Activity and Travel Pattern Scheduling with A Perturbed Utility Route Choice Model

Aoi Watanabe*[1], Mads Paulsen[1], Laurent Cazor[1], and Thomas Kjær Rasmussen[1]

[1]Department of Technology, Management and Economics, Technical University of Denmark, Denmark

Short summary

This paper formulates daily activity scheduling—decisions on when, where, and for what purpose to perform activities and travel—as a Perturbed Utility Route Choice (PURC) problem on a time–space network. Unlike conventional discrete choice models, the proposed framework represents behavior directly as link flows, corresponding to the number of individuals at each location and time. In this framework, optimal flows are sparse, allowing the model to endogenously restrict the effective consideration set without ex ante enumeration while capturing flexible substitution and complementarity patterns across schedules. To address the large size of time–space networks, we propose a column-generation-based estimation algorithm that iteratively expands partial networks. We further introduce a group-based perturbation structure to capture activity-specific correlations beyond those implied by network topology. An empirical application using the Danish National Travel Survey demonstrates that the model can be estimated on realistic networks and reproduces key activity–travel patterns.
Keywords: Activity-based modelling, Behavioral correlation, Column generation, Perturbed utility route choice, Time–space networks

1 Introduction

Modeling daily activity scheduling—how individuals decide when, where, and for what purpose to engage in activities over the course of a day, and how they travel between them—is central to transport and urban policy analysis. Unlike trip-based approaches, activity-based frameworks capture interdependencies across trips, endogenous time allocation, and constraints linking activities and travel within a unified behavioral structure.

Traditionally, activity scheduling has been modeled using hierarchical discrete choice frameworks, where activity type, location, and timing decisions are structured sequentially (Bowman & Ben-Akiva, 2001). While operationally tractable, such models rely on structural decomposition and face increasing difficulty as the spatial–temporal choice space expands. In particular, large activity spaces raise practical challenges related to choice set construction and enumeration.

An important advance is the formulation of activity scheduling as route choice on a time–space network, representing daily behavior as a path through a time–space network (Zimmermann et al., 2018). Mixed recursive logit extensions allow for taste heterogeneity, avoid explicit enumeration of complete daily schedules, and relax the independence of irrelevant alternatives at the unconditional level. However, correlation patterns in these models are determined by the specification of random utility components, and richer behavioral flexibility typically comes at the cost of increased computational complexity. Moreover, the recursive formulation effectively treats the full time–space network as the choice set, although individuals are likely to consider only a small subset of feasible activity schedules.

In route choice modeling, the recently proposed Perturbed Utility Route Choice (PURC) framework (Fosgerau et al., 2022) offers a new foundation. Rather than modeling discrete path choices, PURC represents link flows directly as the solution to a convex perturbed utility optimization problem. This formulation avoids explicit path enumeration, naturally accommodates zero flows on unused links, and generates flexible substitution and complementarity patterns induced directly by network structure, while enabling fast estimation via regression-based methods.

Applying PURC to activity scheduling is appealing but non-trivial. Two main challenges arise. First, time–space networks representing daily activity schedules are orders of magnitude larger than typical route choice networks, making estimation on the full network computationally infeasible.

Second, correlations between activity schedules are often driven by shared activity structures rather than network overlap alone, such as schedules involving similar activity types or sequences (e.g., work followed by an errand) even if the corresponding time–space paths do not share links.
To address these challenges, we propose two methodological components. First, we develop a reduced-cost-based column generation algorithm that enables scalable estimation by iteratively expanding a partial network until it contains the support of the optimal solution. Second, we introduce a group-based perturbation structure that captures correlations between schedules sharing similar activity structures.
By bringing activity scheduling into the PURC framework, the proposed approach enables scalable estimation on large time–space networks while capturing flexible behavioral correlations. Moreover, PURC's ability to represent zero flows naturally restricts the effective consideration set without requiring ex ante enumeration.

