# Heterogeneity in Mobility Behavior:

## Latent Classes from Long-Term Tracking Data

Anne-Valérie Preto Andrea Schneider Michel Bierlaire

Transport and Mobility Laboratory
School of Architecture, Civil and Environmental Engineering
Ecole Polytechnique Fédérale de Lausanne
Swiss Federal Railways (SBB)

26th Swiss Transport Research Conference

# Outline

# Research Question

**How can we extract structurally distinct mobility patterns from rich long-term tracking data, and which socio-economic characteristics predict class membership?**

**The Challenge:**

- Daily mobility is noisy and variable
- Two "50 km/day travelers" can be fundamentally different
- Need to separate signal (structure) from noise

**Why It Matters**

- **Flexible users:** Respond to pricing, shift to off-peak
- **Constrained users:** Drive peaks, captive audience

TRANSP-OR

EPFL

# Position in the Literature

**Mobility behavior is structured, but variable**

- Travel behavior is habitual, but not perfectly repetitive
- Multi-day patterns are difficult to compare
- Long-term data are needed to observe both regularity and variability

Examples: [Schlich and Axhausen, 2003], [Vij et al., 2013], [Walker and Li, 2007]

**Existing typologies only partially address this**

- Clustering identifies profiles, but usually fixes class membership
- Distance-based methods depend strongly on the chosen similarity measure

Examples: [Ben-Gal et al., 2019], [Tsoleridis et al., 2025]

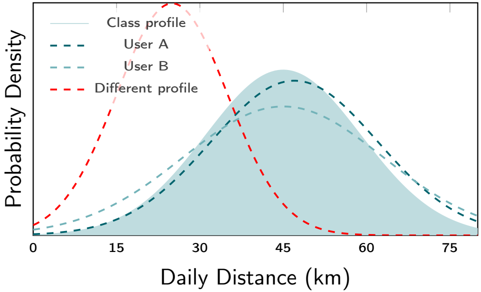
**Research gap:** A probabilistic framework is needed to identify mobility typologies while taking into account the data variability.

# The Methodological Challenge

**Tracking data contain several sources of variability that are difficult to separate.**

Observed variation may come from:

1. differences between mobility lifestyles,
2. persistent differences across individuals,
3. day-to-day fluctuations.




Same observed mobility range can mix within-class variation and overlap across classes.

# Outline

# Data: Continuous Mobility Panel

intervista

### Sample & Tracking

**2,011** respondents
**Oct–Dec 2024**
**96%** tracking completeness
Switzerland-wide CMP panel

### Who?

**50 / 50** women / men
**47 y.o.** median age
**69%** employed
**23%** with children

### Mobility Resources

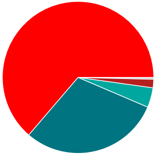
**66%** car available
**40%** Halbtax subscription
**7%** GA (annual pass)

Socio-demographic survey + Long-term daily mobility indicators from passive tracking data

# Key Mobility Metrics

intervista

## Modal Split (by distance)



- Private Vehicle (63.7%)
- Public Transport (29.9%)
- On Foot (4.2%)
- On Bicycle (1.8%)
- Other (0.4%)

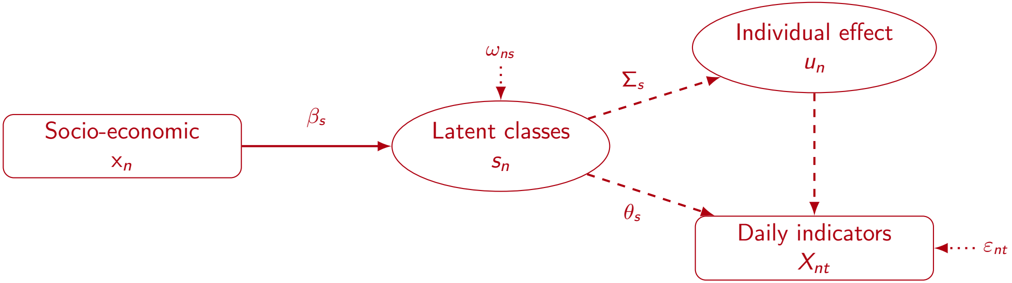
Note: Metrics computed excluding trips made outside Switzerland and by plane

