2025 RESEARCH SEMINAR PRESENTATION

Paper: Path Size Logit route choice models: Issues with current models, a new internally consistent approach, and parameter estimation on a large-scale network with GPS data

Jingran Su
May 20, 2025

Abstract

This paper introduces the **Adaptive Path Size Logit (APSL)** model, which improves upon existing Path Size Logit models by ensuring internal consistency in defining unrealistic routes. APSL weights route contributions based on choice probabilities, addressing inconsistencies in earlier models. The paper proves solution existence/uniqueness, provides an estimation method, and validates APSL with simulations and real GPS data.

Introduction

the **Multinomial Logit (MNL) Random Utility Model (RUM)** often provides unrealistic choice probabilities when applied to real road network route choice. One of the main reasons for this is that the MNL model does not capture the correlation between routes. (IID assumptions)

Category	Typical model	Main feature	Issues
GEV Structures	Cross-Nested Logit (CNL) model, the Paired Combinatorial Logit (PCL) model, the Generalized Nested Logit (GNL)	-Two-level tree structure -Route choice probabilities are complex for large networks	-Estimation of nesting coefficients can be difficult -Some models like CNL and GNL have difficulty with large-scale applications
Mixed Logit	Logit Kernel, Random Parameter Logit, Error Component Logit, and Hybrid Logit.	-Allows correlation between routes due to random errors -Includes random parameter components	-Requires Monte Carlo simulation, which is computationally expensive -Estimating parameters (e.g., covariance) is complex and unstable
MNL-modification Models	C-Logit (CL), Path Size Logit (PSL), and Path Size Correction Logit (PSCL) and Generalised Path Size Logit (GPSL) model.	-Adjusts choice probabilities to approximate correlations -Retains single-level structure of MNL	-While computationally simpler than Mixed Logit, it may require complex estimation -May fail to properly capture route correlations and computational complexity increases with network size
Alternative RUMs	Multinomial Logit (MNL), Multinomial Probit Model (MNP), Multinomial Gamma (MNG)	-Models route choice with random error terms	-Computationally expensive, especially for large-scale networks -Difficult to compute closed-form route probabilities

Definition

Route choice probability:

$$D = \left\{ \boldsymbol{P} \in \mathbb{R}^N_{\geq 0} : 0 \leq P_i \leq 1, \quad \forall i \in R, \sum_{j=1}^{N} P_j = 1 \right\}.$$

Multinomial logit:

$$P_i = \Pr\left(U_i \geq U_j, \quad \forall j \in R, j \neq i\right) = \Pr\left(V_i + \varepsilon_i \geq V_j + \varepsilon_j, \quad \forall j \in R, j \neq i\right).$$

$$P_i = \frac{e^{V_i}}{\sum_{j \in R} e^{V_j}}.$$

$$V_i = -\theta c_i \qquad P_i = \frac{e^{-\theta c_i}}{\sum_{j \in R} e^{-\theta c_j}} = \frac{1}{\sum_{j \in R} e^{-\theta (c_j - c_i)}}.$$

Path size logit models:
Path Size Logit models include correction terms to penalise routes that share links with other routes

$$V_i = -\theta c_i + \kappa_i$$

$$\kappa_i = \beta \ln(\gamma_i), \qquad P_i = \frac{e^{-\theta c_i + \beta \ln(\gamma_i)}}{\sum_{j \in R} e^{-\theta c_j + \beta \ln(\gamma_j)}} = \frac{(\gamma_i)^\beta e^{-\theta c_i}}{\sum_{j \in R} (\gamma_j)^\beta e^{-\theta c_j}} = \frac{1}{\sum_{j \in R} \left(\frac{\gamma_j}{\gamma_i}\right)^\beta e^{-\theta (c_j - c_i)}}.$$

Definition

Path size logit:

$$\gamma_i^{PS} = \sum_{a \in A_i} \frac{t_a}{c_i} \frac{1}{\sum_{k \in R} \delta_{a,k}}.$$

PSL Key Issue: *Unrealistic routes negatively impact the choice probabilities of realistic routes when links are shared.*

The key issue with the PSL model is that all routes contribute equally to path size terms (i.e. the path size contribution factors are simply all 1), and hence the choice probabilities of realistic routes are affected by link sharing with unrealistic routes.

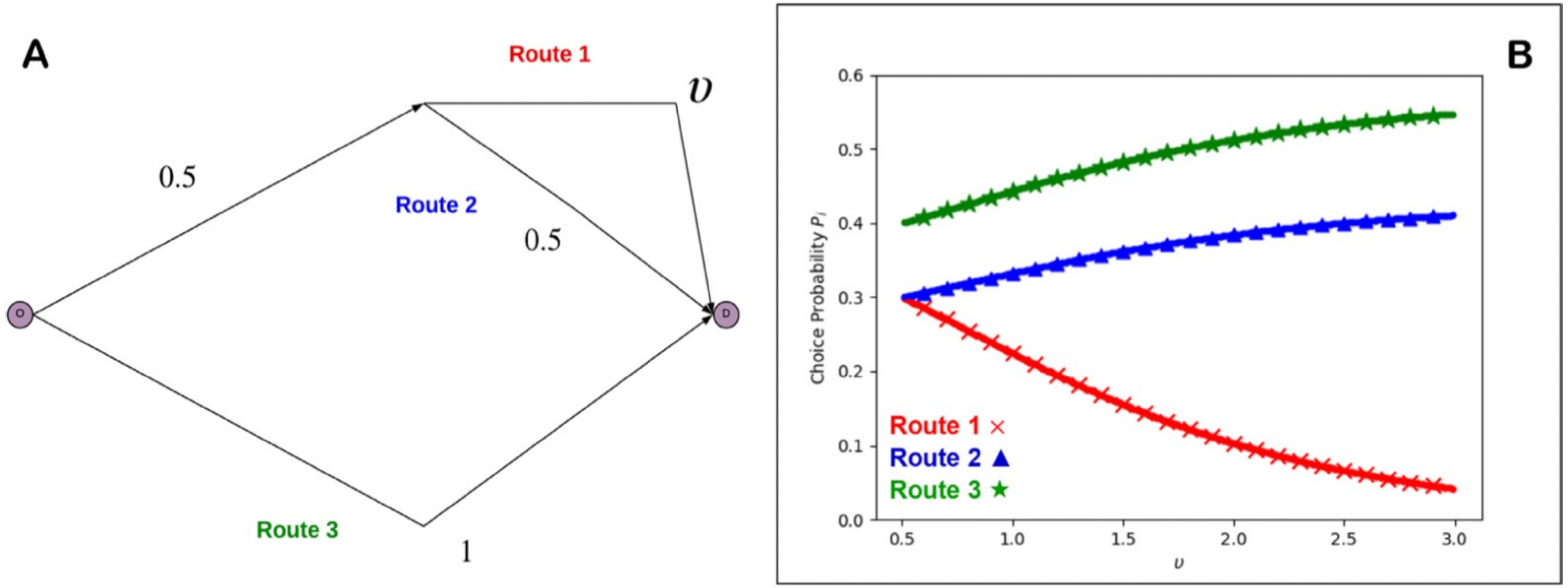


Fig. 2. (A): Example network 1. (B): Example network 1: PSL route choice probabilities for increasing υ; $\theta = \beta = 1$.

Definition

Generalized path size logit:

$$\gamma_i^{GPS} = \sum_{a \in A_i} \frac{t_a}{c_i} \frac{1}{\sum_{k \in R} \left(\frac{c_i}{c_k}\right)^{\lambda} \delta_{a,k}},$$

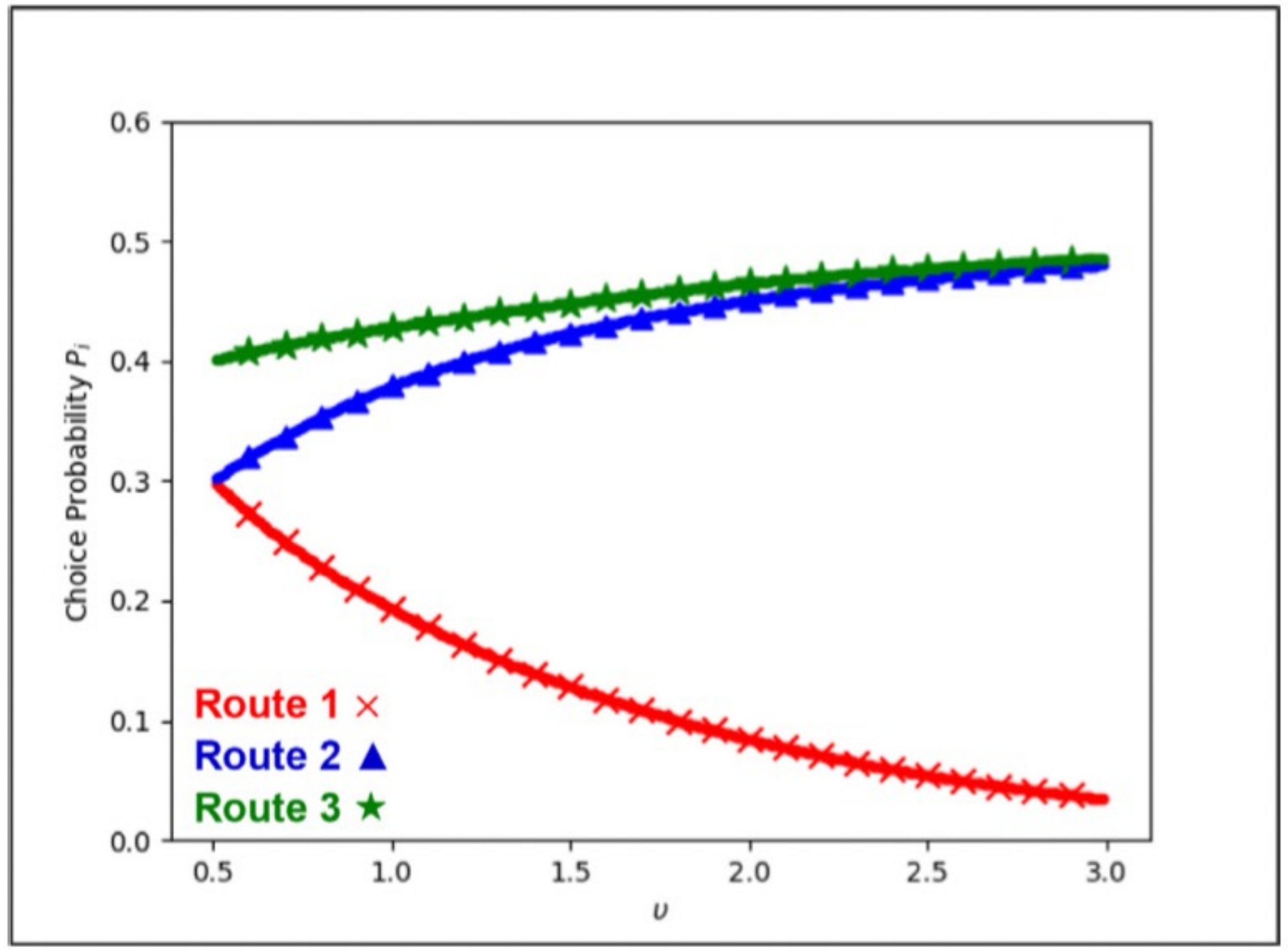


Fig. 3. Example network 1: GPSL route choice probabilities for increasing υ; $\theta = \beta = 1$, $\lambda = 3$.