2 Methodology

Activity Scheduling as a PURC Problem

We represent daily activity scheduling on a time–space network $(\mathcal{V}, \mathcal{E})$. Each node corresponds to a pair (t, z) of discrete time t and spatial zone z. Directed links represent either movements between zones using a specific travel mode or staying at a zone to perform a specific activity. A complete daily activity schedule corresponds to a path from a start node to an end node in this network, typically from home in the morning to home at the end of the day.
Let $\boldsymbol{\beta}$ denote a vector of parameters, and define systematic link costs as $\boldsymbol{c} = Z\boldsymbol{\beta}$, where Z collects link-level attributes. Activity scheduling behavior in the PURC framework is then modeled as the solution to the convex optimization problem

$$\hat{\boldsymbol{x}}(Z\boldsymbol{\beta}, \boldsymbol{b}) = \arg\min_{\boldsymbol{x} \geq \boldsymbol{0}} \left\{ (Z\boldsymbol{\beta})^\top \boldsymbol{x} + F(\boldsymbol{x}) \;\middle|\; A\boldsymbol{x} = \boldsymbol{b} \right\}, \tag{1}$$

where $F(\boldsymbol{x})$ is a strictly convex perturbation function that introduces preference dispersion, allowing multiple schedules to carry positive flow rather than collapsing to a single minimum-cost path. Link flows correspond to quantities of direct behavioral interest, such as the number of individuals engaged in a given activity at a given zone, or traveling between zones using a specific mode, during a given time interval.

A Scalable Estimation Algorithm for Activity Scheduling

For each individual n, we observe $(\boldsymbol{y}_n, Z_n, \boldsymbol{b}_n)$, where $\boldsymbol{y}_n \in \{0,1\}^{|\mathcal{E}|}$ denotes the realized daily activity schedule. Estimation with individual-level data proceeds by iterating between PURC flow prediction and a closed-form parameter update based on the discrepancy between predicted and observed flows (Algorithm 1, Steps 1 and 3; see Fosgerau et al. (2024a) for details).
In activity scheduling applications, however, the underlying time–space networks are extremely large, making repeated solution of (1) on the full network computationally infeasible. To address this challenge, we start from a small initial network and perform estimation on iteratively expanded partial networks. Specifically, as parameter updates proceed, the network is gradually expanded by adding links adjacent to the current partial network whose inclusion is expected to improve the PURC objective (Algorithm 1, Step 2).

Theorem 1. *Consider the PURC problem defined on a full network $(\mathcal{V}, \mathcal{E})$ and a partial network $(\mathcal{V}^{(k)}, \mathcal{E}^{(k)})$ with $\mathcal{V}^{(k)} \subseteq \mathcal{V}$ and $\mathcal{E}^{(k)} \subseteq \mathcal{E}$. Let $(\boldsymbol{x}^{(k)}, \boldsymbol{\lambda}^{(k)})$ denote a primal–dual solution of the PURC problem on the partial network. For any link $e \in \mathcal{E} \setminus \mathcal{E}^{(k)}$ connecting nodes in $\mathcal{V}^{(k)}$, define the reduced cost*

$$\mathrm{rc}_e^{(k)} = c_e + \frac{\partial F(\boldsymbol{x})}{\partial x_e}\Big|_{x_e = 0} - \boldsymbol{a}_e^\top \boldsymbol{\lambda}^{(k)}.$$

Then $\mathrm{rc}_e^{(k)}$ equals the right derivative at zero of the optimal value function associated with relaxing the constraint $x_e = 0$ for link e.

As shown in Theorem 1, reduced costs admit a rigorous interpretation as marginal objective improvements. Accordingly, for links whose endpoints are already present in the partial network (hereafter referred to as internal link candidates), expansion is performed by adding those with negative reduced cost (Algorithm 1, Step 2a).

Exact reduced-cost evaluation requires well-defined dual node potentials and is therefore restricted to such internal link candidates. To explore regions involving nodes not yet included in the partial network (external nodes), we evaluate short patches that start and end at internal nodes while passing through external nodes. At the patch level, intermediate node potentials cancel out, so that the patch score depends only on link costs and the potential difference between its internal endpoints. This provides a reduced-cost-like criterion that can be evaluated using the currently available dual potentials (Algorithm 1, Step 2b).

Once an external node is incorporated through such a patch, it becomes part of the partial network and receives a valid dual potential in the subsequent PURC solve. Its adjacent links can then be evaluated individually using the exact reduced-cost criterion of Theorem 1.

The overall estimation algorithm is summarized in Algorithm 1. In practice, during the column generation step (Step 2), anchor nodes are identified based on the current node potentials and link flows. Internal candidate links and candidate patches involving external nodes are then generated from these anchors, allowing efficient network expansion with reduced computational cost.

