Activity Generation and Scheduling for Travel Demand Models to Account for Telecommuting Behavior

Zur Erlangung des akademischen Grades einer

DOKTORIN DER INGENIEURWISSENSCHAFTEN
(Dr.-Ing.)

von der KIT Fakultät für
Bauingenieur-, Geo- und Umweltwissenschaften
des Karlsruher Instituts für Technologie (KIT)

genehmigte

DISSERTATION

von

Anna Sophie Reiffer, M.Sc.
geb. in Karlsruhe

Tag der mündlichen Prüfung:	23. Juli 2024
Hauptreferent:	Prof. Dr.-Ing. Peter Vortisch
Korreferent:	Prof. Dr.-Ing. Kay W. Axhausen
Korreferent:	Prof. Dr. rer. nat. Stefan Nickel

Karlsruhe 2024

This document is licensed under a Creative Commons
Attribution-Non Commercial 4.0 International License (CC BY-NC 4.0):
https://creativecommons.org/licenses/by-nc/4.0/deed.en

Abstract

The impact of telecommuting on travel demand has been the subject of debate for almost five decades. The assumption, that working from home axiomatically reduces travel demand as individuals commute less into the office, is flawed. Telecommuters can leverage the flexibility gained by working from home to conduct non-work activities, including, for example, leisure activities as well as dropping off children at childcare. Additionally, telecommuting is often conducted only on some days of the week, which means that any travel demand reduction effects might not be spread evenly across all days of the week. It remains unclear to which extent these effects impact travel demand within a transport system. Although activity-based models are theoretically well-suited for evaluating this relationship, including telecommuting behavior adequately requires a model to consider (i) all activities simultaneously, (ii) household interactions, and (iii) day-to-day variability. To date, none of the previously presented activity-based approaches account for all three of these features.

To address this gap, this thesis presents an activity generation and scheduling model that accounts for all three of the listed requirements. The model development is informed by detailed empirical analyses of data from the German Mobility Panel regarding the impact of telecommuting on activity patterns, focusing on different household configurations and roles within the household. A comprehensive dataset is generated that is used for the development of the scheduling model.

The model presented in this work considers all activities at the same time to allow for trade-offs between telecommuting and other planned activities. It generates schedules for one week considering household interaction both at the activity generation as well as scheduling phases of the model.

Zusammenfassung

Der Einfluss von Telearbeit auf die Verkehrsnachfrage wird seit fast fünf Jahrzehnten diskutiert. Die Annahme, dass das Arbeiten von zu Hause aus automatisch die Verkehrsnachfrage reduziert, da weniger Personen ins Büro pendeln, ist zu vereinfacht. Die Arbeit im Homeoffice erhöht die Flexibilität bei der Aktvititätenplanung, sodass zusätzliche, nicht-arbeitsbezogene Aktivitäten durchgeführt werden können, wie zum Beispiel Freizeitaktivitäten oder das Bringen von Kindern zur Kinderbetreuung. Außerdem arbeiten viele nur an einigen Tagen der Woche im Home-Office,wodurch etwaige Reduktionseffekte in der Verkehrsnachfrage möglicherweise nicht gleichmäßig über alle Wochentage verteilt sind. Es bleibt unklar, in welchem Ausmaß diese Effekte die Verkehrsnachfrage innerhalb eines Verkehrssystems beeinflussen. Obwohl aktivitätsbasierte Modelle theoretisch gut geeignet sind, um diese Einflüsse abzubilden und zu bewerten, erfordert die angemessene Berücksichtigung von Telearbeit, dass ein Modell (i) alle Aktivitäten gleichzeitig betrachtet, (ii) Haushaltsinteraktionen berücksichtigt und (iii) die tägliche Variabilität berücksichtigt. Bisher schließt keines der bisher vorgestellten aktivitätsbasierten Ansätze alle drei dieser Merkmale bei der Erstellung von Aktivitätenplänen ein.

Um diese Lücke zu schließen, wird in dieser Arbeit ein Modell zur Aktivitätserzeugung und -planung vorgestellt, das alle drei genannten Anforderungen berücksichtigt. Die Modellentwicklung basiert auf detaillierten empirischen Analysen des Deutschen Mobilitätspanel mit Fokus auf die Auswirkungen von Telearbeit auf Aktivitätsmuster, unter Berücksichtigung unterschiedlicher Haushaltskonfigurationen und Rollen innerhalb des Haushalts. Es wird ein umfassender Datensatz erstellt, der für die Entwicklung des Modells verwendet wird.

Das in dieser Arbeit vorgestellte Modell berücksichtigt alle Aktivitäten gleichzeitig, um Abwägungen zwischen Telearbeit und anderen geplanten Aktivitäten zu ermöglichen. Es erstellt Zeitpläne für eine Woche, wobei sowohl bei der Aktivitätserzeugung als auch in den Planungsphasen des Modells die Haushaltsinteraktionen berücksichtigt werden.

Acknowledgments

Writing and defending a doctoral dissertation is very much like climbing a mountain. At the start, it might seem like a quaint and fun adventure. But when the inevitable fog rolls in and you are trying to plow through knee-deep snow fields, it calls for outside support. In the end, I had to climb the mountain that was this dissertation myself, but I would not have gotten there without the support of many "sherpas", who have supported my PhD journey during my time as a research assistant at the Institute for Transport Studies (IfV) at KIT.

First and foremost, I would like to thank my supervisor Prof. Peter Vortisch. Thank you for the discussions on commercial transport that later shifted to activity-based models, in which we quickly agreed that having a cat around is a good enough reason to work from home. I owe Peter an additional dollop of gratitude as he supported me through the next step of the academic career whirlwind when it came to postdoc applications, for which he wrote many a reference letters, and shared my excitement when I landed the gig at ETH Zurich.

I thank my second supervisors, Prof. Kay W. Axhausen and Prof. Stefan Nickel. Prof. Axhausen, thank you for your valuable feedback and practical advice on my dissertation, and for your support during my postdoc search. Stefan, thank you for saving the day by helping me find my footing in optimization and operations research without ever making me feel out of place.

I thank the KIT House of Young Scientists for awarding me the KHYS Research Travel Grant that made my visit to the University of Leeds possible. I'm incredibly grateful to Prof. Jillian Anable, Prof. Greg Marsden, Dr. Llinos Brown, and Dr. Chris Bretter for making my time at the Institute for Transport studies so memorable.

I want to thank my colleagues at IfV for an unforgettable time. A special shoutout to Michael Heilig for supervising my master's thesis and co-authoring my first ever paper, and to Sebastian Buck, who's not only shaped the institute profoundly but also my caffeine addiction. To my PhD buddies: Tim Wörle, the calm amidst any storm and the best office mate—as long as we agreed on the temperature; Miriam Magdolen, my academic (travel) companion, be that on strolls through Georgetown or kitten-supervised hikes through the Black forest; to Claude Weyland, for always lending an ear during impromptu 308 office coffee breaks; and Nadine Kostorz-Weiss for always being incredibly attentive even from miles away. A big thanks to Team TOP Sonstiges, Lukas Barthelmes and Emre Görgülü, for all the morning coffees and late night kitchen talks at conferences. To Lisa Ecke, for fielding my many German Mobility Panel questions; Jelle Kübler, for always staying calm when project chaos reigned and the collaboration on logiTopp; and Kim Kandler, for cat drawings and surprise pick-me-ups during crunch time.
Many thanks go out to my friends: Shanna, Vicky and Maxi, Markus, Simon, and Jerónimo, for pulling me out of the academic bubble with pub quizzes, game nights and mountain trips; and Yannik and Nils, Tara, and my best friend and sister Lisa for much needed iaro coffee dates, farmers market strolls, and evenings of cooking. My biggest thanks goes out to my parents who have given me a life full of chances, never pushed me but always showed me that hard work pays off.
Last, but by no means least, my deepest gratitude goes to my husband, Dominik, who bore witness to the messiest parts of this PhD journey; who helped me focus when everything got foggy, lifted me back up when my health faltered, took care of our life when I could only function on auto-pilot—and through it all, never once complained or asked me to scale back my ambitions. Finally, an honorable mention goes out to my cats, Janosch and Findus, who would repeatedly blop onto my keyboard to demand belly rubs and thus forced me to take much needed breaks.

Karlsruhe in October of 2024
Anna S. Reiffer

Bypasses are devices which allow some people to dash from point A to point B very fast. People living at point C, being a point directly in between, are often given to wonder what's so great about point A that so many people from point B are so keen to get there. **They often wish that people would just once and for all work out where the hell they want to be.**

Douglas Adams, *The Hitchhiker's Guide to the Galaxy*

Contents

Abbreviations and Symbols

Abbreviations

ABIT	Activity-Based Incremental Transport Model
ADAPTS	Agent-based Dynamic Activity Plannign and Travel Scheduling
Albatross	A Learning Based Transportation Oriented Simulation System
AME	Average Marginal Effect
AMOS	Activity Mobility Simulator
API	Application Programming Interface
CEMDAP	Comprehensive Econometric Microsimulator for Daily Activity-Travel Patterns
C-TAP	Continuous Targetbased Activity Planning
CT-RAMP	Coordinated Travel - Regional Activity Modeling Platform
DAG	Directed Acyclic Graph
Feathers	Forecasting Evolutionary Activity-Travel of Households and their Environmental RepercussionS
FN	False Negative
FP	False Positive
HAPP	Household Activity Pattern Problem

HTS Household Travel Survey

ILP Integer Linear Program

ISA Individual specific alternative

MCC Matthew's correlation coefficient

MCMC Markov Chain Monte Carlo

MDCEV Multiple Discrete Continuous Extreme Value

MDCEV-OP Multiple Discrete Continuous Extreme Value with Ordered Preferences

MDCNEV Multiple Discrete Continuous Nested Extreme Value

METiS Multiagent Estimation of Time Use and Scheduling

MNL Multinomial Logit

MOOP Multi-Objective Optimization Problem

MOP German Mobility Panel

NHTS National Household Travel Survey

OASIS Optimisation-based Activity Scheduling

PCATS Prism-Constrained Activity Travel Simulator

PDPTW Pickup and delivery problem with time windows

RFE Recursive Feature Selection

SEM Structural Equation Model

SHTS Stuttgart household travel survey

SMOTE Synthetic Minority Over-sampling Technique

SVM Support Vector Machines

TCM Transport Control Measures

TN True Negative

TP True Positive

VSURF Variable Selection Using Random Forests

wfh Work From Home

1 Introduction

Telecommuting has been considered a policy measure to reduce travel demand for over 50 years. Nilles et al. (1976) coined the phrase "telecommuting" as the act of working from home using telecommunication technology instead of commuting to work. With every new technological advance, the hope that telecommuting could reduce travel increased. Indeed, the rise of the personal computer and the expansion of internet connectivity removed most of the friction of working from home. Although not every job can be performed remotely, the shift from an industrialized society to a service society in many countries has created favorable conditions for adopting telecommuting. However, until the COVID-19 pandemic, adoption rates of teleworking remained lower than predicted. Even when disregarding that not every type of work is suitable for telecommuting, the issues remain that employers may not allow their employees to work from home and that even if someone is allowed to work from home, they may choose not to do so.

This changed considerably with the COVID-19 pandemic as many governments used working from home as a policy measure to reduce the spread of the virus both during the commute and in the office (Hale et al. 2021). The unparalleled conditions imposed by the pandemic forced employers and employees to disregard any adverse attitudes toward working from home and make telecommuting work. When asked about telecommuting after the pandemic, many employees attested they would like to keep working from home for at least part of the week (de Haas et al. 2020).

Indeed, working from home is one of the pandemic-induced changes which is said to be here to stay. This has re-ignited discussions about telecommuting as a policy measure to reduce travel. In 2021, the German Federal Ministry for Digital and Transport included telecommuting as one of six measures in their emergency program to meet the climate targets set out in the German Climate Change Act (Bundesministerium für Digitales und Verkehr 2022).

The rationale behind this measure is straightforward: when fewer people commute to their workplace and work from home instead, they conduct fewer trips and decrease their traveled mileage. However, there has been an ongoing debate on whether rebound effects nullify or even outweigh the changes in commuting patterns, as telecommuting allows for more temporal and spatial flexibility. Spatial flexibility enables employees to live further away from their work location as the commute becomes less constraining. Temporal flexibility entails that telecommuters can travel outside peak hours and reinvest the time they would usually spend commuting into other activities. While there is an extensive body of research on travel behavior implications of telecommuting, it remains unclear how these effects translate into changes in travel demand on the transport system as only a few studies have regarded telecommuting in travel demand models.

Models are a representation of the real world which take a set of input data and process them to generate information about the system of interest. In the case of travel demand models, the system of interest is the transport system within a defined study area. The output of a travel demand model are the traffic volumes on the network within the study area. Land use data of the study area and a representation of the population living within that area serve as the input of a travel demand model. Travel demand models can be differentiated by the level of aggregation of the considered population and subsequently the output of the model. In macroscopic models, travelers are aggregated into groups of homogenous behavior or market segments, e.g. employed people with cars. Alternatives to this are disaggregated or microscopic models which regard travelers as individuals. Travel demand models can further be differentiated by the demand generation process into trip-based and activity-based models. Although trip-based models are still widely used, especially in practice, they suffer from the drawback that they are not based on the underlying cause of traffic: people's desire to participate in activities that often require them to travel from one place to another. This problem is mitigated by activity-based models in which the travelers' activity patterns are estimated and serve as the cause for travel within the model. While the trip-based approach has been implemented both macroscopically and microscopically, activity-based models are generally disaggregated.

Activity-based travel demand models have evolved considerably over the last decades. They can be differentiated into rule-based and econometric models. While rule-based models are easier to implement, they are limited regarding behavioral realism as the travelers' behavior in the model is solely based on hard-coded rules. In econometric-based approaches, the behavior is based on the theory of utility maximization. The process of generating and scheduling activities is often split into a sequence of choices. This implies a hierarchy between different activity purposes which is imposed solely at the researcher's discretion. Work activities are usually considered early on in the process of activity generation and scheduling as they often pose the largest constraints within a schedule. This influences, how many other activities are chosen and when they are conducted. While this used to be appropriate when most workers were limited to working in the office, telecommuting offers a level of flexibility that requires a more balanced choice of activities. Sequential models would be able to account for this flexibility leading to more or longer non-work activities. However, the opposed effect of a person choosing to work from home *because* they want to conduct more or longer leisure activities cannot be accounted for. In this case, the simultaneous trade-off between different alternative activities is needed, which can be accounted for in Multiple Discrete Continuous Extreme Value (MDCEV) models. In the MDCEV approach, a combination of goods and the level of consumption is considered. This approach has been applied to activity time-use models, in which the combination of activities and the time invested in them is modeled in one step. Most recent developments even consider activity episodes and the order they are conducted, essentially generating and scheduling activities in one step. However, these approaches only consider single-day activity schedules and individual decision-makers. This limits the applicability when accounting for telecommuting behavior. Firstly, telecommuting requires us to consider multiple days as not everyone works remotely every day and travel demand effects could differ throughout the week. Secondly, rebound effects do not necessarily occur on the day one telecommutes but could be observed e.g. at the weekend. Thirdly, telecommuting offers levels of flexibility which may lead to reorganization of household responsibilities in which case the individual-level approach is not sufficient.

1.1 Thesis aims and contributions

The overarching goal of this dissertation is to enhance the precision and applicability of activity-based travel demand models that allow for the representation of telecommuting behavior. Thereby contributing to more effective travel demand management policies. Specifically, the aims of this thesis are:

1. **Understanding the interdependence between telecommuting and activity patterns**: This thesis first seeks to establish a comprehensive understanding of the relationship between telecommuting and daily activity patterns. By reviewing existing literature and conducting empirical analyses, it aims to identify how telecommuting influences the timing, frequency, and type of daily activities, as well as possible household interaction effects.

2. **Generation of a comprehensive data set for model development**: Building on the insights gained from the interdependence analysis, this research aims to generate the requisite datasets that capture detailed aspects of telecommuting and related activity patterns. This includes data on household interactions, telecommuting frequencies, and associated activity characteristics.

3. **Developing a scheduling model for activity-based simulation**: The core contribution of this thesis is the development of a sophisticated scheduling model that accurately generates activity schedules over a one-week period, taking into account household interactions. This model distinguishes itself by incorporating telecommuting behavior not only as an isolated activity but as an integral part of the daily activity pattern of agents in the model.

The model developed in this thesis provides a robust foundation for further research and practical application. By employing the developed model, future studies can simulate and analyze the impact of telecommuting on travel demand, providing empirical evidence on how telecommuting can serve as an effective travel demand management policy, potentially leading to reduced traffic congestion and environmental impacts. This outlook sets the stage for future explorations into the efficacy of work-from-home measures within urban travel demand management strategies.

Definition of telecommuting in this thesis

There is no generally accepted definition of telecommuting. In this thesis, telecommuting is defined as the act of working from home.[1] The terms *telecommuting, telework*, and *work from home* are used interchangeably throughout this thesis. In tables and figures, it is also often abbreviated as *wfh*.

1.2 Thesis plan

The rest of this thesis is structured into the following chapters:

Chapter 2 provides the necessary background to the work in this thesis. Firstly, a literature overview of research on telecommuting and its impact on activity patterns as well as the interdependence between these patterns, the choice to work from home and household interactions is provided. Subsequently, previous approaches to activity generation and scheduling are presented. The literature section concludes with an evaluation of previously presented modeling approaches concerning their capability of accounting for telecommuting behavior. Secondly,

[1] Please note that telecommuting or telework can refer to working from locations other than the traditional office, not just from home. The work in this thesis focuses on working from home due to the limited availability of data for other types of telecommuting.

the main methods applied in the activity generation and scheduling model are introduced, detailing the econometric as well as the optimization methods.

The subsequent Chapter (3), introduces the data sources used in this thesis. The model developed is based on two surveys which are presented first. Subsequently, necessary data preparation steps are detailed, specifically the steps to transform travel diary data into activity schedules. Further, the integration of telecommuting-relevant information into the data set is detailed, including an approach to predict telecommuting engagement in travel diary survey data is presented and evaluated.

Chapter 4 includes empirical analyses of the data described in the previous chapter. The analysis focuses on the relationship between telecommuting and activity patterns, both in the general population of employed individuals as well as by different household types, and household roles.

Subsequently, the modeling framework developed in this work is introduced in Chapter 5. First, an overview of the framework is presented, including how the framework is placed within agent-based travel demand simulation. Subsequently, the steps of the model are presented in detail. The validity of the model approach and its efficacy concerning the representation of telecommuting behavior is analyzed based on the application of the model. The chapter concludes with a discussion of the presented framework, including the limitations of the approach.

Finally, Chapter 6 offers conclusions on the thesis. The main findings of the thesis are summarized and core contributions are presented. Lastly, possible avenues for future work are derived based on the findings and model presented in this thesis.

2 Background

> Der Nutzen ist das große Idol der Zeit, dem alle Kräfte frohnen und alle Talente huldigen sollen.
>
> Friedrich Schiller

2.1 Literature Review

Research on the relationship between telecommuting and travel behavior dates back decades, with Nilles et al. (1976) first exploring to what extent work from home can be used as a travel demand management measure in the 1970s. The majority of these studies are focused on trip-related attributes, vehicle or personal miles traveled. In order to adequately account for telecommuting in travel demand models, we need to take a step back and look at the activities that cause people to (not) travel in the first place. The following section thus presents an overview of studies conducted on the impact of telecommuting on activity patterns. Subsequently, activity-based models are reviewed and their features are explored. The literature review chapter concludes with an evaluation of these activity-based models concerning their capabilities in accounting for telecommuting behavior.

2.1.1 Impact of telecommuting on activity patterns

Studies on the impact of telecommuting on activity patterns mostly focus on the impact of working from home on other activities (and in some cases vice versa).

Although not entirely conclusive, most studies find evidence that telecommuting is associated with increased non-work activity participation. Exploring a Regional Household Travel Survey conducted in New York, New Jersey, and Connecticut in 2010 and 2011 using a Structural Equation Modelling (SEM) approach, Asgari et al. find that telecommuting is associated with an increase in nonmandatory activity participation, which is the highest for full-day telecommuters indicating that the flexibility gained through telecommuting is leveraged to participate in non-work activities (Asgari et al. 2016, Asgari and Jin 2017). Their analysis further indicates a reciprocal effect in that participating in non-work activities such as shopping, maintenance, and discretionary activities, increases the propensity to telework and decreases the propensity to commute to the office. These findings are corroborated by Paleti and Vukovic (2017) who analyze data from the 2009 US National Household Travel Survey (NHTS) using a Poisson model to analyze telecommuting frequency and a Multiple Discrete Continuous Extreme Value Model (MDCEV) to assess activity-time use preferences. Their work shows that flexible work arrangements are associated with a higher likelihood of eating out and maintenance activities. Khaddar et al. (2023) analyze data from the Okanagan Travel Survey conducted in 2018 in the Central Okanagan region of British Columbia, Canada using an MDCEV model with ordered preferences (MDCEV-OP). They also find that telecommuters leverage the gained flexibility to participate in recreational and shopping activities as well as eating-out and escorting activities, i.e., picking someone up or dropping them off. They further find that among telecommuters, the most time is invested into social activities indicating an additional need to seek out social interactions when working from home. Budnitz et al. (2020) analyze data from the UK National Travel Survey from 2002-2016 using a Multinomial Logit Model (MNL), finding that telecommuters conduct more escort, errands, and personal business trips as well as leisure and recreational trips compared to their non-telecommuting counterparts. Additionally, they find that telecommuting is associated with a higher proportion of walking or jogging, corroborating findings from the US indicating that telecommuting increases the likelihood of meeting the recommended duration of moderate exercise (Chakrabarti 2018). Rhee (2008) presents one of the few studies that find that the time saved through telecommuting is almost exclusively invested back into work with almost no time left for additional leisure activities. It should be noted that this study is not based on empirical data but on a spatial equilibrium model including households and firms, which is used for studying the effects of telecommuting and relocation. Therefore, comparing these findings to those of the other studies is not appropriate.

The findings from these studies indicate that telecommuting is at least to some extent leveraged to gain a better work-life balance. This is especially true for households with children for whom previous research has indicated a significantly higher risk of time poverty. Through analysis of data from the American Time Use Study, Bernardo et al. (2015) find that dual-earner households with children run a higher risk of social exclusion given the childcaring and work responsibilities that do not allow them to participate in leisure and social activities. Telecommuting has been shown to improve the reconciliation of work and family demands. An analysis of the California component of the NHTS shows that telecommuting is associated with a higher propensity of escorting trips (Su et al. 2021). Paleti and Vukovic (2017) also find that children in the household are associated with a decreased propensity to commute to work and an increased likelihood of escorting activities. He and Hu (2015) also find that telecommuters conduct more household-related activities based on their analysis of data from the 2007 Chicago Regional Household Travel Inventory. Khaddar et al. (2023) corroborates these findings and shows that especially high-income telecommuters leverage the gained flexibility to accommodate childcare, which is expressed through increased escorting activities. Findings presented by Budnitz et al. (2020) show a significant difference in escorting trips conducted by full-time workers, part-time workers, and teleworkers, in that full-time workers, conduct the fewest pick ups/drop offs and part-time workers the most, followed by telecommuters. These findings indicate that the temporal flexibility gained through telecommuting similar to part-time work is a means to gain more schedule control to reconcile work and family obligations (Kelly et al. 2011). These findings are further supported by studies based on time use data, which compared to traditional travel diary data, allow for analysis of in-home activities in addition to activities associated with travel. Giménez-Nadal et al. (2019) analyze data from the American Time

Use Survey from years 2003-2015 and find that compared to commuters, those who work from home spend more time on unpaid labor, i.e., household chores and childcare.

The described gains in flexibility are not distributed equally and a strong gender effect is identified by numerous studies, in that women who telecommute are more likely to engage in activities associated with household responsibilities whereas telecommuting men have a higher propensity to leverage the gained flexibility of work from home to increase recreational activities. Asgari and Jin (2017) find that women are more likely to report shopping activities. Paleti and Vukovic (2017) presents findings showing that men who telecommute are more likely to spend time eating out and on discretionary activities and less inclined to participate in shopping, maintenance and pick up/drop off activities. These results are similar to those presented by He and Hu (2015) who find that female telecommuters are more likely to conduct drop-off/pick-up activities compared to their male counterparts. These findings are supported by those presented based on time-use data. Losa Rovira et al. (2022) analyze data from the United Kingdom Time Use Survey and find that females have a higher propensity to engage in and allocate more time to shopping and homecare compared to males, who are less likely to allocate time to chores. These results are consistent with findings based on the US Time Use survey, which was analyzed by Wight and Raley (2009) to evaluate how work at home influences time use patterns. They find that while both mothers and fathers are more likely to multitask childcare and paid work activities at home, telecommuting fathers are less likely to engage in primary childcare. Along the same lines, Giménez-Nadal et al. (2019) find that while telecommuting parents allocate more time to childcare compared to commuting parents, in both cases mothers do so more than fathers. These findings show that gender disparities in childcare that have been found among traditional working parents (see e.g. (Offer and Schneider 2011)) also translate to telecommuting parents (Zhang et al. 2020). Interestingly, these disparities seem to depend on society's perception of gender roles and childcaring responsibilities. Kurowska (2020) compares how gender effects differ in countries with different models of labor division using an SEM based on data from the Generations and Gender Survey. She compares models from Poland, which follows a traditional division of labor, and Sweden, in which labor tends to be divided equally among mothers and fathers. Her findings indicate that the gender disparities concerning childcare among telecommuters that have been presented in previous studies are prevalent only in Poland, whereas in Sweden, the effects of telecommuting on childcare and homecare are similar among men and women.

Furthermore, studies find behavior differs depending on the household composition. Based on an analysis of the English National Travel Survey from 2005-2019, Caldarola and Sorrell (2022) find that the propensity for increase in non-work activities is larger for single-worker households compared to dual-worker households indicating a different division of household responsibilities and an intra-household effect of telecommuting depending on the household type. This household-level effect in the form of a shift in household responsibilities towards the telecommuter in the household is also reported by Asgari et al. (2016).

2.1.2 Activity generation and scheduling approaches

Activity-based models have emerged as a response to the limitations of trip-based approaches, in that the latter do not account for the actual mechanism underlying travel, which is driven by the need to conduct activities at different locations. Activity-based models rely on activity schedules which represent an individual's activity pattern throughout a given time frame, often a single day. The activity schedules are the source of the trips that are simulated within an activity-based model. Activity-based models aim to represent realistic behavior and are therefore often complicated. However, this complexity provides a strong foundation to analyze travel behavior within travel demand models (Miller 2023). Activity generation and scheduling models can broadly be categorized into rule-based, utility-based, and optimization-based models. It should be noted that this categorization is not strict. Some rule-based models have econometric parts and some optimization-based models are formulated under random utility theory. The allocation of these models to the three groups is based on their similarities to the other models within that group.

One of the first rule-based approaches is the SCHEDULER model Gärling et al. (1989), Golledge et al. (1994) in which the cognitive process behind activity schedule generation is implemented as a computer program. Activities are assigned a priority and scheduled given they do not violate previously defined constraints. Interestingly, one of the first use cases of the model was the analysis of telecommuting effects on travel behavior. Although an operational model was never developed, this model was the conceptual basis for many other models.
A few years later, Pendyala et al. presented AMOS - Activity Mobility Simulator - which is made up of several modules to generate activity patterns that allow for the analysis of travel demand management policies (Pendyala et al. 1997, 1998). The first module processes individual trip data from travel diaries, ensuring logical consistency and completeness, and establishes a baseline activity-travel pattern for each person. Subsequently, utilizing a neural network trained with data from hypothetical transport control measures (TCM) scenarios, this module predicts the primary behavioral response of individuals to a TCM, based on their baseline activity patterns and socio-economic factors. Based on this response, the next module generates feasible modified activity-travel patterns, taking into account secondary and tertiary behavioral changes induced by the primary response. Next, the utility of the modified activity-travel patterns is assessed to determine whether these new patterns are accepted or rejected. Finally, the output module collects and reports detailed travel data from accepted patterns.
Another extensive rule-based model is Albatross, first presented by Arentze and Timmermans (Arentze et al. 2000, Arentze and Timmermans 2004). Albatross is made up of several control agents that handle e.g., data consistency, simulation of activity patterns, and alternative model scenarios. The model scheduler is at the core of the scheduling process, in which decision trees represent the decision-making process. The subjects of the model are households and thus, household interactions are considered in various parts of the model. These interactions are additionally the focus of subsequent model development and application (Anggraini et al. 2012, Timmermans and Zhang 2009). Albatross was later reimplemented

and published as FEATHERS (Bellemans et al. 2010). The model is developed in a modular way such that the scheduling core can be replaced by other models. The adapted model was applied in the region of Flanders.
Another rule-based model is ADAPTS, presented by Auld and Mohammadian (2009, 2012). Within the model, the activity generation and scheduling are conducted in three subsequent phases. The first phase is activity generation, which involves deciding whether or not to add an activity of a certain type. The second phase is activity planning, where the actual details of the activity are specified. Finally, the third phase is activity scheduling, where the activities are added to the planned schedule and any conflicts are resolved. Compared to the previous model, the decisions at various points in the model are represented by choice models, which is why the model is also referred to as a hybrid model, that combines facets of rule-based and utility-based models.
The aforementioned approaches generate activity schedules for one day. In contrast, Märki et al. (2014), Märki (2014) presents C-TAP, which allows for the consideration of multi-week periods. C-TAP models an agent's underlying motivations through behavioral targets. Discomfort is introduced in the model when an agent is in a condition that deviates from their target. The resulting discomfort can be minimized by conducting activities at different locations. External conditions and constraints are introduced through effectiveness functions, which inform agents of the efficacy the activities and locations have concerning the ability to reduce discomfort. Within C-TAP, activities are modeled continuously, driving the behavior of the agents based on past, current, and (possible) future states.

Utility-based models generate activity schedules based on econometric models, such as Logit and Probit models. One of the earliest advanced utility-based models is the day activity scheduler presented by Bowman and Ben-Akiva (2001). In this model, activity scheduling is considered a multidimensional choice through an extensive Nested Logit formulation, structured into five tiers. At its core, individuals in the model choose whether to conduct an activity pattern that includes traveling from one place to another, selecting from 54 different travel patterns. This choice process begins with a fundamental decision about travel, followed by deeper selections within a nested logit framework where each pattern's utility

incorporates linked decisions from subordinate tours. The model further differentiates between primary and secondary activities. Primary activities, deemed most crucial, dictate the major tour of the day concerning time, destination, and mode of transport. Secondary activities, though less critical, are scheduled similarly. Tours are categorized based on their complexity, number of stops, and purposes, with work tours subdivided into various types, from direct round trips to more complex journeys involving multiple stops or work-based subtours. Similarly, school and other activities follow a simpler categorization.
Around the same time, Kitamura and Fujii (1998) presented the PCATS model, which is short for Prism-Constrained Activity Travel Simulator. The model divides the day into open and blocked periods, with open periods allowing for flexible activities and travel, while blocked periods are reserved for fixed activities at designated locations. Decisions about activities are made sequentially, focusing on immediate past activities within the constraints of these periods, without anticipating future activities beyond them. Activities are categorized as either fixed or flexible based on their timing within these periods. Activity choice is modeled using a two-tier nested logit structure. The initial tier involves choosing among broad categories such as in-home activities, activities near the next fixed location, or other out-of-home activities. The subsequent tier refines these choices further, determining specifics like the type of out-of-home activity or whether to stay at home or move on to another location. Destination and mode choices are also modeled using a nested logit approach, where the first tier selects the destination and the second tier decides the travel mode based on the chosen destination. Finally, the maximum duration of an activity is confined by the size of the time-space prism, considering factors like travel speed and the scheduling of adjacent fixed activities. Durations are adjusted to fit these constraints, ensuring that each activity fits within the available time window by truncating the duration distribution model at the maximum allowable duration.
Along the same lines of assigning a priority to different activities, Bhat et al. (2004) present CEMDAP - the Comprehensive Econometric Microsimulator for Daily Activity-Travel Patterns. Within the model, an individual's activity patterns are first divided into different periods, depending on whether they are working following the assumption that activities are generally scheduled around work. Thus

these periods are defined as before work, home-to-work commute, work period, work-to-home commute, and after work period. The model applies choice models to determine if the individual participates in an activity. These choice models are applied incrementally following a pre-defined order with work being the first activity. The same method is used for students. Subsequently, the simulator includes non-mandatory activity generation.

Another utility-based model is presented by Bradley and Vovsha (2005). CT-RAMP generates activity schedules for the period of one day considering household interactions by explicitly accounting for utilities at the household level. Daily activity patterns and travel decisions are categorized into three types: mandatory, non-mandatory, and at-home activities, with further scheduling for mandatory activities to determine frequency and timing, leaving room for potential non-mandatory or joint activities. Joint household travel is modeled to simulate joint tours, including the number, purpose, participants, destination, and timing. Maintenance and discretionary tours are allocated within households for tasks like shopping or other non-essential outings, deciding their destination and timing.

