO_2, as well as estimated in a different context. Although not included in the WTP-calculations in Table 4, the estimation results from Table 3 reveal that the WTP for last mile delivery is lower for consumers who have free delivery or lower price online than in store as one of their top three most important attributes of online shopping, and higher for consumers who value time savings and agree that reduced consumption is the most important contribution to solving the environmental challenges. These results correspond well with the findings by Vakulenko et al. (2019).

Reduction measures for CO_2-emissions were not specified in the choice experiments in order to avoid influencing respondents' choices and avoid creating potential protest or strategic behavior. Thus, the WTP estimates are generic and allow transport operators and policy makers to design their own solutions for CO_2-mitigation using the findings in this paper. Additionally, policy makers can back up their claims towards freight operators reducing their CO_2-emissions who (along with the online retailers) should be able to transfer (some of) the cost to the consumers. However, when designing measures to support the transformation to environmentally sustainable last mile deliveries, some key points might be considered, as illustrated in Figure 5 and described below.

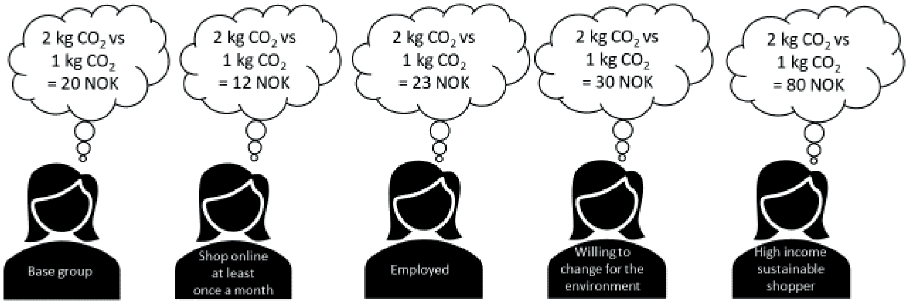

Figure 5: Willingness to pay for CO_2-mitigation for different consumer types.

Firstly, frequent online shoppers have a lower WTP than infrequent shoppers. Frequent shoppers might reflect consumers who are less concerned about their environmental footprint, accustomed to getting free delivery (50% of shoppers were offered free delivery on their last online purchases (Bring Research, 2019)), or fear the potential increase in total expenses of shopping if transportation costs increase. Figures 2 and 3 showed that frequent online shoppers have their homes as the preferred delivery location and value convenient deliveries. Thus, inexpensive measures altering as little as possible, like membership benefits, subscriptions or coupons favoring environmental behavior, might get frequent shoppers choose more climate-friendly transport. Less frequent online shoppers do to a larger extent prefer delivery in a location outside of their homes. Thus crowd-shipping using public transport (or other CO_2-mitigating solutions) might be an alternative for this group, which they might even be willing to pay for. However, few respondents (among both frequent and infrequent online shoppers) ranked delivery location as one of their top three important attributes (Figure 3). This suggests that there might be even more room for adjusting the delivery location than argued above. Females who are employed or have income above 600,000 NOK are willing to pay more than others, but the difference seems to be small. However, data on income and education may be easily collectable from consumers and can be used for price differentiation. A larger difference is found for respondents with environmental concerns, but these aspects are harder to measure and collect from the consumer. A solution is to target the goods: if sustainable products like organic food or clothing rentals are sold, consumers should be offered last mile delivery solutions with low or no CO_2-emissions at a cost.

Independent of consumer type, identifying CO_2-measures adapted to customer attitudes and communicating the effects of the measures is important. Esper et al. (2003) find that transparency and being able to select the carrier increase consumers' willingness to purchase in an e-retail setting, which can be assumed to extend to delivery service preferences (Punel and Stathopoulos, 2017). Further, trust in effective implementation of CO_2-reducing measures makes people more willing to pay and contributes to its success (Yang et al., 2014). Past experience, however, may discourage future participation intention; as found for crowdshipping (Punel and Stathopoulos, 2017), self-collection services (Wang et al., 2019) and parcel lockers (Vakulenko et al., 2019). Hence, when enforcing new solutions for the consumers, professionality is important and defines the future of the solution (Vakulenko et al., 2019). Municipalities and urban planners should facilitate and contribute to reliable and smooth implementations of CO_2-mitigating last mile delivery solutions.