This estimation strategy exploits a key property of the PURC framework: in the optimal solution, only a small subset of links carries positive flow, while the majority of links remain inactive. Consequently, once the partial network contains the support of the full-network optimal solution, further expansion becomes unnecessary, eliminating the need to enumerate the choice set ex ante. From a behavioral perspective, it is also plausible that individuals effectively consider only a limited subset of feasible transitions within the full time–space network.

Algorithm 1 Scalable Estimation Algorithm with Column Generation

Step 0. Choose starting parameter values $\boldsymbol{\beta}^{(0)}$ and initial partial networks $(\mathcal{V}_n^{(0)}, \mathcal{E}_n^{(0)})$.

Step 1. Given $\boldsymbol{\beta}^{(t)}$, solve the PURC problem on the current partial network to obtain $(\hat{\boldsymbol{x}}_n^{(t)}, \boldsymbol{\lambda}_n^{(t)})$:

$$\hat{\boldsymbol{x}}_n^{(t)} = \arg\min_{\boldsymbol{x}\geq \mathbf{0}} \{(Z_n \boldsymbol{\beta}^{(t)})^\top \boldsymbol{x} + F(\boldsymbol{x}) \mid A_n^{(t)} \boldsymbol{x} = \boldsymbol{b}_n\}.$$

Step 2. Define the set of anchor nodes $\mathcal{A}_n^{(t)} \subseteq \mathcal{V}_n^{(t)}$. For each anchor node $v \in \mathcal{A}_n^{(t)}$:

(a) For each feasible transition from v to an internal node $u \in \mathcal{V}_n^{(t)}$, compute the exact reduced cost of the corresponding link

$$\mathrm{rc}_{e,n}^{(t)} = c_{e,n} + F_e'(0) - \boldsymbol{a}_e^\top \boldsymbol{\lambda}_n^{(t)},$$

and add the link if $\mathrm{rc}_{e,n}^{(t)} < 0$.

(b) For each feasible transition from v to an external node $u \notin \mathcal{V}_n^{(t)}$, construct short paths of the form $v \to u \to \cdots \to w$ with $w \in \mathcal{V}_n^{(t)}$. For each such patch p, compute the patch score

$$\mathrm{score}_{p,n}^{(t)} = \sum_{e \in p} \mathrm{rc}_{e,n}^{(t)}$$

and add the patch if $\mathrm{score}_{p,n}^{(t)}$ is below a given threshold.

Step 3. Update parameters according to the closed-form update:

$$\boldsymbol{\beta}^{(t+1)} = -\left(\mathbb{E}_N\left[Z_n^\top \hat{P}_n^{(t)} Z_n\right]\right)^{-1} \mathbb{E}_N\left[Z_n^\top \hat{P}_n^{(t)} \left(\nabla F(\hat{\boldsymbol{x}}_n^{(t)}) + \nabla^2 F(\hat{\boldsymbol{x}}_n^{(t)})(\boldsymbol{y}_n - \hat{\boldsymbol{x}}_n^{(t)})\right)\right],$$

where $\hat{P}_n^{(t)} = \hat{B}_n^{(t)} - \left(A_n^{(t)} \hat{B}_n^{(t)}\right)^+ A_n^{(t)} \hat{B}_n^{(t)}$, $\hat{B}_n^{(t)} = \mathrm{diag}\left(\mathbf{1}_{\hat{\boldsymbol{x}}_n^{(t)} > 0}\right)$. $(\cdot)^+$ denotes the Moore–Penrose pseudoinverse.

Step 4. Repeat Steps 1–3 until convergence.

