Methods in Transportation Econometrics and Statistics (Master)

Winter semester 2023/24, Solutions to Tutorial No. 1

Solution to Problem 1.1: Micro vs. macro models

General By definition, *microscopic* models ("micromodels") go down to the smallest *relevant* system unit:

- Choice tasks: the decisions of any single person
- Traffic: The individual *driver-vehicle units* or the *agents* (latin: *agere*=to act) in an agent-based model/simulation.
- Economy: the single company ("micro economics")
- Physics: Single molecules, atoms, or elementary particles

In contrast , *macroscopic* models (macromodels) generalize or aggregate the microscopic parts to a *collective* or systemic description in terms of macroscopic quantities:

- Choice tasks: percentage for a certain alternative,
- Traffic:
  - traffic flow [veh./h] or density [veh./km] or collective phenomena such as *traffic waves*, or more generally, *traffic flow dynamics*.
  - still one level more macroscopic, one speaks of the daily traffic volume (DTV), origin-destination matrix, or the objects of *transportation planning* in general.
- Economy: macro-economics
- Physics: macroscopic phenomena such as the flowing of liquids and gasses including sound and water waves[1]

[1] Both in traffic and sound/water waves, the particles move completely differently from the waves: In traffic, the particles (vehicles) move *against* the wave propagation direction. In sound waves, the gas particles just oscillate to and fro, and in deep-water waves, they describe closed circles.

Solution to Problem 1.2: Model and system equations of linear regression

(i) The *general linear model equation*

$$Y(\vec{x}, \vec{\beta}) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + ... + \epsilon = \sum_j \beta_j x_j + \epsilon = \hat{y}(\vec{x}, \vec{\beta}) + \epsilon, \tag{1}$$

has one or more linear factors x_j. Writing the exogenous variables x'_k appearing in these factors in terms of the system attributes identifies the system. The fixed (linear or nonlinear) equations making up each factor characterize the general system dynamics. The residual term $\epsilon \sim i.i.d.\, N(0, \sigma^2)$ should satisfy the *Gauß-Markow-assumptions.*

(ii) From the general model, one obtains the *system equations* by applying the former to the data, more specifically, to the n data points

$$\vec{p_i} = \{x_{i0}, x_{i1}, ..., x_{iJ}, y_i\}, \quad i = 1, ..., n$$

Inserting, in turn, each point into the general equation results in the system equations

$$y_i = \sum_{j=0}^{J} x_{ij}\beta_j + \epsilon_i = \left(\underline{\underline{X}}\vec{\beta} + \vec{\epsilon}\right)_i, \quad i = 1, 2, ..., n. \tag{2}$$

Here, the empirical residual term $\epsilon_i(\vec{x})$ denotes the observed deviation between the endogenous variable of the data point i and the model estimate $\sum_j x_{ij}\beta_j$ as a function of the parameter vector $\vec{\beta}$.

(iii) By minimizing the sum of squared errors

$$S(\vec{\beta}) = \vec{\epsilon}(\vec{\beta})'\vec{\epsilon}(\vec{\beta})$$

with respect to $\vec{\beta}$ gives the *estimated/calibrated model*

$$y(\vec{x}) = \hat{y}(\vec{x}) + \epsilon, \quad \hat{y}(\vec{x}) = \sum_j \hat{\beta}_j x_j \tag{3}$$

which is formally like the general model. The difference is that $\vec{\beta}$ is no longer general but has fixed values $\hat{\vec{\beta}}$ obtained from this particular calibration task. The model is *validated* if it applies also new situations with unchanged $\vec{\beta}$.

The fact that the parameters are fixed once the model is calibrated is the basis for the analysis/prediction performance of econometric models.

Solution to Problem 1.3: Model and system equation in a binary choice situation

(a) **Utility function**: The perceived (deterministic) utility function should only depend on the time and ad-hoc cost differences of the two alternatives, $\tilde{x}_1$ and $\tilde{x}_2$, respectively, and on the income $\tilde{x}_3$ such that the time sensitivity increases and the cost sensitivity decreases with the income:

$$\Delta V(\tilde{x}_1, \tilde{x}_2, \tilde{x}_3) = \beta_1 + \beta_2 \tilde{x}_1 + \beta_3 \tilde{x}_2 + \beta_4 \tilde{x}_1 \tilde{x}_3 + \beta_5 \tilde{x}_2 \tilde{x}_3 \tag{4}$$

where

- $\tilde{x}_1 = T_{\text{PT}} - T_{\text{alt}}$ difference in the complex (doorstep-to-doorstep) travel times between the public transport (PT) and the best other alternative,
- $\tilde{x}_2 = C_{\text{PT}} - C_{\text{alt}}$ is the corresponding ad-hoc cost difference
- In contrast to $\tilde{x}_1$ and $\tilde{x}_2$, the income $\tilde{x}_3$ depends only on the person, not the alternatives. Such *socio-economic* variables cannot be used directly (then, the difference in the determinsitic utilities would be zero) but only with alternative-specific parameters and/or multiplication with alternative-specific variables such as $\tilde{x}_1$, $\tilde{x}_2$. Here the context implies the latter.

