DISS. ETH NO. 27419

UNDERTAKING MOBILITY FIELD EXPERIMENTS USING GPS TRACKING

A dissertation submitted to attain the degree of

DOCTOR OF SCIENCES of ETH ZURICH
(Dr. sc. ETH Zurich)

presented by

JOSEPH ANTHONY MOLLOY
M.Sc. Transport Systems, Technical University of Munich

born on 3 April 1990
citizen of Australia

examiners:

Prof. Dr. K. W. Axhausen, examiner
Prof. Dr. E. Verhoef, co-examiner
Prof. Dr. M. Fellendorf, co-examiner

2021

DOI: 10.3929/ethz-b-000497915

To my family and friends

ABSTRACT

The world we live in is increasingly digital. In only the last decade, the humble mobile phone has developed into a powerful portable computer carried in the pockets of almost the entire population. The integration of GPS-receivers into these devices allows the gathering of data on human mobility at an unprecedented scale and detail. This cumulative dissertation deals with several important aspects of the collection, processing and application of such data. The work presented is all in the context of the MOBIS study, a nationwide field-experiment on mobility pricing, using a smartphone-based GPS tracking app, Catch-my-Day.

First, the recruitment and study design of the MOBIS experiment are presented. The meta-behaviour of the respondents in the study is analysed, identifying important factors influencing the rate of attrition in long duration tracking studies. In the second paper, a method for estimating the external costs of road transportation based on GPS data and using the MATSim framework is developed.

The amount of data collected through such tracking studies is massive, presenting a challenge for the application of discrete choice methods. This is the topic of the third paper. Here, a new software package for R, called mixl, is presented. The final paper presents the initial descriptive results of the impacts of the Covid-19 pandemic on mobility behaviour in Switzerland. Here, the original MOBIS-panel was re-invited to use the tracking app. Drastic changes in travel behaviour are identified as a result of the restrictions, and the implications for future transport policy discussed.

In closing, this dissertation presents a mixture of methodological and applied work centered around the use of smartphone-based GPS tracking in transportation research. The methodological contributions will help inform the design of future tracking studies and field experiments, and the insights gained into mobility behaviour during the pandemic will hopefully stimulate further research efforts.

ZUSAMMENFASSUNG

Die Digitalisierung beeinflusst und verändert die Welt, in der wir leben, immer stärker. Erst in den letzten zehn Jahren hat sich das Mobiltelefon zu einem leistungsfähigen Computer entwickelt, welcher von fast jeder Person täglich in der Tasche mitgeführt wird. Die Integration von GPS-Empfängern in solche Geräte ermöglicht die Erfassung von hochauflösenden Mobilitätsdaten im grossen Stil. Diese Dissertation befasst sich tiefgehend mit verschiedenen Aspekten der Sammlung, Verarbeitung und Anwendung dieser Daten. Die vorgestellten Arbeiten sind alle im Zuge der MOBIS-Studie entstanden, einem landesweiten Feldexperiment zur Preisgestaltung für Mobilität unter Verwendung der Smartphone-basierten GPS-Ortungsapplikation «Catch-my-Day».

In der ersten Arbeit wird zunächst der Rekrutierungsvorgang der MOBIS-Teilnehmer erläutert und das Design des Experiments vorgestellt. Im Fokus steht die Analyse des Verhaltens der Teilnehmer während der Studie. Dabei werden wichtige Faktoren identifiziert, welche zum Abbruch der Teilnahme einer Langzeit-Tracking-Studie führen können. Im zweiten Beitrag wird eine Methode zur Schätzung der externen Kosten des Strassenverkehrs präsentiert, die auf der Grundlage von GPS-Daten und unter Verwendung von MATSim entwickelt wurde.

Für die Modellierung der Verkehrsmittelwahl werden häufig diskrete Entscheidungsmodelle verwendet. Die schiere Datenmenge, die durch digitale Mobilitätstagebücher generiert wird, stellt für solche Modelle aber meist eine grosse Herausforderung dar. In der dritten Arbeit wird deshalb ein Paket für die Statistiksoftware R namens «mixl» vorgestellt, welches eigens dafür entwickelt wurde. Der vierte und letzte Beitrag präsentiert die ersten Ergebnisse der Auswirkungen der Covid-19 Pandemie auf das Mobilitätsverhalten in der Schweiz. Hierzu wurden die MOBIS-Teilnehmer erneut eingeladen, das digitale Tagebuch auf freiwilliger Basis während der Pandemie weiterzuführen. Die drastischen Veränderungen des Mobilitätsverhaltens werden unmittelbar als Folge der nationalen Einschränkungen identifiziert und dessen Auswirkungen auf die zukünftige Verkehrspolitik in der Schweiz ausführlich diskutiert.

Diese Dissertation präsentiert eine Reihe von methodischen und angewandten Analysen in der Verkehrsforschung basierend auf der Verwendung von Smartphone-basierten GPS-Daten. Die methodischen Beiträge helfen da-

bei wesentlich zur Weiterentwicklung zukünftiger Tracking-Studien sowie Feldexperimenten im Bereich der Verkehrsforschung. Die Pandemie gibt der Forschung die einmalige Chance, das Mobilitätsverhalten aus neuen Blickwinkeln zu studieren. Die gewonnenen Erkenntnisse sollen helfen, die Forschung zu inspirieren und stetig voranzutreiben.

ACKNOWLEDGEMENTS

There is a theory about fun: that it comes in two forms, type I and type II. Type I activities are fun while you're doing them, while a type II activity is a struggle while it's happening, but fun in retrospect. For me, undertaking a PhD was a validation of this theory. It involved doubt, challenges and setbacks. However, these were overcome with perseverance, teamwork and a good bit of luck. More importantly, I never would have reached this milestone without those who encouraged, supported and challenged me along the way.

Firstly, I have to thank my supervisor, Professor Kay Axhausen, for giving me this opportunity. It was much easier to be ambitious and take initiative knowing that I always had your advocacy, guidance and encyclopedic knowledge of the literature right next door. I would like to thank Professor Erik Verhoef and Professor Martin Fellendorf for being co-examiners of this thesis, I hope it's an interesting read!

Part of what made the PhD great was the jetsetting around Europe and the world. It's been a pleasure meeting and working with so many intelligent colleagues in Basel, Graz, Lausanne, Munich, Sydney, Tartu and at every "winter" seminar.

To my colleagues at IVT - it's been great fun working on the Hönggerberg with you. Special mentions go to Basil and Felix for their work on mixl, to Chris, my "mobis bro", and to Thomi, my office mate in the Jungle. I couldn't think of a better office partner to have, Muchas Gracias!

To Manu, Nadja and Sean. Thank you for the hundreds of pitches of climbing, pre-work ski tours and 17 hour mountain epics (the real type II fun). But mostly, thank you for the friendship and good times.

To Marijke, your optimism and positive attitude is infectious to everyone around you. Thanks for the great adventures, and your kindness and generosity over the last few years.

And finally, but not least, I have to thank my family and friends in Zurich, around the world, and especially back in Australia. To my siblings James, Kate and Tom: Thank you for all the hilarious moments - London, Hvar, Niseko, Whale Beach and Balmain were just some of them. To Mum and Dad, you are to thank for my insatiable wanderlust and thirst for knowledge, which has led me across the world in search of new challenges, and I hope I can continue to make you proud.

CONTENTS

NOTATION

FREQUENTLY USED SYMBOLS

g	gram
t	metric tonne
CHF	Swiss Franc (monetary unit)
GB	Gigabyte

ABBREVIATIONS

API	Application Programming Interface
ARE	Bundesamt für Raumentwicklung, Switzerland
COVID-19	Coronavirus disease 2019
ERP	Electronic Road Pricing
HBEFA	Handbook for Emissions Factor Analysis
GPS	Global Positioning System
MATSim	Multi-Agent Transport Simulatior
MOBIS	MObility Behaviour In Switzerland
OSM	Openstreetmap
PM	Particulate Matter
PT	Public Transport
SBB	Swiss Federal Railways (Schweizerische Bundesbahnen)
SDK	Software Development Kit

1

INTRODUCTION

Technological advances have been at the center of each paradigm shift in transportation throughout history, with inflection points that have spurred massive changes in how we live and interact with our surroundings. However, it can be argued that the data collection methods necessary to model and hence understand the movement of the population within these transport networks are only just now in the middle of their own inflection point, precipitated by the technological advance of the smartphone, and more specifically, the ubiquity of GPS-receivers in these devices.

A key data collection method in transportation research has been the travel diary, which, since its introduction in 1930s has played an essential role in the understanding of transport behaviour (Axhausen, 1995; Stopher and Greaves, 2007). Traditionally, these were collected through mailed out booklets, and then later through telephone based interviews. In the last 20 years, GPS technology has made the streamlined collection of travel diaries possible, at a level of detail infeasible with paper and telephone based methods. Early work using GPS required dedicated tracking units, which the user had to carry with them and return to the researchers. This process is expensive and involved. Initially, GPS devices were only used to validate the results of paper or telephone based surveys (Stopher, FitzGerald, and Xu, 2007; Wolf, Loechl, et al., 2003). Shen and Stopher (2014) provide a comprehensive list of studies using dedicated devices.

It is only in the last two decades that GPS-receivers have been integrated into mobile phones. Furthermore, the consolidation of the smartphone market into essentially two leaders, Apple and Android, has made it possible to develop a single digital travel diary application downloadable and usable by close to 100% of the population in a developed country. Montini et al. (2015) compared the performance of self-contained devices and GPS-enabled smartphones and found the performance varied, with accuracy being better with the dedicated GPS tracker, but sufficient with the mobile phone for routing and trip-diary purposes.

As an alternative, mobile tracking data can also be passive. That is, the location of the device as determined by the network operator, usually through triangulation, based on the connections that the mobile device makes to the antennas in the mobile network.

As such, the work in this thesis is motivated by the desire to understand the capabilities and limits of tracking data in the context of mobility studies, and to utilise the technology in the study of current transport-related challenges facing our society.

Two field-experiments are presented in the following chapters. The first is an empirical investigation into the effectiveness of mobility pricing as tool to internalise the external costs of transport. Economists argue that pricing mechanisms are the most effective way to deal with the continuously growing problem of congestion (Lindsney and Verhoef, 2001). However, to be most effective, prices in the implemented scheme would need to vary according to the severity of congestion (Vickrey, 1963). Furthermore, there is increasing evidence that in multimodal transport networks, it is insufficient just to price road transport, and that the interplay between demand and supply across the multi-modality of the transport network is important (Tirachini, Hensher, and Rose, 2014).

Currently, usage of the road network in most cases is priced through three components: a registration fee, a fuel tax and an optional flat yearly highway-usage fee. None of these charges incentivise users to avoid travelling on congested roads. Fuel taxes are also becoming increasingly ineffective as cars become more fuel-efficient. The gradual electrification of the vehicle fleet will only exacerbate this trend. Levinson (2010) identify additional reasons, which include the need to raise revenue for future investment. While there has been stated-preference work into the effects of mobility pricing (Brownstone and Small, 2005; Ettema, Ashiru, and Polak, 2004; Vrtic et al., 2010), there have only been a small number of real world experiments (Nielsen, 2004; Transurban, 2016). These experiments have been limited to road pricing, and have not considered the external costs of the transport system as a whole.

Finally, the political acceptance of mobility pricing continues to be an issue, and both these experiments, as well as the experience of the London and Stockholm congestion charges showed that participants were overall more accepting of pricing measures after exposure (Transport for London, 2004; Winslott-Hiselius et al., 2009). Bringing these factors together, there is a multitude of urgent reasons to further understand the impacts of mobility pricing in a real world setting, and hopefully accelerate the political acceptance of mobility pricing.

The second experiment, while not by design, is one of the largest sudden disruptions of global mobility in modern history. The onset of the Covid-19 pandemic saw drastic changes in people's daily lives. In Switzerland, a lockdown was imposed in March 2020 to slow the spread of the virus. A rapid shift to remote-work and a slow reopening of the economy followed. By reactivating the dormant GPS tracking panel from the mobility pricing experiment, it has been possible to understand just how people were adapting their mobility behaviour to the pandemic and restrictions. Here, the motivations were two fold: first, there was the potential to aid short-term policy-making during the pandemic, and then second, discover lessons that may inform long-term transport policy-making in the post-pandemic world.

1.1 BACKGROUND AND STATE-OF-THE-ART

Data collected from mobile devices first started to play a role in transportation research in the 2000's, starting with Asakura and Hato (2004), who used call detail records (CDR) to investigate the feasibility of using mobile network data to study metropolitan-scale travel behaviour. At this stage, the first iPhone was still a prototype, and it would not be until 2008 that the first iPhone with integrated GPS was released (Apple, 2008). Further work using aggregated mobile network data continued (Ahas et al., 2010; Anda, Medina, and Fourie, 2018; Gonzalez, Hidalgo, and Barabasi, 2008). However, the benefits of passive mobile network data - namely the non-existent response burden and sample size come with trade-offs which need to be acknowledged. Privacy laws normally ensure that the traces must remain anonymous, meaning that minimal, if any demographics are available. Secondly, particularly in Europe, tracking of participants over consecutive days is rarely allowed without their explicit permission (Cik et al., 2020).

While cellular network data is collected passively by the network operator, the collection of GPS data requires the installation of an app on the device. However, the spatial and temporal accuracy of the data is very good, rarely exceeding 30m in outdoor settings (Zandbergen and Barbeau, 2011). In contrast, the accuracy of cellular network positioning is not comparable to that of GPS data (Widhalm et al., 2015). On mobile phones, the data from the GPS receiver is often augmented with other available sources, such as WiFi, cell tower triangulation and accelerometers to improve accuracy, which are collectively known as *location services*. For the purpose of this thesis these location services will be referred to as GPS, as is common in the literature. GPS data can either be collected anonymously through aggregation services at a large scale (Buck et al., 2014), or through the recruitment of participants, who are asked to install and use a specific app. For studies where a representative sample is required, or behavioural models with socio-demographic variables are to be developed, the recruitment of the participants is naturally the preferred option.

1.1.1 *GPS tracking in transport behaviour research*

Over the last 20 years, GPS has been increasingly used both in the collection of travel diaries and the understanding of daily patterns (Wolf and Guensler, 2000). Many of these studies have been of short duration, i.e. a couple of days (Allström, Kristoffersson, and Susilo, 2017; Greene et al., 2012), due to the acknowledged response burden and resulting attrition observed in these studies (Kohla and Meschik, 2013). As Widhalm et al. (2015) note, the sample size and observation period of most GPS studies is still limited. A table of recent GPS-based travel surveys is presented in chapter 2.

In addition, continuous methodological advancements have been made in the processing of collected GPS data. The process is normally divided into stages: filtering, trip segmentation, transport mode detection and, where necessary, map-matching to a network. For an overview of the various methods and advances, see Shen and Stopher (2014) and Zheng (2015). Graphhopper (Karich and Schröder, 2014) provides an open-source framework for map-matching GPS traces to Openstreetmap (OSM) (Haklay and Weber, 2008) network using the Hidden Markov Model approach developed by Newson and Krumm (2009).

1.1.2 *Theoretical work on external costs of transport*

The cost of travel in transportation can be divided into two categories - individual costs - i.e. those payed by the traveller, and the costs of travel bourn by other individuals, known as external costs. In transport, the external costs are divided into the following groups: congestion, environmental effects, noise damages, and accidents. (Verhoef, 2000).

The theoretical foundations for road pricing were laid by Pigou (1920) in his work on the internalisation of the external costs in a market. He used the example of two roads to suggest that differential taxes can be useful in increasing the overall utility in the simple network where congestion is present. Knight (1924) explored this further, but suggested that a government intervention would be required, and tolls should be set by private operators.

Vickrey (1963) used a bottleneck model to demonstrate how road pricing could influence travellers' choice of route and transport mode. Vickrey demonstrated that in perfect congestion pricing, tolls must match the severity of congestion, and vary by time of day, location, type of vehicle and current conditions. It then follows that transport users should be charged their marginal external costs - costs which would otherwise be absorbed by other transport users and society (Button and Verhoef, 1998).

The methods for internalising these external costs can be categorised by their level of optimality. In first-best pricing, the marginal external cost is charged to the user. In this case, both the charging mechanism and the amount charged need to be optimal (Verhoef, 2000). In second-best pricing, the pricing mechanism is guided by the principle of marginal external costs, but the implemented scheme is simplified (Small, Verhoef, and Lindsey, 2007).

Most of the early work on the pricing of externalities focused on the different transport networks in isolation - the congestion on the road network was not considered in the context of public transport or non-motorized modes. Multiple researchers have identified how this is insufficient. Small (2008) argued that a road congestion charge can act as an effective way to financially support public transport, by raising the cost of car travel, which has been traditionally cheap as the numerous external costs are not paid by the driver. Tirachini and Hensher (2012) also examined the intricate relationship of pricing between car travel, public transport and non-motorized modes, in both a first-best and second-best context. As is acknowledged in the literature, there exist limitations to implementing first-best pricing. Verhoef (2000) identifies both general issues such as the "limited social and political acceptability and the technical feasibility of marginal external cost pricing" and the unlikelihood that the assumptions required for effective Pigouvian taxes apply.

1.1.3 *Mobility pricing in practice*

Although the theoretical challenges of efficiently internalising the external costs in transport have been discussed for decades, it is only more recently that solutions have been implemented for personal vehicle travel, either experimentally or in reality. Schemes for freight traffic are already widespread, with Germany being one example (Link, 2008).

In the AKTA study (Nielsen, 2004), 500 drivers around Copenhagen had a GPS receiver installed in their vehicle (before the widespread availability of GPS-enabled mobile devices). After a control period, they were exposed to different road pricing schemes over 8-12 weeks, with the charges calculated based on the data from the GPS receiver.

The Melbourne road pricing experiment also investigated the feasibility of a road pricing scheme for a sample of customers for an Australian tollway operator in metropolitan Melbourne (Transurban, 2016). It was conducted over 17 months with 1,635 participants, and a range of charging schemes. The results showed that such as system could act as a significant funding source for new transport investment, and help manage demand in congested areas and peak hours. Here, as in the AKTA study, a GPS device was installed in the participant's car for the duration of the study.

The London congestion charge is one of the most well known real world implementations (Leape, 2006; Santos and Shaffer, 2004). First applied in 2003, it has seen numerous extensions since, including additional charges for heavily polluting vehicles and discounts for electric cars. In Stockholm, road pricing has also been implemented in the form of a congestion charge (Eliasson et al., 2009). An analysis by Karlström and Franklin (2009) showed that drivers crossing the toll cordon (perimeter) boundaries were 15% more likely to switch to public transport. In both cases, cars entering a cordon around the central business district have to pay a charge during certain times. These second-best schemes only include road travel, and are relatively blunt instruments.

Over the years, a variety of road pricing schemes have been implemented in Singapore, starting in 1975 with a paper-based peak-hour permit scheme (Chin, 2005). In 1994, the system was revised to offer two levels of licensing. Starting in 1997, the ERP (Electronic Road Pricing) was introduced, in which vehicles are charged each time they pass through a gantry (Agarwal and Koo, 2016). The charges are regularly adjusted to maintain a certain level of service. In 2023 a satellite (GPS) based system will be introduced (Tan, 2020).

Solutions involving tradable permits have also been proposed (Verhoef, Nijkamp, and Rietveld, 1997), but were assumed to be an academic curiosity due to the large number of actors in the system, among other reasons. Brands et al. (2020) demonstrated the feasibility of a tradable permit scheme for road pricing, using a virtual experiment where participants were asked to trade permits for their usual commute throughout the week.

1.1.4 *External costs in agent-based simulation*

Verhoef (2000) notes that the welfare benefits of such systems need to be considered on an individual level. One way to do this is to capture the heterogeneity in individual behaviour through the use of agent-based modelling. Chakirov (2016) used the agent-based transport modelling framework MATSim (Horni, Axhausen, and Nagel, 2016) to investigate mobility pricing, and the interacting effects when congestion charging and dynamic public transport pricing are combined in one setting.

Hülsmann et al. (2011) developed the emissions model for MATSim, which takes the HEBFA database (Keller et al., 2017) and calculates the pollutant emissions for agents in the network. This module was developed further by Kickhöfer, Hülsmann, et al. (2013), and applied to calculate time-dependent, vehicle-specific pollutant exposure tolls in an agent-based scenario (Kickhöfer and Kern, 2015). Here, the factors for the monetisation of the emissions were taken from Maibach et al. (2008).

The internalisation of externalities using agent-based simulation has also been investigated using the MATSim framework. Kaddoura, Kickhöfer, et al. (2015) looked at optimal public transport pricing to internalise the marginal social cost of crowding and waiting on public transport, based on a multi-modal corridor scenario. Kaddoura (2015) applied a marginal external cost (MEC) approach to road congestion for the MATSim greater Berlin scenario, where the delay caused by an agent in the simulation to each other affected agent is calculated at each link exit in the network. The marginal social costs generated by each agent can be calculated based on these delays. Kaddoura, Agarwal, and Kickhöfer (2017) considered the simultaneous congestion, noise and air pollution pricing for the greater Berlin area using MATsim under Pigouvian assumptions, using the aforementioned emissions module and MEC approach to road congestion.

1.1.5 *Mobility monitoring during pandemics*

It has been widely acknowledged that transportation is a key driver in the spread of infectious diseases (Baroyan and Rvachev, 1967; Herrera-Valdez, Cruz-Aponte, and Castillo-Chavez, 2011). The important role of mobility in a pandemic has been demonstrated for historical pandemics such as the Spanish flu in 1918 (Ammon, 2002; Trilla, Trilla, and Daer, 2008). More recently Wesolowski et al. (2015) used mobile phone data to predict the spread of the dengue epidemics in Pakistan. For a comprehensive overview of studies exploring the link between transport and infectious diseases, see Muley et al. (2020).

In the current pandemic, mobile data has shown to be useful in understanding the role of both regional and global mobility during the pandemic. Many countries are now using mobile data to understand the effectiveness of measures, including Austria, Belgium, Chile, China, Germany, France, Italy, Japan, Spain, United Kingdom and the United States (Oliver et al., 2020; Yabe et al., 2020).

During the COVID-19 pandemic, anonymised aggregated mobile data has been a key new technological tool in monitoring mobility. Vinceti et al. (2020) used aggregated mobile data for three regions in Italy (Lombardy, Veneto and Emilia-Romagna) between February 2020 and April 2020 to monitor mobility during the first and second lockdowns. They determined that a relaxed lockdown did not reduce mobility sufficiently to slow the virus.

Iacus et al. (2020) used such data to explore the correlation between mobility and the number of positive tests in regions of France, Spain and Italy. They argue that reducing internal mobility is more important than mobility across provinces in the spread of the disease.

Heiler et al. (2020) investigated the nation-wide changes in mobility during the first European wave for Austria using real-time anonymised mobile phone data. They saw a doubling of the number of persons with a radius of gyration (activity space) of less than 500m, and increased segmentation of the community structure.

In the USA, Badr et al. (2020) found a strong correlation between reduced mobility behaviour and decreased COVID-19 case growth rates. Furthermore, they show evidence that behavioural changes were already observable days to weeks before movement restriction policies were implemented.

There has also been a wealth of work looking at the initial spread of the disease in China, with the help of anonymous mobile phone data (Jia et al., 2020; Kraemer et al., 2020; Xiong et al., 2020; Zhou et al., 2020). However, as the data sets used in these works have a few key limitations: they are less effective at capturing mobility changes at local urban scales, and individual socio-demographic variables are generally not available. This is important for both understanding which groups are most affected by restrictions, and understanding which population segments were best able to adapt to the changed conditions. Indeed, even the quarantine measures introduced to combat the 1636 outbreak of the plague were found to affect different classes of society differently (Newman, 2012).

1.2 METHODS

The use of location tracking apps is the central methodology in this thesis. Harding et al. (2020) compared the performance of a wide range of the tracking apps available. In the following work, the Catch-my-Day app, developed by Motiontag GmBH, was used. Although Catch-my-Day was developed specifically for use in mobility studies undertaken by the Institute for Transport Systems and Planning, ETH Zurich, it uses the same interface and processing backend as Motiontag's publicly available app.

The Motiontag platform is provided as an API and SDK for Android and iOS. The SDK was designed to minimize the use of GPS by taking advantage of other sensors in the phone where possible. The data is sent to the Motiontag servers where it is enriched with Openstreetmap data and segmented into stays (activities) and trips. The trips are then divided into stages using a Deep Recurrent Neural Network (RNN) which identifies the transport modes and transfer points on the trip. The output is a travel diary, constituting stages (labeled with the transport mode) and activities (labeled with the activity purpose). Figure 1.1 shows the main interface screens of the Catch-my-Day app.

1.2.1 *The mobility pricing experiment*

The work in Chapters 3, 4 and 2 needs to be framed in the context of the MOBIS mobility pricing field experiment. The MOBIS experiment was an eight-week mobility pricing study undertaken in Switzerland between September 2019 and January 2020. Participants were invited by post to participate, and offered the incentive of 100 CHF [1] for the completion of the study. In an introduction survey they were asked for socio-demographic information, and were posed questions about their attitudes to transport policies and problems, as well as values using the method developed by

1 around $100 USD

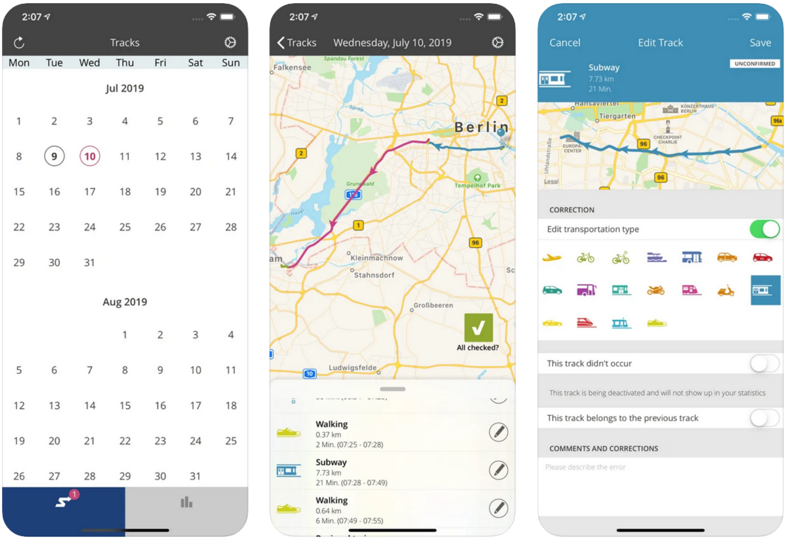


FIGURE 1.1: The Catch-my-Day interface. From left to right: 1) Calendar home page. 2) Daily view showing recorded trips. 3) Editing the mode of a selected trip.

Otte (2008). The introduction survey also screened participants against eligibility criteria. The criteria were:

- aged between 18 and 65
- living in urban agglomerations (based on a postcode list),
- not a professional driver
- able to walk
- have access to a smartphone for downloading the app

Eligible participants who were willing to participate were then given a registration code and directions to install the Catch-my-Day app. They could start the tracking whenever they wanted, and the 8 weeks would start from the first complete tracking day. If participants were identified as not tracking for a certain number of days, they were notified by email, with the aim to increase the quality of the tracking data and reduce the dropout rate. In the first 4 weeks, participants received a weekly email with a summary of their kilometers travelled with each transport mode. Participants who provided less than 12 days of tracking in the first 4 weeks were not allowed to participate in the second phase.

After the 4 week control phase, each participant was randomly assigned to either one of two treatment groups, or a continued control. Those in the treatment groups were informed by email that the conditions of the study were changing. In the first treatment group *information*, participants only received weekly information on their external costs in the previous week. In the *pricing* group, participants received the external cost information, as well as a budget for the treatment phase, where any remaining budget would be transferred to them at the conclusion of the study. The budget for the pricing group in the treatment phase was set at 120% of their external costs in the control phase. Additionally, to discourage gaming of the experiment, the budget was reduced by $1/28^{th}$ for each day where no tracking data was provided over the 4 weeks of the treatment phase. During the treatment phase, participants were provided with a weekly email report updating on their external costs in the previous 7 days, as well as their remaining budget if applicable. Due to the rolling start to the experiment, participants received these reports on different days of the week. As such, a pipeline was developed to download the new data from Motiontag and calculate the external costs and mail the reports each evening. The methodology to calculate external costs is covered in detail in chapter 3.

1.2.2 *External costs in the mobility pricing experiment*

The external costs presented to participants were grouped into 3 categories: CO_2, congestion and health. CO_2 is the external cost of CO_2, N_2O and methane (CH_4), in CO_2 equivalent emissions of a particular mode of travel. Congestion is the link-based time lost caused to other drivers in the case of car travel, and the peak-hour transit surcharge for transit modes. Health is broader, and groups together the impacts of particulate matter (PM), noise, nitrous oxides (NO_x), accident costs and the (positive) health benefits of active mobility such as cycling and walking.

PARTICULATE MATTER Particulate emissions were priced according to the urbanity of the trip. Using the the development zoning classification for Switzerland (ARE, 2017), each link in the MATSim network was designated as urban or rural, and PM emissions were computed for each, and priced accordingly.

NOISE After determining that a dis-aggregated noise model developed by Kaddoura, Kröger, and Nagel (2017) would not scale to a national level, a fixed per-Km cost of noise emissions was used, with the value taken from the Swiss norms.

NO_x The health externalities from NO_x emissions were priced using the CHF/t value in the Swiss norms.

ACCIDENT COSTS The Swiss norms provide accident costs per km travelled for each mode. In contradiction to multiple studies (Gössling et al., 2019; Sælensminde, 2004), the Swiss norms give external accident costs for cycling which outweigh the health benefits, giving an overall negative external cost for cycling.

HEALTH BENEFITS For walking and cycling, the Swiss norms stipulate a health benefit of 0.1870 CHF/km and 0.1863 CHF/km, and accident costs of 0.075 CHF/km and 0.257 CHF/km respectively. This gives a total external cost of -0.112 CHF/km and 0.070 CHF/km for cycling and walking respectively. The calculation of external costs was performed within the agent-based transport framework MATSim (Horni, Axhausen, and Nagel, 2016).

To calculate the emissions from car travel, the emissions module for MATSim (Hülsmann et al., 2011; Kickhöfer, Hülsmann, et al., 2013) was used to calculate the pollutant emissions of a driver on each link based on the Hbefa (De Haan and Keller, 2004; Keller et al., 2017). The congestion externalities were estimated using the marginal-cost method developed by Kaddoura (2015). Here, a constant value of time was used for all agents in the simulation, in this case a necessary violation of the constraints required for true first-best pricing. Walking and cycling externalities were calculated on a per-Km basis using the normative values provided by the relevant Swiss authorities, updated for 2019 (ASTRA, 2017).

1.2.3 *Dynamic public transport pricing*

For public transport externalities, pollution emissions were calculated on a per-Km basis, using the values in Table 1.1. These provided estimates of per-person-km emissions in grams for the key pollutants for which monetary values are available from Table 3.6. In contrast to private car travel, the marginal social cost of public transport (in terms of pollution

Mode	CO_2	PM_{10}	NO_x	Accidents	Noise	Health [a]
Train	0.000066	0.0140	-	0.00066	0.0087	-
Bus	0.0144	0.0437	0.5440	0.0141	0.0257	-
Tram	-	-	-	0.0126	0.0075	-
Bicycle	-	-	-	0.257	-	-0.1870
Walk	-	-	-	0.075	-	-0.1863

TABLE 1.1: Per-km Monetary costs used in the MOBIS experiment. Note the zero pollution values for Tram - None were provided in the NISTRA. This has a minimal effect on the study results, as they would be similar to Train, i.e. Minimal.

a Note that the negative values indicate a benefit

and noise) decreases as the demand increases. On the other hand, crowding affects willingness to pay and can be seen as a form of congestion in public transport, and delay in some circumstances (Tirachini, Hensher, and Rose, 2013). However, crowding effects are extremely heterogeneous, both spatially and temporally. Even in peak hours, crowding can be restricted to particular transit lines during very short periods (Zurich Public Transport (VBZ), 2017). As such, it would be unreasonable to distribute the crowding effects in an aggregate measure across all peak-hour travellers in a specific public transit region. Additionally, for each public transport operator, data would have to be collected separately and collated as it is not available on a national level. As an alternative solution, a zonal peak-hour surcharge pricing scheme was developed for the national public transport network, as a form of second-best pricing. Throughout the experiment, participants had access to a interactive map which showed them where and when the pricing scheme applied (see fig. 1.2. The peak-hour pricing surcharge applied a 0.1 CHF/km surcharge on public transport trips between those zones with a larger demand in peak-hours compared to the off-peak.

The zoning system was based on the *Gemeinde*[2], with large urban municipalities such as Zurich being split into their *Kreise*[3]. The surcharge was applied to transit stages between any two zones which experience peak-hour demand. The peak-hour windows and the affected zone-pairs were determined using the output of the MATSim scenario for Switzerland

2 A *Gemeinde* in Switzerland is roughly translatable as a municipality.
3 A *Kreise* in Switzerland is roughly translatable as a municipal district

(Bösch, Müller, and Ciari, 2016). The peak windows were set as 7am-9am and 5pm-7pm, and not adapted for regional variations in working patterns. Municipality pairs were priced if the maximum hourly transit trip count during peak-hour was greater than three times the average hourly transit trip count during the daily off-peak (9am - 5pm) for that pair. A municipality could also be paired with itself if the the above criteria was met, and the direction of the peak-hour flow is not considered. If the trip was partially in both the peak and off-peak periods, only the proportion of the travel duration that overlaps with the peak period was charged.

1.2.4 *Estimation of discrete choice models*

Discrete choice methods are an important methodological tool for modelling the impacts of transport policies. In particular, they are useful for both generating value-of-time estimates and understanding how a mobility pricing scheme affects mode share. However, the amount of data generated by a tracking study like MOBIS is extremely large. The tracking of 3,680 participants over 8 weeks generated over 1.4 million travel stages in 935,570 trips. Each of these stages or trips would count as an observation in the discrete choice model. Sampling methods have been developed for working with such large data sets (Cranenburgh and Bliemer, 2019). However, even with a 10% sample, the data set is still very large by discrete choice modelling standards. With many available packages, the memory requirements for more complicated models with many random components that require simulated estimation become prohibitive, requiring over 200 GB of memory, as shown in chapter 4. As such, there is currently a methodological gap here, which the work in chapter 4 seeks to fill. This is done by taking advantage of some methods from computer science. Namely the translation and complication of computing code from one syntax to another, and the use of data parallelism to split the problem over multiple processing cores (Subhlok et al., 1993). The high performance parallel computing framework OpenMP (Chapman and Massaioli, 2005) was used to formulate the estimation of the log-likelihood in discrete choice in the data-parallelism framework.

1.3 RESEARCH OBJECTIVES AND QUESTIONS

The MOBIS mobility pricing project had the ambitious goal of determining if the consideration of the external costs - through either information or incentivisation - would lead to significant behavioural change in the Swiss

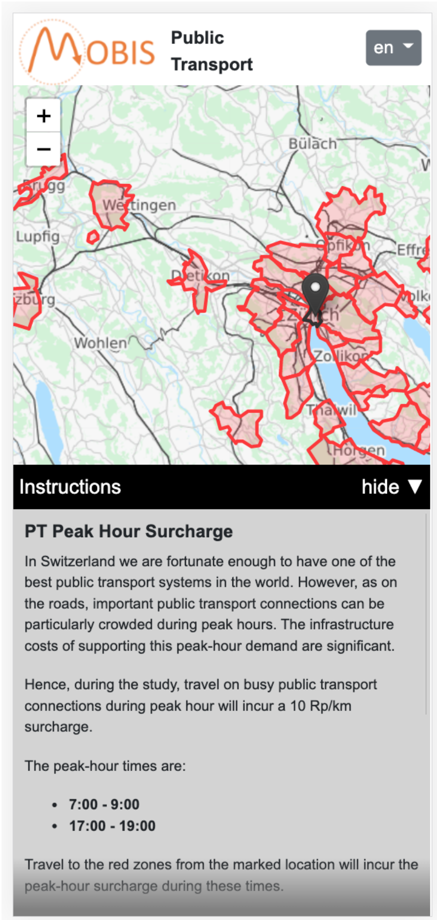


FIGURE 1.2: The PT surcharge guide as viewed on a mobile device. The red areas indicate destination zones which would be charged during peak-hour from the black zone. The map is movable and zoomable by the user.

population. The goal was to do this through a real-world field experiment. This thesis is concerned with some of the research questions that arose as a result of the work on this project. The following list corresponds to the order of the papers presented in this thesis.

1. Which factors affect the attrition rate in a long term GPS tracking study? Does the duration of participation vary across the population or between treatment groups? How does the response rate compare to other studies in the literature? Are app-based tracking studies sufficiently accurate and advanced enough to support such a mobility field experiment?