5 Conclusion

This paper documents an analysis of female consumers' willingness to pay (WTP) for climate friendly home deliveries of online shopping. Their WTP for a reduction in CO_2-emissions compared to other last mile delivery attributes is analyzed, along with how consumer characteristics influence the WTP for CO_2. The findings are seen in relation to the environmental transformation of last mile deliveries from online shopping.

The results show that the WTP for last mile delivery attributes differs with observed and unobserved heterogeneity, where the latter is measured using the mixed logit model. The attributes delivery time, delays and information service are found to have an average WTP ranging from 3 to 8 NOK per level of service. CO_2-emission reduction is the most valued attribute (at the given levels), although there are great variations between consumer types. The average WTP ranges from 12–30 NOK per 1kg of CO_2-mitigation and is significantly higher for respondents with an annual income above 600,000 NOK than below, although the overall difference is small. The results suggest that public policy makers can back up their claims for CO_2-mitigation from last mile deliveries as freight operators and online retailers can transfer (some of) the costs to their female customers.

As females have a slightly higher willingness to pay for green transportation than males (Schniederjans and Starkey, 2014, Achtnicht, 2012), and commodity type might be important for the environmental preferences for last mile deliveries (Collins, 2015), more surveys should be performed on samples including males and for other commodity groups. However, the study provides evidence that at least female online shoppers accept paying for transportation costs, and that they are willing to pay extra to reduce greenhouse gas emissions of last mile deliveries. The results are in line with the observed consumer consciousness with respect to environmentally friendly consumption (Fjeld & Krekling, 2020), and could be used by policy makers in supporting new carbon pricing policies for freight transportation as online shopping has surged during the COVID-19 pandemic (NTB, 2020). This would provide incentives for consumers to choose climate-friendly deliveries and curb greenhouse gas emissions in the age of e-commerce.

Acknowledgements: The authors would like to thank FJONG for constructive inputs to the survey, Stefan Flügel for fruitful discussions on methodology, Jardar Andersen for manuscript review, and Anna Herzog for proofreading.

Funding: This work is undertaken as part of the research project 250432 NORSULP (Sustainable Urban Logistics Plans in Norway), financed by the Research Council of Norway and the Norwegian Public Roads Administration. Ståle Navrud's work on this paper is funded by the project "Fjong 2025 – an endless, sustainable wardrobe. Powered by research in behavioral economy, environmental impact and artificial intelligence" (Research Council of Norway project no. 309977).

Conflicts of interest: None

References

ACHTNICHT, M. 2012. German car buyers' willingness to pay to reduce CO_2 emissions. *Climatic Change,* 113, 679-697.

ALBERINI, A., BIGANO, A., ŠČASNÝ, M. & ZVĚŘINOVÁ, I. 2018. Preferences for Energy Efficiency vs. Renewables: What Is the Willingness to Pay to Reduce CO_2 Emissions? Ecological *Economics,* 144, 171-185.

BJERKAN, K. Y., BJØRGEN, A. & HJELKREM, O. A. 2020. E-commerce and prevalence of last mile practices. *Transportation Research Procedia,* 46, 293-300.

BRING RESEARCH 2019. Slik velger kundene deg. *In:* NORGE, P. (ed.).

CARSON, R. T. & CZAJKOWSKI, M. 2019. A new baseline model for estimating willingness to pay from discrete choice models. *Journal of Environmental Economics and Management,* 95, 57-61.

