Household activity scheduling: choice-set generation and parameter estimation

Negar Rezvany *　　Tim Hillel †　　Michel Bierlaire *

July 2025

Report TRANSP-OR 250712
Transport and Mobility Laboratory
School of Architecture, Civil and Environmental Engineering
Ecole Polytechnique Fédérale de Lausanne
`transp-or.epfl.ch`

*École Polytechnique Fédérale de Lausanne (EPFL), School of Architecture, Civil and Environmental Engineering (ENAC), Transport and Mobility Laboratory, Switzerland, {negar.rezvany,michel.bierlaire}@epfl.ch

†University College London (UCL), Department of Civil, Environmental and Geomatic Engineering (CEGE), United Kingdom, {tim.hillel@ucl.ac.uk}

Abstract

Activity-based travel models that represent both household and individual decision-making involve intricate behavioural interactions and rely on complex simulation tools. As a result, calibrating their parameters presents significant methodological challenges. In this paper, we propose a novel maximum likelihood estimation procedure that explicitly accounts for choice sets defined at the household level. We apply the method to data from the UK National Travel Survey. Beyond illustrating and validating the approach, our results demonstrate that models calibrated at the household level produce daily schedule distributions that more closely reflect observed data than those based solely on individual-level modelling.

Keywords: Activity-based modelling, Intra-household interactions, Choice-set generation, ABM calibration, Discrete choice models, Daily scheduling.

1 Introduction

The scheduling process is a fundamental element of activity-based demand modelling. The predominant approach for scheduling within Activity-based models (ABMs) used in both research and practice is to consider isolated individual agents, whose choices are independent of other decision-makers. This approach neglects the influence of household interactions in the scheduling process, which may lead to biased simulation of activity-travel schedules, leading to inappropriate policy actions and investments as the schedule of household members are mutually dependent. Capturing interpersonal dependencies between individuals belonging to the same household enhances consistency of predicted choices and behaviour.

Individuals do not plan their day in isolation from other members of the household. Real-world decision-making involves considering the activities and schedules of other household members and sometimes other individuals in their social network. Various interactions, time arrangements, and constraints affect the activity schedules of individuals. Intra-household interactions significantly shape daily activity schedules through mechanisms such as: (i) synchronising schedules to enable joint activities, for example family movie nights, (ii) coordinating travel, for instance, escorting children or sharing a ride, and (iii) sharing responsibilities and allocating resources, such as household vehicle usage. However, most ABMs overlook the household decision-making perspectives. Hence, models based solely on individual choices must be revisited and enhanced to take into account intra-household interactions.

There are two major research streams within the scope of ABMs: (i) rule-based or computational process models (e.g., Arentze and Timmermans 2004), and (ii) econometric models (e.g., Nurul Habib 2018). Econometric models assume that individuals choose their activity schedules to maximise their utility. Within these models, activity scheduling and travel behaviour are explained and predicted through discrete choices, modelled either sequentially or jointly, and estimated using econometric methods such as advanced discrete choice models (Bowman and M. E. Ben-Akiva 2001) or micro-simulation approaches (Bhat 2005). Constraints are a critical component in explaining activity-travel behaviour. To produce meaningful behavioural outputs, it is important that such models incorporate the various constraints individuals face—whether temporal, spatial, institutional, or resource-based—which influence not only the set of available actions but also the dynamics and feasibility of activity execution (Rezvany, Michel Bierlaire, and Hillel 2023; Pougala, Hillel, and Michel Bierlaire 2022; Auld and Mohammadian 2011; Arentze, Ettema, and Timmermans 2010). Using discrete choice models implies the need for calibration of maximum likelihood estimators of the parameters of the utility functions. Estimating parameters remains a key challenge in ABMs.

Consistent estimation of parameters requires complete enumeration of the alternatives within the choice set. This necessitates behavioural data for hypothetical or unobserved scenarios in addition to the chosen alternative (revealed preference). Such data are generally unobservable and not available from conventional data sources like travel diary surveys or time-use datasets. Moreover, deriving choice probabilities and likelihood functions requires the modeller to assume a universal choice set that is finite and enumerable. In reality, the full choice set, encompassing all possible activities and their spatio-temporal sequences, is combinatorially large and practically impossible to enumerate, while individuals are indeed only aware of a limited subset. Consequently, it is common practice to estimate parameters using a sampled choice set of alternatives (Guevara and Moshe E. Ben-Akiva 2013). Thus, defining a representative choice set for household activity scheduling is essential for operationalising household random utility models. Identifying and operationalising appropriate techniques for generating consistent and representative sampled choice sets remains a significant challenge.

For ABM parameter estimation, using discrete-choice modelling is not straightforward because the assumption of independent and identically distributed (i.i.d.) errors often does not hold, as the schedule alternatives within a choice set may share overlaps in certain components despite their general distinction. Furthermore, the presence of constraints further complicates the problem, making closed-form probability derivations unattainable. Simultaneous simulation of different scheduling choice dimensions, including: (i) activity participation, (ii) activity location, (iii) activity schedule, (iv) activity duration, (v) activity participation mode (solo/joint), and (vi) transport mode; enhances the behavioural realism of the model and enables capturing trade-offs and interactions. However, this simultaneity increases the complexity of parameter estimation, as properly accounting for correlations between choice dimensions and alternatives within a tractable modelling framework is challenging. Sequential models simplify the parameter estimation process by decomposing it into multiple stages, considering a predefined or arbitrarily constructed choice set, but at the expense of flexibility and behavioural realism (Bowman and M. E. Ben-Akiva 2001).

In order to simulate the daily activity schedules of individuals within a household, the numerous and complex interactions among household members should be explicitly accounted for. In earlier research, we propose an econometric ABM framework for household activity scheduling, contributing to the state-of-the-art in activity-based modelling by explicitly modelling multiple interaction dimensions within the same framework (Rezvany, Michel Bierlaire, and Hillel 2023). In this approach, the scheduling problem is formulated as a mixed-integer utility optimisation problem, where the intra-household interactions are captured through constraints and the objective function. The objective is to maximise the household utility, subject to a set of schedule continuity constraints and household-level constraints, such as (i) allocation of the resources to household members, (ii) sharing household maintenance responsibilities, (iii) joint participation of household members in activities, and (iv) escorting. Another merit of this scheduling model is its simultaneous simulation of different daily scheduling choice dimensions such as activity participation, location choice, scheduling, duration, participation mode (solo/joint), and transport mode, thus effectively capturing the trade-offs between different choice dimensions. This scheduling framework simulates the activity schedules of individuals from a group decision-making point-of-view rather than treating individuals as isolated.

We identify two research questions necessary to operationalise household ABMs: (i) How can the choice set be formulated such that it represents the household scheduling problem? (ii) How to formulate a tractable model specification to estimate the parameters of the household ABMs?

In this paper, we address these research questions by focusing on generation of a consistent choice set and estimating meaningful and significant parameters within household-level activity-based scheduling. In this context, the choice set comprises multiple alternatives, each representing an ensemble of daily activity schedules for all household members. Employing a choice set generation technique based on a Metropolis Hastings (MH) sampling algorithm can be a smart move to efficiently explore the solution space and strategically sample alternatives suitable for econometric calibration of the activity-based model. The generated sample should include plausible alternatives that are competitive with the observed choice and have high enough utility to be realistically considered by the decision-maker. Due to the complexities introduced by intra-household interactions, such as additional choice dimensions, timing constraints, and joint decision-making mechanisms, these aspects must be explicitly incorporated into both the choice set generation process and the subsequent parameter estimation to maintain the consistency. Such integration ensures the consistency and realism of household schedules.

We propose a choice set generation framework for household activity scheduling, producing ensembles of schedules with consistent alternatives for all household members. This paper builds on two recent contributions to the field of activity-based modelling: (i) an ABM designed to capture joint scheduling behaviour of households (Rezvany, Michel Bierlaire, and Hillel 2023), and (ii) a maximum likelihood estimation approach tailored for individual-level ABMs (Pougala, Hillel, and Michel Bierlaire 2023). We bridge these advances by formulating an estimation framework that extends the latter to accommodate the complex, high-dimensional choice sets that arise from household interactions and scheduling constraints. As such, the proposed framework is broadly applicable to ABMs involving structured and constrained decision spaces, beyond the household setting.

Key considerations for household choice set generation are carefully noted and taken into account. Dedicated operators for household-level problem are implemented, including operators that modify whether an activity is performed jointly with other household members or alone. Additionally, the utility function and constraints are explicitly formulated to accommodate the intra-household interactions. Using the generated choice set, we present an estimation procedure to calibrate parameters of the utility-based household scheduling model. The framework is applied to a case study based on the UK National Travel Survey (NTS) dataset. The results are then discussed and analysed, demonstrating how the household-level model enhances behavioural realism and achieves closer alignment with observed empirical patterns.

It should be noted that interpersonal interactions beyond the household are referred to as social interactions and are out of the scope of this study. Furthermore, we explicitly focus on short-term interactions in our framework. Therefore, we assume long-term household decisions such as household car ownership, partnerships, professional occupation status, and home and work locations are exogenous and given.

The remainder of this manuscript is structured as follows. We give a review of the literature in Section 2. In Section 3 the methodology for estimation of model parameters and the generation of choice sets at the household level is explained. Section 4 introduces the case study and provides evidence from observed household data. The choice set generation and estimation methodology is applied to a real-world case study in Section 5. Section 6 presents and analyses the schedule simulation results. Finally we conclude with a discussion on household-level versus individual-level choice set generation models (Section 7.1), followed by outlining directions for future research (Section 7.2).

2 Relevant literature

ABMs consider the demand for travel to be driven by participation in spatially and temporally distributed activities (Bowman and M. E. Ben-Akiva 2001; Chapin 1974; Hagerstrand 1970). By including why trips are derived, ABMs aim to replicate real-world decisions with more behavioural realism compared to the traditional trip-based models focusing on individual trips. This approach has been of interest to modellers and analysts in different domains such as transportation and energy research (Hou, Pawlak, and Sivakumar 2025; Rezvany, Hillel, and Michel Bierlaire 2021; Nurul Habib 2018; Subbiah 2013; Bhat 2005). Individuals do not plan their day in isolation from other household members. The household group decision-making in activity-travel behaviour have mostly been explored either at the top-level of activity generation (Arentze and Timmermans 2009; Bradley and Vovsha 2005), time allocation (Zhang and Fujiwara 2006), or sequential household-level activity pattern generation Bhat et al. 2013. Most models lack simultaneous representation of activity generation, timing, sequence, and schedules, though realistically these should be jointly modelled to generate household activity-travel schedules.

Timmermans and Zhang 2009 review household activity-travel models and emphasise that ABMs should incorporate intra-household interactions, group decision-making processes, contextual factors, and move beyond purely sequential to simultaneous representations of interdependent choice dimensions. The family of Household activity pattern problem (HAPP) models introduced by W. Recker 1995 formulates the household activity scheduling problem as a simultaneous decision process using a Mixed integer linear programming (MILP) framework, drawing inspiration from vehicle routing models; however, their prescriptive nature poses challenges for statistical estimation and empirical calibration from observed behaviour. In Rezvany, Michel Bierlaire, and Hillel 2023, we propose an operational utility-based scheduling framework that explicitly captures multiple intra-household interactions within a single ABM using a simultaneous approach. The model explicitly accommodates complex interactions among household members such as the allocation of private vehicle to household members, escort duties, joint participation in activities, and sharing rides.

One challenge in the utility-based ABMs is parameter estimation and model calibration. This is specifically complex for models which consider all choice dimensions simultaneously, rather than sequential models. Research on household-level activity-based modelling with a specific focus on parameter estimation remains relatively scarce. Parameter estimation can be broadly approached in two ways: using fixed arbitrary parameter values (e.g., Charypar and Nagel 2005) or employing empirical estimation procedures based on data. Estimating parameter based on data remains intricate, as the traditional surveys such as travel diaries are limited to only revealed preferences. The choice set of alternatives is typically latent or unobservable to the analyst. Defining a choice set representative of activity-travel patterns in household activity scheduling problem is necessary for operationalising household-level random utility models.

There are various examples of choice set generation in route-choice modelling literature (Prato and Bekhor 2007; Bekhor, Moshe E. Ben-Akiva, and Ramming 2006; Richardson 1982). The choice set formation models can be broadly categorised into two groups (Pagliara and Timmermans 2009): (i) deterministic, and (ii) stochastic. In deterministic method, a subset of choices is pre-specified for each individual based on the analyst's discretion obtained from decision rules reflecting the domain knowledge. This is often accomplished by either restricting the choice set to include only the alternatives within a pre-specified distance or travel time from each individual's trip origin (Scott and He 2012; Termansen, McClean, and Skov-Petersen 2004), or including all destinations that are chosen by other surveyed individuals residing in proximity to the given individual (Miller and O"Kelly 1983). There exist a large variety of deterministic methods for generating alternative routes in networks such as the k-shortest path (Robert B.Dial 1971), link elimination (Pagliara and Timmermans 2009), link penalty (Barra, Perez, and Anez 1993), and labelling (Moshe Ben-Akiva et al. 1984). Deterministic choice set generation methods are straightforward to implement but rely heavily on analysts' assumptions, potentially leading to biased choice sets. Stochastic approaches such as simulation methods (Flötteröd and Michel Bierlaire 2013; Frejinger, M. Bierlaire, and M. Ben-Akiva 2009; Bovy 2009; Bovy and Fiorenzo-Catalano 2007) introduce randomness into the generation process to better reflect travelers' heterogeneity and uncertainty in route perception (Nielsen 2000).

Choice set generation techniques from route choice modelling have been adopted and applied in ABMs (e.g., Danalet 2015). Additionally, the use of MH algorithms for sampling alternatives in activity-based contexts has been explored in the literature (Pougala, Hillel, and Michel Bierlaire 2023). Typically, the individual-level scheduling process is defined as a discrete choice problem, where parameters are estimated using Maximum Likelihood Estimation. Recent literature has investigated parameter estimation procedures for household-level ABMs. We conduct a comprehensive search across the three largest online publication databases - namely Google Scholar, Web of Science and Scopus - with keywords "household choice set", "household scheduling calibration", and "household ABMs estimation", to identify studies that address estimation procedures especially for household ABMs. Key papers studying calibration methods for household ABMs are selected and discussed in the remainder of this section.

Roorda, Miller, and Kruchten 2006 proposes a Genetic Algorithm (GA)-based procedure for estimating the parameters of a household tour-based mode choice model developed for the Greater Toronto Area, known as Travel Activity Scheduler for Household Agents (TASHA). The Maximum log-likelihood estimation is performed using Monte Carlo simulation, while the search for the optimal parameter set that maximises the log-likelihood function employs a GA procedure. Although TASHA represnts a successful first attempt at operationalising a model based on a group decision-making paradigm, it remains a sequential rule-based microsimulation model.

Recent advances have progress toward estimation of HAPP models (W. Recker 1995). W. Recker, Duan, and Wang 2008 proposes a GA-based procedure to tune model parameters such that the Levenshtein distance of strings from observed pattern and strings generated by HAPP is minimised. An inverse optimisation technique is proposed by Chow and Will W. Recker 2012 to estimate HAPP. Regue, Allahviranloo, and Will Recker 2015 use goal programming to calibrate activity-scheduling priority parameters for different household clusters in HAPP. Although these calibration approaches improve overall model performance, the reliance on comparing simulated and observed patterns limits their capacity to provide insights into non-chosen activity patterns. Xu, Kang, and Chen 2018 develop a choice set generation technique for HAPP grounded in Random Utility Theory, employing the clustering approach developed by Allahviranloo, Regue, and Will Recker 2014. They identify representative patterns from observed activity-travel patterns. A GA is then utilised to sample a pattern from each non-chosen representative cluster, optimising information gain by minimising the D-error of the final sample. Subsequently, goal programming adjusts these sampled alternatives to respect individuals' spatial and temporal constraints, ensuring the feasibility of the generated choice set. This procedure offers behavioural realism grounded in utility theory, however, it introduces bias and endogeneity by favoring high-probability alternatives, leading to potential overfitting and reduced predictive power.

Kim and Parent 2016 develop a spatial multivariate Tobit mode, estimated using Bayesian methods, enabling individuals to account for other household members' willingness to travel when selecting potential destinations. This study offers valuable insights into the impact of intra-household interactions on individual travel decisions. However, the approach does not explicitly incorporate the scheduling and sequencing of activities, limiting its direct applicability to parameter estimation in detailed household-level activity scheduling models.

In contrast to explicit choice set formation methods, implicit approaches that capture choice set availability directly within choice models have also been proposed in the literature. Paleti 2015 propose an implicit choice set generation model that approximates the Manski model (Manski 1977) to understand latent choice set in household auto ownership decisions. Their approach maintain linear complexity relative to the choice set size. However, implicit methods often increase computational complexity, making estimation procedures computationally demanding. Additionally, without explicitly defined choice sets, interpreting results and validating the realism of generated alternatives can become challenging.

Shakeel, Adnan, and Bellemans 2022 model the generation of potential joint leisure activities among household members using a latent class model. Their work specifically focuses on the generation process before the negotiation within household members related to scheduling decisions. They establish the linkage between household and individual attributes affecting joint-activity generation. The latent class model addresses the Independence of Irrelevant Alternatives (IIA) limitation inherent to Multi-nomial logit models by capturing individual heterogeneity through a discrete set of latent classes. Moreover, because it does not require distributional assumptions for parameter estimation — as would be the case with a mixture of logit model — it remains relatively straightforward to estimate. Nevertheless, further research is recommended, particularly in exploring the generation of joint activities, estimating travel parties involved in joint activity, and integrating this approach within operational ABMs.

Although the aforementioned studies provide ample insights into operationalisation of ABMs and intra-household interactions in travel demand modelling, there remains a gap in developing an operational estimation framework for household-level ABMs with simultaneous simulation of choice dimensions. Such a framework should explicitly generate household-level choice sets that ensure consistency across alternatives by accounting for intra-household interactions, while also employing random utility-based parameter estimation with added behavioural value.