Activity-Specific Correlation via Group-Based Perturbation

In standard PURC models, correlation patterns arise endogenously from network topology. While this is a key strength of the framework, it is insufficient in activity scheduling settings where correlations are driven by activity semantics rather than by network structure. For example, participation in the same activity at different locations or times, or activity schedules that share

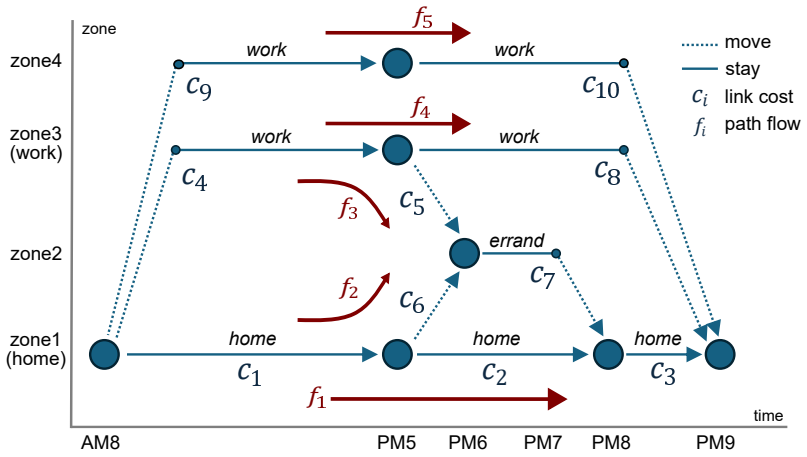


Figure 1: Toy time–space activity network.

the same activity pattern but are shifted in time, may exhibit strong correlations even though the corresponding paths are not overlapping in the time–space network.

To capture such activity-specific correlations that arise independently of network structure, we introduce a group-based perturbation structure. Let $\mathcal{G}$ denote a collection of groups of links corresponding to shared activity-related attributes, such as activity type, activity pattern, or travel mode. We specify the perturbation function as

$$F(\boldsymbol{x}) = \sum_{e \in \mathcal{E}} F_e(x_e) \; + \; \sum_{g \in \mathcal{G}} F_g\left(\sum_{e \in g} x_e\right),$$

where $F_e(\cdot)$ and $F_g(\cdot)$ are strictly convex functions.

This formulation induces correlation among links within the same group, while preserving convexity and tractability of the PURC problem. The proposed group-based perturbation enables the representation of activity-specific correlation patterns that are not implied by network structure alone, and it integrates naturally into the scalable estimation algorithm described in the previous section.

3 Results and Discussion

Toy network

We use a toy network to illustrate the key behavioral properties of the proposed framework through sensitivity analysis. In the PURC model, behavioral responses to changes in link costs can be evaluated analytically. In particular, the Jacobian of link flows with respect to link costs is given by $\frac{\partial \boldsymbol{x}}{\partial \boldsymbol{c}} = -\left(H^{-1} - H^{-1}A^\top (AH^{-1}A^\top)^{+} AH^{-1}\right)$, where $H = \nabla^2 F(\boldsymbol{x})$ (Fosgerau et al., 2024b).

We consider the simple time–space network shown in Figure 1. The network consists of 10 links and 5 representative paths. These paths can be interpreted as stylized daily schedules involving work activities and a mandatory evening errand (e.g., picking up a child). They include: staying at home all day (f_1); staying at home and performing the evening errand (f_2); commuting to work in the morning and performing the evening errand before returning home (f_3); commuting to work and returning home without performing the errand (f_4); and spending the day at a different work location (f_5).

When the perturbation function is quadratic, $F(\boldsymbol{x}) = \frac{1}{2}\boldsymbol{x}^\top \boldsymbol{x}$, we have $H = I$, and the sensitivity of path flows with respect to link costs can be obtained analytically:

$$\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{c}} = \frac{1}{110}$$

	c_1	c_2	c_3	c_4	c_5	c_6	c_7	c_8	c_9	c_{10}
f_1	−20	−50	−20	10	0	30	30	10	10	10
f_2	−21	30	1	16	22	−51	−29	−6	5	5
f_3	22	0	−22	−22	−44	22	−22	22	0	0
f_4	4	10	26	−24	22	−6	16	−46	20	20
f_5	15	10	15	20	0	5	5	20	−35	−35

Table 1: Trip-related cost parameters

Variable	Type	Car	PT	Bike	Walk	Common
Travel time (min)	Continuous	0.0013	0.0007	0.0006	0.0004	–
Monetary cost (DKK)	Continuous	–	–	–	–	0.0002
ASC	Dummy	–	–	0.0080	0.0067	–

The matrix illustrates both substitution and complementarity effects at the schedule level. Substitution is stronger toward paths with similar activity patterns: $\partial f_3/\partial c_1 > \partial f_4/\partial c_1$, since f_3 includes the obligatory evening errand whereas f_4 does not. Complementarity is also captured: $\partial f_4/\partial c_6 < 0$, even though path f_4 does not include link 6. When working at Zone 3, the availability of path f_3, which combines work with the evening errand, makes the work-only schedule f_4 relatively less attractive.