2. How can external costs be accurately imputed at a trip level, taking advantage of the temporal and spatial resolution in GPS data? Does a dis-aggregate link-level approach based on an agent-based framework method better capture the heterogeneity in external costs compared to an average approach?

3. Can the scalability issues with current estimation software for discrete choice models in R be overcome? Is it possible to improve the performance to the extent that complex models on large data sets are feasible to estimate in R, such as those required by the MOBIS project? Can such performance be achieved while providing a simple model specification syntax?

4. The final research questions followed after the completion of the MOBIS tracking period was completed, as a response to the Covid-19 pandemic. How is the mobility behaviour in Switzerland changing in response to the measures introduced to control the pandemic? How do these responses vary by socio-demographics and trip purpose?

1.4 OVERVIEW OF THE THESIS

In this thesis, the above research questions are answered over the span of four research papers, presented in the following chapters.

In chapter 2, The recruitment methodology for the MOBIS study is presented, along with a detailed analysis of the response rates and attrition in the tracking phase of the study. An analysis of the survey-behaviour of the participants follows, and the influence of the mobile operating system on survey participation is quantified.

In chapter 3, a methodology for the link-level calculation of external costs on pre-processed GPS traces is presented. The resulting model is validated against the official average values and various discrepancies discussed. To evaluate the usefulness of the model, the individual external costs calculated for the MOBIS study are compared against a simple per-Km approach to quantify the extra heterogeneity captured using the new methodology.

In chapter 4, a new software package *mixl* for the R programmming language is presented. This package overcomes current scalability limits in the simulated estimation of mixed multinomial logit models, which have been restricting the complexity of the models which could be estimated as the number of observations and random draws increased. It also utilises multiple processing cores better than other open source software to improve the estimation time.

chapter 5 presents a first look at the impact of the restrictions and resulting relaxations on mobility behaviour in Switzerland captured by a re-activation of the MOBIS cohort. The detailed socio-demographic information available for each participant, as well as the travel diaries processed from the collected GPS data give a detailed insight into how mobility behaviour has changed as a result of the pandemic.

Finally, in chapter 6, the four papers are brought together and the limitations and potential for future work presented. The role the methods and results presented in the above chapters have to play in ongoing transportation research is discussed, as well as the immediate and future impact of the work on society.

REFERENCES

Agarwal, Sumit and Kang Mo Koo (2016) Impact of electronic road pricing (ERP) changes on transport modal choice. In: *Regional Science and Urban Economics* 60, pp. 1–11.

Ahas, Rein, Siiri Silm, Olle Järv, Erki Saluveer, and Margus Tiru (2010) Using mobile positioning data to model locations meaningful to users of mobile phones. In: *Journal of urban technology* 17 (1), pp. 3–27.

Allström, Andreas, Ida Kristoffersson, and Yusak Susilo (2017) Smartphone based travel diary collection: Experiences from a field trial in Stockholm. In: *Transportation research procedia* 26, pp. 32–38.

Ammon, CE (2002) Spanish flu epidemic in 1918 in Geneva, Switzerland. In: *Eurosurveillance* 7 (12), pp. 190–192.

Anda, Cuauhtemoc, Sergio A Ordonez Medina, and Pieter Fourie (2018) Multi-agent urban transport simulations using OD matrices from mobile phone data. In: *Procedia computer science* 130, pp. 803–809.

Apple (2008) Apple Newsroom - Apple Introduces the New iPhone 3G. URL: https://www.apple.com/newsroom/2008/06/09Apple-Introduces-the-New-iPhone-3G/ (visited on 11/10/2020).

Asakura, Yasuo and Eiji Hato (2004) Tracking survey for individual travel behaviour using mobile communication instruments. In: *Transportation Research Part C: Emerging Technologies* 12 (3-4), pp. 273–291.

Axhausen, K. W. (1995) Travel diaries: An annotated catalogue (Working paper). In: *Innsbruck, Austria: Institut für Strassenbau und Verkehrsplanung, Leopold-Franzens-Universität*.

Badr, Hamada S, Hongru Du, Maximilian Marshall, Ensheng Dong, Marietta M Squire, and Lauren M Gardner (2020) Association between mobility patterns and COVID-19 transmission in the USA: a mathematical modelling study. In: *The Lancet Infectious Diseases* 20 (11), pp. 1247–1254.

Baroyan, OV and LA Rvachev (1967) Deterministic models of epidemics for a territory with a transport network. In: *Cybernetics* 3 (3), pp. 55–61.

Bösch, Patrick M, Kirill Müller, and Francesco Ciari (2016) The ivt 2015 baseline scenario. In: *16th Swiss Transport Research Conference (STRC 2016)*. Ascona, Switzerland.

Brands, Devi K, Erik Verhoef, Jasper Knockaert, and Paul R Koster (2020) Tradable permits to manage urban mobility: market design and experimental implementation. In: *Transportation Research Part A: Policy and Practice* 137, pp. 34–46.

Brownstone, David and K. A. Small (2005) Valuing time and reliability: assessing the evidence from road pricing demonstrations. In: *Transportation Research Part A: Policy and Practice* 39 (4), pp. 279–293.

Buck, Christoph, Chris Horbel, Tim Kessler, and Claas Christian (2014) Mobile consumer apps: big data brother is watching you. In: *Marketing Review St. Gallen* 31 (1), pp. 26–35.

Bundesamt für Strassen (ASTRA) (2017) Indikatoren für STRAsseninfrastrukturprojekte (NISTRA). URL: https://www.astra.admin.ch/astra/de/home/fachleute/dokumente-nationalstrassen/fachdokumente/nistra.html (visited on 11/10/2019).

Button, Kenneth and Erik Verhoef (1998) Road pricing, traffic congestion and the environment. Edward Elgar Publishing.

Chakirov, Artem (2016) Urban mobility pricing with heterogeneous users. PhD thesis. ETH Zurich.

Chapman, Barbara M. and Federico Massaioli (2005) OpenMP. In: *Parallel Computing* 31 (10), pp. 957–1174.

Chin, Kian-Keong (2005) Road pricing–Singapore's 30 years of experience. In: *CESifo DICE Report* 3 (3), pp. 12–16.

Cik, Michael, Alexandra Lechner, Cornelia Hebenstreit, and Martin Fellendorf (2020) Activity estimation from mobile phone data. In: *Proceedings of the 99th TRB Annual Meeting*.

Cranenburgh, Sander van and Michiel C.J. Bliemer (2019) Information theoretic-based sampling of observations. In: *Journal of Choice Modelling* 31, pp. 181–197.

De Haan, Peter and Mario Keller (2004) Modelling fuel consumption and pollutant emissions based on real-world driving patterns: the HBEFA approach. In: *International journal of environment and pollution* 22 (3), pp. 240–258.

Eliasson, Jonas, Lars Hultkrantz, Lena Nerhagen, and Lena Smidfelt Rosqvist (2009) The Stockholm congestion–charging trial 2006: Overview of effects. In: *Transportation Research Part A: Policy and Practice* 43 (3), pp. 240–250.

Ettema, Dick, Olu Ashiru, and J. W. Polak (2004) Modeling timing and duration of activities and trips in response to road-pricing policies. In: *Transportation Research Record* 1894 (1), pp. 1–10.

Federal Office for Spatial Development (ARE) (2017) Bauzonenstatistik Schweiz. URL: `https://www.are.admin.ch/are/de/home/raumentwicklung-und-raumplanung/grundlagen-und-daten/fakten-und-zahlen/bauzonen.html` (visited on 11/10/2020).

Gonzalez, Marta C, Cesar A Hidalgo, and Albert-Laszlo Barabasi (2008) Understanding individual human mobility patterns. In: *nature* 453 (7196), pp. 779–782.

Gössling, Stefan, Andy Choi, Kaely Dekker, and Daniel Metzler (2019) The social cost of automobility, cycling and walking in the European Union. In: *Ecological Economics* 158, pp. 65–74.

Greene, Elizabeth, Leah Flake, Kevin Hathaway, and Michae Geilich (2012) A Seven-Day Smartphone-Based GPS Household Travel Survey in Indiana. In: *Proceedings of the TRB 95th Annual Meeting*. Transportation Research Board.

Haklay, Mordechai and Patrick Weber (2008) Openstreetmap: User-generated street maps. In: *IEEE Pervasive Computing* 7 (4), pp. 12–18.

Harding, Chris, Ahmadreza Faghih Imani, Siva Srikukenthiran, Eric J Miller, and Khandker Nurul Habib (2020) Are we there yet? Assessing smartphone apps as full-fledged tools for activity-travel surveys. In: *Transportation*, pp. 1–28.

Heiler, Georg, Tobias Reisch, Jan Hurt, Mohammad Forghani, Aida Omani, Allan Hanbury, and Farid Karimipour (2020) Country-wide mobility changes observed using mobile phone data during COVID-19 pandemic. In: *arXiv preprint arXiv:2008.10064*.

Herrera-Valdez, Marco Arieli, Maytee Cruz-Aponte, and Carlos Castillo-Chavez (2011) Multiple outbreaks for the same pandemic: Local transportation and social distancing explain the different" waves" of A-H1N1pdm cases observed in México during 2009. In: *Mathematical Biosciences & Engineering* 8 (1), p. 21.

Horni, Andreas, K. W. Axhausen, and Kai Nagel (2016) The multi-agent transport simulation MATSim. Ubiquity Press.

Hülsmann, Friederike, Regine Gerike, Benjamin Kickhöfer, Kai Nagel, and Raphael Luz (2011) Towards a multi-agent based modeling approach for air pollutants in urban regions. In: *Proceedings of the Conference on "Luftqualität an Straßen". Bundesanstalt für Straßenwesen*.

Iacus, Stefano Maria, Carlos Santamaria, Francesco Sermi, Spyros Spyratos, Dario Tarchi, and Michele Vespe (2020) Human mobility and COVID-19 initial dynamics. In: *Nonlinear Dynamics* 101 (3), pp. 1901–1919.

Jia, Jayson S, Xin Lu, Yun Yuan, Ge Xu, Jianmin Jia, and Nicholas A Christakis (2020) Population flow drives spatio-temporal distribution of COVID-19 in China. In: *Nature*, pp. 1–5.

Kaddoura, Ihab (2015) Marginal congestion cost pricing in a multi-agent simulation investigation of the greater berlin area. In: *Journal of Transport Economics and Policy (JTEP)* 49 (4), pp. 560–578.

Kaddoura, Ihab, Amit Agarwal, and Benjamin Kickhöfer (2017) Simulation-based optimization of congestion, noise and air pollution costs: the impact of transport users' choice dimensions. In: *ITEA Annual Conference and School on Transportation Economics*, pp. 17–19.

Kaddoura, Ihab, Benjamin Kickhöfer, Andreas Neumann, and Alejandro Tirachini (2015) Optimal public transport pricing: Towards an agent-based marginal social cost approach. In: *Journal of Transport Economics and Policy (JTEP)* 49 (2), pp. 200–218.

Kaddoura, Ihab, Lars Kröger, and Kai Nagel (2017) An activity-based and dynamic approach to calculate road traffic noise damages. In: *Transportation Research Part D: Transport and Environment* 54, pp. 335–347.

Karich, Peter and Stefan Schröder (2014) Graphhopper. URL: http://www.graphhopper.com (visited on 10/10/2019).

Karlström, Anders and Joel P Franklin (2009) Behavioral adjustments and equity effects of congestion pricing: Analysis of morning commutes during the Stockholm Trial. In: *Transportation Research Part A: Policy and Practice* 43 (3), pp. 283–296.

Keller, Mario, Stefan Hausberger, Claus Matzer, Philipp Wüthrich, and Benedikt Notter (2017) HBEFA Version 3.3. In: *Background documentation, Bern* 12.

Kickhöfer, Benjamin, Friederike Hülsmann, Regine Gerike, and Kai Nagel (2013) Rising car user costs: comparing aggregated and geo-spatial impacts on travel demand and air pollutant emissions. In: *Smart Transport Networks*. Edward Elgar Publishing.

Kickhöfer, Benjamin and Julia Kern (2015) Pricing local emission exposure of road traffic: An agent-based approach. In: *Transportation Research Part D: Transport and Environment* 37, pp. 14–28.

Knight, Frank H (1924) Some fallacies in the interpretation of social cost. In: *The Quarterly Journal of Economics* 38 (4), pp. 582–606.

Kohla, Birgit and Michael Meschik (2013) Comparing trip diaries with GPS tracking: Results of a comprehensive Austrian study. In: *Transport survey methods: Best practice for decision making*. Emerald Group Publishing Limited.

Kraemer, Moritz UG, Chia-Hung Yang, Bernardo Gutierrez, Chieh-Hsi Wu, Brennan Klein, David M Pigott, Louis Du Plessis, Nuno R Faria, Ruoran Li, William P Hanage, et al. (2020) The effect of human mobility and control measures on the COVID-19 epidemic in China. In: *Science* 368 (6490), pp. 493–497.

Leape, Jonathan (2006) The London congestion charge. In: *Journal of Economic Perspectives* 20 (4), pp. 157–176.

Levinson, David (2010) Equity effects of road pricing: A review. In: *Transport Reviews* 30 (1), pp. 33–57.

Lindsney, Robin and Erik Verhoef (2001) Traffic congestion and congestion pricing. Emerald Group Publishing Limited.

Link, Heike (2008) Acceptability of the German charging scheme for heavy goods vehicles: Empirical evidence from a freight company survey. In: *Transport Reviews* 28 (2), pp. 141–158.

Maibach, Markus, Christoph Schreyer, Daniel Sutter, HP Van Essen, BH Boon, Richard Smokers, Arno Schroten, C Doll, Barbara Pawlowska, and Monika Bak (2008) Handbook on estimation of external costs in the transport sector. In: *Ce Delft* 336.

Montini, Lara, Sebastian Prost, Johann Schrammel, Nadine Rieser-Schüssler, and Kay W Axhausen (2015) Comparison of travel diaries generated from smartphone data and dedicated GPS devices. In: *Transportation Research Procedia* 11, pp. 227–241.

Muley, Deepti, Md Shahin, Charitha Dias, and Muhammad Abdullah (2020) Role of Transport during Outbreak of Infectious Diseases: Evidence from the Past. In: *Sustainability* 12 (18), p. 7367.

Newman, Kira LS (2012) Shutt up: bubonic plague and quarantine in early modern England. In: *Journal of social history* 45 (3), pp. 809–834.

Newson, Paul and John Krumm (2009) Hidden Markov map matching through noise and sparseness. In: *Proceedings of the 17th ACM SIGSPATIAL international conference on advances in geographic information systems*, pp. 336–343.

Nielsen, Otto Anker (2004) Behavioral responses to road pricing schemes: Description of the Danish AKTA experiment. In: *Intelligent Transportation Systems*. Vol. 8. 4. Taylor & Francis, pp. 233–251.

Oliver, Nuria, Bruno Lepri, Harald Sterly, Renaud Lambiotte, Sébastien Deletaille, Marco De Nadai, Emmanuel Letouzé, Albert Ali Salah, Richard Benjamins, Ciro Cattuto, et al. (2020) Mobile phone data for informing public health actions across the COVID-19 pandemic life cycle.

Otte, Gunnar (2008) Sozialstrukturanalysen mit Lebensstilen: eine Studie zur theoretischen und methodischen Neuorientierung der Lebensstilforschung. Springer-Verlag.

Pigou, Arthur Cecil (1920) The economics of welfare.

Sælensminde, Kjartan (2004) Cost–benefit analyses of walking and cycling track networks taking into account insecurity, health effects and external costs of motorized traffic. In: *Transportation Research Part A: Policy and Practice* 38 (8), pp. 593–606.

Santos, Georgina and Blake Shaffer (2004) Preliminary results of the London congestion charging scheme. In: *Public Works Management & Policy* 9 (2), pp. 164–181.

Shen, Li and Peter R Stopher (2014) Review of GPS travel survey and GPS data-processing methods. In: *Transport Reviews* 34 (3), pp. 316–334.

Small, K. A. (2008) Urban transportation policy: A guide and road map. In: *Unraveling the Urban Enigma: City Prospects, City Policies, Conference and Book*. Wharton School, University of Pennsylvania.

Small, K. A., Erik Verhoef, and Robin Lindsey (2007) The economics of urban transportation. Routledge.

Stopher, Peter, Camden FitzGerald, and Min Xu (2007) Assessing the accuracy of the Sydney Household Travel Survey with GPS. In: *Transportation* 34 (6), pp. 723–741.

Stopher, Peter R and Stephen P Greaves (2007) Household travel surveys: Where are we going? In: *Transportation Research Part A: Policy and Practice* 41 (5), pp. 367–381.

Subhlok, Jaspal, James M Stichnoth, David R O'hallaron, and Thomas Gross (1993) Exploiting task and data parallelism on a multicomputer. In: *Proceedings of the fourth ACM SIGPLAN symposium on Principles and practice of parallel programming*, pp. 13–22.

Tan, Christopher (2020) New ERP system to start in 2023 but no distance-based charging yet; replacement of IU from second half of 2021.

Tirachini, Alejandro and David A Hensher (2012) Multimodal transport pricing: first best, second best and extensions to non-motorized transport. In: *Transport Reviews* 32 (2), pp. 181–202.

Tirachini, Alejandro, David A Hensher, and John M Rose (2013) Crowding in public transport systems: effects on users, operation and implications for the estimation of demand. In: *Transportation research part A: policy and practice* 53, pp. 36–52.

Tirachini, Alejandro, David A Hensher, and John M Rose (2014) Multimodal pricing and optimal design of urban public transport: The interplay between traffic congestion and bus crowding. In: *Transportation Research Part B: Methodological* 61, pp. 33–54.

Transport for London (2004) Congestion Charging Central London. Impacts monitoring. Second Annual Report. Tech. rep.

Transurban (2016) Changed conditions ahead. The transport revolution and what it means for Australians, Melbourne Road Usage Study Report. Report. Melbourne.

Trilla, Antoni, Guillem Trilla, and Carolyn Daer (2008) The 1918 "Spanish flu" in Spain. In: *Clinical infectious diseases* 47 (5), pp. 668–673.

Verhoef, Erik (2000) The implementation of marginal external cost pricing in road transport. In: *Papers in Regional Science* 79 (3), pp. 307–332.

Verhoef, Erik, Peter Nijkamp, and Piet Rietveld (1997) Tradeable permits: their potential in the regulation of road transport externalities. In: *Environment and Planning B: Planning and Design* 24 (4), pp. 527–548.

Vickrey, William S (1963) Pricing in urban and suburban transport. In: *The American Economic Review* 53 (2), pp. 452–465.

Vinceti, Marco, Tommaso Filippini, Kenneth J Rothman, Fabrizio Ferrari, Alessia Goffi, Giuseppe Maffeis, and Nicola Orsini (2020) Lockdown timing and efficacy in controlling COVID-19 using mobile phone tracking. In: *EClinicalMedicine* 25.

Vrtic, Milenko, Nadine Schuessler, Alexander Erath, and Kay W Axhausen (2010) The impacts of road pricing on route and mode choice behaviour. In: *Journal of Choice Modelling* 3 (1), pp. 109–126.

Wesolowski, Amy, Taimur Qureshi, Maciej F Boni, Pål Roe Sundsøy, Michael A Johansson, Syed Basit Rasheed, Kenth Engø-Monsen, and Caroline O Buckee (2015) Impact of human mobility on the emergence of dengue epidemics in Pakistan. In: *Proceedings of the National Academy of Sciences* 112 (38), pp. 11887–11892.

Widhalm, Peter, Yingxiang Yang, Michael Ulm, Shounak Athavale, and Marta C González (2015) Discovering urban activity patterns in cell phone data. In: *Transportation* 42 (4), pp. 597–623.

Winslott-Hiselius, Lena, Karin Brundell-Freij, Åsa Vagland, and Camilla Byström (2009) The development of public attitudes towards the Stockholm congestion trial. In: *Transportation Research Part A: Policy and Practice* 43 (3), pp. 269–282.

Wolf, Jean and Randall Guensler (2000) Using GPS data loggers to replace travel diaries in the collection of travel data. PhD thesis. Georgia Institute of Technology, School of Civil and Environmental Engineering.

Wolf, Jean, Michael Loechl, Miriam Thompson, and Carlos Arce (2003) Trip rate analysis in GPS-enhanced personal travel surveys. In: *Transport survey quality and innovation* 28, pp. 483–98.

Xiong, Chenfeng, Songhua Hu, Mofeng Yang, Weiyu Luo, and Lei Zhang (2020) Mobile device data reveal the dynamics in a positive relationship between human mobility and COVID-19 infections. In: *Proceedings of the National Academy of Sciences* 117 (44), pp. 27087–27089.

Yabe, Takahiro, Kota Tsubouchi, Naoya Fujiwara, Takayuki Wada, Yoshihide Sekimoto, and Satish V Ukkusuri (2020) Non-compulsory measures sufficiently reduced human mobility in Tokyo during the COVID-19 epidemic. In: *Scientific Reports* 10 (1), pp. 1–9.

Zandbergen, Paul A and Sean J Barbeau (2011) Positional accuracy of assisted GPS data from high-sensitivity GPS-enabled mobile phones. In: *The Journal of Navigation* 64 (3), pp. 381–399.

Zheng, Yu (2015) Trajectory data mining: an overview. In: *ACM Transactions on Intelligent Systems and Technology (TIST)* 6 (3), pp. 1–41.

Zhou, Ying, Renzhe Xu, Dongsheng Hu, Yang Yue, Qingquan Li, and Jizhe Xia (2020) Effects of human mobility restrictions on the spread of COVID-19 in Shenzhen, China: a modelling study using mobile phone data. In: *The Lancet Digital Health* 2 (8).

Zurich Public Transport (VBZ) (2017) Servicequalität 2017- Bericht zur Servicequalität und Nachfrage auf dem Netz der Verkehrsbetriebe Zürich (german). Yearly Report.

2

MOBIS: RESPONSE RATES AND SURVEY METHOD RESULTS

FULL TITLE

A National-Scale Mobility Pricing Experiment using GPS Tracking and Online Surveys in Switzerland: Response Rates and Survey Method Results

AUTHORSHIP

Joseph Molloy, Alberto Castro, Thomas Götschi, Beaumont Schoeman, Christopher Tchervenkov, Uros Tomic, Beat Hintermann and Kay W. Axhausen

This chapter was accepted for presentation at the *100th Annual Meeting of the Transportation Research Board, January 2021*, and is under review for publication in *Travel Behaviour and Society*

2.1 abstract

This article presents the first results and observations from the MOBIS Study, a nation-wide mobility pricing field experiment in Switzerland. Mobility pricing is widely regarded as a promising policy measure to combat congestion, internalize external costs of transport, and offset decreasing fuel tax revenues. However, the implementation of mobility pricing in Switzerland is hindered by a lack of empirical evidence, among other things. In the field experiment participants participated through the use of a GPS tracking app, Catch-my-Day, which logged their daily travel on different transport modes and imputed the trip segments and modes. The experiment lasted 8 weeks, bookended by online surveys. After the first 4 week control phase, participants were split into three treatment groups. The first continued as a control. The second received information on their external costs, and the third received a real monetary budget, from which their external costs were deducted. The first results show that the technology is capable of supporting such an experiment on both Android and iOS, the two main mobile platforms. Significant differences in the engagement and attrition were observed between iOS and Android participants over the 8 week period. Finally, the attrition rate did not vary between treatment groups.

2.2 introduction

Mobility pricing is widely regarded as a promising policy measure to combat congestion, internalize external costs of transport, and offset decreasing fuel tax revenues. The concept of mobility pricing was first proposed in the 1920's as an example of a corrective tax to internalize congestion externalities (Pigou, 1920). Since then, there has been much study of the topic, including mathematical theory (Small et al., 2004; Verhoef, Nijkamp, and Rietveld, 1996) and simulation experiments (Chakirov, 2016; Kaddoura, 2015; Meyer de Freitas et al., 2017). Most of the research and practical implementations have focused specifically on road pricing, which is a limited form of mobility pricing that focuses on drivers. Despite the theoretical capabilities to maximise infrastructure utilisation, mobility pricing has only been sparsely implemented in practice as it is typically viewed as a 'new tax' and is thus associated with strong political resistance. Schemes in London (Leape, 2006; Santos and Shaffer, 2004) and Stockholm (Eliasson et al., 2009) are two well-known examples where limited mobility pricing has been implemented in the form of congestion charges: Cars entering the central business district during certain hours have to pay a fee. These 'congestion charges' don't reflect all the external costs from all modes of transportation. Schemes have also been implemented in a number of cities including Milan, Paris, Rome, Stuttgart and Singapore.

Although there is evidence on the success of congestion pricing (Eliasson et al., 2009; Leape, 2006; Santos and Shaffer, 2004), understanding the effects of broader mobility pricing schemes remains a challenge. A key challenge is understanding the potential impacts of the proposed policy. Multiple studies have looked at route, mode and destination choice within the context of various pricing schemes using stated-preference experiments (Li and Hensher, 2012; Vrtic, Schuessler, et al., 2010; Washbrook, Haider, and Jaccard, 2006). Work on the acceptance of pricing schemes includes Jakobsson, Fujii, and Gärling (2000) and Vrtic, Schüssler, et al. (2007). More recently, the proliferation of affordable GPS tracking and mobile connectivity has opened up the possibilities to do field experiments exploring transport users' behavioural responses under a pricing scheme, which would have been financially and logically infeasible in the pre cell-phone era. In one of the first examples, Nielsen (2004) equipped 500 cars with a GPS-based device, and monitored participants for a control period before exposing them to a pricing scheme for the Copenhagen region. This study was in the pre-smartphone era and hence limited to a small sample size and no control group. A similar study using car-based GPS loggers was performed in Melbourne, in which 1,400 toll road users experienced different types of congestion charges (Martin and Thornton, 2017; Transurban, 2016). A period of several months was used to monitor baseline behaviour before the pricing schemes were introduced for three quarters of the sample. In both these experiments, only car trips with the primary household vehicle were tracked. Public transport and active modes were not recorded. The Melbourne study did investigate possible modal shifts to rail commuting, by identifying car trips and subsequent parking at railway stations. The study reportd that 30% of participants reported changing their road travel use under the pricing scheme. Until now there have been no studies that have attempted to use smartphone-based GPS tracking to look at road or mobility pricing, limiting the opportunity to understand modal shifts.

The use of GPS tracking for mobility research is now widespread. Multiple studies have identified how traditional travel diaries under-report the number of trips, due to, among other reasons, response burden and memory recall (Janzen et al., 2018; Stopher, FitzGerald, and Xu, 2007; Wolf et al., 2003). Passive tracking mostly mitigates these issues, although the collecting of trip metadata such as detailed trip purpose, fellow passengers and travel expenses mostly still requires more traditional survey methods. Furthermore, the performance of GPS tracking depends on the quality of the GPS traces, and the algorithms used to identify trips, stages and activities, as well as the mode and purpose of travel. Here there has been significant advances in recent years (Marra et al., 2019; Schuessler and Axhausen, 2009). For two comprehensive reviews on the processing of GPS tracking data, the reader is referred to Shen and Stopher (2014) and Nikolic and Bierlaire (2017). Other studies note that the performance of the algorithms is highly dependent on the quality of the GPS data (Burkhard et al., 2020; Harding, 2019; Montini et al., 2015).

One of the key factors influencing the quality of GPS data is the device used. This can be either a dedicated GPS logger, or a smartphone, where the data is collected through an app. The quality of the data can vary between devices, in particular between iOS and Android devices, depending for example on battery saving settings.

Few studies have explored the implications of this iOS/Android dichotomy and the implications for mobility studies using app-based tracking. Harding (2019) compared the performance of trip identification and mode detection by different apps and found that iOS-based apps tended to have a higher accuracy. However, not only is the quality of the recorded data important, but also the attrition rate throughout the study, as this ultimately determines the sample size. This is an open question that has not been widely explored. The market penetration rates of iOS and Android - and even different Android-based manufacturers - varies across regions and, possibly, segments of the population. For studies requiring a representative sample, for example official national travel surveys, an understanding of these factors is important. As the MOBIS study aims to analyze societal impacts of mobility pricing to inform policy and decision making, obtaining a representative sample was a key objective.

We report our experiences undertaking a tri-lingual, national-scale mobility pricing survey and randomized controlled trial in Switzerland, combining traditional survey methods and app-based GPS tracking. MOBIS aims to understand the effects on travel behaviour of a) informing subjects about congestion, health effects, and carbon emissions of their mobiliity, and b) actually charging subjects the external costs associated with these 3 factors under a mobility pricing experiment. To do this, we examine two different treatments - information and pricing. In the current political discourse it is of interest to understand if information measures are found to have a similar impact as mobility pricing. On the other hand, evidence for pricing would support calls to restructure current mobility taxes and subsidies. In this paper, we focus on the survey method and the role of app-based tracking. In particular, contributions include a detailed analysis of the response rate over the duration of the study, and how it was impacted by the differences between iOS and Android devices.

2.3 Methodology

The 8-week study consisted of two consecutive 4-week phases, a control and treatment phase respectively, book-ended by introductory and concluding online surveys. A pretest with a mail-out sample of 1,500 letters was undertaken to estimate the expected response rate for the main study and test the surveys and GPS tracking.

2.3.1 *Initial recruitment*

For the main study, a representative list of 60,000 addresses randomly selected across the major agglomerations (in the German and French speaking parts) of Switzerland from the Swiss Federal Office of Statistics was used. Based on the response rate in the pretest, this address sample was skewed to account for under-represented groups. Additionally, to achieve the desired sample size of 3,500 study participants, a second wave of around 30,000 persons were contacted using addresses from a private vendor, yielding a total of a little over 90,000 invitations. Only people living in an agglomeration area of Switzerland (excluding the Italian-speaking canton Ticino) were invited to participate in the study.

The letter invited the recipients to fill in a screening-survey with transport-related questions and, if they met the inclusion criteria, to participate in a smartphone-based mobility experiment where they would receive 100 CHF ($100 USD) for participating for the entire 8 weeks. Neither the 'mobility pricing' nature of the study nor the focus on the external costs of transport was shared with the participants.

Two reminder letters were also sent in the first wave, 4 and 7 weeks after the invitation letter was received, to those who had not responded to previous letters. No reminders were sent in the second wave as the target number of 3,500 participants had already been achieved.

2.3.2 *Introduction and final surveys*

The initial survey was designed to determine a respondent's eligibility for the main tracking study and collect data that would be needed in the calculation of external costs (such as mobility tool ownership, car type and age, and some general attitudes towards transport policies). The final survey included a series of stated-choice experiments and lifestyle and values questions, as well as awareness questions to gauge if participants understood the experiment and were therefore 'knowledgeable' participants. Completion of the final survey was a condition for receiving the incentive.

2.3.3 *Recruitment for the field experiment*

The participants who completed the introduction survey were assessed against the eligibility criteria for the field experiment. Specifically, participants

- had to use a car at least two days a week
- were restricted to the age of 18 to 65
- must be able to walk without assistance
- must own a smartphone
- were not allowed to drive in a professional capacity - i.e. postman/-woman or taxi driver.

Those who met the requirements for the study and gave consent to participate were sent an email with a unique registration code and a link to download the Catch-My-Day app and participate in the tracking study.

2.3.4 *Tracking app*

The Catch-My-Day app is a location tracker for iOS and Android, which uses the location services of the respective operating system. GPS tracks are stored on the phone and uploaded to the MotionTag analytics platform, where stages, travel modes and activities are imputed. The following modes are detected the by Catch-my-Day app.

- Airplane

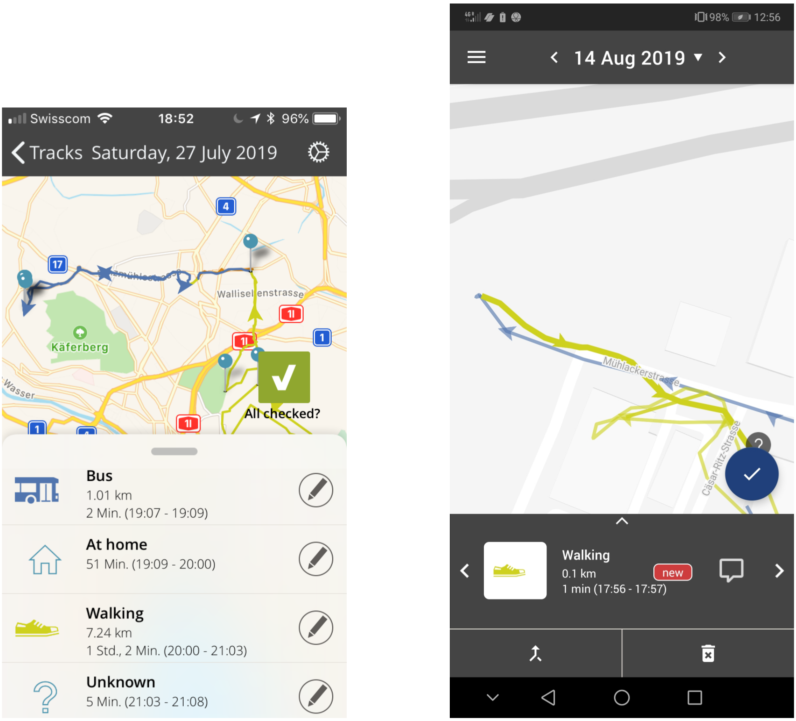
- Bicycle
- Bus
- Car
- Ferry
- S-Bahn (Local train)
- Regional train
- Subway
- Train (other)
- Tram
- Walk

Those marked with an asterisk are not automatically detected, but selectable by the user as a correction

- Boat*
- Carsharing*
- Motorbike/Scooter*
- Taxi/Uber*

Users can view their daily travel patterns on their phone in the form of a logbook, validate the travel mode and activity purpose or indicate if a trip or activity did not take place. The database stores both their correction and the original algorithmic imputation. There are some user-interface differences between the iOS and Android versions, which are most noticeable in the trip validation interface.

Users of could view their daily travel log in the app, and correct any incorrect travel mode imputations. Validation in the treatment phase was still allowed, even for the pricing group. Disabling validation in the treatment phase would have disadvantaged those affected by mis-detection, especially if they had made corrections in the control phase, due to the lower external costs of public transport. To counter any possible 'gaming' of the experiment, an outlier analysis was performed before transferring the incentive to the participants. No clearly suspicious behaviour was observed,



(a) iPhone

(b) Android

Figure 2.1: Trip/validation interface

except for one participant who seemed to switch to riding his e-bike for the entire second phase of the study. Figure 2.1 presents the validation interface of the app for the respective operating systems.

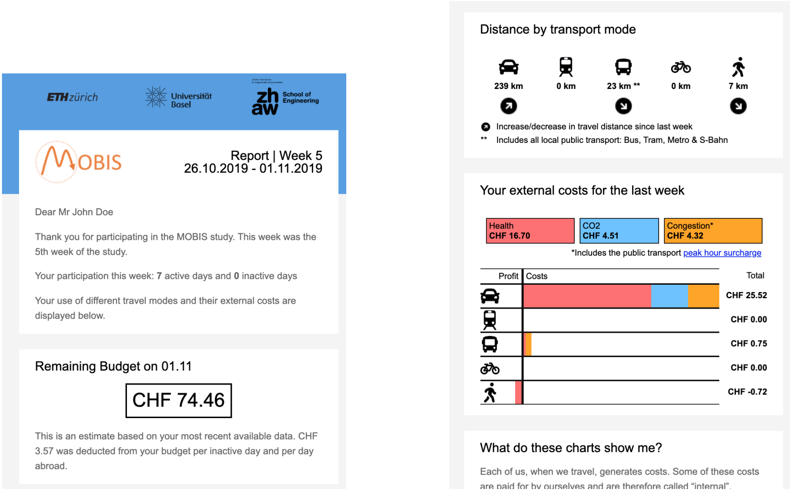
Users were required to activate the app by creating an account, which requires the provision of an email address and the choice of a password, along with the unique registration code provided. Participants are not required to validate their trips and activities, but were informed that this was possible and would be appreciated.

To increase the retention rate, automated reminder emails were sent to participants when they had not activated the app, or no data was recorded for a certain number of days. A help-desk was set up for participants experiencing difficulties. User-guides on correctly configure one's smartphone for the app were provided. Additionally, participants who did not record data on at least 12 of the first 28 days were removed from the study, and notified by email.

