# Discrete choice modeling with anonymized data

**Milos Balac**

**Sebastian Hörl**

**Basil Schmid**

**STRC conference paper 2022** **May 9, 2022**

**STRC** | **22nd Swiss Transport Research Conference**
Monte Verità / Ascona, May 18-20, 2022

# Discrete choice modeling with anonymized data

Milos Balac
IVT
ETH Zürich
CH-8093 Zurich
milos.balac@ivt.baug.ethz.ch

Sebastian Hörl
IRT SystemX

Basil Schmid
IVT
ETH Zurich

May 9, 2022

## Abstract

This paper presents an approach to estimate mode-choice models from spatially anonymized revealed preference travel survey data. We propose an algorithm that aims to find a feasible sequence of activity locations, for each individual, that minimizes the maximum error of each trip's Euclidean distance within the activity chain. The reconstructed activity locations are then used to create unchosen alternatives within the choice set for each individual. This is followed by the mode-choice model estimation. We test our approach on three large-scale travel surveys conducted in Switzerland, Île-de-France and São Paulo. We find that with our approach we can reconstruct activity locations that accurately match trip Euclidean distances, but with location errors that still provide location protection. The discrete mode-choice models estimated on the reconstructed locations perform similarly, in terms of goodness of fit and prediction, to the ones obtained from the observed activity locations.

## Keywords

anonymization, data privacy, travel survey, discrete choice model

## Suggested Citation