Again, these described utility-based models are limited to one day. Hilgert et al. (2017), Hilgert (2019) present actiTopp, a model that generates activity schedules for the period of one week. Similar to the previously presented approaches, the schedules are generated in a sequential manner by applying numerous Logit models. First, the main activity of each day is determined, followed by the number of hours per day and their purpose. Subsequently, the number of activities per tour and their purpose is chosen. Based on a weekly time budget per activity and agent, the duration of each activity is assigned. Based on the personal preferences of the agents, the start of the tour is determined. If joint activities are planned for the week, the schedules of each household member are adapted such that the joint activities align.

In a very similar way (and based on the same data - the German Mobility Panel), Moeckel et al. (2024) presents the Activity-Based Incremental Transport Model (ABIT). In this model, first mandatory activity participation and frequency are determined followed by incremental models to determine discretionary activities for each agent given a hierarchy of the purposes. Each mandatory tour can include subtours, which are determined in the next step. Given the main tours and subtours

by activity purpose, duration and start times for each activity are generated based on a probability distribution from empirical data. For each activity without a fixed destination (i.e., home, education and work), a destination is chosen. Finally, based on a vehicle allocation model, a Nested Logit model is applied to determine the mode that the agent uses to get to each activity.
Another utility-based approach is presented by Ordóñez Medina (2016), who presents a model that generates multi-day schedules of flexible activities given the input of a skeleton activity agenda. The skeleton agenda includes mandatory activities that dictate when flexible activities can be scheduled. The flexible activities are generated based on a binary Logit model determining if the agent conducts a given activity.
The aforementioned utility-based models all utilize some version of a discrete choice model, in which a hierarchy of activities has to be assumed as activities cannot be considered simultaneously. This limits the models' behavioral realism as no trade-offs between the activities are accounted for. This is addressed by models that apply the Multiple Discrete Continuous Extreme Value Model (MDCEV), first presented by Bhat (2005). In this model, all activities are considered simultaneously, and both the activity participation as well as the duration are determined at the same time (see Section 2.2). Applying an MDCEV model does not yield schedules but generates time use patterns without episode-level information. Palma et al. (2021) have extended the model formulation such that not aggregated time use patterns are generated, but episode-level activities during one day. They formulate the model in a way that each episode is considered as a separate alternative within the utility function and apply the forecasting algorithm proposed by Pinjari and Bhat (2010b) to generate multiple activity episodes. However, there is still a scheduling module missing that determines the order in which these episodes are conducted. This is addressed by Saxena et al. (2022) who propose an MDCEV model formulation with ordered preferences. Specifically, the model prevents that time is allocated to activity at position n j_n if no time is allocated to activity j_{n-1}. The model determines activity type, duration, episodes, and their order, thus providing full activity schedules.

The final category of activity generation and scheduling models consists of optimization-based models. These models propose that household or individual activity schedules are based on the solution of an optimization problem subject to a set of constraints. The presented models derived under this approach almost all include multiple objectives. The models differ in (a) the objectives (b) the constraints (c) how the weights associated with the objectives are estimated. The first optimization-based model presented to generate activity schedules was presented by Recker (1995), who formulated and proposed a solution to the so-called Household Activity Pattern Problem (HAPP). In this problem, the activity patterns are generated through solving a pickup and delivery problem with time windows (PDPTW), in which activities are *picked up* by a household member and completed at a given location. Once the activity is completed, it is *delivered* through returning back home. In the original form, the problem is formulated as a multiobjective optimization problem in which all objectives are considered equally important. Later formulations assign weights to each objective thus formulating the problem using the weighted sum method (Recker et al. 2008). The weights in this formulation are determined by applying a genetic algorithm such that the Levenshtein distance between the generated sequence of activities and the observed sequence of activities is minimized.

Similar to the HAPP approach, Allahviranloo and Axhausen (2018) propose a bi-level optimization model where the lower level generates activity schedules based on a PDPTW formulation and the upper level maximizes the accuracy of the generated schedules measured based on survey data. The parameters are again determined using a genetic algorithm.

While these formulations provide an efficient way of handling multiple dimensions of the decision-making process, the focus of a HAPP is on scheduling and not activity generation, which have to be provided as input to the problem. To address this issue, Xu et al. (2017, 2018) propose a variation of the HAPP in which they include the activity participation decision. They develop their problem under the random utility maximization theory. In the proposed framework, for each individual, a choice set is generated by first clustering observed activity patterns and selecting a subset of these patterns. The generated choice set is individualized such that no infeasible patterns are part of the choice set. Finally, the parameters

of the utility function are estimated using a Path Logit model.
A similar approach is presented by Pougala et al. (2023) in the OASIS model - Optimisation-based Activity Scheduling with Integrated Simultaneous choice dimensions. Similarly to the HAPP, the utility is a function of utility components that describe activity-travel behavior including activity participation, the start time of each activity, their durations and a penalty associated with the travel time to get to the location associated with each activity. They face the same problem as Xu et al. in that the complete choice set is too large for the problem. They propose to generate the choice set by applying the Metropolis-Hastings algorithm. Based on the OASIS model, Rezvany et al. (2023) extend the model formulation to explicitly account for interactions between household members. They consider pick-up and drop-off activities, joint activities and a combination of the two, i.e. drop-off and stay. In contrast to the original OASIS formulation, in this formulation, the household utility is maximized under individual and household constraints. OASIS has also been extended to model multiple days, however, not while accounting for household interactions (Rezvany et al. 2023).

2.1.3 Evaluating activity-based models in the context of telecommuting

Based on the findings from the literature review on the effects of telecommuting on activity patterns, this section evaluates the capability of existing activity-based models to account for telecommuting behavior. A common theme across previous studies is that telecommuters gain more flexibility that most are likely to leverage for increased non-work activities and vice versa. This means that increased demand for non-work activity participation can lead some to choose to telecommute to satisfy this demand. This reciprocal relationship between telecommuting and activity engagement challenges the assumption that telework can be adequately modeled in approaches that represent activities hierarchically.
Previous studies have further identified that there is a gender bias regarding how the gained flexibility affects non-work activities. While men are more likely to conduct leisure activities, studies indicate that women are more likely to telecommute to

reconcile household responsibilities. This is especially the case for households with children. If we want the model to adequately represent this behavior including the allocation of household-level activities, the framework has to account for household interactions at the activity generation as well as the scheduling level. Furthermore, telecommuting is not equally spread throughout the week, as not everyone works from home full-time. As the literature indicates (Asgari et al. 2014, Asmussen et al. 2023), there are preferred days of the week on which individuals telecommute. It is therefore integral that any activity-based model accounts for day-to-day variation in telework demand and subsequently travel demand if the results are to inform policy-makers on the efficacy of telecommuting as a travel demand management measure.

Household interactions have been the focus of several model development efforts. For example, Anggraini et al. (2012) analyze the interactions between members of car-deficient households in the case of vehicle allocation for non-work tours by applying the Albatross model. More recently, Rezvany et al. (2023) integrated household interactions and the allocation of chores to the household members into the OASIS model. These developments offer great insights into how household interactions can be regarded in activity-based models, however, they both consider the simulation period for one day and are thus currently not suitable for the representation of telework.

There have also been great strides in multi-day approaches. Both Hilgert (2019) and Moeckel et al. (2024) present model approaches based on the 7-day travel diary data from the German Mobility Panel. Ordóñez Medina (2016) presents a two-phase approach to model multi-day schedules in which first a travel skeleton is created, which is later supplemented with flexible activities. Although the temporal scale of these models would fit the analysis and integration of telecommuting well, the formulation of the choice situations is discrete in all models, which means that there reciprocal effects of activity generation cannot be accounted for. The work by Märki et al. (2014) stands out in that it does not utilize choice models but simulates activities continuously in multi-week scenarios. This model approach seems promising in terms of its capability of integrating telecommuting behavior.

The data needed for the model is extensive multiday travel diary data, which is expensive to acquire, thus the transferability to other modeling contexts is limited.

Lastly, recent modeling efforts have improved the state-of-the-art of activity-based models regarding trade-offs in multiple dimensions. The model proposed by Saxena et al. (2022) models activity episodes, their duration, and order in a given day in a single choice model. The OASIS model (Pougala et al. 2023) and the extension of the HAPP model by Xu et al. (2018) also consider multiple constraints and trade-offs at the same time during activity scheduling. However, they also focus on single-day simulations. Although Pougala (2024) proposes an approach to extend the model to the multi-day case, the runtimes become too large to make the approach feasible in large-scale applications.

Overall, there have been great advances in activity-based models, however, there exists no approach that allows for the representation of the intricacies of telecommuting.

2.2 Random Utility Theory

The work in this thesis on multiple occasions relies on the application of models rooted in random utility theory. More specifically, both discrete and multiple discrete choice models are applied. As readers are likely more familiar with discrete choice models, random utility models are first derived based on the example of discrete choices. Subsequently, multiple discrete continuous choice models are introduced.

2.2.1 Discrete Choices

Discrete choices describe situations in which a decision-maker chooses exactly one option from a set of mutually exclusive discrete alternatives based on how their preferences for each option compare to the other alternatives. This process can be expressed based on utility theory, which, in general, establishes a mathematical relationship between, for example, the goods we consume, transport modes we use or days we work from home and our preferences, assuming that we aim to maximize our utility. Models based on utility theory are derived as follows. Let J be a set of mutually exclusive alternatives and n the decision-maker who faces a choice among those alternatives. The utility obtained by the decision-maker from choosing alternative j is denoted by U_{nj}. The assumption underpinning utility theory is that the decision-maker chooses alternative i over j if and only if $U_{ni} > U_{nj} \forall i \neq j$. The choice can be influenced both by attributes of the alternatives perceived by the decision-maker x_{nj} as well as attributes of the decision-maker s_n.

In traditional utility theory, this choice is assumed to be deterministic, meaning that a decision-maker will always gain the same utility in different choice situations if these situations are characterized by the same factors x_{nj} and s_n. Vice versa, this means that any choice that violates that assumption would be inconsistent and considered irrational. However, the decision-maker does not necessarily consider all attributes of an alternative equally across choice situations and thus may behave inconsistently across choice situations. Moreover, the choice may be influenced by factors that are not captured within x_{nj} and s_n. Both effects can explain inconsistencies in choice behavior while still assuming the rationality of the decision-maker. These inconsistencies can be accounted for in random utility models (Block 1974). Random utility models account for the inconsistencies by introducing an error term ε_{nj}. This error term is added to the deterministic utility, which we shall now refer to as V_{nj}, such that $U_{nj} = V_{nj} + \varepsilon_{nj}$. In the binary case of two alternatives in the choice set, the decision-maker n will choose alternative i over j if the utility of i is greater than that of j, i.e., $V_{ni} + \varepsilon_{ni} > V_{nj} + \varepsilon_{nj}$. Because the random term ε cannot be observed, the choice of the decision-maker cannot be predicted exactly. However, given a probability density function (PDF) and the corresponding cumulative distribution function (CDF) of ε, we can derive the probability for the choice of each alternative in the choice set. Discrete choice models differ in the assumption about these functions. In Logit models, the PDF is assumed to be a Gumbel distribution, whereas Probit models are derived under a Standard normal distribution of the error term. In the following, discrete choice probabilities are derived for the Logit model, as this is mainly applied in this thesis. The probability density function of the standard Gumbel distribution is given by

$$f(x) = e^{-(x+e^{-x})} = e^{-x}e^{-e^{-x}}$$

and the cumulative distribution function is

$$F(x) = e^{-e^{-x}}$$

Based on the Gumbel distribution, the PDF of the unobserved utility component of alternative j is given by:

$$f(\varepsilon_{nj}) = e^{-\varepsilon_{nj}}e^{-e^{-\varepsilon_{nj}}}$$

The CDF is given by:

$$F(\varepsilon_{nj}) = e^{-e^{\varepsilon_{nj}}}$$

Based on the PDF and CDF of the unobserved component of the utility functions, we can derive the choice probabilities. The probability that decision-maker n chooses alternative i over j is defined as:

$$\begin{aligned} P_{ni} &= Pr(V_{ni} + \varepsilon_{ni} > V_{nj} + \varepsilon_{nj}, j \neq i) \\ &= Pr(\varepsilon_{nj} < V_{ni} + \varepsilon_{ni} - V_{nj}, j \neq i) \end{aligned}$$

If we assume ε_{ni} is given, we can determine the conditional probability $P_{ni}|\varepsilon_{ni}$. In this case, the probability that decision-maker n chooses alternative i given ε_{ni}, is that the unobserved utility component of alternative j, ε_{nj}, is smaller than the value of $V_{ni} + \varepsilon_{ni} - V_{nj}, \forall j \neq i$:

$$\begin{aligned}
P_{ni}|\varepsilon_{ni} &= \int_{-\infty}^{V_{ni}+\varepsilon_{ni}-V_{nj}} e^{-x} e^{-e^{-x}} \\
&= e^{-e^{-x}} \Big|_{-\infty}^{V_{ni}+\varepsilon_{ni}-V_{nj}} \\
&= e^{-e^{-(V_{ni}+\varepsilon_{ni}-V_{nj})}} - \underbrace{\lim_{x \to -\infty} e^{-e^{-x}}}_{\to 0} \\
&= e^{-e^{-(V_{ni}+\varepsilon_{ni}-V_{nj})}}
\end{aligned}$$

Moving on from the binary case to the case of multiple j, we have to calculate the probability that $\varepsilon_{nj} < V_{ni} + \varepsilon_{ni} - V_{nj}$ for all $j \neq i$. Because in a Logit model, each ε_{nj} is assumed to be independent, this probability is determined by the product of the individual probabilities:

$$P_{ni}|\varepsilon_{ni} = \prod_{j \neq i} e^{-e^{-(V_{ni}+\varepsilon_{ni}-V_{nj})}}$$

Because ε_{ni} is not given, we must determine the unconditional probability P_{ni}. Applying the law of total probability given by $P(A) = \int_{-\infty}^{\infty} P(A|X = x) f_x dx$, the unconditional choice probability is:

$$\begin{aligned}
P_{ni} &= \int_{-\infty}^{\infty} (P_{ni}|\varepsilon_{ni}) \cdot f_{\varepsilon_{ni}} d\varepsilon_{ni} \\
&= \int_{-\infty}^{\infty} \prod_{j \neq i} e^{-e^{-(V_{ni}+\varepsilon_{ni}-V_{nj})}} \cdot e^{-\varepsilon_{ni}} e^{-e^{-\varepsilon_{ni}}} d\varepsilon_{ni}
\end{aligned}$$

Evaluating this integral results in a closed-form expression:

$$P_{ni} = \frac{e^{V_{ni}}}{\sum_j e^{V_{nj}}}$$

The arithmetic steps to arrive at this expression are given in appendix A.1.

2.2.2 Multiple Discrete Continuous Choices

Discrete choice models have been applied extensively in activity-based modeling frameworks (see 2.1.2). However, they assume that alternatives are perfect substitutes for each other and that they have a hierarchy in which they can be modeled. Moreover, given the discreteness of the models, they rely on the choice of conducting a given activity and the estimation of time use of the respective alternative in separate steps. These assumptions lead to inadequate representation of choice behavior. These shortcomings are addressed in models of “multiple-discreteness”, which have first emerged in marketing literature to account for the fact that buyers in many markets can choose among multiple units of a brand and multiple brands of the same good (Hendel 1999). Bhat (2005) has extended these models to the use case of discretionary activity time-use decisions. Based on the model structure presented by Kim et al. (2002), he introduces the Multiple Discrete Continuous Extreme Value Model (MDCEV), which compared to previous models of this kind, can account for a large number of discrete alternatives. The model is derived under random utility theory in that decision maker n maximizes their utility of consuming k alternatives. The formulation differs from discrete choice models in both the way the utility function is formulated as well as how the probabilities of consumption are derived. The utility is represented by a direct utility function $U(x)$, where x is the consumption quantity vector. In the case of activity time use, this would be the duration spent on each activity k. The overall utility function is given by:

$$U(x) = \sum_{k} \frac{\gamma_k}{\alpha_k} \psi_k((\frac{x_k}{\gamma_k} + 1)^{\alpha_k} - 1) \tag{2.1}$$

subject to the constraint that the household’s time budget E is maintained:

$$E = \sum_{k=1}^{K} x_k \tag{2.2}$$

In the model presented in this work, a one-week time frame is considered, resulting in the household's time budget E being the product of 10,080 minutes and the number of household members. In equation 2.1, K represents the number of alternatives, and x_k denotes the time allocation for activity k. ψ_k represents the base utility of alternative k, i.e., the marginal utility of the alternative when time allocation is 0 minutes. The parameters α and γ serve as saturation parameters, however, with different roles in the model. γ is a satiation parameter as in higher values of γ_k correspond to more time invested into activity k. It also serves as a translation parameter through which the model allows for corner solution, i.e., zero time investment into activity k. In turn, α_k is only associated with satiation. As an exponent of the baseline marginal utility ψ_k, it reduces the marginal utility with increasing time investment into alternative k. A comprehensive overview of the function of each parameter is presented by Bhat (2008). As in Logit models, the expressions ψ and γ can be parametrized as follows:

$$\psi_k = e^{\beta_{k,z} z_h + \varepsilon_k} \tag{2.3}$$

$$\gamma_k = \zeta_{k,z} z_h \tag{2.4}$$

where z_k accounts for the sociodemographic characteristics of household h, $\beta_{k,z}$ and $\zeta_{k,z}$ are estimated parameters, and ε_k is an extreme value error term. Maximizing this utility function based on the budget constraint poses a mathematical optimization problem (see section 2.3) of a nonlinear function, which can be solved by forming the Lagrangian and applying the Karush-Kuhn-Tucker conditions. The Lagrangian of the general utility function in 2.1 is given by[1]:

$$\mathcal{L} = \sum_{k=2}^{K} \frac{\gamma_k}{\alpha_k} \psi_k((\frac{x_k}{\gamma_k} + 1)^{\alpha_k} - 1) - \lambda \left(\sum_{k=1}^{K} x_k - E \right), \tag{2.5}$$

[1] Note: The model derivation presented in Bhat (2008) allows for price normalization of each good. This is unnecessary in activity time use estimation as no price is associated with spending time on an activity. The terms accounting for price normalization are therefore omitted from the equations.

where λ is the Lagrangian multiplier. The Karush-Kuhn-Tucker conditions for the optimal time investment x_k^* are then given by:

$$\psi_k \left(\frac{x_k^*}{\gamma_k} + 1 \right)^{\alpha_k - 1} - \lambda = 0, \quad \text{if } x_k^* > 0, k = 1, 2, \ldots, K \tag{2.6}$$

$$\psi_k \left(\frac{x_k^*}{\gamma_k} + 1 \right)^{\alpha_k - 1} - \lambda < 0, \quad \text{if } x_k^* = 0, k = 1, 2, \ldots, K \tag{2.7}$$

.

Solving for the lagrangian multiplier λ for the first good in equation 2.6, and substituting for the other alternatives in equations 2.6 and 2.7, the Karush-Kuhn-Tucker conditions can be rewritten as:

$$V_k + \varepsilon_k = V_1 + \varepsilon_1, \quad \text{if } x_k^* > 0, k = 2, 3, \ldots, K \tag{2.8}$$

$$V_k + \varepsilon_k < V_1 + \varepsilon_1, \quad \text{if } x_k^* = 0, k = 2, 3, \ldots, K, \text{where} \tag{2.9}$$

$$V_k = \beta_{k,z} z_h + (\alpha_k - 1) \ln \left(\frac{x_k^*}{\gamma_k} + 1 \right), \quad k = 1, 2, 3, \ldots, K \tag{2.10}$$

.

Based on this general utility formulation, Bhat (2008) presents different model structures and formulations for several choice situations. This thesis utilizes a model structure with an outside good, in which one alternative is always chosen; in the model presented in section 5.2, alternative *home* is considered as an outside good. This model is defined as follows:

$$U(x) = \frac{1}{\alpha_1} \psi_1 x_1^{\alpha_1} + \sum_{k=2}^{K} \frac{\gamma_k}{\alpha_k} \psi_k \left(\left(\frac{x_k}{\gamma_k} + 1 \right)^{\alpha_k} - 1 \right) \tag{2.11}$$

Compared to the general utility formulation, there is no γ parameter defined for alternative 1, as this alternative is always consumed and thus no corner solutions of zero consumptions have to be accounted for. Regardless of whether a model is defined with or without an outside good, we cannot estimate both γ_k and α_k at the same time. There are three “profiles” that can be estimated:

- α-profile: the α_k parameters are estimated while defining $\gamma_k = 1$
- γ-profile: the γ_k parameters are estimated while defining $\alpha_k \to 1, \forall k \geq 2$
- α-γ-profile: the γ_k parameters are estimated while $\alpha_k = c$ where c is a constant that does not vary between alternatives, serving as a "pinning effect" from the satiation parameter for the outside good.

Although there are estimation procedures for all three profiles, there currently exists no numeric forecasting procedure for a model based on an α-profile. As the intention of the modeling framework proposed in this work is to be applied within an agent-based demand model, it is essential that there is a suitable forecasting procedure available. Thus, the models in Section 5.2 are based on the α-γ-profile. The probability that M of the K alternatives are consumed is given by:

$$\begin{aligned} &P\left(x_1^*, x_2^*, x_3^*, \ldots, x_M^*, 0, 0, \ldots, 0\right) \\ &= \left[\prod_{i=1}^{M} f_i\right] \left[\sum_{i=1}^{M} \frac{1}{f_i}\right] \left[\frac{\prod_{i=1}^{M} e^{V_i}}{\left(\sum_{k=1}^{K} e^{V_k}\right)^M}\right] (M-1)! \end{aligned} \tag{2.12}$$

where V_k given the α-γ-profile is defined as:

$$\begin{aligned} V_k &= \beta z_k + (\alpha - 1) \ln\left(\frac{x_k^*}{\gamma_k} + 1\right) \forall k \geq 2 \\ V_1 &= (\alpha - 1) \ln(x_k^*) \end{aligned} \tag{2.13}$$

and f_i is defined as:

$$f_i = \frac{1 - \alpha_i}{x_i^* \gamma_i} \tag{2.14}$$

Based on the initial introduction of the MDCEV model by Bhat (2008) as outlined above, there have been many extensions. Most relevant to this work is the Multiple Discrete Continuous Nested Extreme Value Model (MDCNEV) presented by Pinjari and Bhat (2010a), which relaxes the assumptions that the error terms of the alternatives are independently distributed. In this model, the error terms are assumed to be distributed according to a joint extreme value distribution:

$$F\left(\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_K\right) = \exp\left[-\sum_{s=1}^{S_K}\left\{\sum_{i\in \text{ sth nest}} \exp\left(-\frac{\varepsilon_i}{\theta_o}\right)\right\}^{\theta_s}\right] \tag{2.15}$$

where s is attributed to one of the S_k nests. Based on the distribution of the correlated error term, Pinjari and Bhat (2010a) derive the probability of choosing a combination of alternatives in the MDCNEV model, which is defined as follows:

$$\begin{aligned}
P(x_1^*, x_2^*, ..., x_M^*, 0, ..., 0) &= |J| \frac{\prod_{i\in \text{CA}} e^{\frac{V_i}{\theta_i}}}{\prod_{s=1}^{S_M}(\sum_{i\in s^{th} Nest} e^{\frac{V_i}{\theta_i}})^{q_s}} \\
&\cdot \sum_{r_1=1}^{q_1} \cdots \sum_{r_s=1}^{q_s} \cdots \sum_{r_{s_M}=1}^{q_{s_M}} \left\{ \prod_{s=1}^{S_M} \left[\frac{(\sum_{i\in s^{th} Nest} e^{\frac{V_i}{\theta_i}})^{\theta_s}}{\sum_{s=1}^{S_k} \left\{ (\sum_{i\in s^{th} Nest} e^{\frac{V_i}{\theta_i}})^{\theta_s} \right\}} \right]^{q_s - r_s + 1} \right. \\
&\left. \times \Big(\prod_{s=1}^{S_M} \text{sum}(X_{rs}) \Big) \Big(\sum_{s=1}^{S_M} (q_s - r_s + 1) - 1 \Big)! \right\}
\end{aligned} \tag{2.16}$$

Here, θ_i is the parameter for the associated nest i, which is estimated alongside the other parameters in the model. Similarly to a Nested-Logit model (Train 2001), the values should lie between 0 and 1, where a value of 1 indicates no correlation between alternatives within that nest, making the model equivalent to the MDCEV model without a nested structure. The expression $sum(X_{rs})$ represents the sum of elements in a row matrix X_{rs}. For a more detailed description of the matrix, please refer to Pinjari and Bhat (2010a). The variable S_M describes the nest containing M alternatives, and q_{SM} represents the chosen alternatives within it. The set CA represents the set of chosen alternatives. Finally, the variable x_M^* denotes the optimal time allocation, and J stands for the Jacobian matrix.

2.3 Mathematical Optimization

The activity scheduling model presented in this work is formulated as a mathematical optimization problem. This section provides an overview of the methods under which the scheduling problems are formulated.
In mathematical optimization, we aim to optimize, i.e. maximize or minimize, a function under constraints. In its most general form, an optimization problem is defined as:

$$P = \begin{Bmatrix} \max & f(x) \\ \text{s.t.} & x \in X \end{Bmatrix} \tag{2.17}$$

where $f(x)$ is the objective function to be optimized. Note that each maximization problem can be transformed into a minimization problem, thus choosing maximization in the formulation given in 2.17 is without loss of generality. Further, x is the decision variable and X is called the solution space.

Integer Linear Programming

There are many different types of optimization problems, whose classification depends on the type of objective function and the values the variables can assume. The optimization problems formulated in this work are Integer Linear Programs (ILP). In a linear programming problem, both the objective function and the constraints are linear. A problem is called an integer program or integer optimization problem when all variables in the problem can only assume integer values. An integer linear program in canonical form with n decision variables and m constraints is defined as:

$$\begin{aligned} \max \quad & \mathbf{c}^{\mathbf{T}} x \\ \text{s.t.} \quad & \mathbf{A}x \leq \mathbf{b}, \quad x \in \mathbb{Z}^n \end{aligned} \tag{2.18}$$

where

$x \in \mathbb{Z}^n$ is the vector of **decision variables**. This vector is varied in the process of finding the optimum of the objective function.

$\mathbf{c} \in \mathbb{R}^n$ are the **objective function coefficients**.

$\mathbf{A} \in \mathbb{R}^{nxm}$ is called the **coefficient matrix** with $\mathbf{M}$ rows and $\mathbf{N}$ columns

$\mathbf{b} \in \mathbb{R}^m$ is the vector with $\mathbf{M}$ components defining the right-hand side of the linear inequalities.

Multiobjective optimization

The aim of an ILP as presented in Section 2.18 is to optimize exactly one objective. In some cases, including generating activity schedules, the problem at hand involves optimizing multiple, often competing objectives. These problems are referred to as multiobjective optimization problems (MOOP). Multiobjective optimization allows for the consideration (and illustration) of trade-offs when scheduling activities, such as household members wanting to spend as much time together as possible while also having to work or run errands, which in turn should preferably be allocated equally among household members. A multiobjective optimization problem is defined as:

$$\begin{aligned} \max \quad & F(x) \\ \text{s.t.} \quad & \mathbf{A}x \leq \mathbf{b}, \\ & x \geq 0 \end{aligned} \tag{2.19}$$

where $F(x)$ is no longer a single objective function but a vector of p objective functions:

$$F(x) = \begin{pmatrix} F_1(x) \\ \vdots \\ F_p(x) \end{pmatrix} \tag{2.20}$$

As stated above, the objectives of a MOOP are often at least partially conflicting, which means that there usually does not exist a single point that optimizes all objectives at once. The goal is rather to find points that are Pareto optimal: A point is Pareto optimal iff it is impossible to find another point that improves at least one objective. In the case of a single objective, a Pareto optimal point is often unique. In a multiobjective problem, there are often many Pareto optimal solutions. To deal with this issue, we can include preferences among objectives to find a single solution to a multiobjective optimization problem. There are three different types of preference-based optimization methods whose classification is based on the time when the preference is articulated:

A priori methods The preference of the objectives is determined prior to finding a solution to the optimization problem. Based on the defined preference, a priori methods will result in a single solution to the multiobjective optimization problem.

Interactive or progressive methods The decision maker provides input throughout the optimization process and progressively shifts the outcome of the optimization solution towards their preferences.

A posteriori methods In these methods, instead of determining a single solution, a set of solutions within the solution space is generated. The decision-maker expresses their preference by choosing the solution best suited to their needs.

The ultimate aim of the activity modeling framework developed in this thesis is for the generated activity schedules to be simulated within an agent-based model.

Although it would be technically possible and indeed interesting to explore for agents to choose among a set of schedules or have them interact with the scheduler during the optimization process to find a solution, both progressive and a posteriori methods are more computationally expensive than a priori methods in this use case. Therefore, the scheduler in this thesis is based on a priori methods to finding a solution to the multiobjective activity scheduling problem. Among the a priori methods, the **weighted sum approach** is one of the most frequently applied methods (Marler and Arora 2010), and has also been applied in other optimization-based activity scheduling models (see section 2.1.2). Following previous approaches, the scheduling model developed in this thesis also adopts the weighted sum method.
The aim of the weighted sum approach is to transform the problem from a multi-objective to one with a single objective by assigning a weight w_p to each objective function $F_p(x)$ and adding the weighted objectives to one function $\widehat{F}(x)$ which is then optimized:

$$\begin{aligned} \max \quad & \widehat{F}(x) = \sum_{i=1}^{p} w_p F_p(x) \\ \text{s.t.} \quad & \mathbf{A}x \leq \mathbf{b}, \\ & x \geq 0 \end{aligned} \tag{2.21}$$

The weights are restricted to positive values, i.e., $w \geq 0$, to ensure Pareto optimality. Further, the weights have to be set such that $\sum_{i=1}^{p} w_p = 1$ to guarantee a convex combination of the objective functions.

Weight Calibration Approach

The aim of the model presented in this thesis is to simulate activity scheduling behavior as realistically as possible. Thus, the primary goal of the weight calibration step is to tune the model weights so that the results replicate real-world behavior and patterns observed in the target population. This involves adjusting parameters within the model to ensure that the emergent behavior of the simulated agents matches observed behaviors or data from the real world as closely as possible. More formally, the calibration process can be described based on the following:

- $\mathbf{X}$: The set of parameters or rules of the model that need to be calibrated.
- $\mathbf{Y}_{\text{sim}}$: The output or behavior of the simulated model corresponding to parameter set $\mathbf{X}$.
- $\mathbf{Y}_{\text{obs}}$: The observed real-world behavior or data that we aim to replicate.
- $f(\mathbf{X})$: The function representing the model that maps parameter set $\mathbf{X}$ to simulated behavior $\mathbf{Y}_{\text{sim}}$.
- $\mathrm{L}(\mathbf{Y}_{\text{sim}}, \mathbf{Y}_{\text{obs}})$: The loss or discrepancy between simulated behavior $\mathbf{Y}_{\text{sim}}$ and observed behavior $\mathbf{Y}_{\text{obs}}$.

The goal is to find the parameter set $\mathbf{X}^*$ that minimizes the loss function:

$$\mathbf{X}^* = \underset{\mathbf{X}}{\operatorname{argmin}} \, \mathrm{L}(f(\mathbf{X}), \mathbf{Y}_{\text{obs}})$$

Here, the loss function can take various forms depending on the nature of the data and the specific calibration objectives. For instance, it could be the mean absolute error of the number of weekly activity episodes, or another suitable metric measuring the discrepancy between simulated and observed behavior. There are several methods to approach this problem, for instance, Bayesian optimization, Genetic Algorithms, local search, or exhaustive evaluation. In this work, a Bayesian optimization approach is chosen due to its ability to handle expensive evaluations of objective functions (Brochu et al. 2010).

In Bayesian optimization, the above optimization problem is solved iteratively, where at each iteration, the surrogate model is updated based on observed data points (simulated behaviors D_n), and an acquisition function α is used to decide which parameter set to evaluate next in order to balance exploration and exploitation. Figure 2.1 illustrates the Bayesian optimization approach.