Definition

GPSL Key Issue 1: *For large λ, GPSL path size terms are highly sensitive to small differences in route travel cost.*

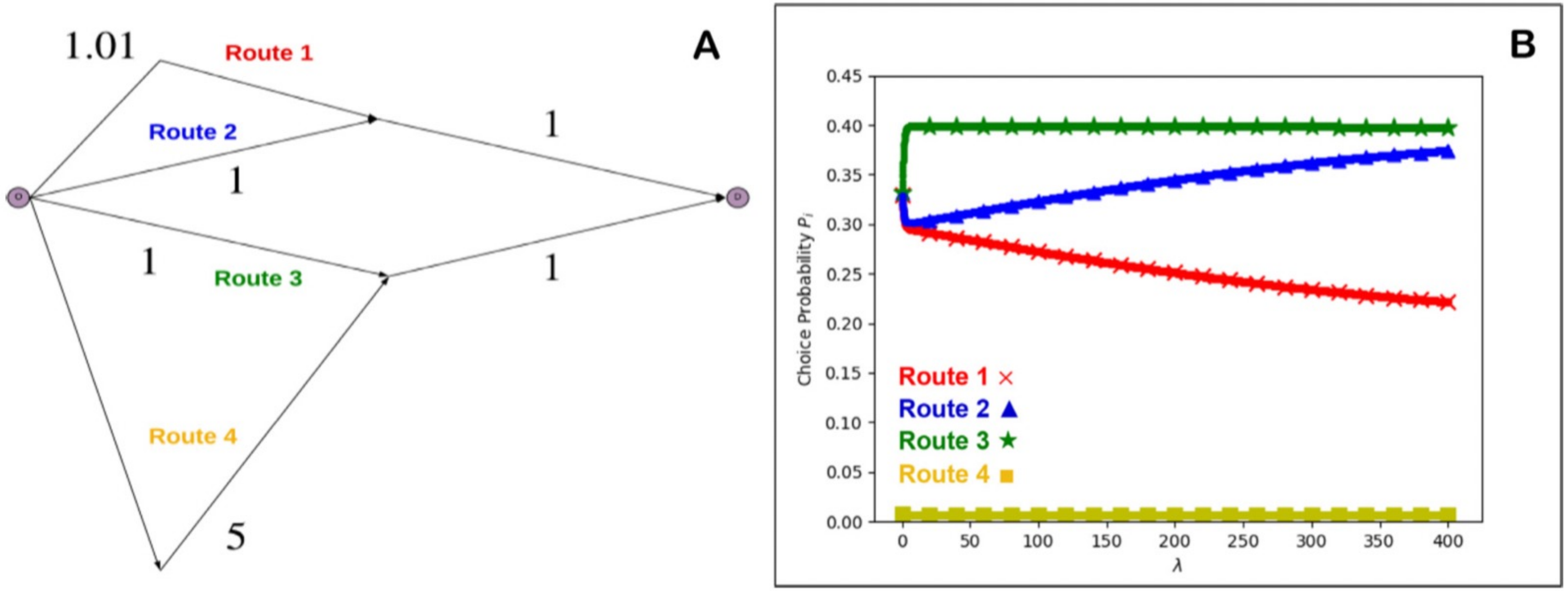
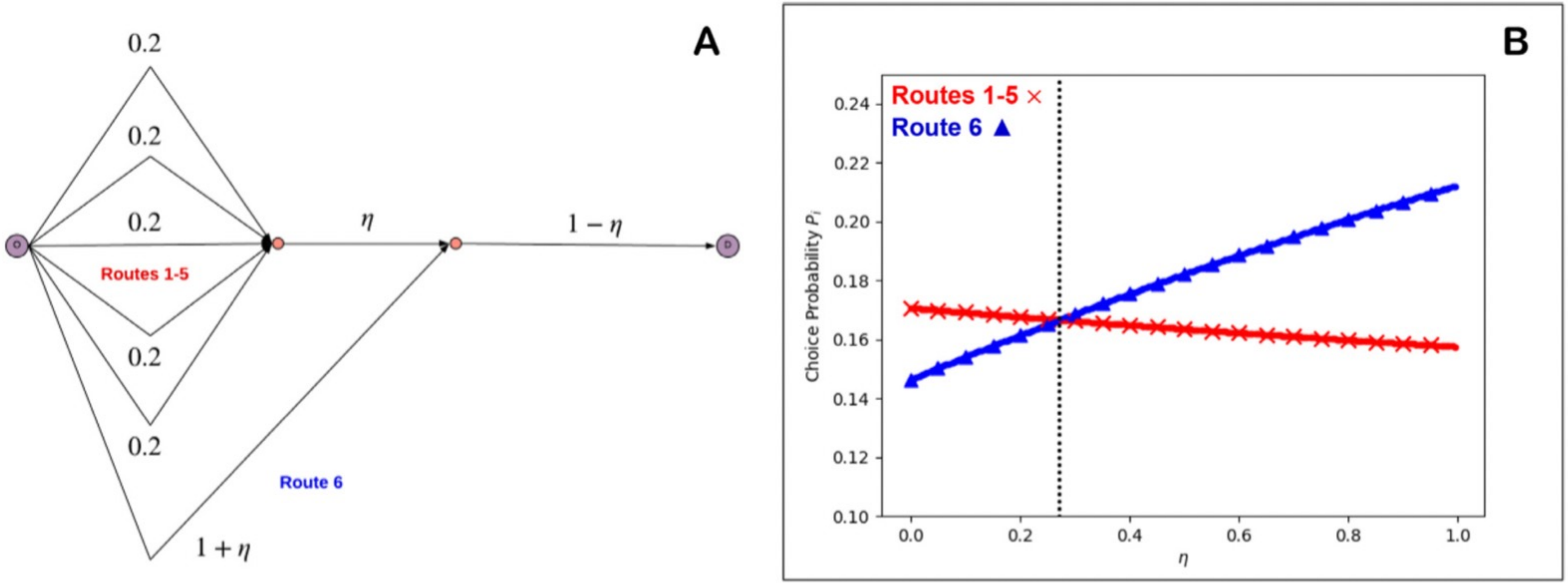


Fig. 4. (A): Example network 2. (B): Example network 2: GPSL route choice probabilities for increasing λ; $\theta = \beta = 1$.

Definition

GPSL Key Issue 2: *Internally inconsistent assessment of the feasibility of routes.*



$\eta = 1$, Route 6 to have a choice probability greater than each of Routes 1–5 due to being distinct.

$\eta = 0$, Route 6 has a smaller choice probability than each of Routes 1–5 due to having the larger detour.

Definition

GPSL Key Issue 2: *Internally inconsistent assessment of the feasibility of routes*.

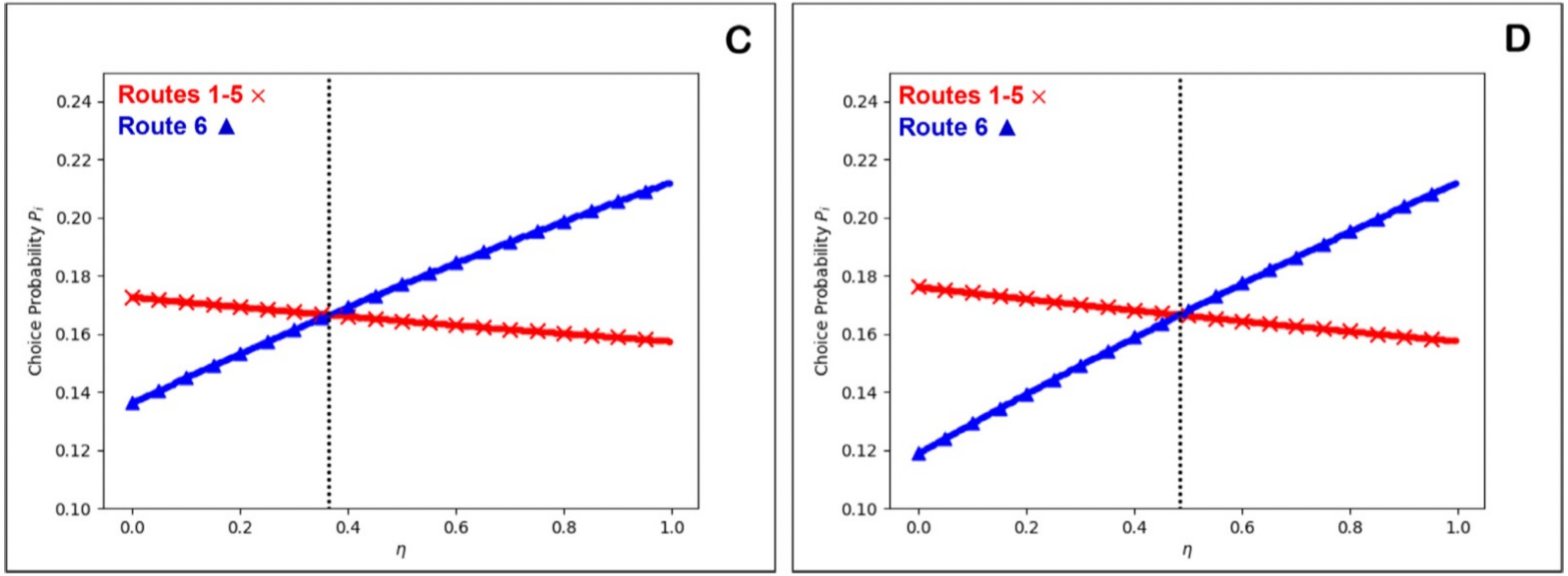
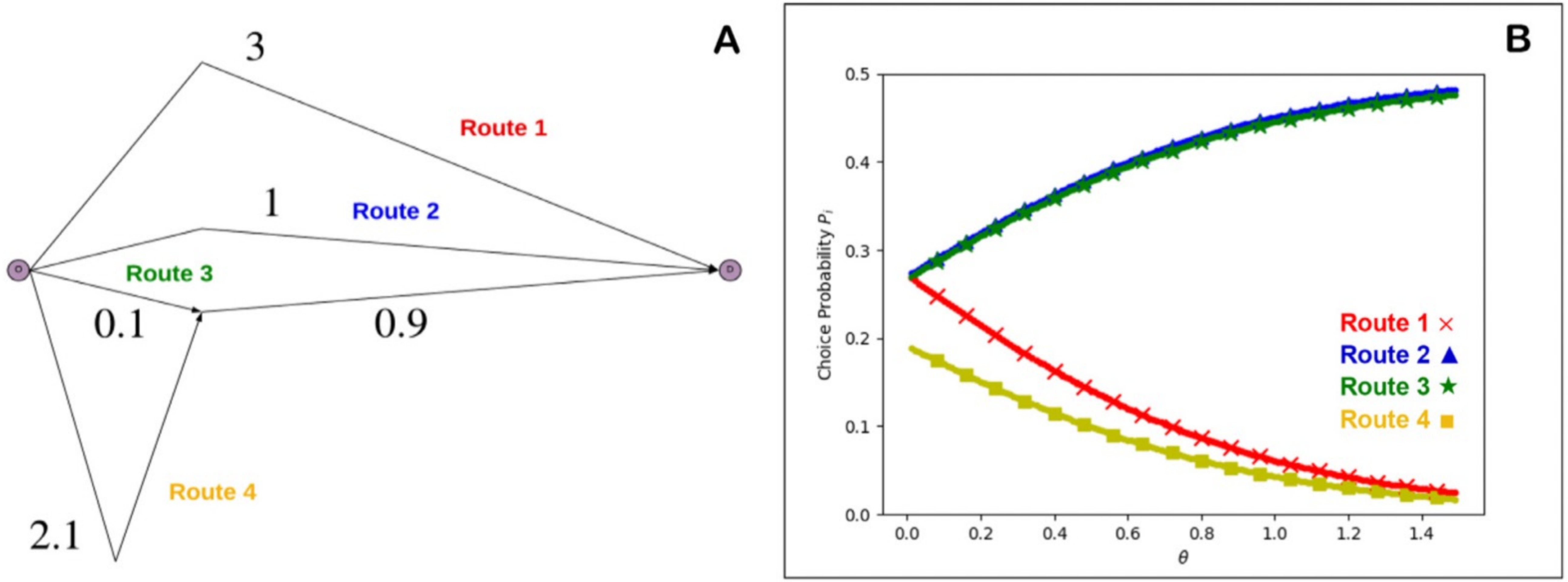


Fig. 5. (A): Example network 3. Example network 3: Route choice probabilities for increasing η; $\theta = \beta = 1$ –(B): PSL, ($\eta_{eq} = 0.27$). (C): GPSL, $\lambda = 1$, ($\eta_{eq} = 0.37$). (D): GPSL, $\lambda = 10$, ($\eta_{eq} = 0.48$).

larger values for λ moves *η_eq* further away from 0.27.