3 Methodology

We introduce an estimation framework for utility-based household scheduling models in which key choice dimensions — such as participation, start time, duration, participation mode, and transport mode — are simulated simultaneously. The framework consists of two main components: (i) a household-level choice set generator, and (ii) a parameter estimation procedure based on discrete choice modelling. The choice-set generator constructs a set of schedule alternatives that individuals within a household may consider when planning their daily activities. It ensures the internal consistency of alternatives across household members by explicitly accounting for intra-household constraints and time arrangements. The choice set generator outputs consistent schedule alternatives ensured through explicit consideration of intra-household interaction constraints. To explore the combinatorial space of feasible household schedules, the generator employs a Metropolis–Hastings (MH) sampling algorithm, following the approach introduced by Pougala, Hillel, and Michel Bierlaire 2023. In the estimation component, the household scheduling process is formulated as a discrete choice problem. Model parameters are estimated by maximising the likelihood function defined over the sampled choice sets.

In the remainder of this section, we first introduce the key definitions for household scheduling problem

(Section 3.1), provide a synopsis of the household activity scheduling simulation model (Section 3.2), describe the parameter estimation procedure (Section 3.3), and finally give a detailed description of the household-level choice set generation framework (Section 3.4).

3.1 Scheduling problem definitions

Consider individual $n^{(h)}$, living in household h, consisting of N_h individuals, such that $n^{(h)} \in \{1, ..., N_h\}$. For simplicity of notation, we omit the superscript (h) in $n^{(h)}$. Each individual n schedules their activities over a time period T (e.g., 24 hours) by considering participating in activities $a_i^{(n)}$ distributed in space and time. For convenience of the notation, we drop the superscript (n) in $a_i^{(n)}$, except where necessary for clarity. Each activity a_i in the considered activity set $\mathcal{A}^n$, is an action taking place at location ℓ_{a_i} with a start time x_{a_i} and duration τ_{a_i}. Each individual also decides whether to participate in the activity jointly with (an)other individual(s) or alone, captured by a binary variable called activity participation mode p_{a_i}, which is 1 if the activity is performed with (an)other member(s) and 0 if solo. Where consecutive actions a_i and a_j are at different locations ℓ_{a_i} and ℓ_{a_j}, activity a_i would be followed by a trip with transport mode m_{a_i} - in other words, each activity is associated with the proceeding trip.

The schedule S_n of individual n is a sequence of activities over a time horizon T. The household schedule S_h is defined as the set of individual schedules $(S_n)_{n \in \mathcal{N}_h}$ for all members n of household h, coordinated over the time horizon T. The actual chosen schedule of individual n is the individual's realised schedule $\overline{S}_n$. The set of schedules realised by all members of household h, is the household's realised schedule $\overline{S}_h$. The realised schedules are a subset of feasible schedules.

In the household model, the joint participation of household members in activities and joint travels is captured, which in turn introduces the concept of household resources. Each household has limited resources r. For example, consider a household who owns N_r cars, which can be used by individuals to travel to participate in activities. In case of joint travel, multiple individuals can use the resource at the same time, with the number of users represented by the resource occupancy O_{e_r}. The maximum number of individuals that can share a resource is defined by the resource capacity C_r. Each resource has an associated schedule consisting of resource events e_r. A household resource does not make independent decisions; it is solely used by, and dependent on, the decision-making individuals. The resource event schedule is constrained by the schedules of the household members and is governed by associated constraints.

A summary of the notation introduced in the scheduling problem definition are presented in Table 1.

We treat the household scheduling as a utility maximisation problem for the household solved as an Mixed integer linear programming (MILP) optimisation problem subject to constraints and decision variables (Section 3.2). Model parameters are estimated from historic data (Section 3.3), through a choice-set generation procedure (Section 3.4).

3.2 Utility-based household activity scheduling model

Household activity scheduling is modelled as a utility optimisation problem. It is formulated as a mixed-integer optimisation framework grounded in random utility theory, jointly considering multiple scheduling decisions. The schedule of each individual is represented as a sequence of activities over a time horizon T, resulting from the individual's choices such as activity participation, activity duration, activity sequence, transportation mode, and whether to do the activity with (an)other household member(s). Each individual schedule S_n is associated with a utility function $U(S_n)$, which captures the utility of the schedule for each individual considering the preferences of the individual as well as household-related interactions, such as utility/disutility from joint activities

Table 1: Notation in problem definition and decision variables

Notation	Name	Description
$n^{(h)} \in \mathcal{N}_h$	Individual	Index of an individual belonging to household h consisting of the set of individuals $\mathcal{N}_h$ of size N_h.
$h \in \mathcal{H}$	Household	Index of a household from the set of all households $\mathcal{H}$.
$a_i^{(n)} \in \mathcal{A}^n$	Activity	Activity $a_i^{(n)}$ that can be performed by individual n, from individual-specific considered set of activities $\mathcal{A}^n$.
T	Time horizon	The time period over which the schedules are generated.
$\ell_{a_i} \in \mathcal{L}_{a_i}$	Activity location	Location for activity a_i from set of possible locations $\mathcal{L}_{a_i}$.
$m_{a_i} \in \mathcal{M}^n$	Transportation mode	The mode of travel from the location of the current activity, ℓ_{a_i}, to the location of the following activity, ℓ_{a_j}, chosen from discrete and finite individual-specific list of considered transport modes $\mathcal{M}^n$.
$x_{a_i} \in [0, T]$	Activity start time	A continuous variable representing the start time of activity a_i.
$\tau_{a_i} \in [0, T]$	Activity duration	A continuous variable representing the duration of activity a_i.
$p_{a_i} \in \{0, 1\}$	Activity participation mode	Equals to 1 if activity a_i is performed jointly with (an)other individual(s), and 0 if performed solo.
$S_n \in \mathcal{F}_n$	Individual schedule	An ordered list of activities for individual n covering a time horizon T, from feasible set of schedules $\mathcal{F}_n$.
$S_h \in \mathcal{F}_h$	Household schedule	An ordered set $(S_1, .., S_{N_h})$ of schedules for all household members from feasible set of schedules $\mathcal{F}_h$.
$\overline{S}_n, \overline{S}_h$	Realised schedules	The actual schedule/set of schedules realised by an individual or household.
$r \in \mathcal{R}_h$	Resource index	Index of a household resource used by its members,from the set of household resources $\mathcal{R}_h$ of size $N_r^{(h)}$. Resources have no independent decision making capabilities and are purely dependent on the decision-making individuals.
$O_{e_r} \in \{0, \ldots, C_r\}$	Resource occupancy	The number of individuals using resource r at the same time, where C_r is the resource capacity.
$e_r \in E^r$	Resource event	Event e_r that can be scheduled for resource r, from the associated resource event set E^r.

or escorting other household members. The household schedule comprises the schedules of all its individual members. The household schedule S_h is associated with a household utility function $U(S_h)$, capturing collective household decision-making, which reflects trade-offs between individual utilities. The scheduling model is defined subject to a set of constraints that ensures the validity of individual schedules and captures inter-personal interactions between household members.

The goal of the decision-making individuals is to maximise the utility of the entire household. Thus, the objective function in household scheduling is defined as Equation 1:

$$\max U(S_h) \tag{1}$$

$$U(S_h) = \sum_{n=1}^{n=N_h} w_n \, U(S_n) \tag{2}$$

where $U(S_h)$ denotes the household utility function. $U(S_n)$ is the utility of the schedule S_n for individual n in household h. $U(S_n)$ can be either positive, negative, or zero. w_n is the individual priority parameter, which captures the relative "power" of each individual in the household-oriented decisions. N_h is the number of individuals in household h. Equation 2 is written for one specific household, but in parameter estimation, the maximum likelihood loops over all households in the dataset, $h \in \mathcal{H}$.

The utility function $U(S_n)$, captures the utility of the schedule for each individual n in the household. Possible interaction aspects are captured in the utility function of individuals. We use the same form of utility form as Pougala, Hillel, and Michel Bierlaire 2022, with added terms incorporated to capture possible interaction aspects in the utility function. $U(S_n)$ is made up of a generic utility, $U^{\text{gen}}_{S_n}$, linked to the entire schedule of the individual and utility components linked to the performed activities, U_{a_i}. The generic utility, $U^{\text{gen}}_{S_n}$, captures schedule-level preferences not directly linked with any specific activity, such as a dislike for overly busy days or a preference for including at least one out-of-home activity. U_{a_i} is specified as the sum of components capturing the individual's activity and travel behaviour (e.g., time sensitivity), as well as capturing possible interaction aspects within the utility function. The general form of $U(S_n)$ is defined as follows:

$$\begin{aligned} U(S_n) &= U^{\text{gen}}_{S_n} + \sum_{a_i \in A^n} U_{a_i} \\ &= U^{\text{gen}}_{S_n} + \sum_{a_i \in A_n} \left(U^{\text{partic}}_{a_i} + U^{\text{start}}_{a_i} + U^{\text{duration}}_{a_i} + \sum_{a_j \in A^n} (U^{\text{travel}}_{a_i,a_j}) \right) + \epsilon_{S_n} \end{aligned} \tag{3}$$

$U^{\text{start}}_{a_i}$ captures the perceived penalty of deviation in start time from the desired start time ($x^*_{a_i}$). $U^{\text{duration}}_{a_i}$ captures the perceived penalty of deviation in duration of activity a_i from the preferred duration ($\tau^*_{a_i}$), which can be either single values or time intervals. $U^{\text{travel}}_{a_i,a_j}$ is a utility term associated with the trip from ℓ_{a_i} to ℓ_{a_j}, including the penalty associated with travel time and other travel variables such as travel cost. The utility terms also include a random error term ϵ_{S_n}, capturing the unobserved variables. The error terms are assumed to be i.i.d. and Extreme Value distributed, with a scale parameter μ fixed to 1 for identification purposes. $U^{\text{partic}}_{a_i}$ is a utility term purely associated with participation in activity a_i, irrespective of its timing and associated trips. Possible interaction terms such as joint activity participation and escorting are considered in $U^{\text{partic}}_{a_i}$ as Equation 4. It is notable that more complex forms of the utility function can be also utilised.

$$U^{\text{partic}}_{a_i} = U^{\text{const}}_{a_i} + U^{\text{joint}}_{a_i} + U^{\text{escort}}_{a_i} \tag{4}$$

where:

- $\mathbf{U^{\text{partic}}_{a_i}}$**:** a utility term, which is purely associated with participation in activity a_i, irrespective of any schedule deviations and travel behaviour.

- $\mathbf{U}^{\mathbf{const}}_{a_i}$**:** an activity specific constant utility term capturing the inherent preference for participation in activity a_i.
- $\mathbf{U}^{\mathbf{joint}}_{a_i}$**:** a utility term for joint engagement, accounting for the (dis)utility of participating in activity a_i jointly with other household member(s).
- $\mathbf{U}^{\mathbf{escort}}_{a_i}$**:** a utility term associated with (dis)utility of doing an escorting task.

The household scheduling problem is subject to a set of constraints that ensure the validity of schedules with respect to both individual-level and household-level restrictions. Individual-level feasibility constraints ensure the continuity of schedules by ensuring:

1. Time budget constraint; the simulated schedules should fit within the individual's time budget and cannot exceed it,
2. Sequence constraints; each activity must start only after the completion of the trip from the preceding activity,
3. Consistent transport modes in tours,
4. Feasible time windows; activities should be scheduled within their feasible time windows, such as the store opening hours for shopping activities.

The inter-personal interactions within a household are captured through a set of household constraints as follows:

$$\omega_{a_i} + m^V_{a_i} \leqslant N^h_V + 1 \qquad \forall a_i \in \mathcal{A}^n, \forall n \in \mathcal{N}_h \tag{5}$$

$$z_{e_r e'_r} + z_{e'_r e_r} \leqslant 1 \qquad \forall e_r, e'_r \in E^r, e_r \neq e'_r \tag{6}$$

$$\left(z_{e_r e'_r} - 1\right) T \leqslant x_{e_r} + \tau_{e_r} - x_{e'_r} \leqslant \left(1 - z_{e_r e'_r}\right) T \qquad \forall e_r, e'_r \in E^r \tag{7}$$

$$\sum_{e_r \in E^r} \frac{\tau_{e_r}}{O_{e_r}} = T \tag{8}$$

$$O_{e_r} \leqslant C_r \qquad \forall e_r \in E^r \tag{9}$$

$$\omega_{e_r} = \omega_{a_i} \qquad \forall e_r \in E^r \cap \mathcal{A}^n, \forall a_i \in \mathcal{A}^n \cap E^r, \forall n \in \{\text{Adults}\} \tag{10}$$

$$x_{e_r} = x_{a_i} + \tau_{a_i} \qquad \forall e_r \in E^r \cap \mathcal{A}^n, \forall a_i \in \mathcal{A}^n \cap E^r, \ell_{a_i} \in \{\text{Home}\}, \forall n \in \{\text{Adults}\} \tag{11}$$

$$\tau_{e_r} = \sum_{a_j \in \mathcal{A}^n} \left(z_{a_i a_j}\, \rho(\ell_{a_i}, \ell_{a_j}, \text{Driving})\right)$$
$$\forall e_r \in E^r \cap \mathcal{A}^n,\ \forall a_i \in \mathcal{A}^n \cap E^r,\ \ell_{a_i} \in \{\text{Home}\},\ \forall n \in \{\text{Adults}\} \tag{12}$$

$$x_{e_r} = x_{a_i} \qquad \forall e_r \in E^r \cap \mathcal{A}^n, \forall a_i \in \mathcal{A}^n \cap E^r, \ell_{a_i} \notin \{\text{Home}\}, \forall n \in \{\text{Adults}\} \tag{13}$$

$$\tau_{e_r} = \tau_{a_i} + \sum_{a_j \in \mathcal{A}^n} \left(z_{a_i a_j}\, \rho(\ell_{a_i}, \ell_{a_j}, \text{Driving})\right)$$
$$\forall\, e_r \in E^r \cap \mathcal{A}^n,\ a_i \in \mathcal{A}^n \cap E^r,\ \ell_{a_i} \notin \{\text{Home}\},\ n \in \{\text{Adults}\} \tag{14}$$

$$\omega_{a_i^n} = \omega_{a_i^{n'}} \qquad \forall a_i \in \mathcal{A}^n \cap \mathcal{A}^{n'}, p_{a_i^n} = p_{a_i^{n'}} = 1, \forall n,\ n' \in \mathcal{N}_h \quad (15)$$

$$x_{a_i^n} = x_{a_i^{n'}} \qquad \forall a_i \in \mathcal{A}^n \cap \mathcal{A}^{n'}, p_{a_i^n} = p_{a_i^{n'}} = 1, \forall n,\ n' \in \mathcal{N}_h \quad (16)$$

$$\tau_{a_i^n} = \tau_{a_i^{n'}} \qquad \forall a_i \in \mathcal{A}^n \cap \mathcal{A}^{n'}, p_{a_i^n} = p_{a_i^{n'}} = 1, \forall n,\ n' \in \mathcal{N}_h \quad (17)$$

$$\sum_{n \in \text{Adults}} \omega_{a_i^n} = \omega_{a_i^{\text{Passenger}}} \qquad \forall a_i \in \mathcal{A}^{\text{Passenger}} \cap \mathcal{A}^{\text{Adults}}, \lambda_{a_i^{\text{Passenger}}} = \lambda_{a_i^n} = 1 \quad (18)$$

$$\sum_{n \in \text{Adults}} x_{a_i^n} = x_{a_i^{\text{Passenger}}} \qquad \forall a_i \in \mathcal{A}^{\text{Passenger}} \cap \mathcal{A}^{\text{Adults}}, \lambda_{a_i^{\text{Passenger}}} = \lambda_{a_i^n} = 1 \quad (19)$$

$$\sum_{n \in \text{Adults}} \tau_{a_i^n} = \tau_{a_i^{\text{Passenger}}} \qquad \forall a_i \in \mathcal{A}^{\text{Passenger}} \cap \mathcal{A}^{\text{Adults}}, \lambda_{a_i^{\text{Passenger}}} = \lambda_{a_i^n} = 1, \chi_{a_i^{\text{Passenger}}} = 0 \quad (20)$$

$$\sum_{n \in \text{Adults}} \sum_{a_j^n \in \mathcal{A}^n} \left(z_{a_j^n a_i^n}\ \ell_{a_j^n} \right) = \sum_{a_j^{\text{Passenger}} \in \mathcal{A}^{\text{Passenger}}} \left(z_{a_j^{\text{Passenger}} a_i^{\text{Passenger}}}\ \ell_{a_j^{\text{Passenger}}} \right)$$
$$\forall a_i \in \mathcal{A}^{\text{Passenger}} \cap \mathcal{A}^{\text{Adults}},\ \lambda_{a_i^{\text{Passenger}}} = \lambda_{a_i^n} = 1,\ \chi_{a_i^{\text{Passenger}}} = 0 \quad (21)$$

$$\sum_{n \in \text{Adults}} \sum_{a_j^n \in \mathcal{A}^n} \left(z_{a_i^n a_j^n}\ \ell_{a_j^n} \right) = \sum_{a_j^{\text{Passenger}} \in \mathcal{A}^{\text{Passenger}}} \left(z_{a_i^{\text{Passenger}} a_j^{\text{Passenger}}}\ \ell_{a_j^{\text{Passenger}}} \right)$$
$$\forall a_i \in \mathcal{A}^{\text{Passenger}} \cap \mathcal{A}^{\text{Adults}},\ \lambda_{a_i^{\text{Passenger}}} = \lambda_{a_i^n} = 1,\ \chi_{a_i^{\text{Passenger}}} = 0 \quad (22)$$

$$\sum_{n \in \text{Adults}} \tau_{a_i^n} = \vartheta\ \omega_{a_i^{\text{Passenger}}} \qquad \forall a_i \in \mathcal{A}^{\text{Passenger}} \cap \mathcal{A}^{\text{Adults}}, \lambda_{a_i^{\text{Passenger}}} = \lambda_{a_i^n} = 1, \chi_{a_i^{\text{Passenger}}} = 1 \quad (23)$$

$$\sum_{n \in \text{Adults}} \sum_{a_j^n \in \mathcal{A}^n} \left(z_{a_i^n a_j^n}\ \ell_{a_j^n} \right) = \sum_{a_j^{\text{Passenger}} \in \mathcal{A}^{\text{Passenger}}} \left(z_{a_i^{\text{Passenger}} a_j^{\text{Passenger}}}\ \ell_{a_j^{\text{Passenger}}} \right)$$
$$\forall a_i \in \mathcal{A}^{\text{Passenger}} \cap \mathcal{A}^{\text{Adults}},\ \lambda_{a_i^{\text{Passenger}}} = \lambda_{a_i^n} = 1,\ \chi_{a_i^{\text{Passenger}}} = 1 \quad (24)$$

$$\sum_{n \in \text{Adults}} \tau_{a_i^n} = \vartheta\ \omega_{a_i^{\text{Passenger}}} \qquad \forall a_i \in \mathcal{A}^{\text{Passenger}} \cap \mathcal{A}^{\text{Adults}}, \lambda_{a_i^{\text{Passenger}}} = \lambda_{a_i^n} = 1, \chi_{a_i^{\text{Passenger}}} = 2 \quad (25)$$

$$\sum_{n \in \text{Adults}} \sum_{a_j^n \in A^n} \left(z_{a_j^n a_i^n}\ \ell_{a_j^n} \right) = \sum_{a_j^{\text{Passenger}} \in \mathcal{A}^{\text{Passenger}}} \left(z_{a_j^{\text{Passenger}} a_i^{\text{Passenger}}}\ \ell_{a_j^{\text{Passenger}}} \right)$$
$$\forall a_i \in \mathcal{A}^{\text{Passenger}} \cap \mathcal{A}^{\text{Adults}},\ \lambda_{a_i^{\text{Passenger}}} = \lambda_{a_i^n} = 1,\ \chi_{a_i^{\text{Passenger}}} = 2 \quad (26)$$

Equation 5 is household private vehicle ownership constraints, such that if a household owns no private vehicles, no member can choose one as a transport mode. $m^V_{a_i}$ is an indicator variable that is 1 if a private mode is chosen for activity a_i and 0 otherwise. N_V^h is number of household private cars.