Bayesian optimization employs a surrogate model, which is typically a Gaussian process (GP). The surrogate model approximates the true objective function (the loss function) based on the observed data points (simulated behaviors) evaluated

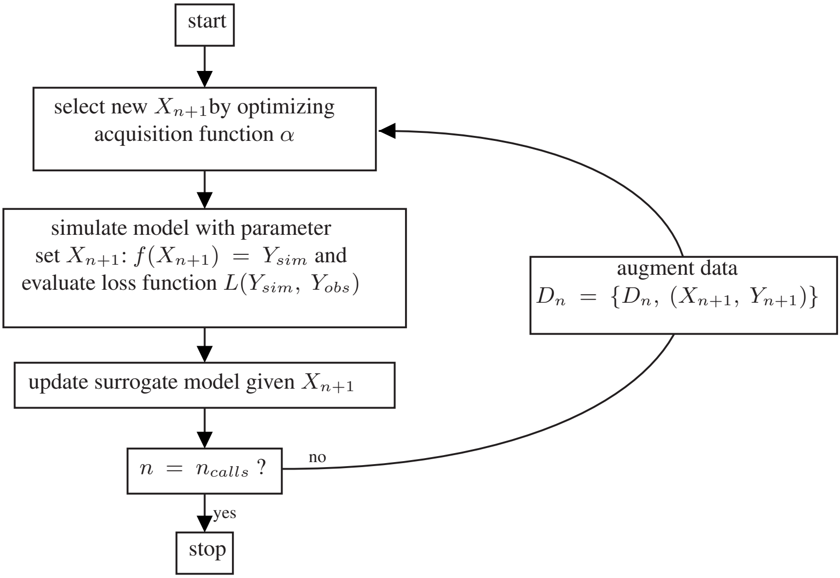


Figure 2.1: Bayesian Optimization process based on Shahriari et al. (2016)

so far. The GP provides a probabilistic representation of the objective function, including an estimate of the mean and uncertainty (covariance) at each point in the parameter space. At each iteration of Bayesian optimization, the surrogate model is updated based on the newly observed data points (simulated behaviors). This update incorporates the information gained from the evaluations of the objective function so far, improving the surrogate model's accuracy and reducing uncertainty. In addition to the surrogate model, Bayesian optimization uses an acquisition function to guide the selection of the next parameter set to evaluate. The acquisition function quantifies the utility or informativeness of evaluating a particular parameter set based on the current state of knowledge provided by the surrogate model. It balances exploration (sampling parameter sets in regions where uncertainty is high or where the model's predictions are uncertain) and exploitation (sampling parameter sets that are expected to yield low loss values based on the surrogate model) (Shahriari et al. 2016).

3 Survey Data and Data Preparation

It is a capital mistake to theorize before one has data. Insensibly one begins to twist facts to suit theories, instead of theories to suit facts.

Sir Arthur Conan Doyle,
The Adventures of Sherlock Holmes - A Scandal in Bohemia

This thesis is based on two main sources of data: the German Mobility Panel (MOP) and a household travel survey (HTS) conducted in the Metropolitan Region of Stuttgart, Germany. This section provides an overview of the two surveys, focussing on telecommuting relevant variables.

3.1 The German Mobility Panel

The German Mobility Panel is a national household travel survey that was conducted annually between 1994 and 2023. It is a longitudinal survey in which travel behavior is reported for seven consecutive days in three consecutive years. Each yearly sample consists of approximately 3,000-3,400 respondents in about 1,800-2,000 households. All household members over the age of 10 in a participating household are asked to keep a travel diary for seven days and to provide information on their socio-demographic characteristics, e.g., age, gender, and job status. Further, household characteristics are provided, such as the number of cars in the household and household income. A question on the telecommuting status was added to the personal questionnaire in 2012, asking whether respondents can work exclusively from home on some working days and if they choose to do so.

Responses are captured as a single choice among the following choice options:

- Yes, and I do so frequently (once a week or more)
- Yes, and I do so infrequently (less than once a week)
- Yes, but I do not choose to work from home
- No, I cannot work from home

Figure 3.1 shows the development of responses to this question over the years since its introduction. The graph shows that the share of telecommuters remained steady between 2012 and 2017, and a slight increase in telecommuters in 2018 and 2019. What can clearly be seen is the effect the COVID-19 pandemic has had on the share of telecommuters, which, as a result, increased by 19 percentage points in 2020 compared to the previous year. It is unclear how representative the telecommuting shares are for the employed population of Germany. Compared to the German microcensus (Statistisches Bundesamt 2024), the shares are higher in the MOP. However, compared to the Global Survey of Working Arrangements (Aksoy et al. 2022), the telecommuting rate presented here is lower. It should be noted that the MOP is biased towards high-income respondents (Kuhnimhof et al. 2006). As those in high-paying occupations are also more likely to be able to telework, telecommuters may be overrepresented in the MOP.

The yearly sample is relatively small compared to other national household travel surveys, which could potentially affect the validity and reliability of the statistical analyses and models developed based on the data. However, this issue can be mitigated by pooling data from multiple years, leading to a larger sample size and thereby enhancing the robustness of the analysis. To balance the need for timely results and sufficient sample size, the analytical and modeling framework presented in the next chapters are based on data acquired between 2018 and 2022, with the exclusion of the 2020 wave to avoid biases resulting from the COVID-19 pandemic.

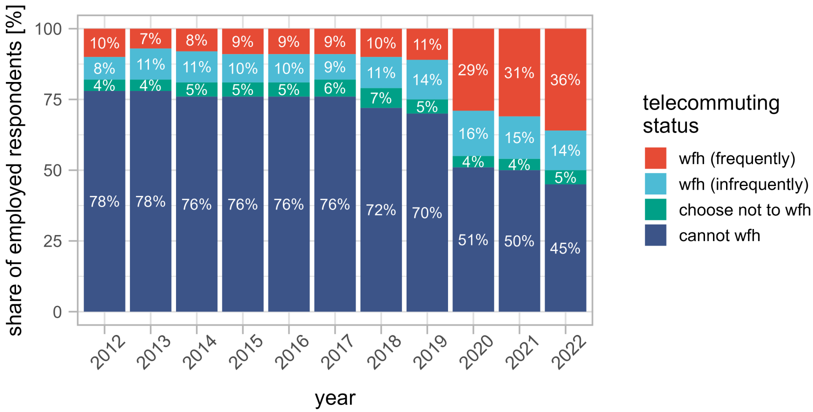


Figure 3.1: Share of telecommuters among employed respondents in the German Mobility Panel over the years. The graph is based on weighted data using the weighting factors at the individual level provided with the data.

Although the MOP provides a sound foundation for analyzing travel behavior and activity patterns, it lacks comprehensive data on telecommuting behavior. Most importantly, the survey does not ask whether participants worked remotely on the day of the survey or how often they did so throughout the week. Therefore, the data is supplemented by information from another household travel survey conducted in the Region of Stuttgart. The following section presents an outline of the survey, followed by an explanation of the supplementation approach.

3.2 Stuttgart Household Travel Survey

The Stuttgart Household Travel Survey (SHTS) was conducted in 2021 as part of the project VenAMo[1] among 9,959 respondents in 5,477 households. The SHTS

[1] Full project title: Verkehrsentlastung durch neue Arbeitsformen und Mobilitätstechnologien (original) / Traffic Reduction through new Forms of Work and Mobility Technologies (translation)

was similar to the MOP and consisted of a travel diary that each household member was requested to fill out, along with questionnaires on socio-demographic information at both the individual and household levels. In contrast to the MOP survey, the travel diary covers only one day and it is a one-off survey. Further, there was no age restriction on who should complete the survey. An adult household member was asked to fill out the surveys representatively for children under the age of 14. Due to the scope of the project, the survey allowed more nuanced questions on telecommuting. Additionally, the survey was supplemented by a voluntary online questionnaire focussing on employed respondents and their working conditions. The supplemented survey was filled out by 659 respondents of which 443 attested they worked from home. Table 3.1 presents an overview of the main telecommuting-related questions in the SHTS. The survey results are similar to those in the MOP wave of the same year (2021). About 51% of employed respondents have the possibility to work from home. Of those who can work from home, about 6% choose to (almost) never work from home, whereas the majority telecommutes frequently. Among those who cannot work from home, the majority indicate that their type of occupation does not allow them to work remotely. About 10% said their employer does not allow telecommuting, and just over 3% of respondents cannot work remotely because their home does not provide sufficient (quiet) space.

3.3 Data Preparation

The data preparation steps serve two purposes: first, transforming the data to create a dataset for estimating parameters for the MDCEV model, and second, using the data as a reference for calibrating schedules. Additionally, the data needs to contain detailed information about telecommuting behavior, specifically the days and times during which an individual engages in telecommuting. This section outlines the performed data preparation steps to generate this data.

Table 3.1: Overview of the telecommuting-related questions in the SHTS. The presented data is weighted based on the weighting factors at the individual level provided with the data.

variable	levels	share	question type
wfh possible	yes	51%	single choice
	no	49%	
wfh frequency if possible	(almost) daily	35%	single choice
	1-3 days per week	40%	
	1-3 days per month	12%	
	Less than 1 day per month	6%	
	(almost) never	6%	
not possible due to	insufficient (quiet) space at home	3%	multiple choice
	employer does not allow wfh	10%	
	type of work does not allow wfh	82%	
	other reasons	9%	

3.3.1 From travel diaries to time-use data and activity schedules

The model presented in this thesis is split into two parts: time use estimation and the scheduling. Both model parts are based on the travel diaries obtained from the MOP. However, the format of the data needs to be adapted before it can be used. For instance, a travel diary may not cover the entire week for a respondent. If a respondent has conducted out-of-home activities for only three days during the survey week, then the travel diary would only reflect those three days instead of the entire week. In such cases, the data is augmented by adding home activities to each diary. This ensures that each diary starts on a Monday at 00:00 and ends on a Sunday at 23:59.

As identified in the literature section, household interactions play a vital role in activity schedules. These interactions are in part reflected by joint activities. Although the MOP does not collect information on whether an activity was conducted jointly, Hilgert et al. present an approach to identify joint trips and activities in the MOP travel diaries (Hilgert et al. 2017, Hilgert 2019). This approach is also applied in this work. For this purpose, activities are identified as joint activities if they have the same purpose, start and end at the same time, and if the trips concerning these activities cover the same distance. The MOP data has a significant drawback in that it lacks geo-coded information on trip destinations. Consequently, this method is subject to uncertainty. However, it does serve as a sufficient foundation for the modeling framework suggested later on, which is set up to be able to account for joint activities once more suitable data becomes available.

Lastly, travel times have to be accounted for in the activity schedules. Because the MOP does not include information on trip destinations, the modeling framework does not include parts on destinations or mode choices. Similar to the work presented by Hilgert (2019), these choices are made subsequently to generate activity schedules. Therefore, the duration of each trip is added to the subsequent activity. This approach simulates the decision-making process of scheduling, as we generally reserve a time block for certain activities without knowing exactly how and when to get there when planning the activity. In order to maintain reasonable ratios between the travel times and the duration of activities, the scheduled duration of the activity can be taken into consideration in the destination choice. This can be ensured by limiting the choice set of destinations to those that would result in an appropriate ratio between travel time and activity duration. Joint activities are limited to certain activity purposes, specifically shopping and leisure activities.

The dataset resulting from conducting the aforementioned preparation steps represents activity schedules in the target format of the modeling framework presented in this thesis. This can be used to calibrate the scheduling module of the model. However, at this point, there is detailed information missing on telecommuting within these schedules. This information is supplemented in two steps, described in subsections 3.3.2 and 3.3.3, respectively.
Additionally, an intermediate step of the proposed framework is the generation of activity time-use at the household level while allowing for individual-specific alternatives that are not interchangeable between household members. Therefore

the data has to be aggregated into a dataset that can be used to estimate the activity time-use model. At this stage, two types of activities are differentiated:

individually specific alternatives Alternatives that are allocated to a specific household member and cannot be interchanged between different members: *work, work-related, telecommuting, education, leisure*

household level alternatives Alternatives that can be allocated interchangeably between different household members or that are conducted jointly and thus pertain to multiple household members: *home, shopping, joint shopping, joint leisure, pick up and drop off*

For household-level alternatives, the activity durations for the respective alternatives are summed over all household members, whereas individually specified alternatives are treated as separate alternatives. Time-use for a household of two household members $i \in 1, 2$ may be defined, for example, for the following activities a_i: *home, shopping, joint leisure,* $work_1$, $leisure_1$, $work_2$, $telecommuting_2$. Given this data, the MDCEV model on time-use can be estimated and the resulting parameters calibrated.

3.3.2 Prediction of telecommuting engagement

The work in this section has been presented in the following contribution:

Reiffer, A.S., Kagerbauer, M., Vortisch, P. *Prediction of Telecommuting Engagement through Machine Learning to Enhance Travel Survey Data* presented at the 103rd Annual Meeting of the Transportation Research Board, Washington D.C., January 2024.

As mentioned at the beginning of this chapter, telecommuting engagement, i.e., whether a respondent worked remotely on a given survey day, is not recorded in the MOP. As a trade-off between the benefit of week-long data and information density from the 1-day SHTS, this section presents an approach to predict telecommuting engagement for each day of the week using machine learning classifiers trained on the SHTS. Figure 3.2 shows an overview of the proposed approach. First, all possibly relevant features are selected from the SHTS; however, they are limited to those that are also available in the MOP data. The SHTS is unbalanced regarding the telecommuting engagement variable, so the Synthetic Minority Over-sampling TEchnique (SMOTE) is used to balance the data (Chawla et al. 2002). The resulting dataset is then split into train and test data. Several feature selection algorithms are applied to the test data, and for each feature set, different machine learning classifiers are trained. The trained models are then tested using the test data set, and their performance is evaluated using several criteria. Finally, the best model is chosen and used to predict telecommuting engagement in the MOP data for each reported day of a telecommuter's travel diary.

It is worth noting that although the MOP is generally limited to one question on telecommuting, in 2022 an additional questionnaire was included that asked respondents to indicate how many days they worked from home in the survey week. This data is also used to test the performance of the classifiers.

In the following, first, all machine learning classifiers are described, followed by the feature selection techniques and finally, the performance metrics are introduced. Subsequently, the results from the feature selection and classification are presented.

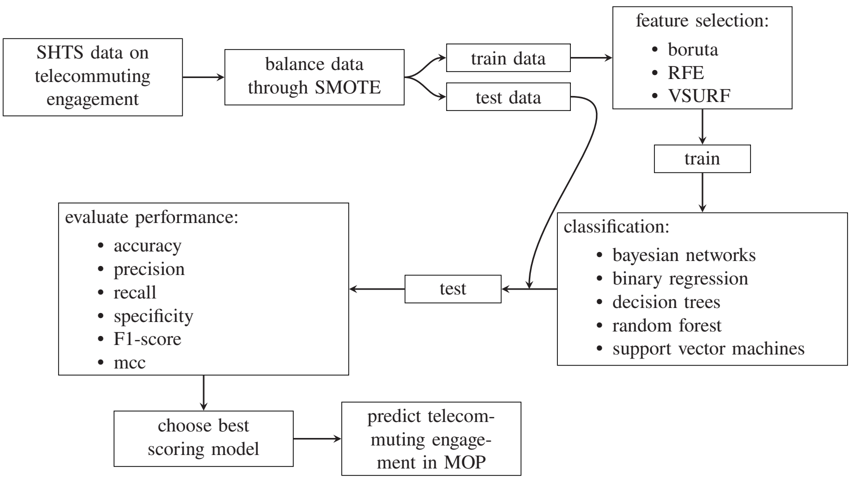


Figure 3.2: Approach to predict telecommuting engagement in MOP data based on machine learning classifiers trained on SHTS

Machine learning classifiers

Bayesian Networks
Bayesian networks are probabilistic graphical models that represent the dependencies between random variables using a directed acyclic graph (DAG) (Pearl 1997, Neapolitan 1990). Each node in the graph corresponds to a random variable, and the edges between nodes encode conditional dependencies. The conditional probability distribution for each variable given its parents is modeled using Bayes' rule. Let X_i denote the i-th random variable, and $\mathrm{Pa}(X_i)$ be the set of parent nodes of X_i in the graph. The joint probability distribution of all variables can be written as:

$$P(X_1, X_2, \ldots, X_n) = \prod_{i=1}^{n} P(X_i \mid \mathrm{Pa}(X_i)) \tag{3.1}$$

Inference in Bayesian networks involves computing probabilities or making predictions based on observed evidence. Popular algorithms for inference include variable elimination and Markov chain Monte Carlo (MCMC) methods.

Binary Regression
Binary regression, also known as logistic regression, is a popular supervised learning algorithm for binary classification tasks. Given a dataset with input-output pairs $(\mathbf{x}_i, y_i)$, where $\mathbf{x}_i$ is a feature vector and $y_i \in \{0, 1\}$ is the binary class label, the goal is to learn a model that estimates the probability of the positive class, i.e., $P(y_i = 1 \mid \mathbf{x}_i)$. The logistic regression model assumes a linear relationship between the features and the log-odds of the positive class:

$$\log\left(\frac{P(y_i = 1 \mid \mathbf{x}_i)}{1 - P(y_i = 1 \mid \mathbf{x}_i)}\right) = \mathbf{w}^T \mathbf{x}_i + b \tag{3.2}$$

where $\mathbf{w}$ is the weight vector and b is the bias term. To obtain probabilistic predictions, the logistic function is applied to the output of the linear model:

$$P(y_i = 1 \mid \mathbf{x}_i) = \frac{1}{1 + e^{-(\mathbf{w}^T \mathbf{x}_i + b)}} \tag{3.3}$$

Decision Trees

Decision trees are non-linear, hierarchical models used for both classification and regression tasks (Dattatreya and Kanal 1984). They recursively split the data into subsets based on the values of individual features, aiming to maximize the information gain or Gini impurity at each split. Each internal node in the tree represents a decision based on a feature, and each leaf node corresponds to a predicted class or regression value. The decision tree can be represented as a set of rules, and the final prediction for a given input is determined by following the path from the root to the appropriate leaf node.

Random Forest

Random forests are ensemble learning methods that combine multiple decision trees to improve predictive performance and reduce overfitting (Tin Kam Ho 1995, 1998). The key idea is to build a collection of decision trees by training each tree on a random subset of the training data (bootstrap sampling) and selecting a random subset of features at each split. The final prediction is made by aggregating the predictions of all individual trees, often using majority voting for classification problems or averaging for regression problems. Random forests tend to be more robust and accurate than individual decision trees, and they can handle high-dimensional data and capture complex relationships between variables.

Support Vector Machines

Support Vector Machines (SVMs) are powerful supervised learning algorithms used for both classification and regression tasks. SVMs aim to find the optimal hyperplane that best separates the data points of different classes while maximizing the margin between the classes (Cortes and Vapnik 1995). In the case of binary classification, given a training dataset $(\mathbf{x}_i, y_i)$, where $\mathbf{x}_i$ is the feature vector and $y_i \in \{-1, 1\}$ is the class label, SVMs find the weight vector $\mathbf{w}$ and bias term b that define the decision boundary:

$$\mathbf{w}^T \mathbf{x}_i + b = 0 \tag{3.4}$$

The margin is computed as the distance between the hyperplane and the closest data points (support vectors) from each class. SVM aims to maximize this margin while penalizing misclassifications. For non-linearly separable data, SVM can use kernel tricks to map the data into a higher-dimensional space, where linear separation becomes possible. Common kernel functions include polynomial, radial basis function (RBF), and sigmoid kernels.

Feature Selection

We further present the feature selection algorithms applied. We have trained and tested all algorithms on the dataset described above using the features determined by the respective feature algorithm.

Boruta
Boruta is a method for determining the importance of variables in a system using random forests. The system involves replicating each descriptive variable and randomly permuting the values of replicated variables across objects (Kursa et al. 2010). The randomization is different for each run of the random forest algorithm. The importance of each variable is computed for each run, and a statistical test is performed to determine if the variable is significant or not. An attribute is considered important for a single run if its level of importance is greater than the highest level of importance among all randomized attributes. If a variable is deemed unimportant, it is removed from the system along with its replicated mirror pair. The procedure is repeated for a predefined number of iterations or until all attributes are either rejected or deemed important. The algorithm was applied using the R package *boruta* (Kursa and Rudnicki 2010).

Variable Selection Using Random Forests - VSURF
Another method based on RF is the VSURF algorithm, which is short for “Variable Selection Using Random Forests” (Genuer et al. 2015). The procedure consists of two steps. In the first step, the variables are ranked based on their importance, estimating a threshold value for variable importance (VI) using the standard deviation of VI for less important variables and retaining only the variables with an averaged VI value above the threshold. In the second step, a sequence of ascending RF models is used to make predictions. Variables are added to each

model only if they significantly decrease the error rate, using a threshold based on the out-of-bag (OOB) error decrease. The final set of variables comes from the last model. In this study, we applied the VSURF method using the R package with the same name (Genuer et al. 2015).

Recursive Feature Elimination - RFE

Recursive Feature Elimination (RFE) was first introduced by Guyon et al. (Guyon et al. 2002). It is a method for feature selection similar to backward feature elimination (Kohavi and John 1997) but allows for the elimination of multiple variables simultaneously instead of having to eliminate one feature at a time through exhaustive enumeration. In the RFE procedure, a model is first built on all features. In the second step, a ranked feature list is created by ranking the combination of each feature. Lastly, features are eliminated if they do not meaningfully contribute to the model. We applied the RFE method using the R package *caret* (Kuhn 2008).

Performance metrics

We utilize several quantitative metrics to assess the performance of the ML classifiers. They all rely on the true positive (TP), true negative (TN), false positive (FP), or false negative (FN) values in one way or another. In the context of this study, these values are defined as:

- true positive (TP): model correctly classifies a telecommuting day as a telecommuting day
- true negative (TN): model correctly classifies a non-telecommuting day as a non-telecommuting day
- false positive (FP): model incorrectly classifies a non-telecommuting day as a telecommuting day
- false negative (FN): model incorrectly classifies a telecommuting day as a non-telecommuting day

Accuracy

The accuracy of a classification model is the percentage of sample objects that are correctly classified and labeled. This is done by calculating the ratio of the total number of true predictions to the sum of all observations.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{3.5}$$

Precision

Precision, also known as the positive predictive value (PPV), is defined as the ratio of correctly classified positive cases over all classified positive cases (Chicco and Jurman 2020).

$$Precision = PPV = \frac{TP}{TP + FP} \tag{3.6}$$

Recall

Recall, also referred to as sensitivity or true positive rate (TPR), is defined as the ratio of correctly classified positive cases overall actually positive cases (Chicco and Jurman 2020). Recall and precision are often trade-offs of each other.

$$Recall = Sensitivity = TPR = \frac{TP}{TP + FN} \tag{3.7}$$

Specificity

Specificity, also known as the true negative rate (TNR), is determined like the TPR, except that negative cases are now relevant. The TNR is defined as the ratio of correctly classified negative cases over all actually negative cases Chicco and Jurman (2020).

$$Specificity = TNR = \frac{TN}{TN + FP} \tag{3.8}$$

F1 Score

The F1-score is determined by calculating the harmonic mean of the precision (eq. 3.6) and recall (eq. 3.7) (Chicco and Jurman 2020). The F1-score can take values between 0 and 1, where 1 constitutes a perfect classification. As can be seen in the formula, this is achieved if the sum of false positives and false negatives is zero.

$$F1 = \frac{2 \cdot TP}{2 \cdot TP + FP + FN} = \frac{2 \cdot Precision \cdot TPR}{Precision + TPR} \tag{3.9}$$

Matthew's Correlation Coefficient

Although many studies use accuracy as the gold standard of model evaluation, it is very sensitive to unbalanced data, which can lead to a false sense of model performance (Chicco and Jurman 2020). A way to counteract the issue of class imbalance when evaluating model performance is to calculate the Matthews Correlation Coefficient (MCC) (Baldi et al. 2000). The MCC is calculated similarly to the Pearson product-moment correlation coefficient based on the confusion matrix of the model. The MCC can take on values between -1 and 1, where -1 is the worst possible value ($TP = TN = 0$) and 1 is the best possible value (i.e., $FP = FN = 0$).

$$MCC = \frac{TP \cdot TN - FP \cdot FN}{\sqrt{(TP + FP) \cdot (TP + FN) \cdot (TN + FP) \cdot (TN + FN)}} \tag{3.10}$$

We first assess the results of the feature selection algorithms, which are presented in 3.2. Most strikingly is the similarity between boruta and rfe. The only difference in the two feature sets is car access (included in rfe but not boruta) and the number of work-related trips (included in boruta but not in rfe). The vsurf feature set is the smallest, consisting of eight features. Features included in all three feature sets are *daily distance traveled, travel time, the hour of the last trip of the day, the number of work trips the time spent at home, spent for leisure activities, spent shopping, and on work-related activities.*. While it is expected that time spent at home is an important feature, we initially assumed that this would also be the case for time spent at work, which was deemed unimportant by the vsurf algorithm.

Table 3.2: Selected features by selection method. Values indicate if a feature was selected by the method (1) or not (0)

value	boruta	rfe	vsurf
age	1	1	0
car access	0	1	0
telecommuting Frequency	1	1	0
distance traveled	1	1	1
travel time	1	1	1
escorting someone	1	1	0
first trip of the day	1	1	0
last trip of the day	1	1	1
home	1	1	0
leisure	1	1	0
other	1	1	0
round trip	1	1	0
shopping	1	1	0
work	1	1	1
work-related	1	0	0
time use escorting someone	1	1	0
time use home	1	1	1
time use leisure	1	1	1
time use round trip	1	1	0
time use shopping	1	1	1
time use work	1	1	0
time use work-related	1	1	1

Figure 3.3: Confusion matrix by classifier and feature selection algorithm

Classification results

After selecting the features, we trained each model with the three feature sets on 70% of the 1-day HTS data. Subsequently, we tested the models on the remaining 30% by predicting whether a respondent in the test data worked from home on the respective survey day. In order to measure the performance of these models, we calculated the confusion matrix of predicted and real values, which provided us with values for true positive, true negative, false positive, and false negative. Figure 3.3 presents the confusion matrices for each classifier and each feature selection method.

The highest true positive values are predicted by the random forest model with 238 correctly classified true values based on the boruta and rfe feature sets, and 231 based on the vsurf feature set, respectively. All models have only few false negative values, with Bayesian networks based on the rfe feature set providing the lowest false negative values. This seems to come at the price of a very high value for false positive predictions. These are comparatively low in all other models,

Table 3.3: Performance metrics by classification model and feature selection algorithm

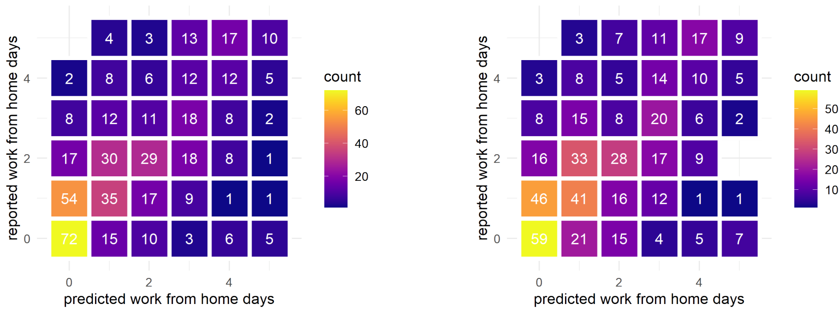
classifier	feature selection	accuracy	precision	sensitivity	specificity	F1-Score	mcc
Bayesian Network	rfe	0.55	0.43	0.96	0.33	0.59	0.34
Random Forest	rfe	0.88	0.79	0.91	0.87	0.84	0.76
Support Vector Machnie	rfe	0.86	0.73	0.92	0.82	0.81	0.71
Binary Regression	rfe	0.86	0.75	0.91	0.84	0.82	0.72
Decision Trees	rfe	0.86	0.72	0.94	0.81	0.82	0.72
Bayesian Network	vsurf	0.86	0.73	0.92	0.83	0.82	0.71
Random Forest	vsurf	0.87	0.75	0.94	0.84	0.83	0.74
Support Vector Machnie	vsurf	0.87	0.74	0.95	0.83	0.83	0.74
Binary Regression	vsurf	0.85	0.72	0.94	0.81	0.81	0.71
Decision Trees	vsurf	0.86	0.72	0.94	0.81	0.82	0.72
Bayesian Network	boruta	0.77	0.65	0.70	0.80	0.68	0.50
Random Forest	boruta	0.89	0.79	0.92	0.88	0.85	0.77
Support Vector Machnie	boruta	0.87	0.74	0.93	0.83	0.82	0.73
Binary Regression	boruta	0.86	0.74	0.91	0.84	0.82	0.71
Decision Trees	boruta	0.86	0.72	0.94	0.81	0.82	0.72

with random forest, again, performing best. Finally, true positive predictions are highest in the Bayesian network model based on the rfe feature set and lowest also for the Bayesian network when considering the boruta feature set. This is an interesting finding, as the boruta and rfe feature sets are almost identical (see Table 3.2), highlighting the need for pre-processing of data as Bayesian networks are often unsuited for continuous data or outliers (Cheng and Greiner 2013). All other models show almost identical confusion matrices for these two feature sets.

To further assess the performance of the model, we put the values of each confusion matrix into context with each other. The performance metrics for each model based on the values in the confusion matrices are presented in table 3.3.

Overall, we can see that almost all models achieved high rates of accuracy with only two models achieving an accuracy below 0.85. In both cases, bayesian network classification performed much worse compared to the other models. The recall/sensitivity metric shows even higher values, with, again, only the bayesian network model based on the features selected through the boruta algorithm reaching values below 0.91. The specificity is not as high as the previous two metrics, but overall, almost consistently values of over .80 are reached. Regarding the F1-Score and the Matthew's Correlation Coefficient (MCC), a similar trend concerning Bayesian networks is detectable: overall, the metric values are relatively high and over 0.80 and 0.70, respectively. However, both metrics are much lower for the Bayesian network model based on the features selected through rfe and boruta. Our analysis indicates that the random forest model utilizing features from the boruta selection exhibited the best performance overall. This model achieved the highest accuracy, specificity, F1-Score, and mcc values in comparison to the other models. However, it did present a comparatively low sensitivity value. On the other hand, the Bayesian network model based on rfe feature selection demonstrated the highest sensitivity value but underperformed in all other metrics.

To further evaluate the performance of our proposed approach to data enhancement, we predicted the telecommuting engagement in a separate HTS. For this purpose, we leveraged data from the German Mobility Panel in which respondents keep a travel diary for seven days. As the random forest models performed best, we used those for prediction. To evaluate how the comparatively large number of features from the best model (boruta feature set) compares to the smaller feature set from the vsurf algorithm, we conducted two out-of-sample predictions. Because information on telecommuting engagement is provided at the week-level and not the day-level, we cannot calculate the aforementioned performance metrics. Instead, we predicted the telecommuting engagement for each day and added them over the week for each respondents to get a comparative measure. At the aggregate level over the entire dataset, the model based on the boruta feature set performs slightly better. This model detects 803 telecommuting days out of the 915 (87.7%) reported days. Whereas the model based on the smaller vsurf feature set predicts 792 work from home days (85.6%) The confusion matrices by the number of telecommuting days are presented in Figure 3.4. On the left (Figure 3.4a), the results based on the boruta feature set are depicted, and on the right (Figure 3.4b) those on the vsurf feature set. Values above the diagonal are likely false negatives while values below the diagonal are most likely false positives.



(a) based on the boruta feature set

(b) based on the vsurf feature set

Figure 3.4: Predicted and actual number of work-from-home days in German Mobility Panel dataset.

The two models perform very similar. Both have a high rate of predicted non-telecommuting days over reported non-telecommuting days but also a relatively large rate of predicted non-telecommuting days over the reported one day of telework per week. In the latter case, vsurf performs slightly better than boruta. This performance difference is negated for a larger number of telework days per week.

Overall, the application of the models on the MOP dataset shows promising results and shows that the approach is viable to be applied to other HTS data. Random forest models are best suited for this approach and perform well even on a relatively small feature set. To the best of our knowledge, this is the first study testing different machine learning models to enhance in-home activity information in HTS data. Thus, we cannot compare our results to other studies. However, other studies comparing the performance of classification methods also find random forests to be one of the best performing methods (Zhang et al. 2017, Chen et al. 2020).

3.3.3 Imputation of telecommuting episodes

As the final data preparation step for the activity schedules, telecommuting episodes have to be identified for days on which a person telecommutes. For this purpose, the online survey supplemental to the SHTS is used as it provides detailed information on start and end times as well as durations of telecommuting. For each work-from-home day, candidate slots are determined among all home episodes in the schedule based on certain criteria. To be a candidate for a telecommuting episode, the home episode has to be between 7:00 AM and 9:00 PM, and it must be at least 60 minutes long.

Once the candidate slots are identified, the work-from-home duration for that day is sampled from the distribution of telecommuting durations provided by the data from the online survey supplemental to the SHTS.

The possible telecommuting slots are then sorted in descending order and relabeled as telecommuting episodes until the sampled work-from-home duration is distributed among the candidate episodes.

For episodes that are not completely replaced by telecommuting, an activity is added, and the activity start and end times of the telecommuting and home activity, respectively, are adjusted accordingly.

4 Telework Effects on Activity Patterns

Time is an illusion. Lunchtime doubly so.

Douglas Adams, *The Hitchhiker's Guide to the Galaxy*

The aim of the modeling framework in this thesis is to generate travel demand based on the underlying motivation of travel by accounting for different activity patterns. To be able to build the model accordingly, we first need an understanding of the effects of telecommuting on activity patterns. This section first presents an analysis of differences in activity patterns of telecommuters and non-telecommuters. Subsequently, it delves into a more detailed analysis by differentiating the regarded sample by household types, and lastly, activity patterns are analyzed by roles in family households, i.e., fathers and mothers. This analysis is based on employed respondents from the MOP waves 2018 through 2022, excluding data from 2020. Activity patterns are not describable by a one-dimensional variable. Instead, they are characterized by a multitude of variables and can be considered at multiple temporal scales. To gain an understanding, the following variables are investigated:

Weekly time use This variable describes how the 10,080 minutes / 168 hours in a week are split over different activities.