Definition

GPSL Key Issue 3: *Internally inconsistent scaling parameters.*



B. GPSL, $\beta = 1, \lambda = 4$.
For low ϑ however, the sensitivity to the differences in travel cost is dampened within the probability relation, yet Route 4's path size term contribution to Route 3 accentuates the difference in cost, and the GPSL model is thus inconsistent.

Definition

GPSL Key Issue 3: *Internally inconsistent scaling parameters.*
Alternative GPSL':

$$\gamma_i^{GPS'} = \sum_{a \in A_i} \frac{t_a}{c_i} \frac{1}{\sum_{k \in R} e^{-\lambda(c_k - c_i)} \delta_{a,k}}.$$

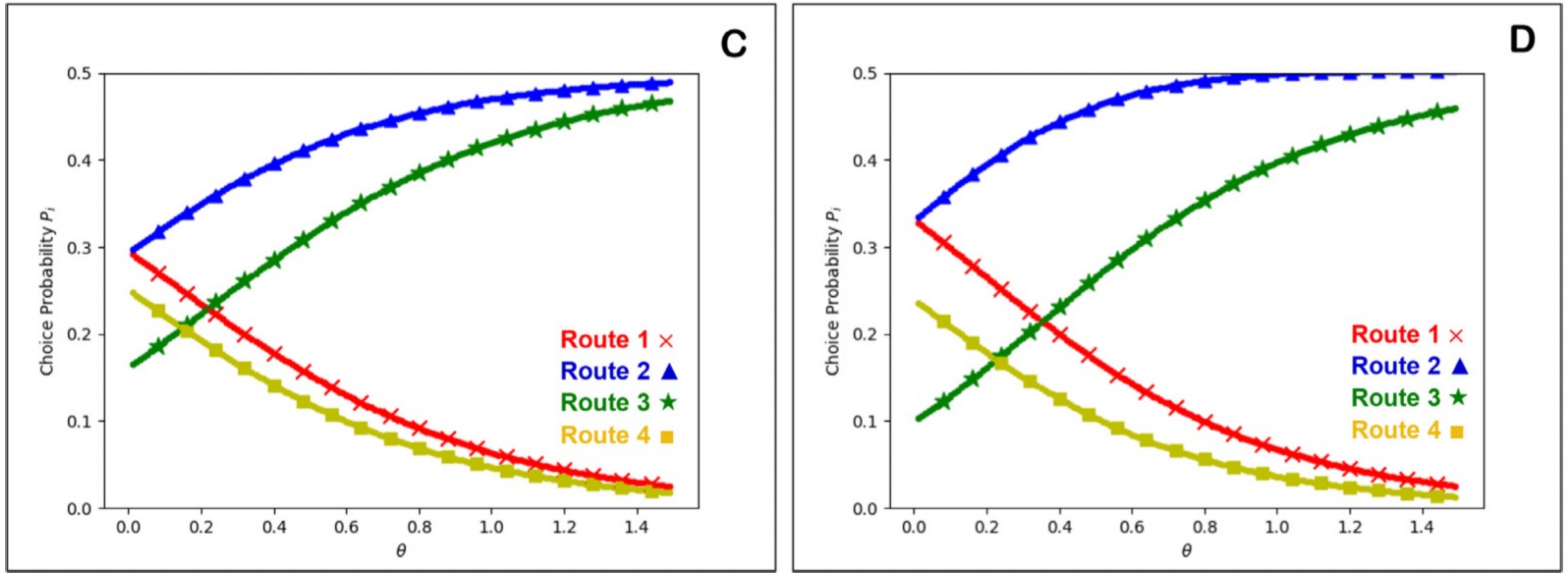


Fig. 6. (A): Example network 4. Example network 4: Route choice probabilities for increasing θ –(B): GPSL, $\beta = 1$, $\lambda = 4$. (C): GPSL′, $\lambda = \theta$, $\beta = 1$. (D): GPSL′, $\lambda = \theta$, $\beta = 2$.

An increase in β increases the sensitivity to distinctiveness within the probability relation. (C&D)
Attractiveness due to distinctiveness is not considered within path size contribution factors.

APSL

The Adaptive Path Size Logit model
The paper thus propose a fully internally consistent PSL model where all components assess the feasibility of routes according to its relative attractiveness due to **travel cost** *and* **distinctiveness**.

Proposed Definition:

The APSL choice probabilities, $\boldsymbol{P}^*$, (for a choice set of size N) are a solution to the fixed-point problem $\boldsymbol{P} = \mathbf{G}(\mathbf{g}(\boldsymbol{\gamma}^{APS}(\boldsymbol{P})))$, where:

$$G_i\Big(g_i\big(\gamma^{APS}(\boldsymbol{P})\big)\Big) = \tau + (1 - N\tau)\cdot g_i\big(\boldsymbol{\gamma}^{APS}(\boldsymbol{P})\big), \tag{11}$$

$$g_i\big(\boldsymbol{\gamma}^{APS}(\boldsymbol{P})\big) = \frac{e^{-\theta c_i + \beta \ln\left(\gamma_i^{APS}(\boldsymbol{P})\right)}}{\sum_{j\in R} e^{-\theta c_j + \beta \ln\left(\gamma_j^{APS}(\boldsymbol{P})\right)}} = \frac{\left(\gamma_i^{APS}(\boldsymbol{P})\right)^{\beta} e^{-\theta c_i}}{\sum_{j\in R}\left(\gamma_j^{APS}(\boldsymbol{P})\right)^{\beta} e^{-\theta c_j}}, \tag{12}$$

$$\gamma_i^{APS}(\boldsymbol{P}) = \sum_{a\in A_i} \frac{t_a}{c_i}\frac{P_i}{\sum_{k\in R} P_k \delta_{a,k}} = \sum_{a\in A_i}\frac{t_a}{c_i}\frac{1}{\sum_{k\in R}\left(\frac{P_k}{P_i}\right)\delta_{a,k}}, \tag{13}$$

$$\forall i \in R, \quad \forall \boldsymbol{P} \in D^{(\tau)},\ \theta > 0,\ \beta \geq 0,\ 0 < \tau \leq \frac{1}{N},$$

$$D^{(\tau)} = \left\{\boldsymbol{P} \in \mathbb{R}^N_{>0} : \tau \leq P_i \leq (1 - (N-1)\tau), \quad \forall i \in R, \sum_{j=1}^{N} P_j = 1\right\}.$$

APSL

The Adaptive Path Size Logit model
Motivation:

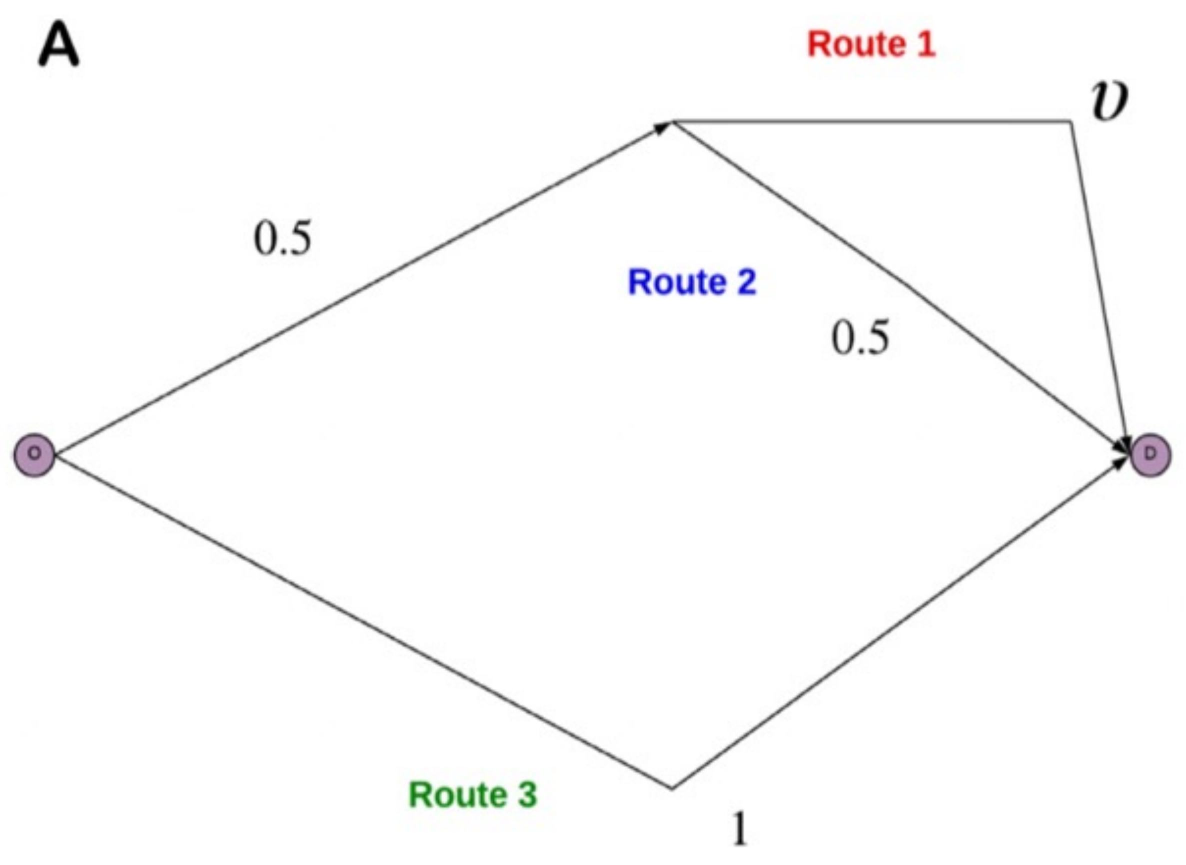
G_i is a fixed-point function, and its construction was motivated by some desired behaviours, as well as some required properties for proving existence and uniqueness:

1. G_i must map into itself.
2. G_i must be continuous for all $\boldsymbol{P}$.
3. G_i must be continuously differentiable with respect to $\boldsymbol{P}$ for all $\boldsymbol{P}$.
4. The domain of G_i must be closed and bounded.
5. The domain of G_i must not allow for zero choice probabilities.
6. G_i should be able to approximate g_i.
7. The domain of G_i should allow for choice probabilities to be approximately close to zero.