We further examine the effect of introducing a group-based perturbation that aggregates links corresponding to primary work-hour participation. Specifically, we consider $F(\boldsymbol{x}) = \frac{1}{2}\boldsymbol{x}^\top\boldsymbol{x} + \frac{\rho}{2}(x_4 + x_9)^2$, where $\rho > 0$ controls the strength of correlation. In this case, $H = I + \rho\boldsymbol{h}\boldsymbol{h}^\top$, with $\boldsymbol{h} = [0\,0\,0\,1\,0\,0\,0\,0\,1\,0]^\top$, which modifies the sensitivity matrix accordingly. For example, $\partial f_5/\partial c_4$ increases from 2/11 when $\rho = 0$ to 31/151 when $\rho = 1$. This confirms that increasing ρ strengthens the correlation between schedules associated with primary work participation.

Real Data Application

We apply the proposed framework to data from the 2023 Danish National Travel Survey. The study area covers 566 contiguous spatial zones in the Capital Region of Denmark. Time is discretized into 10-minute intervals from 5:00 to 23:00, yielding 108 time layers in the time–space network. The network includes four travel modes (car, public transport, bicycle, and walk) and four activity types (home, work, errand, and leisure). Home activity is restricted to the individual's residential zone, while other activities may occur in any zone.

The estimation sample consists of 186 weekday observations for which the reported daily schedule starts and ends at home and is fully contained within the study area. Each observed trajectory is represented as a feasible path in the time–space network, connecting the home zone at 5:00 to the home zone at 23:00.

Travel time and monetary costs are constructed from externally generated Level-of-Service (LOS) matrices defined between zone centroids. Travel times depend only on origin–destination pairs and mode and are assumed constant within the day. For car travel, LOS provide separate travel time and monetary cost components; the latter includes distance-based operating costs and toll charges. For public transport, travel time corresponds to weighted generalized time including in-vehicle, waiting, transfer, and access components. For walk and bicycle, travel time is derived from network-based shortest paths with fixed speeds.

Under this specification, the theoretical full time–space network contains approximately 140 million links per individual. In practice, however, an individual's effective consideration set is expected to comprise only a very small subset of these links, reflecting the sparsity of behaviorally relevant transitions within the full network.

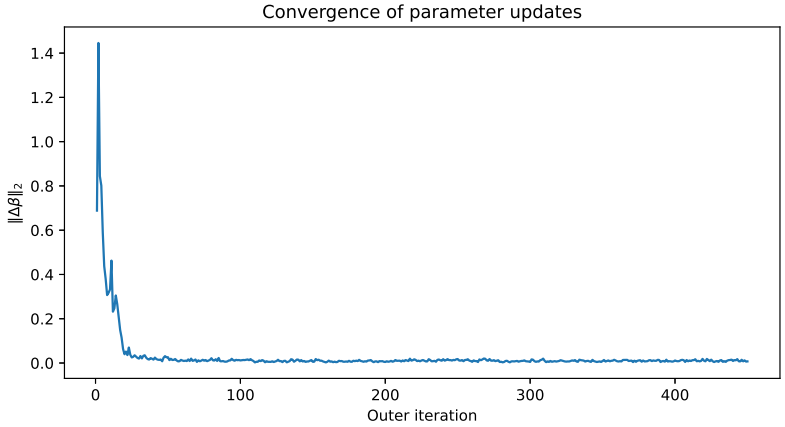
The systematic link cost specification follows Tables 1 and 2. Trip-related attributes include mode-specific travel time coefficients, a common monetary cost parameter, and alternative-specific constants for bicycle and walk. Activity-related components include activity-specific alternative-specific constants, a size term defined as the logarithm of the number of employees in the relevant industry within each zone, and time-of-day dummy variables capturing temporal participation patterns. For work activity, an additional own-work stay parameter captures the preference for performing work at the individual's registered workplace.