Number of episodes The term *activity episode* in the context of an activity schedule refers to a distinct period of time during which a specific activity is performed. The number of episodes is the sum of these events.

Activity duration Duration of individual activity episodes. These are regarded per day, meaning that episodes can take a maximum value of 1,440 minutes.

These variables are analyzed through two types of graphs: The weekly time use and activity duration are illustrated using boxplots, whereas the number of activities is presented using a barplot of the mean number of episodes per activity and the respective analytical variable, with error bars showing the mean standard error of the distribution. The summary statistics underlying each figure are also presented in Appendix A.2.
Additionally, statistical tests are conducted to assess the significance of the difference between different subsamples. To compare the means of the weekly time use and the activity duration, a Wilcoxon signed-rank test was conducted between each group. The number of weekly episodes constitutes count data, which is not suitable for a t-test or a Wilcoxon signed-rank test, thus a Poisson regression was conducted where the number of episodes was considered as the dependent variable and the analytical variable of choice was included as an independent variable. The significance of the test results is depicted in each figure, using the following notation:

- significance at the 1%-level: ***
- significance at the 5%-level: **
- significance at the 10%-level: *
- no significance: n.s.

4.1 Activity patterns by telecommuting status

First, we analyze weekly time investments into different activities by telecommuting status. Figure 4.1 displays a boxplot of time use by telecommuting status, with the brackets at the top of each activity group indicating if the distributions of each telecommuting status by activity are statistically different based on the Wilcoxon signed-rank test.

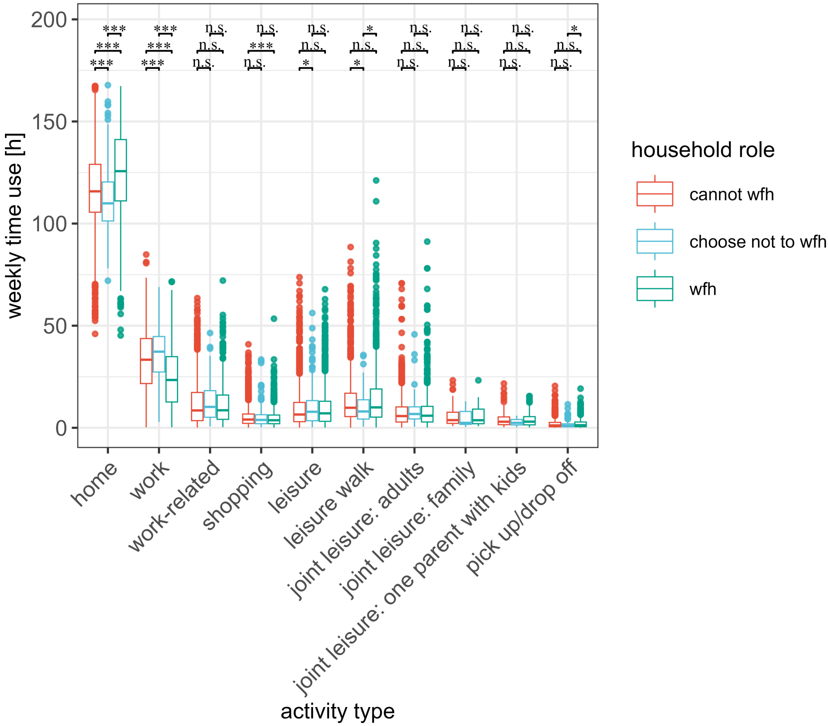


Figure 4.1: Weekly time-use by activity and telecommuting status of employed respondents

The analysis shows that teleworkers spend a significantly higher amount of time at home and less time in the office. Meanwhile, those who choose not to work from home spend the least amount of time at home and invest the most into work. One possible explanation for this is that they prefer to work on-site due to their responsibilities or career aspirations. Additionally, those who choose not to work from home spend the least amount of time on leisure walks. On the other hand, those who work from home spend significantly more time on escorting activities compared to those who choose not to work from home. However, no significant difference can be discerned when compared to those who cannot work from home.

Further analysis of the data reveals differences between the telecommuting groups regarding the number of conducted activities (see Figure 4.2).

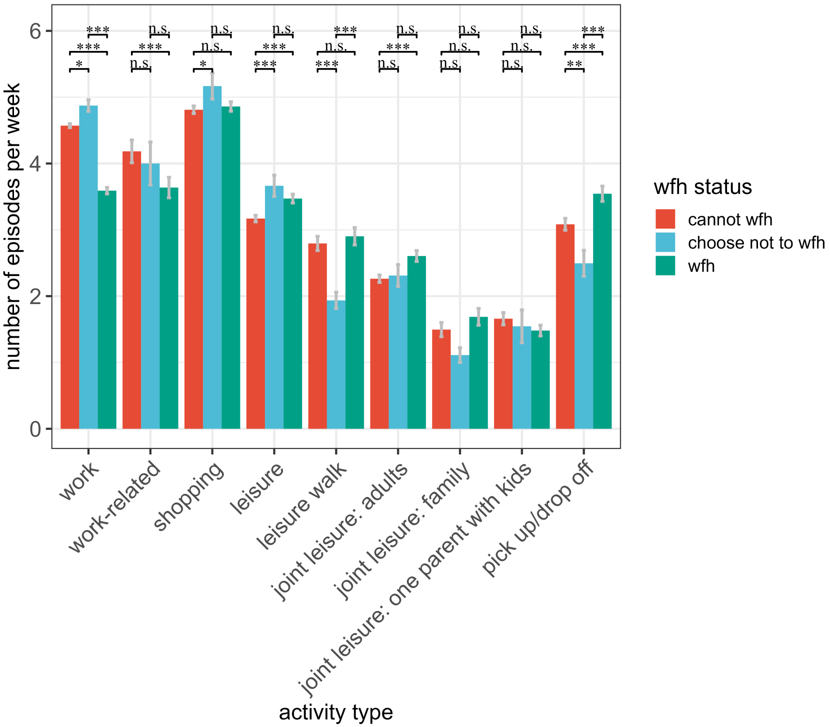


Figure 4.2: Weekly number of episodes by activity and telecommuting status of employed respondents

From the chart, it can be seen that telecommuters conduct the fewest number of work and work-related episodes. Those who choose not to work from home conduct significantly more work episodes than those who cannot telecommute. Further, those who have the option to work from home, regardless of whether they do so, conduct significantly more leisure episodes compared to those who cannot work from home. Furthermore, those who choose not to telework conduct

significantly fewer leisure walks.
There is no significant difference between the groups regarding the number of joint activities, except for a significantly higher rate of joint leisure activities of adults by telecommuters compared to those who do not work from home, indicating that some of the flexibility gained through telecommuting is used to spend more time with other (adult) household members. The figure also shows significant differences regarding the number of escorting activities. Those who choose not to work from home, even though they could, conduct the fewest pickup/drop off activities, while telecommuters conduct the most, showing significant differences compared to the other two groups. Although, those who cannot telecommute conduct significantly fewer escorting episodes, they do conduct more compared to those who choose not to telecommute, even though both groups can be described as non-telecommuters.

Considering work episode duration, we see similar trends as with weekly time use and number of episodes. Figure 4.3 shows the boxplot of the duration of episodes by activity and telecommuting status.

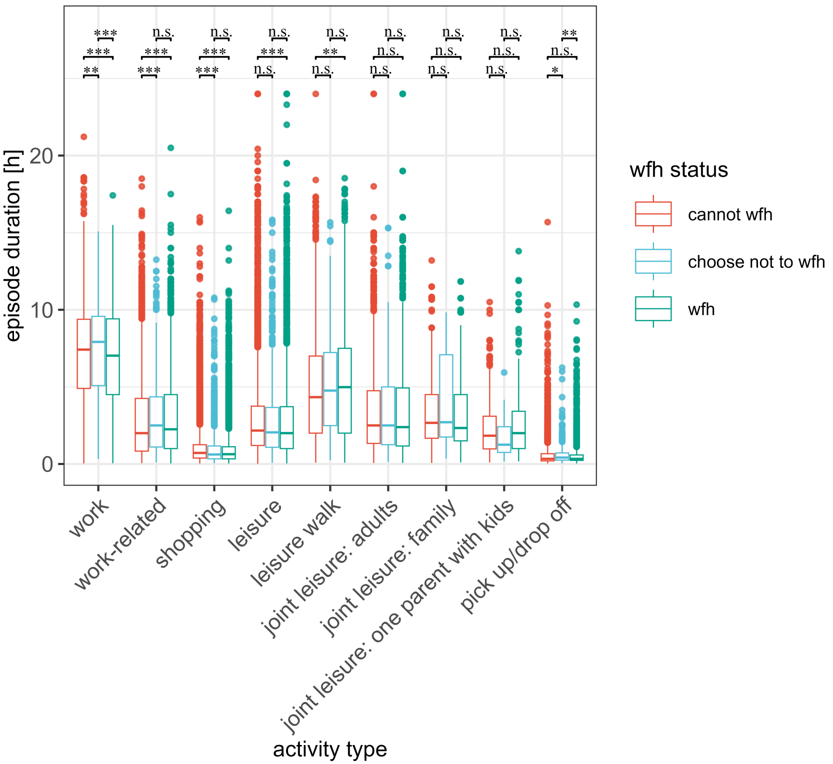


Figure 4.3: Duration of episodes by activity and telecommuting status of employed respondents

Those who work from home have the shortest duration, even if they go to the office. This is likely because they are avoiding peak hours or pushing back the beginning of on-site work to accommodate childcare and escorting activities. Those who choose not to work from home have significantly longer escort episodes compared to the other two groups. These are more likely to involve escorting children to their leisure activities rather than school runs, which tend to be shorter. Lastly, those who cannot work from home have the shortest work-related episodes.

In summary, the data analysis in this section reveals that individuals who work from home unsurprisingly spend the most time at home and the least time at work. However, the activity patterns of those who cannot work from home and those who choose not to telecommute differ significantly although they could both be categorized as non-telecommuters. Those who choose not to work from home spend a significantly longer amount of time at work, both weekly and episodically, compared to those who cannot work from home. This discrepancy is likely due to the fact that telecommuting is limited to specific industries, such as finance, IT, and service industries, while manufacturing, retail, and transportation industries require on-site presence. However, the latter industries also tend to have regulated work hours and fewer career opportunities. The findings suggest that those who prefer working in the office even if they have the option to work from home have responsibilities or career aspirations that are often tied to longer work hours and on-site engagement, an effect that has also been established in earlier studies ((Maruyama and Tietze 2012)).

4.2 Activity patterns across household types and telecommuting status

In this section, we examine the activity patterns of the three telecommuting categories based on household type, distinguishing between couple households, family households, multi-adult households (i.e., those who share accommodations), and single households. Figure 4.4 displays the distribution of weekly time use by household type for each telecommuting status, with brackets indicating the results of the Wilcoxon signed-rank test on whether the distributions of two household types differ significantly from each other.

Among individuals who are unable to work from home, those living in family households spend the longest amount of time at home. One possible explanation for this is that occupations in which employees cannot work from home are often associated with lower incomes, making it less likely for these households to afford external childcare compared to those working in telecommuting-prone industries.

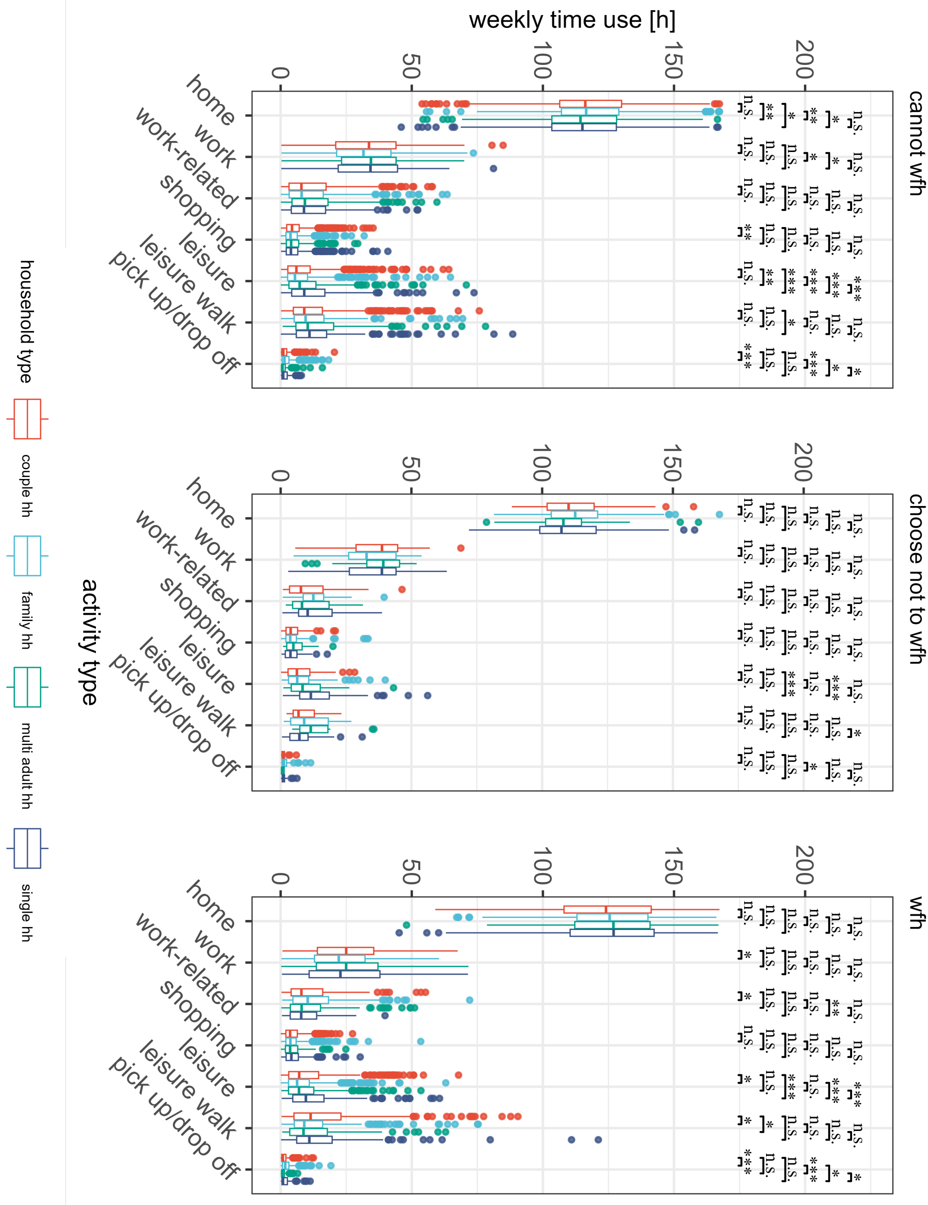


Figure 4.4: Weekly time-use by household type, activity, and telecommuting status of employed respondents

Additionally, those in family households spend significantly more time on pick-up/drop-off activities, except among those who choose not to work from home, who are likely to have more disposable income and can afford adequate childcare. In contrast, those who telecommute might do so because they do not want to outsource childcare, leading to higher time investment in escorting activities.
Single households and those in multi-adult households spend significantly more time on leisure activities, indicating a higher likelihood of seeking out social interaction.

Next, Figure 4.5 displays the average number of episodes conducted by activity and household type for each telecommuting category. The brackets located at the top of the chart represent the outcome of the Poisson regression analysis determining whether two household types within each telecommuting group differ significantly from each other concerning the number of activities they conduct throughout the week.

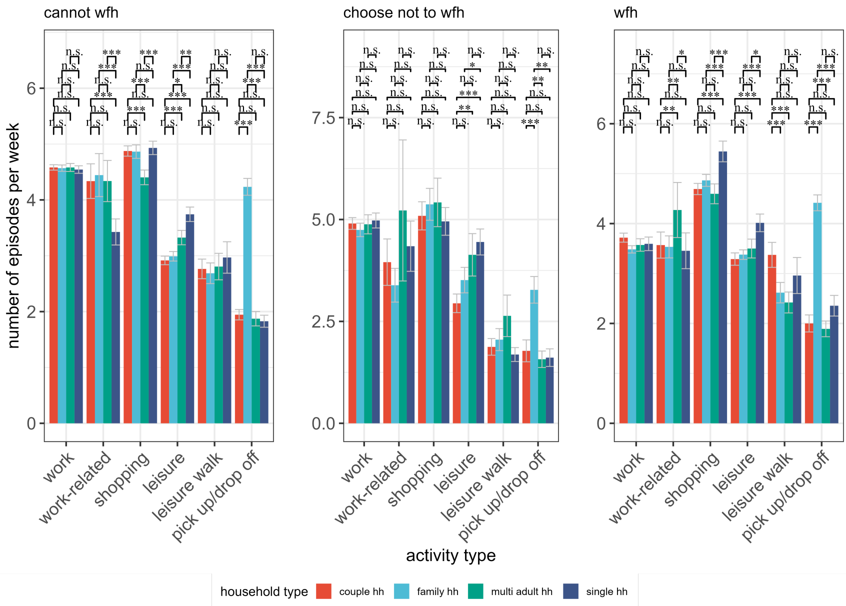


Figure 4.5: Weekly number of episodes by household type, activity, and telecommuting status of employed respondents

The findings of the analysis indicate that individuals who live alone spend more time engaging in leisure activities. Furthermore, they tend to engage in such activities more frequently, an effect that is particularly prevalent among telecommuters who live alone. This is likely because teleworkers who live alone have limited opportunities for social interaction during the day, unlike those who work in offices or live with other household members. Therefore, they are more likely to seek out social interactions through leisure activities, as reflected in the MOP data. Additionally, teleworkers who live alone engage in significantly more shopping activities compared to those who live in family households. On the other hand, individuals who choose not to work from home on average conduct the fewest escorting activities across all telecommuting groups.

Figure 4.6 shows boxplots of episode durations by activity and household type for each telecommuting group with brackets at the top presenting the results from the Wilcoxon signed-rank test comparing the distribution of episode duration of two household types for each activity within each telecommuting group.

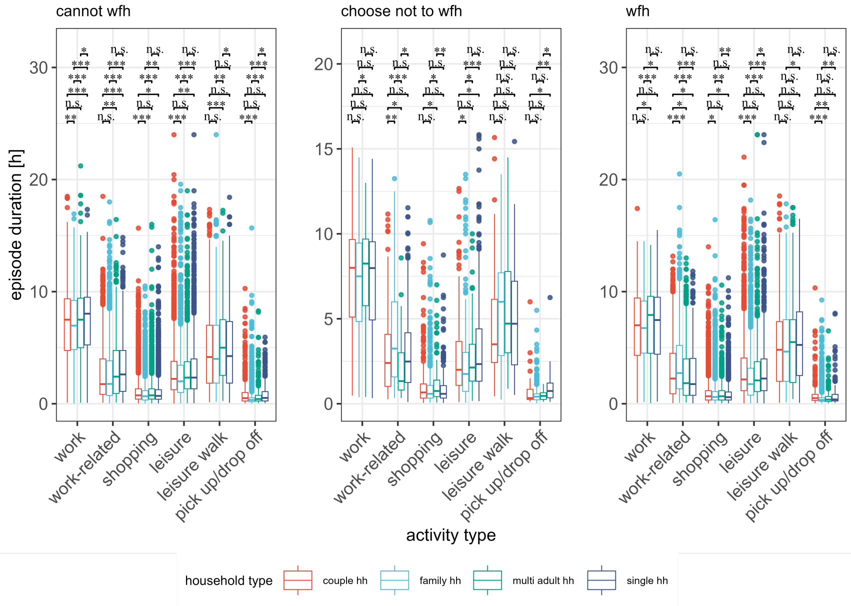


Figure 4.6: Duration of episodes by household type, activity, and telecommuting status of employed respondents

The data reveals that the duration of work and leisure activities vary across different telecommuting groups. The group that lives in family households conducts the shortest work activities. However, the difference in work duration compared to other household types is the smallest among those who choose not to work from home. Interestingly, the same cannot be said for work-related activities. Those who cannot work from home conduct the shortest work-related activities, while those who can telecommute, regardless of whether they choose to do so,

conduct longer work-related activities. This suggests that those working from home might schedule longer work-related activities, such as workshops, which cannot be replaced by online meetings. On the other hand, those who choose not to work from home may be required to attend work-related activities that keep them from teleworking.
Furthermore, the data also shows that those living alone conduct the longest leisure activities, while those in family households conduct the shortest. Additionally, those in family households conduct the shortest pick-up and drop-off activities, which is consistent with school runs. In contrast, those from other households are more likely to pick up and drop off someone at the train station or airport, which typically takes longer.

Overall, the findings highlight noticeable disparities in activity habits across various household compositions. Notably, individuals in single and multi-adult households report significantly higher engagement in leisure activities, indicative of a need for social interactions outside of a household context, which is particularly pronounced among teleworkers living alone. The results in this chapter further indicate that family households, typically with constraints around non-telecommutable jobs and childcare, spend more time at home and engaged in pick-up/drop-off activities. This suggests a strong link between household demands and the flexibility or lack thereof in their employment. Based on prior research regarding the division of unpaid work in family households and the influence of telecommuting, it is important to investigate who is responsible for taking care of children in these households. Therefore, the following section compares the activity patterns of mothers and fathers.

4.3 Activity patterns of mothers and fathers by telecommuting status

The final part of the empirical analysis evaluates differences between the activity patterns of mothers and fathers within each telecommuting group. Figure 4.7 compares the weekly time invested in the different activities.

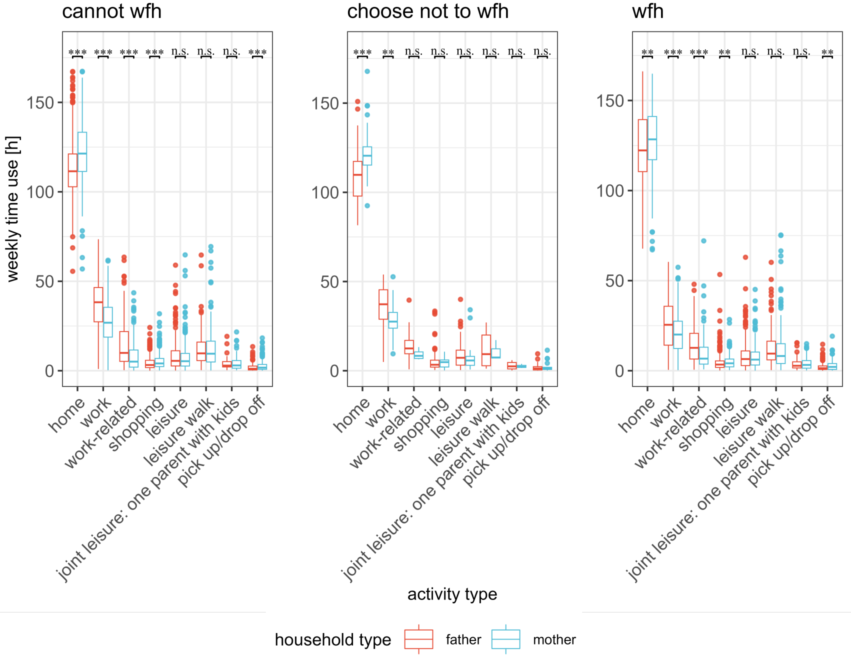


Figure 4.7: Weekly time-use by household role, activity, and telecommuting status of employed respondents

From the graph, it can be seen that mothers tend to spend more time at home and less time at work than fathers across all groups. This pattern is also observed in work-related activities, although the difference is statistically significant only for those who cannot work from home and those who choose to telecommute. For those who choose not to telecommute, no significant difference is evident in the weekly time use of non-work activities between mothers and fathers. However, for mothers compared to fathers who cannot work from home and those who choose to telecommute, a significant difference is observed in the time spent on shopping and escorting activities.

Differences between mothers and fathers regarding the number of weekly episodes by activity are presented in Figure 4.8.

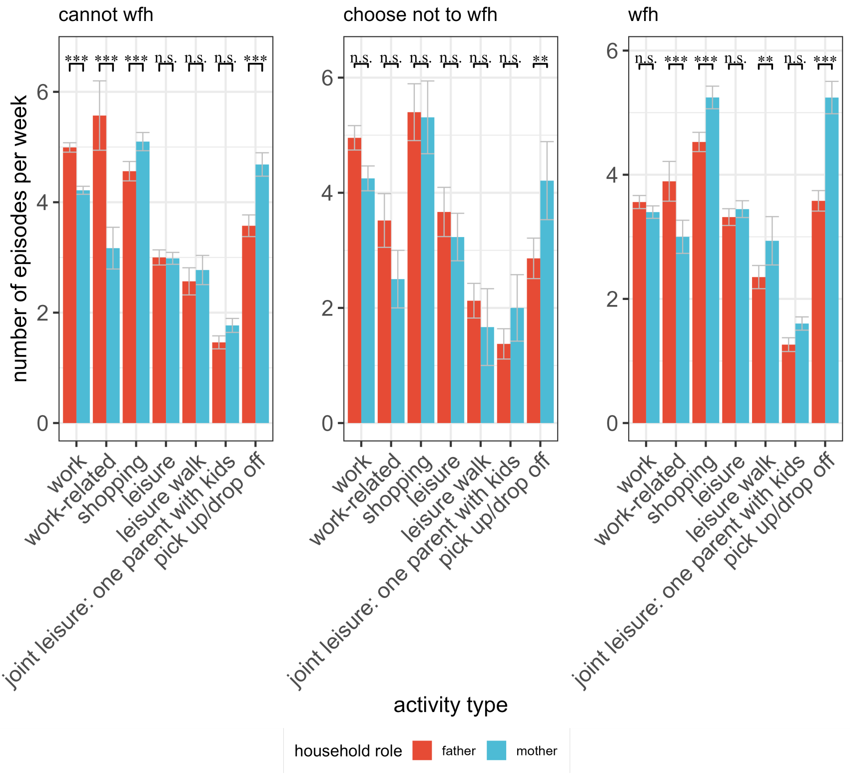


Figure 4.8: Weekly number of episodes by household role, activity, and telecommuting status of employed respondents

Fathers tend to engage in more work-related activities compared to mothers, which is more pronounced among those who cannot work from home. However, among telecommuters, this trend is only true for work-related activities, and the difference is not statistically significant among those who choose not to work from home. In line with the findings from the weekly time use analysis, mothers also conduct significantly more escorting activities, with the largest difference found among telecommuters.

Similar to the results of the two previous analyses, Figure 4.9 indicates that there is a significant difference in the duration of work episodes conducted by mothers and fathers.

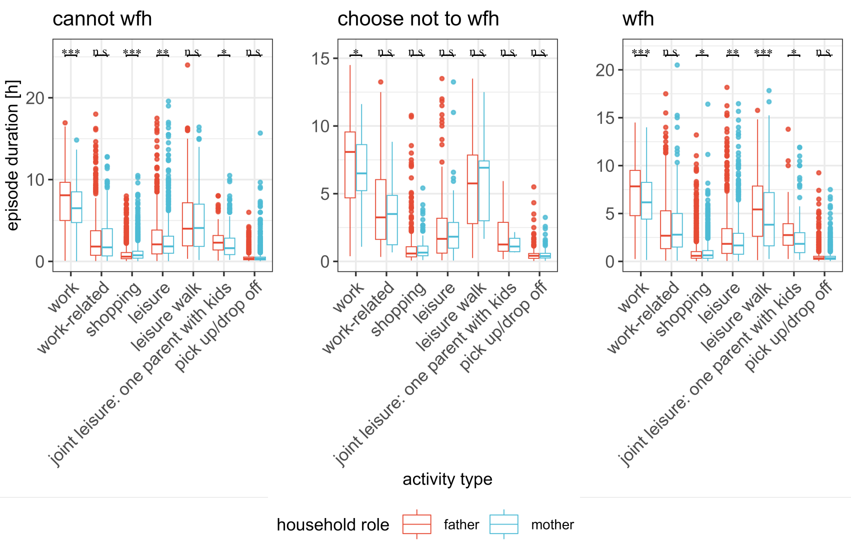


Figure 4.9: Duration of episodes by household role, activity, and telecommuting status of employed respondents

Specifically, mothers tend to conduct significantly shorter work episodes compared to fathers. However, when comparing the duration of work-related episodes, no statistically significant difference can be observed between the two genders. Interestingly, when it comes to leisure activities, fathers tend to conduct significantly longer episodes compared to mothers, particularly among those who cannot work from home and those who choose to telecommute. On the other hand, mothers tend to conduct longer shopping episodes, however, this difference is not statistically significant concerning those who choose not to telecommute.

Another noteworthy finding is that fathers tend to conduct longer joint leisure activities with their children, regardless of whether they choose to telecommute or cannot work from home. This finding is consistent across both groups.

In summary, there are considerable differences in activity patterns between mothers and fathers, with mothers taking on significantly more unpaid labor like shopping and escorting activities compared to fathers, who, in turn, tend to work more and longer hours as well as conduct longer leisure activities, both alone and with their children. Telecommuting does not alleviate the gender difference in care work allocation but rather enhances them, corroborating findings presented in the literature review (see Chapter 2.1.1). The findings based on the MOP reveal that in Germany, there are still gender inequalities concerning care work distribution in the household. Indeed, Germany ranks below average among member states of the European Union concerning the equality of care work (Barbieri et al. 2023).

4.4 Implications for scheduling approach

The results presented here as well as in the literature imply several requirements that an activity scheduling model has to meet to adequately account for telecommuting behavior.

First, the scheduling model must be capable of distinguishing between different types of **work arrangements** and corresponding work location preferences. This involves creating agent profiles that not only specify whether an individual can telecommute but also capture the choice of whether they want to work from home. As suggested by Singh et al. (2013) and more recently by Heimgartner and Axhausen (2024), the actual number of telecommuters in a given setting is the result of a process that encompasses the option of an individual to telecommute, their preference to telecommute and finally the frequency at which employees choose to work from home.

Second, the model must account for **different household types**, recognizing that those living alone may have different desires and face fewer constraints than those living with others. For example, single-person households might have more flexibility in their scheduling and a higher need for social activities to mitigate isolation, whereas family households might prioritize schedules that align with family obligations like childcare.

Third, in order to create a more accurate simulation of daily activity patterns, the model must take into account the intricate dynamics of **household interactions**. This includes the coordination of joint activities and the allocation of household responsibilities, particularly unpaid labor such as shopping and childcare. By incorporating these factors, the model becomes more sensitive to various influences on activity choice behavior. This level of detail is crucial for analyzing the effects of policies, such as those related to childcare, which can significantly alter household routines and individual schedules. By factoring in these interactions, the model not only reflects real-world complexities but also enhances its ability to predict how policy changes might impact telecommuting trends and broader travel demands.

Finally, it is crucial to model the reciprocal influences between telecommuting and other activities within an agent's daily routine. Since telecommuting can alter the frequency and duration of non-work activities — and vice versa — these **activities must be considered concurrently** rather than sequentially in a hierarchical order. This allows the model to account for trade-offs between activities and does not impose the need to prioritize one activity over the other.

5 The METiS Modeling Framework

> But her friend Rhea had asked her to undertake his education and ***Metis*** was never one to betray a trust. For a year she taught him how to look into the hearts and judge the intentions of others. (...) How to make a plan and how to know when a plan needed to be changed or abandoned.
>
> ---
>
> Stephen Fry,
> *Mythos*

Previous versions of the model presented in this chapter have been presented and in part published in proceedings in the following contributions:

Reiffer, A.S. *Generierung von Aktivitätenplänen für agenten-basierte Nachfragemodelle (Translation: Generating Activity Schedules for Agent-Based Travel Demand Models)* presented at the Heureka Congress, Stuttgart, March 2024.

Reiffer, A.S., Vortisch, P. *Framework for Generating 7-Day Activity Schedules Considering Household Interactions* presented at the 103rd Annual Meeting of the Transportation Research Board, Washington D.C., January 2024.

Reiffer, A.S., Vortisch, P. *Estimating Household-Level Time-Use within a Week Activity Scheduling Framework -– Application of the MDCEV Model* presented at the 11th hEART - Symposium of the European Association for Research in Transportation, Zurich, September 2023.

With the requirements for the activity scheduling model established, this section will introduce the modeling framework that integrates these factors. The modeling framework developed in this thesis is called METiS - **M**ultiagent **E**stimation of

Time-Use and **S**cheduling[1]. The model generates weekly activity schedules for each agent of a synthetic population within the context of their household. The next chapter will provide the prerequisites required before detailing the parts of the model. It first presents an overview of the proposed framework and its placement within agent-based travel demand simulation. Further, the choice model of telecommuting is presented which has to be applied before generating the time use of each household as it serves as one of the inputs. The subsequent sections will detail the individual parts of the model, including model formulation, estimation, and calibration procedures as well as results. This chapter concludes with a discussion of the overall framework and results.

5.1 Prerequisites

5.1.1 Model framework overview

The modeling framework and its placement within an agent-based travel demand simulation are presented in Figure 5.1, where the work presented in this thesis is highlighted in blue and the preliminary and subsequent steps of the travel demand simulation are depicted in gray.