APSL

The Adaptive Path Size Logit model

APSL Key Property 1: *Unrealistic routes have a diminished impact upon the choice probabilities of realistic routes when links are shared.*



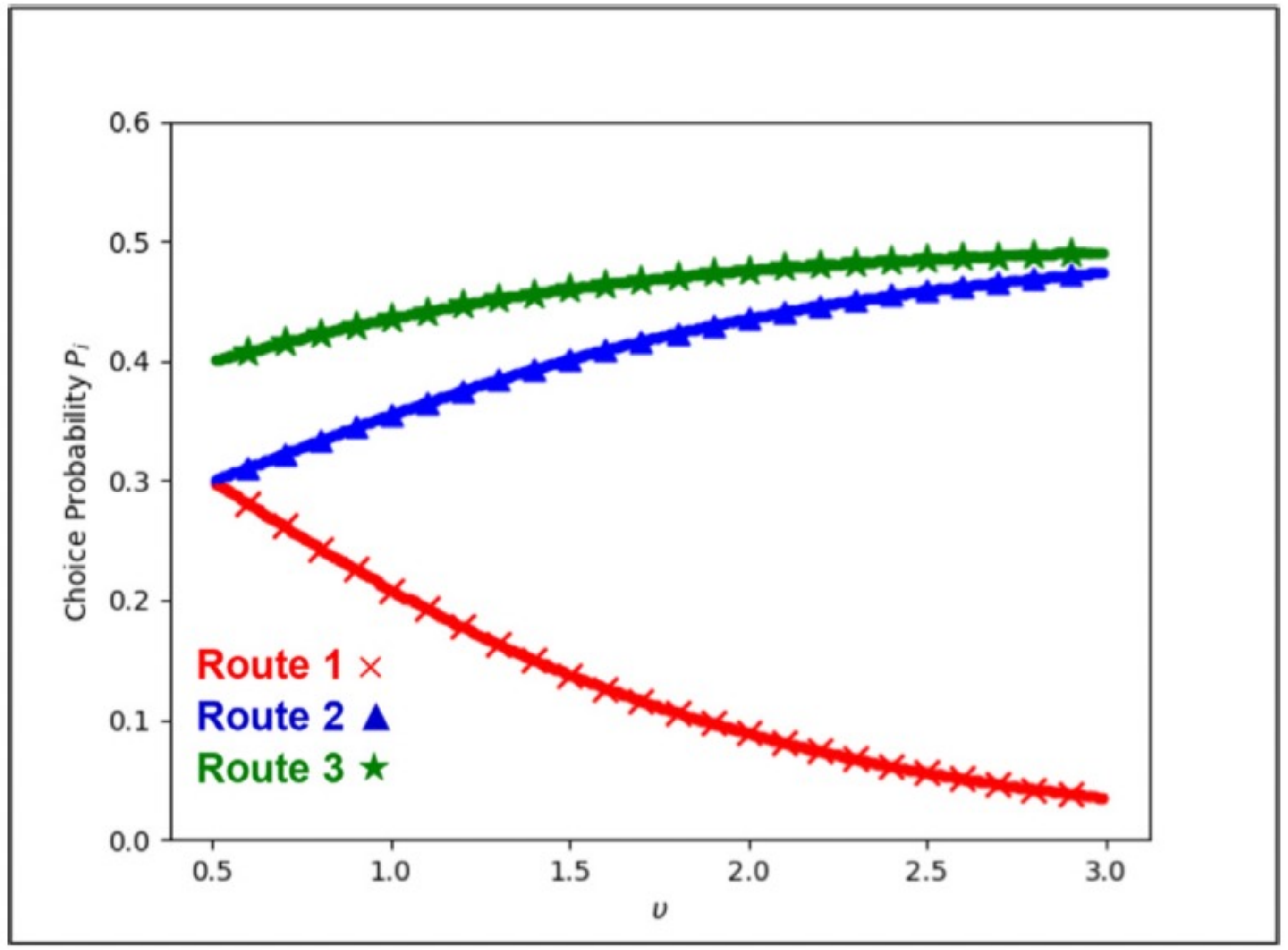


Fig. 7. Example network 1: APSL route choice probabilities for increasing υ; $\theta = \beta = 1$.

APSL

The Adaptive Path Size Logit model
APSL Key Property 2: *APSL path size terms aren't highly sensitive to small differences in route travel cost.*

Table 1
Example network 2: choice probabilities for different PSL models; $\theta = \beta = 1$.

	PSL (= GPSL $\lambda = 0$)	GPSL ($\lambda = 10$)	GPSL ($\lambda = 400$)	APSL
P_1	0.329	0.294	0.222	0.297
P_2	0.332	0.301	0.374	0.301
P_3	0.332	0.399	0.398	0.397
P_4	0.007	0.006	0.006	0.006

the APSL choice probabilities resemble the 'optimised' GPSL choice probabilities for λ = 10.

APSL

The Adaptive Path Size Logit model

APSL Key Property 3: *Internally consistent assessment of the feasibility of routes.*

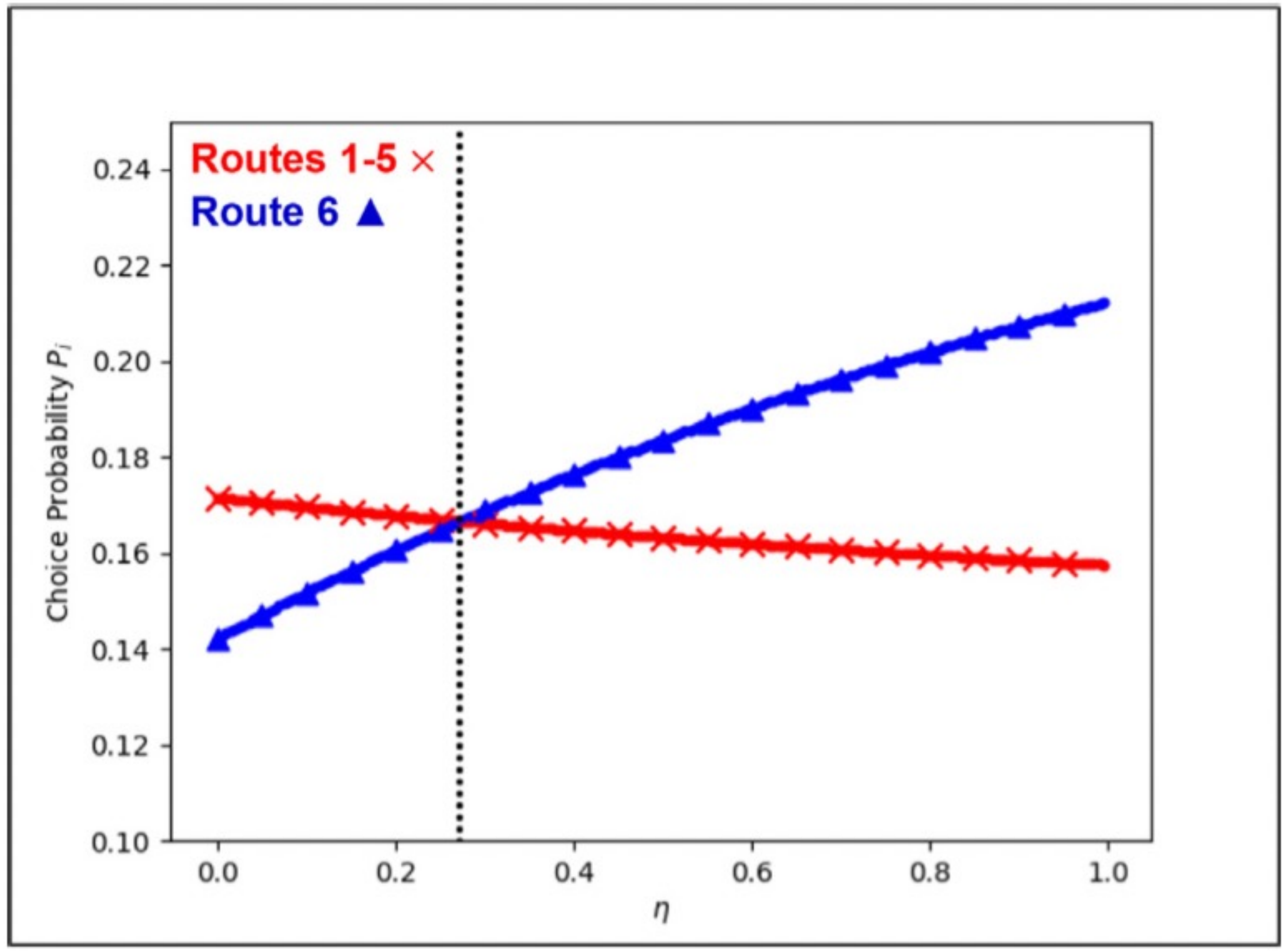


Fig. 8. Example network 3: APSL route choice probabilities for increasing η; $\theta = \beta = 1$.

APSL

The Adaptive Path Size Logit model

APSL Key Property 4: *Internally consistent scaling parameters.*

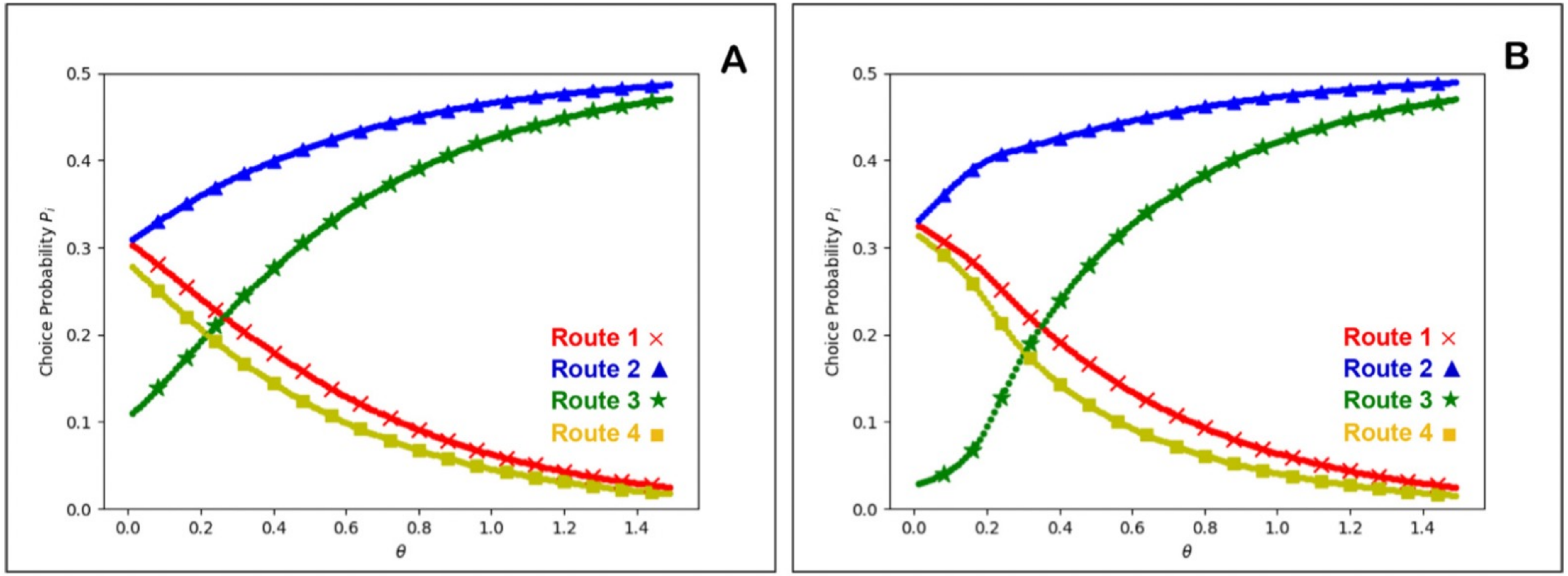


Fig. 9. Example network 4: APSL route choice probabilities for increasing θ – A: $\beta = 1$. B: $\beta = 1.4$.

APSL

The Adaptive Path Size Logit model

Solution method: this paper utilises the simplest fixed-point algorithm available: the Fixed-Point Iteration Method (FPIM)

$$\ln\left(\sum_{i\in R}\left|P_i^{(s+1)} - P_i^{(s)}\right|\right) < \ln\left(10^{-\xi}\right),$$

Existence and uniqueness of APSL solutions

…

Estimation

The Adaptive Path Size Logit model
Definition with multiple OD movements
Model definition
Likelihood formulation:

$$LL(\boldsymbol{\alpha}, \theta, \beta | \boldsymbol{x}) = \ln\left(\prod_{z=1}^{Z} P^*_{m_z,x_z}(\boldsymbol{t}(\boldsymbol{w}; \boldsymbol{\alpha}), \theta, \beta)\right) = \sum_{z=1}^{Z} \ln\left(P^*_{m_z,x_z}(\boldsymbol{t}(\boldsymbol{w}; \boldsymbol{\alpha}), \theta, \beta)\right).$$

Estimation procedure:

Algorithm 1 Pseudo-code for estimating the APSL model.