CHAMP, P. & WELSH, M. 2007. Survey Methodologies for Stated-Choice Studies. *Valuing Environmental Amenities Using Stated Choice Studies: A Common Sense Approach to Theory and Practice.* Arlington, Virgina, USA: Springer.

COLLINS, A. 2015. Behavioural Influences on the Environmental Impact of Collection/Delivery Points. *In:* FAHIMNIA B., BELL M., HENSHER D. & J., S. (eds.) *Green Logistics and Transportation.* Springer, Cham.

COSTA, E., MONTEMURRO, D. & GIULIANI, D. 2019. Consumers' willingness to pay for green cars: a discrete choice analysis in Italy. *Environment, Development and Sustainability,* 21, 2425-2442.

DALY, A., HESS, S. & TRAIN, K. 2012. Assuring finite moments for willingness to pay in random coefficient models. *Transportation,* 39, 19-31.

DAZIANO, R. & ACHTNICHT, M. 2013. Accounting for Uncertainty in Willingness to Pay for Environmental Benefits *ZEW - Centre for European Economic Research Discussion,* Paper No. 13-059.

DE OLIVEIRA, L. K., MORGANTI, E., DABLANC, L. & DE OLIVEIRA, R. L. M. 2017. Analysis of the potential demand of automated delivery stations for e-commerce deliveries in Belo Horizonte, Brazil. *Research in Transportation Economics,* 65, 34-43.

EDWARDS, J., MCKINNON, A., CHERRETT, T., MCLEOD, F. & SONG, L. 2010. Carbon Dioxide Benefits of Using Collection–Delivery Points for Failed Home Deliveries in the United Kingdom. *Transportation Research Record,* 2191, 136-143.

ESPER, T., JENSEN, T., L. TURNIPSEED, F. & BURTON, S. 2003. *The Last Mile: An Examination of Effects of Online Retail Delivery Strategies on Consumers*.

EUROPEAN COMMISSION. *A European Green Deal - Striving to be the first climate-neutral continent* [Online]. European Union. Available: https://ec.europa.eu/info/strategy/priorities-2019-2024/european-green-deal_en [Accessed November 2020].

EUROPEAN COMMISSION 2019a. Clean Energy. https://ec.europa.eu/info/strategy/priorities-2019-2024/european-green-deal/clean-energy_en: European Union.

EUROPEAN COMMISSION 2019b. Sustainable mobility. *In:* UNION, E. (ed.). https://ec.europa.eu/commission/presscorner/detail/en/fs_19_6726: European Commission.

FJELD, I. E. & KREKLING, D. V. 2020. NRK-survey: Half of the Norwegian population say they will reduce their consumption after Corona. *NRK.no*.

GADENNE, D., SHARMA, B., KERR, D. & SMITH, T. 2011. The influence of consumers' environmental beliefs and attitudes on energy saving behaviours. *Energy Policy,* 39, 7684-7694.

GATTA, V., MARCUCCI, E., NIGRO, M., PATELLA, S. & SERAFINI, S. 2018. Public Transport-Based Crowdshipping for Sustainable City Logistics: Assessing Economic and Environmental Impacts. *Sustainability,* 11, 145.

GATTA, V., MARCUCCI, E., NIGRO, M. & SERAFINI, S. 2019. Sustainable urban freight transport adopting public transport-based crowdshipping for B2C deliveries. *European Transport Research Review,* 11, 13.

GATTA, V., MARCUCCI, E. & SCACCIA, L. 2015. On finite sample performance of confidence intervals methods for willingness to pay measures. *Transportation Research Part A: Policy and Practice,* 82, 169-192.

HACKBARTH, A. & MADLENER, R. 2013. Consumer preferences for alternative fuel vehicles: A discrete choice analysis. *Transportation Research Part D: Transport and Environment,* 25, 5-17.