## Trip & Distance Statistics

| Metric | Value |
|---|---|
| Total trips | 686,768 |
| Total legs | 1,156,045 |
| User-days | 154,279 |
| **Daily Distance** | |
| Median | 27.6 km |
| Mean | 52.1 km |
| 90th percentile | 131.2 km |
| **Trips per Day** | |
| Median | 4 |
| Mean | 4.5 |
| 90th percentile | 8 |

# Outline

# Hierarchical Model Structure




**Three sources of variation:**

1. **Between-class** ($\theta_s$): Distinct mobility lifestyles
2. **Within-class** ($u_n$): Individual deviation within the class
3. **Daily** ($\varepsilon_{nt}$): Random fluctuations

**Result:**

- Interpretable classes
- Stable individual effects
- Uncertainty quantification
- Forecasting capability

# Estimation Strategy

**Joint posterior with all three levels.**
The model is estimated by sampling from the joint posterior distribution

$$p(s_n, u_n, \pi, \theta_s, \Sigma_s \mid X) \propto \underbrace{p(X \mid s_n, u_n, \theta_s, \sigma_s)}_{\text{likelihood}} \cdot \underbrace{p(u_n \mid s_n, \Sigma_s)}_{\text{random effects}} \cdot \underbrace{p(s_n \mid \pi_n)}_{\text{class membership}} \cdot p(\pi_n, \theta_s, \Sigma_s)$$

**Why Bayesian?**

- Handles latent classes naturally
- Incorporates expert prior knowledge
- Uncertainty quantification
- Handles hierarchical structure

**Current implementation:**

- Multiple chains with informative priors
- Class-specific prior distributions
- Monitor convergence (Rhat, trace plots)
- Cross-validation for class count

# Measurement Equation

Observed indicators

- For each individual $n$ and day $t$:

$$X_{nt} = (\text{homeonly}, \text{private\_veh}, \text{active\_PT}, \text{trips}, \text{daily\_dist})$$

- Latent class $s_n$ determines the distribution of each indicator
- Conditional independence given $s_n$ and $u_n$

| **Indicator** | **Type** | **Distribution** |
|---|---|---|
| Home-only day | Binary | Bernoulli($p_k^H$) |
| Private vehicle use | Binary | Bernoulli($p_k^V$) |
| Active/PT use | Binary | Bernoulli($p_k^{PT}$) |
| Daily trip frequency | Count | Poisson($\lambda_k$) |
| Daily distance | Continuous | $\mathcal{N}(\mu_k, \sigma_k^2)$ |

# Class Membership Model

- Each individual $n$ has observed characteristics $x_n$
- These predict the probability of belonging to each latent class

$$P(s_n = s \mid x_n) = \frac{\exp(\alpha_s + \beta_s^\top x_n)}{\sum_\ell \exp(\alpha_\ell + \beta_\ell^\top x_n)}.$$

- Multinomial logit specification

| **Variable** | **Type** | **Values/Range** |
|---|---|---|
| Education | Categorical | Low, Medium, High |
| Age | Continuous | Piecewise (15-83) |
| Car availability | Binary | Yes/No |
| PT subscription | Binary | Yes/No (for now) |
| Employment | Categorical | Student, Part-time, Full-time, Unemployed, Retired |

# Outline

# Results: Step-by-Step

**Our approach: iterative model building**

1. **Step 1:** Start simple → Can we identify classes at all?
2. **Step 2:** Diagnose the problem → Why is assignment uncertain?
3. **Step 3:** Add complexity → Individual random effects (under development)
4. **Step 4:** Characterize findings → What do the classes tell us?