Figure 2 shows the convergence of the parameter estimates. Panel 2a indicates that $\|\Delta\boldsymbol{\beta}\|_2$ declines steadily, demonstrating stable convergence of the fixed-point estimation procedure. Panel 2b reports network growth during column generation. While the theoretical network comprises approximately 140 million links per individual, the final partial network contains about 20,000 links, of which roughly 6,000 carry positive flow. This indicates that predicted behavior concentrates on a small subset of the feasible time–space network.

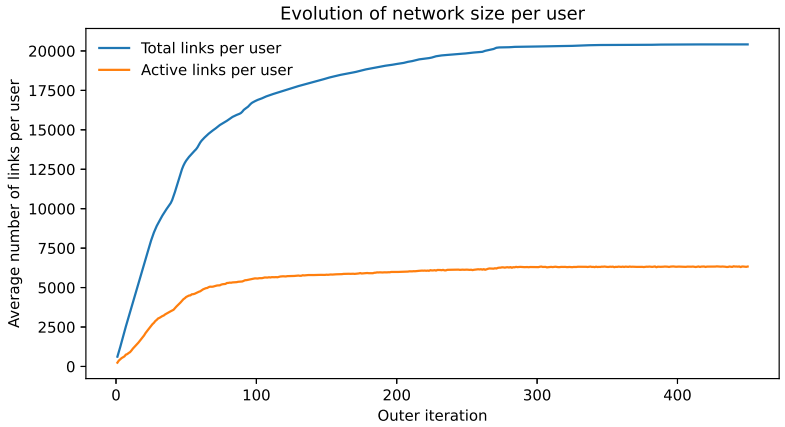
Tables 1 and 2 report the estimated parameters. Consistent with the cost-minimization formulation, travel time and monetary cost coefficients are positive, while the size term and the own-work stay parameter are negative. The estimated signs and the daily variation of time-of-day effects

Table 2: Activity-related cost parameters

Variable	Type	Work	Errand	Leisure	Home
Baseline effects					
Size term (log employees)	Continuous	-0.0009	-0.0008	-0.0008	–
ASC	Dummy	0.0011	-0.0009	-0.0015	-0.1664
Own-work stay	Dummy	-0.6149	–	–	–
Time-of-day effects					
5am–6am	Dummy	0.2651	–	–	-0.9195
6am–7am	Dummy	0.2067	–	–	-0.8043
7am–8am	Dummy	-0.0070	–	–	-0.6742
8am–9am	Dummy	-0.0357	–	–	-0.2610
9am–10am	Dummy	-0.0221	–	–	-0.1487
10am–11am	Dummy	-0.0046	–	–	-0.0819
3pm–4pm	Dummy	0.0520	–	-0.0003	-0.1283
4pm–5pm	Dummy	0.2084	–	0.0000	-0.2873
5pm–6pm	Dummy	0.3705	–	-0.0014	-0.4244
6pm–7pm	Dummy	0.4717	–	-0.0055	-0.5463
7pm–8pm	Dummy	0.4796	–	-0.0044	-0.6274
8pm–9pm	Dummy	0.5355	–	-0.0037	-0.6917
9pm–10pm	Dummy	0.5775	–	-0.0012	-0.7516
10pm–11pm	Dummy	0.6185	–	0.0016	-0.8161

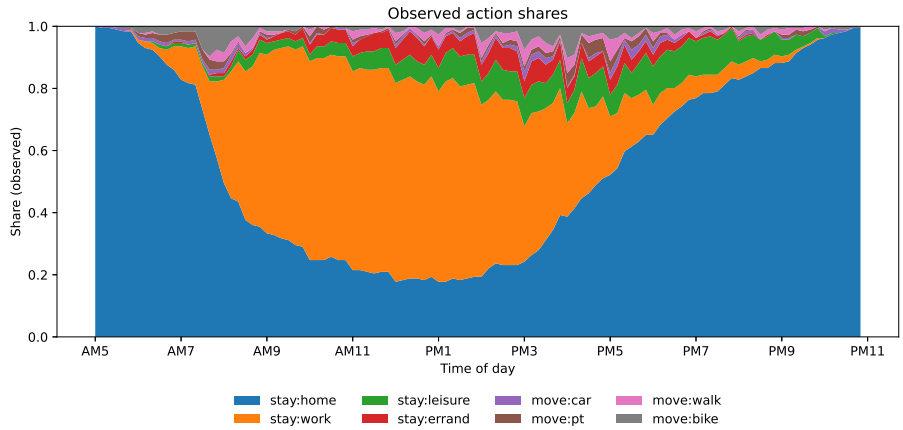


(a) Convergence of parameter updates ($\|\Delta\beta\|_2$).

