Model specifications

Alternative specific constants, variables, and functional forms

Michel Bierlaire

Mathematical Modeling of Behavior

EPFL

Outline

The logit model

Specification

$$
\begin{aligned}
U_{in} &= V_{in} + \eta + \frac{1}{\mu}\varepsilon_{in} \\
&= \sum_k \beta_k z_{ink} + \eta + \frac{1}{\mu}\varepsilon_{in},
\end{aligned}
$$

where ε_{in} are i.i.d. EV(0, 1).

The logit model

Choice probability

$$P(i|V, \mathcal{C}_n; \mu) = \frac{e^{\mu V_{in}}}{\sum_{j \in \mathcal{C}_n} e^{\mu V_{jn}}},$$

or

$$P(i|z_n, \mathcal{C}_n; \beta, \mu) = \frac{e^{\mu(\sum_k \beta_k z_{ink})}}{\sum_{j \in \mathcal{C}_n} e^{\mu(\sum_k \beta_k z_{jnk})}}.$$

The logit model: normalizations

Moneymetric utility: $\beta_c = -1$

$$\mu V_{in} = -\mu \text{cost}_{in} + \sum_k \mu \beta_k z_{ink}.$$

Linear-in-parameters: $\mu = 1$

$$\mu V_{in} = \beta_c \text{cost}_{in} + \sum_k \beta_k z_{ink}.$$

Quantitative attributes

Numerical and continuous

- $z_{ink} \in \mathbb{R}, \forall i, n, k$.
- Associated with a specific unit.
- Vary across both i and n.

Examples

- Auto in-vehicle time (in minutes).
- Transit in-vehicle time (in minutes).
- Auto out-of-pocket cost (in cents).
- Transit fare (in cents).
- Walking time to the bus stop (in minutes).

Straightforward modeling

Quantitative attributes

- β transforms the unit of the corresponding attribute into utility units.
- Example: consider two moneymetric specifications:

$$\begin{aligned} \mu V_{in} &= -\mu \text{cost} + \mu\beta \text{TT}_{in} + \cdots \\ \mu V_{in} &= -\mu \text{cost} + \mu\beta' \text{TT}'_{in} + \cdots \end{aligned}$$

- If TT_{in} is a number of minutes, the unit of β is CHF/minute.
- If TT'_{in} is a number of hours, the unit of β' is CHF/hour.
- Both models are equivalent, but the estimated value of the coefficient will be different

$$\beta \text{TT}_{in} = \beta' \text{TT}'_{in} \Longrightarrow \frac{\beta'}{\beta} = \frac{\text{TT}_{in}}{\text{TT}'_{in}} = 60.$$

Outline

Relax the i.i.d. assumption

i.i.d. assumption

- Same η for all alternatives i: relaxed in this lecture.
- Same η for all observations n: relaxed in this lecture.
- Same μ for all alternatives i: relaxed in the lecture on mixtures.
- Same μ for all observations n: relaxed in this lecture.
- Independence across alternatives i: relaxed in the lecture on MEV models.
- Independence across observations n: relaxed in the lecture on panel data.

Motivation

	iPhone	Samsung
Price	$900	$900
Cellular	5G	5G
Capacity	128GB	128GB
Display	6.6”	6.6”
Market shares US	60 %	27%

Conclusion: we cannnot assume $\mathrm{E}[\varepsilon_{in}] = \mathrm{E}[\varepsilon_{jn}]$.

Source: gs.statcounter.com/vendor-market-share/mobile

Motivation

- How can we explain the difference?
- Only the names of the alternatives are different.
- It means that the names have an important behavioral influence.
- In marketing, it is called “brand equity”.

Modeling

Alternative Specific Constants

- We define new variables:

$$z_{i,\text{iPhone}} = \begin{cases} 1 & \text{if } i \text{ is iPhone,} \\ 0 & \text{otherwise.} \end{cases}$$

$$z_{i,\text{Samsung}} = \begin{cases} 1 & \text{if } i \text{ is Samsung,} \\ 0 & \text{otherwise.} \end{cases}$$

- We include them with coefficients in the utility function.