# Step 1: Single-Layer Model Results

| Parameter | Est. | SE | t-stat | p-value |
|---|---|---|---|---|
| Home-only$_0$ | 0.0417 | 0.0203 | 2.05 | 0.0401 |
| Home-only$_1$ | 0.111 | 0.00824 | 13.5 | $< 0.001$ |
| Home-only$_2$ | 0.00172 | 0.000757 | 2.27 | 0.0229 |
| Home-only$_3$ | 0.001 | 0.0118 | 0.0851 | 0.932 |
| Home-only$_4$ | 0.001 | 0.0112 | 0.0895 | 0.929 |
| Home-only$_5$ | 0.001 | 0.0144 | 0.0694 | 0.945 |
| Private vehicle$_0$ | 0.001 | — | — | — |
| Private vehicle$_1$ | 0.995 | 0.00204 | 488 | $< 0.001$ |
| Private vehicle$_2$ | 0.983 | 0.00422 | 233 | $< 0.001$ |
| Private vehicle$_3$ | 0.001 | 0.0118 | 0.0851 | 0.932 |
| Private vehicle$_4$ | 0.001 | 0.0112 | 0.0896 | 0.929 |
| Private vehicle$_5$ | 0.791 | 0.0104 | 76.1 | $< 0.001$ |
| Active+PT$_0$ | 0.001 | — | — | — |
| Active+PT$_1$ | 0.001 | 0.0154 | 0.0649 | 0.948 |
| Active+PT$_2$ | 0.001 | 0.0144 | 0.0694 | 0.945 |
| Active+PT$_3$ | 0.998 | 0.00262 | 381 | $< 0.001$ |
| Active+PT$_4$ | 0.999 | 0.00744 | 134 | $< 0.001$ |
| Active+PT$_5$ | 0.164 | 0.0114 | 14.4 | $< 0.001$ |
| Trip rate$_0$ | 0.100 | 0.00802 | 12.5 | $< 0.001$ |
| Trip rate$_1$ | 4.41 | 0.0688 | 64.1 | $< 0.001$ |
| Trip rate$_2$ | 7.47 | 0.0400 | 187 | $< 0.001$ |
| Trip rate$_3$ | 7.09 | 0.0722 | 98.2 | $< 0.001$ |
| Trip rate$_4$ | 10.1 | 0.0521 | 194 | $< 0.001$ |
| Trip rate$_5$ | 10.8 | 0.0115 | 939 | $< 0.001$ |

| Parameter | Est. | SE | t-stat | p-value |
|---|---|---|---|---|
| Daily distance$_0$ | 0 | 0 | −0.0184 | 0.985 |
| Daily distance$_1$ | 5.08 | 0.0447 | 114 | $< 0.001$ |
| Daily distance$_2$ | 41.7 | 0.00238 | $1.75 \times 10^4$ | $< 0.001$ |
| Daily distance$_3$ | 15.7 | 0.00418 | $3.74 \times 10^3$ | $< 0.001$ |
| Daily distance$_4$ | 82.2 | 0.000992 | $8.28 \times 10^4$ | $< 0.001$ |
| Daily distance$_5$ | 175 | 0.000104 | $1.68 \times 10^6$ | $< 0.001$ |
| Daily dist (SD)$_0$ | 0 | 0 | 0.78 | 0.435 |
| Daily dist (SD)$_1$ | 3.48 | 0.0623 | 55.9 | $< 0.001$ |
| Daily dist (SD)$_2$ | 29.2 | 0.00118 | $2.47 \times 10^4$ | $< 0.001$ |
| Daily dist (SD)$_3$ | 8.78 | 0.00552 | $1.59 \times 10^3$ | $< 0.001$ |
| Daily dist (SD)$_4$ | 51.0 | 0.000631 | $8.09 \times 10^4$ | $< 0.001$ |
| Daily dist (SD)$_5$ | 109 | 0.000220 | $4.96 \times 10^5$ | $< 0.001$ |

# Step 1: Six Mobility Lifestyles

### Low-mobility

| Socio-econ | Indicators |
|---|---|
| ASC: 0 | Dist: |
| PT: 0 | 0 km |
| Car av.: 0 | Trips: 0.1 |
| | Home 1% |

### Short Car

| Socio-econ | Indicators |
|---|---|
| ASC: 0.19 | Dist: |
| PT: -0.44 | 20±19 km |
| Car av.: 0.50 | Trips: 3.4 |
| | Priv 99% |

### Long Car

| Socio-econ | Indicators |
|---|---|
| ASC: -1.70 | Dist: |
| PT: -0.68 | 51±56 km |
| Car av.: 1.12 | Trips: 8.8 |
| | Priv 99% |

### Short Multimodal

| Socio-econ | Indicators |
|---|---|
| ASC: 0.35 | Dist: |
| PT: 0.75 | 32±38 km |
| Car av.: -1.22 | Trips: 7.5 |
| | PT 99% |

### Long Multimodal

| Socio-econ | Indicators |
|---|---|
| ASC: -1.72 | Dist: |
| PT: 0.21 | 76±57 km |
| Car av.: -0.39 | Trips: 11.2 |
| | PT 100% |