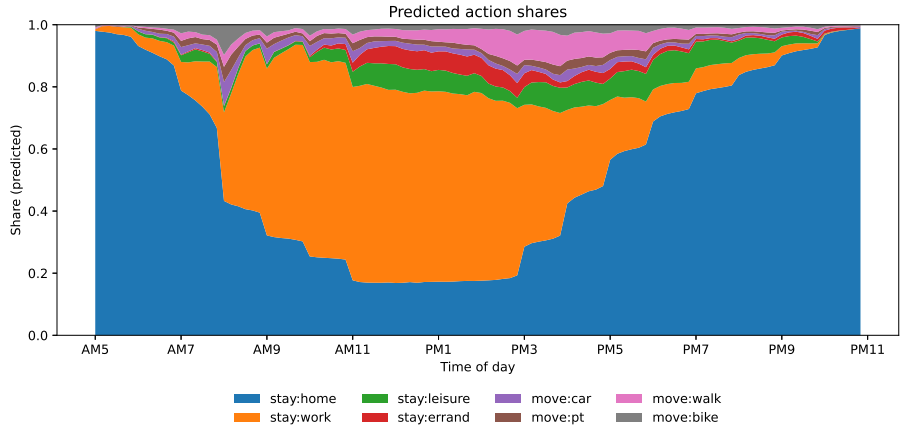


(b) Evolution of network size per user (total vs. active links).

Figure 2: Estimation dynamics over outer iterations.



(a) Observed action shares by time of day.



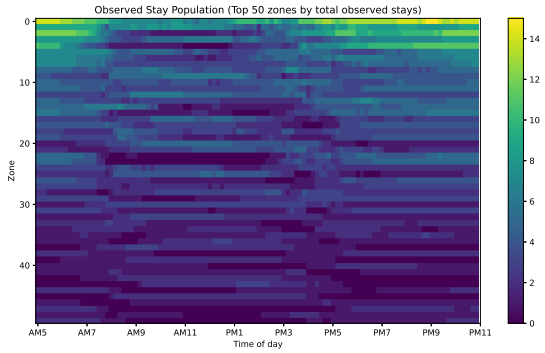
(b) Predicted action shares by time of day.

Figure 3: Time-of-day variation in activity and travel shares.

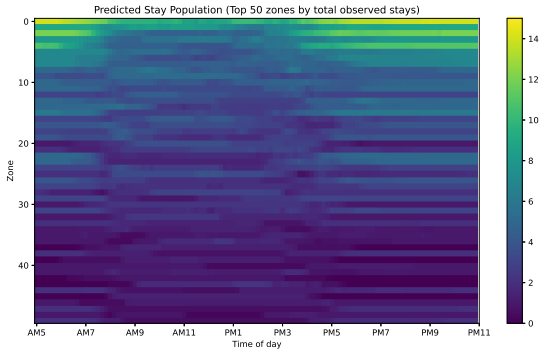
appear behaviorally reasonable. For public transport, the implied value of travel time is approximately 3.18 DKK per minute (about 191 DKK per hour), which lies within the range reported in the Danish national transport appraisal guidelines (The Danish Ministry of Transport, 2025). As the estimate is obtained within a joint activity scheduling framework rather than a pure mode choice model, its interpretation differs somewhat from conventional value-of-time measures.

Figure 3 reports observed and predicted action shares by time of day, while Figure 4 compares observed and predicted stay populations across zones over the day. The model captures the main daily patterns in activity participation and travel, including the morning decline in home activity and the concentration of work during standard working hours. The predicted population distribution also reflects the main concentration patterns across zones. Some discrepancies remain, notably a relatively high predicted share of walking, partly driven by intra-zonal walk links, and a smoother spatio-temporal distribution than observed.