Equations 6 - 14 define the allocation of resources to household members. Household resources refer to shared assets that are used by household members, exemplified by cars in this study, but can be generalised to other resources (e.g., home office space) in future applications. For each resource r, an event schedule is considered. Each resource r has a capacity C_r that limits the maximum number of individuals that can use it at the same time. Number of users using the car at the same time is denoted by O_{e_r} as resource occupancy. Moreover, the moving resources need a driver to move them. Therefore, their schedule is constrained to that of the individuals in the household and additional physical constraints exist for the non-static resources. This is a general approach applicable to any household resource. Equations (6-8 ensure event schedule validity for resource r. Equation 9 is resource capacity constraint ensuring that at each resource event, the occupancy of each resource O_{e_r} do not exceed its capacity C_r. Equations 10 -14 enforce physical constraints on private vehicle resources such that an event can be scheduled for them only if it is accompanied by an adult individual throughout the tour, enforcing consistency of the resource event schedule with the adult individuals in the household.

Equations 15-17 are joint activity participation and ride-share to joint activities constraints. The joint activity participation is considered as a constraint; if there is a joint activity, both members must participate, or the joint activity is canceled (15). Furthermore, the consistency of space (embeded in the chosen activity) and time for all participating members are ensured (16-17).

Equations 18 - 26 define escort activities within household members. Escort is considered as a trip chauffeured by one of the adults in the household with a private vehicle. Escorting by multiple household heads is not included in the presented specification, but can be adopted within the framework. Both types of escort where the core adult picks up/drops off the passenger from/to the activity location (pick-up and drop-off), and where the adult accompanies the passenger throughout the entire tour (escort and stay) are considered. Each escort activity is associated with an indicator variable indicating its type, $\chi_{a_i^n}$. $\chi_{a_i^n}$ is 0 for escort and stay, 1 for the pick-up, and 2 for the drop-off escort type. $\lambda_{a_i^n}$ is a binary variable escort indicator such that for each activity $a_i^{(n)}$, which specifies whether activity $a_i^{(n)}$ is/needs escort or not. $\lambda_{a_i^n}$ is defined as follows:

- for individuals needing escort: $\lambda_{a_i^n}$ specifies whether individual n *needs* to be escorted for activity $a_i^{(n)}$ (1), or not (0), and
- for individuals providing escort: $\lambda_{a_i^n}$ specifies whether activity $a_i^{(n)}$ performed by individual n *is* an escort (1), or not (0).

Variable ϑ is the stop time duration needed to pick-up or drop-off the passenger (e.g., 5 min). Equations 21, 22, 24, and26 ensure location consistency between the passenger and the adult escorting individual.

The framework takes as input the household composition, scheduling preferences, activity flexibilities, household resources and their associated events sets, as well as, a considered activity set including their associated locations, transport modes, and participation modes for each individual in the household. They are utilised to define a distribution over possible schedules from which random realisations can be generated. Due to the stochastic nature of the utility function presented in Equation 3, the model generates empirical instances of the distribution. A simulation technique is used to generate several draws from the distributions of the random terms, and then solve the optimisation problem explicitly for each realisation. The outcome of the schedule simulation model is a realisation from the distributions of valid schedules, presenting the schedules of the individuals in the same household under both individual- and household-level constraints and preferences.

A comprehensive explanation of the household scheduling simulation model can be found in (Rezvany, Michel Bierlaire, and Hillel 2023). The key additional notation used in the household utility scheduling framework are summarised in Table 2.

Table 2: Notation for household scheduling framework

Notation	Name	Description
$U(S_n), U(S_h)$	Schedule utilities	Latent (unobserved) utility of participating in schedule S_n/S_h for individual n/household h, respectively.
$x^*_{a_i}, \tau^*_{a_i}$	Desired start times and durations	A continuous indicator representing the desired start time/duration for activity $a_i^{(n)}$ for individual n.
$\rho(\ell_{a_i}, \ell_{a_j}, m_{a_i})$	Travel time	The travel time between the locations ℓ_{a_i} and ℓ_{a_j} with mode m_{a_i}.
$z_{a_i a_j} \in \{0,1\}$	Activity succession	Equals to 1 if activity a_j is scheduled immediately after activity a_i, and 0 otherwise.
$\omega_{a_i} \in \{0,1\}$	Activity participation	Equals to 1 if individual n participates in activity a_i, and 0 otherwise.
$\chi_{a_i^n} \in \{0,1,2\}$	Escort type indicator	Equals to 0 if escort and stay, 1 if pick-up, and 2 if drop-off.
$\lambda_{a_i^n} \in \{0,1\}$	Escort indicator	Equals to 1 if activity a_i^n needs/is escort, and 0 otherwise.
$\vartheta \in [0, T]$	Stop time duration	Stop time duration needed to pick-up or drop-off the passenger.
w_n	Individual priority parameter	Relative weight capturing the priority that is placed on the schedule utility of each individual in household decision making.

3.3 Parameter estimation

The household scheduling process is defined as a discrete choice problem. Each alternative is a household daily schedule, containing a set of full daily schedules of all household members. Each alternative is associated with a utility, capturing the household utility. The parameters of utility-based scheduling model can be estimated with maximum likelihood estimation on a choice set. The likelihood function is evaluated for each alternative of the choice set. The parameters are derived such that the likelihood function is maximised.

The household members select their daily activity schedules from a choice set containing sets of alternative schedules for all household members, $\mathcal{C}_h$. We consider a daily schedule of household, S_h, to be a discrete alternative, characterised by a set of constraints defining feasible schedules. The feasibility constraints ensure the soundness of the alternative schedule both in terms of schedule continuity such as time budget, activity participation, timing consistency, transport mode consistency, as well as the constraints that appear due to interpersonal dependencies within household members such as schedule synchronisations for joint activities and travels. The full universal choice set for the household, $\mathcal{C}_h$, comprised of all possible combinations of activity scheduling choices for all household members, is combinatorial and cannot be enumerated. Thus, we consider a finite subset of the universal choice set, $\tilde{\mathcal{C}}_h \subset \mathcal{C}_h$, containing a sample of alternative schedules for household h. The sampled choice set $\tilde{\mathcal{C}}_h$ contains alternative household schedules $S_h \in \tilde{\mathcal{C}}_h$, where each alternative is a set of feasible full daily schedules for all individuals in the household. We show how to generate the sampled choice set in Section 3.4.

As the evaluation is carried out on a sample of the full universal choice set, the likelihood function is corrected with probability of sampling the choice set given the chosen alternatives (Moshe E Ben-Akiva and Lerman 1985). $\tilde{\mathcal{C}}_h$ is the sampled choice set for household h. Thus, the probability that household h chooses alternative schedule $\bar{S}_h \in \tilde{\mathcal{C}}_h$, associated with a deterministic utility $V_{\bar{S}_h}$, is defined as follows:

$$P_{\bar{S}_h} = P(\bar{S}_h|\tilde{\mathcal{C}}_h) = \frac{\exp[\mu V_{\bar{S}_h} + \ln P(\tilde{\mathcal{C}}_h|\bar{S}_h)]}{\sum_{S_h \in \tilde{\mathcal{C}}_h} \exp[\mu V_{S_h} + \ln P(\tilde{\mathcal{C}}_h|S_h)]} \tag{27}$$

$\tilde{\mathcal{C}}_h$ is the sampled choice set for household h, which contains sets of schedules for all individuals in the household. V_{S_h} is the deterministic utility of the total household for alternative household schedule S_h. μ is a scale parameter. The alternative-specific correction term accounts for sampling biases, as defined by Moshe E Ben-Akiva and Lerman 1985:

$$P(\tilde{\mathcal{C}}_h|S_h) = \frac{1}{p(S_h)} \prod_{S_h \in \tilde{\mathcal{C}}_h} \left(\sum_{S_h \in \tilde{\mathcal{C}}_h} p(S_h) \right)^{J+1-J^*} \tag{28}$$

where $\tilde{\mathcal{C}}_h$ is the household choice set of size $J+1$ with J^* unique alternatives for household h. Unique alternatives are identified based on the combination of schedules of all household members. S_h represents alternative sampled from the target distribution of the MH algorithm with probability $p(S_h)$. The target distribution depends on the generation protocol for the sample. For each household and each alternative in their respective choice sets, the sample correction term is evaluated to be added to the utility function.

As the sampled choice set contains household socio-demographic characteristics, the utility function can include socio-demographic characteristics, capturing their effect on household scheduling decisions.

A summary of key new notation introduced in parameter estimation is presented in Table 3.

Table 3: Notation used in parameter estimation

Notation	Name	Description
$\mathcal{C}_h$	Household choice set	Universal choice set for household h, which is the complete set of all possible alternatives, each being a set of household member schedules.
$\tilde{\mathcal{C}}_h$	Sampled household choice set	Sampled choice set for household h.
V_{S_n}, V_{S_h}	Deterministic schedule utilities	Deterministic utility component of schedule S_n/S_h for individual n/household h, respectively.
μ	Scale parameter	A strictly positive scale parameter.

3.4 Household-level choice set generation

In our approach, in order to estimate the parameters of the household scheduling model, we require a relatively small sample of alternatives for each observed choice. Each alternative is a set of feasible joint schedules of all members of the household. However, in order for the estimation procedure to effectively capture how households and individuals make trade-offs, it is crucial that the sampled alternatives are competitive with the observed (chosen) one. Thus, the generated schedules should be plausible alternatives — those that could reasonably have been chosen and therefore are likely to have high utility. To achieve this, the method proceeds as follows: (i) We postulate initial values for the unknown parameters based on prior knowledge, existing literature, or engineering intuition. (ii) Using these parameter values, we sample alternatives with a probability proportional to $\exp(U(S_h))$, where $U(S_h)$ denotes the utility of alternative schedule S_h for household h. This sampling is performed using the Metropolis-Hastings algorithm. This strategy ensures that the alternative schedules in the sample are not only representative but also relevant for inferring preferences from observed choices.

Intra-household interactions influence how household members schedule their day, introducing additional choice dimensions, temporal and spatial arrangements, constraints, and group decision-making mechanisms. These factors should be incorporated into the generation of the activity schedule choice set to ensure that the alternatives are compliant with intra-household constraints and arrangement. Since the estimation of ABMs parameters are carried out by enumerating the alternatives in the choice set, using a choice set composed of valid household schedules enhances the behavioural realism of the resulting estimates.

The Metropolis-Hastings sampling approach enables estimation of utility parameters without enumerating the full choice set—which is infeasible due to dimensionality, while maintaining tractable probabilities to compute the sample correction for the likelihood function. Building on the strategic MH sampling algorithm of Pougala, Hillel, and Michel Bierlaire 2023, we generate an ensemble of relevant schedules for the household model, to estimate significant and meaningful parameters. The choice set generation is modelled as a Markov process. The Metropolis-Hastings algorithm performs a random walk over the state space accepting each candidate state with an acceptance probability that guides the chain toward high-probability regions (Hastings 1970). The state space in the context of household ABMs is defined as the set of all feasible joint schedules of all members of the household. This sampling algorithm enables efficient exploration without requiring full enumeration.

Each iteration of the random walk consists of two main steps: generating a candidate point (Section 3.4.1) using operators (Section 3.4.2), and then accepting or rejecting it (Section 3.4.3).

3.4.1 Generation of a candidate state

In the household-level choice-set generation technique, at each step of the random walk, alternative schedules for all individuals within a household are generated in parallel. This approach ensures compatibility among the schedules of household members in the generated alternatives and preserves the relation between individuals and their households. During the random walk process, the household state at step t, denoted as S^t_h, represents the household schedule comprised of a set of feasible schedules of its individual members, $S^t_h = \left(S^t_1, \dots, S^t_n, \dots, S^t_{N_h}\right)$. The state of each individual n in household h, S^t_n, is the individual's activity schedule defined within the time budget T (e.g., 24 hours), discretised in time blocks of duration $\delta \in [\delta_{\text{min}}, T - \delta_{\text{min}}]$, where δ_{min} is the minimum block duration. The discretisation defines the scale of the potential modifications for operators.

The MH for the household problem is decomposed as follows. At each step of the random walk, the new household candidate state S^*_h is generated by applying a heuristic operators ψ to each individual member's schedule, producing a neighbouring joint household schedule $S^*_h = \left(S^*_1, \dots, S^*_n, \dots, S^*_{N_h}\right)$. The operators are chosen from a list of operators $\psi \in \Psi$ defined by the modeller to modify aspects of the current schedule to generate a neighbouring candidate state. Operators can be different for each household member, producing a complex joint proposal. A set of validity constraints should be checked for the generated states to ensure that the choice set only contains feasible schedules in terms of household-level interaction and schedule continuity constraints as defined in Section 3.2.

For each household, the generation algorithm is initialised with a random household schedule S^0_h (e.g., an ensemble of reported schedules from all household members) as the starting point. One individual, denoted as $n^{\text{ref},h}$, is selected as the reference member for household h. The procedure for selecting the reference individual is determined by the modeller and can follow a random or rule-based approach, for example, prioritising individuals based on employment status. At each step of the random walk, the schedules of all household members are checked for consistency with that of the reference individual - such as alignment in time and location for joint activities - to ensure intra-household schedule synchronisation. In other words, the candidate schedule of the reference individual serves as the benchmark for validating the consistency of schedules of household members at each random walk iteration.

At each step of the random walk, the solution space of the reference individual is first explored by applying a randomly selected operator ($\psi \in \Psi$) to the current state of the reference individual ($S^t_{n^{\text{ref},h}}$), while ensuring the continuity of the resulting schedule. If the generated state for the reference individual is infeasible with respect to individual-level constraints, the process is repeated until a feasible candidate state is obtained for the reference individual ($S^*_{n^{\text{ref},h}}$). Once a valid candidate state is generated for the reference individual, the combinatorial solution space of other household members is explored by applying randomly selected operators to the current states of the each individual, ensuring the resulting schedules are validated for continuity and consistency with the reference individual. For example, if the reference individual schedules a joint leisure activity at the cinema with another household member from 17:00 to 19:00, a corresponding joint leisure activity with the same temporal and spatial characteristics must be added to the schedule of the other individual household member(s). If a generated state for any individual is found to violate either individual constraints or consistency with the reference individual, the process is repeated until a feasible candidate state (S^*_n) is produced for all individual members. The set of the proposed states of all household members gives us the proposed new joint household state $S^*_h = \left(S^*_1, \ldots, S^*_n, \ldots, S^*_{N_h}\right)$. Figure 1 illustrates a visual scheme of the procedure for household-level choice set generation at each iteration of the random walk.

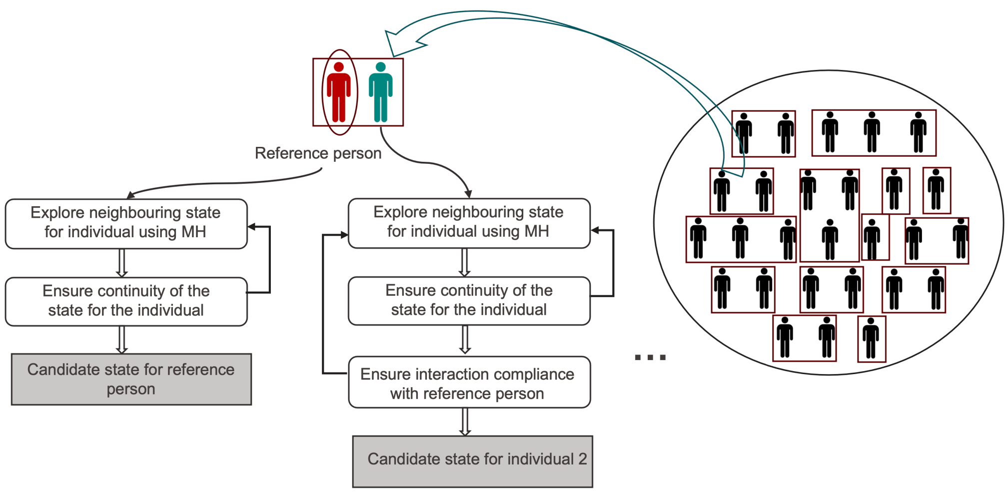


Figure 1: Scheme of household-level choice set generation procedure at each iteration of the random walk.

3.4.2 Operators

Operators $\psi \in \Psi$ are heuristics that modify aspects of the current individual schedule to generate a neighbouring candidate state that differs from the current individual state in one or more dimensions. This dimension may correspond to temporal adjustments, changes in location, modifications in activity participation or companionship during the activity, ensuring that the candidate state remains feasible while exploring the local neighbourhood of the current state. Each household member schedule S^t_n is characterised by one or more anchor nodes ν marking the start of blocks, where each block represents the temporal extent of an operator-induced change. Each operator ψ is selected with a probability P_ψ determined by the modeller. These operators should ensure irreducibility and reversibility of the Markov chain by allowing transitions to any feasible state and enabling backtracking. A set of validity constraints should be checked for the generated states to ensure that the choice set only contains feasible schedules in terms of schedule continuity and intra-household constraints.