### High-mobility

| Socio-econ | Indicators |
|---|---|
| ASC: -4.04 | Dist: |
| PT: -1.00 | 172±168 km |
| Car av.: -1.06 | Trips: 8.8 |
| | PT 75% |

# Step 2: The Overlapping Distributions Problem

**Challenge:** Within-class variance $\approxeq$ Between-class variance

**Implication for individual assignment:**

- Person with 5 km trip: Could belong to Low-mobility OR Short-distance Car
- Class membership probabilities become noisy
- Signal masked by noise

**Key insight:** Need to separate **individual persistence** from **daily noise**

# Why daily aggregation hides hybrid workers

**Expected finding:** Hybrid workers (flexible schedules, remote + occasional long trips)

**What we actually see:** "Local, short-range mobility" class

- Short radius of mobility
- Home-office workers do not dominate this class
  - They still have daily distances on average, maybe for errands, leisure, etc.

**The fundamental problem:**

- Daily mobility indicators miss the purpose of trips
- Without activity type, we see only aggregated distance, not flexibility patterns

# Step 3: Hierarchical Model with Individual Effects

| Class | Users | Class mean | Class SD | Mean $u_{n,d}$ | Credible nonzero |
|---|---|---|---|---|---|
| Low-mobility | 47 | 1.2 | 21.1 | $-18.4$ | 40% |
| Short-distance Car | 124 | 16.9 | 78.0 | $-17.8$ | 25% |
| Long-distance Car | 101 | 41.2 | 38.6 | $-14.1$ | 51% |
| Short-distance Multimodal | 83 | 23.5 | 64.2 | $-16.4$ | 36% |
| Long-distance Multimodal | 31 | 71.8 | 29.0 | $-28.7$ | 65% |
| High-mobility | 14 | 84.2 | 79.2 | 14.1 | 7% |
| All users | 400 | – | – | $-16.4$ | 38% |

Table: Class mean daily distance and individual deviations

- $u_{n,d}$ captures persistent individual deviation from the class mean daily distance
- Multimodal RC far shows the strongest within-class correction

Note: some variance parameters have weak convergence diagnostics and should be interpreted cautiously.

# Step 4: Diagnostic of the Hierarchical Model

**The hierarchical model reveals an identification challenge.**

## What the diagnostics show

- Distance means converge; variances do not
- Low ESS for variance and mode-use parameters
- Class membership needs stronger socio-economic characteristics

This suggests that the model identifies average class profiles more easily than the full decomposition of variability.

## Why sampling is hard

- Classes have overlapping daily profiles
- Binary indicators: limited individual information
- Distributions skewed and highly variable

Different chains explain observations via different effect combinations.

Daily aggregation cannot fully separate lifestyles, individual heterogeneity, and daily noise.

# Outline

# Key Takeaways

1. **Six distinct mobility classes**
   - Clear separation in mode choice and daily distances
2. **Individual random effects are as important as class means**
   - 38% credible individual deviations
   - Within-class SD often competes with between-class differences
3. **Daily aggregation masks activity purpose**
   - Cannot distinguish remote workers from local commuters
4. **Hierarchical decomposition works partially**
   - Reveals class means reliably
   - Needs richer socio-economic predictors

# Next Steps: Strategic Changes

**Class Membership Model enrichment:**

- Workplace locations (inferred from long visits)
- Household characteristics (size, children)
- Geographic context (urban density, transit access)

**Additional layer: Tour-level decomposition**

- **Decompose** days into tours (home-based sequences)
- **Classify** by duration: 8h+ (work), 2–8h (mixed), <2h (shopping)

# Thank You

## Questions?

anne-valerie.preto@epfl.ch

# Individual Level: Persistent Heterogeneity

**Individual random effects.** Each individual is associated with a single random-effect vector:

$$u_n \in \mathbb{R}^K.$$

**Class-dependent heterogeneity.** The distribution of individual deviations depends on the latent class:

$$u_n \mid s_n = s \sim \mathcal{N}(0, \Sigma_s).$$

**Role.** Individual random effects capture persistent deviations from class-level behavior and are shared across all days for individual $n$.