2.3.5 *Treatment groups*

The 8-week study period was divided into two 4-week phases. In the first phase, participants were tracked using the app, and received weekly reports on the kilometers traveled per mode. At the beginning of the second phase, participants were randomly assigned to either the control group, or one of the two treatment groups. The control group continued to receive the same basic information on their behaviour, whereas both the information and pricing groups received additional information on the externalities they caused. Furthermore, participants in the pricing group were provided with a mobility budget, equal to 120% of their external costs in the first phase, from which their external costs in phase 2 were subtracted; with the balance remaining transferred to them, as an incentive to reduce their externalities. An example of the weekly reports is provided in Figure 2.3

These externalities were separated into health, environmental and congestion costs, which were computed using a data pipeline run every evening. For more details on the externality computation, please refer to Tchervenkov, Molloy, and Axhausen (2018). The calculations are based on the HBEFA (Handbook for emissions analysis), relevant Swiss norms and the IVT MATSim scenario for Switzerland (Hörl, Balać, and Axhausen, 2019). Additionally, data collected from the introduction survey was incorporated into the data processing pipeline to improve the computation: Informa-



(a) First sections

(b) Further sections

FIGURE 2.3: Weekly report to participants in the pricing group

tion on the participant's main vehicle was used to calculate individualized external costs.

2.4 RESULTS AND DISCUSSION

In this paper, we present the results in terms of participation and the collection of tracking data. The analysis of the field experiment is still ongoing and will be presented elsewhere.

2.4.1 *Response rates*

Invitations to the study were sent by post to 90,090 persons. From this sample, 23.70% completed the initial survey. This response rate was likely elevated by the prospect of the 100 CHF incentive for the tracking experiment, mentioned in the invitation letter (even though no incentive was provided for participation in the introductory survey on its own). Only 31.89% of those who completed the introduction survey met the criteria for the field experiment. This was predominately due to the minimal car-use requirement. Many people (age 16 and over) in Switzerland neither have access to a car (22%), nor a drivers license (18%) (BFS and ARE, 2017).

The two reminder letters were also effective in the first wave. of the 5320 who registered, 2397 (45%) did so before a reminder letter was sent, and 1793 (34%) and 1245 (23%) did so after the first and second reminder respectively.

Of those who qualified, 78.06% agreed to participate. This compares similarly to the other studies in Table 2.1. At the next stage, out of the remaining 5364 participants, 1146 (21.4%) did not start tracking. Either they either never installed the app, removed it before data was recorded, or were unable to get it to work successfully. Of those who did track, the share with an iOS device was 61%, much higher than the reported 44,4% national market share in 2019 Comparis, 2019, indicating that relatively more Android users were unable or unwilling to use app. Anecdotal evidence from the staff on the study help desk also indicated that more participants had issues installing the app for Android than iOS, and required assistance from the help desk in doing so (Tchervenkov, Molloy, Castro Fernández, et al., 2020).

Finally, 3690 participants successfully completed the 8-week tracking period, giving a completion rate of 69.4% for those that registered, and 4.06% overall. This is somewhere in the middle of the results from previous studies, with the high incentive appropriately offsetting the long tracking period.

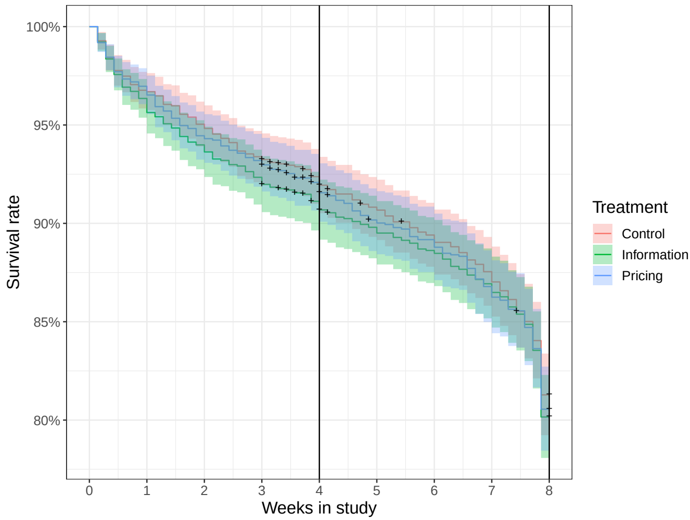


Figure 2.5: Kaplan-Meier survival curve by treatment group. The cross indicates censoring of participants

2.4.2 *Participant retention*

To explore the retention rate of participants in the tracking phase, we performed a survival analysis on the duration of tracking in the study. First, a Kaplan-Meier approach (see Figure 2.5) shows the impact of the treatment on the length of time which participants would track. Participants who were automatically dropped out after phase 1 due to poor tracking compliance but were still tracking at the end of phase 1 were censored (marked by a cross). There is no significant difference between the three treatment groups in their survival curves. A sharp decrease in survival is evident in the last study week. As participants were informed at the end of the study that they could delete the app, the last few days of tracking were sometimes not collected before the app was deleted.

Although the participants in the study had a clear participation goal of 8 weeks, after which they would receive the incentive, the survival curve is extremely linear. One would intuitively expect that the attrition rate would be highest early on in the study, and flatten out as participants neared the 8-week goal. This appears to only slightly be the case, with the dropout rate remaining constant throughout the study, even in the second phase.

Project name	MOBIS	SPOT (1)	IN-THE-MOMENT (2)	ATLAS (3)	AKTA (4)	Cincinnati (5)	Atlanta (6)	Reno (7)	Tel Aviv HTS (8)
Tracker	MotionTag	MEILI	rMove	SITSS	Device	Device	Device	Device/App	App
Country	Switzerland	Sweden	USA	NZ	Denmark	USA	USA	USA	Israel
Year	2019	2015	2015	2014	2001-2003	2010	2011	2015-2016	2016-2017
Tracking days	56	7	7	3	112	3	7	7	2
Min. incentive (USD)	$100		$25		variable	$25	$25	$25	
Validation/annotation	Optional	Yes	Yes	yes	no	yes	yes	yes	yes
Invited persons (N)	90,909	130,000	1,427	186	25,000	11,118	16,374	25,817	67,199
Intro survey (N)	21,571								
% of invited	23.73%								
Qualified (N)	6895	1159	478						38,500
% of invited	7.58%	0.89%	33.50%						57.29%
% of intro completed	31.96%								
Registered (N)	5375	495	295	77	500	4656	1938	602	27,415
% of invited	5.91%	0.38%	20.67%	41.40%			11.84%		40.80%
% of qualified	77.96%	42.71%	61.72%						71.21%
Started tracking (N)	4218	293	295	73	500	4656		602	25,201
% of invited	4.64%	0.23%	20.67%	39.25%	2.00%	41.88%			37.50%
% of qualified	61.17%	25.28%	61.72%						65.46%
Completed tracking (N)	3690	51	240	65	500	3849	1061	312	23,240
% of invited	4.06%	0.04%	16.82%	34.95%	2.00%	34.62%	6.48%	1.21%	34.58%
% of qualified	53.52%	4.40%	50.21%						60.36%
% of registered	68.65%	10.30%	81.36%	84.42%	100.00%	82.67%	54.75%	51.83%	84.77%

1. Allström, Kristoffersson, and Susilo (2017), 2.Greene et al. (2012), 3.Safi et al. (2015), 4.Nielsen (2004),
5.Wargelin, Stopher, et al. (2012), 6.Livingston (2011), 7.Stopher, Daigler, and Griffith (2018), 8.Nahmias-Biran et al. (2018)

Table 2.1: Response rates in various tracking studies

Furthermore, Figure 2.5) shows that the treatment didn't affect the attrition rate in the second phase.

A time-variant Cox proportional hazards model is to investigate the impact of different factors on the participation duration (see Table 2.2 for the model results). To account for time-dependent effects, the study period was stratified into fortnightly windows. Those in high-income brackets (more than 12,000 CHF/month) were more likely to stop tracking. Conversely, those from larger households and those with tertiary education were more likely to track for longer. A significant gender-based difference was only observed in the final fortnight, where females were more likely to remain in the study.

Contrary to expectations, there was no significant effect of age on the hazard rate. This suggests that common concern about the feasibility of tracking studies for older age groups is unfounded, at least up to the age of 65, the age limit in this study.

The coefficient on employment is also time-dependent. Those in the workforce (i.e. excluding students, homeworkers and retirees) were more likely to remain in the study throughout the first fortnight.

The participant's mobile device played a much larger role. Having an Android phone of any model increased the hazard drastically. However, this effect was strongest in the first week. The effects were even larger for Huawei models. The incompatibility of GPS loggers with Android (and particularly Huawei devices) is already well known; however, here the effect is quantified, and seen to be dramatic. The effect was also time-dependent, with the most significant hazard in the first fortnight. At the end of the second fortnight, participants who tracked insufficiently were removed from the study - this explains the reduction in the Android hazard coefficient for the third fortnight, when many of them could have been expected to stop tracking, had they not been removed from the study.

2.4.3 *Post-study retention*

At the end of the tracking study, participants were told that they could delete the app, but were also encouraged to continue using it if they wished. Figure 2.6 shows the dropout rate for the whole study, including the post-study period. The majority of the participants dropped out soon after the study, but even 6 months after the study was completed, around 5% of participants continued to use the app. Anecdotal reports from participants indicated that they enjoyed having an overview of their travel, and that

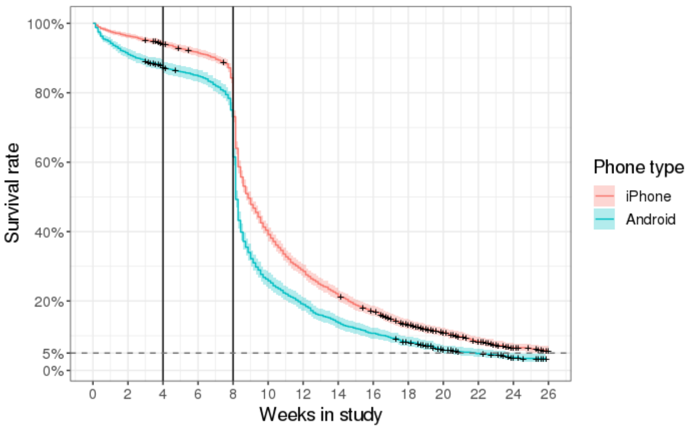


FIGURE 2.6: Post-study participation survival curve

it even continued to inform their mobility decisions. The impacts of the mobile operating system continued even after the study, with the post-study retention rate falling faster for Android users.

2.4.4 *Participant engagement*

Participants in the information and pricing groups were effectively treated through information provided in a weekly email detailing their externalities and the costs incurred. Interactions with the emails were recorded using standard email tacking techniques. Emails that remained unopened were effectively missed treatments. Table 2.3 presents a overview of the engagement with the email communications. The open rate did not change drastically over the duration of the study. Participants in the pricing group viewed their emails much more often than the control or information groups. The information group also opened their emails repeatedly in the first two weeks of phase two, before returning to a pattern similar to the control group, whereas the pricing group continued to repeatedly open their emails.

Participants in the treatment groups likely repeatedly reopened the emails to check their externalities and remaining budget. We suggest that this 'repeat opening' behaviour is a useful indicator to measure the level of engagement with the treatment.

2.4.5 *Trip mode and purpose validation*

Participants were invited to use the validation interface to confirm the detected mode and purpose of their trips and activities. This was optional, but they were encouraged in the weekly email reports to do so. Even in the second phase, participants were trusted to correct the mode detected by the app. As the mode is crucial in determining the external costs deducted from the mobility budget for the pricing group, this consequently gave them the opportunity to 'game' the experiment, by for example 'correcting' car trips to another transport mode. To test for this, a regression analysis using a zero-inflated negative binomial model was performed with the number of corrections for a day as the dependent variable (see Table 2.4). A zero-inflated model was used to accommodate the large number of participants who did not correct any trips. While a significant increase in the number of corrections was observed in phase 2, no increase in the number of corrected trips specific to the pricing group was observed. Conversely, the parameters are insignificant but negative. In fact, the information group saw a significant reduction in the corrections in phase 2. One hypothesis is that by receiving more information on their externalities in the weekly reports in the second phase, participants felt discouraged from correcting their trips in the app. Also, no indication was given to participants that they would be penalised for any suspicious behaviour. The fact that no significant change in the average correction rate was seen between treatment groups, suggests that the trust in the participants was justified.

In recent years, state-of-the-art machine learning algorithms for mode and activity detection have achieved accuracy rates of over 90%, depending on the approach (Nikolic and Bierlaire, 2017; Wu, Yang, and Jing, 2016). Hence, we made validation of the trip purpose and mode optional for participants, in order to ensure a minimal response burden over the 8 weeks. 85.7% of participants confirmed at least 1 of their trips; however, of those who did use the validation functionality, 20.4% of iPhone users and 44.1% of Android users did not make a single correction over the 8 weeks, respectively. Even with state-of-the-art accuracy rates, such a validation behaviour is extremely unlikely. As such, we can assume that these participants did not use or understand the validation interface correctly, and these participants are therefore removed from the following analysis on the mode detection performance. It also indicates that the iPhone validation interface was much more intuitive.

2.4.6 *Mode detection performance*

The mode detection provided by the tracking app was a key component of the MOBIS study. As far as the authors aware, this is the first study to incentivise changes in mobility behaviour based on the output of a mode detection algorithm. As seen in Table 2.5, the algorithm worked exceptionally well on location data from both operating systems. There is small difference in accuracy between iOS and Android, with iOS being on average slightly better (92.23% vs 92.10%) with a p-value of 0.01, test of equal proportions). However, the differences in accuracy are more observable at the categorical level. The iOS performs better on car, local rail, regional rail, tram and walk. However, the differences are only 1-3% in accuracy. Note that 'Rail' groups all rail modes together for conciseness. It is also worth noting that while the accuracy of some individual rail modes is quite low, the overall rail accuracy is very good. The main confusion was between different rail mode types.

Table 2.6 presents the confusion matrix between the modes. Here we can see that the algorithm often mis-detected car travel as bus travel. For conciseness, the category 'Other *' includes those modes which could be manually selected by the participant, but which were not automatically detected. These included: Carsharing, Taxi/Uber, Motorbike/Mopeds, and Gondolas. Most of these were detected as car travel, and the 1,500 'Bicycle' trips which were corrected to 'Other' were predominately trips by motorbike or moped.

These mode detection results confirmed the indications of our pretest that the automatic detection could indeed be used to calculate the external costs of travel with sufficient accuracy and determine the phase 2 budget and deductions based on these. If the accuracy had been too low, more participants would have dropped out of the study, seeing it as 'unfair' if the budget and deductions did not match their travel behaviour.

2.4.7 *Identified mode detection issues*

As previously mentioned, the quality of the mode detection was key to the mobility pricing field experiment. A few issues were identified which are worth considering in future studies that apply algorithmic mode detection.

The first consideration concerns those leisure activities that are movement based over a larger area, such as a bike tour, hiking and skiing. Skiing is especially important in alpine areas: In Switzerland, the percentage of the population that ski regularly is 37% (Statistica, 2018). Gondolas and chairlifts move at between 15 and 50km/h, meaning that these trips are often confused for car travel unless the algorithm has been specifically calibrated. On the downhill, skiers reach similar speeds. Taking a strict definition of a transport trip, such movement-based activities should be excluded from the calculation of external costs. If they were to be included, a person could end up being charged for a long hike in the wilderness on the weekend - which would arguably not be in the spirit of a mobility pricing scheme.

The second consideration is trip chaining. Shen and Stopher (2014) note that all methods to date (albeit in 2014) did not consider trip chains when detecting the transport mode, and only considered each individual stage. While the mode detection provided by the app was sufficient for the purpose of the mobility pricing field experiment, anecdotal evidence indicates that considering trip chains could further improve the performance of the algorithm.

2.5 CONCLUSION

This work makes multiple contributions to the literature on conducting tracking-based mobility studies, and demonstrates the feasibility of running an incentive-based field experiment using a tracking app. We analysed the effect of the mobile device operating system on GPS tracking studies, and identified certain areas where the difference in OS needs to be considered when undertaking such studies. The impact on participant retention is significant. While this effect is strongest at the start of the study, it persists throughout. The on-boarding of Android users into the study took significant resources, and we suggest this be accounted for when planning and budgeting such studies. Correspondence by email was effective, and participant engagement did not decline over the 8 weeks. The mode detection algorithm was also sufficiently accurate to support the calculation of external costs in the field experiment. Finally, concerns that participants would manipulate the study by 'correcting' their trips in the app were unfounded, with participants adhering to the spirit of the study. Socio-demographic differences in the correction rate do, however, indicate that some participants were more engaged than others.

REFERENCES

Allström, Andreas, Ida Kristoffersson, and Yusak Susilo (2017) Smartphone based travel diary collection: Experiences from a field trial in Stockholm. In: *Transportation research procedia* 26, pp. 32–38.

BFS and ARE (2017) Verkehrsverhalten der Bevölkerung – Ergebnisse des Mikrozensus Mobilität und Verkehr 2015.

Burkhard, Oliver, Henrik Becker, Robert Weibel, and Kay W Axhausen (2020) On the requirements on spatial accuracy and sampling rate for transport mode detection in view of a shift to passive signalling data. In: *Transportation Research Part C: Emerging Technologies* 114, pp. 99–117.

Chakirov, Artem (2016) Urban mobility pricing with heterogeneous users. PhD thesis. ETH Zurich.

Comparis (2019) Beliebtheit von iPhones in der Schweiz nimmt ab (German). online, [`https://www.comparis.ch/telecom/handy-gadgets/analyse/smartphone-markt-schweiz`], accessed 13/07/2020.

Eliasson, Jonas, Lars Hultkrantz, Lena Nerhagen, and Lena Smidfelt Rosqvist (2009) The Stockholm congestion–charging trial 2006: Overview of effects. In: *Transportation Research Part A: Policy and Practice* 43 (3), pp. 240–250.

Greene, Elizabeth, Leah Flake, Kevin Hathaway, and Michae Geilich (2012) A Seven-Day Smartphone-Based GPS Household Travel Survey in Indiana. In: *TRB 95th Annual Meeting Compendium of Papers*. Transportation Research Board.

Harding, Chris (2019) From Smartphone Apps to In-Person Data Collection: Modern and Cost-Effective Multimodal Travel Data Collection for Evidence-Based Planning. PhD thesis.

Hörl, Sebastian, Milos Balać, and Kay W Axhausen (2019) Pairing discrete mode choice models and agent-based transport simulation with MATSim. In: *2019 TRB Annual Meeting Online*. Transportation Research Board, pp. 19–02409.

Jakobsson, Cecilia, Satoshi Fujii, and Tommy Gärling (2000) Determinants of private car users' acceptance of road pricing. In: *Transport Policy* 7 (2), pp. 153–158.

Janzen, Maxim, Maarten Vanhoof, Zbigniew Smoreda, and Kay W Axhausen (2018) Closer to the total? Long-distance travel of French mobile phone users. In: *Travel Behaviour and Society* 11, pp. 31–42.

Kaddoura, Ihab (2015) Marginal congestion cost pricing in a multi-agent simulation investigation of the greater Berlin area. In: *Journal of Transport Economics and Policy (JTEP)* 49 (4), pp. 560–578.

Leape, Jonathan (2006) The London congestion charge. In: *Journal of Economic Perspectives* 20 (4), pp. 157–176.

Li, Zheng and David A Hensher (2012) Congestion charging and car use: A review of stated preference and opinion studies and market monitoring evidence. In: *Transport Policy* 20, pp. 47–61.

Livingston, J (2011) Atlanta Regional Commission-Regional Travel Survey-Final Report. In: *PTV NuStata, Austin, TX*.

Marra, Alessio D, Henrik Becker, Kay W Axhausen, and Francesco Corman (2019) Developing a passive GPS tracking system to study long-term travel behavior. In: *Transportation research part C: emerging technologies* 104, pp. 348–368.

Martin, L. A. and Sam Thornton (2017) To Drive or Not to Drive? A Field Experiment in Road Pricing. Working Paper. Parkville, Australia.

Meyer de Freitas, Lucas, Oliver Schuemperlin, Milos Balac, and Francesco Ciari (2017) Equity effects of congestion charges: An exploratory analysis with MATSim. In: *Transportation Research Record* 2670 (1), pp. 75–82.

Montini, Lara, Sebastian Prost, Johann Schrammel, Nadine Rieser-Schüssler, and Kay W Axhausen (2015) Comparison of travel diaries generated from smartphone data and dedicated GPS devices. In: *Transportation Research Procedia* 11, pp. 227–241.

Nahmias-Biran, Bat-hen, Yafei Han, Shlomo Bekhor, Fang Zhao, Christopher Zegras, and Moshe Ben-Akiva (2018) Enriching Activity-Based Models using Smartphone-Based Travel Surveys. In: *Transportation Research Record* 2672 (42), pp. 280–291.

Nielsen, Otto Anker (2004) Behavioral responses to road pricing schemes: Description of the Danish AKTA experiment. In: *Intelligent Transportation Systems*. Vol. 8. 4. Taylor & Francis, pp. 233–251.

Nikolic, Marija and Michel Bierlaire (2017) Review of transportation mode detection approaches based on smartphone data. In: *17th Swiss Transport Research Conference*. CONF.

Pigou, Cecil Arthur (1920) The Economics of Welfare. In: *New Brunswick & Londres: Transaction Publishers* 4.

Safi, Hamid, Behrang Assemi, Mahmoud Mesbah, Luis Ferreira, and Mark Hickman (2015) Design and Implementation of a Smartphone-Based Travel Survey. In: *Transportation Research Record* 2526 (1), pp. 99–107.

Santos, Georgina and Blake Shaffer (2004) Preliminary results of the London congestion charging scheme. In: *Public Works Management & Policy* 9 (2), pp. 164–181.

Schuessler, Nadine and Kay W Axhausen (2009) Processing raw data from global positioning systems without additional information. In: *Transportation Research Record* 2105 (1), pp. 28–36.

Shen, Li and Peter R Stopher (2014) Review of GPS travel survey and GPS data-processing methods. In: *Transport Reviews* 34 (3), pp. 316–334.

Small, Kenneth A et al. (2004) Road pricing and public transport. In: *Road pricing: Theory and evidence* 9, pp. 133–58.

Statistica (2018) Skiing in Europe - Statistics & Facts. online, [https://www.statista.com/topics/3922/skiing-in-europe/], accessed 13/07/2020.

Stopher, Peter, Camden FitzGerald, and Min Xu (2007) Assessing the accuracy of the Sydney Household Travel Survey with GPS. In: *Transportation* 34 (6), pp. 723–741.

Stopher, Peter R., Vivian Daigler, and Sarah Griffith (2018) Smartphone app versus GPS Logger: A comparative study. In: *Transportation Research Procedia* 32. Transport Survey Methods in the era of big data:facing the challenges, pp. 135–145.

Tchervenkov, Christopher, Joseph Molloy, and Kay W Axhausen (2018) Estimating externalities from GPS traces using MATSim. In: *18th Swiss Transport Research Conference (STRC 2018)*. STRC.

Tchervenkov, Christopher, Joseph Molloy, Alberto Castro Fernández, and Kay W Axhausen (2020) MOBIS study - A review of common reported issues. In: *20th Swiss Transport Research Conference (STRC 2020)*. STRC.

Transurban (2016) Changed conditions ahead. The transport revolution and what it means for Australians, Melbourne Road Usage Study Report. Report. Melbourne.

Verhoef, Erik, Peter Nijkamp, and Piet Rietveld (1996) Second-best congestion pricing: the case of an untolled alternative. In: *Journal of Urban Economics* 40 (3), pp. 279–302.

Vrtic, Milenko, Nadine Schuessler, Alexander Erath, and Kay W Axhausen (2010) The impacts of road pricing on route and mode choice behaviour. In: *Journal of Choice Modelling* 3 (1), pp. 109–126.

Vrtic, Milenko, Nadine Schüssler, Alexander Erath, and Kay W Axhausen (2007) Design elements of road pricing schemes and their acceptability. In: *TRB 86th Annual Meeting Compendium of Papers*. Transportation Research Board, pp. 07–1548.

Wargelin, Laurie, Peter Stopher, et al. (2012) GPS-based household interview survey for the Cincinnati, Ohio Region: executive summary report. Tech. rep. Ohio. Dept. of Transportation. Office of Research and Development.

Washbrook, Kevin, Wolfgang Haider, and Mark Jaccard (2006) Estimating commuter mode choice: A discrete choice analysis of the impact of road pricing and parking charges. In: *Transportation* 33 (6), pp. 621–639.
Wolf, Jean, Michael Loechl, Miriam Thompson, and Carlos Arce (2003) Trip rate analysis in GPS-enhanced personal travel surveys. In: *Transport survey quality and innovation* 28, pp. 483–498.
Wu, Linlin, Biao Yang, and Peng Jing (2016) Travel mode detection based on GPS raw data collected by smartphones: A systematic review of the existing methodologies. In: *Information* 7 (4), p. 67.

	Beta (SE)	HR (95% CI)	p
Income > 12,000 CHF	0.28 (0.09)	1.32 (1.10, 1.58)	0.003 **
Household size	-0.07 (0.03)	0.93 (0.87, 1.00)	0.038 *
Age (decades)	0.00 (0.03)	1.00 (0.95, 1.06)	0.883
Tertiary education	-0.19 (0.08)	0.83 (0.70, 0.97)	0.022 *
German speaking	0.03 (0.09)	1.03 (0.87, 1.22)	0.752
Female			
fortnight=1	0.02 (0.15)	1.02 (0.77, 1.35)	0.895
fortnight=2	-0.07 (0.20)	0.93 (0.62, 1.39)	0.721
fortnight=3	-0.04 (0.22)	0.96 (0.62, 1.48)	0.841
fortnight=4	-0.28 (0.12)	0.76 (0.60, 0.96)	0.022 *
Android			
fortnight=1	0.87 (0.16)	2.38 (1.73, 3.26)	0.000 ***
fortnight=2	0.46 (0.22)	1.58 (1.02, 2.45)	0.040 *
fortnight=3	-0.01 (0.25)	0.99 (0.60, 1.62)	0.960
fortnight=4	0.41 (0.13)	1.51 (1.17, 1.94)	0.002 **
Huawei			
fortnight=1	0.38 (0.20)	1.47 (0.99, 2.18)	0.057 .
fortnight=2	0.37 (0.32)	1.45 (0.78, 2.70)	0.239
fortnight=3	0.29 (0.41)	1.33 (0.59, 2.98)	0.487
fortnight=4	0.15 (0.21)	1.16 (0.77, 1.75)	0.465
Employed			
fortnight=1	-0.33 (0.16)	0.72 (0.53, 0.97)	0.033 *
fortnight=2	-0.07 (0.23)	0.94 (0.60, 1.47)	0.775
fortnight=3	0.24 (0.27)	1.27 (0.75, 2.15)	0.369
fortnight=4	0.05 (0.14)	1.05 (0.80, 1.38)	0.718
AIC		10484.33	
Coordance		0.602	
Num. events		655	
PH test		0.76	
Note:		$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$	

TABLE 2.2: Cox porportional-hazard model

Email & Treatment	n	% Opened	Times opened (mean)	Time to open (h) median (IQR)	
Welcome					
-	5475	82.36	2.78	8.50	(2.88 - 20.33)
Report 1					
	4168	84.88	2.13	7.37	(2.53 - 19.22)
Report 2					
	4132	81.03	1.87	6.66	(2.59 - 18.37)
Report 3					
	4105	78.59	1.83	6.19	(2.51 - 17.85)
Report 4					
Control	1247	79.23	1.62	5.40	(2.30 - 14.65)
Info	1262	83.68	1.99	5.40	(2.40 - 16.83)
Pricing	1222	82.90	2.64	6.06	(2.35 - 17.57)
Halfway					
Control	1250	76.80	1.60	5.60	(2.41 - 15.54)
Info	1263	83.29	1.72	5.50	(2.53 - 17.35)
Pricing	1222	80.93	2.17	5.51	(2.24 - 17.15)
Report 5					
Control	1243	76.43	1.55	5.96	(2.42 - 15.37)
Info	1255	80.80	1.90	6.28	(2.42 - 17.29)
Pricing	1213	80.54	2.24	6.94	(2.66 - 19.82)
Report 6					
Control	1238	77.06	1.87	5.78	(2.35 - 16.89)
Info	1252	78.12	1.87	5.87	(2.57 - 17.32)
Pricing	1208	79.22	2.09	6.24	(2.41 - 17.87)
Report 7					
Control	1235	74.98	1.61	5.83	(2.35 - 15.83)
Info	1248	77.64	1.66	6.08	(2.44 - 18.16)
Pricing	1205	80.25	2.02	6.07	(2.33 - 17.49)
Report 8					
Control	1231	79.69	1.50	6.11	(2.55 - 17.01)
Info	1246	78.33	1.46	6.41	(2.49 - 18.85)
Pricing	1200	81.50	2.01	6.55	(2.49 - 18.80)

TABLE 2.3: Engagement with various emails through the study

	Count model (1) Corrections/day		Zeros model (2) Correction/day > 0	
Constant	0.744	(0.032)***	1.504	(0.046)***
Phase 2	0.047	(0.014)**	0.050	(0.020)*
Age (decades)	−0.024	(0.003)***	−0.014	(0.005)**
Male	0.074	(0.012)***	0.047	(0.017)**
Treatment				
Control	-		-	
Information	−0.029	(0.022)	−0.053	(0.032)
Pricing	−0.083	(0.069)	−0.335	(0.103)**
Education				
Mandatory	-		-	
Trade/traineeship (baseline	−0.098	(0.023)***	−0.220	(0.033)***
Higher education	−0.014	(0.023)	−0.321	(0.033)***
Income (CHF per month)				
Less than 4000	-		-	
4000 <= 8000	−0.134	(0.022)***	−0.208	(0.032)***
8000 <= 12,000	−0.203	(0.022)***	−0.324	(0.032)***
12,000 <= 16,000	−0.230	(0.024)***	−0.429	(0.035)***
More than 16,000	−0.124	(0.025)***	−0.360	(0.038)***
Interactions				
Control * male	-		-	
Information * male	−0.027	(0.028)	0.139	(0.040)***
Pricing * male	−0.004	(0.027)	−0.001	(0.040)
pricing * mandatory	-		-	
pricing * trade/traineeship	−0.113	(0.057)	0.099	(0.081)
pricing * higher education	−0.166	(0.057)**	−0.023	(0.082)
pricing * less than 4000	-		-	
pricing * 4000 <= 8000	0.174	(0.059)**	0.278	(0.084)***
pricing * 8000 <= 12,000	0.285	(0.058)***	0.354	(0.083)***
pricing * 12,000 <= 16,000	0.187	(0.065)**	0.456	(0.092)***
pricing * more than 16,000	0.128	(0.068)	0.368	(0.099)***
Observations	147,450			
Log Likelihood	−127,206.400			
Note:	$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$			

TABLE 2.4: Zero inflated negative binomial model of the validation behaviour

Mode	% Correct			
	Android		iOS	
Airplane	99.48%		98.86%	
Bicycle	81.59%		79.14%	
Bus	66.98%		66.82%	
Car	92.98%		93.15%	
Rail	89.50%		91.05%	
Local train		88.67%		90.18%
Regional train		71.35%		73.40%
Subway		93.56%		92.53%
Train		63.13%		63.78%
Tram	95.01%		96.64%	
Walk	95.56%		97.21%	

TABLE 2.5: Comparison of the MotionTag mode detection performance bewteen iOS and Android

	Confirmed mode									
	Airplane	Bicycle	Boat	Bus	Car	Rail	Tram	Walk	Other	**Total**
Predicted										
Airplane	**2,113**	-	-	-	22	-	-	-	-	**2,135**
Bicycle	4	**26,201**	136	438	1,499	177	149	2,771	1,500	**32,875**
Bus	1	435	2	**35,713**	15,085	140	280	889	865	**53,410**
Car	372	2,495	741	8,028	**366,649**	3,314	1,950	2,834	7,433	**393,816**
Rail	64	56	85	1,748	7,298	**60,270**	691	258	298	**70,768**
Tram	-	49	2	128	396	60	**20,174**	149	16	**20,974**
Walk	80	3,807	456	1,224	9,960	868	868	**514,944**	638	**532,845**
	2,634	**33,043**	**1,422**	**47,279**	**400,909**	**64,829**	**24,112**	**521,845**	**10,750**	**1,106,823**

TABLE 2.6: Confusion matrix of mode detection accuracy

3

CALCULATING EXTERNAL COSTS ON GPS TRACES

FULL TITLE

Imputing the external costs of travel using GPS travel diaries: Towards a first-best mobility pricing scheme

AUTHORSHIP

Joseph Molloy, Christopher Tchervenkov and Kay W. Axhausen

This chapter has been submitted to *Transportation Research Part D: Transport and Environment*

3.1 INTRODUCTION

It is increasingly recognized that both the environmental and social costs of travel need to be internalized to manage the demand on already strained transport networks by encouraging shifts in travel patterns. In this direction, there is a growing body of evidence that informal feedback on energy use can encourage more efficient behaviour, both regarding home energy use (Faruqui, Sergici, and Sharif, 2010) and travel behaviour (Fujii and Taniguchi, 2006; Taniguchi et al., 2003). However, providing accurate and individualized feedback on external costs in transport is particularly challenging, primarily due to the heterogeneous nature of travel behaviour and the difficulties inherent in data collection at the individual level.

The main external costs of transportation can be divided into two groups: those that affect other users in the network, namely congestion and accident risks, and those that affect those outside the system such as noise and emissions (Button, 2004). These two categories are called intra- and inter-sectoral, respectively. The impact of congestion is primarily the loss of time spent waiting or slowed down in traffic, whereas emissions and noise have both environmental and health consequences.

Most of the literature on external costs in transportation focuses on road transport, partly because this is where the external costs are the highest (Maibach et al., 2008). The external costs arise because road users only consider their own costs of travel - known as the Marginal Private Cost (MPC) - and do not consider their own contribution to the total societal costs of road usage - known as the Marginal Social Cost (MSC), since the private costs (in time) of each trip rises with the number of drivers in the network. This is the primary intra-sectoral external cost. Combined with the inter-sectoral costs of pollution and noise, the unregulated level of travel exceeds the optimum for the network. As such, by imposing a price or tax equal to the difference between the MPC and MSC, the benefit to society is maximized and under ideal conditions, a pareto-optimum reached, where no one can be better off without disadvantaging another.

Certain constraints are required for this theory to hold. These include complete perfect information on prices and costs, and existence of the network within a broader economic system without market failures. Verhoef (2000) identifies the following components that need to be considered by an optimal charging system for road transport that encapsulates various external costs:

1. The vehicle technology used

2. The actual state and age of this vehicle
3. The time of driving
4. The place of driving
5. The actual route chosen
6. The driving style

Verhoef also notes that in reality, such first-best pricing can only be hypothetical, due to the level of surveillance required, and that practical schemes in transport will require second-best pricing. However, he argues that in the long-run, a scheme will benefit from being designed as closely as possible to these first-best principles.

Currently, when people choose to drive on a road, they usually only have to pay with their private time and for car usage costs, and not the societal costs of their trip, namely, the additional costs the driver imposes on other drivers by increasing demand along the route. From economic theory, internalizing this cost in the form of a tax would reduce congestion and provide an overall social benefit (Arnott and Small, 1994; Pigou, 1920). There are some examples where pricing policies have been implemented to control congestion, such as the area-based congestion charge in London (Leape, 2006) or the cordon system in Stockholm (Eliasson et al., 2009). In the United States, high-occupancy tolling on single highway segments is increasingly common, where one or some of the lanes are priced dynamically according to congestion, with prices changing as regularly as every 3 minutes (Janson and Levinson, 2014). However, the application of dynamic pricing to a whole network faces steep political and technical hurdles, not least of which is the design of a suitable pricing scheme that captures the complex interactions between users experiencing and causing congestion.

As a step towards better understanding these interactions, this paper presents a methodology for estimating the individualized external costs of car travel based on the output from GPS-based travel diaries. It is assumed that the GPS tracks have already been segmented into stages, and the transport mode identified. This approach was developed for the MOBIS mobility pricing study (Molloy et al., 2021), an 8-week study using GPS tracking where participants were presented with their external costs. The resulting software pipeline also supports per-kilometer values for other modes, such as walking, cycling and public transport, where a link-based approach is not feasible.

The paper proceeds as follows: section 3.2 covers the related literature on external costs and various approaches available. Section 3.3 describes the model methodology and its building blocks, referred to from hereon as the *externalities pipeline*, at a high level, before describing the respective methodologies for emissions and congestion in more detail. In Section 3.4, the methodology is validated against reference values from the literature, and applied to GPS data collected during MOBIS study. The results are discussed in section 3.5 and section 3.6 concludes.

3.2 BACKGROUND

Building on Pigou's two road model (**Pigou_1920**), Vickery's bottleneck model (Vickrey, 1963) has become a key model for examining congestion effects in a network (Arnott, De Palma, and Lindsey, 1993; Van Den Berg and Verhoef, 2011), as well as social-optimums via pricing (Chakirov, 2016) and pricing schemes (Laih, 1994). However, Arnott et al. (2001) note that traditional macroscopic models focus on link congestion, while ignoring or simplifying other elements of congestion such as modal congestion, parking, interactions with pedestrians and spillback effects. In particular, the importance of value of time heterogenity among individuals in road pricing models has been recognized by numerous researchers (Small and Yan, 2001; Verhoef and Small, 2004). Modern traffic microsimulation frameworks such as MATSim (Horni, Nagel, and Axhausen, 2016) are specifically designed to incorporate many of these various heterogeneities, making them useful for such modelling.