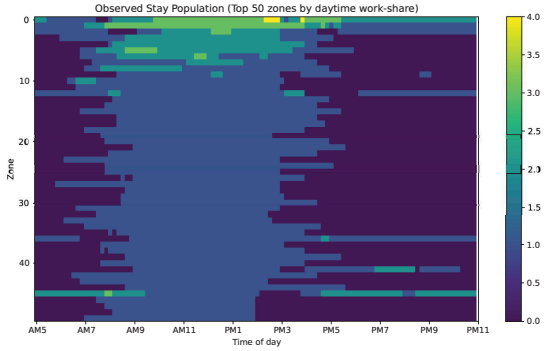
Overall, despite a relatively simple model specification and a modest number of observations, the model reproduces the main empirical regularities in the data. Because PURC relies on regression-type updates of aggregate link flows, the amount of observational information directly affects parameter stability and behavioral resolution. Expanding the set of explanatory variables, such as individual-level covariates, is important; equally important is increasing the estimation sample in order to capture sharper spatio-temporal heterogeneity. In addition, the group-based formulation described in Section 2 enables more flexible correlation structures, for example by specifying separate perturbations for movement and stay components. Once estimation has converged and the active network structure is identified, the analytical sensitivity framework introduced in the previous section becomes directly applicable. This allows quantitative assessment of substitution and complementarity effects within the empirically relevant network.



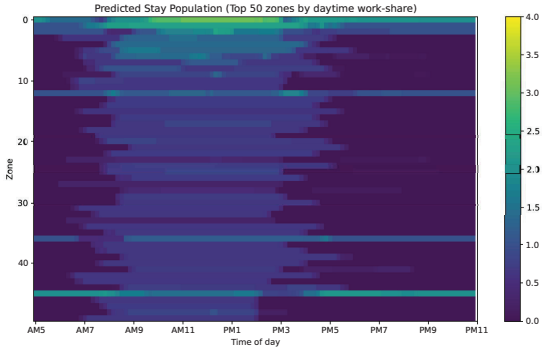
(a) Observed (Top 50 zones by total stays)



(b) Predicted (Top 50 zones by total stays)



(c) Observed (Top 50 zones by daytime work-share)



(d) Predicted (Top 50 zones by daytime work-share)

Figure 4: Observed and predicted stay populations by time of day across zones.

4 Conclusion

This study formulates daily activity scheduling within the Perturbed Utility Route Choice (PURC) framework, representing activity–travel behavior as flows on a time–space network and enabling flexible correlation patterns, including complementarity. Because real-world time–space networks are extremely large, full-network estimation is computationally infeasible. To address this challenge, we propose a column generation algorithm grounded in reduced-cost theory, enabling scalable estimation on iteratively expanded partial networks and eliminating the need for ex ante enumeration of the choice set. The empirical application demonstrates that the method is computationally feasible on realistic networks. In addition, we introduce a group-based perturbation structure that captures activity-specific correlations beyond network topology, and sensitivity analysis on a toy network illustrates how this extension flexibly controls substitution and complementarity patterns. Overall, the proposed framework provides a computationally scalable and behaviorally rich foundation for large-scale activity-based modeling.

References

Bowman, J. L., & Ben-Akiva, M. E. (2001). Activity-based disaggregate travel demand model system with activity schedules. *Transportation research part a: policy and practice*, *35*(1), 1–28.

Fosgerau, M., Nielsen, N., Paulsen, M., Rasmussen, T. K., & Yao, R. (2024a). Estimating the perturbed utility route choice model with individual-level data. In *12th symposium of the european association for research in transportation.*

Fosgerau, M., Nielsen, N., Paulsen, M., Rasmussen, T. K., & Yao, R. (2024b). Substitution in the perturbed utility route choice model. *Available at SSRN 4955133*.

Fosgerau, M., Paulsen, M., & Rasmussen, T. K. (2022). A perturbed utility route choice model. *Transportation Research Part C: Emerging Technologies*, *136*, 103514.

The Danish Ministry of Transport. (2025). *Transport economic unit prices version 2.2.* Retrieved from `https://www.man.dtu.dk/myndighedsbetjening/teresa-og-transportoekonomiske-enhedspriser`

Zimmermann, M., Västberg, O. B., Frejinger, E., & Karlström, A. (2018). Capturing correlation with a mixed recursive logit model for activity-travel scheduling. *Transportation Research Part C: Emerging Technologies*, *93*, 273–291.