# Likelihood with Individual Effects

**Parameters**

- $x_{nt}^{\text{dist}}$: Daily distance for individual $n$, day $t$
- $\mu_{s_n}$: Class mean distance
- $u_{n,\text{dist}}$: Individual deviation (constant across $t$)
- $\sigma_{s_n}$: Day-to-day noise (class-specific)
- $\tau_{s_n}$: Between-individual variance (class-specific)
- $s_n$: Latent class

**Measurement model:**

$$x_{nt}^{\text{dist}} \sim \text{TruncNormal}\left(\mu_{s_n} + u_{n,\text{dist}}, \sigma_{s_n}^2; 0\right)$$

**Random effect:**

$$u_{n,\text{dist}} \mid s_n \sim \mathcal{N}\left(0, \tau_{s_n}^2\right)$$

**Log-likelihood:**

$$\ell_n = \sum_{t=1}^{T_n} \log p(x_{nt}^{\text{dist}} \mid s_n, u_{n,\text{dist}}, \mu_{s_n}, \sigma_{s_n})$$

# Table 1: MLE Estimates (Single-Layer Model)

| Parameter | Estimate | Std. Err. | $t$-stat | $p$-value |
|---|---|---|---|---|
| Home-only$_0$ | 0.0417 | 0.0203 | 2.05 | 0.0401 |
| Home-only$_1$ | 0.111 | 0.00824 | 13.5 | $< 0.001$ |
| Home-only$_2$ | 0.00172 | 0.000757 | 2.27 | 0.0229 |
| Home-only$_3$ | 0.001 | 0.0118 | 0.0851 | 0.932 |
| Home-only$_4$ | 0.001 | 0.0112 | 0.0895 | 0.929 |
| Home-only$_5$ | 0.001 | 0.0144 | 0.0694 | 0.945 |
| Private vehicle$_0$ | 0.001 | — | — | — |
| Private vehicle$_1$ | 0.995 | 0.00204 | 488 | $< 0.001$ |
| Private vehicle$_2$ | 0.983 | 0.00422 | 233 | $< 0.001$ |
| Private vehicle$_3$ | 0.001 | 0.0118 | 0.0851 | 0.932 |
| Private vehicle$_4$ | 0.001 | 0.0112 | 0.0896 | 0.929 |
| Private vehicle$_5$ | 0.791 | 0.0104 | 76.1 | $< 0.001$ |
| Active+PT$_0$ | 0.001 | — | — | — |
| Active+PT$_1$ | 0.001 | 0.0154 | 0.0649 | 0.948 |
| Active+PT$_2$ | 0.001 | 0.0144 | 0.0694 | 0.945 |
| Active+PT$_3$ | 0.998 | 0.00262 | 381 | $< 0.001$ |
| Active+PT$_4$ | 0.999 | 0.00744 | 134 | $< 0.001$ |
| Active+PT$_5$ | 0.164 | 0.0114 | 14.4 | $< 0.001$ |
| Trip rate$_0$ | 0.100 | 0.00802 | 12.5 | $< 0.001$ |
| Trip rate$_1$ | 4.41 | 0.0688 | 64.1 | $< 0.001$ |
| Trip rate$_2$ | 7.47 | 0.0400 | 187 | $< 0.001$ |
| Trip rate$_3$ | 7.09 | 0.0722 | 98.2 | $< 0.001$ |
| Trip rate$_4$ | 10.1 | 0.0521 | 194 | $< 0.001$ |
| Trip rate$_5$ | 10.8 | 0.0115 | 939 | $< 0.001$ |

# Table 1: MLE Estimates (Single-Layer Model)

| Parameter | Estimate | Std. Err. | $t$-stat | $p$-value |
|---|---|---|---|---|
| Daily distance$_0$ | 0 | 0 | –0.0184 | 0.985 |
| Daily distance$_1$ | 5.08 | 0.0447 | 114 | $< 0.001$ |
| Daily distance$_2$ | 41.7 | 0.00238 | $1.75 \times 10^4$ | $< 0.001$ |
| Daily distance$_3$ | 15.7 | 0.00418 | $3.74 \times 10^3$ | $< 0.001$ |
| Daily distance$_4$ | 82.2 | 0.000992 | $8.28 \times 10^4$ | $< 0.001$ |
| Daily distance$_5$ | 175 | 0.000104 | $1.68 \times 10^6$ | $< 0.001$ |
| Daily distance (SD)$_0$ | 0 | 0 | 0.78 | 0.435 |
| Daily distance (SD)$_1$ | 3.48 | 0.0623 | 55.9 | $< 0.001$ |
| Daily distance (SD)$_2$ | 29.2 | 0.00118 | $2.47 \times 10^4$ | $< 0.001$ |
| Daily distance (SD)$_3$ | 8.78 | 0.00552 | $1.59 \times 10^3$ | $< 0.001$ |
| Daily distance (SD)$_4$ | 51.0 | 0.000631 | $8.09 \times 10^4$ | $< 0.001$ |
| Daily distance (SD)$_5$ | 109 | 0.000220 | $4.96 \times 10^5$ | $< 0.001$ |