Fellendorf and Vortisch (2000) developed one of the first approaches for the microsimulation of pollutant emissions. A traffic flow model is used to calculate the speed and acceleration of each vehicle at a 1-second frequency, and available engine maps to calculate the emissions. More recently, Kraschl-Hirschmann et al. (2011) coupled the microscopic traffic flow simulator (VISSIM) with a microscopic emissions model (PHEM) to investigate the impact of traffic signalling on emissions. Ma, Mitchell, and Heppenstall (2015) used the output from a travel-diary survey to build a microsimulation of Beijing, and estimated the CO_2 emissions and possible reductions.

Kaddoura and Kickhöfer (2014) developed an agent-based marginal-cost pricing approach for congestion and applied it successfully to a large-scale scenario of Greater Berlin (Kaddoura, 2015). When considering the internalization of congestion costs, a particular contribution of this work was to assign the external congestion costs to the causing agents. In particular, they note that it is simple to calculate the incurred time-loss through congestion for each agent, but much more challenging to map it back to the causing agents. The approach calculates each agent's contribution to the delays on travelled links using a queue-based node-link model including spillback.

In real networks, such an approach would require knowing the location and of every driver connected to a particular incident of congestion, to determine who was affected. Quantifying the monetary value of the delay would then require knowing each affected driver's willingness to pay for a unit of travel time savings, a measure known as the value of travel time savings (VTTS). This is clearly unrealistic as it would involve tracking a large proportion of the population.

3.2.1 *MATSim and the Switzerland scenario*

MATSim is a powerful tool for performing agent-based transport simulations. A population is represented by a set of agents who try to optimize their daily travel plan over repeated iterations of the model, through a process called re-planning. It can handle scenarios consisting of millions of agents travelling on a city, regional or national transport network. It is designed as a modular event-based framework, where the actions of agents, such as a departure, arrival, link entry or exit, are events which are passed around the framework.

This event-based design makes the traffic flow simulation framework, separate from the re-planning component, well suited for processing individual-level mobility data at the trip level. The traffic flow simulation is based on a first-in first-out queue model where each link is represented as a queue with three attributes: the non-congested (freespeed) travel time t_{free}, the flow capacity c_{flow} and the storage capacity of the link $c_{storage}$. The link queues are updated typically every second, and agents are moved from one link to the next if the freespeed travel time on the link has passed, enough time has passed since the last vehicle left the link (the inverse of c_{flow}), and there is enough capacity $c_{storage}$ on the following link. Importantly, an agent who leaves a link prevents all following agents from leaving that link for the time of $1/c_{flow}$, and the tracking of agents restricted by $c_{storage}$ on the incoming links allows the consideration of spillback on congestion.

The IVT Switzerland Scenario builds on the work of Bösch, Müller, and Ciari (2016) to provide a MATSim scenario for all of Switzerland (Hörl, 2020) with a synthetic population for 2019. It represents a typical working day in Switzerland. As a MATSim scenario, the population consists of individual agents, each with daily travel plans (preferences) and social-demographic characteristics. These agents represent the entire population of Switzerland on a network generated from OpenStreetMap (Haklay and Weber, 2008). The scenario is available in 1%, 10% and 100% samples with respectively increasing runtimes (Hörl, 2020).

The simulations here are carried out using a discrete mode choice approach developed by Hörl, Balać, and Axhausen (2019). The initial travel demand, which comes from census data, is routed along the shortest path based on non-congested travel times. Then, at each iteration, a small fraction of agents are selected for re-planning. Each feasible mode alternative for each agent is routed along the shortest path based on the updated travel times from the previous iteration, and one is selected based on a discrete mode choice model. This process is repeated over several iterations until equilibrium.

3.2.2 *Analysis of road transport externalities in Switzerland*

Numerous sources are available for the analysis of external costs in Switzerland, including standards, government reports and databases. These sources guide and inform the evaluation of new and existing infrastructure projects. The Swiss Federal Office for Spatial Development (ARE) (ARE, 2020) produced a report on the external costs and benefits of transport in Switzerland, built on the methodology developed by Ecoplan / Infras (2019). It presents the most recent external cost-benefit analysis for the Swiss transport system, primarily focusing on external environmental, health and accident-related costs. Specifically, external costs for 12 different cost categories are computed, differentiated according to three different perspectives: transport mode (road/rail/air/water, passenger/freight, vehicle type), transport user and heavy vehicles.

For the modelling of road transport pollutant emissions in Switzerland (and other European countries), emission factors are commonly taken from the Handbook Emission Factors for Road Transport (HBEFA) (De Haan and Keller, 2004; Keller, Hausberger, et al., 2017). The HBEFA database contains emission factors for a range of vehicle categories and traffic situations, differentiated by emission type, pollutant and year. The HBEFA is the standard for road pollutant analysis in Germany, Switzerland and Austria, and is supported by the European Commission.

The Swiss Federal Office for the Environment (FOEN) (FOEN, 2010) also use the HBEFA to provide a detailed analysis of past and predicted future pollutant emissions, covering road transport in Switzerland from 1990 to 2035. Emissions values are calculated for three emissions types: emissions when the engine is in hot operating condition, cold-start emissions and evaporation emissions. The calculation of these values require both traffic volume data as well as the emissions factors from the HBEFA for each emission type. The report models the development of the vehicle fleet composition, vehicle specific mileage and emission standards trends, resulting in traffic volumes (mileage and start/stop processes) differentiated by vehicle category, emission standard and road category. These traffic volumes are then multiplied by the corresponding emissions factors (i.e. for road gradient and temperature) to obtain the final emissions values.

Concerning congestion specifically, Keller and Wüthrich estimate the external traffic delay costs for the years 2009 to 2014 (Keller and Wüthrich, 2016) and then again from 2015 to 2017 (Keller and Wüthrich, 2019). In this study, vehicle hours of delay for Switzerland were estimated and the proportion attributable to heavy vehicles determined. From 2013 onwards, this was achieved by combining and aligning INRIX traffic flow data and traffic demand data from the National Passenger Transport Model. The time lost per road section was calculated by subtracting the free-flow travel time from actual travel time, where traffic jams are considered to occur only when the actual speed is less than 65% of the free-flow speed. This approach only considers flow congestion, and not queuing delays. For the other years, online data from the Swiss Federal Roads Office (FEDRO) counting stations was used. A summary of their results is provided in Table 3.1. The values provide a useful estimate of delay costs in Switzerland. However, the use of an "at-least" approach will tend to underestimate the lost time and resulting associated delay costs (Keller and Wüthrich, 2016). This is particularly the case for non-motorway road segments, where long road lengths and imprecise speed data can influence results.

For the monetization of externalities, the Swiss Association of Road and Transportation Experts (VSS) has published a series of norms (SN 641 82* : Cost Benefit Analysis for Road Traffic) aimed at guiding the assessment of monetary effects and the cost benefit analysis of transport projects, policies and regulations. Norms SN 641 820 (Basic Standard), SN 641 822a (Travel Time Costs for Passenger Traffic) and SN 641 828 (External Costs) are of particular interest in the context of external cost evaluation. They provide standard values for time costs and willingness to pay per vehicle type and

	Congestion costs (M CHF/year)						
	Motorway		Non-motorway		All roads		
Year	LMV	HMV	LMV	HMV	LMV	HMV	Both
2010	608.7	61.4	449.9	17.0	1,058.6	78.4	1,137.0
2011	634.3	63.8	454.4	17.2	1,088.7	80.9	1,169.7
2012	672.2	67.8	458.8	17.3	1,131.0	85.1	1,216.0
2013	645.7	65.5	462.7	17.4	1,108.4	83.0	1,191.3
2014	690.6	70.4	466.6	17.6	1,157.1	87.9	1,245.1
2015	736.6	71.0	468.3	17.7	1,204.9	88.6	1,293.5
2016	780.9	76.9	471.4	17.8	1,252.3	94.7	1,347.0
2017	840.6	87.6	473.8	17.9	1,314.4	105.5	1,419.9

Estimated congestion costs for Light (LMV) and Heavy Motorized Vehicles (HMV), Keller and Wüthrich (2019), p.20

Table 3.1: Estimated congestion costs for 2010-2017

trip purposes as well as standard methods for evaluating the monetary impacts of air pollution and climate impacts.

3.2.3 *Limitations of the aggregate values*

The values from the Swiss standards are only available as CHF/km or CHF/h. Although some external costs are given under a urban/rural or motorway/non-motorway classification, there is no temporal or spatial variation. One hypothesis of this paper is the following: For private car travel, the variation in external costs is significant, and justifies a disaggregate approach to the calculation of external costs. In the following work, this is done for the calculation of private car emissions and congestion delays. For noise, this was found to be too computationally expensive to do on a national scale, and hence per-kilometer values from the norms are still used. Other researchers have previously identified the usefulness of disaggregate noise models (Kaddoura, Kröger, and Nagel, 2017; Kuehnel and Moeckel, 2020). However, their noise calculations were carried out only on a city level and are based on the German RLS-90 approach (Bundesminister für

Verkehr, 1990; Forschungsgesellschaft für des Straßenverkehr (FGSV), 1997), which differs from the norms used in Switzerland.

3.3 METHODOLOGY

In this section, the methodology for estimating externalities on GPS data is presented. The approach requires that the GPS data be already segmented into trip-stages and labelled with the transport mode used. This can be done with one of many methods. For an overview, see (Zheng, 2015). In the case of this paper, the data was segmented and labelled with the transport mode by the GPS tracking app 'Catch-my-Day' (MotionTag GmbH, 2019a), developed by MotionTag GmbH (MotionTag GmbH, 2019b).

The methodology requires a few static data inputs for the calculation of various externalities. Reference values for both emitted air pollutants and caused congestion are required. For the calculation of emissions, the HBEFA database (version 3.3) is used (Keller, Hausberger, et al., 2017). For congestion, average 15-minute-interval values of the delay caused by a vehicle present on that link are calculated for each link using a 10% sample from the 2019 MATSim scenario for Switzerland (see Section 3.2.1). This is done using the approach of Kaddoura and Kickhöfer (2014), described in more detail in section 3.3.5.

A multistage pipeline has been developed for estimating car-based externalities on labelled GPS traces using the MATSim framework. The pipeline consists of the following steps, described in more detail below:

1. Cleaning of GPS data
2. Map matching to the MATSim network using Graphhopper
3. Calculation of link entry and exit times
4. Conversion to MATSim events
5. Estimation of externalities on MATSim events
6. Monetization of the externalities

More broadly, the pipeline is grouped into two stages: the first creates a series of MATSim events representing the map-matched path of the GPS traces; the second processes those events using the previously mentioned reference values to estimate the generated emissions and delays.

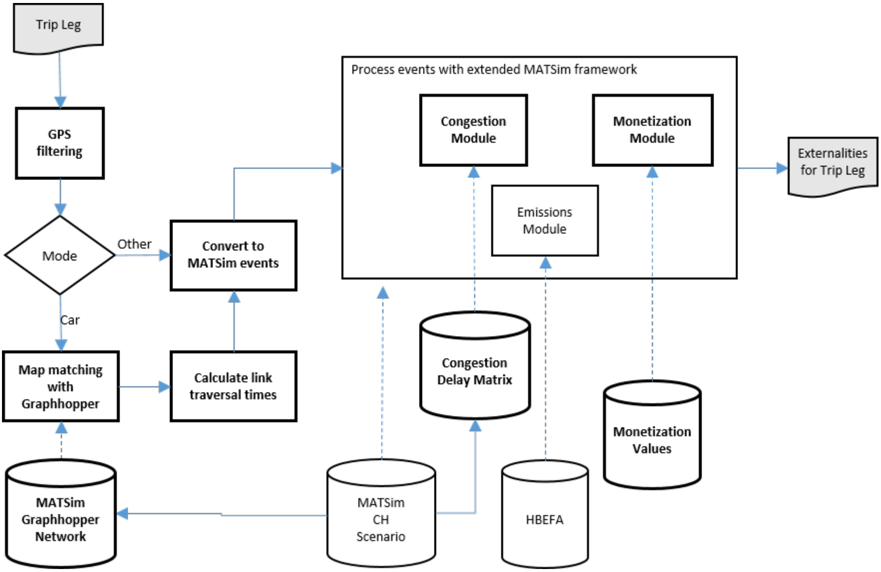


FIGURE 3.1: MATSim-based externalities pipeline

Figure 3.1 illustrates how data flows through the externalities pipeline. The objects in bold are those developed as part of this paper. Dotted lines indicate data inputs from static sources, and solid lines are the flow of the GPS-based trip data through the model. The lack of flows inside the MATSim framework is intentional, as those modules are built on top of the MATSim event framework. This is discussed further in Section 3.3.8.

3.3.1 *Data cleaning*

GPS data accuracy can vary considerably depending on the sensor used, the surrounding environment and even geographical location. Hence, any GPS points not within 200m of a segment of the Swiss road network are removed before map matching.

3.3.2 *Map matching with Graphhopper*

To map trip legs to the MATSim network, the Graphhopper (Graphhopper, 2018) map-matching library was modified to support matching to a MATSim network instead of OpenStreetMap. Graphhopper uses a Hidden

Markov Model (Newson and Krumm, 2009) to identify candidate links for each GPS point, with an error radius, σ, which in our case was set to 200m - equivalent to the filtering distance used to exclude GPS points.

An unlimited distance between consecutive points is allowed. The Graphhopper routing engine then identifies the best route between the set of candidate links, where a minimum of two-matched GPS points are required. However, the standard implementation of Graphhopper does not calculate the entry and exit timestamps for each link in the network, which are needed to calculate the time spent and average speed on each link. Additionally, in the absence of high-frequency GPS measurements or additional sensor information, there may be insufficient GPS measurements to pinpoint the entry and exit times for each link. The MATSim compatible version of Graphhopper was been extended to return the entry and exit times for each link, including links where few or no GPS measurements are available.

3.3.3 *Calculation of link entry and exit times*

A trip leg contains a sequence of links L with the set of GPS points $P(l)$ matched to each link l. For convenience, let the first and last GPS point on each link in the set L be $p_{l,s}$ and $p_{l,e}$ respectively. The start and end links of a trip leg always have at least one GPS point associated with them, while other links may have none or more GPS points. Hence, trip legs are divided into sets of consecutive links L', beginning at l'_1, where $l'_{2..k}$ have no GPS points. The GPS recorded travel time over the links in L' is then proportionally allocated based on the non-congested (freespeed) travel time of each link L' for which there are no measurements, where l_{k+1} is the next non-empty link.

Let the projection of GPS point $p_{l,i}$ onto link l be $p'_{l,i}$. $t(p)$ gives the time at point p. $t\_link(l)$ gives the travel time on link l, $t\_gps(a, b)$ gives the recorded time between two GPS points a and b and $t\_network(a, b)$ gives the summed travel time over a set of links in the network. A helper function $t\_between(a, b)$ returns the time needed to travel between projected points and the vertices of a link l, travelling at the freeflow speed for that link: for example, from $p'_{l,e}$ to the end of the link; or the start of the link to the first projected point on that link $p'_{l,s}$.

In MATSim, the assumptions hold that an agent always starts and ends somewhere on a link. Hence, only the exit time for the first link and the entry time for the last link need to be calculated. Additionally, $entry\_t(l_j) =$

$exit\_t(l_{j-1}), \forall j = 1..n$. As such, the algorithm can be separated into two cases:

- **First Link** For the first link l_1,
  $exit\_time(l_1) = t(p'_{l_1,e}) + t\_between(p'_{l_1,e}, l_1)$
- **Other Links**
  $entry\_time(l_j) = exit\_time(l_{j-1})$
  if $P(l_j) = \emptyset$
  then
  $exit\_time(l_j) = entry\_time(l_j) + t\_gps(p'_{l_{j-1},e}, p'_{l_k,s}) \cdot \frac{t\_link(l_j)}{t\_network(\{i,...,j,...,k\})}$
  where l_i and l_k are the most recent and next link with $P(l_k) \neq \emptyset$, respectively
  else $exit\_t(l_j) = entry\_t(l_j) + t\_between(l_j, p'_{l_j,e})$

The sequence of links with entry and exit times are then converted to valid MATSim events for each person and date.

3.3.4 *Estimation of emission externalities*

To estimate the externalities of each trip leg, the generated events are processed using the MATSim framework, extended with two additional modules. The first, developed by Hülsmann et al. (2011) and Kickhöfer and Nagel (2016), applies the HBEFA factors to calculate the emitted pollutant amounts incurred on each link, based on the observed travel speed on that link. The emissions factors are taken from the HBEFA database (version 3.3). To this module a few extensions have also been made. The module was originally designed to work with simulation output from MATSim, where real world boundary conditions (speeding) and data artifacts are not present. Hence, average speeds on each link are now capped at the freespeed of the link. Furthermore, the road types for assigning emissions factors are extracted from OpenStreetMap, rather than a VISUM model, as was done in the original Berlin Scenario. These improvements have been contributed back to the MATSim codebase, in accordance with open-source principles of the MATSim framework.

The HBEFA provides four traffic states, free-flow, heavy, saturated and stop&go, while MATSim considers only two in its queuing model - free-flow or queuing (to exit the link). Hülsmann et al. (2011) align these by assigning the difference between the actual travel time and the free-flow travel time on a link (the congestion) to the HBEFA stop&go traffic state, and the rest to free-flow. In doing so they ignore the heavy and saturated states. However, in the original paper, they also suggest an alternative version, which accommodates all 4 HBEFA traffic states, using the average speeds of each traffic state provided in the HBEFA. In this paper, we implement this method, allowing for all 4 HBEFA traffic states to be considered in the emissions model.

The emissions module outputs quantities in non-monetized terms. These are then converted to monetary damages using the most current norm values for Switzerland derived from the "Nachhaltigkeits - Indikatoren für Strasseninfrastrukturprojekte" (NISTRA) (FEDRO, 2017), which is itself based on the Swiss Standard SN 641 820 (VSS, 2013). For this work, the values were revised for the year 2019, and the values used are presented in Table 3.2. For PM_{10} emissions, distinct normative values were available for urban and rural areas. Links in the network were assigned the rural or urban classification based on the Swiss building codes (ARE, 2017). Links in unbuilt areas were assigned as rural, and all others as urban. The assignment was done based on the midpoint of the link.

Emission	Aspect	Value	Unit
Scenario year		2019	
CO_2	Climate Costs	136.08	CHF/ton [ab]
PM_{10} Costs (Healthcare)	Rural	515,497	CHF/ton [ab]
	Urban	1,358,461	CHF/ton [ab]
NO_x	Regional	7,109	CHF/ton [ab]
VTTS		25.77	CHF/h [c]

[a] FEDRO (2017) - updated for 2019,
[b] metric tons, [c] scaled nominal wage rate from [a]

TABLE 3.2: External costs of emissions

The NISTRA does not specify whether its monetization values are average or marginal. However, it is widely recognized that for air pollution costs, the marginal costs are virtually the same as the average costs, as numerous epidemiological studies have shown that the relationship between pollutants and health effects are almost linear (Van Essen et al., 2019).

3.3.5 *Estimation of congestion externalities*

The calculation of the experienced delay is a simple affair, if one makes the broad assumption that all delay is attributable to other users in the system, and not external causes such as signal control, rogue pedestrians and extraordinary events. However, calculating the true caused delay to other users in the network would require GPS traces for all users of the transport network. As such we use an average model of caused delay from the output of the MATSim scenario for Switzerland. This method gives the *average* marginal external cost for travelling on each link in a certain time window.

The approach of Kaddoura and Kickhöfer (2014) is used to calculate the caused congestion on each link by an agent. The approach has a number of diverging implementations, and in this paper we apply version 3, where the delays caused to each agent are allocated to the agents ahead in the queue until the delay is fully internalized. For each link in the network, the entry and exit times of each simulated agent are stored as a queue of potential delay-causing agents. Each time an agent exits a link with a delay, this queue is iterated through, and each causing agent pays for $1/c_{flow}$, the delay they caused on that link, until the delay is internalized. If an agent exits a link without delay, the previously stored queue on that link is reset. Any remaining non-internalized delay is considered to be a result of $c_{storage}$, which is carried over to the next link closer to the bottleneck and then distributed to the stored queue for that link. In this manner, the delays caused by an agent to other agents on other links in the network can be accounted for. For a specific example of how this algorithm works, the reader is referred to Section 3.2 of (Kaddoura and Kickhöfer, 2014).

A 30-hour MATSim simulation period is used to allow all trips to conclude, and the average delay caused by a vehicle for each link in the network over a set of time windows covering an entire day (24 hours) is computed. Let $x_{l,t,a}$ be the delay caused on link l at time t by agent a to all other agents in the network which might have been affected on other links. $A_{l,t}$ is the number of agents who passed through link l in time period t. The average delay caused by travelling on link l at time t is then given by

$$x_{l,t} = \frac{\sum_a x_{l,t,a}}{A_{l,t}}$$

This gives a matrix of dimensions L x $(1440/T)$ where L is the number of links in the MATSim network, and T the size of the time period in minutes. The value of 1440 corresponds to 24 hours in minutes. For a trip matched to the MATSim network, it is then trivial to obtain the average caused marginal delay on each traversed link and calculate the average marginal delay caused by the trip. In an ideal world, the time loss caused to each agent in the simulation would be monetized individually based on the VTTS of the affected agent during that trip, before the aggregation was performed. However, as this information is not contained in the MATSim scenario, congestion externalities were monetized using the Swiss reference value of time (VOT), and the monetization factor can be applied to the aggregated values.

3.3.6 *Other external costs*

Where monetization by discrete link based quantities is not implemented or possible, externalities are calculated directly using available CHF/km values. Ideally, noise emissions from car travel would be calculated based on the surrounding population, other noise emitters and the presence of noise reduction features. The Swiss norms (SN 641 828) provide decibel thresholds, above which health-related costs for the affected persons are to be considered. Using the MATSim scenario for Switzerland in a manner analogous to congestion, noise cost values per link per time could be calculated. However, it was unfeasible within the scope of this project to develop and integrate a nation-wide noise model. Instead, a normative value for noise emissions in CHF/km (differentiated by mode) is used. An in-depth spatially and temporally variant consideration of noise for all motorized modes is left as future work to improve the described methodology.

For walking and cycling modes, there are no pollutants or significant congestion externalities to calculate (at least in Switzerland, and excluding E-bikes). Hence, health benefits and damages are calculated in the pipeline on a CHF/km basis.

The pipeline is designed to support the marginal external cost calculations for other modes for which only per-kilometer values are available. Furthermore, the pipeline is adaptable to support the mapping of public transit trips to links in a MATSim transit network. This would enable the calculation of link-level congestion externalities on public transport, if data on crowding was available. Maibach et al. (2008) estimate crowding externalities to be roughly 50% of the VOT. However, they also note that the VOT varies greatly by transport mode, trip purpose and travel distance. The externality would also depend on the definition of "crowding".

3.3.7 *Calibration of the congestion model*

Congestion externalities were computed using the MATSim framework, as described in section 3.2.1. A 10% scenario for Switzerland for 2019 was first simulated for 40 iterations such as to reach equilibrium in terms of mode shares. Since MATSim is a stochastic simulation, this equilibrium is one within a distribution of possible outcomes. Therefore, the average caused congestion per link per time was computed as described in section 3.3.5 over a further 30 iterations, where the agents could only reroute their trips during re-planning. Delays only contribute to the average congestion if the corresponding travel speed was less than 65% of the free-flow travel speed on that link, consistent with the methodology used by Keller and Wüthrich (2016).

The median congestion across these additional 30 iterations was then computed and converted to a per-kilometer cost (fig. 3.2). During each additional MATSim iteration, 10% of agents are allowed to modify their chosen routes. This occasionally results in multiple agents simultaneously choosing to travel on previously non-congested links in the next iteration, and if this number of agents is large, it might take several iterations before they are randomly selected and routed along other less-congested paths. Thus, these oscillations in route choice can still result in high median congestion costs on a few links at certain times; in the current case, the maximum median cost per kilometer is nearly CHF 4,500. Therefore, the congestion costs per kilometer were capped to the 95 percentile value of the distribution, corresponding to a maximum cost per kilometer of just under CHF 2. After capping, the average cost per kilometer over all links exhibiting congestion is 0.22 CHF/km. A comparison of different capping thresholds is presented in Section 3.4.3.

3.3.8 *Software Architecture*

To enable compatibility with both MATSim and Graphhopper, the pipeline has been built to run on the JAVA virtual machine. Java version 1.8 or above is required. MATSim version 11.0 and Graphhopper 0.12.0 are used. A script has been developed to divide the run into multiple instances of the pipeline to allow computation in parallel, allowing the expedient generation of results from the pipeline.

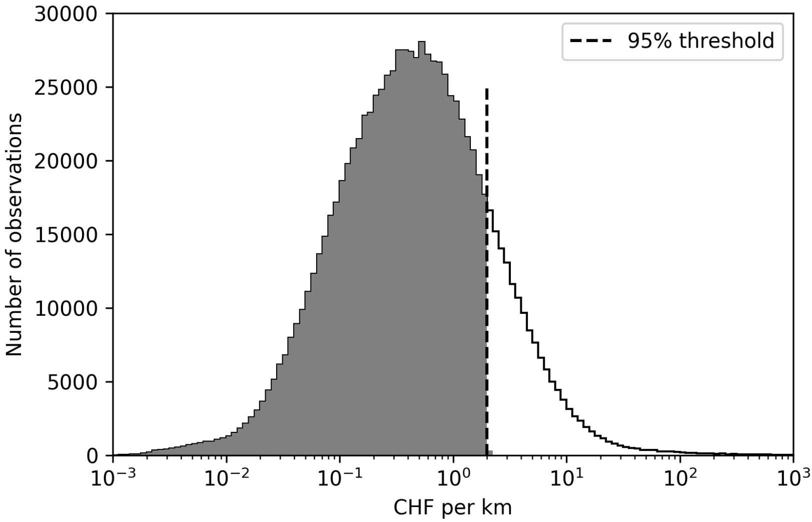


FIGURE 3.2: Distribution of per-kilometer congestion costs in the MATSim scenario with 95% threshold (log-scaled horizontal axis).

The input and output stages of the pipeline are also modular, meaning that data can be read or written from either a database or JSON files, allowing for the integration with different data sources and other models.

The externalities pipeline described in this paper takes advantage of the event-based framework in MATSim. Each module is designed as an *EventHandler* in a event-listener framework, where a function is called when certain events are fired, such as when a vehicle enters or exits a link. Sequences of these events are used to determine values such as travel times and average speeds on links. These handlers in turn generate new events such as a *WarmEmissionEvent*, containing the amounts of various pollutants produced by an agent travelling on a link.

On receiving a link-exit event, the congestion module determines the estimated congestion caused on that link based on the link id and exit time. The monetization module processes arrival events - which denote the end of a trip - to tally up all the produced externalities, and compute the monetary damages. All externalities are available on a trip leg level.

3.4 RESULTS

Using the output of the Swiss MATSim scenario, which represents a simulated day for a synthetic population of Switzerland, the externalities pipeline is first validated by comparing the computed externalities with emission and congestion estimates from previous Swiss external cost reports. The pipeline is then applied to GPS tracking data collected within the course of the MOBIS mobility pricing study (Molloy et al., 2021) to demonstrate the heterogeneity in external costs that can be observed in the data.

3.4.1 *Emissions*

To validate the estimation of pollution externalities using MATSim, the total emissions are calculated using a 10% MATSim scenario for Switzerland and compared to the reference values available in the literature. To accommodate the new vehicle registration statistics according to Blessing and Burgener (2013) and Bianchetti et al. (2016), the personal vehicle fleet composition was adjusted to match the mileage-weighted fleet composition projected by the Swiss Federal Office for the Environment (FOEN) (FOEN, 2010) for 2020 (table 3.3).

Emission values are estimated for the following pollutants: CO_2, CH_4, N_2O, PM_{10} (exhaust and non-exhaust) and NO_x. Figure 3.3 shows hourly

	Mileage-weighted share	
Emission level	Petrol	Diesel
Euro-0	0.00	0.00
Euro-1	0.00	0.00
Euro-2	0.02	0.00
Euro-3	0.03	0.01
Euro-4	0.13	0.07
Euro-5	0.13	0.11
Euro-6	0.26	0.23

TABLE 3.3: Passenger vehicle fleet composition by emission concepts (FOEN, 2010)

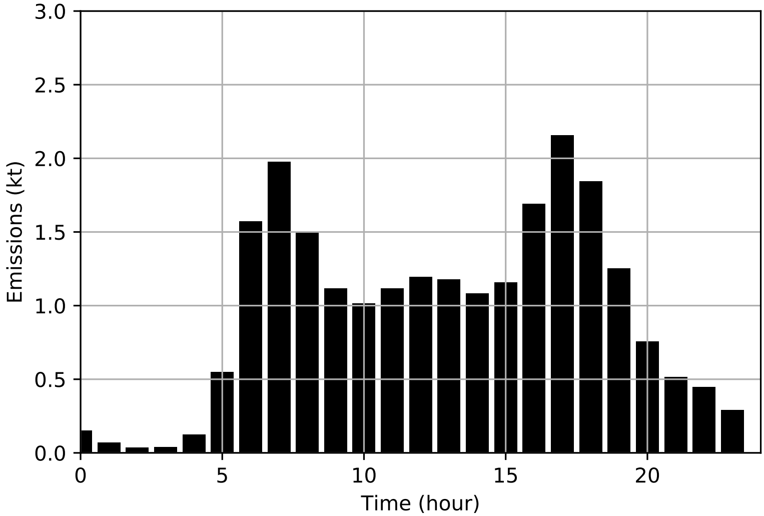
CO_2 emissions estimated from the MATSim scenario. As expected, these correlate with typical commuter patterns: two distinct peaks during the morning and evening rush-hour, low emissions in the morning and at night and higher values around noon.

The MATSim computed emissions values are then compared to those estimated by FOEN for 2020. Since MATSim simulates a single workday, the MATSim emissions values are scaled such as to match the total yearly travel distance by car reported by FOEN for 2020. Table 3.4 compares the total estimated emissions values for both MATSim and FOEN in metric tons per year. Deviations are likely due to the fact that emissions factors depend on the exact type of petrol or diesel engine.

3.4.2 *Congestion*

Contrary to emissions, congestion caused cannot directly be estimated and assigned to the causers from GPS traces alone, since information on how many other drivers were present on the road at that given moment is lacking. Hence, MATSim is used to estimate the marginal congestion externalities during a typical workday.

To assess the suitability of this approach, we compare the total calculated congestion costs over the course of a 10% simulation of the MATSim scenario with those computed by Keller and Wüthrich (2019). As a calibration step to account for unresolved oscillation effects in the MATSim scenario, the per-kilometer congestion costs are limited to the 95% percentile (see



Hourly CO_2 emission values for Switzerland in kilotons (metric). Pollutants other than total CO_2 are omitted as their values are negligible in comparison.

FIGURE 3.3: Hourly CO_2 emissions for Switzerland

Section 3.3.7). The total congestion costs from the scenario are scaled to be equivalent to the total yearly travel distance reported by Keller and Wüthrich, for motorways and non-motorways respectively. The comparison between the congestion costs in the MATSim scenario for Switzerland and the reference values is presented in Table 3.5 for different percentile thresholds, with the 95% percentile threshold in bold font. The corresponding per-kilometer congestion cost for each threshold is also reported.

The total yearly congestion values and thus the resulting costs estimated in MATSim for motorway and non-motorway road segments are lower respectively higher than those estimated by Keller and Wüthrich. This may be due to several factors. On the MATSim side, the model simulates passengers vehicles as well as trucks during a typical workday, and therefore does not account for seasonal variations in travel demand nor extraordinary circumstances such as large events, accidents and holiday traffic which also impact yearly congestion. Unlike emissions, which mainly depend on the to-

Pollutant	FOEN (tons/year)	MATSim (tons/year)	Ratio (MATSim/FOEN)
CO_2	10,167,283	8,740,076	0.86
CH_4	380	361	0.95
N_2O	159	103	0.65
PM_{10} (exhaust)	234	218	0.93
PM_{10} (non-exhaust)	2,381	2,638	1.11
NO_x	12,344	12,045	0.98

Table 3.4: MATSim and FOEN estimated emission totals comparison

tal distance travelled, congestion is highly dependent on the actual demand patterns, road infrastructure and travel behaviour. Thus, any deviations in the mode shares, route choice or road capacities affect the computed congestion values. In addition, the grouping of the MATSim estimates by road segment type is based on OSM data, which might differ from the classification used by Keller and Wüthrich. Finally, Keller and Wüthrich state that they have taken an "at-least" approach in estimating delays and that the values for non-motorway segments are highly underestimated. A combination of these effects likely contributes to the underlying cause of the deviation between the estimates.

3.4.3 *Sensitivity analysis*

Taking the 95% percentile provides a good calibration against the overall total costs calculated by Keller and Wüthrich. However, the MATSim congestion cost estimates are sensitive to the chosen threshold. The sensitivity is evident for both motorway and non-motorway road types. We propose that this sensitivity stems from the long tail in the distribution of the per-kilometer congestion cost values (see fig. 3.2), caused at certain bottlenecks in the MATSim network for Switzerland, where route-choice oscillations result in an *all-or-nothing* switching behaviour between routes. This oscillatory behaviour remains an open problem. It may be that once this is solved, the thresholds are no longer needed. The sensitivity analysis suggests that before applying this methodological approach in other study

Road Type	MATSim						Reference
	90%	92.5%	**95%**	97.5%	99%	100%	
Motorway	245.3	284.9	**332.3**	398.0	461.5	595.3	840.6
	(0.29)	(0.34)	(0.40)	(0.47)	(0.55)	(0.71)	
Non-motorway	788.7	924.4	**1,090.0**	1,326.3	1,580.9	3,169.5	473.8
	(1.66)	(1.95)	(2.30)	(2.80)	(3.34)	(6.69)	
Total	1,034.0	1,209.3	**1,422.3**	1,724.3	2,042.4	3,764.8	1,314.4
	(0.79)	(0.92)	(1.08)	(1.31)	(1.55)	(2.86)	
Max (CHF/km)	0.84	1.23	**1.94**	3.73	7.69	4,476.58	

TABLE 3.5: Comparison between reference congestion totals (M CHF) (Keller and Wüthrich, 2019) and estimates with MATSim for different per-kilometer thresholds. Ratio between MATSim and reference values in parentheses.

regions, individual characteristics of the scenario and network need to be taken into account through a calibration step as was performed above.

3.4.4 *Capturing the heterogeneity in external costs*

As noted in the introduction, is it important to capture both the temporal, spatial, and individual variation in external costs when assessing the policy implications of proposed measures to tackle emissions and congestion. Using a set of over 1.6 million car trips collected from 3,680 participants during the MOBIS study, the external costs for each trip were calculated using the methodology presented in this paper.

fig. 3.4 demonstrates the heterogeneity in external costs observed in the GPS data, as opposed to the available average per-kilometer reference values taken from Table 3.6. The range of the external costs is smaller for pollution emissions than for congestion. The mean values are still consistent with the reference values. Using the map-matching computed with graphhopper, the motorway-share of the trip is computed to allow the application of values for highway and non-highway kilometers separately, rather than just an average.

In fig. 3.5, the hourly variation between the two methods is compared, on a trip level. In subplot (a), a roughly constant split between highway

External cost	Value (CHF/km)
Congestion (average) [a]	0.00939
Highway	0.0206
Non-highway	0.0015
Emissions [b]	
CO_2	0.0233
PM_{10}	0.0388
NO_x	0.00939

[a] Keller and Wüthrich (2016) and Keller and Wüthrich (2019), [b] FEDRO (2017) and Rexeis et al. (2013)

Table 3.6: Average monetization values

and non-highway travel throughout the day results in a nearly constant average hourly external cost with the ARE method, even during the middle of the night. On the other hand, average values with the MATSim method vary between 0 and 0.05 CHF/km depending on the time of day. The small spike around hour 24 is due to agent behaviour in the MATSim scenario, and could benefit from further calibration. Subplot (b) shows the total externalities caused over the observation period. Hence, the increase in traffic during the peak-hours does lead to some temporal variation in the total emissions with the ARE method, but much less than with the MATSim method.

In Table 3.7, a summary of the external costs for different periods of the day are presented for the MATSim method. The morning and evening peaks cover 7:30am to 9:30am and 4:30pm to 7:30pm respectively. Here one can see that while the maximum marginal external costs are high (i.e. 15 CHF for one trip in the evening peak), the 95% percentile is much lower. The average external cost per kilometer in the morning peak is 1.63 times higher than the daily off-peak average, and the evening peak is 2.22 times higher. The per-trip values have similar ratios. The minimum values are not shown, being zero for all time periods.