Different operators can be created according to the modeller's needs and specifications. We consider a set of operators described in Pougala, Hillel, and Michel Bierlaire 2023, which account for the scheduling aspects of each household member, along with additional operator(s) specifically designed for the household context.

- Anchor (ψ_{anchor}); inserts an anchor node ν into the schedule without altering the activity sequence, but shifts the position where other operators may apply modifications.
- Assign (ψ_{assign}); changes the activity type in a time block.

- Swap (ψ_{swap}); randomly swaps the activities of two adjacent blocks.
- Inflate/Deflate ($\psi_{\text{inf/def}}$); modifies durations of activities. It enables a shift in the schedule by randomly increasing (i.e., adding a block of length δ) or decreasing (i.e., removing a block of length δ) the duration of an activity.
- Location (ψ_{loc}); modifies the location of a randomly selected activity.
- Mode (ψ_{mode}); modifies the travel mode of outbound trip of a randomly selected activity.
- Block (ψ_{block}); changes temporal resolution by changing the length of time discretisation blocks, δ, allowing changes in the scale of potential modifications of other operators.

Dedicated operators are implemented for the household context. For instance, we define an operator that changes whether an activity is performed jointly with other member(s) of the household or alone. This operator is called participation mode operator, $\psi_{\text{partic\_mode}}$. When the participation mode of an activity is modified, the algorithm checks for schedule synchronisation among household members. The corresponding activity is added to the schedule of the accompanying member(s) with identical timing, location, and participation mode. To ensure schedule validity, the resulting schedules must start and end at home, and the participation mode of home activity cannot be changed. Thus, the chosen block to apply the change, cannot exceed the time budget by more than the minimum time at home, 2δ. The transition probability associated with this change is defined as the product of three components: the probability of selecting the participation mode operator, the probability of choosing a valid time block, and the probability of selecting one of the possible participation modes. The participation mode is selected according to a probability distribution (P_π), which is assumed to be exogenous to the choice set generation process. This transition probability can be expressed as follows:

$$q(S_n^{t+1}|S_n^t) = q(S_n^t|S_n^{t+1}) = \begin{cases} P_{\text{partic\_mode}} P_\pi \frac{T-2\delta}{T\delta}, & \text{if } b_i \notin \{b_o, b_T\} \\ 0, & \text{otherwise} \end{cases} \tag{29}$$

where S_n^t and S_n^{t+1} are the schedule states of individual n belonging to household h, at step t and $t+1$, respectively.

A combination of two or more distinct operators can be also utilised through a meta-operator ψ_{meta}, where the transition probabilities of the change are the combined forward and backward probabilities of the selected operators.

3.4.3 Acceptance of candidate point

The MH algorithm iteratively samples from the set of feasible household schedule $S_h \in \mathcal{F}_h$. At each iteration t, a new candidate schedule S_h^* is proposed from a proposal distribution $q(S_h^* \mid S_h^t)$, where S_h^t is the current household schedule state. The acceptance of a candidate household state in the Metropolis-Hastings algorithm is governed by the acceptance probability, which compares the proposed schedule S_h^* with the current schedule S_h^t, considering both their likelihoods and the proposal distributions. Formally, the acceptance probability $\alpha(S_h^t, S_h^*)$ for moving from the current household state S_h^t to the proposed state S_h^* is given by:

$$\alpha(S_h^t, S_h^*) = \min\left(1, \frac{p(S_h^*)q(S_h^t|S_h^*)}{p(S_h^t)q(S_h^*|S_h^t)}\right) \tag{30}$$

where $p\left(S_h^t\right)$ is the target sampling probability which is proportional to household utility function. Inspired by Danalet and Michel Bierlaire 2015, for each household state S_h^t, the target weight is defined as $p\left(S_h^t\right) = \exp\left(\hat{U}(S_h^t)\right)$. $\hat{U}(S_h^t)$ is the household utility function whose parameters are postulated based on the literature, prior knowledge, or engineering intuition. To compute the total utility for the household, the utility of individual household individuals should be combined, depending on the nature of the group decision-making strategy.

$$\hat{U}(S_h) = f(\hat{U}(S_1), \ldots, \hat{U}(S_n), \ldots, \hat{U}(S_{N_h})) \tag{31}$$

For example, in Additive-type household, the household utility is defined as the weighted sum of the utility that each individual in the household as Equation 2.

$q\left(S_h^* \mid S_h^t\right)$ is the joint proposal distribution for candidate household states to go from one schedule state to the candidate state. The proposal distribution is obtained from the applied operator(s). As the state of household members are conditional on those of the reference member, the joint proposal distribution for the household is:

$$q\left(S_h^* \mid S_h^t\right) = q\left(S_{n^{\mathrm{ref},h}}^* \mid S_{n^{\mathrm{ref},h}}^t\right) \prod_{n \neq n^{\mathrm{ref},h}} q\left(S_n^* \mid S_n^t, S_{n^{\mathrm{ref},h}}^*\right) \tag{32}$$

As the proposals for non-reference members depend on the reference state proposal, the joint transition distribution is a conditional factorisation reflecting the dependence induced by synchronising schedules on the reference member. $q\left(S_h^t \mid S_h^*\right)$ is the reverse proposal from a reverse application of the operators. The reverse proposal is determined based on: (i) the reverse of the operator applied, where the reverse proposal would involve undoing applied operators to each individual member to return to the previous schedule, and (ii) proposal dependence, where if the operators are dependent or there is coordination between schedules, the reverse proposal must respect this coordination to ensure that S_h^t is a valid household schedule. Given the setup of the household model with a reference person, the full reverse proposal distribution is as follows, ensuring that the transitions of the entire household's state are handled coherently and consistently, while respecting the interdependencies among the individual schedules:

$$q\left(S_h^t \mid S_h^*\right) = q\left(S_{n^{\mathrm{ref},h}}^t \mid S_{n^{\mathrm{ref},h}}^*\right) \prod_{n \neq n^{\mathrm{ref},h}} q\left(S_n^t \mid S_n^*, S_{n^{\mathrm{ref},h}}^t\right) \tag{33}$$

The candidate point is accepted based on a predefined acceptance/rejection rule, where the candidate schedule state is accepted as the state of the next step with probability $\alpha(S_h^t, S_h^*)$, and otherwise the state of the next step is remained as S_h^t. The acceptance probability in the Metropolis-Hastings algorithm ensures convergence to the target distribution, by maintaining detailed balance such that it controls how the chain moves through the state space in such a way that the stationary distribution of the chain matches the target distribution.

The output of the generator is an ensemble comprising sets of schedules for all individuals within each household. The procedure for household choice set formation is summarised in Algorithm 1. Importantly, socio-demographic characteristics of individuals and their households — such as household structure, employment status, and car ownership — are preserved throughout the choice set generation process. These attributes are explicitly captured and embedded within the generated alternatives in the choice set. This feature prevents information loss and enables investigating more behavioural implications explaining the choice of schedules through estimating model specifications with socio-demographic variables.

3.4.4 Implementation notes

The choice set generator framework takes as input the composition of households, selected schedules of individuals in each household, number of alternatives to generate for each household (N_{alt}), and the configuration of the MH algorithm. This includes the number of iterations for the random walk (N_{iter}), number of initial samples to discard as a burn-in period to reduce the effect of initial conditions (N_{burn}), number of samples to skip between accepted samples to reduce autocorrelation (thinning) (N_{skip}), the set of heuristic operators used during the random walk (Ψ), probability distribution of selecting each operator (P_ψ), and the distribution of target weights. These inputs are used to generate a choice set for each household, consisting N_{alt} alternative household schedules, along with the list of accepted operators, and the corresponding probabilities of the generated alternatives.

A subset of accepted schedules after a burn-in phase, N_{burn}, forms the household choice set $\tilde{\mathcal{C}}_h$ for each household h. This MH-based choice set generation technique enables a tractable and behaviourally meaningful

Algorithm 1 Household-level choice-set generation for ABMs with MH

$t \leftarrow 0$, initialise household state with random *household schedule* $S_h^t \leftarrow S_h^0$
▷ A household is comprised of N_h individuals, with each individual having a schedule state S_n^t at iteration t.
Initialise *household utility function* with postulated parameters $\hat{U}(S_h^t)$
for $t = 0, 1, 2, \ldots$ **do**
  Choose individual $n^{\mathrm{ref},h}$ as reference
  for $n = n^{\mathrm{ref},h}$ **do**
    Choose operator $\psi \in \Psi$ with probability P_ψ
    $S^*_{n^{\mathrm{ref},h}}, q(S^*_{n^{\mathrm{ref},h}}|S^t_{n^{\mathrm{ref},h}}), q(S^t_{n^{\mathrm{ref},h}}|S^*_{n^{\mathrm{ref},h}}) \leftarrow \textsc{ApplyChange}(\psi, S^t_{n^{\mathrm{ref},h}})$
    function $\textsc{ApplyChange}(\psi, S_n^t)$
      return new state S'_n, transition probability $q(S'_n|S^t_n)$, backward probability $q(S^t_n|S'_n)$
    end function
    Check $S^*_{n^{\mathrm{ref},h}}$ feasibility in terms of continuity (no gaps in time or space)
    for $n \in \mathcal{N}_h \setminus \{n^{\mathrm{ref},h}\}$ **do**
      repeat
        Choose operator ψ with probability P_ψ
        $S^*_n, q(S^*_n|S^t_n), q(S^t_n|S^*_n) \leftarrow \textsc{ApplyChange}(\psi, S^t_n)$
      until S^*_n is feasible and compliant with $S^*_{n^{\mathrm{ref},h}}$
      return new state S^*_n, transition probability $q(S^*_n|S^t_n)$, backward probability $q(S^t_n|S^*_n)$
    end for
  end for
  Compute target weight $p(S^*_h) = \exp\left(\hat{U}(S^*_h)\right)$
  Compute acceptance probability $\alpha(S^t_h, S^*_h) = \min\left(1, \frac{p(S^*_h)q(S^t_h|S^*_h)}{p(S^t_h)q(S^*_h|S^t_h)}\right) = \min\left(1, \frac{p(S^*_h)\, q\left(S^t_{n^{\mathrm{ref},h}}|S^*_{n^{\mathrm{ref},h}}\right) \prod_{n \neq n^{\mathrm{ref},h}} q\left(S^t_n|S^*_n, S^t_{n^{\mathrm{ref},h}}\right)}{p(S^t_h)\, q\left(S^*_{n^{\mathrm{ref},h}}|S^t_{n^{\mathrm{ref},h}}\right) \prod_{n \neq n^{\mathrm{ref},h}} q\left(S^*_n|S^t_n, S^*_{n^{\mathrm{ref},h}}\right)}\right)$
  With probability $\alpha(S^t_h, S^*_h)$, set $S^{t+1}_h \leftarrow S^*_h$; else $S^{t+1}_h \leftarrow S^t_h$
end for
return $\tilde{\mathcal{C}}_h$: Ensemble containing sets of schedules for all N_h individuals in household h.

estimation of utility parameters in household ABMs, where enumeration is infeasible and realistic schedule variation is essential for capturing trade-offs across multiple dimensions.

The key new notation introduced in the household choice set generation are summarised in Table 4.

4 Case study and evidence

To demonstrate the applicability of our proposed framework, we conduct a case study using a real-world dataset. In this section, we introduce the data used in empirical investigation and present illustrative evidence of intra-household interactions observed in the data.

Data from the 2018 UK NTS (Department for Transport 2024) is used to apply the estimation framework to a real-world case study. The NTS is a rolling annual household survey of personal travel by residents of England travelling within the UK. It contains information on socio-economic characteristics of individuals and their household, collected through face-to-face interviews, as well as detailed daily trips collected through a travel diary over multiple days (up to 7 days). The 2018 survey data contains 13’944 individuals, belonging to 5’896 households, with a total of 70’341 reported daily travel diaries.

This is the first step in addressing a complex problem, approached through a mathematically rigorous framework. In this paper we focus on a sample of the dataset, applying our approach on two-adult-member households.

Table 4: Notation for household choice set generation

Notation	Name	Description
S_n^t	Individual state	State (schedule) of individual n in household h at step t.
S_n^*	Neighbouring individual state	A schedule that can be reached in one step by applying an operator to the current schedule of individual n^h.
S_h^t	Household state	State of household h at step t, which is the household schedule comprised of a cluster of schedules of its individual members; $\left[S_1^t, \ldots, S_{N_h}^t\right]$.
S_h^*	Neighbouring household state	A household state (cluster of schedules of household individuals) that can be reached in one step by applying operators to the current state of its individual members.
$\psi \in \Psi$	Operator	A heuristic that modifies specific aspects of the schedule (i.e., time, space, participation, or activity participation mode (solo, joint)). Ψ is the set of possible operators.
$n^{\text{ref},h}$	Reference individual index	Index for the reference individual in household h, chosen as the benchmark in household choice set generation to compare the schedule of other individuals in the household for compliancy with household constraints.
$N_{\text{operators}}$	Number of operators	Number of implemented operators to modify the schedules.
P_ψ	Operator selection probability	Probability to select operator ψ.
δ	Time block	Each schedule is discretised into blocks of duration δ.
N_{iter}	Number of iterations	Number of random walk iterations in the Metropolis-Hastings algorithm.
N_{burn}	Number of initial burns	Number of initial samples to discard (burn-in period) to reduce the effect of initial conditions in Metropolis-Hastings algorithm.
N_{skip}	Number of skips	Number of samples to skip between accepted samples to reduce autocorrelation (thinning) in Metropolis-Hastings algorithm.
N_{alt}	Number of alternatives	Number of alternatives in the generated choice set.

This concept can be extended to households with three or more members, but this is not implemented here. Figure 2 presents the distribution of household size and structure in the 2018 UK NTS dataset. Two-person households represent the majority, accounting for 39% of all households and 54% of multi-member households. Among two-person households, the majority (95%) are composed of two adults.

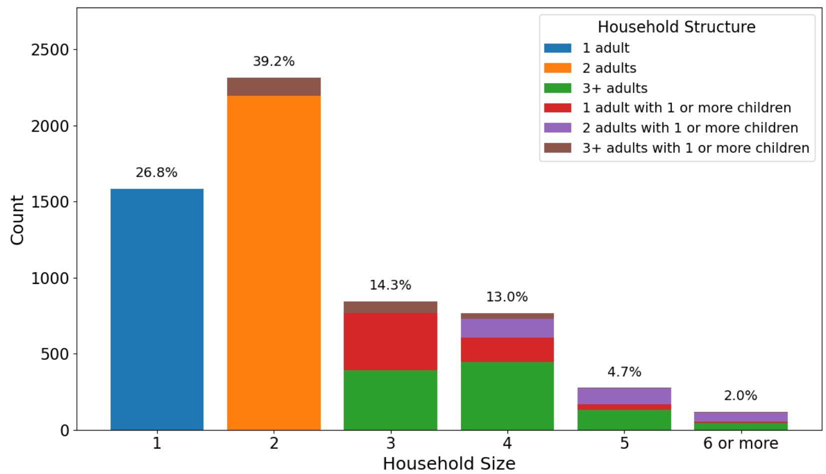


Figure 2: Distribution of household size and structure in the 2018 UK NTS dataset

We clean the sampled data to ensure that all individuals within the surveyed households have reported their travel diaries for the corresponding days. The data is processed to convert trip diaries into daily activity schedules. Data points with missing information are excluded. For this case study, we use a cleaned sample of 3'834 individuals from the 2018 survey, belonging to 1'917 two-adult households, with no missing variables. The sample includes 13'604 daily diaries, providing multiple observations per individual and household across multiple days.

Activities are grouped into six categories: Home, Work, Education, Leisure, Shopping, and Personal business (e.g., eating/drinking, using services such as medical appointments). The mode of start times and durations for each activity, computed from the distribution observed among two-adult households, are used as indicators of desired start and duration times in the model. These values are tailored for individuals based on their employment status (Full-time, Part-time, Not working) and are derived from the surveyed data. Table 5 summarises these scheduling preferences. The scheduling preferences are assumed to be homogeneous across individuals with the same employment type.

Table 5: Scheduling preferences tailored to employment status

Activity	**Full-time**		**Part-time**		**Not-working**	
	Start time	Duration	Start time	Duration	Start time	Duration
Work	09:00	08:30	09:00	04:00	10:00	00:05
Education	10:10	00:30	10:15	00:45	11:20	01:00
Leisure_solo	18:00	01:55	10:15	01:50	10:00	02:00
Leisure_joint	18:30	02:00	11:10	02:00	11:30	01:58
Shopping	11:00	00:30	14:30	00:30	11:15	00:30
Personal business	09:00	00:50	09:50	00:55	10:30	01:00

4.1 Identification of joint trips and activities

In order to identify joint trips in the NTS dataset, a set of rules inspired by Ho and Mulley 2013 is defined to extract individuals within the same household who travel together.

Joint trips can be either fully- or partially- joint. A fully joint trip is one in which all participants travel together from origin to destination, whereas a partially joint trip includes only certain segments where participants travel together, with individuals joining or separating at intermediate locations. The criteria for identifying a fully-joint trip are as follows: (i) the individuals belong to the same household; (ii) the trip occurs on the same day for all individuals; (iii) the trip start time is identical; (iv) the duration of the trip is identical; (v) the purpose of the trip is the same; (vi) the members involved in joint trip travel in the same mode; (vii) the origin of the trip is identical; and (viii) the destination of the trip is identical. Trips made for the purpose of escorting another household member are not classified as joint trips but are instead treated as escort trips. Escort trips refer to instances where an individual merely accompanies another person. Therefore, sharing the same travel purpose and thus activity is not implied. The same identification mechanism used for joint trips can be applied to derive escort trips.

Activity participation modes (solo/joint) are extracted from the data, using a set of rules designed to identify joint participation within households. As with trips, joint activities can be considered either fully- or partially-joint. A fully joint activity is performed together by all participants from start to finish, whereas a partially joint activity involves only partial overlap, with individuals joining or leaving at different times. The identification procedure for joint activities follows a similar approach to that used for joint trips. An activity is considered fully-joint with another household member if the following conditions are met: (i) the purpose of the activities are the same; (ii) the start time and duration of the activities are identical; (iii) the activities occur at the same location; (iv) the individuals belong to the same household; and (v) the activities take place on the same day.