Step 1: Initialisation. For each route observation $z = 1, \ldots, Z$, generate the corresponding universal choice set and store the link attributes and link-route information. Define an initial set of parameter values $(\tilde{\boldsymbol{\alpha}}^{(1)}, \tilde{\theta}^{(1)}, \tilde{\beta}^{(1)})$ for MLE, and set $n = 1$.
Step 2: Recalculate choice probabilities and LL. Given the set of parameter values $(\tilde{\boldsymbol{\alpha}}^{(n)}, \tilde{\theta}^{(n)}, \tilde{\beta}^{(n)})$ for iteration n, calculate the link costs $\boldsymbol{t}(\boldsymbol{w}; \tilde{\boldsymbol{\alpha}}^{(n)})$ and solve each of the fixed-point problems

$$\boldsymbol{P}_{m_z} = \boldsymbol{G}_{m_z}(\boldsymbol{g}_{m_z}(\boldsymbol{c}_{m_z}(\boldsymbol{t}(\boldsymbol{w}; \tilde{\boldsymbol{\alpha}}^{(n)})), \boldsymbol{\gamma}^{APS}_{m_z}(\boldsymbol{t}(\boldsymbol{w}; \tilde{\boldsymbol{\alpha}}^{(n)}), \boldsymbol{P}_{m_z}); \tilde{\theta}^{(n)}, \tilde{\beta}^{(n)}))$$

for $z = 1, \ldots, Z$. Given the fixed-point choice probability solutions $P^*_{m_z,x_z}$ for each of the route observations $z = 1, \ldots, Z$, calculate the Log-Likelihood $LL^{(n)}(\tilde{\boldsymbol{\alpha}}^{(n)}, \tilde{\theta}^{(n)}, \tilde{\beta}^{(n)} | \boldsymbol{x})$ for iteration n.
Step 3: Compute new set of parameters. Based on $LL^{(s)}$ and the associated parameters $(\tilde{\boldsymbol{\alpha}}^{(s)}, \tilde{\theta}^{(s)}, \tilde{\beta}^{(s)})$ for all $s \leq n$, compute a new set of parameters $(\tilde{\boldsymbol{\alpha}}^{(n+1)}, \tilde{\theta}^{(n+1)}, \tilde{\beta}^{(n+1)})$ to test in the following iteration.
Step 4: Stopping criteria. If $|LL^{(n)} - LL^{(n-1)}| < \zeta$, stop. Otherwise, set $n = n + 1$ and return to Step 2.

Simulation

Three steps for the experiment:

- Postulate a true APSL choice model including specification and parameter values.
- Sample a set of observed route choices according to the true model using the specified link travel costs.

-Apply MLE approach to obtain parameter estimates based on the observed route choices.

Sioux Falls application

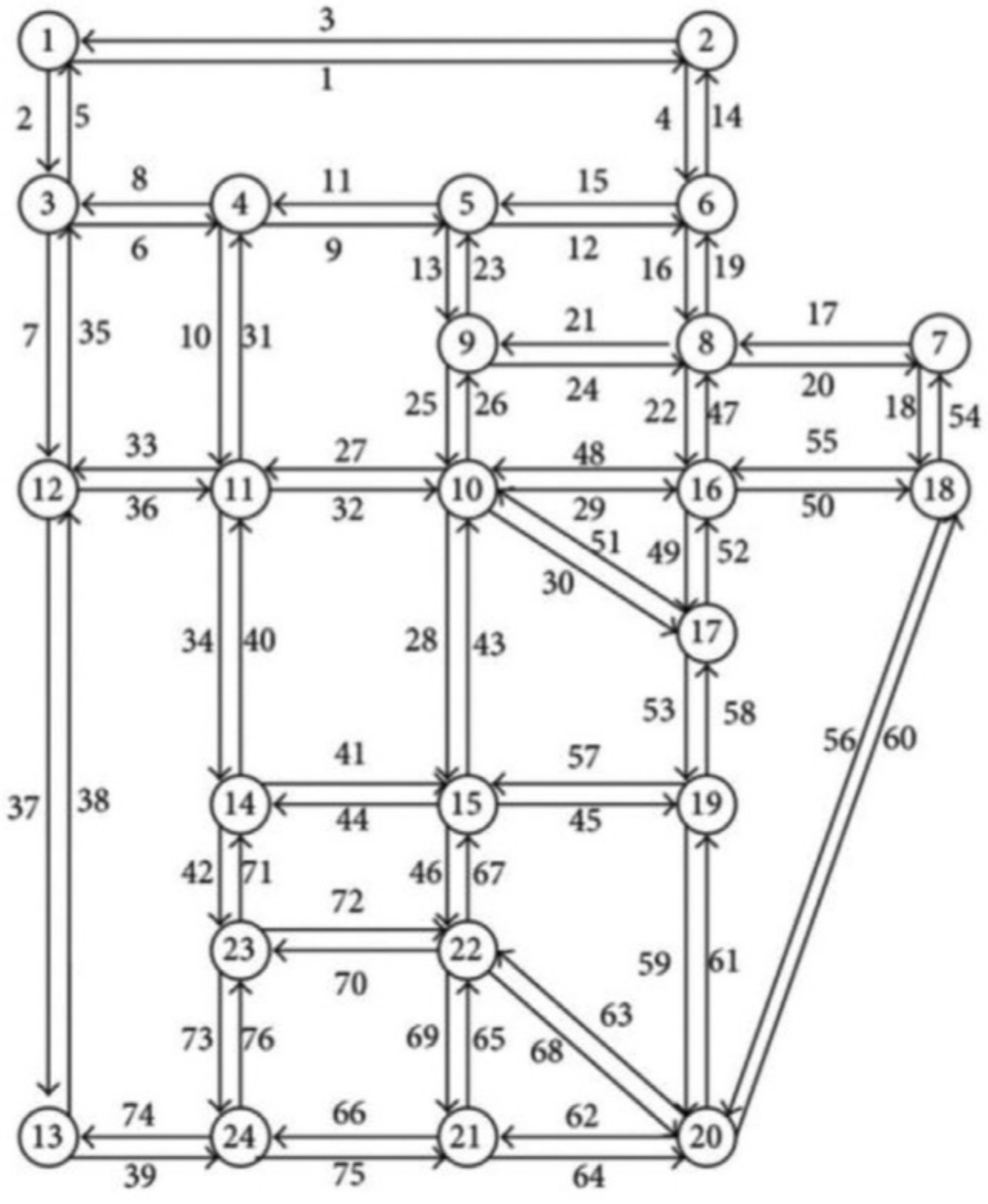


Fig. 13. Sioux Falls network.

$$t_a(\boldsymbol{w}_a; \boldsymbol{\alpha}) = w_{a,1} \cdot \alpha_1,$$

α_1, β, θ=1

Simulation

Experimental result:

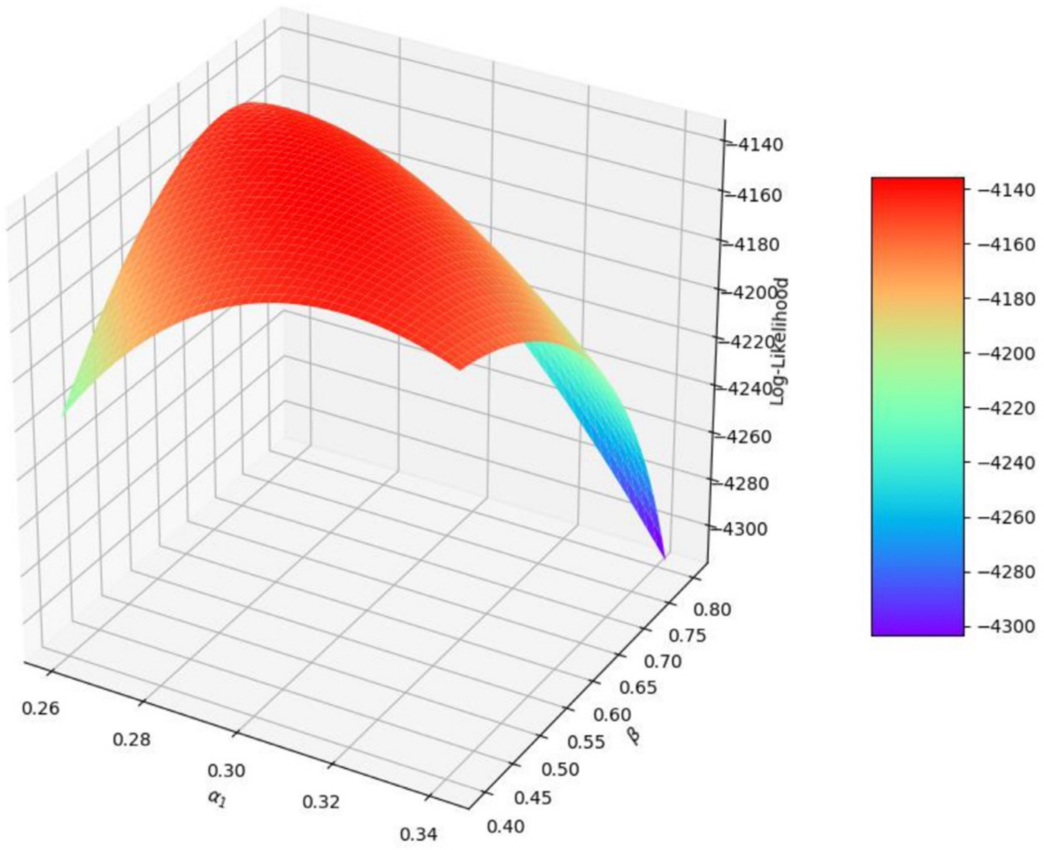


Fig. 14. Sioux Falls simulation study: Log-Likelihood surface; $\theta_{true} = 0.3$, $\beta_{true} = 0.6$, $Z = 2000$.

Table 2
Sioux Falls simulation study: Stability of estimated parameters across multiple experiment replications; $Z = 1000$, $q = 100$.

α_1^{true}	β_{true}	$\hat{\alpha}_1$			$\hat{\beta}$			$cov(\hat{\boldsymbol{\alpha}}_1, \hat{\boldsymbol{\beta}})$
		μ	σ	MSE	μ	σ	MSE	
0.1	0.8	0.1000	0.0032	0.0000	0.8003	0.0136	0.0002	−0.00003
0.2	0.7	0.2016	0.0068	0.0001	0.6972	0.0305	0.0009	−0.00014
0.3	0.6	0.3021	0.0105	0.0001	0.5876	0.0485	0.0025	−0.00034
0.4	0.4	0.4012	0.0153	0.0002	0.3984	0.0817	0.0067	−0.00096

Simulation

Experimental result:

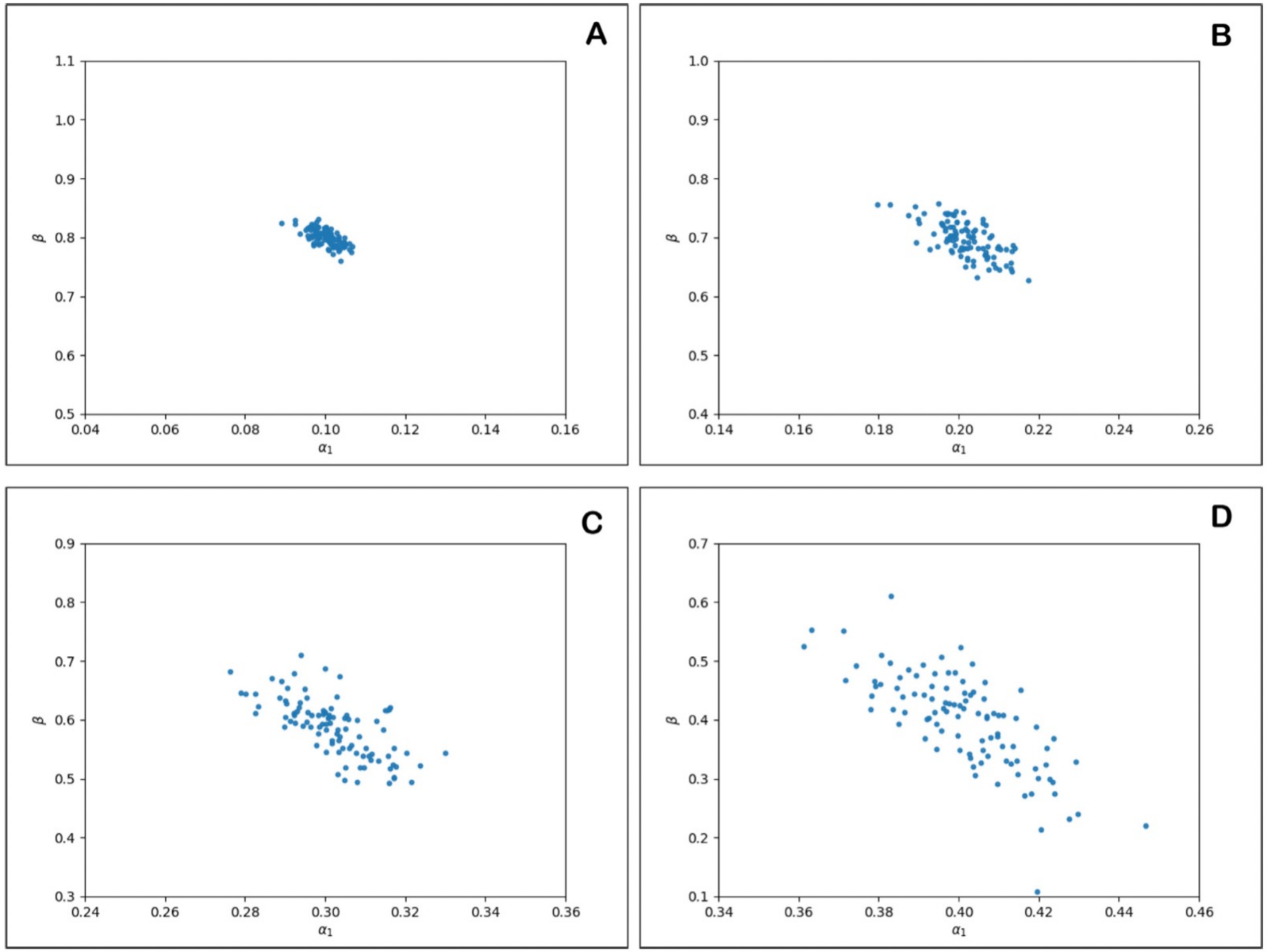


Fig. 15. Sioux Falls simulation study: Distribution of estimated parameters after multiple experiment replications; $Z = 1000$, $q = 100$ – A: $\alpha_1^{true} = 0.1$ $\beta_{true} = 0.8$. B: $\alpha_1^{true} = 0.2$ $\beta_{true} = 0.7$. C: $\alpha_1^{true} = 0.3$ $\beta_{true} = 0.6$. B: $\alpha_1^{true} = 0.4$ $\beta_{true} = 0.4$.

Simulation

Computation analysis:

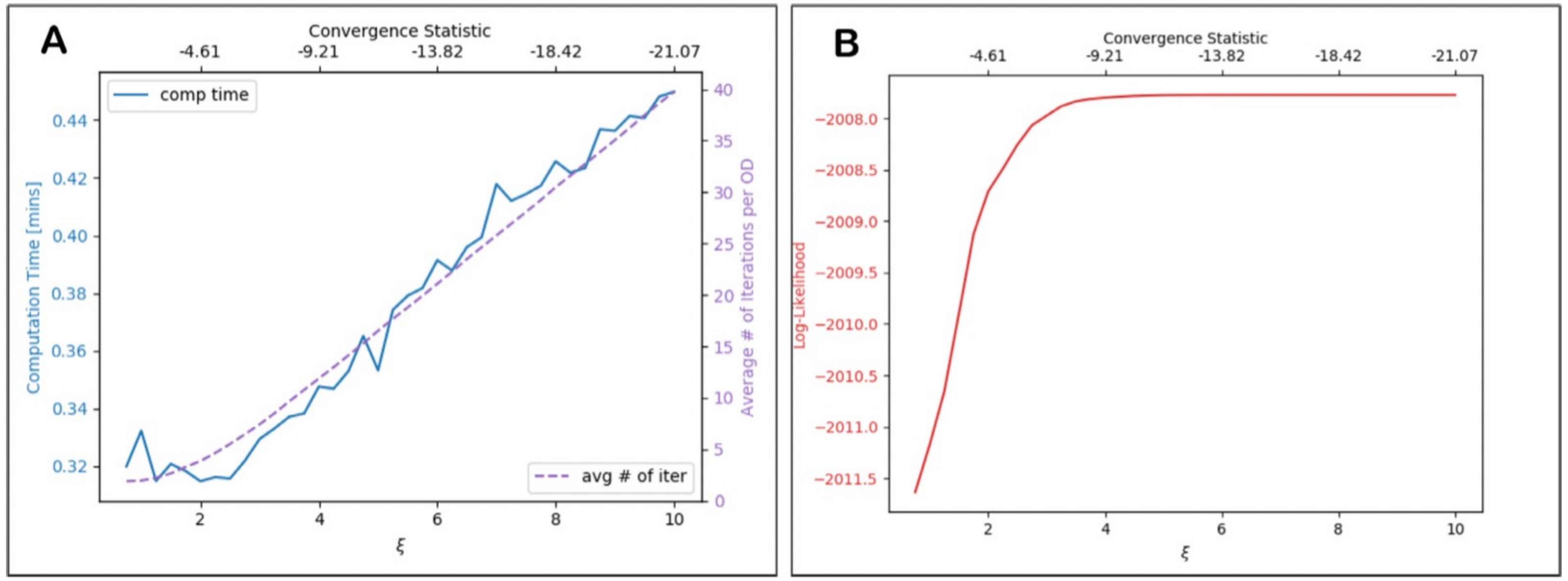


Fig. 16. Sioux Falls simulation study: Computational statistics for calculating APSL Log-Likelihoods as the APSL choice probability convergence parameter ξ is increased –(A): Average number of fixed-point iterations per OD / computation time [mins]. (B): Log-Likelihood value.

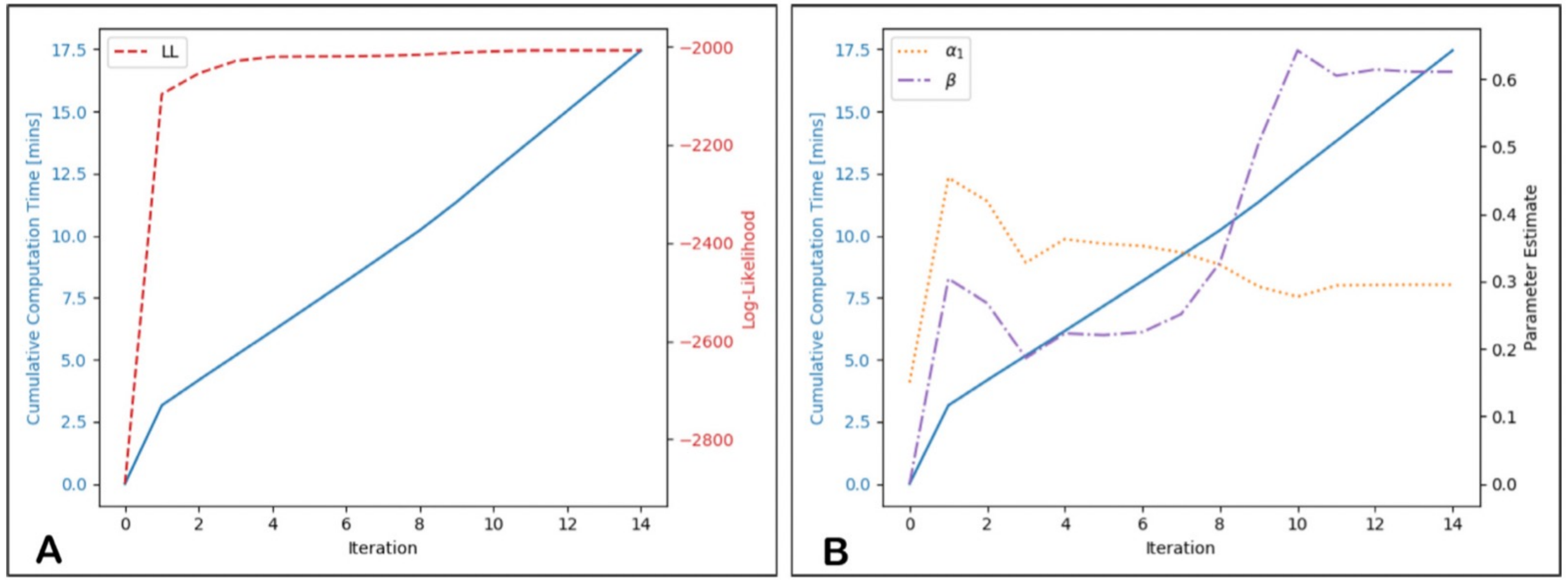


Fig. 18. Sioux Falls simulation study: Cumulative computation time at each iteration of a single MLE, and MLE statistics; $\xi = 7$ –(A): Log-Likelihood. (B): Parameter estimates.

Case study

The data has been collected amongst drivers in the eastern part of Denmark in 2011, and includes a total of 17,115 observed routes. After a filtering to include only trips where the sum of travel time (in minutes) and length (in km) is at least 10, a total of 8696 observations remain.

The GPS traces are map matched to a network, for which corresponding time-of-day dependant travel times are available on the entire network. The network is large-scale, representing all of Denmark, and thus includes 34,251 links.

The authors approximate the universal choice set by generating a choice set for each observed route by applying the doubly stochastic approach. Up to 10 0 unique paths are generated for each observation.

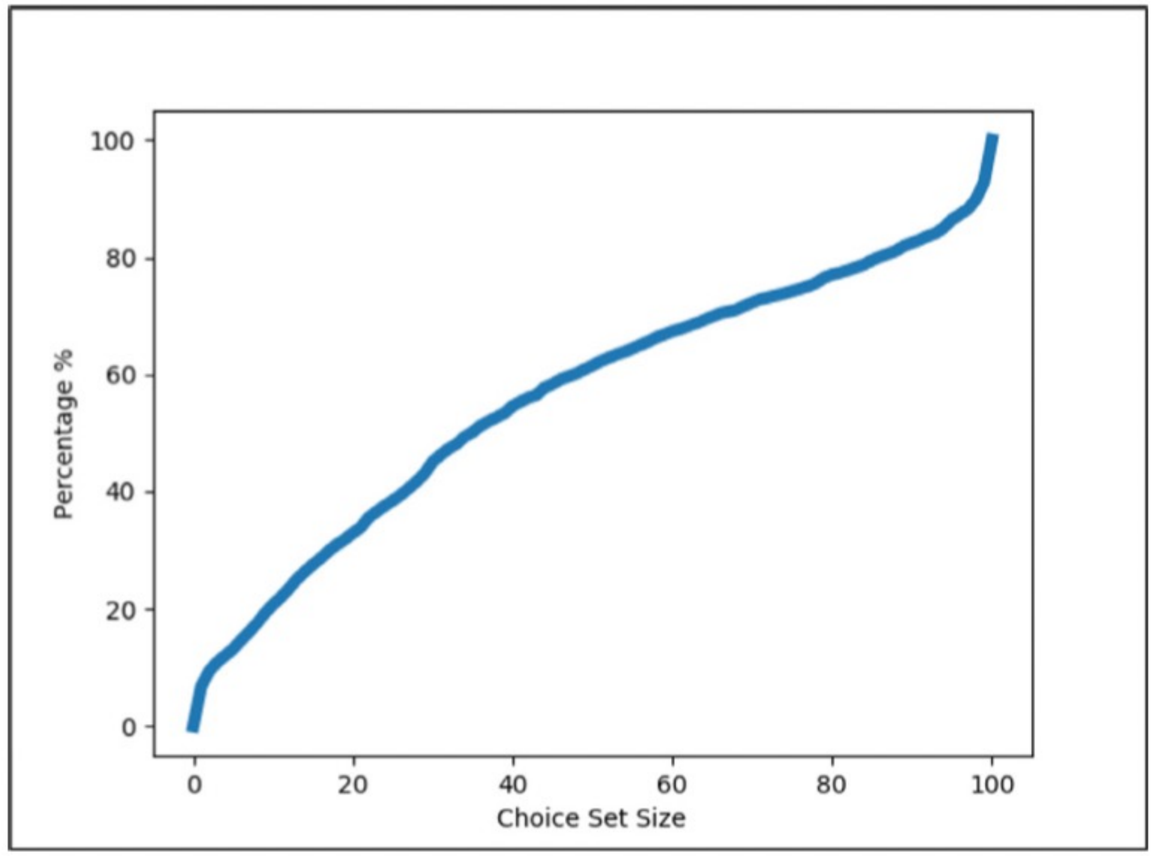


Fig. 22. Real-life case-study: Cumulative distribution of the choice set sizes for the 8696 observations.