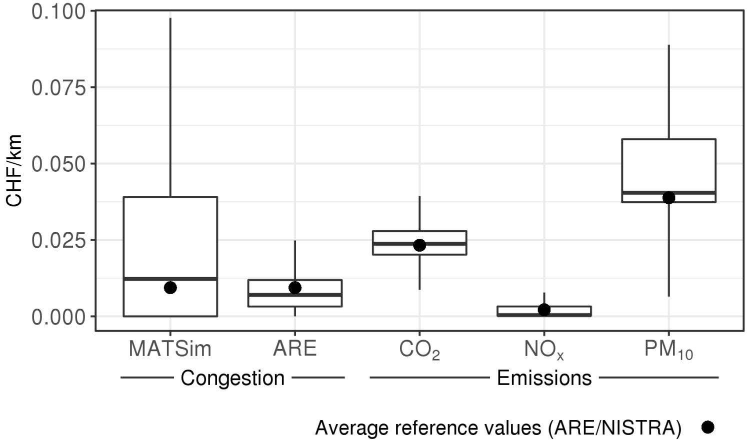


FIGURE 3.4: External cost (CHF/km) by trip. The ARE congestion distribution applies a highway/non-highway classification, and hence still requires map matching to a network.

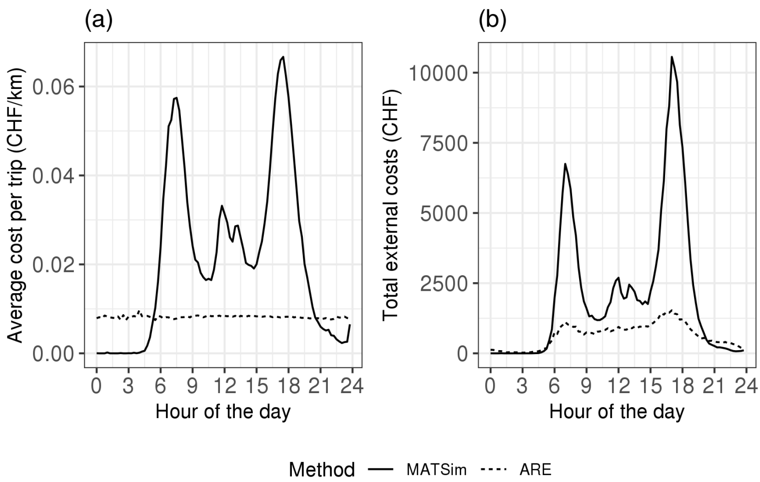


FIGURE 3.5: Total external cost by time of day (15 minute resolution). a) the average cost per trip at the starting time of the trip. b) The total external cost of trips starting at that time.

	Period	Median	Mean	95%	Max
Per-Km	Morning peak	0.016	0.036	0.140	0.937
	Day off-peak	0.011	0.022	0.083	0.769
	Evening peak	0.029	0.049	0.169	1.056
	Night off-peak	0.004	0.019	0.091	0.625
Per-Trip	Morning peak	0.197	0.434	1.699	8.059
	Day off-peak	0.129	0.275	1.060	13.858
	Evening peak	0.328	0.652	2.384	15.363
	Night off-peak	0.049	0.297	1.493	9.554

TABLE 3.7: Summary of congestion externalities for the MATsim-based method

3.5 DISCUSSION

As Verhoef (2002) stipulates, it is important to consider external costs on an individual level. Not only is this important in understanding the spatial and temporal distribution of the external costs within a transport network, it is also an important step towards understanding the potential impact of pricing policies. The method presented here shows how an agent-based transport microsimulation can be applied to real-world trip data to calculate external costs at an individual trip level. This approach captures much more variation than the use of average values.

The validation of the model in section 3.4 identifies some discrepancies between the reference values and the output of the proposed method, resulting from the various limitations of both approaches. However, an exploration of the trip-level heterogeneity on the real-world data indicates that the mean per-kilometer averages for both congestion and emissions are very close to the reference averages.

The main insights follow from an exploration of the temporal variation captured by the congestion model. The mean values hide much of the temporal variation in external costs, and there are implications for policy analysis and transport planning. While average external cost values for congestion and emissions are currently used for the cost-benefit-analysis of new transport projects in Switzerland, the analysis in this paper clearly shows that the respective total societal congestion costs and benefits would be distorted for policies aiming at reducing peak-hour congestion. One such policy might be a mobility pricing scheme, where the use of the average congestion values would lead either to an ineffective price structure, or the benefits of the scheme being undervalued.

The influence of the emissions model on the variation in external costs is less pronounced, but an effect is still evident. In particular, the external costs of PM_{10} emissions are a large component of the estimated externalities, and these vary greatly depending on the age, size and engine type of the vehicle (Rexeis et al., 2013). These effects could be considered using the vehicle data reported by the study participants. The results indicate that reliance on the average per-kilometer values of external costs neglects the large variation around the average per-kilometer values, which undervalues the use of more efficient vehicles.

There are some limitations to the approach used. The performance of the map-matching step is reliant on the quality of the GPS input, the segmentation and mode detection. In a small minority of cases, the map matching is not consistent with the route chosen, which may lead to an over- or underestimation of the external costs. The assumption is also made that the owner used their reported vehicle, and not another one. Additionally, the MATSim model used to estimate the link-level external costs does not take into account other variations in demand that may affect the delay caused by a driver. These include accidents, changes in road conditions, and day-to-day variation in traffic. The external cost of scheduling delays is also not incorporated into the calculation of external costs, though this would be possible using the MATSim framework. As Arnott, De Palma, and Lindsey (1990) note, the scheduling delay costs can be equal to the costs from congestion delays. The model also relies on the assumption that the MATSim scenario accurately reflects average conditions on the network, although thresholds were needed to account for outliers resulting from the oscillating behaviour in the scenario. In future work, observed travel times and delays from real-world GPS data could be used to make link-level adjustments to the model on a day-to-day basis to account for these variations.

3.6 CONCLUSION

This paper presents a methodology for estimating the externalities on GPS traces using the MATSim framework. A MATSim scenario for Switzerland is used to provide aggregate estimates of caused congestion for 15-minute time periods. Pollutant emission factors are taken from the HBEFA. The suitability of the MATSim scenario for this purpose is evaluated by validating the Switzerland-wide externalities against published reference values. The agent-based aspect of MATSim allows for a much finer calculation of externalities by taking into account the heterogeneity in both the population and travel behaviour. The validation step indicates that the aggregated congestion model calculated on the MATSim scenario for Switzerland is suitable for this purpose, with some caveats. Although the total external costs of congestion obtained from the scenario for motorways and other roads are lower respectively higher than the reference values when considered separately, the combined estimates are within 8% of the published values. An analysis of the heterogeneity in the external costs shows that the approach captures important variation in the external costs for different externalities, around mean values which are consistent with the published values for Switzerland. These results indicate that a proper consideration of the individual, spatial and temporal variation in external costs, and not just the mean values, is important to the analysis of potential transport policies and projects.

REFERENCES

Arnott, Richard et al. (2001) The economic theory of urban traffic congestion: a microscopic research agenda. In: *Workshop on Environmental Economics and the Economics of Congestion: Coping with Externalities, Venice International University, Venice Summer Institute, San Servolo, Italy. Found online at http://FMWWWW. bc. edu/ECV/Arnott. fac. html*.

Arnott, Richard, Andre De Palma, and Robin Lindsey (1990) Economics of a bottleneck. In: *Journal of Urban Economics* 27 (1), pp. 111–130.

Arnott, Richard, Andre De Palma, and Robin Lindsey (1993) A structural model of peak-period congestion: A traffic bottleneck with elastic demand. In: *The American Economic Review*, pp. 161–179.

Arnott, Richard and Kenneth Small (1994) The economics of traffic congestion. In: *American scientist*, pp. 446–455.

Bianchetti, Roberto, Peter de Haan, Michel Müller, and Sebastian Dickenmann (2016) Energieverbrauch und Energieeffizienz der neuen Personenwagen 2015 : 20. Berichterstattung im Rahmen der Energieverordnung. Tech. rep. Bern: Ernst Basler + Partner AG.

Blessing and Burgener (2013) 17. Berichterstattung im Rahmen der Energieverordnung über die Absenkung des spezifischen Treibstoff-Normverbrauchs von Personenwagen 2012. Tech. rep. Bern: Auto Schweiz.

Bösch, Patrick M, Kirill Müller, and Francesco Ciari (2016) The IVT 2015 baseline scenario. In: *16th Swiss Transport Research Conference (STRC 2016)*. 16th Swiss Transport Research Conference (STRC 2016).

Bundesminister für Verkehr (1990) Richtlinien für den Lärmschutz an Straßen RLS-90.

Button, Kenneth (2004) The rationale for road pricing: standard theory and latest advances. In: *Road pricing: theory and evidence* 9, pp. 3–25.

Chakirov, Artem (2016) Urban mobility pricing with heterogeneous users. PhD thesis. ETH Zurich.

De Haan, Peter and Mario Keller (2004) Modelling fuel consumption and pollutant emissions based on real-world driving patterns: the HBEFA approach. In: *International journal of environment and pollution* 22 (3), pp. 240–258.

Ecoplan / Infras (2019) Externe Effekte des Verkehrs 2015. Aktualisierung der Berechnungen von Umwelt-, Unfall-und Gesundheitseffektendes Strassen-, Schienen-, Luft-und Schiffsverkehrs 2010 bis 2015. Technical Report. Bern.

Eliasson, Jonas, Lars Hultkrantz, Lena Nerhagen, and Lena Smidfelt Rosqvist (2009) The Stockholm congestion–charging trial 2006: Overview of effects. In: *Transportation Research Part A: Policy and Practice* 43 (3), pp. 240–250.

Faruqui, Ahmad, Sanem Sergici, and Ahmed Sharif (2010) The impact of informational feedback on energy consumption-A survey of the experimental evidence. In: *Energy* 35 (4), pp. 1598–1608.

Federal Office for Spatial Development (ARE) (2017) Bauzonenstatistik Schweiz.

Federal Office for Spatial Development (ARE) (2020) Externe Kosten und Nutzen des Verkehrs in der Schweiz. Strassen-, Schienen-, Luft- und Schiffsverkehr 2017. Technical Report. Bern: Federal Office for Spatial Development (ARE).

Federal Office for the Environment (FOEN) (2010) Pollutant Emissions from Road Transport, 1990 to 2035. Updated in 2010. Technical Report 1021. Bern, p. 129.

Federal Roads Office (FEDRO) (2017) Handbook NISTRA 2017. Technical Report. Bern: Federal Roads Office (FEDRO).

Fellendorf, Martin and Peter Vortisch (2000) Integrated modeling of transport demand, route choice, traffic flow and traffic emissions. In: *79th Annual Meeting of the Transportation Research Board*.

Forschungsgesellschaft für des Straßenverkehr (FGSV) (1997) Empfehlungen für Wirtschaftlichkeitsuntersuchungen an Straßen EWS–Entwurf.

Fujii, Satoshi and Ayako Taniguchi (2006) Determinants of the effectiveness of travel feedback programs-a review of communicative mobility management measures for changing travel behaviour in Japan. In: *Transport policy* 13 (5), pp. 339–348.

Graphhopper (2018) Graphhopper. In:

Haklay, Mordechai and Patrick Weber (2008) Openstreetmap: User-generated street maps. In: *IEEE Pervasive Computing* 7 (4), pp. 12–18.

Hörl, Sebastian (2020) Dynamic Demand Simulation for Automated Mobility on Demand. PhD thesis. ETH Zurich.

Hörl, Sebastian, Miloš Balać, and Kay W Axhausen (2019) Pairing discrete mode choice models and agent-based transport simulation with MATSim. In: *98th Annual Meeting of the Transportation Research Board*. Washington, DC.

Horni, Andreas, Kai Nagel, and Kay W Axhausen (2016) The multi-agent transport simulation MATSim. Ubiquity Press, London.

Hülsmann, Friederike, Regine Gerike, Benjamin Kickhöfer, Kai Nagel, and Raphael Luz (2011) Towards a multi-agent based modeling approach for air pollutants in urban regions. In: *Proceedings of the Conference on "Luftqualität an Straßen"*, pp. 144–166.

Janson, Michael and David Levinson (2014) HOT or not: Driver elasticity to price on the MnPASS HOT lanes. In: *Research in Transportation Economics* 44, pp. 21–32.

Kaddoura, Ihab (2015) Marginal Congestion Cost Pricing in a Multi-agent Simulation Investigation of the Greater Berlin Area. In: *Journal of Transport Economics and Policy (JTEP)* 49 (4), pp. 560–578.

Kaddoura, Ihab and Benjamin Kickhöfer (2014) Optimal road pricing: Towards an agent-based marginal social cost approach. Tech. rep. Working Paper 14-1, TU Berlin, Transport Systems Planning and Transport Telematics.

Kaddoura, Ihab, Lars Kröger, and Kai Nagel (2017) An activity-based and dynamic approach to calculate road traffic noise damages. In: *Transportation Research Part D: Transport and Environment* 54, pp. 335–347.

Keller, Mario, Stefan Hausberger, Claus Matzer, Philipp Wüthrich, and Benedikt Notter (2017) HBEFA Version 3.3. In: *Background documentation, Berne* 12.

Keller, Mario and Philipp Wüthrich (2016) Neuberechnung Staukosten Schweiz 2010-2014. Technical Report. MK Consulting and INFRAS.

Keller, Mario and Philipp Wüthrich (2019) Neuberechnung Staukosten Schweiz 2015-2017. Technical Report. MK Consulting and INFRAS.

Kickhöfer, Benjamin and Kai Nagel (2016) Towards high-resolution first-best air pollution tolls. In: *Networks and Spatial Economics* 16 (1), pp. 175–198.

Kraschl-Hirschmann, Karin, M Zallinger, Raphael Luz, Martin Fellendorf, and Stefan Hausberger (2011) A method for emission estimation for microscopic traffic flow simulation. In: *2011 IEEE Forum on Integrated and Sustainable Transportation Systems*. IEEE, pp. 300–305.

Kuehnel, Nico and Rolf Moeckel (2020) Impact of simulation-based traffic noise on rent prices. In: *Transportation Research Part D: Transport and Environment* 78, p. 102191.

Laih, Chen-Hsiu (1994) Queueing at a bottleneck with single-and multi-step tolls. In: *Transportation Research Part A: Policy and Practice* 28 (3), pp. 197–208.

Leape, Jonathan (2006) The London congestion charge. In: *Journal of Economic Perspectives* 20 (4), pp. 157–176.

Ma, Jing, Gordon Mitchell, and Alison Heppenstall (2015) Exploring transport carbon futures using population microsimulation and travel diaries: Beijing to 2030. In: *Transportation Research Part D: Transport and Environment* 37, pp. 108–122.

Maibach, Markus, C Schreyer, D Sutter, HP Van Essen, BH Boon, R Smokers, A Schroten, C Doll, B Pawlowska, and M Bak (2008) Handbook on estimation of external costs in the transport sector. In: *CE Delft*.

Molloy, Joseph, Alberto Castro Fernández, Thomas Götschi, Beaumont Schoeman, Christopher Tchervenkov, Uros Tomic, Beat Hintermann, and Kay W. Axhausen (2021) A national-scale mobility pricing experiment using GPS tracking and online surveys in Switzerland: Response rates and survey method results. In: *100th Annual Meeting of the Transportation Research Board*. Washington, DC.

MotionTag GmbH (2019a) Catch My Day. In: Apple App Store, Google Play.

MotionTag GmbH (2019b) MotionTag. In:

Newson, Paul and John Krumm (2009) Hidden Markov map matching through noise and sparseness. In: *Proceedings of the 17th ACM SIGSPATIAL international conference on advances in geographic information systems*. ACM, pp. 336–343.

Pigou, Arthur Cecil (1920) The economics of welfare.

Rexeis, Martin, Stefan Hausberger, Jörg Kühlwein, and Raphael Luz (2013) Update of Emission Factors for EURO 5 and EURO 6 vehicles for the HBEFA Version 3.2. Final Report. Graz: TU Graz.

Small, Kenneth A and Jia Yan (2001) The value of "value pricing" of roads: Second-best pricing and product differentiation. In: *Journal of Urban Economics* 49 (2), pp. 310–336.

Swiss Association for Road and Transportation Professionals (VSS) (2013) Cost Benefit Analysis for Road Traffic. Technical Report. Zurich: Swiss Association for Road and Transportation Professionals (VSS).

Taniguchi, Ayako, Fumihiro Hara, Shin'ei Takano, Sei'ichi Kagaya, and Satoshi Fujii (2003) Psychological and behavioral effects of travel feedback program for travel behavior modification. In: *Transportation Research Record: Journal of the Transportation Research Board* (1839), pp. 182–190.

Van Den Berg, Vincent and Erik T Verhoef (2011) Winning or losing from dynamic bottleneck congestion pricing?: The distributional effects of road pricing with heterogeneity in values of time and schedule delay. In: *Journal of Public Economics* 95 (7-8), pp. 983–992.

Van Essen, Huib, Lisanne van Wijngaarden, Arno Schroten, Daniel Sutter, Cuno Bieler, Silvia Maffii, Marco Brambilla, Davide Fiorello, Francesca Fermi, Riccardo Parolin, et al. (2019) Handbook on the external costs of transport, version 2019. 18.4 K83. 131.

Verhoef, Erik T (2000) The implementation of marginal external cost pricing in road transport. In: *Papers in Regional Science* 79 (3), pp. 307–332.

Verhoef, Erik T (2002) Second-best congestion pricing in general networks. Heuristic algorithms for finding second-best optimal toll levels and toll points. In: *Transportation Research Part B: Methodological* 36 (8), pp. 707–729.

Verhoef, Erik T and Kenneth A Small (2004) Product differentiation on roads. In: *Journal of Transport Economics and Policy (JTEP)* 38 (1), pp. 127–156.

Vickrey, William S (1963) Pricing in urban and suburban transport. In: *The American Economic Review* 53 (2), pp. 452–465.

Zheng, Yu (2015) Trajectory data mining: an overview. In: *ACM Transactions on Intelligent Systems and Technology (TIST)* 6 (3), pp. 1–41.

4

MIXL

FULL TITLE

mixl: An open-source R package for estimating complex choice models on large datasets

AUTHORSHIP

Joseph Molloy, Felix Becker, Basil Schmid and Kay W. Axhausen

This chapter is under review for publication in the *Journal of Choice Modelling*

4.1 ABSTRACT

This paper introduces mixl, a new R package for the estimation of advanced choice models. The estimation of such models typically relies on simulation methods with a large number of random draws to obtain stable results. mixl uses inherent properties of the log-likelihood problem structure to greatly reduce both the memory usage and runtime of the estimation procedure for specific types of mixed multinomial logit models. Functions for prediction and posterior analysis are included. Parallel computing is also supported, with near linear speedups observed on up to 24 cores. mixl is directly accessible from R, available on CRAN. We show that mixl is fast and easy to use, and scales to very large data sets. This paper presents the architecture and performance of the package, details its use, and presents some results using real world data and models.

4.2 INTRODUCTION

Choice modelling is an important tool in many fields, including transportation, marketing, psychology, and economics. For example in transportation, individuals must make many repeated decisions, whether to travel or not, which mode of transport to take, and which route or transit line to travel on. All of these decisions involve a choice of one alternative from multiple options: a discrete choice. Since the 1970's multinomial logit models have applied random utility theory to model how decision makers compare and evaluate alternatives (McFadden, 1974). However, for a closed-form solution some mathematical constraints of the decision need to be accepted. However, these restrictions, although computationally convenient, limit the realism of the models. As such, increasingly more complex models and model families have been proposed, including Mixed MNL (MMNL) (McFadden and Train, 2000), Generalized Extreme Value (McFadden, 1980) and Hybrid Choice Models (HCM) (Ben-Akiva, McFadden, Train, et al., 2002; Walker and Ben-Akiva, 2001). These models relax the behavioural restrictions on the classic MNL model, but in the process, require simulation methods to estimate.

In particular, Mixed MNL and hybrid choice models are increasingly popular tools used to explain human behaviour (e.g. Schmid and Axhausen, 2019). These modelling approaches incorporate random parameters and require simulation to estimate. Runtime generally increases linearly with the number of draws selected by the modeller. Additionally, the size of the datasets that researchers are working with are also becoming larger. Hence, it is now common that complex model formulations can take many hours or even days or weeks to estimate when using simulation methods. Cranenburgh and Bliemer (2019) developed the approach Sampling of Observations (SoO) to scale down choice datasets while producing similar results, noting that large models can be otherwise computationally infeasible. In particular they note that the size of the problem estimable is bound by the amount of memory required.

There are various software packages for estimating such models, including Biogeme (Bierlaire, 2018), ALOGIT (ALOGIT, 2016) and routines coded for using within Stata [1], MATLAB [2] [3], and Gauss[4]3. Ideally, the model estimation process integrates seamlessly with the workflow of the modeller, which is commonly the R language. The available R software packages for choice model estimation are limited either in functionality (mlogit, mnlogit) or by their reliance on the R language (Apollo, gmnl, RSGHB), making them both computationally slow and unable to handle large datasets when using simulation methods with a large number of draws. As such, there is a need for open-source R software that can simultaneously estimate complex choice models on large datasets, while taking advantage of modern multi-core machines to improvement performance.

This paper presents a new open-source R package, mixl, for the estimation of MNL, MMNL and HCM. All code is opensource and available at `https://github.com/joemolloy/fast-mixed-mnl`. The estimation process is fast, and scales well on multicore machines, with a 10x speedup with 24 processing cores when compared to a single core, to handle extremely large problems. mixl also avoids the memory limitations of other R packages. Furthermore, models can be estimated with two lines of code, and a straight-forward syntax. Through these contributions, mixl provides the tools needed by both those estimating extremely large models, and those new to choice modelling.

The paper proceeds as follows: In Section 4.3, the necessary background on log-likelihood computation and estimation, including the available software for doing so, is covered. Section 4.4 describes the software architecture and design decisions behind the mixl package. Section 4.5 explains how to use the package, Section 4.6 presents some of the additional features of the package that may be useful to the modeller, and Section 4.7 discusses the the performance of the package. Section 4.10 concludes.

1 `https://www.sheffield.ac.uk/economics/people/hole/stata`
2 `https://github.com/czaj/dce`
3 `https://eml.berkeley.edu/~train/software.html`

4.3 BACKGROUND

In this section a brief overview of the multinomial logit model, the log-likelihood calculation and estimation are presented. For more detail, the reader is referred to Train (2009) and Ben-Akiva and Lerman (1987). Let us assume that a decision maker n is trying to maximize his utility in the form $U_{njt} = V_{njt} + \epsilon_{njt}$ where V_{njt} is the observed utility in choice scenario t for alternative j, and ϵ_{njt} represents unobserved factors (Ben-Akiva and Lerman, 1987). This leads to the following succinct closed formed expression for the choice probability, where V_{njt} can be expressed as $\exp(X_{njt}\beta)$ (Train, 2009):

$$P_{nit} = \frac{\exp(X_{nit}\beta)}{\sum_j \exp(X_{njt}\beta)} \tag{4.1}$$

where n is the decision maker, i the chosen alternative in choice scenario t, and T_n is the number of choice tasks for individual n. The total number of alternatives for a person and situation is specified by $I_{t,n}$. The vector β represents the model parameters to be estimated, and X_{nit} the vector of observed variables. With a sample of N decision makers, the probability that the choice of person n is observed can be represented as

$$L(\beta) = \prod_n^N \prod_t^{T_n} \prod_i^{I_{t,n}} (P_{nit})^{y_{nit}} \tag{4.2}$$

where $y_{ni} = 1$ if person n chose i and zero otherwise. Computationally, it is beneficial to remove the product operators, giving the log-likelihood function, with the β that optimises this function.

$$LL(\beta) = \sum_n^N \sum_t^{T_n} \sum_i^{I_{t,n}} y_{nit} \log(P_{nit}) \tag{4.3}$$

In the mixed case, Equation 4.2 is extended to include the distribution of the parameters, where $f(\beta)$ is the density function, and θ the parameters that describe the density of β:

$$L_n = \int \prod_t^{T_n} P_{nit} f(\beta|\theta) d\beta \tag{4.4}$$

These mixed models can be formulated in several ways, with two main derivations using random coefficients and/or error components. They vary over decision makers with the density $f(\beta)$, a function of the parameters θ. As such, in a mixed model, the parameters β vary over the population.

The calculation of the likelihood for a simple MNL model is straightforward. The likelihood is simply the product of the chosen probabilities for each individual. For panel data, the logs of the probabilities are summed up over all observations for each individual. To calculate the probabilities in a mixed model, R values β^r are drawn from $f(\beta|\theta)$ and used to calculate the likelihood. In equation 4.1, β is hence replaced by β_r to calculate P^r_{nit}. The simulated probability $\hat{P}_n$ for individual n is given by the average over all R draws, and the simulated log-likelihood ($\hat{LL}$) for the sample follows:

$$\hat{P}_n = \frac{1}{R}\sum_r^R \prod_t^{T_n} \prod_i (P^r_{nit})^{y_{nit}} \tag{4.5}$$

$$\hat{LL} = \sum_n \ln\left(\hat{P}_n\right) \tag{4.6}$$

The summation in Equation 4.5 requires that $\prod_t^{T_n} P^r_{nit}$ be calculated R times, and hence, the execution time increases linearly with the number of draws.

4.3.1 *Maximum likelihood estimation*

The process of estimating the optimal value of the parameters β is called maximum likelihood estimation (MLE), or maximum simulated likelihood estimation (MSLE). In this process, an optimization routine tries to find the set of parameters that give the maximum log-likelihood by repeatedly calculating the log-likelihood with different β until the process converges. The speed of convergence relies on repeated improved guesses of β by determining the gradient of the function. The most common approach, when an analytical gradient is not available is to calculate a numerical gradient f' of f with respect to the vector β:

$$f'(\beta) = \frac{f(\beta - \Delta) + f(\beta + \Delta)}{2\Delta} \tag{4.7}$$

To calculate the gradient at each optimization step, the objective function, namely the log-likelihood, needs to be calculated $2k + 1$ times, where k is the number of free parameters to be estimated and Δ is a suitably small number to give an estimate of the derivative for the vector β'. A method of gradient descent is then used to find a minimum (or maximum) of the function. Two of the most popular approaches are the Broyden–Fletcher–Goldfarb–Shanno (BFGS) algorithm and Newton-Raphson method. However, they are not guaranteed to converge for non-linear problems, of which MMNL is an example. As such the selection of the initial start values (the initial β) can influence the final log-likelihood and parameter estimates. One benefit of the BFGS method is that, as a Quasi-Newton method, the Hessian matrix of the second derivative does not need to be computed during estimation, a very computationally expensive process.

Analytical gradients are implemented in many packages, such as gmnl, Apollo (version 0.2.0 onwards) and Biogeme. With analytical gradients, $f'(\beta)$ can be calculated in one step, instead of with two evaluations of $f(\beta)$. Since the evaluation of $f'(\beta)$ is the most common operation in the MLE, this leads to a rougly 2x speedup over numerical derivation, all other things being equal.

4.3.2 *Overview of available choice modelling software*

There is a wide variety of software available for choice model estimation, including but not limited to *Biogeme* (Bierlaire, 2018), the *mlogit* (Croissant, 2015) and *mnlogit* (Hasan, Zhiyu, and Mahani, 2014) packages for R, *ALOGIT* (ALOGIT, 2016), and *NLOGIT* (Greene, 2002). Recently, a new R package for choice modelling, Apollo (Hess and Palma, 2019a; Hess and Palma, 2019b) has also been published. Each of these packages handle various model types and mixed models. Table 4.1 summarizes the attributes of the packages.

The two most well known proprietary packages which are designed spefically for choice modelling are ALOGIT and NLOGIT, are well established, but are neither open-source, nor freely available to researchers.

There is also routines available in most statistical software packages, such as MATLAB, STATA and Gauss. In particular, recently released routines for MATLAB have been shown to be 5-10x faster than previously available R Code for choice modelling Czajkowski and Budziński (2019).

Arguably the two leading open source softwares for discrete choice modeling are Biogeme and Apollo. Both estimate an extremely wide range of parametric models. Biogeme has been under development for many years and is a robust and stable software, now with the latest version, pandasBiogeme directly usable from python. In this paper it will be referred to just as Biogeme. The user can specify arbitrary utility functions and the likelihood formulation. Additionally, it takes advantage of both compilation to C++ and automatic derivation to achieve excellent performance. As a package for the python programming language, it is not directly usable from R.

In R, the mlogit and mnlogit packages provide the most accessible tools for working with MNL models. They support basic mixed models, but don't provide the syntax to specify more complicated models. They also struggle on larger models. The mnlogit package provides significant speed improvements over mlogit by optimizing the calculation of the Hessian matrix and using the Newton-Raphson method for MSLE. However, Nocedal and Wright (2000) observed that quasi-Newton methods such as BFGS perform better on larger MSLE problems. Both these R packages rely on the R *formula* package and syntax for specifying the utility function. Also, mnlogit does not support random coefficients. Both mlogit and mnlogit packages do not support hybrid choice models or other more advanced model formulations. More recently, the gmnl (Sarrias, Daziano, et al., 2017), RSGHB (Dumont, Keller, and Carpenter, 2015) and Apollo Hess and Palma, 2019b have been published. They support more advanced model specifications, including Latent Class models. Of these, only Apollo supports parallel execution on multiple cores. However, as shown in Section 4.8 there has some performance limitations with Apollo when a large number of random draws are used in combination with a large dataset.

	Open-source	R	Mixed models	HCM	Large problems
Apollo	yes	yes	yes	yes	no
gmnl	yes	yes	yes	no	no
RSGHB	yes	yes	yes	yes	no
mlogit	yes	yes	yes	no	no
mnlogit	yes	yes	no	no	yes
Biogeme	yes	no	yes	yes	yes
ALOGIT	no	no	yes	no	yes
NLOGIT	no	no	yes	no	yes
Stata	no	no	yes	yes	no
Gauss	no	no	yes	yes	no
MATLAB	no	no	yes	yes	yes

Table 4.1: Comparison of main software packages for multinomial logit modelling

4.3.3 *Potential and limitations of the R language*

The open-source statistical software *R* is an incredibly powerful and popular platform for data processing, analysis and visualization. However, it is well known for its liberal use of memory. A particular performance bottleneck in R is iteration. *for* loops have a significant overhead in the R language. As such, for many common operations, functions available in R and its packages are written in the programming language C++ for better performance, or 'vectorized' to work on vectors or matrices to avoid R-based iteration.

The *Rcpp* package (Eddelbuettel and Francois, 2011) is most commonly used to improve the performance of R scripts by rewriting critical functions in C++. Functions written in Rcpp accept and return R datatypes such as Vectors and Matrices. Code written in C++ and called from R is often many times faster that the equivalent R code. However, C++ code must be compiled before execution, either when the package is created, or inside the script itself. The process of writing C++ code is also more involved, requiring the developer to generally make a tradeoff between coding effort and execution time when switching from R to C++.

4.4 SOFTWARE ARCHITECTURE

In section 4.3.1 it was noted that while the utility function must be calculated many times during the MSLE process, all the data used by the function except the parameters to be estimated do not change. Furthermore, each observation can be seen as independent from a calculation perspective. For every observation, the utility of each alternative is calculated. From there the log of the probability of the chosen alternative is simple to calculate, which are then summed over each individual for repeated observations. Since this operation is associative, the order in which the observations are processed is not important.

This is exploited to reduce the memory required during estimation. Normally, the log-probability for every single observation is calculated, resulting in a matrix of size $(N * T)$ x R where N is the number of individuals, T the number of observations per individual, and R is the number of draws used. The resulting matrix is then summed over the individuals. Instead, a running logsum of the probabilities over the observations for each individual is kept, requiring a smaller a matrix P of size N x R. For datasets with a large panel structure, and models with a large number of draws, this saves a significant amount of memory - a factor equal to the average number of observations per individual.

In R, this requires using iteration constructs rather than the optimised vectorized linear algebra operations, as the dataset and draws matrix are of different sizes. As mentioned in Section 4.3.3, this results in an unacceptable performance bottleneck. The solution is to code the log-likelihood function in C++. It then follows that $f(\beta)$ must also be written and compiled in C++.

To compile the loglikihood function to C++, a pre-compiler is implemented in mixl to takes a model specification written in plain text, and convert it to a C++ objective function callable from R and optimisation routines. Section 4.5 details how model specifications need to be written. The pre-compiler validates the specification against the dataset to check that all variables are present, and automatically identifies model properties such as mixed effects and hybrid choice components. mixl detects errors in the model specification and reports them to the user.

Since P_{nit} is calculated for each observation t separately, the calculation of the log-likelihood is an *embarrassingly parallel problem* which can be handled efficiently using data parallelism. While there are packages to perform data parallelism in R, for example *parallel* and *foreach*, they require significant communication overheads as new processes are spawned. Since in mixl the log-likelihood function is implemented in C++, the openMP (Chapman and Massaioli, 2005) framework is used to efficiently parallelize the for-loop over observations. Since all data except the intermediate utilities of each alternative are shared between all cores, no copying of the data across cores is needed to run the log-likelihood function in parallel. Furthermore, these intermediate data structures are generated once for each estimation, meaning that new memory does not need to be allocated for each likelihood computation. Compiling with the openMP framework even provides a performance boost on a single core, due to certain optimizations the framework enables in the compiler.

4.5 Using the mixl package

We start by implementing the model denoted in Equations 4.8 to 4.10. Equations 4.8 and 4.9 show the utility functions of the alternatives public transport and car for every decision maker n for every choice situation t. In addition to the price, the travel time, and the number of changes (or transfers), a random ASC for public transport is included in the model.

$$V_{PT,n,t} = ASC_{PT} + \psi_{PT,n} + \beta_{price} * x_{price,PT,n,t} + \beta_{time,PT} * x_{time,PT,n,t}/60 + \beta_{change} * x_{change,PT,n,t} \tag{4.8}$$

$$V_{Car,n,t} = ASC_{Car} + \beta_{price} * x_{price,Car,n,t} + \beta_{time,Car} * x_{time,Car,n,t}/60 \tag{4.9}$$

$$\psi_{PT} \sim \mathcal{N}(0, \sigma^2_{PT}) \tag{4.10}$$

For the implementation in mixl, the modeller describes the utility functions in terms of mathematical relationships between variables in the data (prefixed with a $ sign), and parameters (prefixed with an @ symbol):

LISTING 4.1: Example of mixl syntax.

```
1 ASC_PT_RND = @ASC_PT + draw_1 * @SIGMA_PT;
2
3 U_pt = ASC_PT_RND + @B_price * $price_PT + @B_timePT * $time_PT / 60
4       + @B_change * $change_PT;
5
6 U_car = @ASC_Car + @B_price * $price_Car + @B_timeCar * $time_Car / 60;
```

In listing 4.1, an example of the syntax is provided. All words prefixed with @ (ASC_PT, SIGMA_PT, ...) are parameters to be estimated. Those with $ are variables of the observations available in the data. In this example one random parameter is required, indicated by the draw_ variable. One intermediate variable (line 1) is also calculated, ASC_PT_RND, which is then used in the utility function of public transport. The _RND suffix indicates that this is a random coefficient, for which posteriors can be automatically calculated using the posteriors function. The availabilities must be supplied as a separate matrix, with one row for each observation, and one column for each alternative used in the model specification.

To aid both the specification of models and improve error reporting to the modeller, a small amount of syntax is required. This is covered in more detail in the user-guide.

- Variables from the dataset must be prefixed with a $
- Coefficients to be prefixed with a @
- Every statement must end with a ;
- Intermediate variables that are calculated don't require a prefix
- The utility functions are prefixed by U_xxx, where x is the choice id. "U_1" or "U_car" are, for example, valid. The order in which they are specified must correlate with the numbering of the choices in CHOICE

- Draws are prefixed by *draw*. The same naming rules as for the utility functions apply. If the nDraws parameter is passed into the `specify_model` function, a set of draws will be generated automatically. Currently, this defaults to a halton sequence. Alternatively, any set of draws (Sobol (Sobol, 1967), MLHS (Hess, Train, and Polak, 2006), etc) can be passed in as an argument, as long as the matrix is large enough to accommodate the number of individuals and random parameters.

- Standard mathematical functions such as addition, multiplication, exponentiation, and equality comparisons are allowed with the appropriate operator.

4.5.1 *Iterative development of a model with mixl*

In this section we present the iterative development of a model, starting with a basic MNL model, followed by a mixed MNL model. The dataset is the 'Electricity' dataset available in the *mlogit* package, which Train uses for MMNL examples in Train (2009). This provides a good example of how the successively more advanced models can be iteratively developed using the mixl package.