Figure 3 presents the distribution of activity participation modes across different activity types, segmented by whether the activity was performed alone, jointly with another household member but travelled alone, or jointly with joint travel. Based on the 2018 UK NTS diaries for two-adult households, 59% of out-of-home activities are performed jointly with another household member, illustrating an example of intra-household interactions. Leisure accounts for the biggest proportion (44%) in joint out-of-home activities, with 51% of all leisure activities performed jointly. However, only 8% of trips to joint activities are made jointly. This suggests that while leisure is highly social, people often travel separately to meet for these activities. Overall, joint trips represent a small share compared to trips made individually to joint activities, implying that coordination in shared travel is relatively rare even when the activity is shared. These differences highlight how the social nature of an activity influences both participation and travel behaviour.

Interestingly, as observed from the analysis of data, 60% of education activities are done jointly with another household member, which may appear counterintuitive. According to the data, this is primarily driven by retired individuals attending informal or extracurricular classes together, such as painting, dance, or music lessons. In addition, it is partly explained by students living together in shared flats and pursuing the same degree. As expected, work activities are almost exclusively performed alone, reflecting the individualised nature of work routines. Moreover, activities done alone rarely involve joint trips, as travelling jointly for an activity that is done alone is uncommon.

While education and shopping activities also exhibit high rates of joint participation, leisure stands out both in its relative proportion and in its overall contribution to joint activities. As shown in Figure 3, leisure constitutes the largest volume of jointly performed activities across all activity types. Because incorporating joint participation specifications significantly increases model complexity, we focus solely on leisure activities. To reflect this behavioural significance, the estimation and scheduling framework, explicitly allows leisure to be scheduled as a joint or solo activity. This ensures that the model captures the flexibility observed in real-world leisure participation.

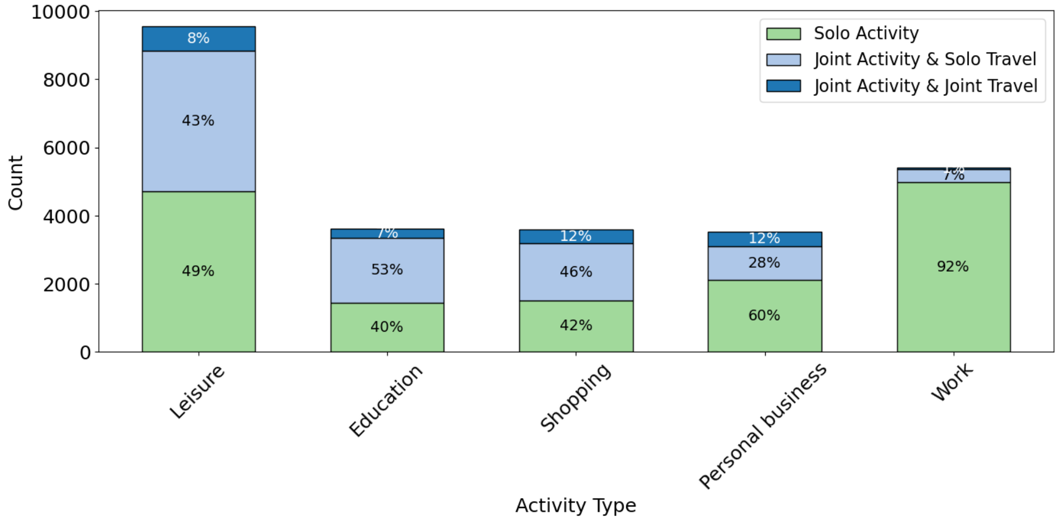


Figure 3: Distribution of activity types, segmented by trip and activity participation mode (solo vs. joint), based on the reported diaries of two-adult households in the 2018 UK NTS dataset.

5 Model estimation

We apply our proposed framework to the data sample from the NTS described in Section 4. We begin by generating a choice set for each household in the sample using the household-level choice set generation method described in Section 3.4. These generated choice sets are then used to estimate the utility parameters of the household scheduling model for the sample.

5.1 Generated choice set

First, the sample data is split into two groups: 80% for the training set and 20% for the test set. The training set is used for choice set generation and parameter estimation, while the test set is reserved for out-of-sample validation, where schedules are simulated using the household-level scheduling model.

For each household, a choice set consisting of 10 alternatives, including the observed household activity schedule, is generated. The experimental set up of the random walk is as follows: (i) The ensemble of observed schedules of household members is used as the initial state of the random walk. (ii) A set of 7 operators (Partic_mode, Block, Assign, Anchor, Swap, Inflate/Deflate, and MetaOperator that combines the actions of two or more operators) are implemented to modify the schedules to generate new states in the random walk. Each operator has equal probability of being chosen, denoted as P_ψ. The target distribution of the random walk is proportional to the exponential of household utility function (Equation 2), with parameters postulated based on literature. (iii) Each alternative in the random generated choice set is constructed by sequentially assigning random activities, modes, participation modes, locations, and durations within the daily time budget, with start times adjusted according to travel and activity sequences. (iv) All household members participate in the household joint decision-making process with the same weight ($w_n = 1/N_h$).

We run $N_{\text{iter}} = 2000$ iterations of the algorithm on the train sample. The initial $N_{\text{warm-up}}$ iterations serve as a warm-up period to stabilise the distribution from which the choice set is sampled. The accepted schedules are sampled after a warm-up period of $N_{\text{warm-up}} = 50$ iterations.

5.2 Parameter estimation: Model specifications and results

Using the generated choice set, the household scheduling model is estimated on the training sample. For identification purposes, "Home" is used as reference. Home is interpreted as absence of activity in this study due to absence of information on in-home activities in the dataset, which can be relaxed with richer data containing in-home activities such as time-use surveys. The magnitudes and signs of the other constants are relative to the baseline behaviour which is staying at home. As precise location information is not available in the data, travel parameters are not estimated. Estimating such parameters would require detailed location and network data to compute travel attributes for both chosen and unchosen alternatives. Therefore, the estimation focuses exclusively on activity scheduling parameters.

The models are estimated with PandasBiogeme (Michel Bierlaire 2020). Since the choice set includes multiple observations per household, a panel specification has been considered. Given the panel structure of the sampled choice set, serial correlation is expected, as error terms associated with observations from the same household are likely to share unobserved factors. To account for this, a panel effect is incorporated into the model estimation.

Different model specifications are tested, varying in the structure of the utility function. We present three specifications: a model including only activity- and scheduling-specific attributes (Section 5.2.1), and two models extended with socio-demographic characteristics (Section 5.2.2 and 5.2.3). As a benchmark, we present model estimations based on individual-level model in Section 5.2.4.

5.2.1 Model 1: household utility with activity-specific parameters

In this specification, the attributes used in the model are related to the activity-specific constants and parameters, as well as scheduling deviation penalties. For each alternative, the household utility function is defined as Equation 2. In this case-study w_n is set to $1/N_h$ for all individuals in the household, indicating identical relative influence for household members.

For each individual n, the utility function for each schedule alternative is defined as follows:

$$U(S_n) = \sum_{a_i^n \in \mathcal{A}^n} U_{a_i^n} \tag{34}$$

where $U(S_n)$ is the utility associated with schedule alternative S_n for individual n. $U(S_n)$ is made up of utility components linked to the performed activities ($U_{a_i^n}$). The activity-specific utility function for each activity a_i^n of individual n is defined as follows:

$$\begin{aligned} U_{a_i^n}^{\text{act. sp.}} = \gamma_{a_i} \quad &+ \quad \theta_{a_i}^{\text{early}} \max(0, x_{a_i^n}^* - x_{a_i^n}) \quad + \quad \theta_{a_i}^{\text{late}} \max(0, x_{a_i^n} - x_{a_i^n}^*) \\ &+ \quad \theta_{a_i}^{\text{short}} \max(0, \tau_{a_i^n}^* - \tau_{a_i^n}) \quad + \quad \theta_{a_i^n}^{\text{long}} \max(0, \tau_{a_i^n} - \tau_{a_i^n}^*) \quad + \quad \theta_{a_i}^{\text{jnt}} p_{a_i^n} + \quad \epsilon_n \end{aligned} \tag{35}$$

where γ_{a_i} is the activity-specific constants, $\theta_{a_i}^{\text{early}}$ and $\theta_{a_i}^{\text{late}}$ are start time penalty parameters for deviations from preference, $\theta_{a_i}^{\text{short}}$ and $\theta_{a_i}^{\text{long}}$ are duration penalty parameters for deviations from preference. $x_{a_i^n}$ is start time of activity a_i^n. $x_{a_i^n}^*$ is preferred start time for activity a_i^n. $\tau_{a_i^n}$ and $\tau_{a_i^n}^*$ are duration and preferred duration of activity a_i^n, respectively. $\theta_{a_i}^{\text{jnt}}$ is joint activity participation parameter for activity a_i, capturing the (dis)utility of joint activity engagement. $p_{a_i^n}$ is the participation mode of activity a_i^n, which is 1 if the individual performs the activity jointly with other member(s), and 0 otherwise. ϵ_n is an error term capturing unobserved variables for individual n.

Table 6 summarises the estimation results for model with activity-specific parameters on the train sample. Home activity is set as a reference, thus magnitudes and signs of coefficients are relative to the home baseline.

Table 6: Estimation results for model with activity-specific parameters on train sample

Name	Value	Rob. Std err	Rob. t-test	Rob. p-value
Education:constant	6.99	0.642	10.9	0
Education:early	-0.988	0.48	-2.06	0.0396
Education:late	-0.553	0.135	-4.1	4.06e-05
Education:long	-0.933	0.242	-3.86	0.000114
Education:short	-6.57	3.02	-2.17	0.0299
Leisure:constant	5.73	0.584	9.8	0
Leisure:early	-0.318	0.0261	-12.2	0
Leisure:joint_partic	0.866	0.15	5.78	7.65e-09
Leisure:late	-0.644	0.052	-12.4	0
Leisure:long	-0.493	0.0287	-17.2	0
Leisure:short	-1.05	0.164	-6.43	1.28e-10
Personal business:constant	5.08	0.471	10.8	0
Personal business:early	-0.553	0.0472	-11.7	0
Personal business:late	-0.595	0.0553	-10.8	0
Personal business:long	-0.434	0.0403	-10.8	0
Personal business:short	-0.782	1.16	-0.672	0.502
Shopping:constant	8.39	0.595	14.1	0
Shopping:early	-0.41	0.0268	-15.3	0
Shopping:late	-0.41	0.0388	-10.6	0
Shopping:long	-0.842	0.0775	-10.9	0
Shopping:short	-3.32	0.939	-3.54	0.000401
Work:constant	17	1.42	12	0
Work:early	-0.661	0.186	-3.56	0.000374
Work:late	-0.789	0.179	-4.4	1.07e-05
Work:long	-1.04	0.176	-5.88	3.98e-09
Work:short	-0.513	0.143	-3.6	0.000322

Summary of statistics
Number of estimated parameters = 26
$L(0)$ = - 35834.54
$L(\hat{\beta})$ = -978.0281
AIC = 2008.056
BIC = 2208.243

The estimated parameters are behaviourally sensible. The activity-specific constants are all positive, indicating a baseline preference for doing an out-of-home activity rather than staying at home, all else being equal. Work activities bring the most utility per time unit followed by Shopping, Education, Leisure, and Personal business.

Most parameter estimates are significant at 95% confidence level. The penalty parameters have a negative sign, indicating a decrease in utility when activities deviate from preferred times or durations. Joint participation in leisure activities has a significant positive coefficient, underscoring the social value of joint leisure. We can observe from the estimation results that the parameter associated with shorter personal business activities is not statistically significant. This may be due to the small number of such instances in the choice set, leading to limited variation and reduced estimation precision. Additionally, the flexible nature of personal business activities may make them less likely to involve trade-offs in scheduling.

Shorter durations for shopping are penalised approximately four times more than long durations, suggesting a strong disutility associated with unmet shopping needs. These significant negative estimates likely reflect the importance of fulfilling both individual and household requirements, which impacts overall satisfaction and well-being. The improvement in log-likelihood relative to the null log-likelihood indicates that the estimated parameters offer a significantly better fit to the observed choices. Overall, the model demonstrates consistent parameter signs and high statistical significance, supporting its behavioural realism.

5.2.2 Model 2: interaction between household car ownership and activity participation modes

The previous model includes only variables that describe attributes of the alternatives (schedules), assuming a homogeneous population in which all individuals share the same taste parameters. However, it is reasonable to expect that individuals have different preferences. In the context of choice models, this implies that the value of the parameters may vary depending on the socio-economic characteristics of the decision-makers. Since socio-economic characteristics do not vary across alternatives, their role in the model is to account for taste heterogeneity. We now investigate the presence of such heterogeneity in the population.

We now introduce a model specification that captures taste heterogeneity in the utility of joint activity participation based on household car ownership. Car ownership is treated as a categorical variable with three mutually exclusive categories: (i) households with “No car" (reference category), (ii) households with “One car", and (iii) households with “Two or more cars". To account for differences in preferences toward joint participation across these groups, interaction terms between joint participation and car ownership dummies are included in the utility function for each activity. The utility specification for each activity a_i^n is as follows:

$$\begin{aligned}
U_{a_i^n}^{\text{car.}} = {} & \gamma_{a_i} + \theta_{a_i}^{\text{early}} \max(0, x_{a_i^n}^* - x_{a_i^n}) + \theta_{a_i}^{\text{late}} \max(0, x_{a_i^n} - x_{a_i^n}^*) \\
& + \theta_{a_i}^{\text{short}} \max(0, \tau_{a_i^n}^* - \tau_{a_i^n}) + \theta_{a_i}^{\text{long}} \max(0, \tau_{a_i^n} - \tau_{a_i^n}^*) \\
& + \theta_{a_i}^{\text{jnt},0} \cdot p_{a_i^n} + \theta_{a_i}^{\text{jnt},1} \cdot p_{a_i^n} \cdot \delta_n^{\text{1car}} + \theta_{a_i}^{\text{jnt},2} \cdot p_{a_i^n} \cdot \delta_n^{\text{2+ cars}} + \epsilon_n
\end{aligned}$$

where, $p_{a_i^n}$ is a binary indicator equal to 1 if activity a_i^n is jointly performed; δ_n^{1car} and $\delta_n^{\text{2+ cars}}$ are dummy variables indicating whether the household to which the individual belongs, owns one car, or two or more cars, respectively. The parameter $\theta_{a_i}^{\text{jnt},0}$ captures the base utility of joint participation for households without a car, while $\theta_{a_i}^{\text{jnt},1}$ and $\theta_{a_i}^{\text{jnt},2}$ capture the incremental effects for one-car and two-or-more-car households, respectively.

The estimated parameters are summarised in Table 7. Most parameter estimates are significant at 95% confidence level. The main joint participation coefficient for leisure activities is positive and highly significant, confirming the social utility of shared leisure. Joint participation in activities can be motivated by considerations such as (i) efficiency; which can be gained from time and/or money savings, (ii) altruism, which is a selfless regard in which an individual gains utility by benefiting someone other than oneself, and (iii) companionship. The interaction effects reveal notable heterogeneity based on household car ownership. The interaction terms reveal heterogeneity based on household car ownership: households with one car exhibit an additional positive and significant utility for joint participation, while those with two or more cars show a negative and significant interaction effect, suggesting lower marginal utility from joint leisure as mobility becomes more independent. This can be explained as access to multiple vehicles enables members to maintain independent schedules and reduces the need for coordinated activity participation.

All penalty parameters remain negative and significant, indicating disutility for deviating from preferred activity schedules. The short-duration parameter for personal business remains insignificant, possibly due to data sparsity in that category.

To assess whether the inclusion of interaction terms between joint participation and household car ownership improves model fit, we perform a likelihood ratio test comparing with Model 1 (Section 5.2.1). Comparing to Model 1, the results show that there is an improvement in the final log likelihood (from -978.0281 to -953.4198). Using a likelihood ratio test (with a test statistics of $49.22 > \chi^2_{0.95,2} = 5.991$), we can conclude that interaction terms significantly improve model fit at the 5% significance level. Household car ownership plays an important role in moderating preferences for joint activity participation.

Table 7: Estimation results for Model 2, with interaction between number of cars and joint activity participation on train sample

Name	Value	Rob. Std Err	Rob. t-test	Rob. p-value
Education:constant	6.37	0.729	8.74	0
Education:early	-1.1	0.632	-1.74	0.0824
Education:late	-0.39	0.131	-2.98	0.00287
Education:long	-0.761	0.267	-2.85	0.00433
Education:short	-4.95	3.64	-1.36	0.174
Leisure:constant	5.64	0.642	8.79	0
Leisure:early	-0.346	0.0249	-13.9	0
Leisure:joint_partic	1.64	0.186	8.79	0
Leisure:joint_partic_SingleCar	0.855	0.307	2.78	0.00537
Leisure:joint_partic_TwoOrMoreCar	-0.98	0.164	-5.96	2.49e-09
Leisure:late	-0.664	0.057	-11.6	0
Leisure:long	-0.496	0.0308	-16.1	0
Leisure:short	-0.971	0.177	-5.5	3.88e-08
Personal business:constant	5.19	0.522	9.94	0
Personal business:early	-0.574	0.0516	-11.1	0
Personal business:late	-0.619	0.0607	-10.2	0
Personal business:long	-0.457	0.0459	-9.95	0
Personal business:short	-0.821	1.21	-0.68	0.496
Shopping:constant	8.75	0.736	11.9	0
Shopping:early	-0.429	0.0302	-14.2	0
Shopping:late	-0.461	0.049	-9.4	0
Shopping:long	-0.871	0.081	-10.7	0
Shopping:short	-2.97	1.07	-2.78	0.00543
Work:constant	17.8	1.49	12	0
Work:early	-0.603	0.192	-3.13	0.00174
Work:late	-0.745	0.193	-3.86	0.000112
Work:long	-1.04	0.196	-5.28	1.32e-07
Work:short	-0.508	0.154	-3.3	0.000963

Summary of statistics
Number of estimated parameters = 28
$\mathcal{L}(0)$ = -35834.54
$\mathcal{L}(\hat{\beta})$ = -953.4198
AIC = 1962.84
BIC = 2178.425

5.2.3 Model 3: employment type and household car ownership interactions

To capture behavioural heterogeneity, the model is extended in two ways. First, we segment the activity-specific constants based on the individual's employment type. Employment type is treated as a categorical variable with three mutually exclusive categories: (i) Full-time (reference category), (ii) Part-time, and (iii) Not working. This segmentation allows the baseline utility of activities to vary across different employment status. Second, we include interaction terms between joint participation and household car ownership to reflect the role of mobility resources in participating in joint activities.