HACKBARTH, A. & MADLENER, R. 2016. Willingness-to-pay for alternative fuel vehicle characteristics: A stated choice study for Germany. *Transportation Research Part A: Policy and Practice,* 85, 89-111.

HAMARI, J., SJÖKLINT, M. & UKKONEN, A. 2016. The sharing economy: Why people participate in collaborative consumption. *Journal of the Association for Information Science and Technology,* 67, 2047-2059.

HARRISON, G. 2007. Making Choice Studies Incentive Compatible. *Valuing Environmental Amenities Using Stated Choice Studies: A Common Sense Approach to Theory and Practice.* Arlington, Virginia, USA: Springer.

HENSHER, D. 2007. Attribute Processing in Choice Experiments and Implications on Willingness to Pay. *Valuing Environmental Amenities Using Stated Choice Studies: A Common Sense Approach to Theory and Practice.* Arlington, Virginia, USA: Springer.

HESS, S. & ROSE, J. M. 2009. Allowing for intra-respondent variations in coefficients estimated on repeated choice data. *Transportation Research Part B: Methodological,* 43, 708-719.

HIDRUE, M. K., PARSONS, G. R., KEMPTON, W. & GARDNER, M. P. 2011. Willingness to pay for electric vehicles and their attributes. *Resource and Energy Economics,* 33, 686-705.

JOHNSTON, R. J., BOYLE, K. J., ADAMOWICZ, W. V., BENNETT, J., BROUWER, R., CAMERON, T. A., HANEMANN, W. M., HANLEY, N., RYAN, M., SCARPA, R., TOURANGEAU, R. & VOSSLER, C. A. 2017. Contemporary Guidance for Stated Preference Studies. *Journal of the Association of Environmental and Resource Economists,* 4, 319-405.

KRUPNICK, A. & ADAMOWICZ, W. 2007. Supporting Questions in Stated Choice Studies. *Valuing Environmental Amenities Using Stated Choice Studies: A Common Sense Approach to Theory and Practice.* Arlington, Virginia, USA: Springer.

LERA-LÓPEZ, F., SÁNCHEZ, M., FAULIN, J. & CACCIOLATTI, L. 2014. Rural environment stakeholders and policy making: Willingness to pay to reduce road transportation pollution impact in the Western Pyrenees. *Transportation Research Part D: Transport and Environment,* 32, 129-142.

MABIT, S. L. & FOSGERAU, M. 2011. Demand for alternative-fuel vehicles when registration taxes are high. *Transportation Research Part D: Transport and Environment,* 16, 225-231.

MANSFIELD, C. & PATTANAYAK, S. 2007. Getting Started. *In:* KANNINEN, B. (ed.) *Valuing Environmental Amenities Using Stated Choice Studies: A Common Sense Approach to Theory and Practice.* Arlington, Virginia, USA: Springer.

MASIERO, L. & HENSHER, D. A. 2010. Analyzing loss aversion and diminishing sensitivity in a freight transport stated choice experiment. *Transportation Research Part A: Policy and Practice,* 44**,** 349-358.

MATHEWS, K., FREEMAN, M. & DESVOUSGES, W. 2007. How and How Much? *Valuing Environmental Amenities Using Stated Choice Studies: A Common Sense Approach to Theory and Practice.* Arlington, Virginia, USA: Springer.

MCFADDEN, D. & TRAIN, K. 2000. Mixed MNL models for discrete response. *Journal of Applied Econometrics,* 15**,** 447-470.

NTB. 2020. After Corona: Norwegians will shop more online and reduce their visits to stores. *NTB.no*.

POLINORI, P., MARCUCCI, E., GATTA, V., BIGERNA, S., BOLLINO, C. A. & MICHELI, S. 2018. Eco-labeling and sustainable urban freight transport: How much are people willing to pay for green logistics? *Rivista Internazionale di Economia dei Transporti / International Journal of Transport Economics,* XLV**,** 631-658.

