Emphasizing the Progression of Knowledge Over the Individual Work

Carlos Guirado, David Ory, Joan Walker

3rd Symposium on Activity-Based Modeling
TUM Science & Study Center Raitenhaslach
December 2024

IATBR *"Are we in crisis?"* workshop feedback

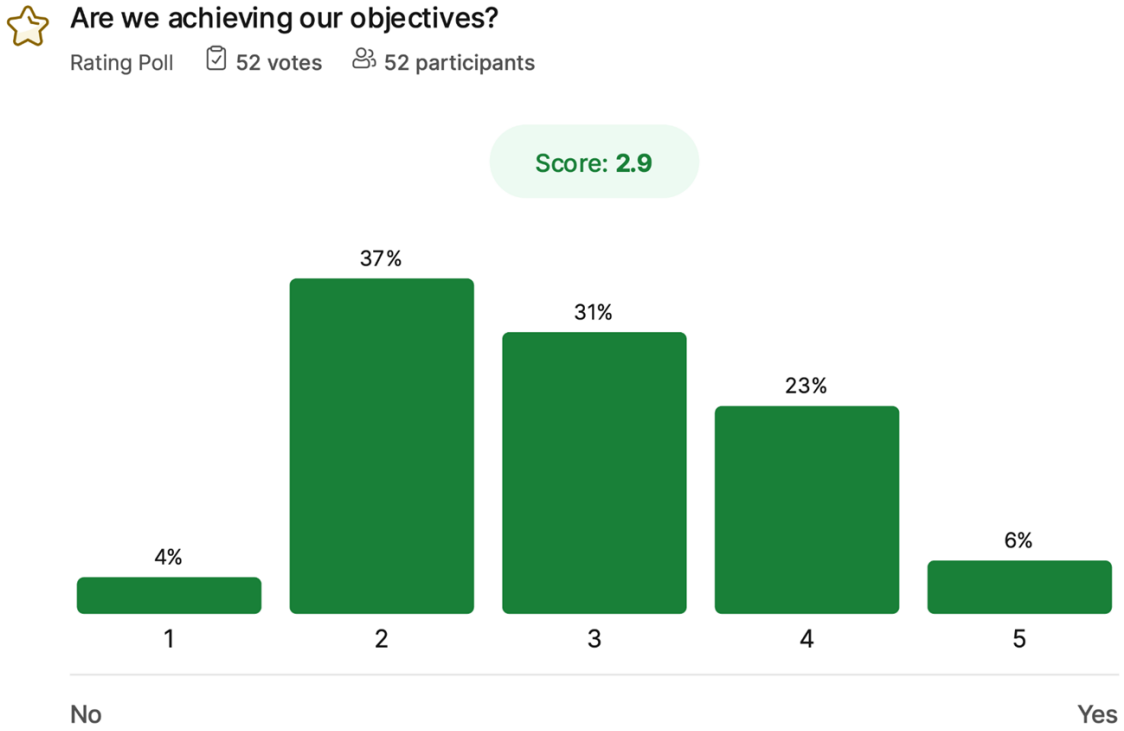
We asked:

"What is the objective of our field?"

Majority responded:

"To inform policy and decision-making"

Are we getting there?



Lessons from other fields

Machine learning:

Standardised benchmarks

Open datasets

Economics:

Robustness checks

Psychology:

Replicability

IATBR "*Are we in crisis?*" workshop feedback (II)

Top priorities identified at IATBR:

- Common datasets
- Open-source infrastructure

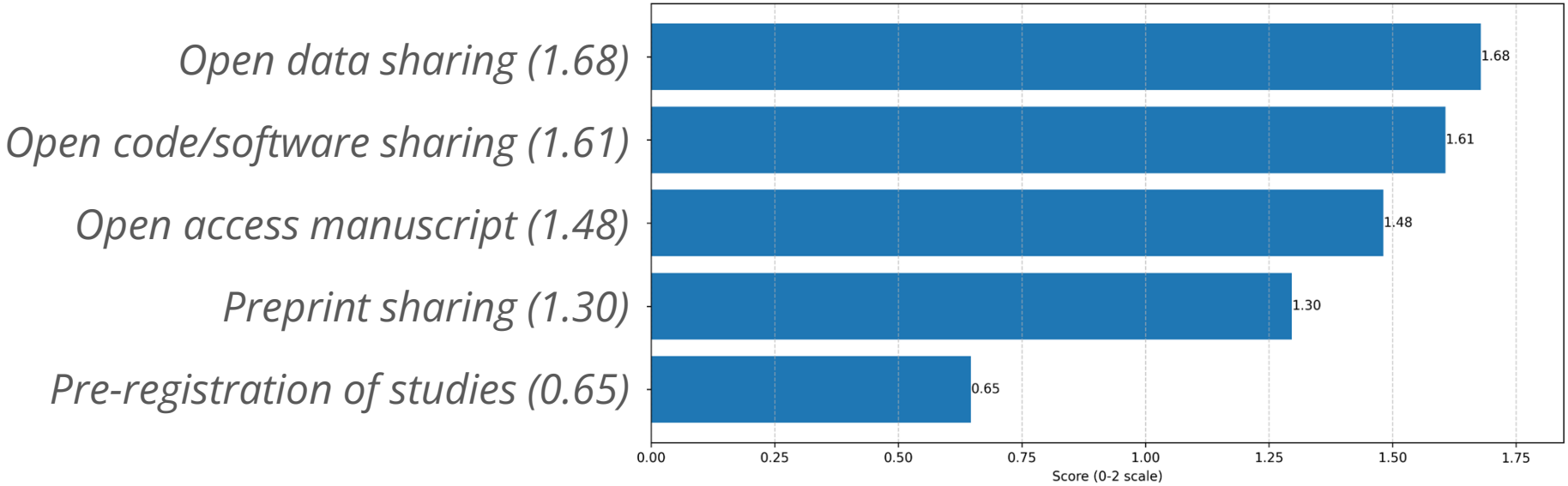
OUTLINE TODAY

- Effort #1: Open-science initiative
- Effort #2: Benchmarking infrastructure for mode choice