The model is applied to all agents created through a population synthesis. The population synthesis generates a set of agents within their household context, representing individuals within a model region, including their socio-demographic characteristics, e.g., age, gender, household income, and number of cars. To account for telecommuting, additional information on employment is necessary. For this purpose, firms in the model region have to be represented as individual entities within the model, including information on their industry sector and size. Based on the characteristics of both the agent and the firms, agents can be allocated to firms within the model area. A possible allocation procedure is presented by

[1] In Greek mythology, Metis was regarded as the goddess of wisdom, foresight and strategic planning, making her a fitting namesake for this model.

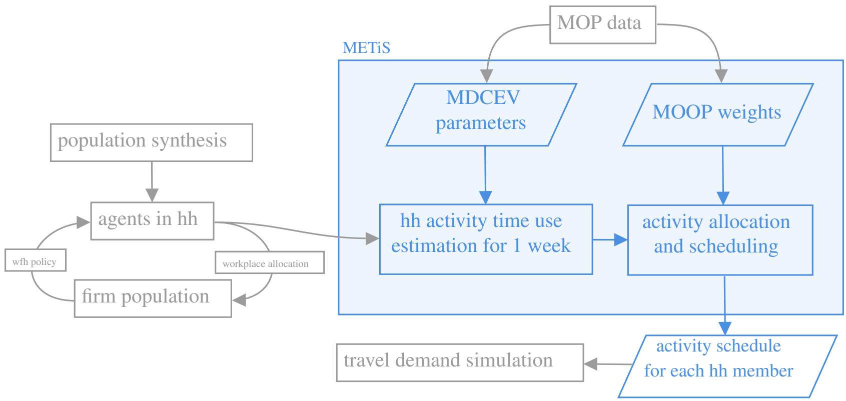


Figure 5.1: METiS modeling framework

Agriesti et al. (2022). Given the assignment of agents to a firm in the model, each firm can assign a work-from-home arrangement to each agent, representing the option of each agent to telework. If the modeling framework is applied based on a population in which this information is not integrated with adequate detail, the telework option can be assigned to each agent probabilistically or based on an econometric model of the agent's socio-demographic characteristics.

For each agent who has the option to work from home, a discrete choice model on the decision of whether they generally choose to telework is applied, regardless of if they do so in the modeling week. This distinction is made because the overall preference to telecommute is different from the teleworking frequency choice (Singh et al. 2013, Heimgartner and Axhausen 2024). Opting to work from home can be regarded as a long-term decision as it includes decisions on whether to equip the home with a workplace and invest in technical gear in case the employer does not provide it. The choice of how often to telecommute depends on other activites to be scheduled and the schedules of other household members. This choice is thus represented separately in the model. These steps conclude the preliminary steps necessary to apply the modeling framework.

Based on this established input the activity schedules of the agents are generated in two steps. Although this modeling framework is constructed in a stepwise manner, the steps don't pertain to generating individual activities that are added in each step but to the level of detail of the schedules. In the first step, the activity time use of the entire household in a week is generated using a Multiple Discrete Continuous Extreme Value Model. The result of the MDCEV model application is the aggregated household time use for the period of one week. Subsequently, the activities are allocated to household members and scheduled throughout the week based on two back-to-back optimization problems. Firstly, activity schedules are created on a coarse level of detail using a multi-objective optimization problem, resulting in a *schedule frame*. In this problem, each time slot of a week can be assigned exactly one activity. This assignment is performed simultaneously for all household members thus regarding interactions and trade-offs in activity allocation. Secondly, the schedule frames are finetuned to generate schedules in which activities are represented with an exact duration and start time. This is achieved through constraint programming. It should be noted that only the activity type is modeled and not the location or mode. These steps have to be conducted in the subsequent travel demand simulation. Although activity scheduling is not conducted separately from location and mode choices, the data used to develop the METiS model does not include information on the locations at which the activities are conducted. Therefore, these choices have to be modeled after generating the activity schedules.

The subsequent sections of this chapter are structured as follows: First, the calibration method is introduced. Then the MDCEV model formulation, the forecasting approach, and the results are presented. Following the order of the model, subsequently, the schedule frame model formulation and the result of the weight calibration are presented. The final section is concerned with the formulation of the schedule fine-tuning problem.

5.1.2 Telecommuting choice model

The telecommuting choice model reflects the decision of an agent to work from home if the telecommuting policy of their employer gives them the option to do so. Within the travel demand simulation, the telecommuting choice precedes the activity generation and scheduling model and determines if telecommuting is available as an alternative in the time-use estimation and all subsequent steps. The variables that can be included in the model are limited to the characteristics of the agents that are generated at the time the model is applied. These include household and individual characteristics of the agents. All variables available were included in the model and retained in the final model if they were statistically significant or if their results were intuitive and consistent with results from the literature.

The model is formulated as a binary logit model, with choosing versus not choosing to telecommute as the two alternatives. The sample is a subsample of the dataset described in the data chapter (section 3.3), consisting of those employed respondents in the MOP who have the option to telecommute. The results of the model estimation are presented in Table 5.1.

Table 5.1: Results of the binary logit model on the choice to telecommute

Parameter	variable	Estimate	Robust t-ratio
alternative specific constant	wfh coice	1.914	6.163
household role	female without young children*	0.175	1.126
	mother with young children*	0.631	2.241
	father with young children*	-0.191	-0.903
home location	urban settlement structure	0.254	1.309
household car ownership	yes	-0.275	-1.176
transit ticket ownership	yes	-0.101	-0.592
household type	single household	-0.699	-4.044

$N = 2,297$
$\rho^2 = 0.4234$
log-likelihood (0) $= 1,592.16$
log-likelihood (final) $= 918.07$

*Note: young children are those aged 10 and younger

The alternative specific constant is significant and positive, indicating - all things equal - there is a preference for choosing to telecommute over choosing not to telecommute if given the option. Consistent with findings presented in the literature and the descriptive analyses on telecommuting of parents (see section 4), mothers with young children are most likely to choose to telecommute compared to other household roles, whereas fathers of young children are least likely to do so. Living in an urban area has a positive effect on the choice to telecommute. This is consistent with previously presented research by Beck et al. (2020), Haider and Anwar (2023), who argue that those living in urban and densely populated areas offer sufficient opportunities for social interactions which those living in rural areas might seek out at their workplace. Both car and public transit pass ownership are negatively associated with telecommuting which is intuitive as well as consistent with literature. Budnitz et al. (2020) also finds that telecommuters have a slightly lower car ownership rate compared to commuters. These results corroborate those presented by Habib (2020) concerning transit pass ownership. The negative relationship is logical, as these tickets are most beneficial for regular commuters who frequently use public transport. Finally, those living alone are less likely to choose to telecommute compared to those living in other household compositions, corroborating findings that those in single households look for social interactions at work which they miss out on when working from home.

These findings are underscored by the average marginal effects (AME) of the model variable, which are presented in Figure 5.2. The AME values highlight that being a mother with young children significantly increases the likelihood of telecommuting, while being a father with young children reduces it. This difference likely reflects traditional gender roles and childcare responsibilities, where mothers may be more inclined or required to work from home. Further, living in a single-person household is strongly associated with a lower likelihood of telecommuting, highlighting that those living alone are most likely to benefit from social interactions in the office.

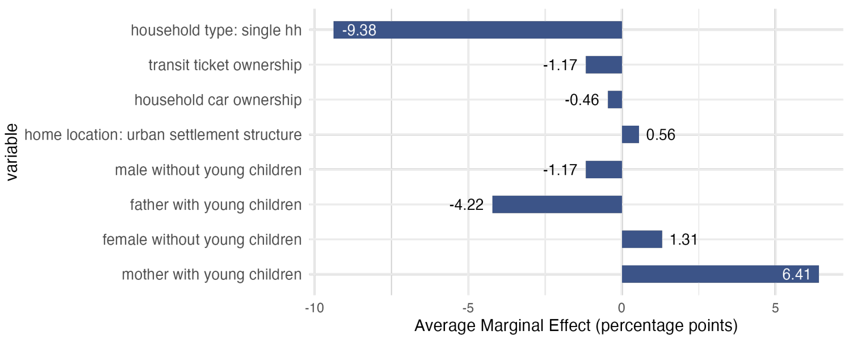


Figure 5.2: Average marginal effects for each explanatory variable in the telecommuting choice model

Finally, owning a car or a transit ticket slightly reduces the likelihood of telecommuting, indicating that those that are well equipped to commute to work are less likely to work remotely.

Although the findings are consistent with those presented in the literature, the model does not necessarily reflect causation. Especially the findings on residential loctation and mobility tool ownership may be the result of self-selection into (not) working from home. For example, the results could reflect a true impact of car ownership on not telecommuting or there could be a simultaneous effect of a preference for driving on the choice to own a car and the choice to commute to work. These effects have been identified in past studies regarding residential location, mobility tool ownership, and mode choice (see e.g. Schmid et al. (2023)). Future research could further investigate these self-selection effects and disentangle causality from correlation in telecommuting behavior.

5.2 Activity Time-Use Estimation

In the METiS model, the activity time use of a household for one week is generated econometrically based on an MDCEV model. The model accounts for socio-demographic information of each household member as well as characteristics of the household. The parameter estimation was conducted using the R software (R Core Team 2022), using the Apollo package (Hess and Palma 2019, Hess and Pamla 2021). The parameters were estimated using maximum likelihood estimation based on the bgw algorithm presented by Bunch et al. (1993). The following sections describe the model specifications for single households, couple households, and finally family households. Subsequently, the forecasting approach applied in the modeling framework is presented. In this work, multi-adult households are not regarded as a separate household type. The majority of these households consist of students sharing accommodation. Thus, the prevalence of telecommuting in these households is relatively low. Additionally, little is known about the household interactions within these households, especially regarding the allocation of chores. Individuals from multi-adult households are thus regarded as independent individuals and modeled as if they lived in single households.

5.2.1 Model specifications

As presented in the literature review as well as the empirical analyses of the MOP data, it is imperative to account for different household types within this model step. In this modeling framework, separate models are estimated for each household type. Although the household type could be considered as a characteristic of the household in the model specification, accounting for nuanced influences of the household type on activities potentially depending on other characteristics, which means that numerous interaction parameters would have to be defined and tested.

As previously established, it is further important to account for household interactions. The approach presented in this work estimates the time use of the entire household while taking into account individual-specific alternatives (ISA) for activities that cannot be shared between household members. For example, the work activities of household member h_1 cannot be conducted by h_2. However, activities like shopping can be done by all household members, and the allocation will depend on the constraints of each household member. In households with more than one member, the ISA are assumed to be separate activities influenced by the socio-demographic characteristics of the respective individual. The parameters associated with ISA, however, are the same across the activity type. For example, in a 2-person household in which both household members are employed, work activities for household members h_1 and h_2 are regarded as alternatives k_{w1} and k_{w2}. The parameterization of the baseline utility ψ_{w1} and ψ_{w2} accounting for age and gender in this example are then given by:

$$\begin{aligned}\psi_{w1} &= \beta_w + \beta_{w,age} \cdot \text{age}_1 + \beta_{w,gender} \cdot \text{gender}_1 \\ \psi_{w2} &= \beta_w + \beta_{w,age} \cdot \text{age}_2 + \beta_{w,gender} \cdot \text{gender}_2\end{aligned} \tag{5.1}$$

By not modeling activity time use independently for each household member, this approach is less likely to overestimate overall household-level activities.
All models are formulated based on a α-γ-profile and parameterization of both the baseline utility ψ_k as well as the satiation parameters γ_k were tested to account for socio-demographic nuances. However, parameterization of the baseline utility has led the models to generate poor results, thus only constants are estimated for ψ_k. The household time use budget depends on its size. Given a household $H = \{h_1, h_2, ..., h_n\}$, then the household budget is determined by $E = 10,080 \cdot |H|$, i.e., the number of minutes in a week multiplied by the household size. In all models, the following socio-demographic characteristics of the household or its members were included as parameters in the estimation and kept either based on the statistical significance of the t-ratio or based on whether they provided a meaningful impact on the result:

- income (binary, low/high)
- car ownership (binary, yes/no)
- region type of the home (binary, urban/rural)
- age (categorized)
- gender (binary, male/female)
- employment (binary, full-time/part-time)
- public transit ticket ownership (binary, yes/no)

Although the data would allow for the model to account for different types of employment (full-time, part-time, unemployed, etc.), the best model results were achieved by only considering the extent of employment, i.e. full-time and part-time. Similarly, the effect of different education levels was tested in the models, however, this did not meaningfully impact the results. This may seem unexpected, given that working from home is often linked with higher levels of education. However, a higher level of education primarily determines the types of jobs that can be done remotely. The models presented here suggest that education does not impact how much someone telecommutes.

As proposed by Bhat (2008), it is useful to constrain the values that the satiation parameters can assume during estimation. As the γ-parameters have to be strictly positive, it is defined in the models as $\gamma^* = e^\gamma$. Further, α has to be ≤ 1 and > 0. This is achieved through defining $\alpha^* = \frac{1}{1+e^\alpha}$.

Single households

The simplest model is the single household model, as only one person's activity time use is estimated. The household budget is $E = 10,080$, i.e., the number of minutes in a week. During model development, several configurations of the model were tested, with an MDCNEV model with home as the outside good providing the best results. In this model, all work activities are grouped into one nest. The final nesting structure is illustrated in Figure 5.3.

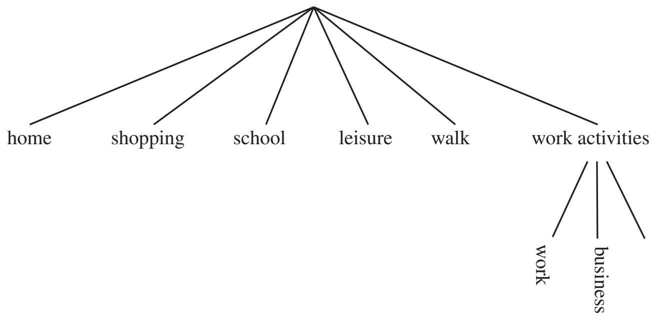


Figure 5.3: Nest structure of the MDCNEV model for single households

Some activities are not available to everyone and their availability is regarded in the model specification. For example, as described earlier, work from home is only available to those who have the option and choose to do so. The activities with limited availability are *work, wfh, business, school*.

The estimation results of the MDCNEV model for single households are presented in Table 5.2. The model identifies shopping as the activity most likely to be chosen, evidenced by its highest baseline utility value. However, shopping also exhibits the smallest satiation parameter, suggesting that while it is frequently selected, individuals tend not to spend a significant amount of time on it. In contrast, mandatory activities such as work and education are less likely to be chosen, indicated by lower baseline utility values, but display high satiation parameters, meaning considerable time is invested in these activities once they are selected.

Table 5.2: Estimation results of the single household MDCNEV model

Parameter	Activity	Estimate	Rob. t-stat	Parameter	Estimate	Rob. t-stat
satiation				baseline utility		
	work	3.320	31.421		-4.509	-123.925
	business	2.734	12.654		-5.393	-122.414
	wfh	3.619	21.105		-5.558	-133.032
	school	3.674	19.960		-8.724	-40.929
	shopping	-0.717	-6.871		-1.826	-20.314
	leisure	0.729	8.916		-3.208	-79.067
	leisure walk	2.360	21.865		-4.849	-121.870
age: under 35	shopping	-0.285	-3.625			
	leisure	0.255	2.683			
	leisure walk	-0.228	-1.339			
age: over 60	work	-0.110	-1.272			
	business	0.262	1.775			
	wfh	0.141	0.832			
	shopping	0.269	4.543			
	leisure	0.105	1.806			
	leisure walk	0.311	2.741			
full-time employed	Work	0.665	8.800			
	business	0.535	3.680			
	wfh	0.089	0.604			
	shopping	-0.502	-8.560			
	leisure walk	-0.247	-2.264			
high income	work	-0.232	-3.328			
	business	-0.261	-2.060			
	wfh	-0.235	-1.939			
	shopping	-0.098	-2.055			
	leisure	-0.063	-1.055			
	leisure walk	0.104	1.126			
urban home location	work	-0.153	-1.669			
	business	-0.505	-2.803			
	leisure	0.179	2.449			
car	wfh	-0.239	-1.491			
θ_{work}		0.588	29.243			
$\theta_{non-work}$		1.000	-			
α		-24.349	-			
N:	2,312					
log-likelihood (start):	-45627.86					
log-likelihood (final):	-31069.1					

Notably, walks, though less likely to be chosen as an activity, receive more time investment compared to other leisure activities, underscoring the distinct nature of this activity. Socio-demographic factors also significantly influence how individuals allocate their time across activities. Age appears to play a crucial role; younger individuals (below 35 years) are less likely to invest time in leisure walks compared to those over 35, with individuals over 60 showing the highest propensity to engage in such walks. This trend is similarly observed in shopping activities. Regarding work, older individuals generally invest less time in on-site work but more in work-from-home compared to their younger counterparts. Income levels inversely affect the time spent on almost all activities except leisure walks, suggesting that higher-income individuals, likely living in more walkable and aesthetically pleasing areas, prefer walking.

Urban residents tend to spend less time on work or business activities, a reasonable outcome considering shorter commute times in urban settings as captured in the time use data. These residents are also more inclined to engage in leisure activities, benefitting from greater accessibility to such opportunities. Additionally, car ownership influences the allocation of time; those with a car are less likely to work from home, potentially perceiving their commute as less burdensome compared to those without access to a car and thus gaining less utility from telecommuting.

Couple households

Compared to single households, the couple household model is more complex as it has to account for both individual-specific alternatives as well as household-level alternatives. The ISA are activities that can only be allocated to a specific household member both in the activity generation as well as the scheduling part. To allow for correlation among the ISA, a nest is defined for each household member. The θ-parameter is the same across both nests. Estimation with a nest for the household level parameters yielded a θ-parameter of 1, indicating that the nest is unwarranted. Thus, this was kept fixed at 1 during estimation. In addition to the ISA, this model also accounts for joint activities, namely joint shopping and joint leisure. The final nest structure of the MDCNEV model on couple households is presented in Figure 5.4. At this stage, shopping activities are only regarded at the household level and not specified as ISA. The shopping activities generated in this step are allocated to the specific household members in the subsequent scheduling model.

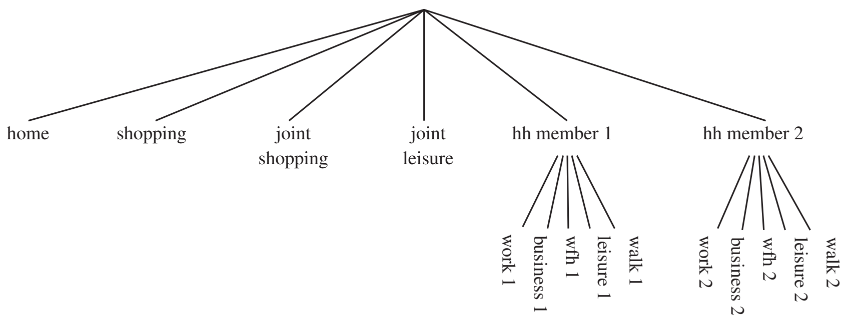


Figure 5.4: Nest structure of the MDCNEV model for couple households

In the MDCEV model on couple households (see Table 5.3), the outcomes for shopping and work closely align with those observed in the single household model in terms of baseline utility, which indicates the likelihood of choosing these activities, and the time dedicated to them once chosen. Similarly, leisure walks remain less frequently selected, but significantly more time is devoted to them compared to other leisure activities, both performed alone or jointly. A new insight in the couple model is that joint shopping is less preferred, but if chosen, it generally involves a longer duration.

Socio-demographic variables exhibit noteworthy impacts on activity participation and duration. Income negatively influences the time dedicated to working from home, suggesting that individuals with higher incomes, likely possessing greater career responsibilities or aspirations, prefer working on-site. Car ownership continues to reduce the time spent telecommuting and is also negatively correlated with time spent on shopping and leisure activities; however, those with cars tend to allocate more time to business activities. Echoing the single household findings, urban dwellers spend less time on work and shopping, which could be attributed to shorter commute times compared to their rural counterparts.

Table 5.3: Estimation results of the couple household MDCNEV model

Parameter	Activity	Estimate	Rob. t-stat	Parameter	Estimate	Rob. t-stat
satiation				baseline utility		
	work	2.479	30.113		-3.131	-81.330
	business	0.845	3.188		-3.969	-82.072
	wfh	2.998	11.776		-4.055	-91.002
	school	2.423	11.490		-7.240	-34.156
	shopping	-1.699	-8.375		-0.608	-7.624
	leisure	0.121	0.955		-2.194	-68.982
	leisure walk	1.344	20.850		-3.487	-81.610
	joint shopping	0.241	1.793		-2.622	-72.218
	joint leisure	0.476	2.266		-1.971	-51.974
high income	wfh	-0.379	-3.043			
	business	-0.015	-0.107			
car	wfh	-0.077	-0.312			
	business	0.336	1.525			
	leisure	-0.269	-2.118			
	shopping	-0.200	-1.520			
	joint shopping	-0.099	-0.711			
	joint leisure	-0.183	-0.859			
urban home location	work	-0.150	-1.869			
	shopping	-0.124	-1.575			
age: over 60	business	-0.231	-1.498			
full-time employed	work	0.208	2.779			
	wfh	-0.504	-4.467			
	business	0.457	2.908			
	leisure	0.058	0.949			
	leisure walk	-0.209	-2.465			
gender: male	work	-0.015	-0.246			
	business	0.243	2.101			
	leisure	-0.016	-0.335			
	leisure walk	0.088	1.235			
θ_{ISA}		0.593	28.973			
θ_{hh}		1	-			
α		0	-			
N:	1,484					
log-likelihood (start):	-50518.17					
log-likelihood (final):	-42170.09					

Age differences are particularly pronounced in business activities, where older individuals (above 60 years) tend to invest less time compared to younger ones (under 60 years). Full-time employment is logically linked to more time spent on work and business activities, and previous studies have documented a negative association with working from home. Gender differences also emerge, with men slightly less likely to spend time on work but more on business activities compared to women. Additionally, men are more inclined to engage in leisure walks

Family households

By far the most complex model is the MDCNEV model for family household time use. There are additional joint activities that are regarded in the model as well as ISA for up to five household members. The nest structure of the family time use model is illustrated in Figure 5.5

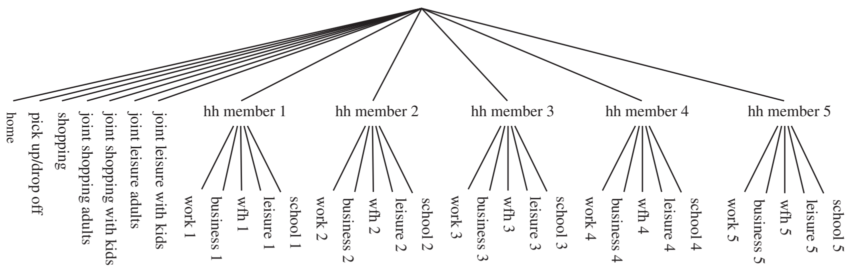


Figure 5.5: Nest structure of the MDCNEV model for family households

The model, again, accounts for joint shopping and joint leisure activities. These are in turn divided into joint activities, in which only adults participate, and those in which all family members can participate. In an earlier model, these were also split into joint activities with one adult and children, and joint activities, in which the entire family participates. However, there are comparatively few observations of joint family activities. Given the already high complexity of the model and the many alternatives that are regarded, these joint activities were merged into one alternative. This model further explicitly includes pick-up/drop-off activities, which, as the empirical analysis has shown, are especially important in the context of telecommuting in households with children. The results of the parameter estimation of the MDCNEV model on family households are presented in Table 5.4.

Table 5.4: Estimation results of the family household MDCNEV model

Parameter	Activity	Estimate	Rob. t-stat	Parameter	Estimate	Rob. t-stat
satiation				baseline utility		
	work	3.525	99.795		-6.053	-296.059
	business	2.293	17.581		-6.770	-180.530
	wfh	3.387	45.655		-6.693	-184.037
	school	4.128	132.149		-6.534	-307.268
	shopping	1.188	-4.046		-2.181	-8.450
	pick up/drop off	0.045	-0.668		-4.466	-69.817
	leisure	2.144	22.260		-5.520	-306.566
	joint shopping adults	1.744	12.505		-5.820	-83.564
	joint shopping family	1.941	20.830		-6.040	-84.502
	joint leisure adults	1.037	8.770		-6.351	-76.907
	joint leisure family	0.797	10.371		-6.291	-80.109
age: under 35	leisure	0.35771	5.24522			
age: over 60	leisure	-0.02132	-0.09868			
high income	shopping	-0.23926	-2.95650			
	joint shopping adults	0.11336	0.69461			
	joint leisure adults	-0.26677	-1.78417			
full-time employed	wfh	-0.20224	-1.85387			
	business	0.37799	2.19276			
	leisure	-0.29916	-3.89804			
gender: male	work	0.21578	4.58039			
	wfh	0.10849	1.03498			
	business	0.40596	2.83517			
	leisure	0.12512	2.18662			
urban home location	shopping	0.16796	1.56775			
	leisure	-0.14909	-1.64583			
θ_{ISA}		0.519	49.537			
θ_{hh}		1	-			
α		0	-			
N:	662					
log-likelihood (start):	-38970.34					
log-likelihood (final):	-30557.04					

In the analysis of baseline parameters for family household activities, the MDCEV model highlights some distinctive trends. Shopping again surfaces with the highest baseline utility, indicating a high likelihood of being chosen, yet it is accompanied by one of the lowest satiation parameters, signifying that minimal time is typically devoted to it. A similar pattern is observed in pick-up/drop-off activities, which hold the second-highest baseline utility after shopping but exhibit even lower satiation, reflecting the brief nature of these engagements.

Work activities, in contrast, show much higher satiation parameters compared to discretionary activities such as (joint) leisure or joint shopping, suggesting that significantly more time is allocated to work once these activities are selected.

Socio-demographic factors continue to play a crucial role in activity engagement. Younger individuals, particularly those under 35 years, tend to spend more time on leisure activities compared to older counterparts. Income levels influence shopping behaviors distinctly; higher-income individuals dedicate more time to joint shopping activities but less to individual shopping, and similarly, less time is spent on joint leisure activities among parents. This suggests a shift in priority towards shared experiences in shopping but a reduction in leisure time spent together.

Full-time employment reinforces its influence on activity choices, positively affecting the time spent on business activities while negatively impacting the likelihood of telecommuting. Gender differences also emerge within family households; males tend to allocate more time to both work and leisure activities, indicating varying commitments based on gender roles.

Urban living affects activity patterns distinctly, with urban dwellers spending more time on shopping activities and less on individual leisure activities. This could be attributed to the accessibility of shopping venues. Further, increased accessibility decreases travel time which is reflected by a lower satiation parameter in the model.

5.2.2 Forecasting

The application of the MDCEV or MDCNEV model within an activity-based modeling framework requires an efficient forecasting procedure. Pinjari and Bhat

(2010b) present a forecasting algorithm based on the Karush-Kuhn-Tucker conditions (see section 2.2.2) for models formulated as γ- and α-γ-profiles. One of the required inputs to the algorithm is a set of draws from the distribution of the error terms. The MDCEV model is based on the assumption that the error terms follow a type I extreme value distribution and thus generating the draws is simple. This gets more complex for the MDCNEV model as the error terms now have to be drawn from a generalized extreme value distribution. Calastri et al. (2017) present an approach to approximate these draws. The algorithm proposed by Pinjari and Bhat (2010b) and the extension for the MDCNEV model presented by Calastri et al. (2017) serve as the basis for the forecasting algorithm applied in this model. However, the general estimation and forecasting approaches have some shortcomings, the most notable one being the lack of considering bounds in consumption or time investments, which means that the model generates activity time use patterns where the time use of an activity can be unrealistically low or high. This is hard to control for in all cases given the nature of the error terms. There have been approaches to mitigate this issue. Bhat et al. (2020) investigate bounds on unobserved budgets and propose to truncate the distribution of the error draws. Again, this is simple enough for MDCEV models but complex for GEV distributions. Saxena et al. (2021) integrate additional constraints into the Karush-Kuhn-Tucker conditions to impose an upper bound on time allocation. In addition to the model formulation, they also present a forecasting algorithm that can account for the imposed bounds. Ideally, the model would be both estimated and forecasted using this formulation. However, there is no open-source software package available that allows for the estimation of this model. As a first step, in this work, the forecasting procedure is thus implemented based using parameters estimated based on the original MDCEV and MDCNEV formulation. The forecasting algorithm generates a set of chosen alternatives and computes the optimal time allocation for each activity. The bounds are accounted for by checking whether any activity is assigned more time than a previously determined maximum time limit. If this is the case, the alternative is added to a set of chosen alternatives and the duration of this alternative is set to the maximum duration. Then the algorithm is re-initialized, however, the previously chosen alternatives are removed from the choice set and the budget is adapted accordingly. All steps of the forecasting procedure applied in this work are presented in algorithm 1. The time-use forecasting algorithm is implemented in Python 3.

Algorithm 1 : MDCEV forecasting algorithm

Data : Input data, model parameters, η, E
Result : Optimal allocations and choice set $\mathbb{M}$

Initialize $\mathbb{K} = \{1, 2, \ldots, K\}$, the set of choice alternatives;
Simulate error terms ε_k and compute baseline utilities ψ_k and translation parameters γ_k for all $k \in \mathbb{K}$;
Read available budget E;
Arrange ψ_k in decreasing order, placing the utility of the outside good first;
Arrange indices k of choice alternatives in the same order;
Assume the outside good is chosen: Initialize $\mathbb{M} = \{1\}$ and set $M = 1$;
for $k = 2$ **to** K **do**
    Compute λ: $\lambda = \frac{\psi_1 + \sum_{k=2}^{M} \psi_k \gamma_k}{E - \sum_{k=2}^{M} x_k^{\min} + \sum_{k=2}^{M} \gamma_k}$
    if $\lambda > \psi_k$ *or* $k = K$ **then**
        Break;
    end
    $M = k$;
end
for *each* $m \in \mathbb{M} \setminus \{1\}$ **do**
    Compute inside good allocations: $t_m^* = \gamma_m \left(\frac{\psi_m}{\lambda} - 1 \right)$
end
Compute outside good allocation: $t_1^* = \frac{\psi_1}{\lambda}$
Set $t_k^* = 0$ for all $k = M + 1, \ldots, K$;
for *each* $m \in \mathbb{M} \setminus \{1\}$ **do**
    if $t_m^* > t_k^{\max}$ **then**
        Set $t_m^* = t_m^{\max}$;
        Add m to set $\mathbb{A}$;
    end
end
if $\mathbb{A}$ *is non-empty* **then**
    Reset budget $E = E - \sum_{k \in \mathbb{A}} x_k^{\max}$;
    Reinitialize algorithm with updated budget $\mathbb{K} = \mathbb{K} \setminus \mathbb{A}$;
else
    Return t_k^* for all $k \in \mathbb{K}$ and exit;
end

5.3 Activity Scheduling

After determining the aggregated weekly time use of a household, the model moves on to generate episodes and a schedule for each household member. This is done in two steps. In the first step, a schedule frame for each household member is created, which is fine-tuned in the second step. In both steps, an optimization problem is solved by considering all household members' schedules at the same time to allow for trade-offs between scheduling choices. Both optimization problems are solved using IBM ILOG CPLEX Optimization Studio, version 22.1.1.0 (ibm 2022) and called using the API for Python 3.

5.3.1 Schedule Frames

Based on the previously determined time allocation per household tu_a, the assignment of activities to individual household members is carried out. The formulation differs depending on the household type, as, for example, single individuals do not require coordination with other household members, and households with children have different needs, preferences, and constraints compared to 2-person households. For brevity, the following focuses on the model for couple households.

The goal of the model is to capture the motivation and underlying behaviors in the assignment and scheduling of activities within a household. Various goals conflict with each other, such as household chores limiting the number or duration of leisure activities. To address (partially) conflicting goals, multi-criteria optimization models are suitable, where multiple variables are optimized in the objective function. Table 5.5 provides an overview of the sets, parameters, and variables included in the schedule frame model.