# Table 2: Posterior Estimates (Hierarchical Bayesian Model)

| Parameter | Mean | SD | $\hat{R}$ | $ESS_{bulk}$ |
|---|---|---|---|---|
| Home-only$_0$ | 0.070 | 0.021 | 1.65 | 7.0 |
| Home-only$_1$ | 0.015 | 0.003 | 1.17 | 15.0 |
| Home-only$_2$ | 0.021 | 0.009 | 1.53 | 7.0 |
| Home-only$_3$ | 0.010 | 0.003 | 1.12 | 23.0 |
| Home-only$_4$ | 0.004 | 0.002 | 1.03 | 84.0 |
| Home-only$_5$ | 0.008 | 0.006 | 1.53 | 7.0 |
| Private vehicle$_0$ | 0.675 | 0.064 | 2.26 | 5.0 |
| Private vehicle$_1$ | 0.677 | 0.057 | 2.36 | 5.0 |
| Private vehicle$_2$ | 0.918 | 0.014 | 1.63 | 7.0 |
| Private vehicle$_3$ | 0.156 | 0.040 | 1.70 | 6.0 |
| Private vehicle$_4$ | 0.072 | 0.028 | 1.54 | 7.0 |
| Private vehicle$_5$ | 0.603 | 0.041 | 1.59 | 7.0 |
| Active+PT$_0$ | 0.308 | 0.062 | 2.36 | 5.0 |
| Active+PT$_1$ | 0.310 | 0.058 | 2.40 | 5.0 |
| Active+PT$_2$ | 0.074 | 0.013 | 1.59 | 7.0 |
| Active+PT$_3$ | 0.842 | 0.041 | 1.72 | 6.0 |
| Active+PT$_4$ | 0.926 | 0.029 | 1.55 | 7.0 |
| Active+PT$_5$ | 0.383 | 0.040 | 1.49 | 7.0 |
| Trip rate$_0$ | 6.681 | 0.474 | 1.62 | 7.0 |
| Trip rate$_1$ | 7.598 | 0.390 | 1.64 | 7.0 |
| Trip rate$_2$ | 7.008 | 0.309 | 1.34 | 10.0 |
| Trip rate$_3$ | 8.697 | 0.383 | 1.43 | 8.0 |
| Trip rate$_4$ | 8.899 | 0.646 | 1.26 | 12.0 |
| Trip rate$_5$ | 9.691 | 1.325 | 1.71 | 6.0 |

# Table 2: Posterior Estimates (Hierarchical Bayesian Model)

| Parameter | Mean | SD | $\hat{R}$ | ESS$_{bulk}$ |
|---|---|---|---|---|
| Daily distance$_0$ | 1.200 | 0.953 | 1.05 | 71.0 |
| Daily distance$_1$ | 16.880 | 1.049 | 1.08 | 34.0 |
| Daily distance$_2$ | 41.243 | 2.012 | 1.16 | 17.0 |
| Daily distance$_3$ | 23.527 | 1.145 | 1.18 | 15.0 |
| Daily distance$_4$ | 71.779 | 1.930 | 1.06 | 45.0 |
| Daily distance$_5$ | 84.244 | 7.208 | 1.81 | 6.0 |
| Daily distance (SD)$_0$ | 21.112 | 2.216 | 2.28 | 5.0 |
| Daily distance (SD)$_1$ | 78.024 | 8.518 | 2.14 | 5.0 |
| Daily distance (SD)$_2$ | 38.597 | 4.118 | 2.52 | 5.0 |
| Daily distance (SD)$_3$ | 64.201 | 22.811 | 2.14 | 5.0 |
| Daily distance (SD)$_4$ | 28.998 | 28.863 | 1.55 | 7.0 |
| Daily distance (SD)$_5$ | 79.190 | 48.334 | 2.30 | 5.0 |