The utility specification for each activity a_i^n is as follows:

$$
\begin{aligned}
U_{a_i^n}^{\text{Empl. car.}} = &\gamma_{a_i}^{\text{Full}} + \gamma_{a_i}^{\text{Part}} \cdot \delta_n^{\text{Part}} + \gamma_{a_i}^{\text{NoWork}} \cdot \delta_n^{\text{NoWork}} \\
&+ \theta_{a_i}^{\text{early}} \cdot \max(0, x_{a_i^n}^* - x_{a_i^n}) + \theta_{a_i}^{\text{late}} \cdot \max(0, x_{a_i^n} - x_{a_i^n}^*) \\
&+ \theta_{a_i}^{\text{short}} \cdot \max(0, \tau_{a_i^n}^* - \tau_{a_i^n}) + \theta_{a_i}^{\text{long}} \cdot \max(0, \tau_{a_i^n} - \tau_{a_i^n}^*) \\
&+ \theta_{a_i}^{\text{jnt,0}} \cdot p_{a_i^n} + \theta_{a_i}^{\text{jnt,1}} \cdot p_{a_i^n} \cdot \delta_n^{\text{1car}} + \theta_{a_i}^{\text{jnt,2}} \cdot p_{a_i^n} \cdot \delta_n^{\text{2+cars}} + \epsilon_n
\end{aligned}
$$

Where δ_n^{Part} and δ_n^{NoWork} are dummy variables for employment type: Part-time and Not working. Full-time serves as the reference category. $\gamma_{a_i}^{\text{Full}}$, $\gamma_{a_i}^{\text{Part}}$, and $\gamma_{a_i}^{\text{NoWork}}$ are activity-specific constants for each employment group. δ_n^{1car} and $\delta_n^{\text{2+cars}}$ are dummy variables for household car ownership. $\theta_{a_i}^{\text{jnt,0}}$ captures the base utility of joint participation for households without a car. $\theta_{a_i}^{\text{jnt,1}}$ and $\theta_{a_i}^{\text{jnt,2}}$ represent incremental effects for households with one car and two or more cars, respectively.

Table 8 present the estimation results. The results highlight significant behavioural heterogeneity across both employment type and household car ownership. Individuals who are not working exhibit substantially higher baseline utilities for activities such as education, leisure, and personal business, reflecting their greater scheduling flexibility. The strong and significant positive coefficient for joint participation confirms the social incentive for joint leisure, while the large negative interaction with having two or more cars suggests that access to multiple vehicles reduces the likelihood or value of coordinating joint activities. Most scheduling penalty parameters remain negative and significant, reinforcing the disutility of deviating from preferred start times and durations.

To assess whether the inclusion of employment-type segmentation improves model fit, we perform a likelihood ratio test comparing it with Model 2 (Section 5.2.2). Comparing to Model 2, the results show an improvement in the final log-likelihood (from -953.4198 to -881.9286). Using a likelihood ratio test (with a test statistic of $142.98 > \chi^2_{0.95,10} = 18.31$), we conclude that the inclusion of employment-based segmentation significantly improves model fit at the 5% significance level. This suggests that employment status meaningfully influences baseline preferences for activity participation.

Among the estimation specifications, this model provides the best statistical fit to the data, as confirmed by likelihood ratio tests as well as the lowest AIC and BIC values, and is therefore selected for application in the household schedule simulations presented in the next section.

Table 8: Estimation results for Model 3, with segmentation based on employment type and considering interaction between number of cars and joint activity participation on train sample

Name	Value	Rob. Std err	Rob. t-test	Rob. p-value
Education:constant	6.22	1.02	6.09	1.16e-09
Education:constant_NOT_WORKING	10.3	1.17	8.78	0
Education:constant_PART_TIME	1.32	0.984	1.35	0.178
Education:early	-1.23	0.511	-2.42	0.0157
Education:late	-0.686	0.222	-3.09	0.002
Education:long	-0.688	0.0997	-6.89	5.43e-12
Education:short	-6.75	2.14	-3.15	0.00161
Leisure:constant	4.07	0.574	7.09	1.29e-12
Leisure:constant_NOT_WORKING	7.73	0.532	14.5	0
Leisure:constant_PART_TIME	11.5	3.38	3.39	0.000687
Leisure:early	-0.381	0.0302	-12.6	0
Leisure:joint_partic	3.76	0.235	16	0
Leisure:joint_partic_SingleCar	-0.296	0.423	-0.7	0.484
Leisure:joint_partic_TwoOrMoreCar	-2.87	0.327	-8.79	0
Leisure:late	-0.724	0.0696	-10.4	0
Leisure:long	-0.477	0.0273	-17.5	0
Leisure:short	-0.564	0.159	-3.54	0.0004
Personal business:constant	3.57	0.486	7.35	1.97e-13
Personal business:constant_NOT_WORKING	8.44	0.589	14.3	0
Personal business:constant_PART_TIME	5.71	0.838	6.81	9.49e-12
Personal business:early	-0.494	0.0339	-14.6	0
Personal business:late	-0.419	0.0323	-13	0
Personal business:long	-0.404	0.0366	-11	0
Personal business:short	-11.1	1.18	-9.39	0
Shopping:constant	10.7	0.636	16.9	0
Shopping:constant_NOT_WORKING	5.82	0.663	8.78	0
Shopping:constant_PART_TIME	-1.24	0.935	-1.33	0.185
Shopping:early	-0.476	0.0514	-9.26	0
Shopping:late	-0.425	0.0323	-13.2	0
Shopping:long	-0.838	0.0665	-12.6	0
Shopping:short	-1.35	1.18	-1.15	0.252
Work:constant	40.5	5.84	6.94	3.98e-12
Work:constant_NOT_WORKING	-2.65	1.34e-12	-1.98e+12	0
Work:constant_PART_TIME	-0.876	8.78e-08	-9.97e+06	0
Work:early	-3.42	1.24	-2.75	0.00589
Work:late	-2.21	0.579	-3.81	0.000138
Work:long	-2.43	0.228	-10.6	0
Work:short	-1.31	0.183	-7.13	1.02e-12

Summary of statistics
Number of estimated parameters = 38
$\mathcal{L}(0)$ = -35834.54
$\mathcal{L}(\hat{\beta})$ = -881.9286
AIC = 1839.857
BIC = 2132.437

5.2.4 Benchmark model

As a benchmark, we estimate model parameters based on individual-level model without accounting for household interactions or constraints. The estimated parameters for individual-level model are presented in Table 9.

Table 9: Estimation results for individual-level model on train sample

Name	**Value**	**Rob. Std err**	**Rob. t-test**	**Rob. p-value**
Education:constant	4.67	0.210	22.2	0
Education:early	-0.504	0.0341	-14.8	0
Education:late	-0.721	0.0472	-15.3	0
Education:long	-0.846	0.0821	-10.3	0
Education:short	-4.25	0.648	-6.56	5.28e-11
Leisure:constant	4.56	0.162	28.1	0
Leisure:early	-0.224	0.0187	-11.9	0
Leisure:late	-0.958	0.0588	-16.3	0
Leisure:long	-0.378	0.0186	-20.4	0
Leisure:short	-4.62	0.317	-14.6	0
Personal business:constant	3.30	0.165	20.0	0
Personal business:early	-1.05	0.0679	-15.4	0
Personal business:late	-0.329	0.0311	-10.6	0
Personal business:long	-0.487	0.0324	-15.0	0
Personal business:short	-5.44	1.79	-3.04	0.00235
Shopping:constant	4.51	0.181	25.0	0
Shopping:early	-0.882	0.0487	-18.1	0
Shopping:late	-0.554	0.0433	-12.8	0
Shopping:long	-0.544	0.0309	-17.6	0
Shopping:short	-12.8	1.73	-7.41	1.27e-13
Work:constant	5.98	0.174	34.4	0
Work:early	-0.923	0.0359	-25.7	0
Work:late	-0.613	0.0509	-12.0	0
Work:long	-0.840	0.0363	-23.2	0
Work:short	-0.699	0.0312	-22.4	0

Summary of statistics

Number of estimated parameters = 25

$\mathcal{L}(0)$ = -16901.05

$\mathcal{L}(\hat{\beta})$ = -3138.885

AIC = 6327.77

BIC = 6501.469

6 Simulation and results

Using the estimated parameters discussed in Section 5, household schedules are simulated for a sample drawn from the test dataset. The schedule simulation and results are presented in this section.

We simulate daily schedules for 260 individuals belonging to 130 households in the test sample. For each household in the sample, 20 realisations are drawn from the underlying schedule distribution. We present two simulation result examples: (i) Schedule simulation with household-level scheduling model, considering the intra-household interactions. (ii) Schedule simulation with individual-level scheduling model, not considering the intra-household interactions.

For the household scheduling model, the estimated parameters from Model 3 (Section 5.2.3) are used for schedule simulation. For the individual-level scheduling simulation, the estimated parameters reported as benchmark model (in Section 5.2.4) are used.

In this section, we illustrate the individual- and household- model results by comparing their schedule distributions and distributions of start times and durations results with the observed distribution from the dataset as well as an analysis of descriptive statistics between the results of the model at an aggregate level (Section 6.1). We then showcase the addition that the household model brings at the disaggregate level and considering intra-household interactions with an example in Section 6.2

6.1 Aggregated simulation results

First, we compare daily averages of out-of-home activity duration and proportion of scheduled activity duration. Tables 10 and 11 compare household and individual simulation model outputs with observed data in terms of the average duration of out-of-home activities and the overall share of time spent on various activity types among 260 individuals in the test set. The household model replicates observed activity durations more closely than the individual model, particularly for work and leisure. The individual model tends to overestimate leisure, shopping, personal business, and education durations, leading to an overestimation of total out-of-home time. Table 11 further confirms that the household model provides a better match to the empirical distribution of activity time shares.

Table 10: Average duration of out-of-home activities among 260 individuals in test sample.

Activity	**Data**	**Individual model**	**Household model**
Work	07:15	06:57	07:20
Education	00:40	00:55	00:31
Leisure	02:13	03:13	02:07
Personal business	00:53	01:15	01:28
Shopping	01:27	01:49	01:31
Total out-of-home	12:30	14:08	12:58

We then visualise and compare the time of day participation of scheduled activities between the household model, individual model, and data. Figure 4 presents the distribution of scheduled activities in the course of a day, in the sampled test data (Figure 4a), and resulting from the simulator framework using individual-level model (Figure 4b), and the household-level algorithm (Figure 4c) on the sampled test data. The distributions are for schedules including at least one activity out of home. The height of each bar represents the proportion of the sample that is participating in each activity at a given moment of time.

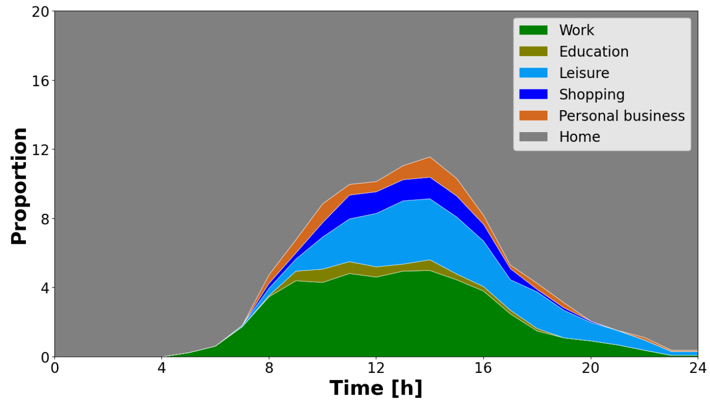
Table 11: Proportion of scheduled activities (based on duration of activities) among 260 individuals in test sample.

Activity	**Data**	**Individual model**	**Household model**
Work	10.60%	8.26%	9.63%
Education	0.51%	0.65%	0.64%
Leisure	5.01%	6.13%	5.30%
Personal business	0.92%	0.95%	0.88%
Shopping	1.44%	1.53%	1.40%
Home	81.52%	82.48%	82.15%

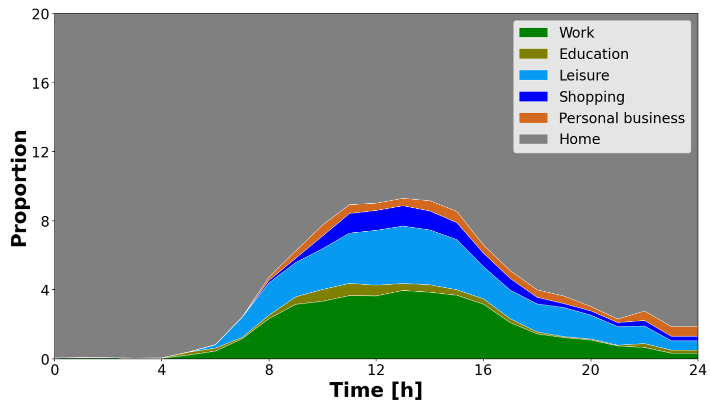
The household model provides a better fit to the observed time-of-day patterns. It more accurately captures the morning and afternoon peaks of out-of-home activities, including the sharper decline in activity participation after 18h. The isolated individual model tends to spread activities more uniformly throughout the day, overestimating leisure and shopping during late hours. The household model better aligns with the timing and intensity of activity transitions, particularly the early return to home, suggesting it more effectively captures the coordinated and constrained nature of real-world household scheduling behaviour.

We compare the distributions of simulated start times and durations, for each activity across the observed data, the individual model, and the household model. Figure 5 and Figure 6 present the distribution of start times and durations, respectively. In each subfigure, the Kolmogorov-Smirnov (KS) statistic compared to the observed data is included, quantifying the difference between simulated and empirical distributions. Across both start times and durations, the household model generally produces distributions that more closely align with the empirical data, reflected in lower KS statistics in most cases. This highlights that accounting for intra-household interactions can improve the model's ability to replicate observed temporal activity patterns more closely.

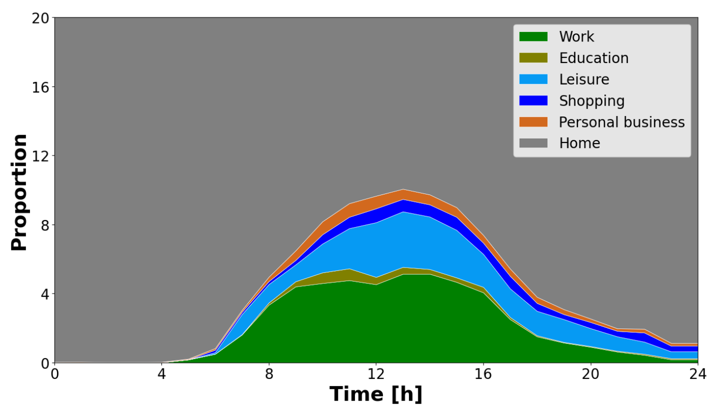
Figure 7 compares the joint distributions of activity start times and durations across the observed data, the household model, and the individual model. The household model outperforms the individual model, particularly in activities that involve shared timing or regular routines — such as being at home together, going to work, or engaging in leisure activities. Activities like personal business and leisure remain challenging for both model. This suggests that these activities are harder to predict or simulate well, possibly due to more variability in when or how they are scheduled.



(a) Data

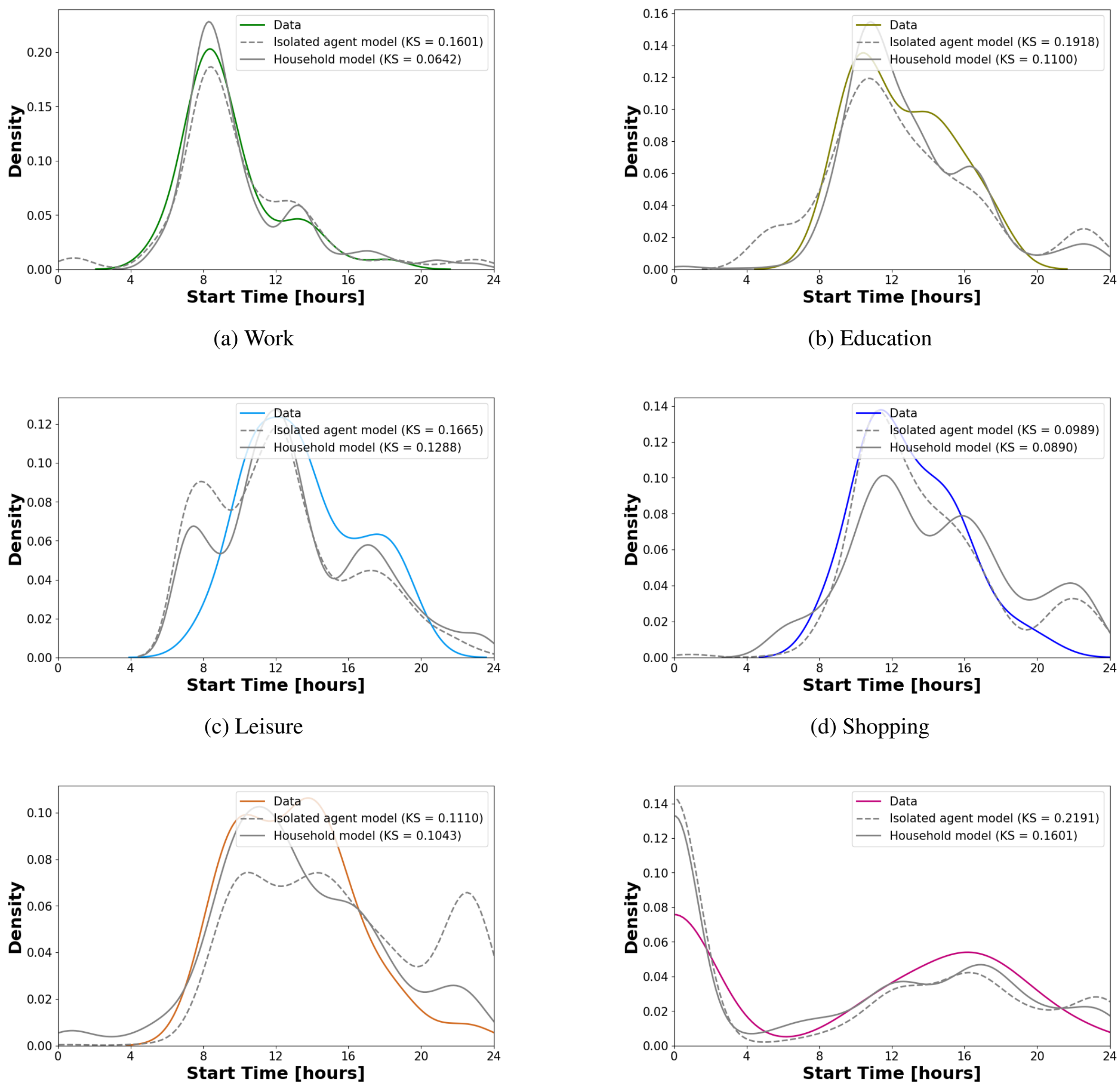


(b) Individual model



(c) Household model

Figure 4: Time of day activity participation, across 260 individuals in test sample