The utility functions of the basic MNL model are denoted in Equation 4.12. The different alternatives represent different heating options and are differentiated by price (x_{pf}), whether the supplier is local (x_{loc}) or well known (x_{wk}), and whether the supplier offers time-of-day (x_{tod}) or seasonal rates (x_{seas}). As the experiment is unlabelled, generic parameters are estimated.

$$V_j = \beta_{pf} * x_{pf,j} + \beta_{cl} * x_{cl,j} + \beta_{loc} * x_{loc,j} + \beta_{wk} * x_{wk,j} + \beta_{tod} * x_{tod,j} + \tag{4.11}$$

$$\beta_{seas} * x_{seas,j}, j \in \{1, ..., 4\} \tag{4.12}$$

Using the 'Electricity' dataset from the *mlogit* package, it is straightforward to set up the data for our model in lines 2-4 in listing 4.2. Only the `ID` and `CHOICE` variables need to be converted to continuous values starting from 1. We then specify the above model as a string of text in R as follows (6-11). Variables in the dataset are prefixed with $ and β to be estimated with @. We then call `specify_model` to convert this model specification to a log-likelihood function. The dataset is passed in so that the variables in the model can be verified.

The starting values are specified on line 15 and the availabilities on line 18. Here we use a default function with two parameters: the dataset and the number of alternatives . The `estimate` function is then called (line 20) on the model specification, and the results presented to the user using the `summary` command on line 21.

To specify a model alternatives are available, then a matrix of n columns and a rows is required, where n is the number of total choice observations, and a the number of utility functions. A value of 1 indicates that that alternative is available for that observation.

Listing 4.2: A simple MNL model on the Electricity dataset.

```
1  library(mixl)
2  data("Electricity", package = "mlogit")
3  Electricity$ID <- Electricity$id
4  Electricity$CHOICE <- as.numeric(Electricity$choice)
5
6  mnl_test <- ‘
7     U_1 = @pf * $pf1 + @cl * $cl1 + @loc * $loc1 + @wk * $wk1 + @tod * $tod1 +
        @seas * $seas1;
8     U_2 = @pf * $pf2 + @cl * $cl2 + @loc * $loc2 + @wk * $wk2 + @tod * $tod2 +
        @seas * $seas2;
9     U_3 = @pf * $pf3 + @cl * $cl3 + @loc * $loc3 + @wk * $wk3 + @tod * $tod3 +
        @seas * $seas3;
10    U_4 = @pf * $pf4 + @cl * $cl4 + @loc * $loc4 + @wk * $wk4 + @tod * $tod4 +
        @seas * $seas4;
11 ’
12
13 model_spec <- specify_model(mnl_test, Electricity)
14
15 est <- stats::setNames(c(0,0,0,0,0,0),
16    c("pf", "cl", "loc", "wk", "tod", "seas"))
17
18 availabilities <- mixl::generate_default_availabilities(Electricity, 4)
19
20 model <- estimate(model_spec, est, Electricity, availabilities)
21 summary(model)
```

For estimation, mixl wraps the maximum likelihood optimisation routine from the *maxLik* package (Henningsen and Toomet, 2011). As Train (2009) suggests, the BFGS (Witzgall and Fletcher, 1989) optimization procedure is used as default. The interface is designed so that all possible parameters to *maxLik* can be passed through, including the choice of optimisation routine, Hessian function, and a limit on the number of iterations. The fixing of parameter values is also supported. The robust standard errors are caculated using the sandwich package (Zeileis, 2006; Zeileis, Köll, and Graham, 2020).

The output from a estimated model is presented in the console as follows:

Listing 4.3: Sample output from mixl for listing 4.2.

```
Model diagnosis: successful convergence

Number of decision makers: 361
Number of observations: 4308

LL(null):  -5972.156
LL(init):  -5972.156
LL(final):  -4958.649
Rho2:  0.1697054

AIC:  9929.3
AICc:  9929.54
BIC:  9967.51
Estimated parameters:  6

Estimates:
         est     se trat_0  trat_1  robse robtrat_0 robtrat_1 rob_pval0
   rob_pval1
pf   -0.6253 0.0232 -26.93  -69.99 0.0334    -18.70    -48.60         0
   0.00
cl   -0.1083 0.0082 -13.13 -134.43 0.0140     -7.74    -79.18         0
   0.00
loc   1.4421 0.0506  28.52    8.74 0.0788     18.31      5.61         0
   0.00
wk    0.9954 0.0448  22.23   -0.10 0.0638     15.61     -0.07         0
   0.94
tod  -5.4636 0.1837 -29.74  -35.18 0.2778    -19.67    -23.27         0
   0.00
seas -5.8408 0.1867 -31.29  -36.64 0.2723    -21.45    -25.12         0
   0.00
```

In the listing 4.3, for each estimated coefficient, multiple values are provided. `est` is the estimated value, `se` the standard error. `trat` is short for the t-ratio, and pval for the p-Value. The `rob` prefix indicates the robust estimates. The _1 values are to be considered when testing against the null hypothesis that the parameter equals one (for example, a scale parameter).

In order to illustrate how random parameters are incorporated into the model, a new model is specified for the Electricity data set, see Equations 4.13 to 4.19. Equation 4.13 shows the utility function for every alternative j, every individual n, and every choice situation t. Unlike in the previous standard MNL model, all parameters follow the normal distribution, see Equations 4.14 to 4.19.

$$V_{j,n,t} = \beta_{pf,rnd,n} * x_{pf,j,n,t} + \beta_{cl,rnd,n} * x_{cl,j,n,t} + \beta_{loc,rnd,n} * x_{loc,j,n,t} + \beta_{wk,rnd,n} * x_{wk,j,n,t} + \beta_{tod,rnd,n} * x_{tod,j,n,t} + \beta_{seas,rnd,n} * x_{seas,j,n,t}, j \in \{1, ..., 4\} \tag{4.13}$$

$$\beta_{pf,rnd} \sim \mathcal{N}(\beta_{pf}, \sigma^2_{pf}) \tag{4.14}$$

$$\beta_{cl,rnd} \sim \mathcal{N}(\beta_{cl}, \sigma^2_{cl}) \tag{4.15}$$

$$\beta_{loc,rnd} \sim \mathcal{N}(\beta_{loc}, \sigma^2_{loc}) \tag{4.16}$$

$$\beta_{wk,rnd} \sim \mathcal{N}(\beta_{wk}, \sigma^2_{wk}) \tag{4.17}$$

$$\beta_{tod,rnd} \sim \mathcal{N}(\beta_{tod}, \sigma^2_{tod}) \tag{4.18}$$

$$\beta_{seas,rnd} \sim \mathcal{N}(\beta_{seas}, \sigma^2_{seas}) \tag{4.19}$$

The associated mixl code is presented in listing 4.4 below. As can be inferred, random draws are prefixed with `draw_`. By default, all parameters follow the normal distribution if included in the way below. Naming the random parameters with _RND is only necessary if one wishes to estimate posteriors using the post-estimation functions (see Section 4.6.2.

Listing 4.4: A Mixed MNL model in mixl syntax.

```
mnl_test <- '
   pf_RND = @pf + draw1 * @sigma_pf;
   cl_RND = @cl + draw2 * @sigma_cl;
   loc_RND = @loc + draw3 * @sigma_loc;
   wk_RND = @wk + draw4 * @sigma_wk;
   tod_RND = @tod + draw5 * @sigma_tod;
   seas_RND = @seas + draw6 * @sigma_seas;

   U_1 = pf_RND * $pf1 + cl_RND * $cl1 + loc_RND * $loc1 +
         wk_RND * $wk1 + tod_RND * $tod1 + seas_RND * $seas1;

   U_2 = pf_RND * $pf2 + cl_RND * $cl2 + loc_RND * $loc2 +
         wk_RND * $wk2 + tod_RND * $tod2 + seas_RND * $seas2;

   U_3 = pf_RND * $pf3 + cl_RND * $cl3 + loc_RND * $loc3 +
         wk_RND * $wk3 + tod_RND * $tod3 + seas_RND * $seas3;

   U_4 = pf_RND * $pf4 + cl_RND * $cl4 + loc_RND * $loc4 +
         wk_RND * $wk4 + tod_RND * $tod4 + seas_RND * $seas4;
'
```

The same code as before is used to estimate the model with the addition of the required number of Halton draws (in this example 100) which has to be indicated in the estimate function. By including a random alternative specific constant, the log-likelihood improves by almost 800 units and doubles the McFadden R^2. Table 4.2 shows the results of the MMNL model, available as a function provided in mixl using the *texreg* package.

Unlike in other software such as Apollo and Biogeme, the likelihood function cannot be modified by the user. For a large range of problems, however, it is sufficient to encode the behaviour in the utility functions, and use a standard log-likelihood function for panel data as described in Section 4.3.

4.6 FURTHER FEATURES

4.6.1 *Estimation of hybrid choice models*

The package also supports hybrid choice models or to be more exact integrated choice and latent variable models. The current example uses a continuous measurement equation (Linear Regression) for the latent variable, which can be easily extended to any (discrete) specification (e.g. Ordered Logit). Due to the size of the actual code input, we want to focus on the crucial parts. The relevant model components are presented in Equations 4.20 to 4.24.

$$I = \bar{i} + x^{lv}\zeta_I^{lv} + v \tag{4.20}$$

$$v_r \sim \mathcal{N}(0, \sigma_{I,r}^2),\ r \in \{1,2\} \tag{4.21}$$

$$x^{lv} = \beta_{Age}^{LV} * x_{Age} + \beta_{Inc}^{LV} * x_{Inc} + \psi_{LV} \tag{4.22}$$

$$\psi_{LV} \sim \mathcal{N}(0, \sigma_{lv,struct}{}^2) \tag{4.23}$$

$$V_1 = ASC_1 + \psi_1 + x^{lv}\beta_{LV,u1} \tag{4.24}$$

$$\psi_1 \sim \mathcal{N}(0, \sigma_1^2) \tag{4.25}$$

Equation 4.20 denotes the measurement equation for the latent variable, where $\bar{i}$ denotes the mean of this indicator. The indicator is treated continuously. The structural equation of the latent variable is shown in Equation 4.22. Both Age and Income are used to explain the latent variable. The utility function of the choice model, which includes the latent variable as an explanatory variable, is presented in Equation 4.25.

	Model 1
pf	-0.86^{***}
	(0.04)
cl	-0.21^{***}
	(0.02)
loc	1.90^{***}
	(0.11)
wk	1.38^{***}
	(0.08)
tod	-7.92^{***}
	(0.40)
seas	-8.14^{***}
	(0.37)
sigma_pf	0.18^{***}
	(0.02)
sigma_cl	0.32^{***}
	(0.03)
sigma_loc	1.21^{***}
	(0.12)
sigma_wk	0.30^{*}
	(0.17)
sigma_tod	2.01^{***}
	(0.24)
sigma_seas	0.89^{***}
	(0.11)
# estimated parameters	12.00
Number of respondents	361.00
Number of choice observations	4308.00
Number of draws	20.00
LL(null)	-5972.16
LL(final)	-4108.61
McFadden R2	0.31
AIC	8241.23
AICc	8242.13
BIC	8317.65

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

TABLE 4.2: Latex model output from a mixed MNL model estimated with mixl

First, in line 1 of the subsequent code snippet, we define the structural equation of the latent variable (LV). In lines 2 and 3, we specify the measurement equations. While `$I_1` relates to the actual indicator in the data, `$I_1_m` refers to its pre-computed sample mean. For reasons of identification, we set the parameter of the first indicator to 1. For the second indicator, we estimate the parameter `@LV_ZETA_I_2`. `@LV_SIGMA_I_1` and `@LV_SIGMA_I_2` refer to the variances of the error of the respective linear regressions. In line 4, the latent variable is included in utility function `U_1`.

```
LV = @LV_AGE * $Age + @LV_Inc * $income + @LV_SIGMA_Struct * draw_2;
P_indic_1 = R::dnorm($I_1 - $I_1_m - 1 * LV,0,(@LV_SIGMA_I_1),0);
P_indic_2 = R::dnorm($I_2 - $I_2_m - @LV_ZETA_I_2 * LV,0,(@LV_SIGMA_I_2),0);
U_1 = @ASC_1  + draw_1 * @SIGMA_1 + @LV_U1 * LV;
```

If models include variable definitions with the prefix `P_indic_`, the model is assumed to have hybrid components, and the `P_indic_` variables will be considered as probability indicators for each observation. `P_indic_` should be defined from 1 to k, the number of indicators to be used in the model. The pre-compiler detects these automatically, and generates the code to include the product of the probability indicators in the log-likelihood as such:

```
p_choice = log(chosen_utility / sum(utilities));
p_indic_total = P_indic_1 * P_indic_2 * .... P_indic_k;
p_choice = p_choice + (1/count) * log(p_indic_total);
```

The count variable, used to normalize the choice indicator, represents number of choice observations per individual and must also be included in the data. On convergence, the model estimation will return both the choice log-likelihood and the model log-likelihood. One extra column is required in the dataset to enable hybrid choice, namely a 'count' column with the total number of observations for the individual making the choice.

4.6.2 *Post processing*

The mixl package provides some key post processing functions for working with an estimated model. The estimation results include all the expected components, such as the (robust) co-variance matrix, table of coefficients, standard errors, Hessian matrix, etc. The following functions are provided:

The `posterior` function allows the model to calculate the posteriors for models with mixed distributions. Random variables in the model specification are automatically detected (those with an equation ending by `_RND` and including draws), and the posterior function returns a labeled matrix of the posteriors for each individual and random variable. This can all be done inside the R language, as the results of th estimation are always returned as either matrices or dataframes.

The `probabilities` function is useful for calculating e.g. elasticities, since it calculates the probabilities for each alternative given the estimated parameters. Different scenarios can be easily evaluated by changing specific values (e.g. by 1%) of the input dataset.

`summary_tex` outputs the model results in latex syntax for insertion into a paper.

4.7 PERFORMANCE AND MULTICORE SCALABILITY

With the widespread availability of multi-core machines and computing clusters, parallel scalability is an important consideration. mixl achieves consistent speedups even on large numbers of cores. Figure 4.1 and Table 4.3 show the performance of the isolated log-likelihood function for different numbers of draws over an increasing number of cores. To demonstrate the speed up, a pooled RP/SP MMNL model on a large panel dataset with 491 individuals and 17,120 observations is used (Schmid, Balac, and Axhausen, 2019). The model contains 27 free parameters, 15 linear-additive utility functions related to four different data/experiment types (mode choice RP, mode choice SP, route choice public transport SP and route choice carsharing SP) and 12 random parameters (six for the ASCs and another six for the mode-specific travel times). The utility functions are a simplified version of the ones presented in Schmid (2019), Chapter 4, including all level-of-service attributes without socioeconomic effects and accounting for scale effects between the data/experiment types.

Each timing was repeated 50 times to obtain an average. For the 10 draw configuration, communication costs dominate and a maximum speedup of 4.78x is observed. As the number of draws used is increased, so do the benefits of using parallel computing. On 24 processing cores for 10,000 draws, a speedup of 19.3x is observed. It is worth noting that only the utility calculation has been parallelized, and potential performance improvements remain in other parts of the log-likelihood function in future work. The results in Table 4.3 show a super-linear speedup in the 4 core case. This can be attributed to cache-effects on the processor.

It is worth noting that the speedups obtained in Table 4.3 will not necessarily replicated on real problems, but rather show the upper bound of obtainable performance. This is due to Amdahl's law, which defines how sequential parts of the program limit the potential gains obtainable with parallel computing. Again, larger problem sizes extract better relative performance from more cores. More complicated utility functions, i.e. with more alternatives or parameters, will also see a benefit from increased parallelism.

	Processors					
Draws	1	2	4	8	16	24
10	1	1.56	2.81	2.73	3.51	4.78
100	1	1.92	3.83	5.09	6.96	14.76
500	1	1.94	4.04	5.91	9.09	17.40
1000	1	1.95	4.04	6.16	9.60	18.49
5000	1	1.96	4.11	6.28	10.22	19.58
10,000	1	1.97	4.21	6.31	10.25	19.35

Table 4.3: Speedup over multiple cores for the log-likelihood calculation

4.7.1 *Further example using the Grapes dataset*

In the following example, we estimation of an MMNL model based on a modified version of the Grapes dataset (Ben-Akiva, McFadden, Train, et al., 2019). Table 4.4 presents the attribute values in the dataset. All binary attributes were sampled from independent uniform distributions. There are three grape alternatives that participants can choose from, in addition to the possibility of not choosing any of them. Datasets were generated for 1000, 4000, and 16,000 individuals with 8 observations per respondent.

The model specification is shown in Equations 4.26 to 4.31. The utility for alternatives 1,2, and 3 is presented in Equation 4.26. n denotes the individual decision maker and t the choice task. Four random parameters are included, i.e. one for each variable. As can be inferred from Equations 4.28 to 4.31, all parameters follow the univariate normal distribution. The opt-out alternative is depicted in Equation 4.27.

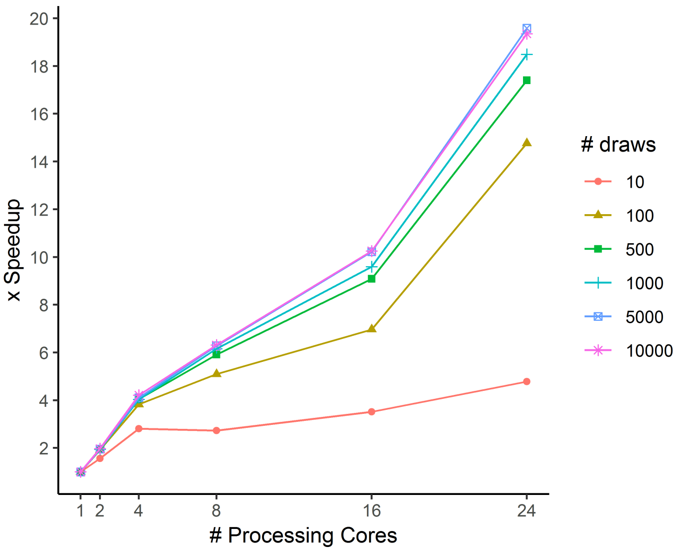


FIGURE 4.1: Performance of the inner log-likelihood function over multiple cores

$$V_{j,n,t} = \beta_{S,V,n} * x_{S,j,n} + \beta_{C,V,n} * x_{c,j,n,t} + \beta_{L,V,n} * x_{l,j,n,t} + \beta_{O,V,n} * x_{o,j,n,t}, j \in \{1, ..., 3\} \tag{4.26}$$

$$V_4 = 0 \tag{4.27}$$

$$\beta_{S,V} \sim \mathcal{N}(\beta_S, \sigma_S^2) \tag{4.28}$$

$$\beta_{C,V} \sim \mathcal{N}(\beta_C, \sigma_C^2) \tag{4.29}$$

$$\beta_{L,V} \sim \mathcal{N}(\beta_L, \sigma_L^2) \tag{4.30}$$

$$\beta_{O,V} \sim \mathcal{N}(\beta_O, \sigma_O^2) \tag{4.31}$$

The true underlying parameters, are reported in Table 4.5. The model is specified in mixl syntax in Listing 4.5.

Attribute	Symbol	Levels
Sweetness	S	Sweet (1) or Tart (0)
Crispness	C	Crisp (1) or Soft (0)
Size	L	Large (1) or Small (0)
Organic	O	Organic (1) or Non-organic (0)

Table 4.4: Grape dataset attributes and levels (Ben-Akiva, McFadden, Train, et al., 2019)

Parameter	Mean	Standard Deviation
B_S	1.00	0.40
B_C	0.90	0.30
B_L	2.50	1.00
B_O	1.50	0.50

Table 4.5: Parameters of the Grapes choice model.

Listing 4.5: MMNL model using the Grapes dataset

```
B_S_V = @B_S + @B_S_S * draws_1;
B_C_V = @B_C + @B_C_S * draws_2;
B_L_V = @B_L + @B_L_S * draws_3;
B_O_V = @B_O + @B_O_S * draws_4;

U_1 = (B_S_V * $S_1 + B_C_V * $C_1 + B_L_V * $L_1 + B_O_V * $O_1);
U_2 = (B_S_V * $S_2 + B_C_V * $C_2 + B_L_V * $L_2 + B_O_V * $O_2);
U_3 = (B_S_V * $S_3 + B_C_V * $C_3 + B_L_V * $L_3 + B_O_V * $O_3);
U_4 = 0;
```

Table 4.6 illustrates the runtime for the estimation using MIXL of an MMNL model based on a modified version of the Grapes dataset [5] (Ben-Akiva, McFadden, Train, et al., 2019). Using 24 cores can reduce the execution time by nearly 90%. Furthermore, the speedup is almost linear, implying that further reductions in execution time are possible if more cores are used.

5 The datasets are available on request should the creators of other packages wish to benchmark their software.

	Processors						
Draws	1	2	4	8	12	16	24
1	00:00:48	00:00:18	00:00:15	00:00:12	00:00:20	00:00:17	00:00:09
		(2.55)	(3.06)	(3.83)	(2.31)	(2.79)	(4.96)
100	00:20:11	00:10:25	00:06:25	00:03:56	00:03:02	00:02:54	00:02:21
		(1.94)	(3.14)	(5.13)	(6.63)	(6.95)	(8.58)
500	01:54:33	01:02:51	00:32:53	00:21:00	00:13:29	00:15:00	00:09:47
		(1.82)	(3.48)	(5.46)	(8.49)	(7.64)	(11.7)
1000	04:07:52	02:14:21	01:12:57	00:50:42	00:37:46	00:24:03	00:16:10
		(1.84)	(3.4)	(4.89)	(6.56)	(10.3)	(15.32)
5000	18:39:44	09:57:14	05:36:15	03:17:41	02:31:20	01:44:46	01:26:45
		(1.87)	(3.33)	(5.66)	(7.4)	(10.69)	(12.91)
10,000	60:16:29	35:00:28	17:58:11	10:33:33	08:57:10	06:38:54	05:27:31
		(1.72)	(3.35)	(5.71)	(6.73)	(9.07)	(11.04)

Table 4.6: mixl performance on a large dataset of 128,000 choice observations - runtime in hh:mm:ss, speedup in brackets

4.8 comparsions to other open-source software

The performance of the mixl (v 1.1.3) is compared against Apollo (v0.2.0) and pandasBiogeme (3.2.5), using the same grapes dataset and model specified above in listing 4.5. The demonstrated difference between the performance of Biogeme, mixl and Apollo is primarily due to the inclusion or exclusion of two optimisations: compilation of the loglikelihood function to C++, and the implementation of symbolic derivation. In version 0.2.0 of Apollo, analytical gradients have been implemented, giving a 2x speedup over version 0.1.0 on this model and dataset. Section 4.3 shows why this is the case. However, its reliance on the R language for the utility functions means that it is still slower than mixl. Biogeme implements both symbolic derivation to generate analytical gradients, and compilation to C++, hence it has the best performance. Table 4.7 shows the performance of the compared programs with respect to the number of processors used, taking the average of different respondent samples (1000, 4000, 16000) and numbers of random draws (10, 100, 500, 1000, 5000, 10,000). With more processing cores, mixl is as fast as Biogeme, and up to 3.5x faster than Apollo.

	Processors						
Program	1	2	4	8	12	16	24
Mixl 1.1.3	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Apollo 0.2.0	0.87	0.73	0.76	0.51	0.50	0.52	0.36
Biogeme 3.2.5	2.16	1.66	1.50	1.35	1.34	1.30	1.08

Table 4.7: Speedup relative to mixl by number of processors(higher values are faster)

The same performance tests were also run for GMNL, which implements analytical gradients, as Apollo and Biogeme do. On the smallest example (1000 respondents and 1000 draws), GMNL had a similar runtime to mixl. However, on a middle size problem with 16,000 respondents and 1000 draws, the code was still running without result after 72 hours. For comparison, Biogeme, mixl and Apollo return results in 2 hours, 4 hours and 5 hours respectively, when using one processing core.

Figure 4.2 and illustrates how mixl handles larger problems using multi-core processing in comparison to Apollo and Biogeme on two program sets, a small one (n=1000, 1000 draws), and a medium size one (n=16,000, 1000 draws). Although Biogeme is faster when fewer processors are available, the use of openMP for the parllelisation in mixl makes it competetive with Biogeme when more cores are used, especially on larger problems. Similarly, we can see how the effectiveness of the R-based parallelisation used by Apollo is more effective on larger problems sizes, but doesn't provide much benefit on more than 4 cores for smaller problems. Figure 4.3 shows this from another perspective, with the speedup on the vertical axis. The speedup value is the performance improvement relative to one processing core. For the three programs, we can see how better speed ups are achieved on larger problems, as the communication costs become less significant. Also visible is the ability of mixl to better utilise the a large number of processing cores - particularly when more than 12 cores are available.

Figure 4.4 shows how the memory usage of the different programs compare. A dashed line indicates the predicted memory usage, as the tests were limited by a 200GB memory ceiling. Apollo does take advantage of R vectorization to avoid R's unoptimised iteration constructs and achieve good results on smaller problems. However, with this technique, the draws must be replicated for each choice task of an individual. The ammount

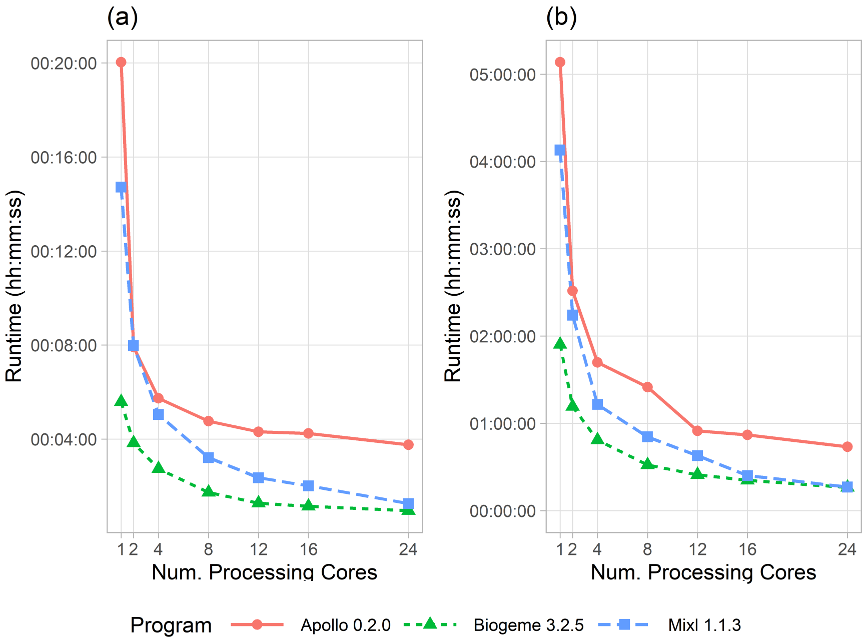


Figure 4.2: Performance of the sample mixed MNL model with 8 parameters and 1000 random draws. (a) 1000 individuals. (b) 16,000 individuals.

of memory required for the duplication means that this approach breaks down for large panel datasets, especially as the number of draws increases, which is required for complex model specifications. The approach used in the mixl package (as with Biogeme) avoids this by accessing the required draws using the ID of the individual. To do this without R's performance penalty is possible by using compiled code. For smaller problems the performance is bound by other sequential parts of the program such as the compilation of the likelihood function. Essentially, mixl is not bound by the number of individuals or repeated choices in the dataset, number of random dimensions or the number of draws used, as long as enough memory is available to store both the data and draws matrix.

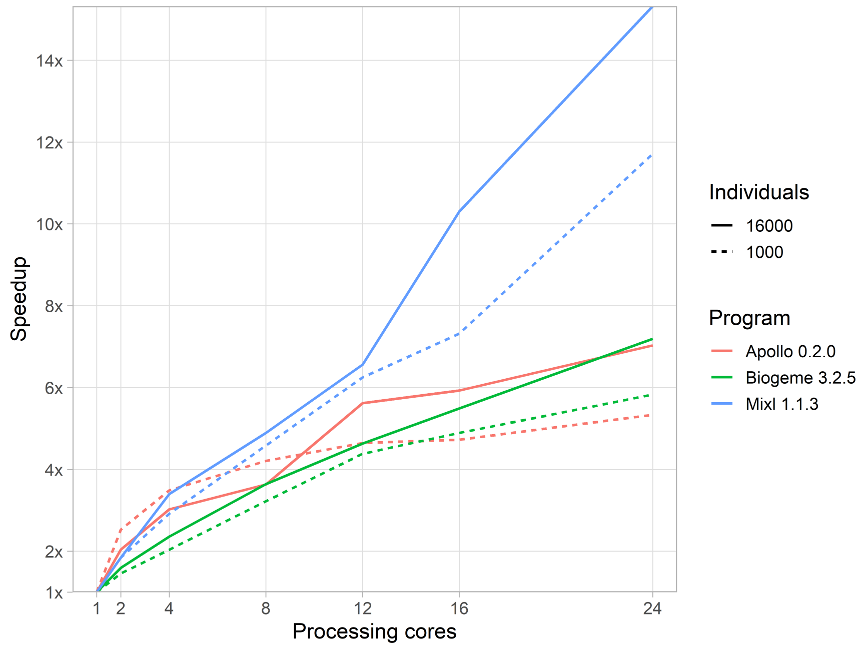


FIGURE 4.3: Speedup over multiple cores for mixl, Apollo and Biogeme on small (n=1000) and large (n=16,000) datasets

4.9 LIMITATIONS

Mixl is designed to support a core group of models, which in the authors experience are used the majority of the time. These include standard MNL, Mixed MNL and also, those less commonly used, hybrid choice models. Furthermore, a design decision was made to only support random heterogeneity across individuals, and not intra-respondent variations. With large datasets and a large amount of draws - even a relatively small number of draws massively increases the memory required, as an extra dimension is required in the draw matrix. Both Apollo and Biogeme support intra-respondent heterogenetity. The second limitation is that mixl currently only supports the logit kernel. As the code is open-source and simply structured, mixl could be extended with with additional model types in the future.

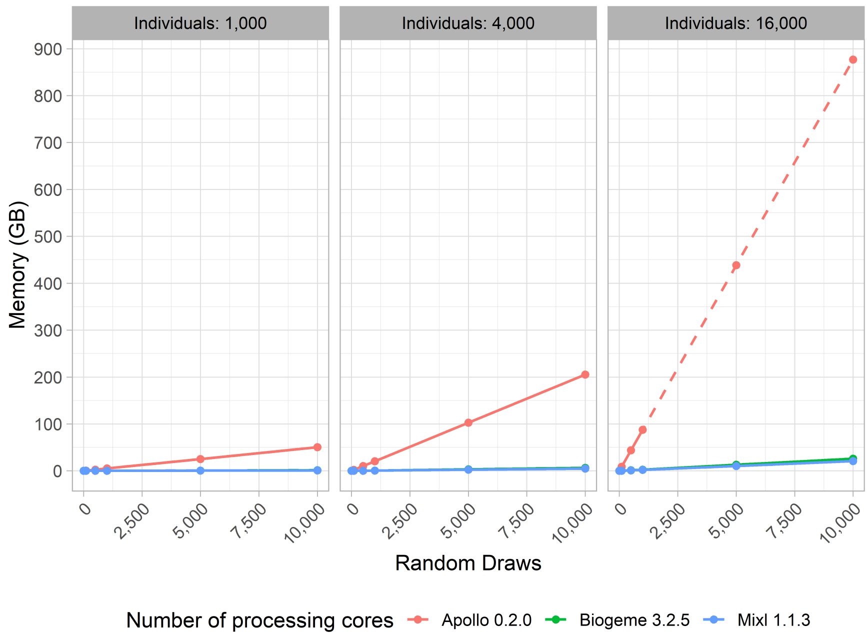


FIGURE 4.4: Comparison of memory usage for mixl, Apollo and Biogeme

4.10 CONCLUSION

This paper presents the mixl R package for estimating flexible multinomial logit models. Mixed models and hybrid choice models are supported through a flexible and intuitive syntax. The package has been designed to have an intuitive model specification syntax, and is engineered with both large datasets and complex mixed MNL models in mind. For R practitioners looking to use other model formulations, Apollo is much more flexible, albeit with performance drawbacks and a different syntax. mixl combines compilation to C++ code with efficient data structures to allow the estimation of models on large datasets that are not feasible with other R packages, especially if the dataset or number of draws used is large. For large problems, parallel computing is an attractive way to gain significant speed increases, and mixlmakes it easy for the user to take advantage of this. The paper presents performance indicators on a complex mixed MNL model estimated on a large dataset with over 128,000 observations demonstrating speedups in model estimation of over 10x when using 24 cores as opposed to a single core, with no increase in memory usage. The package has also already been used in modelling projects (among them, Schmid, Aschauer, et al., 2019) with hundreds of thousands of observations and 10,000 random draws, indicating its robustness and scalability. Future work will aim to integrate the work with other estimation packages such as Apollo, and support more model types.

INSTALLING THE PACKAGE

The estimation software is provided as an R package on CRAN `https://rdrr.io/cran/mixl/`. The code is open-source and shared through Github. A user-guide and documentation are provided with the package.

4.11 ACKNOWLEDGEMENTS

The authors would like to Stephane Hess, who's original R script for discrete choice modeling was the inspiration for this R package. The work was undertaken with the financial support of the SNF and the Deutsche Bahn AG.

4.12 COMPETING INTERESTS

The authors have no competing interests to declare.

REFERENCES

ALOGIT (2016) ALOGIT Software & Analysis Ltd.

Ben-Akiva, Moshe, Daniel McFadden, Kenneth Train, et al. (2019) Foundations of stated preference elicitation: Consumer behavior and choice-based conjoint analysis. In: *Foundations and Trends in Econometrics* 10 (1-2), pp. 1–144.

Ben-Akiva, Moshe, Daniel McFadden, Kenneth E. Train, Joan Walker, Chandra Bhat, Michel Bierlaire, Denis Bolduc, Axel Boersch-Supan, David Brownstone, David S. Bunch, Andrew Daly, André de Palma, Dinesh Gopinath, Anders Karlstrom, and Marcela a. Munizaga (2002) Hybrid Choice Models : Progress and Challenges. In: *Marketing Letters* 13 (3), pp. 163–175.

Ben-Akiva, Moshe E. and Steven R. Lerman (1987) Discrete Choice Analysis: Theory and Application to Predict Travel Demand.

Bierlaire, Michel (2018) PandasBiogeme: a short introduction. Tech. rep. Transport and Mobility Laboratory, EPFL, Switzerland.

Chapman, Barbara M. and Federico Massaioli (2005) OpenMP. In: *Parallel Computing* 31 (10), pp. 957–1174.

Cranenburgh, Sander van and Michiel C.J. Bliemer (2019) Information theoretic-based sampling of observations. In: *Journal of Choice Modelling* 31, pp. 181–197.

Croissant, Yves (2015) Estimation of multinomial logit models in R: The mlogit Packages An introductory example. In: *Data Management*.

Czajkowski, Mikołaj and Wiktor Budziński (2019) Simulation error in maximum likelihood estimation of discrete choice models. In: *Journal of choice modelling* 31, pp. 73–85.

Dumont, Jeff, Jeff Keller, and Chase Carpenter (2015) RSGHB: functions for hierarchical Bayesian estimation: a flexible approach. In: *R package version* 1 (2).

Eddelbuettel, Dirk and Romain Francois (2011) Rcpp: Seamless R and C ++ integration. In: *Journal Of Statistical Software* 40 (8), pp. 1–18.

Greene, William H. (2002) NLOGIT reference guide : version 3.0. Plainview, NY.

Hasan, Asad, Wang Zhiyu, and Alireza S Mahani (2014) Fast Estimation of Multinomial Logit Models: R Package mnlogit. In: *Journal of statistical software* 75 (3).

Henningsen, Arne and Ott Toomet (2011) MaxLik: A package for maximum likelihood estimation in R. In: *Computational Statistics* 26 (3), pp. 443–458.

Hess, Stephane and David Palma (2019a) Apollo version 0.0.9, user manual. Tech. rep.

Hess, Stephane and David Palma (2019b) Apollo: A flexible, powerful and customisable freeware package for choice model estimation and application. In: *Journal of choice modelling* 32, p. 100170.

Hess, Stephane, Kenneth E. Train, and John W. Polak (2006) On the use of a Modified Latin Hypercube Sampling (MLHS) method in the estimation of a Mixed Logit Model for vehicle choice. In: *Transportation Research Part B: Methodological*.

McFadden, Daniel (1974) Conditional logit analysis of qualitative choice behavior. In: *Frontiers in Econometrics*. Ed. by P Zarembka. Academic Press, pp. 105–142.

McFadden, Daniel (1980) Econometric Models of Probabilistic Choice among Products. In: *Journal of Business* 53 (3), pp. 13–29.

McFadden, Daniel and Kenneth Train (2000) Mixed MNL models for discrete response. In: *Journal of Applied Econometrics* 15 (5), pp. 447–470.

Nocedal, Jorge and Stephen J. Wright (2000) Numerical optimization 2nd edition. Springer.