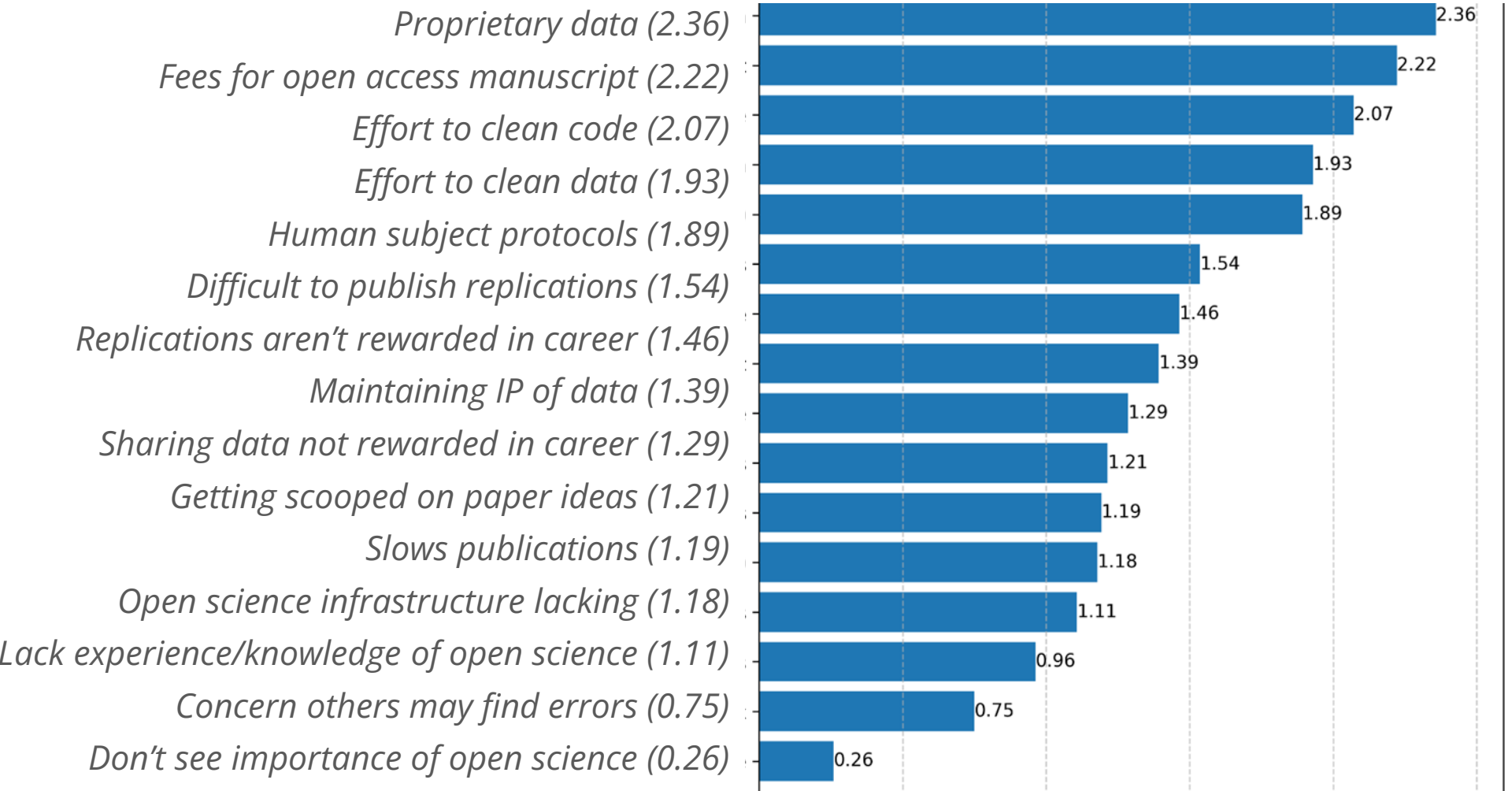
Open Science: collaborating differently

n=28 of you!

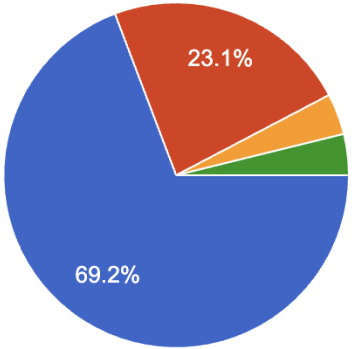
How important do you think this is for travel behavior research?
(composite score 0 = all say not important… 2 = all respond very important)



What limits your ability to practice open science? (0-3 point composite score)