POSTNORD 2020. Netthandel i Norden - Oppsummering 2019. *In:* POSTNORD (ed.) *Netthandel i Norden.* PostNord.

PUNEL, A. & STATHOPOULOS, A. 2017. Modeling the acceptability of crowdsourced goods deliveries: Role of context and experience effects. *Transportation Research Part E: Logistics and Transportation Review,* 105**,** 18-38.

R CORE TEAM R: A Language and Environment for Statistical Computing and Graphics. Vienna, Austria: R Foundation for Statistical Computing

RANIERI, L., DIGIESI, S., SILVESTRI, B. & ROCCOTELLI, M. 2018. A Review of Last Mile Logistics Innovations in an Externalities Cost Reduction Vision. *Sustainability* [Online], 10.

REVELT, D. & TRAIN, K. 1998. Mixed logit with repeated choices: households' choices of appliance efficiency level. *Review of economics and statistics,* 80**,** 647-657.

SANTOS, G. 2017. Road transport and CO_2 emissions: What are the challenges? *Transport Policy,* 59**,** 71-74.

SARRIAS, M. & DAZIANO, R. 2017. Multinomial Logit Models with Continuous and Discrete Individual Heterogeneity in R: The gmnl Package. *Journal of Statistical Software,* 72**,** 1-46.

SCHNIEDERJANS, D. G. & STARKEY, C. M. 2014. Intention and willingness to pay for green freight transportation: An empirical examination. *Transportation Research Part D: Transport and Environment,* 31**,** 116-125.

SCHULTE, J. & NY, H. 2018. Electric Road Systems: Strategic Stepping Stone on the Way towards Sustainable Freight Transport? *Sustainability* [Online], 10. Available: https://www.mdpi.com/2071-1050/10/4/1148.

SILLANO, M. & ORTÚZAR, J. D. D. 2005. Willingness-to-pay estimation with mixed logit models: Some new evidence. *Environment and Planning A,* 37**,** 525-550.

TANAKA, M., IDA, T., MURAKAMI, K. & FRIEDMAN, L. 2014. Consumers' willingness to pay for alternative fuel vehicles: A comparative discrete choice analysis between the US and Japan. *Transportation Research Part A: Policy and Practice,* 70**,** 194-209.

TEOH, T., KUNZE, O., TEO, C.-C. & WONG, Y. D. 2018. Decarbonisation of Urban Freight Transport Using Electric Vehicles and Opportunity Charging.

Sustainability [Online], 10. Available: https://www.mdpi.com/2071-1050/10/9/3258.
THALER, R. H. & SUNSTEIN, C. R. 2009. *Nudge: Improving decisions about health, wealth and happiness,* London, Penguin Books.
TIANYU, J. & MENG, L. 2020. Does education increase pro-environmental willingness to pay? Evidence from Chinese household survey. *Journal of Cleaner Production,* 275, 122713.
VAKULENKO, Y., SHAMS, P., HELLSTRÖM, D. & HJORT, K. 2019. Service innovation in e-commerce last mile delivery: Mapping the e-customer journey. *Journal of Business Research,* 101, 461-468.
WANG, X., YUEN KUM, F., WONG YIIK, D. & TEO, C.-C. 2019. Consumer participation in last-mile logistics service: an investigation on cognitions and affects. *International Journal of Physical Distribution & Logistics Management,* 49, 217-238.
WORLD ECONOMIC FORUM 2020. The Future of the Last-Mile Ecosystem. *In:* FORUM, W. E. (ed.) *Transition Roadmaps for Public- and Private-Sector Players.* www.weforum.org.
YANG, J., ZOU, L., LIN, T., WU, Y. & WANG, H. 2014. Public willingness to pay for CO_2 mitigation and the determinants under climate change: A case study of Suzhou, China. *Journal of Environmental Management,* 146, 1-8.