Table 5.5: Sets, parameters, and variables in the schedule frame problem

$\mathcal{A}$	Activity type
$\mathcal{A}_i \subseteq \mathcal{A}$	Individual specific activity
$\mathcal{A}_w \subseteq \mathcal{A}$	Work activities
$\mathcal{A}_{wfh} \subseteq \mathcal{A}_w$	Work from home activities
$\mathcal{A}_s \subseteq \mathcal{A}$	Shopping activities
$\mathcal{A}_{ch} \subseteq \mathcal{A}$	Chores
$\mathcal{A}_h \subseteq \mathcal{A}$	In-home activities
$\mathcal{A}_l \subseteq \mathcal{A}$	Leisure activities
$\mathcal{A}_{j,l} \subseteq \mathcal{A}$	Joint leisure activities
$\mathcal{A}_{j,sh} \subseteq \mathcal{A}$	Joint shopping activities
$\mathcal{A}_{j,a} \subseteq \mathcal{A}$	Joint activities of adults
$\mathcal{A}_{j,f} \subseteq \mathcal{A}$	Joint activities of adults and children
$\mathcal{H}$	Household members
$\mathcal{H}_w \subseteq \mathcal{H}$	Employed household members
$\mathcal{H}_{wfh} \subseteq \mathcal{H}_w$	Household members working from home
$\mathcal{H}_a \subseteq \mathcal{H}$	Adult household members
$\mathcal{H}_k \subseteq \mathcal{H}$	Underage household members
$\mathcal{D}$	Day of the week
$\mathcal{D}_w \subseteq \mathcal{D}$	Work days
$\mathcal{D}_{PD,wfh} \subseteq \mathcal{D}$	Preferred days to work from home
$\mathcal{D}_{UPw} \subseteq \mathcal{D}$	Non-preferred office days
$\mathcal{D}_s \subseteq \mathcal{D}$	Days of the week where shops are open
$\mathcal{T}$	Time of day
$\mathcal{T}_s \subseteq \mathcal{T}$	Time of day when shops are open
$x_{a,h,d,t} \in \{1,0\}$	Decision variable; 1 if a is assigned to h on d at t
$\lambda_{h,i,d,t} \in \{1,0\}$	Auxiliary variable; is 1 if household member h and household member i are at home on the same day d at the same time t
tu_a	household time-use by activity a
WT^{max}	maximum daily work hours
HT^{min}	minimum duration per day spent at home
AS_a	Activity switches of activity a
CHT	common time spent at home
WDIFF	difference in work hours between days
CHDIFF	difference in allocated chores
OOHT	daily out of home time
WD_{end}	Deviation of realized end of workday compared to preferred end of workday
PWFHD	Time spent working from home on preferred work-from-home days
UPWD	Time spent working in the office on non-preferred office days

In the model approach presented here, the generation of activity plans is formulated as a multi-objective linear optimization problem.

The decision variable to be determined by the optimization, $x_{a,h,d,t}$, takes the value of 1 when activity a is assigned to household member h on day d in time slot t, and 0 otherwise. A graphical representation of the decision variable and its dimensions is presented in Figure 5.6. Theoretically, t can represent different time units, but solving the problem for time units finer than an hour becomes too large to provide results for application in an agent-based model in a reasonable time. Since the result of the previous step indicates time allocation in minutes, it needs to be rounded to the nearest hour for the first optimization problem. Subsequently, the time allocation for home activities is adjusted to match the weekly time budget.

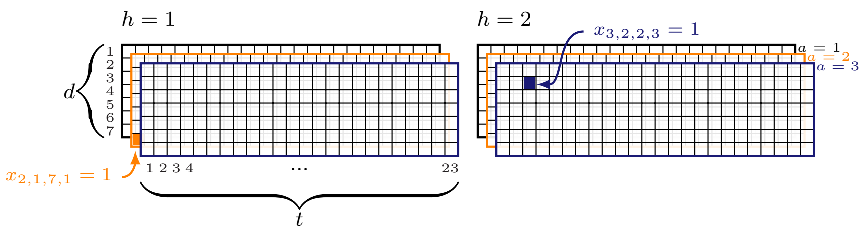


Figure 5.6: Graphical illustration of the different dimensions of the decision variable in the schedule frame model

The following sections will first explain the fundamental constraints needed mathematically to develop activity programs and those imposed by external circumstances. Then, the variables and associated constraints that directly impact the objective function will be listed.

The assignment of activities to time slots in an agent's schedule is unique, meaning that only one activity can be assigned to each time slot on a given day. This requirement is ensured by the following equality constraint:

$$\sum_{a \in \mathcal{A}} x_{a,h,d,t} = 1 \qquad \forall h \in \mathcal{H}, d \in \mathcal{D}, t \in \mathcal{T} \tag{5.2}$$

The following constraint ensures that the sum over all assigned activities equals the previously determined (rounded) time allocation per activity tu_a of the household:

$$\sum_{h \in \mathcal{H}} \sum_{d \in \mathcal{D}} \sum_{t \in \mathcal{T}} x_{a,h,d,t} = \mathrm{tu}_a \qquad \forall a \in \mathcal{A} \tag{5.3}$$

Furthermore, it must be ensured that individually specific activities (work, business, work from home, education, leisure, and walks) can only be assigned to the corresponding household member. This means that an activity a from the set of unique individual alternatives $\mathcal{A}_i$ of household member i cannot be assigned to household member h.

$$x_{a,h,d,t} = 0 \qquad \forall a \in A_i, h, i \in \mathcal{H} : h \neq i, d \in \mathcal{D}, t \in \mathcal{T} \tag{5.4}$$

The execution of joint activities can only occur at the same time, meaning that these activities must be assigned to the same time slot t on a day d in the plans of the involved agents. This is achieved through constraint 5.5, which ensures that the sum of assigned joint activities over all household members is either 0 or > 1. There are two different types of joint activities, those where only adults participate and those where both children and adults participate. Constraint 5.6 ensures, that activities are not assigned to children in the household. Constraint 5.7 handles joint family activities, in which at least one child and one adult have to participate in the activity.

$$\sum_{h \in \mathcal{H}} x_{a,h,d,t} \neq 1 \qquad \forall a \in \mathcal{A}_{joint}, d \in \mathcal{D}, t \in \mathcal{T} \tag{5.5}$$

$$\sum_{h \in \mathcal{H}_k} x_{a,h,d,t} = 0 \qquad \forall a \in \mathcal{A}_{joint,adults}, d \in \mathcal{D}, t \in \mathcal{T} \tag{5.6}$$

$$\sum_{h \in \mathcal{H}_k} x_{a,h,d,t} \geq 1 \Leftrightarrow \sum_{h \in \mathcal{H}_a} x_{a,h,d,t} \geq 1 \quad \forall a \in \mathcal{A}_{joint,family}, d \in \mathcal{D}, t \in \mathcal{T} \tag{5.7}$$

The scheduling of some activities is determined by external influences, which are also reflected in the constraints. For example, shopping activities a_s cannot be performed outside of store opening hours:

$$x_{a,h,d,t} = 0 \qquad \forall a \in \mathcal{A}_s, h \in \mathcal{H}, d \in \mathcal{D} \setminus \mathcal{D}_s, t \in \mathcal{T} \setminus \mathcal{T}_s \tag{5.8}$$

Further, in households with children, the model only allocates shopping activities to adults.

$$x_{a,h,d,t} = 0 \qquad \forall a \in \mathcal{A}_s, h \in \mathcal{H}_k, d \in \mathcal{D}, t \in \mathcal{T} \tag{5.9}$$

Work activities are typically limited to weekdays, from Monday to Friday, and must be performed within a specified time window. Additionally, the daily working time (WT) is legally restricted to eight or ten hours[1]. Note that this can be set

[1] Despite legal working time regulations, data shows clear exceedances of daily working hours. Therefore, the limit is adjusted individually for each agent in the model based on weekly working time tu$_w$.

dynamically for each agent, thus allowing for differentiated work policies in the agent population. These constraints in the planning of work activities are considered through the following constraints:

$$x_{a,h,d,t} = 0 \qquad \forall a \in \mathcal{A}_w, h \in \mathcal{H}, d \in \mathcal{D} \setminus \mathcal{D}_w, t \in \mathcal{T} \setminus \mathcal{T}_w \tag{5.10}$$

$$\sum_{t=1}^{24} x_{a,h,d,t} \leq \mathrm{WT}^{max} \qquad \forall a \in \mathcal{A}_w, h \in \mathcal{H}, d \in \mathcal{D} \tag{5.11}$$

Another constraint that needs to be considered when planning activities is that individuals spend a certain amount of time at home during the day (HT), for activities like physiological recovery or household management.

$$\sum_{t=1}^{24} x_{a,h,d,t} \geq \mathrm{HT}^{min} \qquad \forall a \in \mathcal{A}_h, h \in \mathcal{H}, d \in \mathcal{D}, t \in \mathcal{T} \tag{5.12}$$

Based on the decision variable $x_{a,h,d,t}$, multiple objectives are defined which are to be minimized in the model. The scalarized objective function considers eight variables n that affect the objective function value with weights ω_n:

minimize

$$\begin{aligned}
&\omega_1 \mathrm{AS}_w + \omega_2 \left(\mathrm{AS}_l + \mathrm{AS}_{sh} + \mathrm{AS}_{l,joint} + \mathrm{AS}_{sh,joint}\right) \\
&- \omega_3 \mathrm{CHT} \\
&+ \omega_4 \mathrm{WDIFF} \\
&+ \omega_5 \mathrm{CHDIFF} \\
&+ \omega_6 \mathrm{OOHT} \\
&+ \omega_7 \mathrm{WD}_{end} \\
&- \omega_8 \left(\mathrm{PWFHD} - \mathrm{UPWD}\right)
\end{aligned} \tag{5.13}$$

The first objective is driven by the fact that individuals tend to minimize activity switches to a certain extent. Therefore, this objective aims to minimize switches between activities that require traveling to a new location. Throughout model development, it has become apparent that determining and minimizing constraints for *work, leisure,* and *shopping* yields reasonable results while keeping the model as simple as possible. In the model, the variable AS_a keeps track of the switches between different activities by comparing activities assigned in time slot t and time slot $t+1$. Different weights are associated with work activities and non-work activities to allow for different considerations of the objectives.

$$
\begin{aligned}
\mathrm{AS}_w &= \sum_{a\in\mathcal{A}_w}\sum_{h\in\mathcal{H}}\sum_{t_1\in\mathcal{T}}\sum_{t_2\in\mathcal{T}} |x_{a,h,d,t_1} - x_{a,h,d,t_2}| \\
\mathrm{AS}_l &= \sum_{a\in\mathcal{A}_l}\sum_{h\in\mathcal{H}}\sum_{t_1\in\mathcal{T}}\sum_{t_2\in\mathcal{T}} |x_{a,h,d,t_1} - x_{a,h,d,t_2}| \\
\mathrm{AS}_{sh} &= \sum_{a\in\mathcal{A}_{sh}}\sum_{h\in\mathcal{H}}\sum_{t_1\in\mathcal{T}}\sum_{t_2\in\mathcal{T}} |x_{a,h,d,t_1} - x_{a,h,d,t_2}| \\
\mathrm{AS}_{l,joint} &= \sum_{a\in\mathcal{A}_{l,joint}}\sum_{h\in\mathcal{H}}\sum_{t_1\in\mathcal{T}}\sum_{t_2\in\mathcal{T}} |x_{a,h,d,t_1} - x_{a,h,d,t_2}| \\
\mathrm{AS}_{sh,joint} &= \sum_{a\in\mathcal{A}_{sh,joint}}\sum_{h\in\mathcal{H}}\sum_{t_1\in\mathcal{T}}\sum_{t_2\in\mathcal{T}} |x_{a,h,d,t_1} - x_{a,h,d,t_2}| \\
&\forall\{t_1, t_2 \in \mathcal{T} : t_2 = t_1 + 1\}
\end{aligned}
\tag{5.14}
$$

The second objective in planning activities in households with multiple members is to maximize the time spent together at home (Vuk et al. 2016). This time is expressed by the variable $\lambda_{h,i,d,t}$, which takes the value of 1 when two household members, i and h, have planned a home activity at the same time, and 0 otherwise.

$$
\begin{aligned}
\lambda_{h,i,d,t} &\leq x_{a,h,d,t} \\
\lambda_{h,i,d,t} &\leq x_{a,i,d,t} \\
x_{a,h,d,t} + x_{a,i,d,t} - 1 &\leq \lambda_{h,i,d,t} \\
\mathrm{CHT} = \sum_{h\in\mathcal{H}}\sum_{i\in\mathcal{H}}\sum_{d\in\mathcal{D}}\sum_{t\in\mathcal{T}} &\lambda_{h,i,d,t} \\
&\forall\{a \in \mathcal{A}_h, h, i \in \mathcal{H}, d \in \mathcal{D}, t \in \mathcal{T} : h \neq i\}
\end{aligned}
\tag{5.15}
$$

It is also known from the literature that individuals show a high degree of stability in the duration of their planned work activities (Hilgert et al. 2017), meaning that individuals minimize the difference between the duration of work activity on different days, which is represented by WDIFF:

$$\left| \sum_{t \in \mathcal{T}} x_{a,h,d,t} - \sum_{t \in \mathcal{T}} x_{a,h,e,t} \right| \leq \text{WDIFF} \quad \forall \{a \in \mathcal{A}_w, h \in \mathcal{H}, d, e \in \mathcal{D} : d \neq e\} \tag{5.16}$$

The established need for the model to account for the allocation of household chores to household members is explicitly included in the model. The variable CHDIFF accounts for the difference in allocated chores between the members of a household. Through the weight in the objective function, the importance of the objective can be accounted for. This means that chores are not necessarily distributed equally.

$$\left| \sum_{d \in \mathcal{D}} \sum_{t \in \mathcal{T}} x_{a,h,d,t} - \sum_{d \in \mathcal{D}} \sum_{t \in \mathcal{T}} x_{a,i,d,t} \right| \leq \text{CHDIFF} \quad \forall \{a \in \mathcal{A}_{ch}, h, i \in \mathcal{H} : h \neq i\} \tag{5.17}$$

Furthermore, each agent has a preference on when they would like to start their day, and when they would end it, i.e., when the first out-of-home activity starts and the last one ends, respectively. In the model, this is expressed by the variable OOHT, which is minimized as part of the objective function.

$$\begin{gathered} \text{OOHT} \geq \sum_{t=1}^{\mathrm{P}_s} x_{a,h,d,t} + \sum_{t=\mathrm{P}_e}^{24} x_{a,h,d,t} \\ \forall \{a \in \mathcal{A} \setminus \mathcal{A}_h, h \in \mathcal{H}, d \in \mathcal{D}\} \end{gathered} \tag{5.18}$$

Similarly, the preferred end of the workday can be accounted for. In this case, the desired start does not have to be separately accounted for, because as the work activity is often the first activity of the day, this is already accounted for through the previous objective. The preferred end of the workday is defined as analogous to the preferred end of the last out-of-home activity:

$$\begin{gathered}\mathrm{WD}_{end} \geq \sum_{t=\mathrm{P}_{wd,e}}^{24} x_{a,h,d,t} \\ \forall\{a \in \mathcal{A}_w, h \in \mathcal{H}, d \in \mathcal{D}\}\end{gathered} \tag{5.19}$$

Finally, the model can account for a preference regarding the day on which to work from home. This has been shown to not be distributed equally throughout the week. Considering the day-to-day heterogeneity in the activity scheduler is vitally important as any travel demand management policy concerning telecommuting will need to account for the unequal distribution throughout the week. Although Asgari et al. (2014) has identified that telecommuters are more likely to choose to work from home on mid-week days, more recent findings indicate that this has changed since the COVID-19 pandemic. For example Asmussen et al. (2023) identify Mondays and Fridays as the days on which most choose to telecommute. However, there are no representative statistics available for the distribution of telecommuting days in Germany. Thus, the general preference is based on an analysis of the generated MOP data, which shows that individuals in the sample prefer to work from home on Thursdays and Fridays. This preference is included in the following variable:

$$\begin{gathered}\mathrm{PWFHD} = \sum_{t}\sum_{a}\sum_{d}\sum_{h} x_{a,h,d,t} \\ \forall a \in \mathcal{A}_{wfh}, d \in \mathcal{D}_{PDwfh}, h \in \mathcal{H}_{wfh}\end{gathered} \tag{5.20}$$

Further, to account for the fact that these individuals also tend to have full telecommuting days, and do not commute into the office, the scheduling of work activities on those days is also controlled through the variable *unpreferred work day* (UPWD):

$$\begin{aligned}
\text{UPWD}_h &= \sum_t \sum_a x_{a,h,d,t} \\
&\qquad \forall a \in \mathcal{A}_w, d \in \mathcal{D}_{UPw}, h \in \mathcal{H}_{wfh} \\
\text{UPWD}_h &= 0 \\
&\qquad \forall a \in \mathcal{A}_w, d \in \mathcal{D}_{UPw}, h \in \mathcal{H} \setminus \mathcal{H}_{wfh} \\
\text{UPWD} &= \sum_h \text{UPWD}_h
\end{aligned} \tag{5.21}$$

Preference order of the objectives

The schedule frame model is formulated as a MOOP using the weighted sum method, which implies a preference order the individual places on the objectives. To generate activity schedules that are representative of real behavior, the weights are generated by calibrating them against observed data from the MOP. This is achieved through Bayesian Optimization aiming to minimize the difference between the simulated and the real schedules. The difference evaluation is based on the mean and the variance of the number of episodes generated for each of the main activities. The calibration process is implemented in Python 3 using the package scikit-optimize (Head et al. 2021). Table 5.6 shows the result of the weight estimation.

Table 5.6: Estimated weights of the schedule frame model

ω_n	description	value	preference rank
ω_1	work activity fragmentation	.23	1
ω_2	other activity fragmentation	.095	3
ω_3	common home time	.167	2
ω_4	difference in daily work duration	.23	1
ω_5	difference in allocated chores	.023	4
ω_6	out of home time	.023	4
ω_7	end of workday preference	.023	4
ω_8	preferred wfh days	.023	4

The highest importance is placed on weights 1 and 4. The result of weight 4 corroborates the findings of the literature that individuals behave consistently throughout the week regarding the daily work duration (Hilgert et al. 2017). The value for ω_1 indicates that individuals do not fragment work activities but conduct them in large blocks of time. This is an inherent result of work policies, which dictate most employees to be working during a given period of the day, during which private activities cannot be conducted outside of breaks.
The smaller value for ω_2 shows that this is less important concerning fragmentation of other activities. Indeed, these are more likely to be short and spread throughout the schedule. This result highlights the need to account for activity fragmentation differently depending on the activity type. The objective with the second highest importance is the one placed on time spent together at home. This finding is again consistent with results presented in the literature (Vuk et al. 2016). All other objectives are valued equally and lower compared to the aforementioned objectives.

5.3.2 Schedule Fine-Tuning

The solution to the optimization problem for generating activity schedules consists of a list of hourly episodes per activity purpose, including their order, the day they are performed, and the assignment to household members. In the next step, the discrete schedules serve as input for the fine-tuning of the schedule, which converts them into minute-based schedules to determine the exact duration and start time of the episodes. This step is also formulated as an optimization problem but is less complex. At this point, a constraint-based optimization problem is sufficient, where the goal is to find a solution while adhering to the constraints without optimizing an objective function. Table 5.7 lists all sets, parameters, and variables used in the schedule fine-tuning model.

Table 5.7: Sets, parameters, and variables in the schedule fine-tuning problem

$\mathcal{A}$	Activity type
$\mathcal{H}$	Household members
$\mathcal{E}$	Episodes to be scheduled
$\psi_{h,a,e}$	duration in minutes of episode e with activity purpose a assigned to household member h
$\tau_{h,a,e}$	start time of episode e with activity purpose a assigned to household member h in minutes since 00:00
APPROXDUR_e	approximate duration of episode e determined by schedule frame model
day_e	day on which episode e is scheduled
sh^o	shop opening time
sh^{cl}	shop closing time

In this step, the temporal resolution is increased, moving from aggregated time-use at the hour level, to minute-based time-use. Some constraints from the first step are transferred, and additional constraints are introduced based on the agents' preferences regarding the start time of work activities. These will be explained in more detail below.

The decision variable in this model is $\tau_{h,a,e}$, representing the duration of episode e associated with activity purpose a for household member h. This leads to the second decision variable, $\psi_{h,a,e}$, which determines the start time of the episodes. The start times of each episode are given in minutes and refer to the beginning of the week, for example, $\psi_{h,a,e} = 4993$ refers to the 4,993rd minute of the week, which corresponds to Thursday at 11:13 PM. A graphical illustration of the decision variables in the model is provided in Figure 5.7. Episodes can be extracted from the activity schedule framework in various ways. Indices are generated as a sorted sequence over all episodes, a sorted sequence over episodes of the same activity purpose, or a sorted sequence over a day and an activity purpose. This allows for various comparisons between the two episodes.

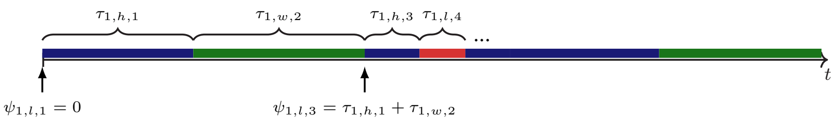


Figure 5.7: Graphical illustration of the decision variables in the schedule fine-tuning model

The determination of the start time of an episode is based on the following equality constraints. In essence, constraint (5.23) sets the starting point for the first episode in a schedule to 00:00 and the start time for all following episodes is determined by the sum of the duration of the previous episodes.

$$\psi_{h,a,e} = \sum_{1}^{e} \sum_{a} \tau_{h,a,e} \qquad \forall h \in \mathcal{H} \tag{5.22}$$

$$\psi_{h,a,e} = 0 \qquad \forall h \in \mathcal{H}, e = 0. \tag{5.23}$$

Further, the previously exact duration of time spent on each activity must be adhered to. Similar to the activity framework model, a constraint ensures that the sum of the durations of all episodes corresponds to the household's time usage as determined by the MDCNEV model.

$$\sum_h \sum_e \tau_{h,a,e} = \mathrm{tu}_a \qquad \forall a \in A. \tag{5.24}$$

As in the previous step of generating activity frameworks, constraints must also be introduced in this step to ensure that joint activities occur at the same time. This is achieved by coordinating the durations of activities for the respective household members h and i and their corresponding start times:

$$\psi_{h,a,e} = \psi_{i,a,f} \tag{5.25}$$

$$\tau_{h,a,e} = \tau_{i,a,f} \qquad \forall h, i \in \mathcal{H}, a \in \mathcal{A}_j, e, f \in \mathcal{E}_a : h \neq i, e = f \tag{5.26}$$

Furthermore, to carry over the results from the schedule frame model on the generated durations, the deviation from the determined duration $\tau_{h,a,e}$ is limited to 30 minutes:

$$|\tau_{h,a,e} - \mathrm{APPROXDUR}_e| \leq 30, e \in \mathcal{E} \tag{5.27}$$

Finally, in this model part, it must also be considered that shopping activities can only take place during store opening hours, i.e., between opening time sh_o and closing time sh_{cl}.

$$\psi_{h,a,e} \geq \mathrm{sh}^o \cdot 60 + \mathrm{day}_e \cdot 1440 \tag{5.28}$$

$$\psi_{h,a,e} + \tau_{h,a,e} \geq \mathrm{sh}^{cl} \cdot 60 + \mathrm{day}_e \cdot 1440 \qquad \forall h \in \mathcal{H}, a \in \mathcal{A}_s, e \in \mathcal{E} \tag{5.29}$$

The result of this modeling step consists of episodes with precise start times and durations for each agent in a household of a synthetic population. The activity plans generated in this way conform to the format of the actiTopp model developed by Hilgert et al. (2017), Hilgert (2019), ensuring compatibility with the agent-based demand model mobiTopp (Mallig et al. 2013) and MATSim (Briem et al. 2019).

5.4 Model Application

To assess the validity and efficacy of the proposed model to account for telecommuting behavior and difference, this section presents results from a small-scale application based on the hold-out sample of the MOP that was used neither for model estimation nor weight calibration. In the first part, the results are compared to observed data. Subsequently, the simulated telecommuting behavior is analyzed.

5.4.1 Validation

First, the results of the time-use model step are evaluated. Because the model combines the discrete and continuous choice, two values are evaluated: the simulated time use of non-zero investment and the choice to conduct the activity at all, which are labeled in the following plots as *mean time use* and *ratio of times chosen*, respectively.

Time use models

First, the time use model for single households is analyzed. The results are presented in Figure 5.8. From the plot, we can see that the discrete choice is simulated well, with minor differences compared to the real values for telecommuting, business, and leisure activities. A more discernible deviation from the real values can be seen for shopping and work activities. In both cases, the simulation underestimates the choice to engage in the activities. Regarding the time use invested in these activities, we can see that for work activities also the continuous choice is underestimated. This indicates, that the baseline utility should be higher compared to the estimated value. For shopping, on the other hand, the simulation overestimates the time use. This result is more difficult to address as the continuous and the discrete choice are interrelated and increasing, e.g., the baseline utility of shopping to increase the discrete choice would also increase the continuous choice, thus enlarging the difference. This has to be balanced out through the satiation parameter. Therefore, a thorough calibration approach should be applied before large-scale model application.

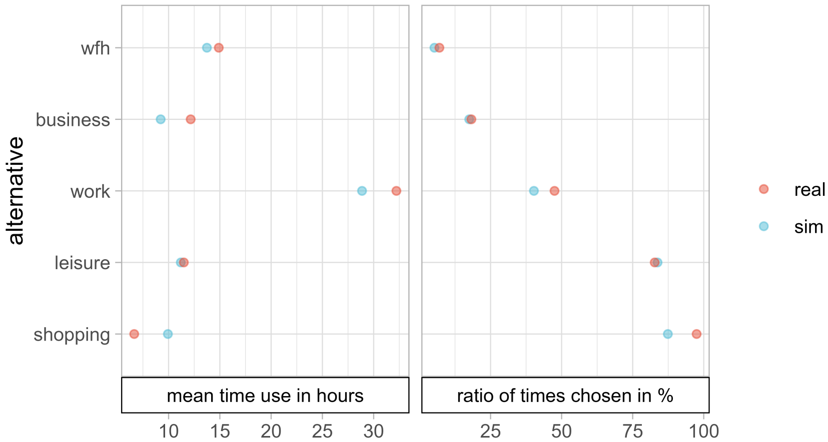


Figure 5.8: Simulated vs. real values of the single time use model

Figure 5.9 presents the result of the time use model for couple households. The ratios of chosen alternatives over all choices in the population show only small deviations across the alternatives, except for the shopping alternative. In this case, the model underestimates the number of times shopping is chosen in the population. At the same time, the model overestimates the time spent on shopping. The largest difference in simulated versus real time-use can be seen for joint leisure activities, where the model considerably underestimates the time invested. All other differences are moderate, and the simulation findings for telecommuting closely align with the actual values.

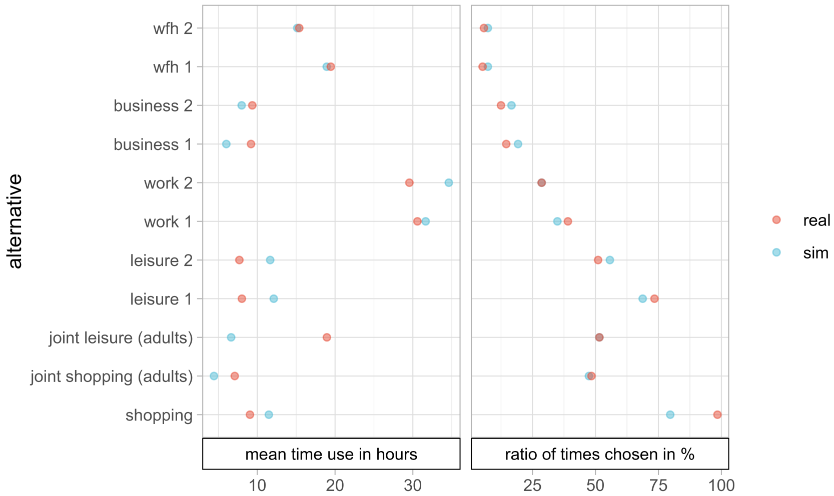


Figure 5.9: Simulated vs. real values of the couple time use model

Moving on to the results of the family time use model (Figure 5.10), we can see, that again the ratio of chosen alternatives fits the real data well. Differences are mostly discernible among household variables, for which the simulated ratio is lower, except for pick-up/drop-off activities. The analysis of the mean time-use of activities shows some larger differences between simulated and real values. Shopping activities show the largest difference, in that the model overestimates the time use. Further, work 3 and work 4 show considerably different values, in this case, the model underestimates the time use.

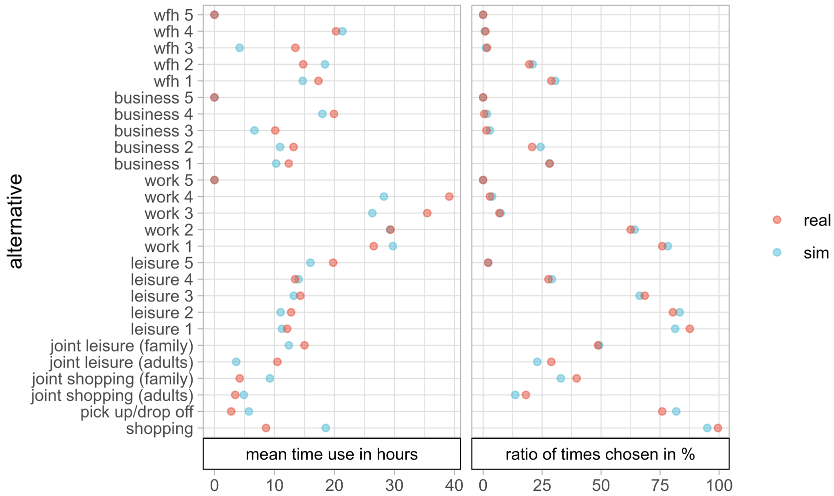


Figure 5.10: Simulated vs. real values of the family time use model

The results of the time-use model are promising. Especially the discrete choice, i.e., engaging in an activity in the week with non-zero time investment is modeled well. The overall results for the average time spent, if an activity is chosen, are also close to the observed values. However, for some activities, considerable difference lies between simulated and real values. This warrants a better calibration of the parameters. Additionally, it could be fruitful to include interaction parameters in the model, especially for multi-person households to better understand and represent intra- and inter-personal heterogeneity. Further, other multiple discrete continuous approaches could be tested, to improve the results. Especially the approach presented by Bhat (2018) could be promising. He proposes a model structure that breaks the tight link between discrete and continuous choice, which is especially helpful if a good is consumed at large ratios but low values, such as the shopping activities in this work.

Scheduling

Turning now to the validation of the scheduling approach, we compare the mean and variance of the number of daily episodes generated in the scheduling model. The results are presented in Table 5.8. The table shows the simulated, real values, and their squared differences for activities that are conducted by all household types. The result shows a very good match between the simulated and real values. Across all activities, the squared difference is lower than .1. The mean number of leisure and work activities is replicated almost perfectly.

Table 5.8: Comparison of statistical moments on the number of daily activities for various activities

Activity	Moment	Simulated	Real	Squared Difference
Home	Mean	2.031	2.207	0.0311
Home	Variance	0.920	0.682	0.0567
Work	Mean	1.096	1.076	0.0004
Work	Variance	0.087	0.076	0.0001
Business	Mean	1.303	1.072	0.0531
Business	Variance	0.316	0.067	0.0617
Wfh	Mean	1.669	1.600	0.0048
Wfh	Variance	0.573	0.557	0.0003
Leisure	Mean	1.147	1.142	0.0000
Leisure	Variance	0.189	0.141	0.0023
Shopping	Mean	1.066	1.230	0.0269
Shopping	Variance	0.082	0.223	0.0199

The overall validation results show sound results. Although the time-use models show some discrepancies, these are likely handled through parameter calibration. The scheduling results highlight the validity of the formulated problems to generate activity schedules.

5.4.2 Analysis of simulated telecommuting behavior

This section explores how sensitive and realistic the model represents telecommuting behavior. The analysis focuses on employed individuals, more specifically, those who conducted at least one work or work-from-home activity during the simulation week.
Figure 5.11 compares the start times of out-of-home activities for employed individuals by their engagement in telecommuting during the simulation week on days they worked, either from home and/or in the office. Those who do not work from home at all show the distinctive patterns of the morning and afternoon peak. It can be seen that during the morning peak hour, telecommuters travel less to conduct out-of-home activities, however, over the day starting around noon, they are more active. A small peak can be discerned for telecommuters, which is attributed to the fact that individuals do not necessarily work from home every day of the week and also not always the entire day.

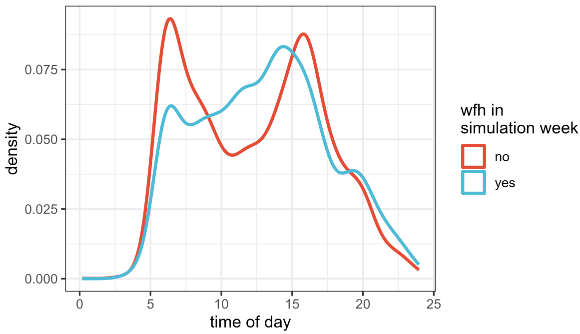


Figure 5.11: Distribution of out-of-home activity start times of employed individuals on days they worked (from home)

The model further allows for the analysis of how the degree to which somebody teleworks impacts travel demand. Figure 5.12 further differentiates those who work from home based on whether day do so for the entire day versus just for part of the day, i.e., a hybrid form of telework. Those who work from home only part of the day and part of the time in the office, show multiple peaks throughout the day. A very discernible morning peak, one in the early afternoon and one later in the evening, highlighting that the model accounts for the flexibility that working from home offers resulting in diverse activity patterns observed for telecommuters. On the other hand, those who telecommute the entire day conduct increased out-of-home activities later in the day.