(a) Work

(b) Education

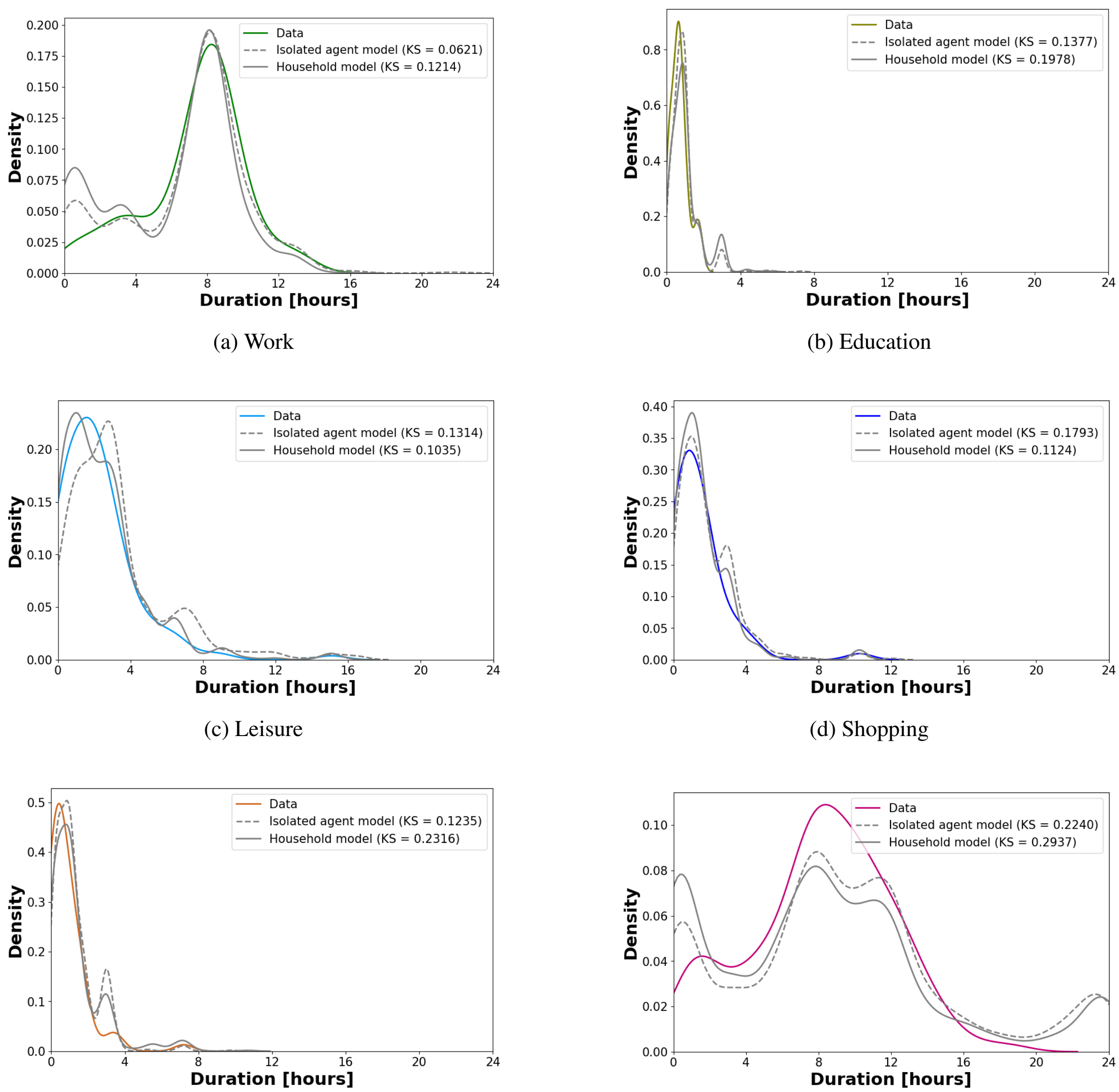
(c) Leisure

(d) Shopping

(e) Personal business

(f) Home

Figure 5: Simulated start times, per model and activity, across 260 individuals in test sample.



(a) Work

(b) Education

(c) Leisure

(d) Shopping

(e) Personal business

(f) Home

Figure 6: Simulated durations, per model and activity, across 260 individuals in test sample.

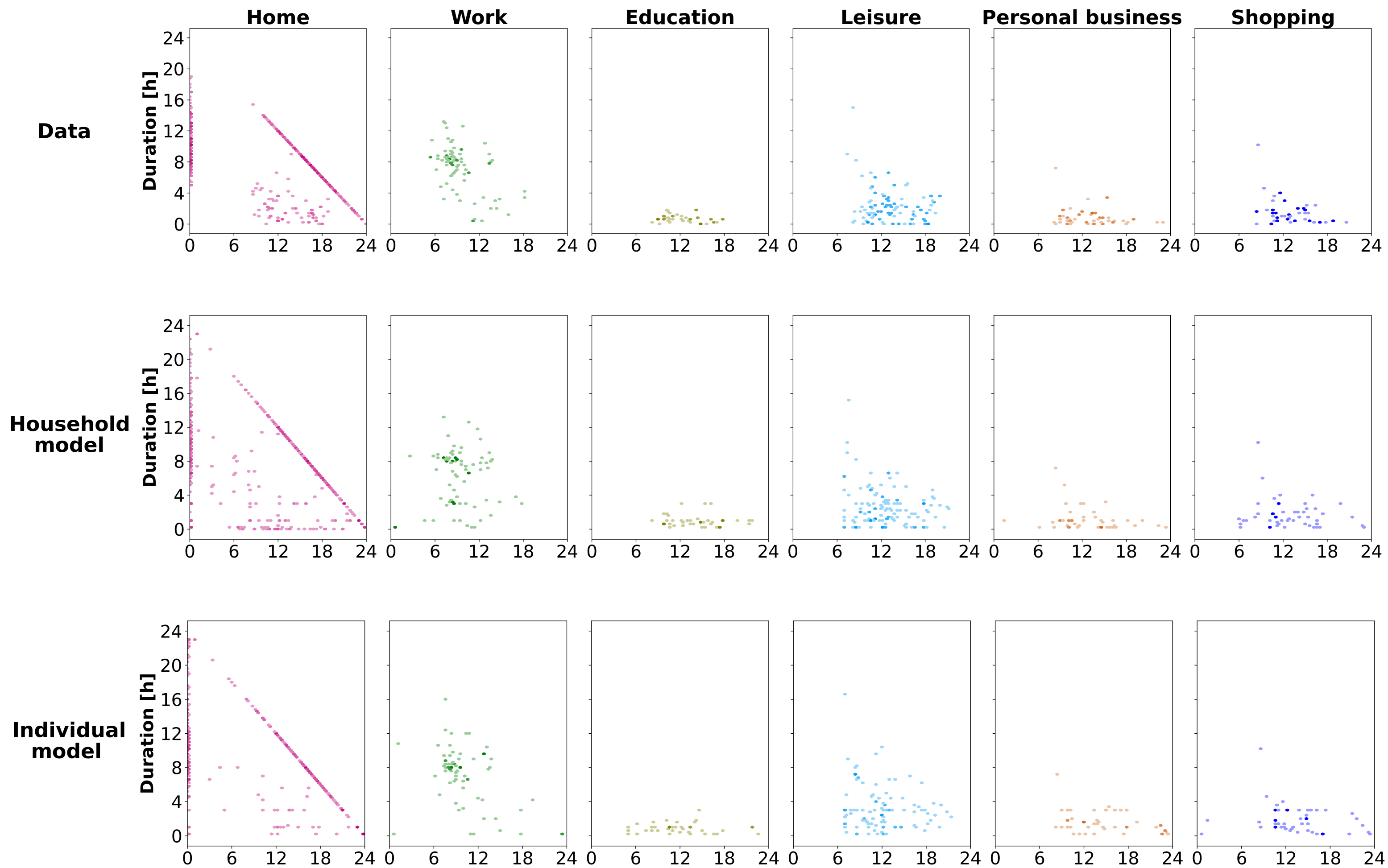


Figure 7: Joint density plots of start time and duration per model and activity, across 260 individuals in test sample.

6.2 Disaggregated simulation results

The explicit and simultaneous simulation of interactions and scheduling choice dimensions, ensure consistency of choices between individuals in the household. In this section, we illustrate, through an example, the synchronisation of household members' schedules generated using the household framework model.

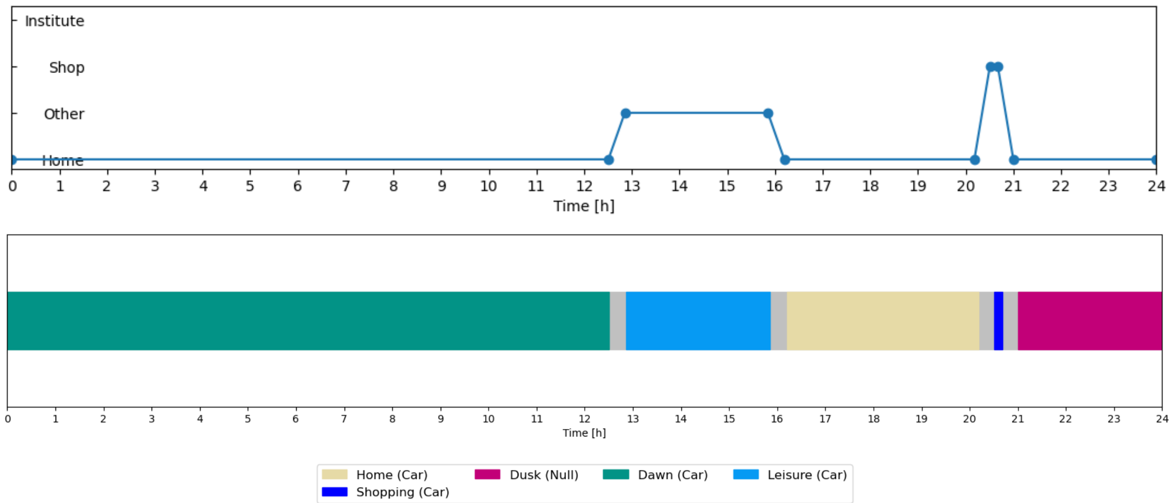
Consider an example of a two-person household with one car on a weekend. Figure 8 presents an arbitrarily selected realisation from the distribution of generated schedules using household-level model. The corresponding location sequence for the car and household members is shown as location plot in the same figure. This example showcases the advantage of the household simulation model in ensuring consistency across disaggregated results. Interactions within the household — such as car allocation among members, joint activity participation, and shared rides—are captured in the simulation outcomes. Examining the generated location sequence for the car (Figure 8c) alongside the simulated location sequences of the household members reveals that the location and mode choices of the household members are consistent with the availability and allocation of the household car.

The household model ensures compatibility of schedules for individuals in multi-member households by accounting for complex behaviours and interactions among household members. We compare descriptive statistics derived from the observed data with those obtained from household schedule simulations, which explicitly capture intra-household interactions and the synchronisation required for joint participation. In Table 12, the columns grouped under "Activity" present the proportion of leisure activity durations performed solo versus jointly with another household member, comparing the observed data with the simulated results from the household model. The columns grouped under "Trip" show the proportion of joint trips undertaken for leisure purposes across the observed data and the simulated results. The household model generates the participation modes for leisure activities with good alignment to the observed data, though it slightly overestimates joint activity participation and underestimates joint trips. Incorporating household-level interactions enables the model to capture social coordination in leisure activities. This highlights the behavioural relevance of modelling joint decisions explicitly in the household model.

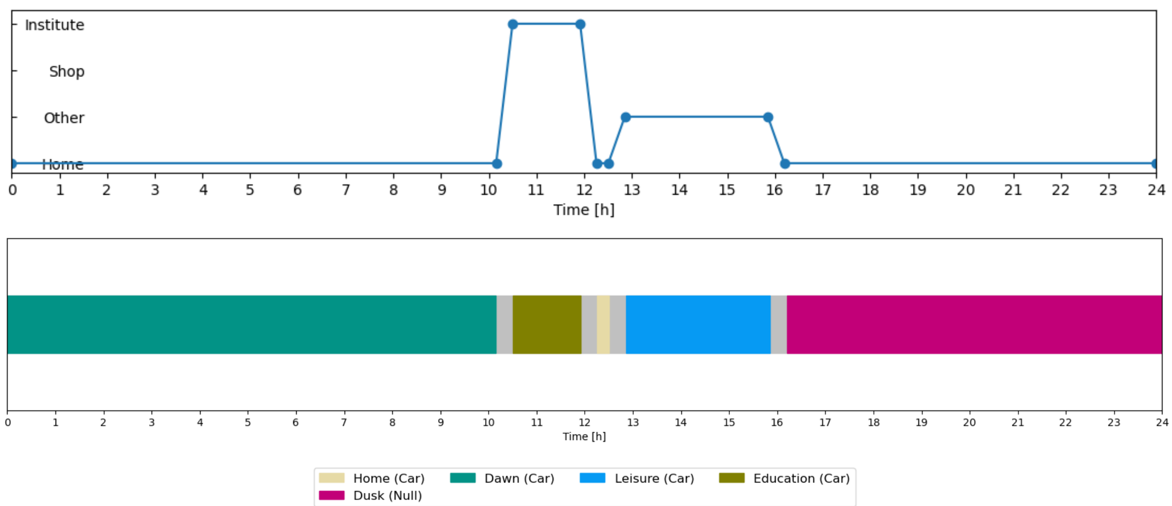
Table 12: Proportion of joint activity and trip for Leisure among 260 individuals in test sample.

	Activity		**Trip**	
	Data	**Household model**	**Data**	**Household model**
Solo	39.53%	33.34%	92.48%	93.34%
Joint	60.47%	66.66%	7.52%	6.66%

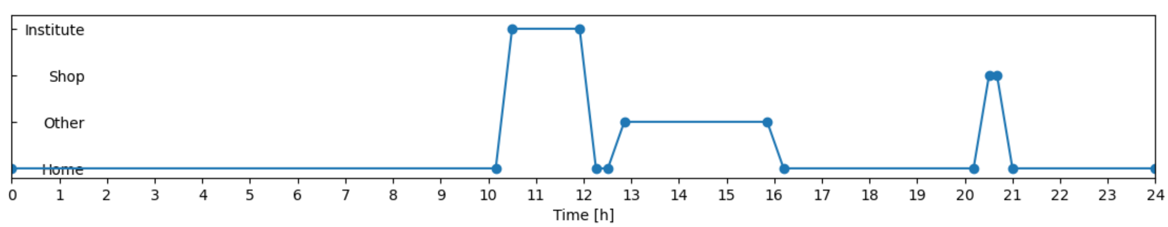
Figure 9 presents the distribution of start times and durations for solo and joint leisure activities, across 260 individuals in test sample. Overall, the household model effectively reproduces the underlying temporal distributions for solo and joint leisure activities. For solo leisure activities, the model captures the uni-modal pattern of start times centered around late morning and the sharp decline in durations beyond two hours. Figure 9b shows how the model is able to capture the multi-modal nature of the start time distribution for joint leisure activities. This multimodality occurs as, when performing leisure activities jointly, individuals have to deal with constraints for multiple participants. We speculate that the model identifies two common scheduling solutions avoiding standard work and education hours for joint leisure participation; either early morning before work or evening after work. In practice, society tends to favour evening leisure activities over early morning. However, as there is no explicit penalty representing this preference in the model, this results in early morning activities that do not happen in reality.



(a) Generated schedule and location sequence for household member 1

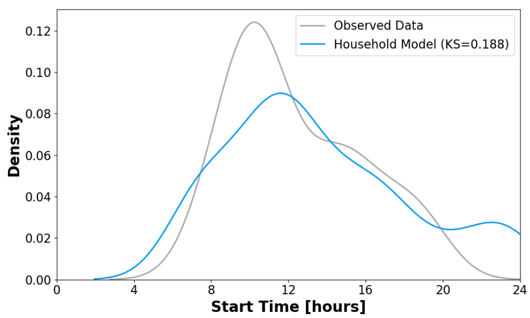


(b) Generated schedule and location sequence for household member 2

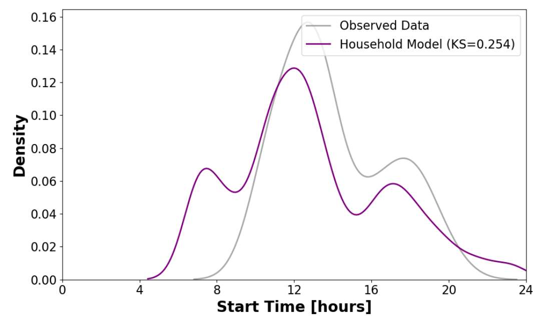


(c) Generated location sequence for the car

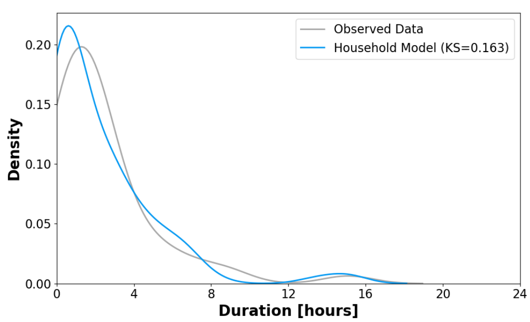
Figure 8: Generated schedules and location sequences of household members and the car in the example of family of 2



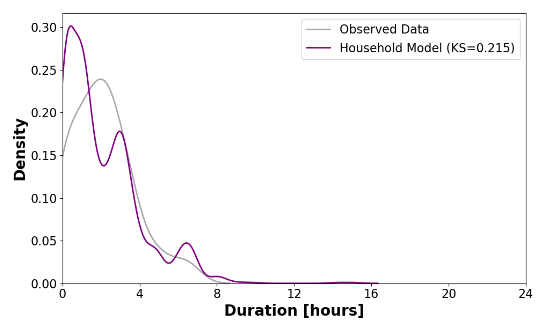
(a) Distribution of start time for **solo** leisure activity



(b) Distribution of start time for **joint** leisure activity



(c) Distribution of duration for **solo** leisure activity



(d) Distribution of duration for **joint** leisure activity

Figure 9: Distribution of start time and duration for solo and joint leisure activities, across 260 individuals in test sample.