Appendix A

Frequency of choices from the discrete choice experiment. The 460 respondents received 9 choice sets each, resulting in a total of 4140 choices (observations). Table A1 reveals that nearly equal fractions of respondents' chose alternatives 1 or 2, while less than 18% of the choices were opt-outs in terms of choosing alternative 3 (i.e. the alternative stating: "I would not shop if these were the only delivery options").

Table A1: Alternatives chosen in the choice tasks, frequency distribution and percentage. N= 4140

Alternative	1	2	3 (Opt-Out)
Frequency	1717	1689	734

Appendix B

Table B1: Pearson correlation between individual specific variables, including time use in the Discrete Choice Experiment (DCE) part of the questionnaire (Time Use DCE) and inconsistencies between the rating of the assertion "Society should take more care of the environment than what is done today" before and after DCE (Swap)

	Top 25 pop	*Gen Z or Mill*	*Income > 600,000 NOK*	*Empl.*	*Freq. shopper*	*Red cons.*	*Change habits*	*Lower price*	*Free delivery*	*Sustainable shopping*	*Time savings*	*Suppl. info*	*Time use DCE*	*Swap*
College or university	0.19	-0.04	0.19	0.19	0.13	-0.03	0.04	-0.11	0.00	0.04	0.02	0.04	0.02	-0.13
Top 25 pop		0.11	0.01	0.02	0.11	0.02	0.08	0.00	0.05	0.06	0.01	0.04	-0.06	-0.02
Generation Z or Millennials			-0.18	-0.06	0.14	-0.06	-0.04	0.04	0.08	-0.02	-0.12	-0.09	-0.04	-0.03
Income > 600,000 NOK				0.28	0.03	0.02	0.07	-0.11	0.07	0.03	0.12	0.09	-0.01	0.02
Employed					0.06	0.02	-0.10	-0.08	0.10	-0.04	0.07	0.03	-0.05	-0.01
Frequent online shopper						0.02	0.04	-0.02	-0.07	0.05	0.11	0.08	0.06	-0.06
Reduced consumption							0.38	-0.08	-0.09	0.12	0.07	0.02	0.04	-0.11
Change habits								-0.08	-0.20	0.14	0.07	-0.02	0.05	-0.10
Lower price									0.09	-0.20	-0.09	0.02	-0.06	0.11
Free delivery										-0.15	-0.14	0.01	-0.05	0.00
Sustainable shopping											-0.01	-0.02	-0.02	-0.03
Time savings												0.04	0.15	0.05
Supplementary information													0.08	0.05
Time use DCE														-0.02

Appendix C

Table C1: Mixed logit model with price following a normal and a log-normal distribution.