Case study

Travel cost of link a:

$$t_a(\mathbf{w}_a; \boldsymbol{\alpha}) = w_{a,1} \cdot \alpha_1 + w_{a,2} \cdot \alpha_2,$$

α_1 is parameter of travel time in minutes and α_2 is parameter of length in kilometers.

$$c_{m,i}(\mathbf{t}(\mathbf{w}; \boldsymbol{\alpha})) = \sum_{a \in A_{m,i}} t_a(\mathbf{w}_a; \boldsymbol{\alpha}) = \sum_{a \in A_{m,i}} (w_{a,1} \cdot \alpha_1 + w_{a,2} \cdot \alpha_2) = \alpha_1 \sum_{a \in A_{m,i}} w_{a,1} + \alpha_2 \sum_{a \in A_{m,i}} w_{a,2}.$$

The model requires the specification of four parameters:

$\alpha_1, \alpha_2, \beta, \theta$=1

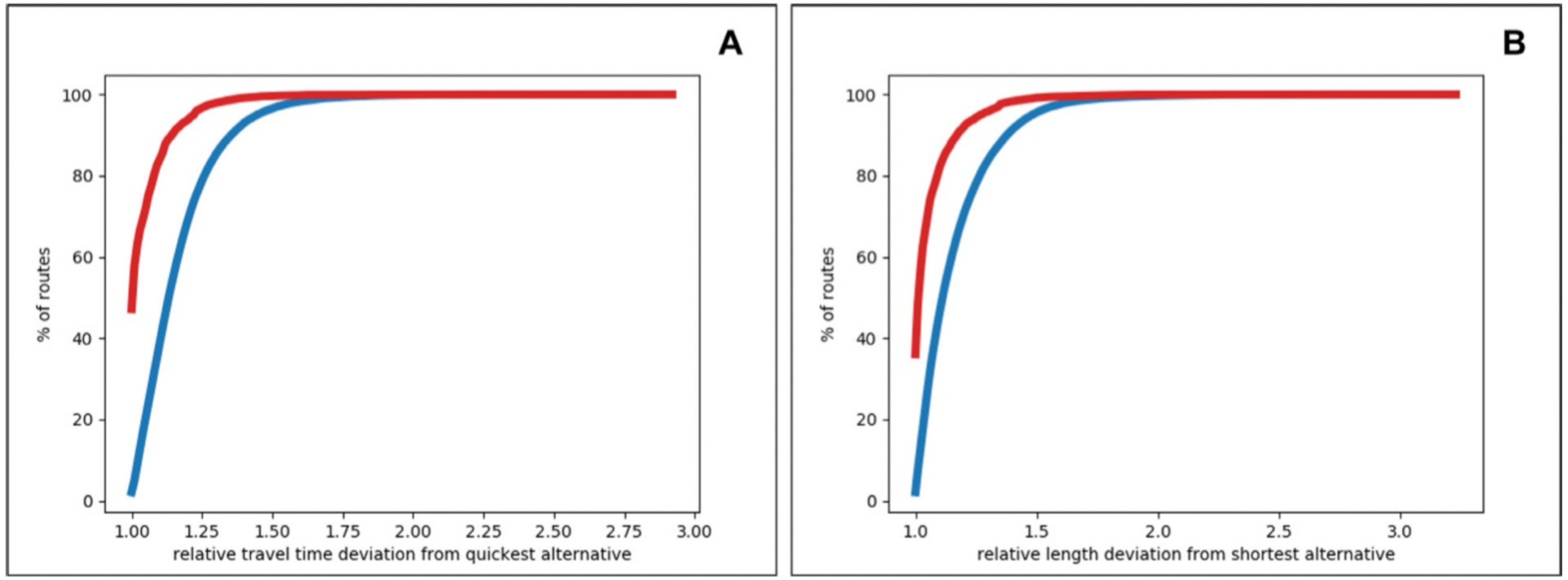


Fig. 23. Real-life case-study: Relative deviations away from quickest/shortest routes in the choice sets for the observed routes (red) and alternative routes generated (blue) – (A): Travel time. (B): Length. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Case study

Results:

Loglikelihood surfaces:

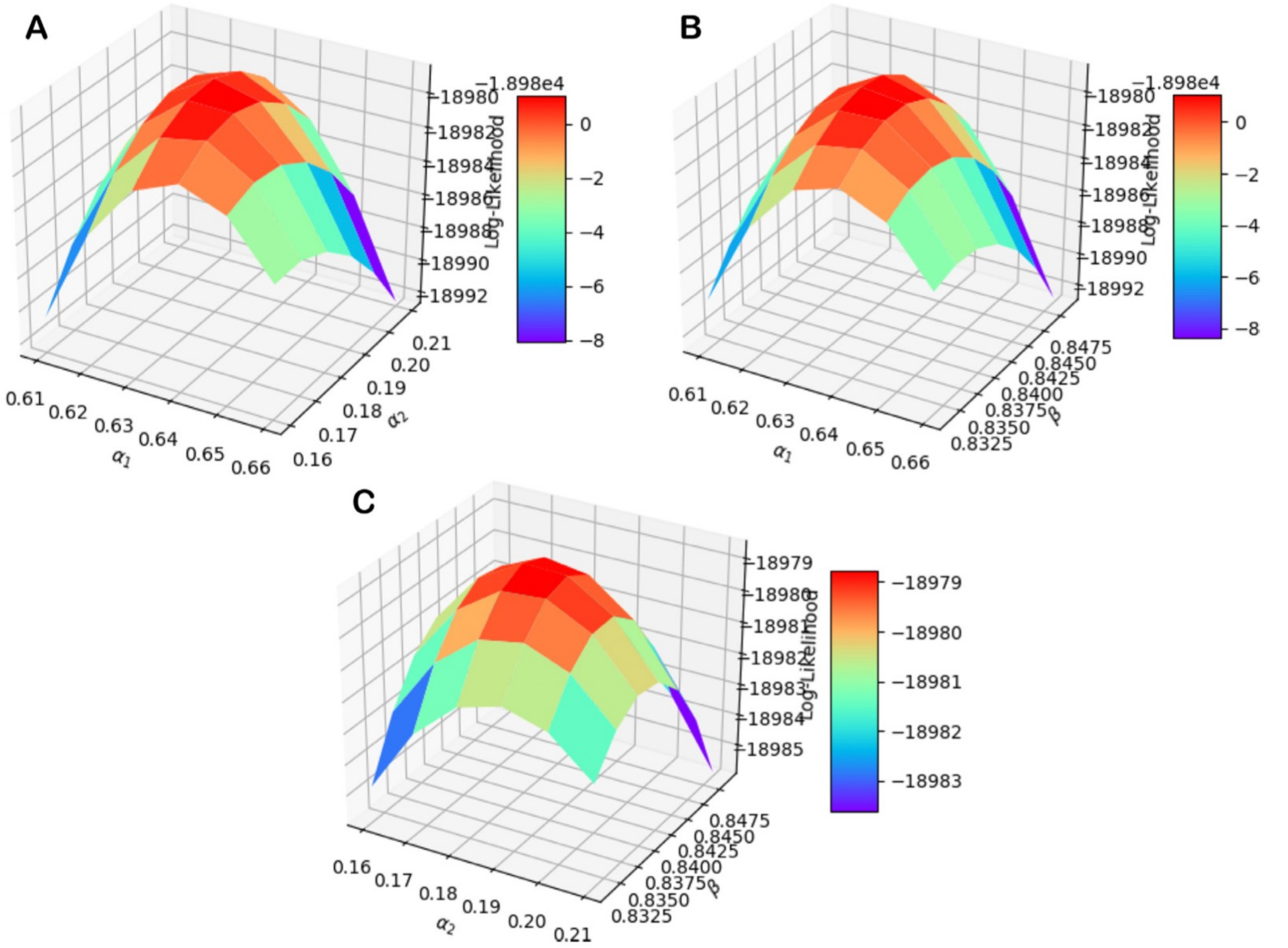


Fig. 24. Real-life case-study: APSL Log-Likelihood surfaces around parameter estimates in Table 3 – A: α_1, α_2. B: α_1, β. C: α_2, β.

Case study

Results:
Comparison of models

Table 4
Real-life case-study: Estimation results and Log-Likelihood values.

	$\hat{\alpha}_1$	$\hat{\alpha}_2$	$\hat{\beta}$	$\hat{\lambda}$	LL
MNL	0.777	0.330			−21,308
PSL	0.966	0.306	1.347		−20,581
GPSL	0.415	0.085	1.186	91.95	−17,874
$GPSL'_{(\lambda=\theta)}$	0.691	0.154	1.807		−19,152
APSL	0.633	0.184	0.840		−18,978

Table 5
Real-life case-study: Comparison of fit between models based on non-nested Horowitz type tests.

h	K_h	$\bar{\rho}_h^2$	y_h	Y_h
PSL	3	0.27006	0.02573	−38.1144
GPSL	4	0.36604	0.12171	−82.8673
$GPSL'_{(\lambda=\theta)}$	3	0.32074	0.07642	−65.6583
APSL	3	0.32690	0.08258	−68.2537

Case study

Results:

Comparison of models

Table 6

Real-life case-study: Comparison of fit between models based on penalised-likelihood criteria.

	AIC	BIC	CAIC
MNL	42,621	42,635	42,637
PSL	41,169	41,190	41,193
GPSL	35,756	35,784	35,788
$GPSL'_{(\lambda = \theta)}$	38,311	38,332	38,335
APSL	37,963	37,984	37,987

In both tests, the GPSL’($\lambda = \vartheta$) and APSL models outperform the PSL model with the same number of parameters, where APSL outperforms GPSL’($\lambda = \vartheta$). This suggests that there is value in including a measure of distinctiveness within the path size contribution factors. The GPSL model outperforms all models due to the greater flexibility the λ parameter provides.

Table 8

Real-life case-study: Estimation results including GPSL with λ restricted to $\lambda \leq 10$.