7 Discussion and conclusions

7.1 Household-level vs individual-level choice set generation

In this section, we compare and discuss the household-level with individual-level choice set formation technique. Within-household interactions lead to additional complexities in the household scheduling. In the household-level choice-set generation technique, these aspects can be broadly classified as: (i) additional choice dimensions; activity participation mode; whether an individual participates in an activity solo or jointly with another household member, (ii) time arrangements; schedule synchronisation between participating individuals in joint activities, (iii) constraints; such as resource availability and limitation, (iv) group decision-making mechanism; moving from schedule utility of isolated individuals to household utility function, reflected in the MH algorithm through the target distribution and target weight of each candidate state (state = set of schedules of individuals in a household). Table 13 presents a summary comparison of household-level and individual-level choice set formation.

Table 13: Summary comparison of household-level vs individual-level choice-set formation

Feature	**Individual-level**	**Household-level**
Initialisation	Individual schedule	Schedule of all household members
Generation procedure	Separate	In parallel
Target weight	Schedule utility	Household utility
Decision variables	Specific to activities schedule and location	Specific to activities schedule, location and companionship
Constraints	Schedule continuity	Schedule continuity and household synchronisations
Operators	Specific to scheduling dimensions	Specific to companionship and scheduling dimensions
Output	Ensemble of schedules	Ensemble of sets of schedules

Choice-set generation technique for household scheduling, generates an ensemble of schedules with consistent alternatives for all household members, forming choice set of all individuals in a household in parallel. This ensures inter-agent validity of alternatives in the choice-set, enhancing model realism in capturing household dynamics. Whereas the relation between individuals and their household is lost in individual-level choice-set formations, leading to separate choice set formation procedures with no feedback between them.

Analysing the generated choice-set with the household-level algorithm, the frequency of leisure activities with activity participation type chosen as joint, is identical for both individuals in the household. This information is not captured in individual-level choice-set formation technique. The observed compatibility between the generated schedules in the choice-set, both through observations from randomly selected alternatives and also aggregated checks on the whole choice set, ensures the soundness of the algorithm logic.

As an empirical investigation of the added value of estimating parameters with the household-level model compared to the individual-level choice set generation model, we simulate schedules for a sample test data using parameter estimates from both algorithms, presented in Figure 4. From the results, the household-level model provide realistic results, closer to the observations in the data.

7.2 Conclusion and future work

This paper addresses two core research questions in household ABMs: (i) how to formulate choice sets that reflect intra-household dynamics, and (ii) how to estimate behaviourally meaningful parameters for household ABMs. We propose a framework that integrates household-level constraints and decision-making into choice set generation, and enables the estimation of a utility-based household scheduling model using a discrete choice modelling approach. Our household-level choice set generation builds on a MH based sampling algorithm to generate consistent and interdependent household schedules that capture complex joint behaviours.

The main characteristics of our household choice-set generation framework can be summarised as: (i) alternatives are generated in parallel for all household members to preserve inter-agent dependencies, (ii) we move from individual utility function to household utility function, (iii) dedicated operators are introduced to modify choice dimensions specific to household scheduling, (iv) generated schedules are constrained to satisfy both household-level and individual-level feasibility requirements; (v) the algorithm returns an ensemble containing clusters of schedules for individuals in household, and (vi) individual and household socio-demographic characteristics are preserved and reported in the generated choice sets, allowing for the inclusion of socio-demographic variables in model specifications.

The proposed framework is applied to a real-world case study. Utilising the choice set generation technique, the parameters of a utility-based ABM, the household-level Optimisation-based Activity Scheduling Integrating Simultaneous choice dimensions (OASIS) are estimated. The estimated scheduling preference parameters are well-identified and statistically significant, even with a relatively small number of alternatives in the choice set. Using the estimated parameters, the household schedules are simulated. This paper contributes to the state of the art in ABMs by explicitly simulating joint activities. The household approach yields more behaviourally realistic representations of daily activity scheduling, ensuring consistency of scheduling choices between household members and capturing joint activity synchronisation.

There are further extensions and improvements of the current work, suggesting avenues for future research. While the present specification captures observed heterogeneity through socio-demographic segmentation, it assumes homogeneity within each segment. Future research can explore alternative formulations of scheduling preferences such as incorporating unobserved continuous heterogeneity via random coefficients, or discrete heterogeneity through latent class models. The computational and implementation costs of different approaches should be considered. Another interesting extension to the choice set generation algorithm, is capturing inter-day correlations in scheduling behaviour, enabling multi-day analyses. Moreover, complex travel-related interaction dimensions within household members such as shared resource constraints (e.g., car availability) can be investigated in the choice set generation step. Estimating travel-related parameters within this framework will require access to detailed spatial and network data. Additionally, future research could explore strategies for determining the optimal number of alternatives to generate during the choice set generation step, balancing the trade-off between computational efficiency and the accuracy of parameter estimates.

References

Allahviranloo, Mahdieh, Robert Regue, and Will Recker (2014). “Pattern clustering and activity inference”. In: *Transp. Res. Board 93rd Annu. Meet.* Washington DC, United States.

Arentze, Theo A., D. Ettema, and Harry J.P. Timmermans (2010). “Incorporating time and income constraints in dynamic agent-based models of activity generation and time use: Approach and illustration”. In: *Transp. Res. Part C Emerg. Technol.* 18.1, pp. 71–83. ISSN: 0968090X. DOI: 10.1016/j.trc.2009.04.016.

Arentze, Theo A. and Harry J.P. Timmermans (2004). “A learning-based transportation oriented simulation system”. In: *Transp. Res. Part B Methodol.* 38.7, pp. 613–633. ISSN: 01912615. DOI: 10.1016/j.trb.2002.10.001.

— (2009). “A need-based model of multi-day, multi-person activity generation”. In: *Transp. Res. Part B Methodol.* 43.2, pp. 251–265. DOI: 10.1016/j.trb.2008.05.007.

Auld, Joshua and Abolfazl Mohammadian (2011). “Planning-constrained destination choice in activity-based model: Agent-based dynamic activity planning and travel scheduling”. In: *Transp. Res. Rec.* 1.2254, pp. 170–179. ISSN: 03611981. DOI: 10.3141/2254-18.

Barra, Tomas de la, B Perez, and J Anez (1993). “Multidimensional path search and assignment”. In: *Proc., 21st PTRC Summer Annu. Meet. PTRC Educ. Res. Serv. Ltd.* London.

Bekhor, Shlomo, Moshe E. Ben-Akiva, and M. Scott Ramming (2006). “Evaluation of choice set generation algorithms for route choice models”. In: *Ann. Oper. Res.* 144.1, pp. 235–247. ISSN: 02545330. DOI: 10.1007/s10479-006-0009-8.

Ben-Akiva, Moshe et al. (1984). “Modeling inter urban route choice behaviour”. In: *Proc. from Ninth Int. Symp. Transp. Traffic Theory*. Delft, Netherlands, pp. 299–330.

Ben-Akiva, Moshe E and Steven R Lerman (1985). *Discrete choice analysis : theory and application to travel demand*. MIT Press. ISBN: 9780262536400.

Bhat, Chandra R. (2005). “A multiple discrete-continuous extreme value model: Formulation and application to discretionary time-use decisions”. In: *Transp. Res. Part B Methodol.* 39.8, pp. 679–707. DOI: 10.1016/j.trb.2004.08.003.

Bhat, Chandra R. et al. (2013). “A household-level activity pattern generation model with an application for Southern California”. In: *Transportation (Amst).* 40.5, pp. 1063–1086. DOI: 10.1007/s11116-013-9452-y.

Bierlaire, Michel (2020). *A short introduction to PandasBiogeme*. Tech. rep.

Bovy, Piet H.L. (2009). “On modelling route choice sets in transportation networks: A synthesis”. In: *Transp. Rev.* 29.1, pp. 43–68. ISSN: 01441647. DOI: 10.1080/01441640802078673.

Bovy, Piet H.L. and Stella Fiorenzo-Catalano (2007). “Stochastic route choice set generation: Behavioral and probabilistic foundations”. In: *Transportmetrica* 3.3, pp. 173–189. ISSN: 18128602. DOI: 10.1080/18128600708685672.

Bowman, J. L. and M. E. Ben-Akiva (2001). “Activity-based disaggregate travel demand model system with activity schedules”. In: *Transp. Res. Part A Policy Pract.* 35.1, pp. 1–28. ISSN: 09658564. DOI: 10.1016/S0965-8564(99)00043-9.

Bradley, Mark and Peter Vovsha (2005). “A model for joint choice of daily activity pattern types of household members”. In: *Transportation (Amst).* 32.5, pp. 545–571. DOI: 10.1007/s11116-005-5761-0.

Chapin, S. (1974). *Human Activity Patterns in the City: Thing People Do in Time and Space*. Wiley, New York, USA.

Charypar, David and Kai Nagel (2005). “Generating complete all-day activity plans with genetic algorithms”. In: *Transportation (Amst).* 32.4, pp. 369–397. ISSN: 00494488. DOI: 10.1007/s11116-004-8287-y.

Chow, Joseph Y.J. and Will W. Recker (2012). “Inverse optimization with endogenous arrival time constraints to calibrate the household activity pattern problem”. In: *Transp. Res. Part B* 46.3, pp. 463–479. DOI: 10.1016/j.trb.2011.11.005.

Danalet, Antonin (2015). “Activity choice modeling for pedestrian facilities”. PhD thesis. EPFL. ISBN: 2000211410.

Danalet, Antonin and Michel Bierlaire (2015). “Importance sampling for activity path choice”. In: *15th Swiss Transp. Res. Conf.* Ascona, Switzerland.

Department for Transport (2024). *National Travel Survey, 2002-2023 [data collection]*. DOI: 10.5255/UKDA-SN-5340-14.
Flötteröd, Gunnar and Michel Bierlaire (2013). "Metropolis-Hastings sampling of paths". In: *Transp. Res. Part B Methodol.* 48, pp. 53–66. ISSN: 01912615. DOI: 10.1016/j.trb.2012.11.002.
Frejinger, E., M. Bierlaire, and M. Ben-Akiva (2009). "Sampling of alternatives for route choice modeling". In: *Transp. Res. Part B Methodol.* 43.10, pp. 984–994. ISSN: 01912615. DOI: 10.1016/j.trb.2009.03.001.
Guevara, C. Angelo and Moshe E. Ben-Akiva (2013). "Sampling of alternatives in Logit Mixture models". In: *Transp. Res. Part B Methodol.* 58, pp. 185–198. ISSN: 01912615. DOI: 10.1016/j.trb.2013.08.011.
Hagerstrand, T. (1970). "What about people in regional science?" In: *Reg. Sci. Assoc. Pap.* 24.1, pp. 6–21.
Hastings, W. K. (1970). "Monte carlo sampling methods using Markov chains and their applications". In: *Biometrika* 57.1, pp. 97–109. ISSN: 00063444. DOI: 10.1093/biomet/57.1.97.
Ho, Chinh and Corinne Mulley (2013). "Tour-based mode choice of joint household travel patterns on weekend and weekday". In: *Transportation (Amst).* 40.4, pp. 789–811. ISSN: 00494488. DOI: 10.1007/s11116-013-9479-0.
Hou, Huiqiao, Jacek Pawlak, and Aruna Sivakumar (2025). "Integrating activity-based transport and building occupancy models for campus-scale energy management". In: *Environ. Plan. B Urban Anal. City Sci.*, pp. 1–21. ISSN: 23998091. DOI: 10.1177/23998083251343715.
Kim, Changjoo and Olivier Parent (2016). "Modeling individual travel behaviors based on intra-household interactions". In: *Reg. Sci. Urban Econ.* 57, pp. 1–11. ISSN: 18792308. DOI: 10.1016/j.regsciurbeco.2015.12.002.
Manski, Charles F. (1977). "The Structure of Random Utility Models". In: *Theory Decis.* 8.3, pp. 229–254.
Miller, Eric J. and Morton E. O"Kelly (1983). "Estimating shopping destination choice models from travel diary data". In: *Prof. Geogr.* 35.4, pp. 440–449. ISSN: 14679272. DOI: 10.1111/j.0033-0124.1983.00440.x.
Nielsen, Otto Anker (2000). "A stochastic transit assignment model considering differences in passengers utility functions". In: *Transp. Res. Part B Methodol.* 34.5, pp. 377–402. ISSN: 01912615. DOI: 10.1016/S0191-2615(99)00029-6.
Nurul Habib, Khandker (2018). "A comprehensive utility-based system of activity-travel scheduling options modelling (CUSTOM) for worker's daily activity scheduling processes". In: *Transp. A Transp. Sci.* 14.4, pp. 292–315. ISSN: 23249943. DOI: 10.1080/23249935.2017.1385656.
Pagliara, Francesca and Harry J.P. Timmermans (2009). "Choice set generation in spatial contexts: A review". In: *Transp. Lett.* 1.3, pp. 181–196. ISSN: 19427875. DOI: 10.3328/TL.2009.01.03.181-196.
Paleti, Rajesh (2015). "Implicit choice set generation in discrete choice models: Application to household auto ownership decisions". In: *Transp. Res. Part B Methodol.* 80, pp. 132–149. ISSN: 01912615. DOI: 10.1016/j.trb.2015.06.015.
Pougala, Janody, Tim Hillel, and Michel Bierlaire (2022). "Capturing trade-offs between daily scheduling choices". In: *J. Choice Model.* 43. DOI: 10.1016/j.jocm.2022.100354.
— (2023). "OASIS: Optimisation-based Activity Scheduling with Integrated Simultaneous choice dimensions". In: *Transp. Res. Part C Emerg. Technol.* 155. ISSN: 0968090X. DOI: 10.1016/j.trc.2023.104291.
Prato, Carlo Giacomo and Shlomo Bekhor (2007). "Modeling route choice behavior: How relevant is the composition of choice set?" In: *Transp. Res. Rec.* 1, pp. 64–73. ISSN: 03611981. DOI: 10.3141/2003-09.
Recker, W. (1995). "The household activity pattern problem: General formulation and solution". In: *Transp. Res. Part B* 29.1, pp. 61–77. ISSN: 01912615. DOI: 10.1016/0191-2615(94)00023-S.
Recker, W., J. Duan, and H. Wang (2008). "Development of an estimation procedure for an activity-based travel demand model". In: *Comput. Civ. Infrastruct. Eng.* 23.7, pp. 483–501. ISSN: 10939687. DOI: 10.1111/j.1467-8667.2008.00555.x.
Regue, Robert, Mahdieh Allahviranloo, and Will Recker (2015). "Understanding Household Priorities when Scheduling Activities". In: *94th Annu. Meet. Transp. Res. Board*, p. 2992.
Rezvany, Negar, Michel Bierlaire, and Tim Hillel (2023). "Simulating intra-household interactions for in- and out-of-home activity scheduling". In: *Transp. Res. Part C Emerg. Technol.* 157. DOI: https://doi.org/10.1016/j.trc.2023.104362.
Rezvany, Negar, Tim Hillel, and Michel Bierlaire (2021). "Integrated models of transport and energy demand: A literature review and framework". In: *21st Swiss Transp. Res. Conf.*

Richardson, Anthony (1982). "Search models and choice set generation". In: *Transp. Res. Part A Gen.* 16.5-6, pp. 403–419. ISSN: 01912607. DOI: 10.1016/0191-2607(82)90068-1.

Robert B.Dial (1971). "A probabilistic multipath traffic assignment model which obviates path enumeration". In: *Transp. Res.* 5, pp. 83–111.

Roorda, Matthew, Eric J. Miller, and Nicolas Kruchten (2006). "Incorporating within-household interactions into mode choice model with genetic algorithm for parameter estimation". In: *Transp. Res. Rec.* 1, pp. 171–179. ISSN: 03611981. DOI: 10.3141/1985-19.

Scott, Darren M. and Sylvia Y. He (2012). "Modeling constrained destination choice for shopping: A GIS-based, time-geographic approach". In: *J. Transp. Geogr.* 23, pp. 60–71. ISSN: 09666923. DOI: 10.1016/j.jtrangeo.2012.03.021.

Shakeel, Farhan, Muhammad Adnan, and Tom Bellemans (2022). "Joint-Activities Generation among Household Members using a Latent Class Model". In: *Transp. Res. Procedia* 62.2021, pp. 557–564. ISSN: 23521465. DOI: 10.1016/j.trpro.2022.02.069.

Subbiah, Rajesh (2013). "An activity-based energy demand modeling framework for buildings: A bottom-up approach". PhD thesis. Virginia Polytechnic Institute and State University.

Termansen, Mette, Colin J. McClean, and Hans Skov-Petersen (2004). "Recreational site choice modelling using high-resolution spatial data". In: *Environ. Plan. A* 36.6, pp. 1085–1099. ISSN: 0308518X. DOI: 10.1068/a36164.

Timmermans, Harry J.P. and Junyi Zhang (2009). "Modeling household activity travel behavior: Examples of state of the art modeling approaches and research agenda". In: *Transp. Res. Part B Methodol.* 43.2, pp. 187–190. DOI: 10.1016/j.trb.2008.06.004.

Xu, Zhiheng, Jee Eun Kang, and Roger Chen (2018). "A random utility based estimation framework for the household activity pattern problem". In: *Transp. Res. Part A Policy Pract.* 114, pp. 321–337. ISSN: 09658564. DOI: 10.1016/j.tra.2018.01.036.

Zhang, Junyi and Akimasa Fujiwara (2006). "Representing household time allocation behavior by endogenously incorporating diverse intra-household interactions: A case study in the context of elderly couples". In: *Transp. Res. Part B Methodol.* 40.1, pp. 54–74. DOI: 10.1016/j.trb.2005.01.004.