# Table 3: Multinomial Logit Coefficients

| Class | Covariate | Mean | SD | 95% CI Lower | 95% CI Upper |
|---|---|---|---|---|---|
| Short-dist car | Education | 0.030 | 0.293 | −0.544 | 0.604 |
| | Age | −0.221 | 0.225 | −0.663 | 0.220 |
| | Car availability | 0.766 | 0.392 | −0.002 | 1.535 |
| | Transit subscription | 0.854 | 0.359 | 0.151 | 1.557 |
| | Job: other | 0.078 | 0.547 | −0.995 | 1.152 |
| | Job: retired | 0.867 | 0.514 | −0.141 | 1.874 |
| | Job: student | −0.603 | 0.559 | −1.699 | 0.492 |
| Long-dist car | Education | 0.134 | 0.211 | −0.279 | 0.546 |
| | Age | 0.028 | 0.239 | −0.440 | 0.496 |
| | Car availability | 1.557 | 0.464 | 0.647 | 2.466 |
| | Transit subscription | −0.854 | 0.364 | −1.568 | −0.140 |
| | Job: other | −0.499 | 0.559 | −1.595 | 0.598 |
| | Job: retired | −0.586 | 0.597 | −1.755 | 0.583 |
| | Job: student | 0.081 | 0.658 | −1.208 | 1.371 |
| Short-dist modal | Education | −0.194 | 0.270 | −0.723 | 0.335 |
| | Age | 0.095 | 0.249 | −0.393 | 0.583 |
| | Car availability | −0.978 | 0.388 | −1.738 | −0.217 |
| | Transit subscription | 0.523 | 0.504 | −0.466 | 1.512 |
| | Job: other | 0.033 | 0.538 | −1.022 | 1.088 |
| | Job: retired | −0.350 | 0.690 | −1.702 | 1.003 |
| | Job: student | 0.476 | 0.614 | −0.727 | 1.679 |

# Table 3: Multinomial Logit Coefficients

| Class | Covariate | Mean | SD | 95% CI Lower | 95% CI Upper |
|---|---|---|---|---|---|
| Long-dist PT | Education | −0.170 | 0.268 | −0.695 | 0.355 |
| | Age | −0.127 | 0.285 | −0.686 | 0.431 |
| | Car availability | −0.951 | 0.445 | −1.823 | −0.079 |
| | Transit subscription | 0.351 | 0.607 | −0.838 | 1.541 |
| | Job: other | −0.282 | 0.669 | −1.593 | 1.028 |
| | Job: retired | −0.382 | 0.781 | −1.912 | 1.149 |
| | Job: student | −0.062 | 0.710 | −1.453 | 1.329 |
| High-mobility | Education | 0.020 | 0.466 | −0.895 | 0.934 |
| | Age | −0.274 | 0.324 | −0.909 | 0.361 |
| | Car availability | 0.658 | 0.573 | −0.465 | 1.780 |
| | Transit subscription | −0.108 | 0.573 | −1.232 | 1.016 |
| | Job: other | −0.897 | 0.800 | −2.465 | 0.671 |
| | Job: retired | −0.714 | 0.822 | −2.324 | 0.897 |
| | Job: student | 0.044 | 0.698 | −1.325 | 1.412 |

# Bibliography I


Ben-Gal, I., Weinstock, S., Singer, G., and Bambos, N. (2019).
Clustering users by their mobility behavioral patterns.
*ACM Transactions on Knowledge Discovery from Data*, 13(4):45.

Schlich, R. and Axhausen, K. W. (2003).
Habitual travel behaviour: Evidence from a six-week travel diary.
*Transportation*, 30(1):13–36.

Tsoleridis, P., Choudhury, C. F., and Hess, S. (2025).
Using probabilistic clustering techniques as a specification tool for capturing heterogeneity in choice models.
*Transportation Research Part C: Emerging Technologies*, 179:105289.

# Bibliography II

Vij, A., Carrel, A., and Walker, J. L. (2013).
Incorporating the influence of latent modal preferences on travel mode choice behavior.
*Transportation Research Part A: Policy and Practice*, 54:164–178.

Walker, J. L. and Li, J. (2007).
Latent lifestyle preferences and household location decisions.
*Journal of Geographical Systems*, 9(1):77–101.