	$\hat{\alpha}_1$	$\hat{\alpha}_2$	$\hat{\beta}$	$\hat{\lambda}$	LL
MNL	0.777	0.330			−21,308
PSL	0.966	0.306	1.347		−20,581
GPSL ($\lambda \leq 10$)	0.769	0.160	1.943	10	−19,312
GPSL	0.415	0.085	1.186	91.95	−17,874
$GPSL'_{(\lambda = \theta)}$	0.691	0.154	1.807		−19,152
APSL	0.633	0.184	0.840		−18,978

Case study

Results:
Comparison of models

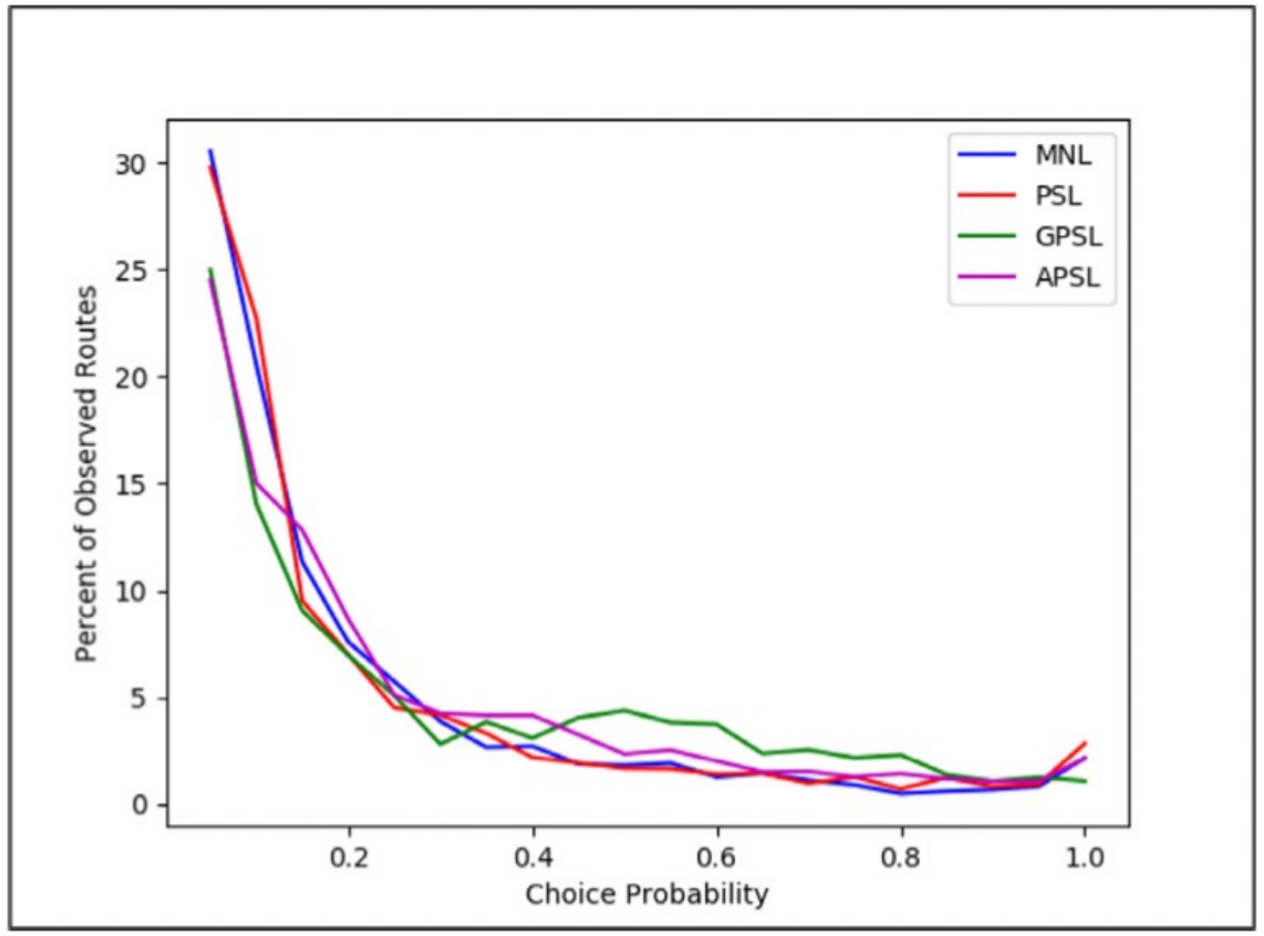
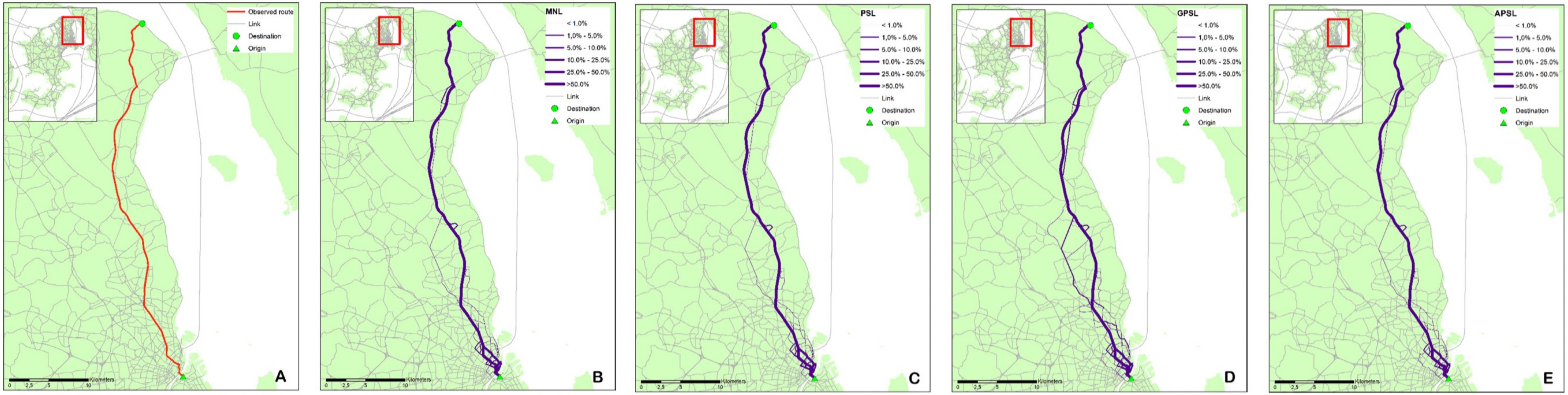


Fig. 29. Real-life case study: Choice probability distribution of the observed routes under the different estimated models.

For all models, a large percentage of the observed routes have small choice probabilities. This seems to also suggest that there are many observations where an unattractive route was chosen.

Case study

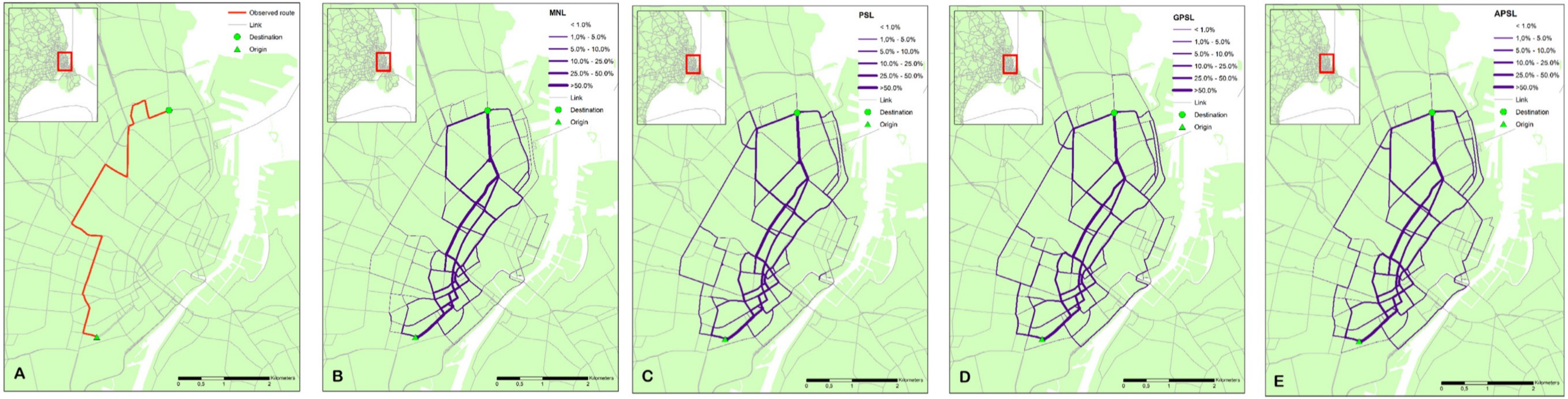
Results:
Single OD 1



OD 1: –(A): Observed route. (B): MNL. (C): PSL. (D): GPSL. (E): APSL.

Case study

Results:
Single OD 2



OD 2: –(A): Observed route. (B): MNL. (C): PSL. (D): GPSL. (E): APSL.

Case study

Results:
OD 1&2

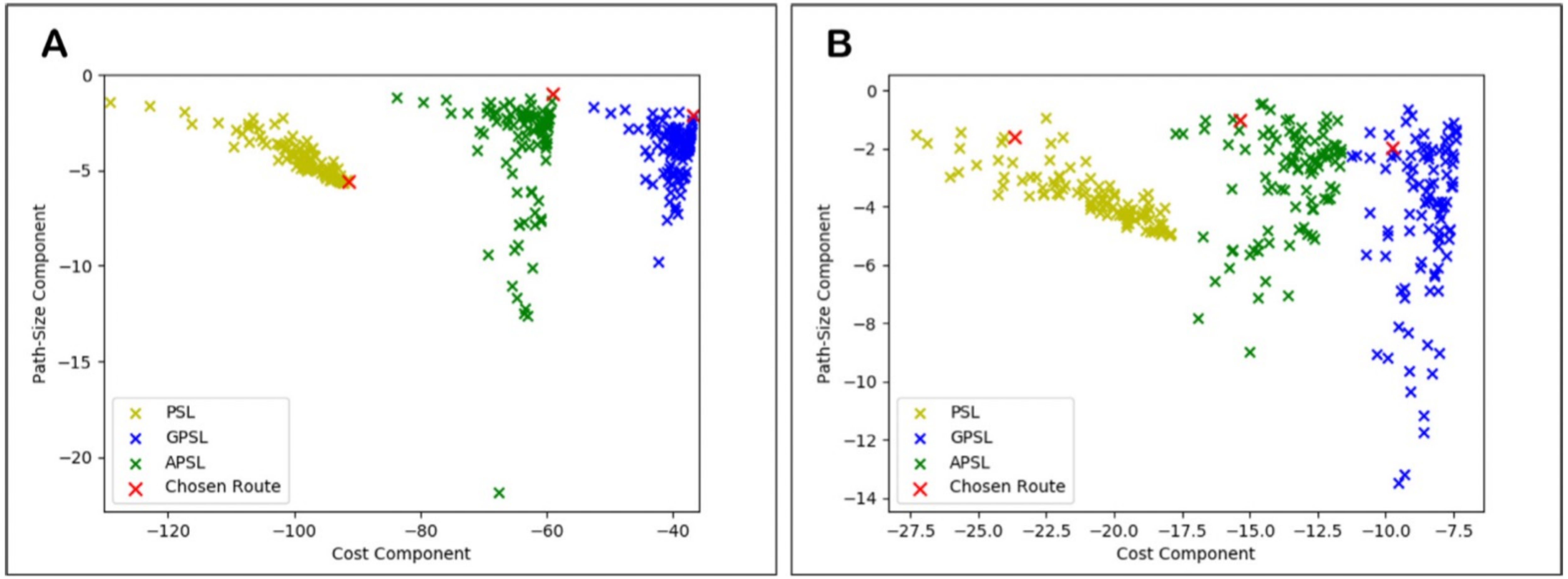


Fig. 32. Real-life case study: Cost / Path-Size components of the route utilities from PSL/GPSL/APSL – (A): OD 1. (B): OD 2.

In both cases, universal distinctiveness tends to increase as travel cost increases.
The APSL model thus provides the best fit for this observation: the path size contribution factors consider probability ratios and hence the observed route is considered the most attractive and distinct compared to its overlapping routes.

THANK YOU