	MMNL - price is normal			*MMNL - price is log-normal*		
	Mean	Median	Standard deviation	Mean	Median	Standard deviation
α1	5.660***			5.581***		
	(0.192)			(0.186)		
α2	5.519***			5.460***		
	(0.187)			(0.183)		
Price	-0.034***		0.025***	-0.037***	-0.027***	0.033***
	(0.002)		(0.002)	(0.003)	(0.002)	(0.004)
Information services	0.219***		0.564***	0.227***		0.482***
	(0.066)		(0.105)	(0.062)		(0.112)
Delivery time	-0.295***	-0.156***	0.471***	-0.275***	-0.145***	0.444***
	(0.024)	(0.012)	(0.077)	(0.02)	(0.011)	(0.062)
Delays	-0.45***	-0.099*	1.997*	-0.991	-0.039	25.272
	(0.072)	(0.045)	(0.8)	(2.216)	(0.035)	(90.281)
CO_2	-1.856***	-0.587***	5.57***	-1.773***	-0.563***	5.294***
	(0.251)	(0.091)	(1.582)	(0.234)	(0.089)	(1.433)
	Model estimate	*Shift in CO_2 mean*	*Shift in CO_2 median*	*Model estimate*	*Shift in CO_2 mean*	*Shift in CO_2 median*
CO_2 x Supplementary information	-0.091	*0.161*	*0.051*	0.029	*-0.052*	*-0.017*
	(0.491)			(0.413)		
CO_2 x Sustainable shopping	1.284*	*-4.849*	*-1.533*	1.427*	*-5.611*	*-1.782*
	(0.539)			(0.710)		
CO_2 x Frequent online shopper	-0.712	*0.945*	*0.299*	-0.619	*0.818*	*0.260*
	(0.428)			(0.587)		
CO_2 x Change habits	0.660	*-1.734*	*-0.548*	0.399	*-0.870*	*-0.276*
	(0.566)			(0.504)		
CO_2 x Employed	-0.075	*0.134*	*0.042*	0.149	*-0.284*	*-0.090*
	(0.420)			(0.573)		
Price x Income > 600 000 NOK	1.58E-04*			-0.005	*0.00018*	*0.00013*
	8.72E-05			(0.003)		
Price x Free delivery	-1.45E-04*			0.004	*-0.00014*	*-0.00010*
	6.92E-05			(0.002)		
Price x Lower Price	-1.35E-04*			0.004*	*-0.00015*	*-0.00011*
	6.20E-05			(0.002)		
Price x Time savings	1.21E-04			-0.005	*0.00018*	*0.00013*
	1.00E-04			(0.003)		
Price x Reduced consumption	8.99E-05			-0.003	*0.00012*	*0.00009*
	5.55E-05			(0.002)		
AIC	5722.504			5731.338		
BIC	5861.729			5870.564		
Log-likelihood	-2839.252			-2843.669		
N	4140			4140		

Elise Caspersen

Foto: T Lauluten

School of Economics and Business
Norwegian University of Life Sciences (NMBU)
P.O Box 5003
N-1432 Ås, Norway

Telephone: +47 6496 5700
Telefax: +47 6496 5701
e-mail: *hh@nmbu.no*
http:/www.nmbu.no/hh

ISSN: 1503-1667
ISBN: 978-82-575-0893-7

Elise Caspersen was born in Oslo, Norway in 1989. She is a Master of Economics from the Norwegian University of Science and Technology (NTNU) and is working as a research economist at the Institute of Transport Economics (TØI) in Oslo, Norway. At TØI she specializes in the field of urban freight transport, logistics and innovation.

Her thesis focuses on freight trip generation and externalities in urban areas and asks the overall research questions i) what is the magnitude of freight trip generation by consumers compared to freight trip generation by establishments, and ii) can externalities of urban freight traffic be reduced by targeting consumer last mile delivery alternatives? These questions have been sought answered through four papers focusing on freight trip attraction from establishments, freight trip attraction from consumers, as well as consumer preferences for environmentally sustainable last mile delivery transport.

The overall finding is that freight trip generation per consumer is still smaller than freight trip generation per establishment (measured using freight trip attraction models for retail activities), but that the large numbers of consumers compared to retail stores makes their contribution non-negligible. Further, the (increasing) role of consumers as generators of freight traffic offers new opportunities to reduce its negative impacts. This thesis shows that one such opportunity is to offer consumers sustainable last mile delivery choices, which several consumer groups seem to prefer. The main contribution is thus that consumers are a key agent in understanding urban freight traffic and should be included in the transition towards environmentally sustainable last mile deliveries, along with establishments and freight operators.

With this thesis, Elise aims to contribute knowledge to municipalities and urban planners working on measures to reduce externalities from urban road freight transport.

Supervisor: Ståle Navrud

E-mail: elc@toi.no
Web page: https://www.toi.no/staff/caspersen-elise-article31929-27.html

ISBN: 978-82-575-1791-5
ISSN: 1894-6402

Norwegian University
of Life Sciences

Postboks 5003
NO-1432 Ås, Norway
+47 67 23 00 00
www.nmbu.no

109532 / ANDVORDGRAFISK.NO