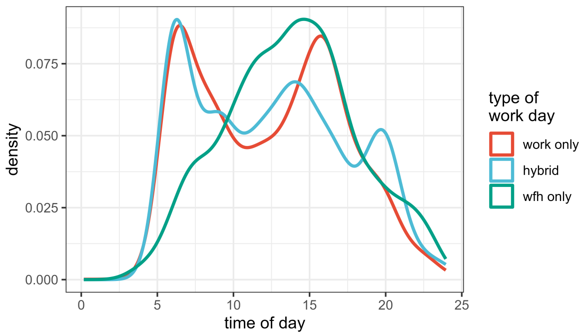


Figure 5.12: Distribution of out-of-home activity start times of employed individuals on days they worked (from home) by differentiated by full-day office work, hybrid office and work from home, and full-day work from home.

Although the analyses show that especially morning peak-hour travel is reduced through telecommuting, this is not spread equally throughout the week. Figure 5.13 shows that the model is sensitive to the preference of telecommuting days. Currently, it is calibrated such that agents prefer to telecommute on Thursdays and Fridays. However, this can be adapted to fit observed data as well as be subject to scenario analysis, e.g., to assess which workplace policies would need to be in place regarding the days of the week on which people can telework such that travel reductions become apparent. As a comparison, work is scheduled evenly throughout the simulation week, indicating that the model generates realistic results.

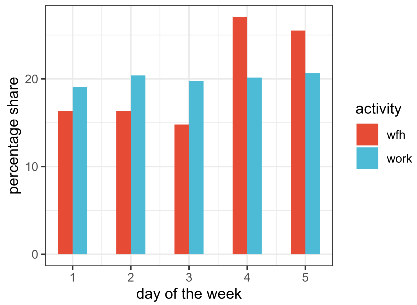


Figure 5.13: Share of work from home days and work days by day of the week

The application results further show that telecommuting is taken into account during the allocation of household activities. As Figure 5.14 shows, those who telecommute during the simulation week spend considerably more time on pick-up/drop-off activities. Recall, that while telecommuting is an individual-specific variable, time-use for escorting activities is generated at the household level and allocated to the household members in the schedule frame model. The allocation is thus the result of the different constraints and objectives of the model and the results indicate that this model generates a realistic allocation of household tasks.

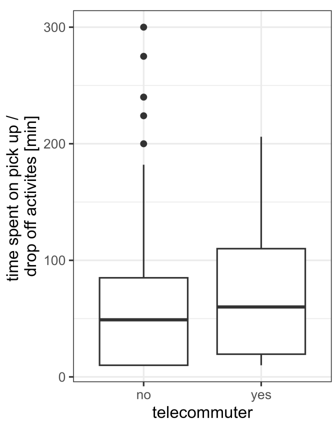


Figure 5.14: Time spent on pick up/drop off activities by telecommuting activities

5.4.3 Runtime analysis

In this section, runtimes of the simulation are presented to evaluate the scalability of the model. CPLEX allows for several settings regarding the runtime and acceptable solution. A time limit can be set at which the program terminates with a possibly suboptimal solution or no solution at all if none is found within the time limit. Pougala (2024) set the time limit of the solver to 1,800 seconds to test the scalability of the multi-day model. A similar approach was taken in this work. Two setups were tested, regarding the timelimit. To allow for comparability with the multi-day OASIS formulation, in the first setup, the time limit is also set to 1,800 seconds. The second setup puts a more restrictive time limit on all households, accepting suboptimal solutions, especially for family households. The analysis is based on 100 simulations per run. The time limit is only imposed on the schedule frame model. The runtimes are presented for all modeling steps of the METiS model. The model was run on a 3,70 GHz 8-core intel E-2288 with 128 GB of RAM.

Table 5.9: Simulation runtimes

time limit	average runtime per simulation
1,800s	44.8s (single and couple households) 1,807s (family households)
20s	5.8s (all households)

Analysis of the results given the first setup indicates the time limit is only ever reached when scheduling activities of family households. This is intuitive as the dimensions of the problem become larger with more household members to consider. For both single and couple households, the runtimes lie far below that. Regarding only the runtimes of scheduling for single and couple households, the average runtime of each model is 44.8 seconds. This is the average time taken to run the METiS model where the optimal solution for the schedule frame model of single and couple households is determined. Accepting suboptimal solutions by enforcing a tighter time limit for the model results in an average runtime of each simulation of 5.8 seconds, taking all households into account.

These results show that the model is suitable for large-scale applications given we accept suboptimal solutions, at least for family households. It should be noted, that this might even reflect a more realistic representation of reality as humans are unlikely to process all options and information (Märki et al. 2014). This warrants further investigations into which degree of optimality represents observed behavior best. Further speed-ups could be reached through increased computing power. Given that the models are run separately for each household, the activity schedules can be generated separately from each other. Although CPLEX already utilizes parallelization techniques, additional measures could be containerization of the software and running it on multiple machines could decrease runtimes further.

5.5 Discussion on the modeling framework

This thesis presents a modeling approach that generates activity schedules for the period of one week while considering household interactions. The overall framework is similar to the HAPP model (Recker 1995). However, instead of simulating schedules for one day, METiS allows for the consideration of a simulation week. Additionally, activities are generated as part of the model and the activity episodes do not have to be provided beforehand. Although this issue has been addressed by Pougala et al. (2023), the proposed model, OASIS, is not suitable for large-scale applications when considering activity schedules for a week, as runtimes become too large. Considering the week instead of a day is a vital feature necessary for the analysis of telecommuting effects. The results of the model application show that it is sensitive to individual characteristics, preferences, and household interactions concerning telecommuting behavior. Although the results are promising, there are some limitations worth noting.

First, the model only crudely considers travel or more precisely destinations for the respective activities and thus largely ignores agents' accessibility in activity generation and scheduling. Travel times to the locations where the activity is conducted are included in the duration of the activity, thus reserving an appropriate *budget* for travel in the model. This can be used in the subsequent destination choice model to generate a choice set of locations that provide a reasonable ratio of travel time to activity duration.

Second, the data on telecommuting on multiple occasions relies on imputed data. This is not an issue of the model, as parameters can be re-estimated with better-quality data. However, at the current state, any imputation errors will skew the results of the model input data and thus the model results. Therefore, it is advisable (if possible) to validate telecommuting-related variables in any application of the model against more detailed data.

Third, the MDCNEV models show room for improvement regarding their forecasting ability. It is worth noting that the open-source software available to estimate MDCEV parameters at all is very limited. Although theoretically great improvements to the initial formulations have been made, these have not yet made it into statistical modeling software. The results of the model application warrant a deeper analysis of whether more sophisticated MDC approaches would improve the model quality.

Fourth, the current implementation of the scheduler is based on the proprietary software IBM CPLEX. This conflicts with the FAIR (Findable, Accessible, Interoperable and Reusable) principles for research software. Although IBM provides free software licenses for academics, it should be tested how the model runtime increases using an open-source solver. At least a small reproducible example based solely on open-source software should be made available.

Finally, similar to many other activity scheduling models, METiS generates schedules at a specific point in time. This is a simplification of the decision process. The patterns we observe in most survey data are the result of planning and optimizing activities over time. The decision to participate in an activity during the week is not necessarily made at the same time the decision on all other activities is made, indeed it is most unlikely that such a large choice set is considered at one point in time. However, simply regarding one day is also not the solution to the problem, as many effects of activity scheduling behavior will be disregarded, as presented in this thesis. A better approach would be to build the schedule little by little over time while tightening constraints with each step in time and then generating a final schedule. The framework of the METiS schedule is well suited for such an approach, it follows a similar idea of first generating relatively crude schedules which are later finetuned. However, this requires data collection efforts that go far beyond the currently conducted travel surveys.

6 Conclusions

> I'm smart enough now to know I'm stupid. That's progress.
>
> Andy Weir,
> *Project Hail Mary*

Telecommuting is often seen as an easy way to reduce travel demand, especially during peak hours. However, to evaluate how working from home can be leveraged as a travel demand management measure, we have to move beyond the rule-of-three intuition, which leads many to believe that teleworking axiomatically reduces travel because individuals are staying at home instead of commuting to work. The complexity of teleworking as a travel demand reduction measure starts with the fact that it is not a transport policy at all, but first and foremost a workplace policy. Further, it is not a decision employees make solely based on their commute and travel conditions, but interwoven with their demand for conducting other activities as well as their needs to arrange their everyday lives, often interacting with other household members, especially when children are involved. The work in this thesis presents a modeling approach that allows for the consideration of these minute intricacies and complexities of telecommuting behavior. This last chapter provides a summary of this work, including the contributions made to the state of the art. Finally, avenues for future work are elaborated.

6.1 Thesis summary and contributions

The work in this thesis presents a modeling framework for the generation and scheduling of activities that allows for the analysis of telecommuting behavior and its effects on travel demand. This thesis investigates the complex dynamics introduced by telecommuting through an empirical analysis of data from the German Mobility Panel, highlighting substantial behavioral deviations among telecommuters compared to non-telecommuters. The research reveals that telecommuters reallocate their saved commuting time to other activities, with these decisions deeply influenced by individual household contexts. Notably, while telecommuting mothers often merge care responsibilities with work, fathers typically gain personal time. These findings emphasize the necessity of incorporating household interactions into activity generation and scheduling models.

The lack of models that can account for telecommuting behavior can at least in part be explained by the limited data availability. In this research, a machine-learning-based approach is presented showing how telecommuting data can be incorporated into survey datasets where it was previously unrecorded. This method has proven effective within this study and holds potential for broader applications, such as merging time use and travel diary data.

The core of this thesis is the development of the METiS model, which is designed to generate activity schedules over a week while explicitly considering household interactions. The model integrates a specification of the MDCEV model to estimate activity time use at the household level for one week. It effectively captures the trade-offs that occur not just at the individual activity level, but across the household collectively. These interactions are also considered in the scheduling module, which is divided into two parts. First, schedule frames at the hourly level are generated. The model is formulated as a multi-objective optimization problem applying the weighted sum method to represent the preferences of the objectives. The weights are calibrated against real data, thus providing a valid base for scenario analysis. The results of the schedule frame model are activity episodes scheduled throughout the week with approximate start times and durations. These are finetuned in the second part of the model, which is formulated such that constraints are kept consistent across the two model parts.

Application of the model has proven that it can successfully account for the intricacies of telecommuting behavior at different levels. The model results clearly show different daily patterns of telecommuters versus non-telecommuters. They also show that behavior differs depending on the degree to which an individual teleworks. And lastly, it can account for the uneven distribution of telework throughout the week. These features allow for a comprehensive analysis of the efficacy of telecommuting as a policy measure to reduce travel demand and related emissions through demand simulation based on the results of the METiS model.

Central to the METiS model is its capacity to manage multiple scheduling objectives, each weighted to reflect different priorities. The calibration of these weights against observed data using Bayesian optimization for instance reveals that currently low importance is put on the equal allocation of chores. This insight could suggest that such an objective might be deemed unnecessary in the scheduler based on current norms. However, the main contribution of the model lies in its ability to simulate different scenarios that offer a unique avenue for societal exploration. As the weights can be dynamically adapted, different scheduling preferences can be evaluated.

6.2 Suggestions for future work

The METiS model offers interesting paths for future research. Based on the comprehensive empirical analyses and model development, the work presented in this thesis can substantially guide data collection efforts to include more nuanced details about telecommuting habits, such as the intensity and regularity of telework across different sectors and job types. This could enable the refinement of the METiS model to better reflect the variability in telecommuting practices and their impacts on activity scheduling and travel demand.

Furthermore, additional in-home activities could be regarded, granted suitable data is available. This would increase analytic capabilities and allow for a more nuanced definition of some constraints in the model. Along the same lines, it would also be valuable to include online shopping behavior as, similar to telecommuting, previous research has identified the rebound effect of increased in-store shopping with a propensity to buy goods online. Integration and analysis of these effects would be a suitable use case of the model.

Additionally, there is a notable opportunity to explore the long-term effects of telecommuting on urban planning and transportation infrastructure. Future studies could use the findings from this thesis as a foundation to simulate various future scenarios where telecommuting becomes more prevalent, assessing impacts on traffic congestion, public transit usage, and urban sprawl. Integrating a land-use model with the METiS model could provide deeper insights into the long-term effects of working from home, particularly regarding residential relocation and workplace choice.

Finally, considering the differential impacts of telecommuting on gender and family roles as identified in this work, further research could focus on developing targeted strategies that address these disparities. For instance, creating models that specifically examine the allocation of time to caregiving and household responsibilities could inform policies aimed at promoting gender equity and work-life balance in the context of telecommuting. Such studies would not only add depth to the academic understanding of telecommuting dynamics but also offer practical insights for employers and policymakers aiming to optimize the benefits of telework arrangements. This could further open new avenues to address the commute gender gap, which describes the difference in commuting distance or time between male and female working individuals. Similar to the gender wage gap, research has shown that females have shorter commutes than their male counterparts, thereby limiting their access to higher-paying jobs in the labor market. As the model can account for the constraints that care work puts on women, scenarios can be designed and evaluated to unravel which policies would help reduce the commute gender gap.

A Appendix

A.1 Logit Choice Probabilities

The following covers the arithmetic steps to get from the general formulation of the Logit probabilities to the closed-form expression. As derived in section 2.2.1, the probability that decision-maker n chooses alternative i over j assuming a Gumbel distribution of the unobserved utility component is given by:

$$P_{ni} = \int_{-\infty}^{\infty} \left(\prod_{j \neq i} e^{-e^{-(\varepsilon_{ni}+v_{ni}-v_{nj})}} \right) e^{-\varepsilon_{ni}} e^{-e^{-\varepsilon_{ni}}} d\varepsilon_{ni}$$

For simplicity, replace ε_{ni} with s:

$$\begin{aligned}
P_{ni} &= \int_{-\infty}^{\infty} \left(\prod_{j \neq i} e^{-e^{-(s+V_{ni}-V_{nj})}} \right) e^{-s} e^{-e^{-s}} ds \\
&= \int_{-\infty}^{\infty} \left(\prod_{j} e^{-e^{-(s+V_{ni}-v_{nj})}} \right) \cdot \frac{1}{e^{-(s+V_{ni}-V_{nj})}} \cdot e^{-s} e^{-e^{-s}} ds \\
&= \int_{-\infty}^{\infty} \left(\prod_{j} e^{-e^{-(s+V_{ni}-V_{nj})}} \right) \cdot e^{-s} ds \\
&= \int_{-\infty}^{\infty} e^{-\sum_j e^{-(s+V_{ni}-V_{nj})}} \cdot e^{-s} ds \\
&= \int_{-\infty}^{\infty} e^{-e^{-s} \sum_j e^{-(V_{ni}-V_{nj})}} \cdot e^{-s} ds
\end{aligned}$$

Define $t = e^{-s}$ such that $dt = -e^{-s}ds$, then

$$P_{ni} = \int_{-\infty}^{\infty} \left(e^{-t\sum_j e^{-(V_{ni}-V_{nj})}} \right) (-dt)$$

Note that $\lim_{s\to\infty} t = \lim_{s\to\infty} e^{-s} = 0$ and $\lim_{s\to\infty} t = \lim_{s\to-\infty} e^{-s} = \infty$ Given these integral limits, we get:

$$\begin{aligned} P_{ni} &= \int_{\infty}^{0} \left(e^{-t\sum_j e^{-(v_{ni}-v_{nj})}} \right) (-dt) \\ &= \int_{0}^{\infty} e^{-t\sum_j e^{-(v_{ni}-v_{nj})}} dt \\ &= e^{-t\sum_j e^{-(v_{ni}-v_{nj})}} \cdot \frac{1}{-\sum_j e^{-(v_{ni}-v_{nj})}} \Bigg|_0^{\infty} \end{aligned}$$

With $\lim_{t\to\infty} e^{-t\sum_j e^{-(v_{ni}-v_{nj})}} = 0$ follows:

$$\begin{aligned} P_{ni} &= 0 - e^{0\left(\sum_j e^{-(v_{ni}-v_{nj})}\right)} \cdot \frac{1}{-\sum_j e^{-(v_{ni}-v_{nj})}} \\ &= \frac{1}{\sum_j e^{-(V_{ni}-V_{nj})}} = \frac{e^{V_{ni}}}{\sum_j e^{V_{nj}}} \end{aligned}$$

A.2 Numerical Data Supporting Figures in Chapter 4

Table A.1: Summary statistics for weekly time use of employed respondents in MOP by activity, and telecommuting status

activity	wfh	N	mean	median	SD
home	cannot wfh	3556	117.28	115.78	18.68
home	choose not to wfh	342	111.27	109.89	15.67
home	wfh	2131	125.25	125.67	21.01
work	cannot wfh	3168	32.15	33.33	14.07
work	choose not to wfh	324	35.45	37.28	12.00
work	wfh	1489	24.55	23.42	13.95
work-related	cannot wfh	877	12.45	8.50	12.12
work-related	choose not to wfh	112	12.80	10.22	9.82
work-related	wfh	742	11.62	8.57	10.23
shopping	cannot wfh	3319	5.14	4.03	4.38
shopping	choose not to wfh	323	4.97	3.85	4.66
shopping	wfh	2008	4.73	3.69	4.04
leisure	cannot wfh	2616	9.17	6.50	8.92
leisure	choose not to wfh	271	10.26	7.83	9.13
leisure	wfh	1663	9.73	7.08	9.38
leisure walk	cannot wfh	944	13.84	9.81	13.18
leisure walk	choose not to wfh	94	10.11	8.00	7.79
leisure walk	wfh	697	15.22	10.00	16.10
joint leisure: adults	cannot wfh	1178	7.74	5.75	7.78
joint leisure: adults	choose not to wfh	103	8.16	6.75	7.01
joint leisure: adults	wfh	770	8.92	5.92	10.41
joint leisure: family	cannot wfh	113	5.37	3.75	4.54
joint leisure: family	choose not to wfh	9	4.65	2.42	4.46
joint leisure: family	wfh	80	5.75	3.71	4.70
joint leisure: one parent with kids	cannot wfh	147	4.03	2.97	3.72
joint leisure: one parent with kids	choose not to wfh	11	2.73	2.42	1.97
joint leisure: one parent with kids	wfh	106	4.27	3.00	3.84
pick up/drop off	cannot wfh	1153	2.01	1.17	2.35
pick up/drop off	choose not to wfh	121	1.65	1.00	1.91
pick up/drop off	wfh	801	2.09	1.28	2.27

Table A.2: Summary statistics for weekly time use of employed respondents in MOP by activity, household type, and telecommuting status

		cannot wfh				choose not to wfh				wfh			
activity	household type	N	mean	median	SD	N	mean	median	SD	N	mean	median	SD
home	couple hh	1340	118.10	116.17	19.05	118	111.76	110.10	13.85	748	124.01	124.07	22.51
home	family hh	968	118.19	116.50	17.48	99	113.01	112.63	16.27	756	125.75	125.47	19.34
home	multi adult hh	571	115.42	114.33	18.85	36	109.02	108.11	16.71	300	126.26	126.90	19.66
home	single hh	677	115.93	115.07	19.26	89	109.60	107.38	16.78	327	125.64	126.97	21.98
work	couple hh	1162	32.12	33.68	14.14	113	36.14	38.75	12.00	493	25.29	25.00	14.08
work	family hh	872	31.33	31.57	13.83	94	33.76	32.85	11.89	564	23.43	22.17	12.68
work	multi adult hh	518	32.65	34.42	14.35	34	37.07	39.25	11.23	210	25.25	24.98	15.17
work	single hh	616	32.92	34.25	14.00	83	35.78	38.72	12.44	222	25.07	22.79	15.38
work-related	couple hh	315	12.20	7.92	12.16	43	11.68	7.75	10.86	274	11.06	7.92	9.75
work-related	family hh	241	12.23	8.00	12.52	31	13.46	12.50	7.80	253	12.92	10.33	10.66
work-related	multi adult hh	145	14.06	9.25	12.94	9	11.68	8.17	9.85	103	12.26	8.00	12.25
work-related	single hh	176	11.84	8.90	10.70	29	14.09	10.33	10.38	112	9.44	7.83	7.67
shopping	couple hh	1263	5.38	4.33	4.47	111	4.70	3.83	3.91	709	4.74	3.60	3.86
shopping	family hh	901	4.72	3.75	3.81	94	5.52	3.66	6.27	712	4.65	3.71	4.28
shopping	multi adult hh	513	5.09	4.00	4.19	31	6.07	4.87	4.54	270	4.58	3.61	3.65
shopping	single hh	642	5.29	3.92	4.99	87	4.32	3.75	3.25	317	5.05	4.08	4.16
leisure	couple hh	935	8.55	6.00	8.61	89	7.81	6.22	6.33	550	10.10	7.06	10.06
leisure	family hh	732	7.74	5.46	7.73	74	8.55	6.33	8.06	605	8.11	6.27	7.33
leisure	multi adult hh	421	9.89	7.25	9.52	30	10.97	8.37	9.33	232	9.92	7.00	9.76
leisure	single hh	528	11.68	9.00	9.92	78	14.42	11.55	11.15	276	12.40	9.62	10.89
leisure walk	couple hh	361	13.12	9.00	12.64	32	8.96	6.83	5.78	233	17.42	11.43	17.95
leisure walk	family hh	233	13.06	9.52	12.19	19	11.64	9.00	9.04	240	13.29	9.00	13.32
leisure walk	multi adult hh	155	15.21	10.33	14.23	11	15.07	11.50	10.98	100	13.22	8.71	13.39
leisure walk	single hh	195	15.01	11.00	14.29	32	8.65	7.17	6.98	124	16.42	10.94	18.66
joint leisure: adults	couple hh	624	8.80	6.62	8.28	56	9.01	8.00	6.98	382	10.88	6.92	12.29
joint leisure: adults	family hh	336	5.97	4.61	5.85	37	7.58	6.00	7.65	279	6.27	4.42	6.40
joint leisure: adults	multi adult hh	218	7.43	5.79	8.42	10	5.55	6.25	3.32	109	8.86	6.00	10.09
joint leisure: family	family hh	113	5.37	3.75	4.54	9	4.65	2.42	4.46	80	5.75	3.71	4.70
joint leisure: one parent with kids	family hh	147	4.03	2.97	3.72	11	2.73	2.42	1.97	106	4.27	3.00	3.84
pick up/drop off	couple hh	289	1.79	1.03	2.32	27	1.26	0.85	1.27	140	1.66	0.83	2.15
pick up/drop off	family hh	585	2.30	1.45	2.50	62	2.03	1.26	2.29	507	2.39	1.75	2.35
pick up/drop off	multi adult hh	134	1.48	0.71	2.13	14	0.78	0.58	0.53	92	1.19	0.65	1.28
pick up/drop off	single hh	145	1.74	0.98	1.77	18	1.60	1.00	1.72	62	2.04	0.90	2.57

Table A.3: Summary statistics for weekly time use of employed respondents in MOP by activity, household role, and telecommuting status

		cannot wfh				choose not to wfh				wfh			
activity	household role	N	mean	median	SD	N	mean	median	SD	N	mean	median	SD
home	father	437	113.06	111.50	16.93	69	109.02	109.80	15.06	407	123.98	122.30	20.21
home	mother	531	122.41	121.38	16.80	30	122.17	120.51	15.44	349	127.81	128.43	18.10
work	father	396	36.26	38.28	13.98	66	36.00	37.26	11.76	300	25.55	25.50	13.43
work	mother	476	27.23	26.83	12.30	28	28.48	27.53	10.63	264	21.02	20.08	11.31
work-related	father	128	15.42	9.96	13.99	27	14.10	12.50	8.10	151	14.54	12.70	10.45
work-related	mother	113	8.62	5.08	9.45	4	9.08	8.33	3.12	102	10.52	6.72	10.56
shopping	father	393	4.19	3.20	3.48	65	5.85	3.35	7.26	377	4.45	3.42	4.73
shopping	mother	508	5.13	4.08	4.00	29	4.80	4.82	3.06	335	4.87	4.13	3.72
leisure	father	312	8.27	5.52	8.31	48	9.20	7.25	8.20	318	8.45	6.47	7.84
leisure	mother	420	7.34	5.33	7.25	26	7.35	5.74	7.79	287	7.72	6.17	6.72
leisure walk	father	97	12.89	9.72	11.12	16	11.86	9.30	9.66	131	12.91	9.50	10.95
leisure walk	mother	136	13.19	9.51	12.93	3	10.50	7.42	5.78	109	13.74	8.17	15.73
joint leisure: one parent with kids	father	52	3.87	2.88	3.47	8	2.82	2.46	2.26	38	4.23	2.75	4.13
joint leisure: one parent with kids	mother	95	4.11	2.97	3.87	3	2.51	2.17	1.19	68	4.29	3.22	3.70
pick up/drop off	father	237	1.86	1.10	1.96	43	1.86	1.25	1.94	252	2.04	1.42	2.05
pick up/drop off	mother	348	2.61	1.73	2.77	19	2.41	1.45	2.95	255	2.73	2.12	2.56

Table A.4: Summary statistics for weekly number of episodes of employed respondents in MOP by activity, and telecommuting status

activity	wfh	N	mean	median	SD
work	cannot wfh	3168	4.57	5.00	1.66
work	choose not to wfh	324	4.87	5.00	1.57
work	wfh	1489	3.59	4.00	1.89
work-related	cannot wfh	877	4.18	2.00	5.09
work-related	choose not to wfh	112	4.00	3.00	3.42
work-related	wfh	742	3.64	2.00	4.20
shopping	cannot wfh	3319	4.81	4.00	3.30
shopping	choose not to wfh	323	5.17	4.00	3.52
shopping	wfh	2008	4.86	4.00	3.23
leisure	cannot wfh	2616	3.17	2.00	2.54
leisure	choose not to wfh	271	3.66	3.00	2.64
leisure	wfh	1663	3.47	3.00	2.73
leisure walk	cannot wfh	944	2.79	1.00	3.29
leisure walk	choose not to wfh	94	1.94	1.00	1.19
leisure walk	wfh	697	2.90	2.00	3.45
joint leisure: adults	cannot wfh	1178	2.26	2.00	1.92
joint leisure: adults	choose not to wfh	103	2.31	2.00	1.67
joint leisure: adults	wfh	770	2.61	2.00	2.27
joint leisure: family	cannot wfh	113	1.50	1.00	1.13
joint leisure: family	choose not to wfh	9	1.11	1.00	0.33
joint leisure: family	wfh	80	1.69	1.00	1.13
joint leisure: one parent with kids	cannot wfh	147	1.66	1.00	1.12
joint leisure: one parent with kids	choose not to wfh	11	1.55	1.00	0.82
joint leisure: one parent with kids	wfh	106	1.48	1.00	0.83
pick up/drop off	cannot wfh	1153	3.08	2.00	3.03
pick up/drop off	choose not to wfh	121	2.50	2.00	2.15
pick up/drop off	wfh	801	3.54	2.00	3.26

Table A.5: Summary statistics for weekly number of episodes of employed respondents in MOP by activity, household type, and telecommuting status

		cannot wfh				choose not to wfh				wfh			
activity	household type	N	mean	median	SD	N	mean	median	SD	N	mean	median	SD
work	couple hh	1162	4.58	5.00	1.66	113	4.90	5.00	1.53	493	3.72	4.00	2.01
work	family hh	872	4.57	5.00	1.68	94	4.74	5.00	1.60	564	3.48	4.00	1.74
work	multi adult hh	518	4.58	5.00	1.62	34	4.88	5.00	1.37	210	3.57	4.00	1.81
work	single hh	616	4.54	5.00	1.67	83	4.98	5.00	1.67	222	3.59	4.00	2.03
work-related	couple hh	315	4.34	2.00	5.50	43	3.95	2.00	3.73	274	3.57	2.00	4.33
work-related	family hh	241	4.44	2.00	5.98	31	3.39	3.00	2.30	253	3.53	2.00	3.51
work-related	multi adult hh	145	4.34	3.00	4.44	9	5.22	2.00	5.19	103	4.27	2.00	5.60
work-related	single hh	176	3.43	2.00	3.09	29	4.34	4.00	3.32	112	3.46	2.00	3.79
shopping	couple hh	1263	4.88	4.00	3.29	111	5.09	4.00	3.64	709	4.69	4.00	2.99
shopping	family hh	901	4.86	4.00	3.63	94	5.37	4.50	3.79	712	4.87	4.00	3.19
shopping	multi adult hh	513	4.40	4.00	2.97	31	5.42	5.00	3.31	270	4.60	4.00	3.27
shopping	single hh	642	4.93	4.00	3.04	87	4.95	4.00	3.18	317	5.44	5.00	3.69
leisure	couple hh	935	2.92	2.00	2.35	89	2.94	2.00	2.16	550	3.29	2.00	2.91
leisure	family hh	732	2.99	2.00	2.30	74	3.51	3.00	2.69	605	3.38	3.00	2.35
leisure	multi adult hh	421	3.33	3.00	2.64	30	4.13	3.00	2.85	232	3.50	3.00	2.88
leisure	single hh	528	3.74	3.00	2.98	78	4.45	4.00	2.81	276	4.01	3.00	2.93
leisure walk	couple hh	361	2.76	1.00	3.33	32	1.88	1.00	1.16	233	3.37	2.00	3.82
leisure walk	family hh	233	2.69	2.00	2.83	19	2.05	2.00	1.18	240	2.62	1.00	3.17
leisure walk	multi adult hh	155	2.81	2.00	2.95	11	2.64	2.00	1.69	100	2.42	2.00	2.09
leisure walk	single hh	195	2.97	2.00	3.94	32	1.69	1.00	0.97	124	2.96	2.00	4.02
joint leisure: adults	couple hh	624	2.46	2.00	2.07	56	2.41	2.00	1.67	382	3.00	2.00	2.55
joint leisure: adults	family hh	336	1.98	1.00	1.69	37	2.24	1.00	1.85	279	2.15	1.00	1.77
joint leisure: adults	multi adult hh	218	2.13	2.00	1.76	10	2.00	2.00	0.94	109	2.39	2.00	2.11
joint leisure: family	family hh	113	1.50	1.00	1.13	9	1.11	1.00	0.33	80	1.69	1.00	1.13
joint leisure: one parent with kids	family hh	147	1.66	1.00	1.12	11	1.55	1.00	0.82	106	1.48	1.00	0.83
pick up/drop off	couple hh	289	1.94	1.00	1.58	27	1.78	1.00	1.40	140	2.00	1.00	2.02
pick up/drop off	family hh	585	4.23	3.00	3.65	62	3.27	2.00	2.57	507	4.42	4.00	3.58
pick up/drop off	multi adult hh	134	1.87	1.00	1.49	14	1.57	1.00	0.76	92	1.89	1.00	1.51
pick up/drop off	single hh	145	1.83	1.00	1.30	18	1.61	1.00	0.92	62	2.35	2.00	1.63

Table A.6: Summary statistics for weekly number of episodes of employed respondents in MOP by activity, household role, and telecommuting status

		cannot wfh				choose not to wfh				wfh			
activity	household type	N	mean	median	SD	N	mean	median	SD	N	mean	median	SD
work	father	396	4.99	5.00	1.70	66	4.95	5.00	1.72	300	3.56	4.00	1.83
work	mother	476	4.22	4.00	1.58	28	4.25	4.00	1.14	264	3.40	3.00	1.63
work-related	father	128	5.57	3.00	7.11	27	3.52	4.00	2.42	151	3.89	3.00	3.94
work-related	mother	113	3.17	2.00	4.02	4	2.50	2.00	1.00	102	3.00	2.00	2.70
shopping	father	393	4.56	4.00	3.48	65	5.40	5.00	3.98	377	4.53	4.00	3.01
shopping	mother	508	5.10	4.00	3.72	29	5.31	4.00	3.40	335	5.24	5.00	3.34
leisure	father	312	3.00	2.00	2.42	48	3.67	3.00	2.96	318	3.32	3.00	2.41
leisure	mother	420	2.98	2.00	2.21	26	3.23	3.00	2.10	287	3.45	3.00	2.29
leisure walk	father	97	2.57	2.00	2.42	16	2.12	2.00	1.20	131	2.35	1.00	2.14
leisure walk	mother	136	2.77	2.00	3.09	3	1.67	1.00	1.15	109	2.94	1.00	4.06
joint leisure: one parent with kids	father	52	1.46	1.00	0.85	8	1.38	1.00	0.74	38	1.26	1.00	0.69
joint leisure: one parent with kids	mother	95	1.77	1.00	1.23	3	2.00	2.00	1.00	68	1.60	1.00	0.88
pick up/drop off	father	237	3.57	2.00	3.03	43	2.86	2.00	2.31	252	3.58	3.00	2.62
pick up/drop off	mother	348	4.68	3.00	3.96	19	4.21	4.00	2.95	255	5.24	4.00	4.16