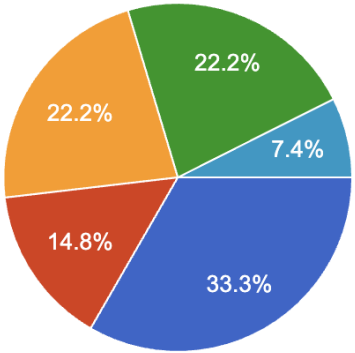


Great intentions to share…



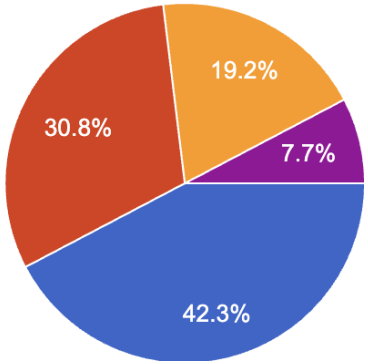
Manuscript

92% yes



Data

48% yes

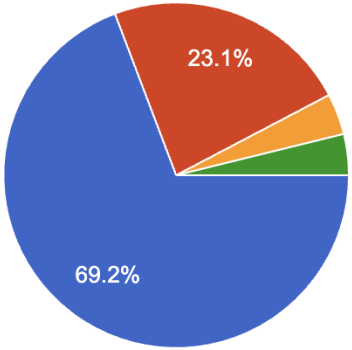


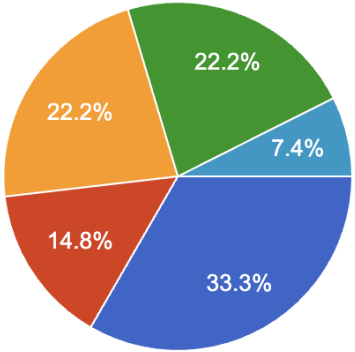
Code

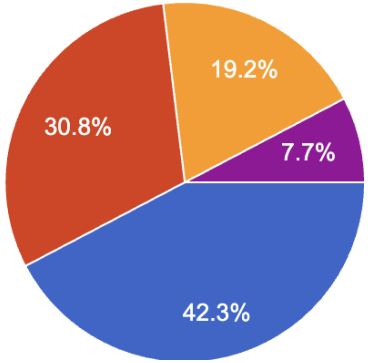
73% yes

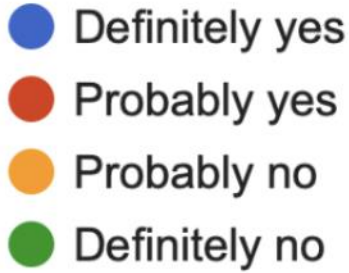
n=26 of you

… not so great results









Manuscript

92% yes

n=26 of you

65% yes

Data

48% yes

15% yes

Code

73% yes

5% yes

Effort #1
IATBR Operation 100% OPEN SCIENCE

Open data and code is first, critical step.

Targeting 11 special issues from IATBR Vienna 2024.

Objectives

- Creating processes, definitions, and practices for open science in our field
- Increasing open science participation by nudging and working through concerns of researchers (e.g., intellectual property rights)
- Documenting limitations of open science in our field (e.g., proprietary software and data, protection of human subjects)

Effort #2: Rethinking our modeling practice

What would a benchmarking suite look like for our field?

Let's tackle mode choice!

Using PMLB Python interface

Install

Most users should install PMLB from the Python Package Index (PyPI) using `pip`:

```
$ pip install pmlb
```

For access to new features or datasets that are not yet merged into an official release, you can instead install from the GitHub sources:

```
$ git clone https://github.com/EpistasisLab/pmlb
$ cd penn-ml-benchamrks/
$ pip install .
```

Usage

```
from pmlb import fetch_data

# Returns a pandas DataFrame
mushroom = fetch_data('mushroom')
mushroom.describe().transpose()
```

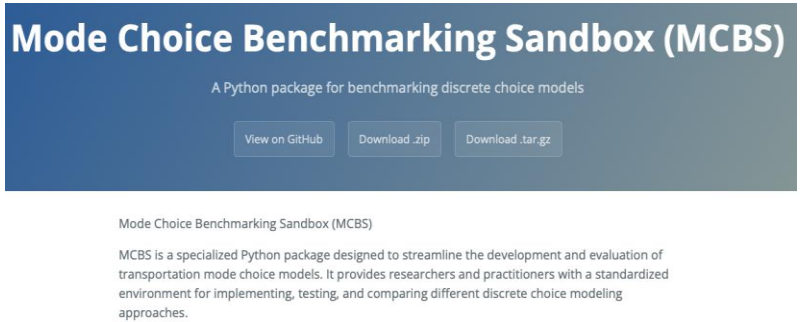
Source: Penn Machine Learning Benchmarks (PMLB)

Mode choice benchmarking sandbox (MCBS)

Python library built on Biogeme

COMPILES
- Datasets
- Methodologies

GENERATES DIRECT COMPARISONS
- Across datasets
  (Swissmetro, London, Canada)
- Across model formulations
  (Logit, Nested Logit, Mixed Logit)
- Across multiple metrics
  (Goodness of fit, Value of Time, Forecast)