Sarrias, Mauricio, Ricardo Daziano, et al. (2017) Multinomial logit models with continuous and discrete individual heterogeneity in R: the gmnl package. In: *Journal of Statistical Software* 79 (2), pp. 1–46.

Schmid, B., F Aschauer, S Jokubauskaite, S Peer, R Hössinger, R Gerike, S R Jara-Diaz, and Kay W Axhausen (2019) A pooled RP/SP mode, route and destination choice model to investigate mode and user-type effects in the value of travel time savings. In: *Transportation Research Part A: Policy and Practice* 124, pp. 262–294.

Schmid, B., M. Balac, and K. W. Axhausen (2019) Post-Car World: Data collection methods and response behavior in a multi-stage travel survey. In: *Transportation* 46 (2), pp. 425–492.

Schmid, Bas and Kay Werner Axhausen (2019) In-store or online shopping of search and experience goods: A Hybrid choice approach. In: *Journal of Choice Modelling* 31, pp. 156–180.

Schmid, Basil (2019) Connecting Time-Use, Travel and Shopping Behavior: Results of a Multi-Stage Household Survey. PhD thesis. Zurich: IVT, ETH Zurich.

Sobol, I M (1967) On the distribution of points in a cube and the approximate evaluation of integrals. In: *USSR Computational Mathematics and Mathematical Physics* 7 (4), pp. 86–112.

Train, Kenneth E (2009) Discrete Choice Methods with Simulation. 2 Ed. Cambridge university press.

Walker, Joan and Moshe Ben-Akiva (2001) Extensions of the Random Utility Model. Tech. rep. MIT.

Witzgall, Christoph and R. Fletcher (1989) Practical Methods of Optimization. In: *Mathematics of Computation*.

Zeileis, Achim (2006) Object-Oriented Computation of Sandwich Estimators. In: *Journal of Statistical Software* 16 (9), pp. 1–16.

Zeileis, Achim, Susanne Köll, and Nathaniel Graham (2020) Various Versatile Variances: An Object-Oriented Implementation of Clustered Covariances in R. In: *Journal of Statistical Software* 95 (1), pp. 1–36.

5

MOBILITY & COVID-19 IN SWITZERLAND

FULL TITLE

Observed impacts of COVID-19 on travel behaviour in Switzerland based on a large GPS panel

AUTHORSHIP

Joseph Molloy, Thomas Schatzmann, Beaumont Schoeman, Christopher Tchervenkov, Beat Hintermann and Kay W. Axhausen

This chapter was accepted for presentation at the *100th Annual Meeting of the Transportation Research Board, January 2021*, and is under review for publication in *Transport Policy*

5.1 ABSTRACT

In Switzerland, strict measures as a response to the COVID-19 pandemic were imposed on March 16, 2020, before being gradually relaxed from May 11 onwards. We report the impact of these measures on mobility behaviour based on a GPS tracking panel of 1,439 Swiss residents. The participants were also exposed to online questionnaires. The impact of both the lockdown and the relaxation of the measures are presented. Reductions of around 60% in the average daily distance were observed, with decreases of over 90% for public transport. Cycling increased in mode share drastically. Behavioural shifts can even be observed in response to the announcement of the measures and relaxation, a week before they came in to place. Long-term implications for policy are discussed, in particular the increased preference for cycling as a result of the pandemic.

5.2 INTRODUCTION

The sudden onset of the Covid-19 pandemic in early 2020 dramatically altered the rhythms of daily life around the world. Within the space of a few weeks, borders were closed, lockdowns imposed and local economies brought to a stand-still in the attempt to contain the spread of the virus. Schools were closed, and children had their first experience of remote learning. After decades of unfulfilled predictions of the move towards teleworking, firms were forced to make the shift almost overnight.Those under lockdown orders were generally only allowed outside for shopping and small amounts of exercise. As a consequence, mobility behaviour became unrecognisable.

Understanding the impact these restrictions have had on mobility has been challenging, given the pace of the change. New policies were being implemented constantly; classic survey methods would not have kept pace with the dynamic changes to people's travel patterns. In this paper we present both the method and descriptive results from an app-based mobility survey panel of over 1500 people which was recruited as the pandemic took hold in Switzerland in mid-March. Uniquely, this panel was recruited from a previous mobility survey undertaken in autumn 2019. Thus, a baseline behaviour for each participant in the study is available against which their mobility behaviour under these extra-ordinary circumstances can be understood.

The paper is structured as followed. First a timeline of the pandemic as it developed in Switzerland is presented, including the various restrictions and relaxations which occurred. The period up to the end of July is covered in this paper. A brief overview of GPS-based mobility studies in the literature is presented, as well as other available sources of Covid-19 mobility data. In the methods and data section, the recruitment and app are presented. The results section presents some of the key insights from the mobility tracking. In the discussion the implications of these findings for transport policy in the near future are discussed.

5.2.1 *Timeline of coronavirus in Switzerland*

The first confirmed case of COVID-19 was registered in Switzerland on February 25, 2020. The onset followed that in Northern Italy, which was one of the earliest hotspots outside China. The situation deteriorated quickly and by March 9, over 100 people had been infected. At the time of writing, July 31 2020, 791,725 people have been tested in Switzerland, over 35,000 people have been infected with the virus and more than 1,700 have died, which translates to infection and death rates per 100,000 inhabitants of 38 and 2.4, respectively.

On March 16, the Federal Council declared an extraordinary situation, which gives it widespread competencies that allowed it to order the introduction of uniform measures in all cantons. Non-essential businesses were closed, along with schools, recreational facilities and public parks. Only food stores, post offices and healthcare institutions were not affected. Furthermore, checks on the borders to Germany, Austria, France and Italy were introduced. Entry to Switzerland from its four large neighbours was only possible for Swiss citizens, persons holding a residence permit for Switzerland and those who have to enter Switzerland for work-related reasons. Employers were urged to reorganise the working hours of their employees to avoid rush hour travel. Wherever possible, home-office was to be implemented. To minimise the risk of infection, the Federal Office of Public Health (FOPH) recommended avoiding the use of public transport wherever possible. People who had symptoms of respiratory disease and people over 65 years of age were told to not use public transport. National and regional passenger transport services were maintained to support the functioning of the economy and society. International passenger transport was maintained to countries with open borders.

The number of confirmed COVID-19 cases in Switzerland increased rapidly. On March 20, more than 4,800 infections and 43 fatalities were reported. The canton of Ticino on the border to Northern Italy was hit especially hard by the corona virus. The Federal Council decided to step up measures and forbid gatherings of more than 5 people in public places, and an inter-personal distance of 2 meters was mandated for groups with fewer than 5 people. However, unlike in neighboring countries as for example in Austria, Belgium, France, Italy, Luxembourg and Spain, people were allowed to leave the house during the lockdown. Public transport continued operation, though at reduced frequency. The peak rate of daily infections and COVID19-related deaths was in late March and early April. The statistics were subject to weekly fluctuations and showed fewer cases towards the end of the week. On April 5, Switzerland reported a total 158,000 COVID-19 tests, 15% (21,100) of which were positive, and 559 deaths, one of the highest incident rates in Europe. After that, infections started to decrease as a consequence of the containing measures. Despite the fact that the measures put in place to combat the virus were being followed well by the public and were having the desired effect, the Federal Council extended the measures until April 26.

On April 27, hospitals were able to resume all medical procedures, including non-urgent procedures, and a first group of businesses was allowed to re-open (garden centers, florists, hairdressing salons and cosmetic studios). The Federal Council's strategy to emerge from the lockdown was structured into three phases, with transitions depending on the number of new infections, hospital admissions and deaths as well as hospital occupancy rates. In a first phase from April 27 to May 11, measures were eased on businesses where a low level of direct contact is possible, where precautionary measures can easily be put in place, and where there will be no significant movements of people. As opposed to other countries, there was no general obligation for healthy people to wear a mask. Keeping a safe distance and washing hands was seen as the most effective protective measures. Still, sick people were advised to stay home in isolation.

On May 11, Switzerland moved to phase 2 by further easing measures as the spread of the corona virus had continued to slow. Even though rules on hygiene and social distancing still applied, most types of businesses opened again, including restaurants, along with mandatory school for students up to grade 9. Museums and libraries were allowed to re-open and sports activities in small groups of up to 5 people were permitted as well. In parallel, restrictions on entering Switzerland were relaxed and scheduled public transport services increased significantly. Furthermore, the Federal Council decided to introduce contact tracing for new infections with COVID-19.

Entering phase 3 of the emergency plan on June 6, all events of up to 300 persons and spontaneous gatherings for up to 30 people were allowed again. High schools and universities were able to resume, all leisure and entertainment businesses plus tourist attractions re-opened. Switzerland reported a cumulative total of 30,988 infections, 1,663 COVID-related deaths and only 16 new cases. The number of new infections, hospitalisations and deaths continued to fall despite measures being eased, and stabilised at a low level. The extended powers of the federal government expired on June 19.

From June 22 onwards, the start of the fourth phase, most of the measures put in place were lifted and only the ban on large-scale events remains in place until at least the end of August. Interestingly, the Federal council reduced the minimum distance that should be kept between two people from 2 metres to 1.5 metres and strongly recommended to wear a face mask when using public transport if it is not possible to maintain the necessary distance. It also lifted the recommendation to work from home, leaving the decision to the employer, which is required to protect the health of staff by putting in place appropriate measures. In contrast to the first wave, the prime responsibility in the event of a renewed increase in COVID-19 cases now rests with the 26 cantons in Switzerland.

In response to an increase in new daily infections in mid-June and increased ridership on public transport, the Federal Council made wearing masks compulsory on public transport throughout Switzerland starting from July 6. Since mid-June, the number of new corona virus cases has been rising in Switzerland as infected persons have entered the country from countries both within and outside the Schengen area. Consequently, travellers entering Switzerland from certain regions must quarantine for ten days upon their return.

5.2.2 *Tracking studies in the literature*

The use of GPS tracking in mobility studies is becoming an integral data collection method in transportation. Where traditionally, such surveys required participants to carry a GPS logging device which had to be returned at the end of the study for the data to be collected, more recent methods allow users to install an app on their smartphone, which uses the GPS receiver and other location services on the phone to detect the location. These apps run in the background, and depending on the phone model, have a minimal effect on battery life, something which used to be extremely detrimental to such studies. For example, in the original MOBIS study sample on which this study is based, only 12.5% of participants who started tracking were lost over the 8 week study duration (Molloy et al., 2020).

In recent years there have been an increasing number of studies which have used either a separate GPS device (Livingston, 2011; Nielsen, 2004; Wargelin, Stopher, et al., 2012), or a smartphone app (Allström, Kristoffersson, and Susilo, 2017; Greene et al., 2012; Nahmias-Biran et al., 2018; Safi et al., 2015; Stopher, Daigler, and Griffith, 2018). This data is then usually segmented into trips (or sometimes lower level segments) and activities, on which the mode and trip purpose are imputed. The usefulness of this method is indicated by multiple sources that identify trip under-reporting in traditional paper-based travel diaries (Janzen et al., 2018; Stopher, FitzGerald, and Xu, 2007; Wolf et al., 2003).

5.2.3 *Other COVID-19 mobility data sources*

The rapid onset of the COVID-19 pandemic meant that there was little opportunity to set up new tracking studies to measure changes in mobility. However, there were a number of data sources which were already available. The most prominent were those from Apple (Apple, 2020) and COVID-19 Community Mobility Reports (Google, 2020). The Apple reports monitor the number of search requests by transport mode (driving, walking, public transport) as an indicator for mobility. The Google reports use records of visits to land use types (i.e. parks), based on anonymised data collected from google apps and mobile devices. Both these reports present only overall aggregate numbers and do not include analysis of socio-demographic differences. For the baseline, the Apple reports take January 13, 2020, and Google uses the period January 3 – February 6, 2020. Another available resource for Switzerland is produced by Intervista AG, in collaboration with the Zurich Statistics Bureau (Intervista AG, 2020). The Intervista sample consists of 2561 persons (daily average) and is broadly representative, with socio-demographic attributes for the participants. Data is also available since the start of 2020.

5.2.4 *The impacts of pandemics on mobility behaviour*

It has been widely acknowledged that transportation is a key driver in the spread of infectious diseases (Baroyan and Rvachev, 1967; Herrera-Valdez, Cruz-Aponte, and Castillo-Chavez, 2011). The important role of mobility in a pandemic has been demonstrated for historical pandemics such as the Spanish flu in 1918 (Ammon, 2002; Trilla, Trilla, and Daer, 2008). For a comprehensive overview of studies exploring the link between transport and infectious diseases, see Muley et al. (2020). However, the impact of measures to overcome pandemics also have an impact on mobility. In the much of the recently published transport literature on the Covid-19 pandemic, the focus has been drawing conclusions on the impact of restrictions on the spread of the disease, with little focus on the specific impacts on the restrictions on mobility itself within the population. In the following paragraphs a few studies which do focus on mobility are highlighted.

In the current pandemic, mobile data has played a key role in understanding the changes in both regional and global mobility during the pandemic. Many countries are now using mobile data to understand the effectiveness of measures, including Austria, Belgium, Chile, China, Germany, France, Italy, Japan, Spain, United Kingdom and the United States (Oliver et al., 2020; Yabe et al., 2020). In particular, Heiler et al. (2020) used real-time anonymised mobile phone data to understand the changes in mobility behaviour as a result of the introduced measures in Austria during the first wave. They saw a doubling of the number of persons with a radius of gyration (activity space) of less than 500m, and increased segmentation of the community structure.

In the USA, Badr et al. (2020) found a strong correlation between reduced mobility behaviour and decreased COVID-19 case growth rates. Furthermore, they show evidence that behavioural changes were already observable days to weeks before movement restriction policies were implemented, indicating that it was not just the introduced policy measures that restricted mobility, but the desire of persons to avoid the pandemic.

All these data sets use data collected from mobile phone providers, which are less effective at capturing mobility changes at local urban scales. Socio-demographic differences (e.g. across age, gender and household-size) can also not be explored with this data. This is important for understanding how the mobility of different groups was affected by the pandemic.

Other work has used travel surveys to explore the impact of the COVID-19 pandemic on mobility. Beck and Hensher (2020) performed a household survey of 1457 respondents. The main conclusions of this paper focus on telework (working-from-home) and suggest that there could be significant implications for localised transport networks. The shift towards more flexible working conditions may also lead to a spreading of peak-hours.

Huang et al. (2020) used location data from Baidu in China to examine the impacts on mobility and activity patterns as a result of the pandemic. They suggest that the government should promote cycling as means of transport and the construction of bike lanes.

5.3 methods and data

5.3.1 *Recruitment*

Starting in September 2019, a sample of 5,375 persons living in Switzerland were recruited to a mobility pricing field experiment called MOBIS (MObility Behaviour in Switzerland). The study required the participants to use an app (available for both Android and iPhone) which used the phones location services to track their location, and identify trip stages and activities. They were tracked for 8 weeks to investigate their response to a conceptual mobility pricing scheme. 3680 persons completed the 8 weeks, with the last person finishing in January. They did not have to uninstall the app afterwards, and could continue using it. At the start of March 2020, around 300 people were still using the app. The tracking panel for the MOBIS-COVID19 study was recruited from those 3680 participants who completed the MOBIS study. Of these participants, around 1,600 volunteered to reactivate the tracking app, Catch-my-Day, developed by Motion-Tag. table 5.1 presents the distribution of the sample by different socio-demographics, showing that the MOBIS-COVID19 sample is broadly representative. Our sample is slightly more educated and more likely to be employed - but this is primarily since those over 65 years of age were not invited to the original MOBIS study. Participants were only eligible to participate in the MOBIS study if they used a car at least 2 days a week. This also skewed our sample towards car drivers, compared to the Swiss population.

Weekly reports have been produced online in three languages (English, German and French) since the start of the MOBIS-COVID19 study, to which both participants and the public have access. Participants also have access to a custom dashboard of their own mobility behaviour during the crisis, as long as they continue tracking. This was designed as a form of non-monetary incentive to encourage continued participation.

5.3.2 *App based tracking*

The Catch-my-Day app works similarly to other tracking apps. It runs in the background, collecting location data from the phone, with the participant's consent. The tracking can be turned on and off at any time. The data is transferred to the server when the phone is connected to a WiFi network, where the location points are processed to detect the transport stages, activity locations, mode of travel and where possible, trip purpose. The trip purpose detection is based on the activity type nominated by the participant for other activities at the same location. The detection algorithms are based on Openstreetmap data (OpenStreetMap contributors, 2017) and available public transport schedules. Users are able to validate the imputations made by the algorithms and make corrections.

Since not all activities have an assigned purpose in the data collected from the app, we impute the missing activity purposes for the rest of the data using a random forest approach, using the reported activity purposes as the ground-truth dataset.

5.3.3 *Online surveys*

The tracked data was supplemented by online surveys, which the participants were invited to participate in by email. Initially, socio-demographic information about participants was collected in the original MOBIS study. Participants were also asked in May to complete another 2 surveys, one during the lockdown, and one after, asking about their working conditions and experience during the lockdown.

5.3.4 *Sample weighting*

As mentioned in above, our sample is skewed towards car-users in urban areas, due to the sample composition of the original MOBIS study. To make some correcrtion for this, the MOBIS-Covid sample is weighted against the original 22,000 participants which filled out the introductry survey in the MOBIS Study. This allowed a weighting of the MOBIS-Covid results by age, gender, income, education, mobility tool ownership and accessibility. Weights were calculated for each person-week in both the MOBIS and

MOBIS-Covid study periods, against the sample present in that week. This accounts for the changing sample composition over the weeks of the study. The weights are both reasonable and stable across the study period, with the weights in the range 0.5-3.

5.4 RESULTS

Across all segments of the population, a downward trend in the average travelled Kilomters per day was observed before the lockdown measures were put in place on March 16. This indicates that the Swiss population preempted the preventative measures, and were proactively taking steps to protect themselves from the virus by reducing their mobility. This is also a consequence of the clear messaging from the Swiss federal government, which mostly gave forewarning of at least a week before measures were implemented.

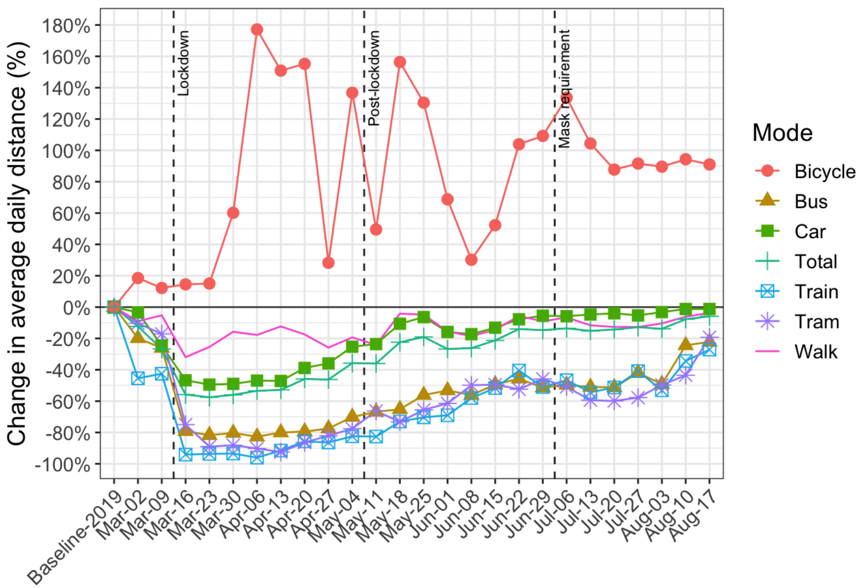


FIGURE 5.1: Weekly change in average daily Km travelled

Mobility behaviour was most suppressed in the first week of the lockdown, however, the average daily distance and the number of trips per day immediately started increasing, demonstrating the challenge of sustaining a suppressed mobility demand over a long period through non-policed lockdown measures.

5.4.1 *Impact of home-office and Kurzarbeit (Short-work)*

During lockdown, 20% of participants reported being on kurzarbeit (Short-work). kurzarbeit is an industrial-relations policy that exists predominately in German speaking countries, where a employee in the scheme has their hours reduced by between 20 and 95% due to extreme economic circumstances. They are then paid the majority of their usual wage, with the difference between hours worked and hours paid covered by government finances.

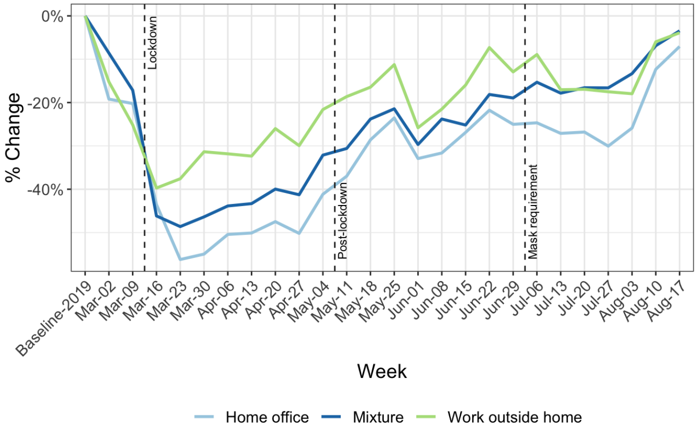
Surprisingly, being on kurzarbeit had little effect on the number of trips per day during the lockdown, compared to those on normal conditions. The amount of home-office (telework) a participant was performing is much more influential (see fig. 5.2). After the lockdown was lifted, the difference between the groups has reduced as more of those who were allowed home-office have returned to the workplace. However, differences still persist, clearly indicating that home-office is effective for suppressing travel demand.

5.4.2 *The cycling boom*

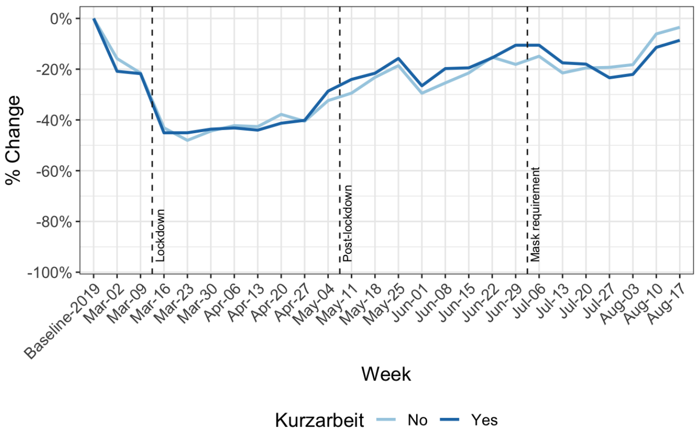
Significant changes in modal split were observed, particularly towards cycling, where a large increase in the average daily distance travelled was observed (see fig. 5.1). The change in average daily distance is, on some weeks, greater than 100%, which is well beyond what the seasonal causes would imply. The magnitude of the variations fluctuates throughout the weeks, depending on the weather, which was excellent during the lockdown period, but variable afterwards.

A large increase of the number of cycling trips per day was also observed, as can be seen in the temporal patterns by time-of-day and type of day in fig. 5.4. The red curve shows the average number of cycling trips started for each hour of the day during the 2019 baseline period, whereas the blue curves show this for the last two weeks respectively. All intermediate weeks are plotted in grey.

These temporal patterns indicate that cycling was primarily used as a leisure tool. First, the increase primarily occurred on weekends, indicating



(a) By home-office arrangement



(b) By kurzarbeit status

FIGURE 5.2: Trips/day

a leisure objective. Second, although an increase in the number of trips was also observed during the week, much of this occurred over midday, hinting that these trips were not conducted for commuting purpose. Indeed, considering that a large number of commuting and station-access/egress trips did not take place during the lockdown, this suggests that the use of cycling for other purposes has increased.

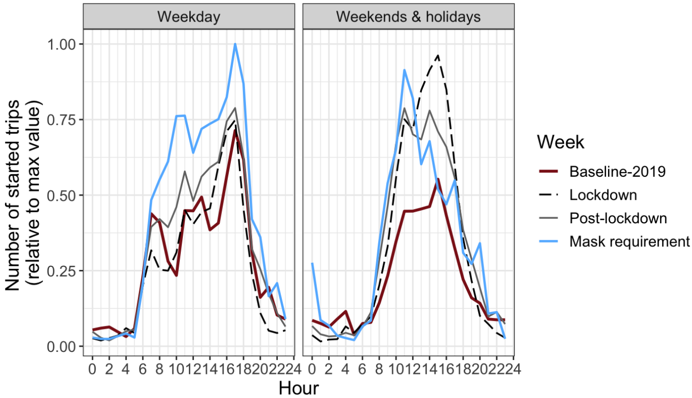


FIGURE 5.4: Hourly bicycle trip counts

In Figure 5.5, the increase in mode share for different trip purposes is presented. The mode share compared to the baseline period has increased for all trip purposes. The influence of the weather is visible across all purposes. The increases were largest during the lockdown, but the trend has stabilised since then, at around a 75% increase over the baseline period.

5.4.3 *Socio-demographic variations during and after the lockdown*

As indicated by the analysis of working conditions, particularly home-office capability, certain segments of the working population were able to reduce their travel more than others. fig. 5.6 shows the reduction in the average daily travelled kilometers by education level, as compared to the reference period in September/October 2019. Those with a tertiary education (i.e. from a university or technical college) had a larger decrease in daily kilometers than the less educated. Towards the end of the lockdown and post-lockdown, this difference became more pronounced. Those with the mandatory education level experienced a rapid increase in their daily kilometers. This coincides with the return to work for those in service

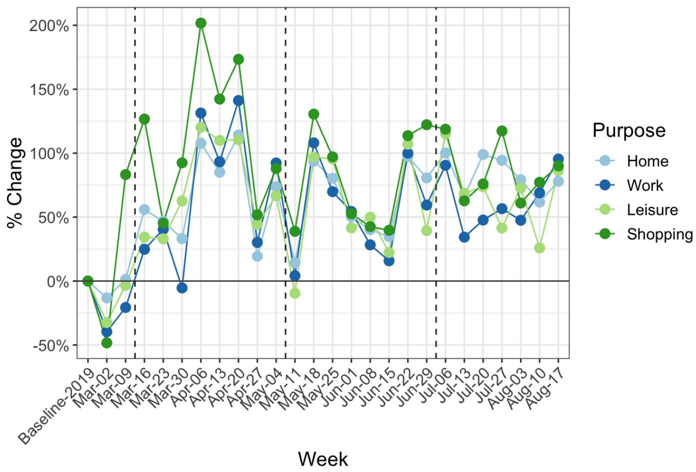


FIGURE 5.5: Increase in bicycle mode share by trip purpose

industries, where professional degrees are not required. Travel per day for this segment has plateaued in the post-lockdown period, indicating that a new daily pattern for this cohort has been reached.

The relaxation of the lockdown measures were met with a continued increase in travelled distance. All socio-demographic groups saw a spike in the average daily distance as the lockdown measures were lifted, to varying degrees, before a return to a trend similar to that observed during the lockdown.

The various adaptions to the lockdown measures by different household sizes is also insightful, as seen in fig. 5.7. Again, the total average weekly kilometers dropped by approximately 60%, with the reduction in travel correlated with household size throughout the lockdown period. On one had, larger households had more travel constraints (e.g. taking care of the children who now have to stay at home). Furthermore, those in smaller households, particularly single person households had the smallest reduction in daily travel. It is important to note that no household had more than one person participating in the study, hence we cannot say how tasks such as shopping were redistributed among the household members to reduce public exposure. It can also be reasonably inferred that those in larger households had a larger incentive to reduce their daily travel. If one household member were to contract the virus, it would have spread to the whole household. This ties to the idea of social pressure, with larger house-

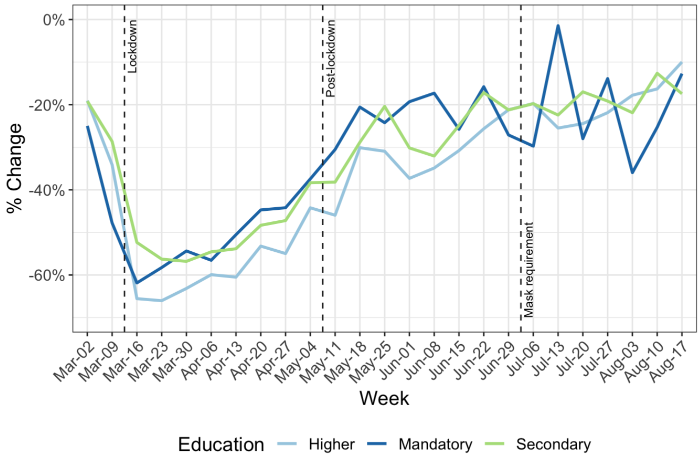


FIGURE 5.6: Change in kilometers travelled by education level

holds resulting in a level of normative behaviour within the household community.

The relaxation of the lockdown measures not only impacted travelled distance, but also had an effect on the average number of trips per day and the activity space. Indeed, at the beginning of the lockdown, the number of daily trips plummeted by nearly half and remained low during the entirety of the lockdown, whereas they have essentially returned to pre-COVID numbers during the post-lockdown period (see fig. 5.9a). The evolution, however, was quite gradual. The effect of the relaxation is much more evident on the activity spaces, as can be seen in fig. 5.9b. The lockdown caused a reduction in the activity space to around $50km^2$, which only slowly increased during the lockdown - indicating that while people were travelling more, they were doing so within their local neighbourhood. Immediately after the lockdown was lifted, the area of the weekend and holiday activity spaces increased. A short downward trend over 3 weeks followed, before the activity space started growing again. This post-lockdown downward trend was also observed on weekdays, which one would not expect to be compatible with the rhythms of the 5-day work week. The size of the activity spaces are now similar to what they were pre-COVID, both during the week and on weekends.

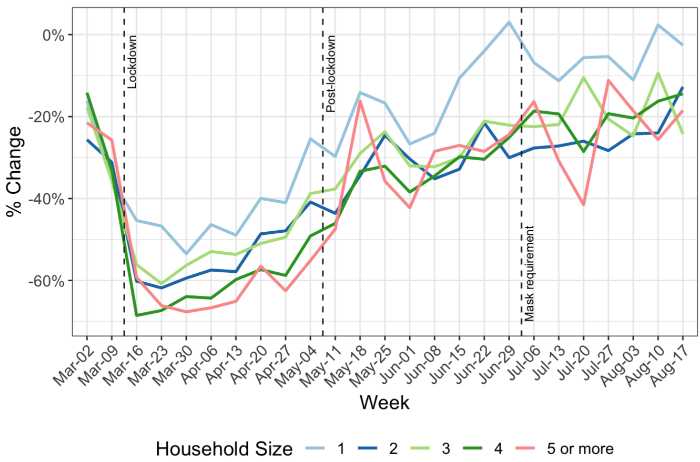


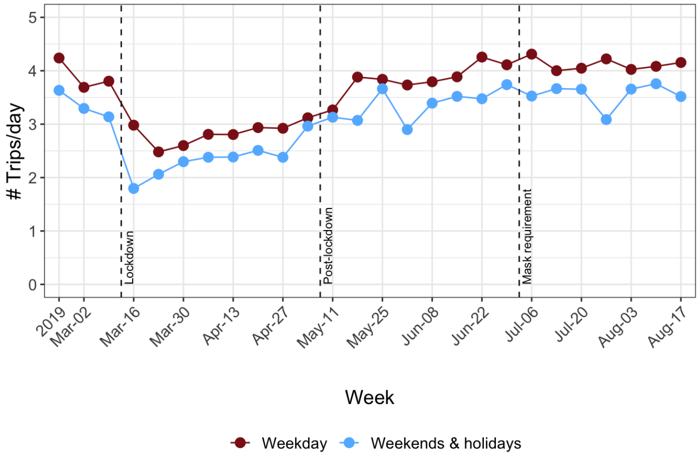
FIGURE 5.7: Change in kilometers travelled by household size

5.4.4 *Change in road speeds*

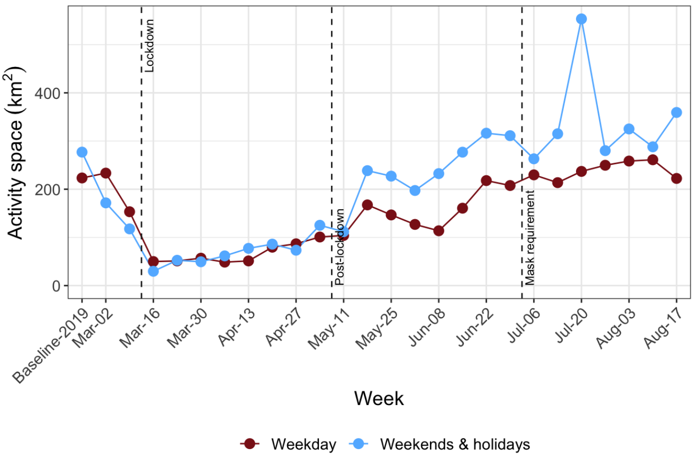
The general reduction in travel observed during the lockdown period also included a significant reduction in car travel. With less vehicles on the roads, congestion levels dropped during the lockdown, and average travel speeds across different trip-distance classes correspondingly increased.

fig. 5.10 shows the effect of the COVID-19 crisis on median car travel speeds during the week, i.e. excluding weekends and holidays. The black line indicates the baseline travel speeds from the baseline period in 2019. Car trips of over 500 km, which correspond to driving across Switzerland, or with an average speed of over 180 km/h were considered unrealistic and removed from the analysis.

During the lockdown period from March 16 to May 11, an increase by up to 15 km/h in the peak-hour speeds was observed, indicating a decrease in overall congestion. Reductions in congestion were observed during both morning and evening peaks. However, no reduction was observed in the middle of day. Since the relaxation of the lockdown measures on May 11, peak-hour speeds have returned to almost pre-COVID-19 values, a sign that congestion is returning to usual levels in peak periods. This is consistent with the fact that public transport use is still suppressed, while car mileage is roughly back to the baseline (see fig. 5.1). Post-lockdown, non-peak travel speeds during the day are now lower than they were during the



(a) Average number of daily trips



(b) Activity space (95% CI Ellipse)

Figure 5.8: Changes in activity space and number of trips

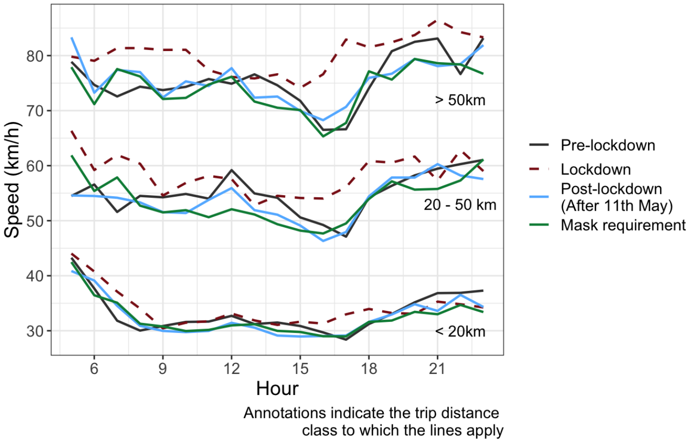


FIGURE 5.10: Road travel speeds

baseline period, particularly for the 20-50km trip-distance class. Viewed in combination with fig. 5.11, which shows the hourly counts for car trip departures in the two weeks post-lockdown, it is evident that more car trips are being generated during the middle of the day than during the baseline period. This is a natural consequence of the reduction in public transport usage, and has concerning implications for the coming months if public transport demand remains suppressed.

5.4.5 *Transport mode share shifts*

The COVID-19 crisis also had had a clear impact on mode shares. While the overall distance travelled has recovered since the lockdown and is now only 15% below pre-COVID values, public transport has been most affected. The daily distance travelled on public transport modes is still 50% below the reference values, 8 weeks after the lockdown ended. In contrast, the daily car distance is now essentially back to pre-COVID levels. The biggest winner has been cycling, which has seen an increase in travelled distance throughout the crisis. However, it is also necessary to look at the transport modal shifts that occurred.