Table A.7: Summary statistics for episode durations of employed respondents in MOP by activity, and telecommuting status

activity	wfh	N	mean	median	SD
work	cannot wfh	14480	7.03	7.42	2.92
work	choose not to wfh	1579	7.28	7.92	2.87
work	wfh	5346	6.84	7.03	3.01
work-related	cannot wfh	3669	2.97	2.00	2.82
work-related	choose not to wfh	448	3.20	2.50	2.65
work-related	wfh	2699	3.19	2.25	2.87
shopping	cannot wfh	15966	1.07	0.72	1.19
shopping	choose not to wfh	1669	0.96	0.62	1.15
shopping	wfh	9758	0.97	0.63	1.15
leisure	cannot wfh	8292	2.89	2.17	2.61
leisure	choose not to wfh	993	2.80	2.05	2.52
leisure	wfh	5773	2.80	2.00	2.71
leisure walk	cannot wfh	2638	4.95	4.33	3.55
leisure walk	choose not to wfh	182	5.22	4.76	3.43
leisure walk	wfh	2023	5.24	4.98	3.64
joint leisure: adults	cannot wfh	2665	3.42	2.50	2.89
joint leisure: adults	choose not to wfh	238	3.53	2.50	3.03
joint leisure: adults	wfh	2006	3.42	2.39	3.14
joint leisure: family	cannot wfh	169	3.59	2.67	2.84
joint leisure: family	choose not to wfh	10	4.18	2.71	3.28
joint leisure: family	wfh	135	3.41	2.33	2.87
joint leisure: one parent with kids	cannot wfh	244	2.43	1.83	2.09
joint leisure: one parent with kids	choose not to wfh	17	1.77	1.25	1.53
joint leisure: one parent with kids	wfh	157	2.88	2.00	2.80
pick up/drop off	cannot wfh	3555	0.65	0.33	0.96
pick up/drop off	choose not to wfh	302	0.66	0.42	0.81
pick up/drop off	wfh	2839	0.59	0.33	0.82

Table A.8: Summary statistics for episode durations of employed respondents in MOP by activity, household type, and telecommuting status

		cannot wfh				choose not to wfh				wfh			
activity	household type	N	mean	median	SD	N	mean	median	SD	N	mean	median	SD
work	couple hh	5324	7.01	7.50	2.92	554	7.37	8.00	2.90	1833	6.80	7.00	3.05
work	family hh	3984	6.86	6.97	2.89	446	7.12	7.50	2.85	1965	6.73	6.75	2.87
work	multi adult hh	2373	7.13	7.50	2.94	166	7.59	8.25	2.66	750	7.07	7.92	3.16
work	single hh	2799	7.24	8.03	2.92	413	7.19	7.98	2.93	798	6.97	7.47	3.09
work-related	couple hh	1366	2.81	1.75	2.68	170	2.96	2.40	2.38	978	3.10	2.25	2.77
work-related	family hh	1071	2.75	1.77	2.81	105	3.97	3.25	3.04	894	3.66	2.73	3.07
work-related	multi adult hh	629	3.24	2.42	2.93	47	2.24	1.33	1.98	440	2.87	1.83	2.81
work-related	single hh	603	3.46	2.62	2.95	126	3.24	2.48	2.74	387	2.73	1.75	2.55
shopping	couple hh	6158	1.10	0.75	1.19	565	0.92	0.67	1.08	3327	1.01	0.67	1.19
shopping	family hh	4383	0.97	0.67	1.00	505	1.03	0.58	1.36	3464	0.96	0.60	1.17
shopping	multi adult hh	2259	1.15	0.75	1.35	168	1.12	0.75	1.14	1241	1.00	0.67	1.08
shopping	single hh	3166	1.07	0.70	1.28	431	0.87	0.58	0.93	1726	0.93	0.58	1.09
leisure	couple hh	2727	2.93	2.22	2.63	262	2.65	2.00	2.23	1809	3.07	2.17	2.90
leisure	family hh	2189	2.59	2.00	2.49	260	2.43	1.75	2.50	2044	2.40	1.75	2.36
leisure	multi adult hh	1400	2.97	2.33	2.50	124	2.65	2.13	1.96	812	2.83	2.08	2.74
leisure	single hh	1976	3.12	2.33	2.76	347	3.24	2.33	2.84	1108	3.09	2.25	2.89
leisure walk	couple hh	998	4.74	4.17	3.43	60	4.78	3.50	3.43	786	5.16	4.82	3.62
leisure walk	family hh	626	4.86	4.00	3.59	39	5.67	6.00	3.28	628	5.08	4.65	3.55
leisure walk	multi adult hh	435	5.42	5.00	3.40	29	5.71	4.72	3.66	242	5.46	5.50	3.95
leisure walk	single hh	579	5.06	4.25	3.77	54	5.12	4.71	3.43	367	5.55	5.25	3.63
pick up/drop off	couple hh	562	0.92	0.50	1.27	48	0.71	0.33	0.99	280	0.83	0.50	1.13
pick up/drop off	family hh	2477	0.54	0.33	0.81	203	0.62	0.42	0.73	2239	0.54	0.33	0.72
pick up/drop off	multi adult hh	251	0.79	0.42	1.15	22	0.50	0.46	0.29	174	0.63	0.37	0.84
pick up/drop off	single hh	265	0.95	0.52	1.11	29	0.99	0.75	1.17	146	0.87	0.38	1.30

Table A.9: Summary statistics for episode durations of employed respondents in MOP by activity, household role, and telecommuting status

		cannot wfh				choose not to wfh				wfh			
activity	household type	N	mean	median	SD	N	mean	median	SD	N	mean	median	SD
work	father	1977	7.26	8.08	3.15	327	7.27	8.08	3.00	1068	7.18	7.83	2.93
work	mother	2007	6.46	6.50	2.55	119	6.70	6.50	2.32	897	6.19	6.17	2.70
work-related	father	713	2.77	1.82	2.90	95	4.01	3.25	3.08	588	3.74	2.68	3.15
work-related	mother	358	2.72	1.72	2.64	10	3.63	3.50	2.67	306	3.51	2.79	2.91
shopping	father	1793	0.92	0.58	1.00	351	1.08	0.58	1.53	1707	0.98	0.58	1.29
shopping	mother	2590	1.01	0.75	1.00	154	0.90	0.65	0.84	1757	0.93	0.63	1.02
leisure	father	936	2.76	2.08	2.57	176	2.51	1.67	2.65	1055	2.55	1.83	2.46
leisure	mother	1253	2.46	1.83	2.43	84	2.28	1.82	2.15	989	2.24	1.67	2.24
leisure walk	father	249	5.02	4.00	3.88	34	5.58	5.76	3.18	308	5.49	5.43	3.39
leisure walk	mother	377	4.76	4.08	3.40	5	6.30	6.92	4.26	320	4.68	3.83	3.65
joint leisure: one parent with kids	father	76	2.65	2.29	1.94	11	2.05	1.25	1.82	48	3.35	2.75	2.74
joint leisure: one parent with kids	mother	168	2.33	1.62	2.15	6	1.26	1.10	0.63	109	2.68	1.83	2.81
pick up/drop off	father	847	0.52	0.33	0.63	123	0.65	0.42	0.81	902	0.57	0.33	0.81
pick up/drop off	mother	1630	0.56	0.33	0.89	80	0.57	0.39	0.58	1337	0.52	0.33	0.66

A.3 Difference between real and simulated values of time use models

Table A.10: Difference between real and simulated values for the single household time use model

activity	simulated mean	real mean	difference of means	simulated ratio	real ratio	difference of ratios
shopping	9.93	6.65	3.29	87.38	97.50	-10.12
home	71.55	71.25	0.30	100.00	99.50	0.50
leisure	11.19	11.49	-0.30	83.75	82.75	1.00
work	28.87	32.22	-3.35	40.25	47.50	-7.25
wfh	13.74	14.89	-1.16	5.25	7.00	-1.75
business	9.23	12.16	-2.93	17.50	18.25	-0.75

Table A.11: Difference between real and simulated values for the couples time use model

activity	simulated mean	real mean	difference of means	simulated ratio	real ratio	difference of ratios
shopping	11.49	9.06	2.43	79.69	98.44	18.75
joint shopping (adults)	4.44	7.11	2.67	47.40	48.44	1.04
joint leisure (adults)	6.65	18.95	12.30	51.56	51.56	0.00
home	116.89	129.15	12.26	100.00	93.75	6.25
leisure 1	12.12	8.03	4.09	68.75	73.44	4.69
leisure 2	11.66	7.70	3.96	55.73	51.04	4.69
work 1	31.64	30.59	1.06	34.90	39.06	4.17
wfh 1	18.93	19.45	0.52	7.29	5.21	2.08
business 1	6.03	9.20	3.17	19.27	14.58	4.69
work 2	34.62	29.55	5.07	28.65	28.65	0.00
wfh 2	15.14	15.40	0.25	7.29	5.73	1.56
business 2	8.00	9.36	1.36	16.67	12.50	4.17

Table A.12: Difference between real and simulated values for the family household time use model

activity	simulated mean	real mean	difference of means	simulated ratio	real ratio	difference of ratios
shopping	18.56	8.60	9.96	94.99	99.52	4.53
pick up/drop off	5.73	2.79	2.94	81.86	75.89	5.97
joint shopping (adults)	4.89	3.46	1.43	13.60	18.14	4.53
joint shopping (family)	9.24	4.21	5.03	32.94	39.62	6.68
joint leisure (adults)	3.62	10.49	6.86	22.91	28.88	5.97
joint leisure (family)	12.40	15.01	2.62	49.16	48.69	0.48
home	198.90	200.01	1.11	100.00	100.00	0.00
leisure 1	11.25	12.12	0.87	81.38	87.59	6.21
leisure 2	11.03	12.78	1.74	83.29	80.43	2.86
leisure 3	13.22	14.33	1.11	66.35	68.50	2.15
leisure 4	14.02	13.44	0.59	29.12	27.68	1.43
leisure 5	16.00	19.79	3.79	2.15	2.15	0.00
work 1	29.75	26.55	3.20	78.28	75.89	2.39
work 2	29.29	29.34	0.05	64.20	62.53	1.67
work 3	26.32	35.48	9.15	7.40	6.92	0.48
work 4	28.25	39.16	10.91	3.82	2.86	0.95
work 5	0.00	0.00	0.00	0.00	0.00	0.00
business 1	10.29	12.38	2.10	28.16	28.16	0.00
business 2	10.94	13.19	2.24	24.34	20.76	3.58
business 3	6.67	10.13	3.47	2.86	1.43	1.43
business 4	18.00	19.93	1.93	1.67	0.48	1.19
business 5	0.00	0.00	0.00	0.00	0.00	0.00
wfh 1	14.72	17.34	2.62	30.55	28.88	1.67
wfh 2	18.42	14.80	3.62	21.00	19.57	1.43
wfh 3	4.20	13.49	9.29	1.19	1.67	0.48
wfh 4	21.33	20.27	1.07	0.72	0.95	0.24
wfh 5	0.00	0.00	0.00	0.00	0.00	0.00

List of Figures

List of Tables

Bibliography

IBM ILOG CPLEX Optimization Studio. IBM, New York, 2022.

Serio Agriesti, Claudio Roncoli, and Bat-hen Nahmias-Biran. Assignment of a Synthetic Population for Activity-Based Modeling Employing Publicly Available Data. *ISPRS International Journal of Geo-Information*, 11(2):148, 2022.

Cevat Giray Aksoy, Jose Maria Barrero, Nicholas Bloom, Steven J. Davis, Mathias Dolls, and Pablo Zarate. Working from home around the world. *Brookings Papers on Economic Activity*, 2022(2):281–360, 2022.

Mahdieh Allahviranloo and Kay W. Axhausen. An optimization model to measure utility of joint and solo activities. *Transportation Research Part B: Methodological*, 108:172–187, 2018.

Renni Anggraini, Theo A. Arentze, and Harry J. P. Timmermans. Car allocation decisions in car-deficient households: The case of non-work tours. *Transportmetrica*, 8(3):209–224, 2012.

Theo A. Arentze and Harry J. P. Timmermans. A learning-based transportation oriented simulation system. *Transportation Research Part B: Methodological*, 38(7):613–633, 2004.

Theo A. Arentze, Frank Hofman, Henk Van Mourik, and Harry J. P. Timmermans. ALBATROSS: Multiagent, Rule-Based Model of Activity Pattern Decisions. *Transportation Research Record: Journal of the Transportation Research Board*, 1706(1):136–144, 2000.

Hamidreza Asgari and Xia Jin. Impacts of Telecommuting on Nonmandatory Activity Participation: Role of Endogeneity. *Transportation Research Record: Journal of the Transportation Research Board*, 2666(1):47–57, 2017.

Hamidreza Asgari, Xia Jin, and Ali Mohseni. Choice, Frequency, and Engagement: Framework for Telecommuting Behavior Analysis and Modeling. *Transportation Research Record: Journal of the Transportation Research Board*, 2413(1):101–109, 2014.

Hamidreza Asgari, Xia Jin, and Yiman Du. Examination of the Impacts of Telecommuting on the Time Use of Nonmandatory Activities. *Transportation Research Record: Journal of the Transportation Research Board*, 2566(1): 83–92, 2016.

Katherine E. Asmussen, Aupal Mondal, and Chandra R. Bhat. The interplay between teleworking choice and commute distance. Technical paper, Department of Civil, Architectural and Environmental Engineering, The University of Texas at Austin, 2023.

Joshua Auld and Abolfazl Mohammadian. Framework for the development of the Agent-based Dynamic Activity Planning and Travel Scheduling (ADAPTS) model. *Transportation Letters*, 1(3):245–255, 2009.

Joshua Auld and Abolfazl Mohammadian. Activity planning processes in the Agent-based Dynamic Activity Planning and Travel Scheduling (ADAPTS) model. *Transportation Research Part A: Policy and Practice*, 46(8):1386–1403, 2012.

Pierre Baldi, Søren Brunak, Yves Chauvin, Claus A. F. Andersen, and Henrik Nielsen. Assessing the accuracy of prediction algorithms for classification: An overview. *Bioinformatics*, 16(5):412–424, 2000.

Davide Barbieri, Carla Bernardo, Eva L. Eldermans, Anita Mezza, Blandine Mollard, Mia Nahrgang, Vytautas Peciukonis, Jolanta Reingarde, Irene Riobóo Lestón, Lina Salanauskaite, and Annabel Wildschut. Gender Equality Index 2023: Towards a green transition in transport and energy. Publications Office of the European Union, 2023.

Matthew J. Beck, David A. Hensher, and Edward Wei. Slowly coming out of COVID-19 restrictions in Australia: Implications for working from home and commuting trips by car and public transport. *Journal of Transport Geography*, 88:102846, 2020.

Tom Bellemans, Bruno Kochan, Davy Janssens, Geert Wets, Theo A. Arentze, and Harry J. P. Timmermans. Implementation Framework and Development Trajectory of FEATHERS Activity-Based Simulation Platform. *Transportation Research Record: Journal of the Transportation Research Board*, 2175(1):111–119, 2010.

Christina Bernardo, Rajesh Paleti, Megan Hoklas, and Chandra R. Bhat. An empirical investigation into the time-use and activity patterns of dual-earner couples with and without young children. *Transportation Research Part A: Policy and Practice*, 76:71–91, 2015.

Chandra R. Bhat. A multiple discrete–continuous extreme value model: Formulation and application to discretionary time-use decisions. *Transportation Research Part B: Methodological*, 39(8):679–707, 2005.

Chandra R. Bhat. The multiple discrete-continuous extreme value (MDCEV) model: Role of utility function parameters, identification considerations, and model extensions. *Transportation Research Part B: Methodological*, 42(3), 2008.

Chandra R. Bhat. A new flexible multiple discrete–continuous extreme value (MDCEV) choice model. *Transportation Research Part B: Methodological*, 110:261–279, 2018.

Chandra R. Bhat, Jessica Y. Guo, Sivaramakrishnan Srinivasan, and Aruna Sivakumar. Comprehensive Econometric Microsimulator for Daily Activity-Travel Patterns. *Transportation Research Record: Journal of the Transportation Research Board*, 1894(1):57–66, 2004.

Chandra R. Bhat, Aupal Mondal, Katherine E. Asmussen, and Aarti C. Bhat. A multiple discrete extreme value choice model with grouped consumption data and unobserved budgets. *Transportation Research Part B: Methodological*, 141:196–222, 2020.

Henry. D. Block. *Random Orderings and Stochastic Theories of Responses*, pages 172–217. Springer Netherlands, Dordrecht, 1974.

John L. Bowman and Moshe E. Ben-Akiva. Activity-based disaggregate travel demand model system with activity schedules. *Transportation Research Part A: Policy and Practice*, 35(1):1–28, 2001.

Mark Bradley and Peter Vovsha. A model for joint choice of daily activity pattern types of household members. *Transportation*, 32(5):545–571, 2005.

Lars Briem, Nicolai Mallig, and Peter Vortisch. Creating an integrated agent-based travel demand model by combining mobiTopp and MATSim. *Procedia Computer Science*, 151:776–781, 2019.

Eric Brochu, Vlad M. Cora, and Nando de Freitas. A Tutorial on Bayesian Optimization of Expensive Cost Functions, with Application to Active User Modeling and Hierarchical Reinforcement Learning. 2010.

Hannah Budnitz, Emmanouil Tranos, and Lee Chapman. Telecommuting and other trips: An English case study. *Journal of Transport Geography*, 85:102713, 2020.

David S. Bunch, David M. Gay, and Roy E. Welsch. Algorithm 717: Subroutines for maximum likelihood and quasi-likelihood estimation of parameters in non-linear regression models. *ACM Transactions on Mathematical Software*, 19(1): 109–130, 1993.

Bundesministerium für Digitales und Verkehr. Sofortprogramm für den Sektor Verkehr aufgrund einer Überschreitung der zulässigen Jahresemissionsmenge für das Jahr 2021 auf Grundlage von § 8 Absatz 1 KSG, 2022.

Chiara Calastri, Stephane Hess, Andrew Daly, and Juan Antonio Carrasco. Does the social context help with understanding and predicting the choice of activity type and duration? An application of the Multiple Discrete-Continuous Nested

Extreme Value model to activity diary data. *Transportation Research Part A: Policy and Practice*, 104:1–20, 2017.

Bernardo Caldarola and Steve Sorrell. Do teleworkers travel less? Evidence from the English National Travel Survey. *Transportation Research Part A: Policy and Practice*, 159:282–303, 2022.

Sandip Chakrabarti. Does telecommuting promote sustainable travel and physical activity? *Journal of Transport & Health*, 9:19–33, 2018.

Nitesh V. Chawla, Kevin W. Bowyer, Lawrence O. Hall, and William P. Kegelmeyer. SMOTE: Synthetic Minority Over-sampling Technique. *Journal of Artificial Intelligence Research*, 16:321–357, 2002.

Rung-Ching Chen, Christine Dewi, Su-Wen Huang, and Rezzy Eko Caraka. Selecting critical features for data classification based on machine learning methods. *Journal of Big Data*, 7(1):52, 2020.

Jie Cheng and Russell Greiner. Comparing Bayesian Network Classifiers. 2013.

Davide Chicco and Giuseppe Jurman. The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. *BMC Genomics*, 21(1):6, 2020.

Corinna Cortes and Vladimir Vapnik. Support-vector networks. *Machine Learning*, 20(3):273–297, 1995.

G. R. Dattatreya and Laveen N. Kanal. Decision Trees in Pattern Recognition. University of Maryland, 1984.

Mathijs de Haas, Roel Faber, and Marije Hamersma. How COVID-19 and the Dutch ‘intelligent lockdown’ change activities, work and travel behaviour: Evidence from longitudinal data in the Netherlands. *Transportation Research Interdisciplinary Perspectives*, 6:100150, 2020.

Tommy Gärling, Kurt Brännäs, Jürgen Garvill, Reginald G. Golledge, Sucharita Gopal, S Holm, and Erik Lindberg. Household Activity Scheduling. In *Selected*

Proceedings of the Fifth World Conference on Transort Research, volume 4, pages 235–248, Ventura, California, 1989.

Robin Genuer, Jean-Michel Poggi, and Christine Tuleau-Malot. VSURF: An R package for variable selection using random forests. *The R Journal*, 7(2):19–33, 2015.

José Ignacio Giménez-Nadal, José Alberto Molina, and Jorge Velilla. Work time and well-being for workers at home: Evidence from the American Time Use Survey. *International Journal of Manpower*, 41(2):184–206, 2019.

Reginald G. Golledge, Mei-Po Kwan, and Tommy Gärling. Computational Process Modeling of Household Travel Decisions using a Geographical Information System. *Papers in Regional Science*, 73(2):99–117, 1994.

Isabelle Guyon, Jason Weston, Stephen Barnhill, and Vladimir Vapnik. Gene Selection for Cancer Classification using Support Vector Machines. *Machine Learning*, 46(1/3):389–422, 2002.

Khandker Nurul Habib. On the Factors Influencing the Choices of Weekly Telecommuting Frequencies of Post-secondary Students in Toronto. 2020.

Murtaza Haider and Amar Iqbal Anwar. The prevalence of telework under Covid-19 in Canada. *Information Technology & People*, 36(1):196–223, 2023.

Thomas Hale, Noam Angrist, Rafael Goldszmidt, Beatriz Kira, Anna Petherick, Toby Phillips, Samuel Webster, Emily Cameron-Blake, Laura Hallas, Saptarshi Majumdar, and Helen Tatlow. A global panel database of pandemic policies (Oxford COVID-19 Government Response Tracker). *Nature Human Behaviour*, 5(4):529–538, 2021.

Sylvia Y. He and Lingqian Hu. Telecommuting, income, and out-of-home activities. *Travel Behaviour and Society*, 2(3):131–147, 2015.

Tim Head, Manoj Kumar, Holger Nahrstaedt, Gilles Louppe, and Iaroslav Shcherbatyi. Scikit-optimize/scikit-optimize, 2021.

Daniel Heimgartner and Kay W. Axhausen. Hybrid work arrangement choices and its implications for home office frequencies. *103rd Annual Meeting of the Transportation Research Board*, 2024.

Igal Hendel. Estimating Multiple-Discrete Choice Models: An Application to Computerization Returns. *Review of Economic Studies*, 66(2):423–446, 1999.

Stephane Hess and David Palma. Apollo: A flexible, powerful and customisable freeware package for choice model estimation and application. *Journal of Choice Modelling*, 32:100170, 2019.

Stephane Hess and David Pamla. Apollo version 0.2.4, user manual. Choice Modelling Center, University of Leeds, Leeds, 2021.

Tim Hilgert. *Erstellung von Wochenaktivitätenplänen für Verkehrsnachfragemodelle*. PhD thesis, KIT Scientific Publishing, Karlsruhe, 2019.

Tim Hilgert, Michael Heilig, Martin Kagerbauer, and Peter Vortisch. Modeling Week Activity Schedules for Travel Demand Models. *Transportation Research Record: Journal of the Transportation Research Board*, 2666(1):69–77, 2017.

Erin L. Kelly, Phyllis Moen, and Eric Tranby. Changing Workplaces to Reduce Work-Family Conflict: Schedule Control in a White-Collar Organization. *American Sociological Review*, 76(2):265–290, 2011.

Shivam Khaddar, Mahmudur Rahman Fatmi, and Trevor Nikodym. Are Telecommuters' and Non-Telecommuters' Daily Time-Use Behaviors Different? An Episode-Level Model for Non-Mandatory Activities. *Transportation Research Record: Journal of the Transportation Research Board*, 2677(11):96–110, 2023.

Jaehwan Kim, Allenby, and Peter E. Rossi. Modeling Consumer Demand for Variety. *Marketing Science*, 21(3):229–250, 2002.

Ryuichi Kitamura and Satoshi Fujii. Two computational process models of activity-travel behavior. *Theoretical foundations of travel choice modeling*, pages 251–279, 1998.

Ron Kohavi and George H. John. Wrappers for feature subset selection. *Artificial Intelligence*, 97(1-2):273–324, 1997.

Max Kuhn. Building Predictive Models in *R* Using the **caret** Package. *Journal of Statistical Software*, 28(5), 2008.

Tobias Kuhnimhof, Bastian Chlond, and Dirk Zumkeller. Nonresponse, selectivity, and data quality in travel surveys: Experiences from analyzing recruitment for the german mobility panel. *Transportation research record*, 1972(1):29–37, 2006.

Anna Kurowska. Gendered Effects of Home-Based Work on Parents' Capability to Balance Work with Non-work: Two Countries with Different Models of Division of Labour Compared. *Social Indicators Research*, 151(2):405–425, 2020.

Miron B. Kursa and Witold R. Rudnicki. Feature selection with the Boruta package. *Journal of Statistical Software*, 36:1–13, 2010.

Miron B. Kursa, Aleksander Jankowski, and Witold R. Rudnicki. Boruta – A System for Feature Selection. *Fundamenta Informaticae*, 101(4):271–285, 2010.

Yago Losa Rovira, Ahmadreza Faghih Imani, Aruna Sivakumar, and Jacek Pawlak. Do in-home and virtual activities impact out-of-home activity participation? Investigating end-user activity behaviour and time use for residential energy applications. *Energy and Buildings*, 257:111764, 2022.

Nicolai Mallig, Martin Kagerbauer, and Peter Vortisch. mobiTopp – A Modular Agent-based Travel Demand Modelling Framework. *Procedia Computer Science*, 19:854–859, 2013.

Fabian Märki. *An Agent-Based Model for Continuous Activity Planning of Multi-Week Scenarios*. PhD thesis, ETH Zürich, Zürich, 2014.

Fabian Märki, David Charypar, and Kay W. Axhausen. Agent-based model for continuous activity planning with an open planning horizon. *Transportation*, 41(4):905–922, 2014.

R. Timothy Marler and Jasbir S. Arora. The weighted sum method for multi-objective optimization: New insights. *Structural and Multidisciplinary Optimization*, 41(6):853–862, 2010.

Takao Maruyama and Susanne Tietze. From anxiety to assurance: Concerns and outcomes of telework. *Personnel Review*, 41(4):450–469, 2012.

Eric Miller. The current state of activity-based travel demand modelling and some possible next steps. *Transport Reviews*, 43(4):565–570, 2023.

Rolf Moeckel, Wei-Chieh Huang, Joanna Ji, Carlos Llorca, Ana Tsui Moreno, Corin Staves, Qin Zhang, and Gregory D. Erhardt. The Activity-based model ABIT: Modeling 24 hours, 7 days a week. *Transportation Research Procedia*, 78:499–506, 2024.

Richard E Neapolitan. *Probabilistic reasoning in expert systems: theory and algorithms*. John Wiley & Sons, Inc., 1990.

Jack Nilles, Frederic Carlson, Paul Gray, and Gerhard Hanneman. Telecommuting - An Alternative to Urban Transportation Congestion. *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-6(2):77–84, 1976.

Shira Offer and Barbara Schneider. Revisiting the Gender Gap in Time-Use Patterns: Multitasking and Well-Being among Mothers and Fathers in Dual-Earner Families. *American Sociological Review*, 76(6):809–833, 2011.

Sergio A. Ordóñez Medina. *Simulating Work-Leisure Cycles in Large Scale Scenarios: Models and Implementation*. PhD thesis, ETH Zürich, Zürich, 2016.

Rajesh Paleti and Ivana Vukovic. Telecommuting and its impact on activity–time use patterns of dual-earner households. *Transportation Research Record*, 2658 (1):17–25, 2017.

David Palma, Annesha Enam, Stephane Hess, Chiara Calastri, and Romain Crastes dit Sourd. Modelling multiple occurrences of activities during a day: An extension of the MDCEV model. *Transportmetrica B: Transport Dynamics*, 9 (1):456–478, 2021.

Judea Pearl. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. The Morgan Kaufmann Series in Representation and Reasoning. Morgan Kaufmann, San Francisco, Calif, rev. 2nd ed. edition, 1997.

Ram M. Pendyala, Ryuichi Kitamura, Cynthia Chen, and Eric I. Pas. An activity-based microsimulation analysis of transportation control measures. *Transport Policy*, 4(3):183–192, 1997.

Ram M. Pendyala, Ryuichi Kitamura, and D. V. G. Prasuna Reddy. Application of an activity-based travel-demand model incorporating a rule-based algorithm. *Environment and Planning B: Planning and Design*, 25(5):753–772, 1998.

Abdul Rawoof Pinjari and Chandra R. Bhat. A multiple discrete–continuous nested extreme value (MDCNEV) model: Formulation and application to non-worker activity time-use and timing behavior on weekdays. *Transportation Research Part B: Methodological*, 44(4):562–583, 2010a.

Abdul Rawoof Pinjari and Chandra R. Bhat. An efficient forecasting procedure for Kuhn-Tucker consumer demand model systems. *Technical paper. Department of Civil & Environmental Engineering, University of South Florida*, 2010b.

Janody Pougala. In *OASIS: An Integrated Optimisation Framework for Activity Scheduling*. PhD thesis, EPFL, Lausanne, 2024.

Janody Pougala, Tim Hillel, and Michel Bierlaire. OASIS: Optimisation-based Activity Scheduling with Integrated Simultaneous choice dimensions. *Transportation Research Part C: Emerging Technologies*, 155:104291, 2023.

R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna, Austria, 2022.

Wilfred W. Recker. The household activity pattern problem: General formulation and solution. *Transportation Research Part B: Methodological*, 29(1):61–77, 1995.

Wilfred W. Recker, J. Duan, and H. Wang. Development of an Estimation Procedure for an Activity-Based Travel Demand Model. *Computer-Aided Civil and Infrastructure Engineering*, 23(7):483–501, 2008.

Negar Rezvany, Michel Bierlaire, and Tim Hillel. Simulating intra-household interactions for in- and out-of-home activity scheduling. *Transportation Research Part C: Emerging Technologies*, 157:104362, 2023.

Hyok-Joo Rhee. Home-based telecommuting and commuting behavior. *Journal of Urban Economics*, 63(1):198–216, 2008.

Shobhit Saxena, Abdul Rawoof Pinjari, Ananya Roy, and Rajesh Paleti. Multiple discrete-continuous choice models with bounds on consumptions. *Transportation Research Part A: Policy and Practice*, 149:237–265, 2021.

Shobhit Saxena, Abdul Rawoof Pinjari, and Rajesh Paleti. A multiple discrete-continuous extreme value model with ordered preferences (MDCEV-OP): Modelling framework for episode-level activity participation and time-use analysis. *Transportation Research Part B: Methodological*, 166:259–283, 2022.

Basil Schmid, Felix Becker, Kay W. Axhausen, Paul Widmer, and Petra Stein. A simultaneous model of residential location, mobility tool ownership and mode choice using latent variables. *Transportation research part A: policy and practice*, 178:103867, 2023.

Bobak Shahriari, Kevin Swersky, Ziyu Wang, Ryan P. Adams, and Nando De Freitas. Taking the Human Out of the Loop: A Review of Bayesian Optimization. *Proceedings of the IEEE*, 104(1):148–175, 2016.

Palvinder Singh, Rajesh Paleti, Syndney Jenkins, and Chandra R. Bhat. On modeling telecommuting behavior: Option, choice, and frequency. *Transportation*, 40:373–396, 2013.

Statistisches Bundesamt. Erwerbstätige, die von zu hause aus arbeiten, 2024. URL `https://www.destatis.de/DE/Themen/Arbeit/Arbeitsmarkt/Qualitaet-Arbeit/Dimension-3/home-office.html`. [Accessed 10-09-2024].

Rongxiang Su, Elizabeth C. McBride, and Konstadinos G. Goulias. Unveiling daily activity pattern differences between telecommuters and commuters using human mobility motifs and sequence analysis. *Transportation Research Part A: Policy and Practice*, 147, 2021.

Harry J. P. Timmermans and Junyi Zhang. Modeling household activity travel behavior: Examples of state of the art modeling approaches and research agenda. *Transportation Research Part B: Methodological*, 43(2):187–190, 2009.

Tin Kam Ho. Random decision forests. In *Proceedings of 3rd International Conference on Document Analysis and Recognition*, volume 1, pages 278–282, Montreal, Que., Canada, 1995. IEEE Comput. Soc. Press.

Tin Kam Ho. The random subspace method for constructing decision forests. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(8):832–844, 1998.

Kenneth E. Train. *Discrete Choice Methods with Simulation*. Cambridge University Press, 2nd edition, 2001.

Goran Vuk, John L. Bowman, Andrew Daly, and Stephane Hess. Impact of family in-home quality time on person travel demand. *Transportation*, 43(4):705–724, 2016.

Vanessa R. Wight and Sara B. Raley. When Home Becomes Work: Work and Family Time among Workers at Home. *Social Indicators Research*, 93(1): 197–202, 2009.

Zhiheng Xu, Jee Eun Kang, and Roger Chen. A Random Utility Based Estimation Framework for the Household Activity Pattern Problem. *Transportation Research Procedia*, 23:809–826, 2017.

Zhiheng Xu, Jee Eun Kang, and Roger Chen. A random utility based estimation framework for the household activity pattern problem. *Transportation Research Part A: Policy and Practice*, 114:321–337, 2018.

Chongsheng Zhang, Changchang Liu, Xiangliang Zhang, and George Almpanidis. An up-to-date comparison of state-of-the-art classification algorithms. *Expert Systems with Applications*, 82:128–150, 2017.

Shihang Zhang, Rolf Moeckel, Ana Tsui Moreno, Bin Shuai, and Jie Gao. A work-life conflict perspective on telework. *Transportation Research Part A: Policy and Practice*, 141:51–68, 2020.