https://carlosguirado.github.io/mode-choice-benchmarking-sandbox/

Easy setup

Quick Start

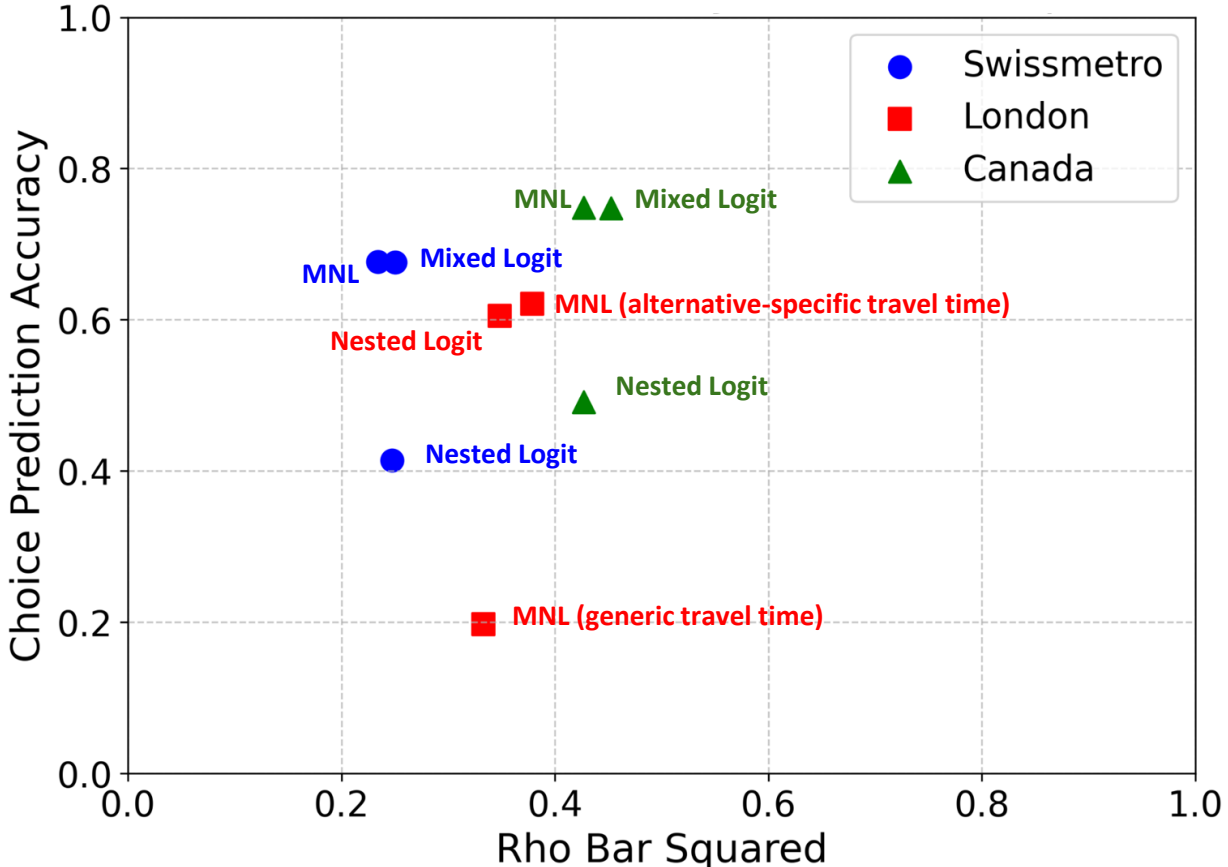
Install MCBS using pip:

```
pip install mcbs
```