The change in mode shares over the course of the COVID-19 crisis for different types of public transport subscriptions - GA (National), Halbtax (50% discount) and other - can be best visualized with the help of ternary

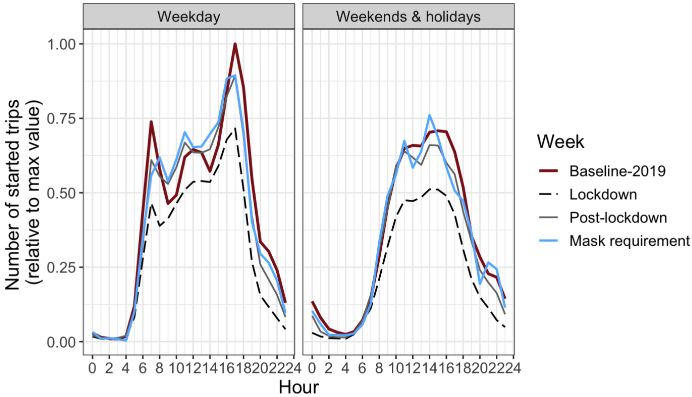


FIGURE 5.11: Car travel hourly counts

diagrams. fig. 5.12 plots the shifts in mode shares by distance. The modes are grouped into the following categories:

- Motorized individual transport (car, motorbike, taxi, Uber)
- Public transport (bus, tram, ferry, metro, train)
- Non-motorized transport (walk, bike)

During the lockdown, a higher share of kilometers were performed using motorized individual and non-motorized modes as compared to the reference period. After the lockdown, the share of public transport has increased and the share of non-motorized modes has decreased slightly. The share of motorized individual modes it still greater than during the reference period. The modal shifts were much larger for those with a full national subscription. Those with other types of subscription - which are usually for a particular zone or connection - have not returned to public transport in the same way that national ticket holders have.

5.5 DISCUSSION

The reductions in mobility have been drastic. The average daily kilometers travelled fell by 60% and 95% for car and train modes respectively (see fig. 5.1). Car travel has recovered slowly, back to the Autumn values within 2 weeks of the lockdown measures being relaxed. However, public transport

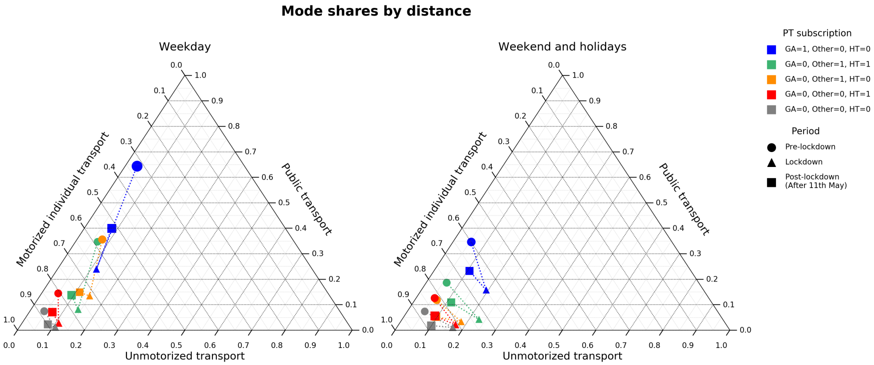


FIGURE 5.12: Mode shares by distance during the COVID-19 crisis

ridership at the same point was still low, at 20% of pre-COVID levels, despite the importance of public transport as a commuting mode in Switzerland. This is likely a result of many people still working from home. Indeed, further results show that the morning and evening peeks are only slowly returning - and this return is dominated by car travel.

The fall in the number of new corona cases in Switzerland since the introduction of the measures (Federal Office of Public Health, 2020) shows that restricting social interaction is a key requirement in dealing with the pandemic. The home-office requirement aimed to reduce the possibility for disease transmission both in the workplace and on public transport. Indeed, the daily number of public transport trips has fallen by around 90%.

An important question is to what extend home-office will be encouraged post-pandemic, and the implications for transport policy. A reduction in peak-hour travel may reduce crowding and thus postpone or negate the need for further infrastructure investment and dynamic demand measures, such as mobility pricing.

Furthermore, it remains to be seen if some of the behavioural shifts observed during this period will become permanent. These shifts have been both positive and negative.

The increase in car usage during the day, with midday off-peak travelled kilometers now slightly above the 2019 baseline (see fig. 5.11), suggests that people are driving more to avoid public transport. It remains to be seen if this observed trend continues in the coming months. If and when workers eventually return to the workplace, this shift from public transport to car usage would lead to additional congestion and be financially challenging for the public transport operators. The obligatory wearing of masks was introduced on the July 6, in order to reduce transmission on public transport. This will also be important for encouraging a return to public transport. Passengers may see public transport as safe enough when masks are required. The alternative would have been other measures to increase public transport capacity to allow social distancing, such as increasing the train length or providing more services. These policy measures haven't been implemented, with mask wearing deemed to be sufficient. Furthermore, the cost burden is shifted to the passengers, who need to provide their own mask. If disposable masks are used, this effectively increases the per trip cost of travel, which may further suppress public transport demand in the coming months. Tirachini and Cats (2020) suggest that there is a possible risk that if the public transportation is seen as unsafe during the pandemic due to crowding and a lack of social distancing measures, formed negative perceptions of public transportation will become more prevalent and could persist after the pandemic, resulting in the formation of new undesirable habits.

Home-office will continue to play a role. Globally, many companies are continuing to allow or even encourage home-office working (The most well publicised examples being the big tech companies). This is also the case in Switzerland. In the post-lock period, we have already seen how congestion has increased in certain times of the day.

The Cycling boom observed since March also raises policy challenges for Switzerland. Other research has highlighted the negative effect of car traffic on willingness to cycle (Kaplan, Luria, and Prato, 2019), The lack of car traffic during the lockdown undoubtedly made cycling more attractive. Sales of cycles increases dramatically, not just in Switzerland (Matthias Heim, 2020), but worldwide (Roger Harrabin, 2020). Sustaining the observed increase in cycling would be associated with health benefits (De Hartog et al., 2010), but also with an increase in accidents. The Covid-19 pandemic has shown a willingness in the population to take up cycling, and policy measures have already been mooted or implemented in other jurisdictions to further encourage this trend (Keohane and Abboud, 2020). In the countries there has been moves to introduce new temporary and permanent cycling infrastructure as a result of the pandemic. One of the key justifications is to simultaneously reduce demand for public transport and try persuade public transport users to shift to active modes rather than use the car, which would increase congestion. Such a policy has been suggested for Zurich, the largest city in Switzerland and the initiative was approved in the last elections in September 2020. This is a good step towards making these changes permanent.

In Handy, Van Wee, and Kroesen (2014), the key research needs for promoting cycling are discussed. In particular, they note that even in pro-cycling cities, strong evidence around cycling behaviour is important for effectively directly resources. To this aim, the wealth of data collected in the MOBIS-COVID19 study will prove invaluable in highlighting areas which are underserved by cycling infrastructure. Lanzendorf and Busch-Geertsema (2014) identify the importance of local government policy in increasing the cycling mode share. They also note that in the case studies presented, the governments took advantage of favourable general conditions. With this in mind, governments should take advantage of the observed cycling uptake as an opportunity to further develop their cycling strategies.

5.6 CONCLUSION

The application of app-based tracking, combined with online surveys of participants allowed the study of the changes in mobility patterns caused by the Covid-19 pandemic. The study sample is broadly representative, with a slight over representation of car owners. Important variations in travel reduction were observed across different socio-demographic groups, which are obscured in other data sources. Finally, the implications for transport policy in Switzerland are widespread, especially if further increases in congestion are to be avoided, and the observed uptake of cycling to be made habitual.

REFERENCES

Allström, Andreas, Ida Kristoffersson, and Yusak Susilo (2017) Smartphone based travel diary collection: Experiences from a field trial in Stockholm. In: *Transportation research procedia* 26, pp. 32–38.

Ammon, CE (2002) Spanish flu epidemic in 1918 in Geneva, Switzerland. In: *Eurosurveillance* 7 (12), pp. 190–192.

Apple (2020) Covid-19 Mobility Trends Reports. online, [`https://www.apple.com/covid19/mobility`], accessed 31/07/2020.

Badr, Hamada S, Hongru Du, Maximilian Marshall, Ensheng Dong, Marietta M Squire, and Lauren M Gardner (2020) Association between mobility patterns and COVID-19 transmission in the USA: a mathematical modelling study. In: *The Lancet Infectious Diseases* 20 (11), pp. 1247–1254.

Baroyan, OV and LA Rvachev (1967) Deterministic models of epidemics for a territory with a transport network. In: *Cybernetics* 3 (3), pp. 55–61.

Beck, Matthew J and David A Hensher (2020) Insights into the impact of COVID-19 on household travel and activities in Australia–The early days of easing restrictions. In: *Transport policy* 99, pp. 95–119.

De Hartog, Jeroen Johan, Hanna Boogaard, Hans Nijland, and Gerard Hoek (2010) Do the health benefits of cycling outweigh the risks? In: *Environmental health perspectives* 118 (8), pp. 1109–1116.

Federal Office of Public Health (2020) New coronavirus: Situation in Switzerland. available online at: `https://www.bag.admin.ch/bag/en/home/krankheiten/ausbrueche-epidemien-pandemien/aktuelle-ausbrueche-epidemien/novel-cov/situation-schweiz-und-international.html`. Accessed: 2020-05-28.

Google (2020) Covid-19 Community Mobility Reports. online, [`https://www.google.com/covid19/mobility/`], accessed 31/07/2020.

Greene, Elizabeth, Leah Flake, Kevin Hathaway, and Michae Geilich (2012) A Seven-Day Smartphone-Based GPS Household Travel Survey in Indiana. In: *TRB 95th Annual Meeting Compendium of Papers*. Transportation Research Board.

Handy, Susan, Bert Van Wee, and Maarten Kroesen (2014) Promoting cycling for transport: research needs and challenges. In: *Transport reviews* 34 (1), pp. 4–24.

Heiler, Georg, Tobias Reisch, Jan Hurt, Mohammad Forghani, Aida Omani, Allan Hanbury, and Farid Karimipour (2020) Country-wide mobility changes observed using mobile phone data during COVID-19 pandemic. In: *arXiv preprint arXiv:2008.10064*.

Herrera-Valdez, Marco Arieli, Maytee Cruz-Aponte, and Carlos Castillo-Chavez (2011) Multiple outbreaks for the same pandemic: Local transportation and social distancing explain the different" waves" of A-H1N1pdm cases observed in México during 2009. In: *Mathematical Biosciences & Engineering* 8 (1), p. 21.

Huang, Jizhou, Haifeng Wang, Miao Fan, An Zhuo, Yibo Sun, and Ying Li (2020) Understanding the impact of the COVID-19 pandemic on transportation-related behaviors with human mobility data. In: *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pp. 3443–3450.

Intervista AG (2020) Mobilität in der Schweiz während der Covid-19-Pandemie (german). online, [https://www.intervista.ch/mobilitaets-monitoring-covid-19], accessed 31/07/2020.

Janzen, Maxim, Maarten Vanhoof, Zbigniew Smoreda, and Kay W Axhausen (2018) Closer to the total? Long-distance travel of French mobile phone users. In: *Travel Behaviour and Society* 11, pp. 31–42.

Kaplan, Sigal, Ravid Luria, and Carlo G Prato (2019) The relation between cyclists' perceptions of drivers, self-concepts and their willingness to cycle in mixed traffic. In: *Transportation research part F: traffic psychology and behaviour* 62, pp. 45–57.

Keohane, David and Leila Abboud (2020) Cycling lanes, wider pavements: How EU cities rethink public transport. In: *Financial Times*.

Lanzendorf, Martin and Annika Busch-Geertsema (2014) The cycling boom in large German cities—Empirical evidence for successful cycling campaigns. In: *Transport policy* 36, pp. 26–33.

Livingston, J (2011) Atlanta Regional Commission-Regional Travel Survey-Final Report. In: *PTV NuStata, Austin, TX*.

Matthias Heim (2020) Die Schweiz hat umgesattelt (german). In: *SRF (Swiss Radio and Television)*.

Molloy, Joseph, Christopher Tchervenkov, Beat Hintermann, and Kay W. Axhausen (2020) MOBIS: An 8-week mobility pricing field experiment using GPS tracking. presented at Mobile Tartu.

Muley, Deepti, Md Shahin, Charitha Dias, and Muhammad Abdullah (2020) Role of Transport during Outbreak of Infectious Diseases: Evidence from the Past. In: *Sustainability* 12 (18), p. 7367.

Nahmias-Biran, Bat-hen, Yafei Han, Shlomo Bekhor, Fang Zhao, Christopher Zegras, and Moshe Ben-Akiva (2018) Enriching Activity-Based Models using Smartphone-Based Travel Surveys. In: *Transportation Research Record* 2672 (42), pp. 280–291.

Nielsen, Otto Anker (2004) Behavioral responses to road pricing schemes: Description of the Danish AKTA experiment. In: *Intelligent Transportation Systems*. Vol. 8. 4. Taylor & Francis, pp. 233–251.

Oliver, Nuria, Bruno Lepri, Harald Sterly, Renaud Lambiotte, Sébastien Deletaille, Marco De Nadai, Emmanuel Letouzé, Albert Ali Salah, Richard Benjamins, Ciro Cattuto, et al. (2020) Mobile phone data for informing public health actions across the COVID-19 pandemic life cycle.

OpenStreetMap contributors (2017) OpenStreetMap. url: https://www.openstreetmap.org (visited on 11/30/2020).

Roger Harrabin (2020) Coronavirus: Boom time for bikes as virus changes lifestyles. In: *BBC*.

Safi, Hamid, Behrang Assemi, Mahmoud Mesbah, Luis Ferreira, and Mark Hickman (2015) Design and Implementation of a Smartphone-Based Travel Survey. In: *Transportation Research Record* 2526 (1), pp. 99–107.

Stopher, Peter, Camden FitzGerald, and Min Xu (2007) Assessing the accuracy of the Sydney Household Travel Survey with GPS. In: *Transportation* 34 (6), pp. 723–741.

Stopher, Peter R., Vivian Daigler, and Sarah Griffith (2018) Smartphone app versus GPS Logger: A comparative study. In: *Transportation Research Procedia* 32. Transport Survey Methods in the era of big data:facing the challenges, pp. 135–145.

Tirachini, Alejandro and Oded Cats (2020) COVID-19 and public transportation: Current assessment, prospects, and research needs. In: *Journal of Public Transportation* 22 (1), p. 1.

Trilla, Antoni, Guillem Trilla, and Carolyn Daer (2008) The 1918 "Spanish flu" in Spain. In: *Clinical infectious diseases* 47 (5), pp. 668–673.

Wargelin, Laurie, Peter Stopher, et al. (2012) GPS-based household interview survey for the Cincinnati, Ohio Region: executive summary report. Tech. rep. Ohio. Dept. of Transportation. Office of Research and Development.

Wolf, Jean, Michael Loechl, Miriam Thompson, and Carlos Arce (2003) Trip rate analysis in GPS-enhanced personal travel surveys. In: *Transport survey quality and innovation* 28, pp. 483–498.

Yabe, Takahiro, Kota Tsubouchi, Naoya Fujiwara, Takayuki Wada, Yoshihide Sekimoto, and Satish V Ukkusuri (2020) Non-compulsory measures sufficiently reduced human mobility in Tokyo during the COVID-19 epidemic. In: *Scientific Reports* 10 (1), pp. 1–9.

Variable	Value	N MOBIS	N MZ	% MOBIS	% MZ
Access to car	No	13	2,587	0.8	6.2
	Yes	1,480	31,824	90.4	75.8
	Sometimes	144	7,584	8.8	18.1
Age	Under 18	-	7,546	-	13.2
	(18,25]	259	5,126	16.9	9.0
	(25,35]	272	8,117	17.7	14.2
	(35,45]	314	8,789	20.4	15.4
	(45,55]	359	9,560	23.4	16.7
	(55,65]	333	7,384	21.7	12.9
	66+	-	10,568	-	18.5
Correspondence language	German	1,191	39,023	72.8	68.4
	French	446	14,450	27.2	25.3
	Italian	-	3,617	-	6.3
Education	Mandatory	96	9,954	5.9	19.3
	Secondary	758	25,594	46.3	49.5
	Higher Ed.	783	16,124	47.8	31.2
Gender	Female	828	28,928	50.6	50.7
	Male	809	28,162	49.4	49.3
Hosuehold size	1	209	19,439	12.8	34.0
	2	546	20,222	33.4	35.4
	3	331	7,400	20.2	13.0
	4	417	7,133	25.5	12.5
	5 or more	134	2,897	8.2	5.1
Income	4 000 CHF or less	92	10,139	5.6	17.8
	4 001 - 8 000 CHF	470	18,728	28.7	32.8
	8 001 - 12 000 CHF	495	9,945	30.3	17.4
	12 001 - 16 000 CHF	267	3,878	16.3	6.8
	More than 16 000 CHF	168	2,593	10.3	4.5
	Prefer not to say	143	11,807	8.7	20.7
Main employment	Employed	1,196	27,521	73.1	48.2
	Self-employed	111	4,133	6.8	7.2
	Apprentice	10	1,488	0.6	2.6
	Unemployed	59	1,423	3.6	2.5
	Student	83	2,114	5.1	3.7
	Other	110	9,421	6.7	16.5
	Retired	68	10,990	4.2	19.3
Nationality	Swiss	1,517	43,347	92.7	75.9
	Other	120	13,743	7.3	24.1

TABLE 5.1: Comparison of MOBIS-Covid19 (MOBIS) sample with the last national travel diary Mikrozensus (MZ) 2015

6 DISCUSSION AND OUTLOOK

The previous chapters in this thesis presented four pieces of work under the umbrella of mobility field experiments using GPS tracking apps. This discussion brings the papers together, discussing the main results, key limitations, and the possible directions for future work.

In the chapter 2, the methodology for the use of app-based GPS tracking in mobility studies was presented and evaluated. In answer to the research questions, such methods are indeed suitable for such experiments, and the trip and mode detection algorithms are now mature enough to reliably support a mobility pricing experiment that covers multiple transport modes. The key role that participants mobile device plays on the attrition rate was identified, which is important contribution towards future studies which may wish to use the same methods. It will also be interesting to follow this up in future studies, both with Catch-my-Day and other Apps, to see if this situation changes. From the analysis presented, no evidence was found of any manipulation of the app or trip validation interface by the pricing treatment group.

Furthermore, as all participants were subject to a four-week control period, and the control treatment group continued as so for another four weeks, more 20,000 person weeks of "untreated" travel diary data are now available for the further research of travel behaviour. This data set is already stimulating much further research, not least the understanding of mobility during the Corona pandemic presented in chapter 5.

Chapter 3 covers the design of a link-level approach to calculating the external costs of travel on GPS traces. The pipeline focuses on private car travel, taking into account the speed, route and car type of the participant, by combining various parts of the MATSim framework and Graphhopper. An analysis of the application of the pipeline to the data collected from the Catch-my-Day app shows that this method captures much more heterogeneity in the external costs than the use of normative values, while still being consistent on the average level. As mentioned in chapter 1, mobility pricing has the potential to play a key role in future of our transport networks, and the work in sections 2 and 3 are key components of the MOBIS study, which aims to further our understanding of the potential impacts of mobility pricing in Switzerland. Since the completion of this work, the pipeline has been applied in the nationwide mobility pricing experiment referenced in chapter 2, and combined with zone-based public transport peak-hour charging and per-km external costs for non-car modes, forming the basis of a mobility pricing scheme in the MOBIS experiment. The pricing scheme covers all main modes of transport, incentivises active modes and public transport over car travel, and should encourage both modal and departure time shifts. The analysis of the treatment effects in experiment is still ongoing.

Developing models with the amount data collected from such a large tracking experiment also requires new tools, and to this end, *mixl* was developed for the R programming environment. Chapter 4 details the development and use of this new package, which has already been applied in multiple pieces of published research (Marra and Corman, 2020; Schmid, 2019; Wicki et al., 2019), and is used as a teaching tool in the course "measurement and modelling" at the Institute for Transport Systems and Planning, ETH Zurich. Arguably more importantly, it has provided a performance benchmark for discrete choice estimation software in R, which has motivated steady performance improvements in other discrete choice estimation package for R, namely, Apollo. Apollo of course provides a much larger modelling toolkit than *mixl*, although the way the utility functions are specified is roughly similar. Valuable future work would be to integrate the two packages, so that the performance benefits of *mixl* are available to Apollo users in cases where their models are supported. While the performance of *mixl* is impressive, there is still room for further reductions in the estimation time. Further gains could be made through the implementation of symbolic differentiation, and the adaptation of *mixl* to support GPUs (graphical programming units), which contain hundreds of cores on a chip. The data parallel problem structure of mixed multinomial logit structures is perfectly suited to GPU architecture, and speed improvements of an order of magnitude would be feasible.

Finally, chapter 5 presents a first analysis of mobility behaviour in the first months of the Covid-19 pandemic in Switzerland, using the Catch-my-Day app and the MOBIS panel of respondents. This work demonstrates the power of the GPS tracking methods to capture unexpected changes in mobility behaviour. It took only a few days to invite the MOBIS participants to reactive the app, and as they were already familiar with the technology, and the data pipelines were still in place, tracking could begin immediately. Furthermore, as the original MOBIS tracking data was available from Autumn 2019, a baseline could be established against which to measure the impacts of the restrictions on mobility in Switzerland. A report of the mobility was provided online weekly, in English, German and French. This work became an important reference point both in Switzerland and more widely across Europe. The results were reported widely in the Swiss media throughout 2020. It was also utilised across the transport industry, from local transport authorities, to consulting companies and the automotive industry in understanding the impacts of the pandemic. For a list of the media coverage, see chapter 6. The MOBIS-Covid19 data is also being used to develop mode choice models (utilising *mixl*) for the Swiss railways (SBB) to incorporate the changes in passenger behaviour and the shift towards telework (also known as working-from-home or remote-work) in their in-house planning models.

The identification of the cycling boom in chapter 5 was also a surprising result, which was only possible due to the availability of baseline data from 2019 MOBIS study. If the MOBIS-Covid19 tracking had started from null at the beginning of the pandemic, it is likely that the capture of the cycling boom in the data would have been missed, or only identified much later on comparison with other data sources. The data on cycling in the dataset is still largely under-utilised, but there is still exciting work to be done on the behavioural influences on cycling using the MOBIS-Covid19 data. It is also worth mentioning that the MOBIS-Covid19 data collection is still on going, and providing insights into the behavioural responses during the Second wave in Autumn 2020.

6.1 OUTLOOK

The pandemic has clearly accelerated the shift towards telework that was predicted to take many more years to occur. Indeed, tech companies such as Facebook and Twitter now allow employees to work from home permanently, and Facebook's CEO anticipates that half of Facebook's workforce will permanently work from home by 2030 (Sandler, 2020). It will be particularly interesting to see how this develops in the near future, and how it shapes our transport systems. It is feasible that for many office based workers, their residential location choice will no longer be constrained by a necessary proximity to the workplace. This could lead to a restructuring of our cities, and a flattening of the peak-hour demand curves. On the other-side, the lack of peak-hour commuting flows on public transport may have drastic consequences for the financial viability of such systems.

Already in 2020, the SBB has reported a loss of over 479 Million in the first half of 2020 as a result of the pandemic (Burroughs, 2020).

More work needs to be undertaken to explore how best to adapt our transport systems with a post-pandemic world, and further analysis on data collected during this thesis will play a key role here. The results will apply not only to the Swiss context but more widely, as many of the insights will be potentially transferable to comparable regions, especially around Europe.

As mentioned in the introduction, road pricing schemes using GPS are already being implemented in parts of the world (Tan, 2020), and practical implementations have been shown to be effective in London (Leape, 2006), Copenhagen (Nielsen, 2004) and Stockholm (Eliasson et al., 2009), among others. However, where GPS devices has been used to implement or monitor such a system, they have been installed in the car. Designing a *mobility*-pricing scheme with GPS, covering all transport modes, as opposed to just a *road*-pricing scheme, would potentially require almost complete surveillance of the population, either through use of a tracking app, or a combination of GPS technology and smartcards (such as those already common in public transport networks). In some respects, we are already living in such a dystopian future, where our economic system is built around the surveillance of population and monetisation of their data (Zuboff, 2019). It is possible that the reluctant acceptance of this new reality might make the introduction of mobility-pricing schemes more acceptable, if users of the system are convinced that they would overall be better off, as has been the promise of dynamic transport pricing since its conception.

There is also is the possibility to both replicate and extend the results of works on human mobility scales, such as those by (Gonzalez, Hidalgo, and Barabasi, 2008) and Alessandretti, Aslak, and Lehmann (2020). Such "laws" of mobility have been applied to both large datasets of hundreds of thousands of participants, and smaller ones of 1000 students. Here, the long-term nature of the MOBIS dataset may be useful, where hundreds of the participants have been tracking for well over a year. In particular, these models could be used to expore the shifts to telework, and the influence of household size on the scales of mobility. Such contributions have the potential to improve our transport models, by improving our understanding of how mobility patterns vary across different segments of the population.

There are still limitations to the methods, most of which were identified in the previous chapters. In particular, the effectiveness of tracking apps is at the mercy of the mobile phone manufacturers. Certain default settings, such as those on Huawei Android phones can render the tracking Apps almost useless unless they are changed, and such requirements are only becoming more common as device makers strive for longer battery life and better performance. Furthermore, the difference in attrition between iOS and Android may make GPS-based mobility studies less feasible in regions with a higher share of cheaper Android devices, such as developing countries. During the course of the MOBIS study, Apple released a new software update that completely redesigned their location privacy framework. This could have been a death-knell for the study, but fortunately it was handled well by the app developers, Motiontag. It did however, emphasise the fragility inherent in GPS tracking studies, which by necessity rely on a large stack of technology provided by third-parties. It also needs to be acknowledged that the usefulness of GPS tracking in urban areas with large underground transport networks may be less effective.

It was proposed in the introduction that, after decades of reliance on paper and telephone based methods for mobility studies and travel diary collection, the field is experiencing an inflection point. While the potential applications for GPS tracking were previously clear, the widespread application was hindered by a reliance on physical GPS-receivers and a lack of methods for processing the data. However, with the use of mobile phone tracking apps, which are easy to install and run in the background, such as Catch-my-Day, coupled with the accuracy of the machine learning methods for identifying trips and modes, the stage is set for these GPS data to play a ever more central role in transport planning.

REFERENCES

Alessandretti, Laura, Ulf Aslak, and Sune Lehmann (2020) The scales of human mobility. In: *Nature* 587 (7834), pp. 402–407.

Burroughs, David (2020) SBB records SFr 479m loss as passenger numbers drop. URL: `https://www.railjournal.com/financial/sbb-records-sfr-479m-loss-as-passenger-numbers-drop` (visited on 11/24/2020).

Eliasson, Jonas, Lars Hultkrantz, Lena Nerhagen, and Lena Smidfelt Rosqvist (2009) The Stockholm congestion–charging trial 2006: Overview of effects. In: *Transportation Research Part A: Policy and Practice* 43 (3), pp. 240–250.

Gonzalez, Marta C, Cesar A Hidalgo, and Albert-Laszlo Barabasi (2008) Understanding individual human mobility patterns. In: *nature* 453 (7196), pp. 779–782.

Leape, Jonathan (2006) The London congestion charge. In: *Journal of Economic Perspectives* 20 (4), pp. 157–176.

Marra, Alessio Daniele and Francesco Corman (2020) Determining an efficient and precise choice set for public transport based on tracking data. In: *Transportation Research Part A: Policy and Practice* 142, pp. 168–186.

Nielsen, Otto Anker (2004) Behavioral responses to road pricing schemes: Description of the Danish AKTA experiment. In: *Intelligent Transportation Systems*. Vol. 8. 4. Taylor & Francis, pp. 233–251.

Sandler, Rachel (2020) Half Of Facebook's Employees May Permanently Work From Home By 2030, Zuckerberg Says. URL: `https://www.forbes.com/sites/rachelsandler/2020/05/21/half-of-facebooks-employees-may-permanently-work-from-home-by-2030-zuckerberg-says` (visited on 11/20/2020).

Schmid, Basil (2019) Connecting Time-Use, Travel and Shopping Behavior: Results of a Multi-Stage Household Survey. PhD thesis. ETH Zurich.

Tan, Christopher (2020) New ERP system to start in 2023 but no distance-based charging yet; replacement of IU from second half of 2021.

Wicki, Michael, Sergio Guidon, Felix Becker, Kay Axhausen, and Thomas Bernauer (2019) How technology commitment affects mode choice for a self-driving shuttle service. In: *Research in Transportation Business & Management* 32, p. 100458.

Zuboff, Shoshana (2019) The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power. Profile Books.

CURRICULUM VITAE

PERSONAL DATA

Name	Joseph Anthony Molloy
Date of Birth	April 3, 1990
Place of Birth	Sydney, Australia
Citizenship	Australia, United Kingdom, Ireland

EDUCATION

2017 – 2021	Eidgenössisches Technische Hochschule, Zürich, Switzerland
2014 – 2017	Technical University Munich, Germany *Final degree:* M.Sc Transport Systems
2009 – 2013	University of New South Wales, Australia (UNSW) *Final degree:* Bachelor of Engineering (Software), Honours *Awards:* 2010, 2011, 2012 – Dean's Honours List 2010 – CiSRA 2nd Year Prize for Excellence in Computing Studies
2001 – 2008	St Ignatius College, Riverview, Australia Higher School Certificate

EMPLOYMENT

March 2017 –	Research Assistant *Institute for Transport Planning and Systems*, Zurich, Switzerland
Dec 2015 – Dec 2017	Student Researcher *Technical University Munich*, Munich, Germany

May 2014 – Sept 2014	Graduate Consultant (Enterprise Intelligence team) *Ernst & Young*, Sydney, Australia
Dec 2012 – Feb 2013	Vacationer (Enterprise Intelligence team), 12 week placement *Ernst & Young*, Sydney, Australia
Feb 2012 – Aug 2013	Software Development Intern, JIRA product team, 6 month placement *Atlassian*, Sydney, Australia
2012 Academic Year	Academic Tutor *New College, UNSW*, Sydney, Australia
Jun 2010 – Dec 2011	Academic Teaching Assistant and Mentor *School of Computer Science and Engineering, UNSW*, Sydney, Australia

PUBLICATIONS

ARTICLES IN PEER-REVIEWED JOURNALS

Llorca, Carlos, Joseph Molloy, Joanna Ji, and Rolf Moeckel (2019) Estimation of a long-distance travel demand model using trip surveys, location based big data and trip planning services. In: *Transportation Research Record* 2672 (47), pp. 103–113.

Molloy, Joseph and Rolf Moeckel (2017a) Automated design of gradual zone systems: An iterative algorithm to design optimally sized spatial zones suitable for spatial modeling, while respecting municipal boundaries. In: *Open Geospatial Data, Software and Standards* 2.

Molloy, Joseph and Rolf Moeckel (2017b) Improving destination choice modeling using location-based big data. In: *ISPRS International Journal of Geo-Information* 6, p. 291.

Molloy, Joseph, Christopher Tchervenkov, Beat Hintermann, and Kay W. Axhausen (2020) Tracing the Sars-CoV-2 impact: The first month in Switzerland - March to April 2020. In: *Transport Findings*.

CONFERENCE CONTRIBUTIONS

Hörl, Sebastian and Joseph Molloy (2020) A ticket-based public transport pricing model for Switzerland. In: *20th Swiss Transport Research Confernce (virtual)*.

Molloy, Joseph (2017) Using high resolution passive mobile positioning data in transport models: Aims, challenges and current progress. In: *Workshop on data processing and analytics of smartphone and GPS data*. Tartu, Estonia.

Molloy, Joseph and Kay W. Axhausen (2018) Microsimulation of time variant road pricing for Kanton Zug using MATSim. In: *18th Swiss Transport Research Conference*. Ascona, Switzerland.

Molloy, Joseph and Kay W. Axhausen (2020a) Generating Synthetic mobile phone datasets using MATSim. In: *20th Swiss Transport Research Confernce (virtual)*.

Molloy, Joseph and Kay W. Axhausen (2020b) MOBIS study: A review of common reported issues. In: *20th Swiss Transport Research Confernce (virtual)*.

Molloy, Joseph, Alberto Castro Fernández, Thomas Götschi, Beaumont Schoeman, Christopher Tchervenkov, Uros Tomic, Beat Hintermann, and Kay W. Axhausen (2021) A national-scale mobility pricing experiment using GPS tracking and online surveys in Switzerland: Response rates and survey method results. In: Washington, D.C.

Molloy, Joseph, Thomas Schatzmann, Beaumont Schoeman, Christopher Tchervenkov, Beat Hintermann, and Kay W. Axhausen (2021) Observed impacts of COVID-19 on travel behaviour in Switzerland based on a large GPS panel. In: Washington, D.C.

Molloy, Joseph, Basil Schmid, and Felix Becker (2019) mixl: An open-source R package for estimating complex choice models on large datasets. In: *19th Swiss Transport Research Confernce*. Ascona, Switzerland.

Molloy, Joseph, Siiri Silm, Rein Ahas, and Kay W. Axhausen (2018a) Comparison of passive mobile traces and GPS data for the calculation of mobility indicators. In: *15th International Conference on Travel Behavior Research (IATBR 2018)*. Santa Barbara, CA, USA.

Molloy, Joseph, Siiri Silm, Rein Ahas, and Kay W. Axhausen (2018b) Mobile Tartu 2018: Mobile Data, Geography, LBS. In: *15th International Conference on Travel Behavior Research (IATBR 2018)*. Tartu, Estonia.

Molloy, Joseph, Christopher Tchervenkov, and Kay W. Axhausen (2018a) Berechnen der Externalitäten von GPS Tracks mittels Anwendung von Mikrosimulation. In: *39. Universitätstagung Verkehrswesen 2018*. Obergurgl, Austria.

Molloy, Joseph, Christopher Tchervenkov, and Kay W. Axhausen (2018b) Estimating the externalities of a sustainable mobility platform using GPS traces. In: *mobil.TUM 2018 - Urban Mobility: Shaping the Future Together*. Munich.

Schmid, Basil, Joseph Molloy, Simona Jokubauskaite, Florian Aschauer, Stefanie Peer, Reinhard Hoessinger, Regine Gerike, Sergio R. Jara-Diaz, and Kay W. Axhausen (2019) Individual-level decomposition of the value of travel time savings into the value of leisure and the value of time assigned to travel. In: Budapest, Hungary.

Tchervenkov, Christopher, Joseph Molloy, and Kay W. Axhausen (2018) Estimating externalities from GPS traces using MATSim. In: *18th Swiss Transport Research Confernce*. Ascona, Switzerland.

invited presentations

Molloy, Joseph (2020a) Führt die Coronakrise zu einer neuen Mobilitätslandschaft in der Schweiz? 13. BMM-Tag 2020, Vaduz, Liechtenstein. Vaduz, Lichtenstein.

other contributions

Galleguillos, Marcello, Adrienne Grêt-Regamey, Joseph Molloy, and Kay W. Axhausen (2020) Are our cities green enough in times of a pandemic? url: `https://www.nsl.ethz.ch/are-our-cities-green-enough-in-times-of-pandemics/` (visited on 09/30/2020).

Molloy, Joseph (2018) Connecting large scale transport models and mobility trace data. url: `https://www.research-collection.ethz.ch/handle/20.500.11850/334287` (visited on 09/14/2020).

Molloy, Joseph (2020b) Mobilitätsverhalten vor/nach/während Covid-19. url: `https://www.nsl.ethz.ch/mobilitaetsverhalten-vor-nach-waehrend-covid-19/` (visited on 06/29/2020).

Molloy, Joseph (2020c) Mobility behaviour in Switzerland: Coronavirus study. url: `https://www.nsl.ethz.ch/en/mobility-behaviour-in-switzerland-coronavirus-study/` (visited on 06/01/2020).

in the media

Balmer, D. (2020) Corona transforms Switzerland into a bike country. In: *Tages Anzeiger*.

Fritz, D. (2020) Wie Corona die Mobilität verändert. In: *Liechtensteiner Volksblatt*.

Gross, N. (2020) Grundvoraussetzung für die Zukunft. In: *Liechtensteiner Vaterland*.

Müller, A. (2020) Angst vor dem «Peak Zürich»: Setzt das Coronavirus der langen Hochphase der Grossstadt ein Ende? url: `https://www.nzz.ch/zuerich/coronavirus-in-zuerich-droht-wegen-der-pandemie-der-exodus-ld.1565780`.

Platter, M. (2020) Das Virus bewegt aufs Velo. In: *Neue Zürcher Zeitung*.

Radio Liechtenstein (2020) Unternehmen in die Verantwortung nehmen, interview mit Joseph Molloy.

Radio SRF (SRF4 News) (2020) Interview with Joseph Molloy about mobility regulations and changes during the first and the second wave.

Schenk, M. (2020) Mobility-Pricing: Die Suche nach dem flüssigsten System. In: *Automobil Revue* (17), pp. 2–3.

Schuler, C (2020) Mehr Velos unterwegs – das ist ein «riesiges Potenzial. URL: `https://www.zentralplus.ch/mehr-velos-unterwegs-das-ist-ein-riesiges-potenzial-1862219/`.

Siegrist, P. (2020) Wie der öffentliche Verkehr genesen soll. In: *Tages Anzeiger*.

Zeller, R (2020) Comeback der Strasse. URL: `https://www.weltwoche.ch/ausgaben/2020-35/diese-woche/comeback-der-strasse-die-weltwoche-ausgabe-35-2020.html`.