Meaning of the parameters:

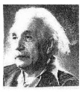
- β_1: global *ad-hoc bonus* for the public transport (generally, for alternative 1 if the other alternative is the reference)
- $\beta_2 < 0$ describes the time sensitivity
- $\beta_3 < 0$ describes the cost sensitivity
- $\beta_4 < 0$ describes the increase of the time sensitivity with income. It is negative since, with growing income, people become more time sensitive, i.e., the time sensitivity becomes more negative. This becomes evident if writing the β_2 and β_4 terms as

$$\beta_2 \tilde{x}_1 + \beta_4 \tilde{x}_1 \tilde{x}_3 = (\beta_2 + \beta_4 \tilde{x}_3) \tilde{x}_1.$$

  The term in parantheses is the effective time sensitivity for a given income.
- In analogy, $\beta_5 > 0$ describes the change of the sensitivity to costs with the income. Since, obviously, people with high income are less sensitive to proces than those with low income, i.e., the effecitive price sensitivity $\beta_3 + \beta_5 \tilde{x}_3$ becomes less negative, we should have $\beta_5 > 0$. However, the model comes at its limits for "super-earners" where $\beta_3 + \beta_5 \tilde{x}_3$ becomes positive which is nonsense: Nobody loves high prices per se (maybe *Gucci* bags are an exception of that).

Setting up the (abstract) model equations, i.e., the model *structure* is known as **model specification**. Two aspects matter:

(i) The **principle of parsimony**, a.k.a. **Occam's Razor** is best paraphrased in the words of Albert Einstein:



Make it as simple as possible.
But not simpler.

In our example: If linearity is enough, keep to it. If some nonlinearity is immanent to the system (here: a multiplicative interaction) restrict to the simplest form, i.e. quadratic or multiplicative terms.

(ii) Check qualitative consistency of the parameters in form of plausible signs (need a calibration first). Here,

- $\beta_2 + \beta_4 \tilde{x}_3 < 0$ (additional non-leisure time is always a disutility),
- $\beta_3 + \beta_5 \tilde{x}_3 < 0$ (More money for the same product/service under the same conditions is always bad).

Hint: In the sense of *Occam's Razor*, One could check to simplify the model further by ignoring one of the multiplicative interactions, e.g., the β_5 term. The β_4 interaction term alone ensures that rich (more precisely much earning) people value time more in terms of Euros or Dollars than poor people. In particular, the **value of time (VoT)**

$$\text{VoT} = \frac{\beta_2 + \beta_4 \tilde{x}_3}{\beta_3} > 0$$

increases with the income if $\beta_2 < 0$, $\beta_3 < 0$, and $\beta_4 < 0$ and, as a bonus, the problem of a positive price sensitivity of the super-earners is also gone.

(b) **System equations**

Assume data from a survey for the last trips home → university for n persons. Each person $i \in \{1, 2, ..., n\}$ can either chose the public transport ($y_{n1} = 1, y_{n2} = 0$) or other transport modes ($y_{n1} = 0, y_{n2} = 1$). The public tranmsport and the best of the other modes also have certain **characteristics** (alternative-dependent attributes), at least total travel time T_{ni} and ad-hoc costs C_{ni}. Each person also may have **socio-economic** attributes, e.g., the income I_i in multiples of 1 000 €/year: Defining the factors

$$\tilde{x}_{1i} = T_{1i} - T_{2i}, \tilde{x}_{2i} = T_{1i} - T_{2i}, \tilde{x}_{3i} = I_i,$$

we have following data matrix:

i	T_{PT}	T_{alt}	$\tilde{x}_{1i}$	C_{PT}	C_{alt}	$\tilde{x}_{2i}$	$\tilde{x}_{3i} = I_i$	y_{1i}	y_{2i}
1	20	30	-10	2	0	2	30	1	0
2	40	25	15	2	3	-1	35	0	1
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

By the way: the parameter vector $\vec{\beta}$ will be calibrated by the **maximum likelihood method** instead of minimizing the **sum of squared errors (SSE)** (we will come to that later). This means, we maximize the conditional probability that the model predicts the data (the discrete decisions), given the parameter vector, with respect to this vector.