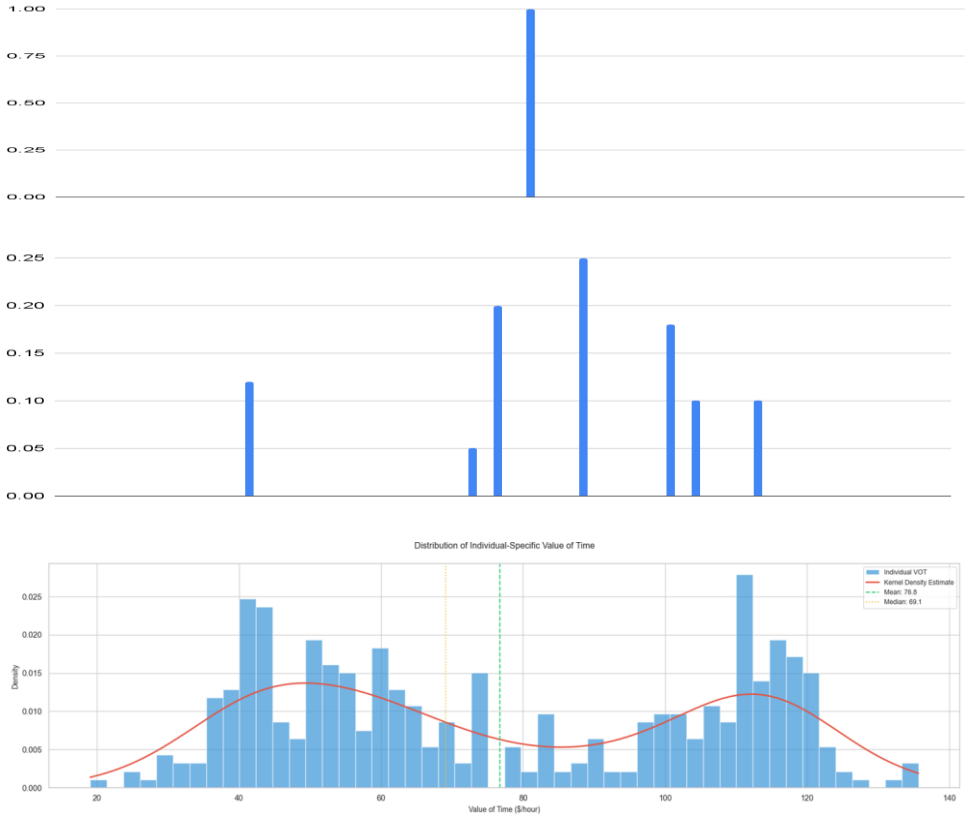
Basic usage example:

```
from mcbs import Benchmark

# Initialize benchmark with dataset
benchmark = Benchmark("swissmetro_dataset")

# Run your models
results = benchmark.run(models)
benchmark.compare_results(results)
```

METRIC #1: Goodness of Fit



METRIC #2: Value of Time



Sample density

Value of Time

MODEL A:
Logit - Homogeneous

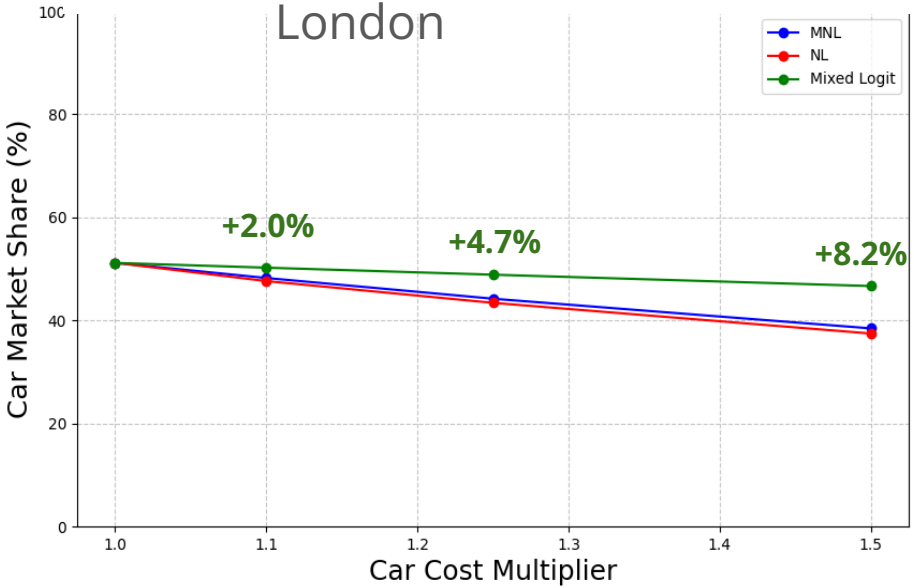
MODEL B:
Logit- Systematic heterogeneity

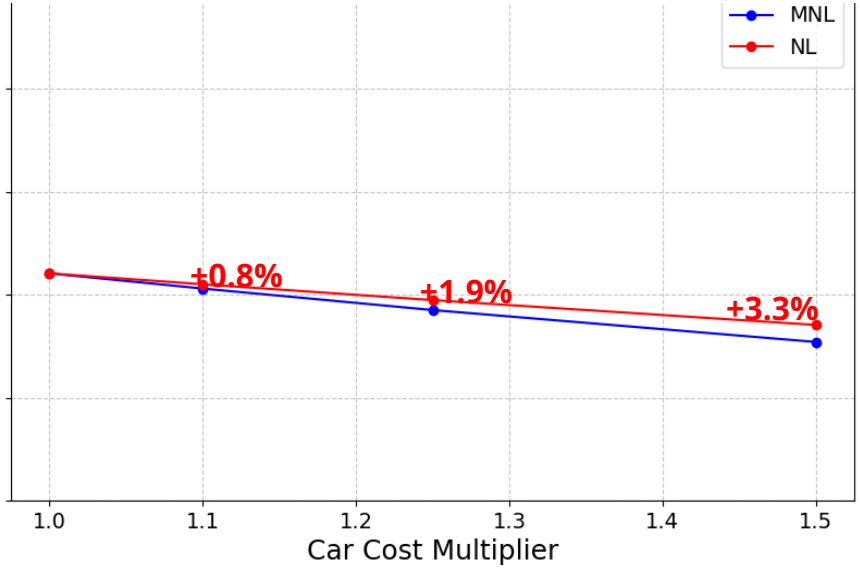
MODEL C:
Mixed logit (posterior)

METRIC #3: Forecast (Congestion Pricing)

Canada

London





Reflection

Building such a sandbox isn't easy… but Rome wasn't built in a day.

ChatGPT estimates 10,000-20,000 scholarly papers on mode choice.

- What would our (mode choice) models look like today if we had such an infrastructure or explicit comparative culture for the last 50+ years?
- How would new modes such as Shared Mobility and Autonomous Vehicles have been integrated?

What would our impact be if we prioritized credibility?

Habit models from computational neuroscience

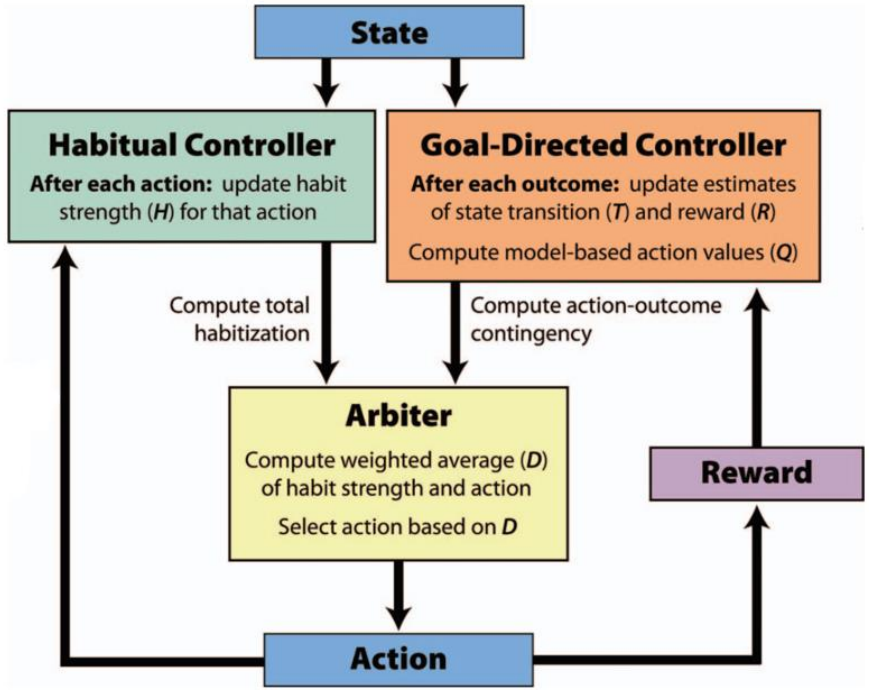
Qianhua Luo, Berkeley PhD student

Cool model!

- Dual system theory
- Models habit formation
- Application to Telecommute Decisions through the Pandemic

Open science reflection:

- Smartphone data, yet will publish estimation dataset/code – It's a pain!
- Obvious papers that should be compared on same dataset (Cherchi, Bansal) – Not done or valued!



Miller et al. (2019)

Discussion

How could the work in this symposium be coordinated to better achieve our objective of informing policy & planning?

What would a benchmarking sandbox for evolutionary behavior look like?