Alternative Specific Constants

Original formulation

$$U_{in} = V_{in} + \eta + \frac{1}{\mu}\varepsilon_{in}.$$

New formulation

$$U_{in} = V_{in} + (\beta_{\text{iPhone}} z_{i,\text{iPhone}} + \beta_{\text{Samsung}} z_{i,\text{Samsung}}) + \eta + \frac{1}{\mu}\varepsilon_{in}.$$

Alternative Specific Constants

New formulation

$$
\begin{aligned}
U_{\text{iPhone},n} &= V_{\text{iPhone},n} + (\beta_{\text{iPhone}} + \eta) + \frac{1}{\mu}\varepsilon_{\text{iPhone},n}. \\
U_{\text{Samsung},n} &= V_{\text{Samsung},n} + (\beta_{\text{Samsung}} + \eta) + \frac{1}{\mu}\varepsilon_{\text{Samsung},n}.
\end{aligned}
$$

Location parameters

$$
\begin{aligned}
\eta_{\text{iPhone}} &= \beta_{\text{iPhone}} + \eta \\
\eta_{\text{Samsung}} &= \beta_{\text{Samsung}} + \eta
\end{aligned}
$$

Alternative Specific Constants

New formulation

$$U_{\text{iPhone},n} = V_{\text{iPhone},n} + \eta_{\text{iPhone}} + \frac{1}{\mu}\varepsilon_{\text{iPhone},n}.$$

$$U_{\text{Samsung},n} = V_{\text{Samsung},n} + \eta_{\text{Samsung}} + \frac{1}{\mu}\varepsilon_{\text{Samsung},n}.$$

Comments

- The use of ASCs relaxes the assumption that the location parameters are the same across alternatives.
- Abuse of language: the coefficients of the ASCs are often called ASCs themselves.
- In practice, the numbering of alternatives is often consistent with their names (iPhone is always alternative 1, etc.) But it is not required.
- There is an identification issue.

Identification

Issue

For any value of η,

$$P_n(i|\mathcal{C}_n) = \Pr(V_{in} + \eta + \eta_i + \frac{1}{\mu}\varepsilon_{in} \geq V_{jn} + \eta + \eta_j + \frac{1}{\mu}\varepsilon_{jn}, \ \forall j \in \mathcal{C}_n).$$

Normalization

- Select any arbitrary alternative j, and normalize $\eta = -\eta_j$.
- It is equivalent to normalize the ASC of alternative j to zero.
- If there are J_n alternatives, only $J_n - 1$ ASCs are identified.

Example: binary choice

Normalization of the iPhone constant

$$U_{\text{iPhone},n} = V_{\text{iPhone},n} + \frac{1}{\mu}\varepsilon_{\text{iPhone},n}.$$

$$U_{\text{Samsung},n} = V_{\text{Samsung},n} + \beta_{\text{Samsung}} + \frac{1}{\mu}\varepsilon_{\text{Samsung},n}.$$

Normalization of the Samsung constant

$$U_{\text{iPhone},n} = V_{\text{iPhone},n} + \beta_{\text{iPhone}} + \frac{1}{\mu}\varepsilon_{\text{iPhone},n}.$$

$$U_{\text{Samsung},n} = V_{\text{Samsung},n} + \frac{1}{\mu}\varepsilon_{\text{Samsung},n}.$$

Relax the i.i.d. assumption

i.i.d. assumption

- ✓ Same η for all alternatives i: relaxed in this lecture.
- Same η for all observations n: relaxed in this lecture.
- Same μ for all alternatives i: relaxed in the lecture on mixtures.
- Same μ for all observations n: relaxed in this lecture.
- Independence across alternatives i: relaxed in the lecture on MEV models.
- Independence across observations n: relaxed in the lecture on panel data.

Motivation

	iPhone	Samsung
Price	\$900	\$900
Cellular	5G	5G
Capacity	128GB	128GB
Display	6.6"	6.6"
Market shares (US)	60 %	27%
Market shares (Europe)	36 %	31%
Market shares (CH)	58 %	27%

Conclusion: ASC may vary across n. Heterogenous population.

Source: gs.statcounter.com/vendor-market-share/mobile

Modeling heterogeneity

Behavioral assumption

- Individuals have different taste parameters.
- We replace β_{iPhone} and β_{Samsung} by $\beta_{\text{iPhone},n}$ and $\beta_{\text{Samsung},n}$.
- In practice, it is in general statistically impossible to estimate a different β per individual.
- Instead, the heterogeneity is captured by socio-economic characteristics:

$$\beta_{\text{iPhone,n}} = \beta_{\text{iPhone},n}(\text{country}_n; \theta)$$

Modeling heterogeneity

ASC function of the socio-economic characteristics

$$\beta_{in} = \theta_{i1}\ s_n + \theta_{i2}\ (1 - s_n),$$

where $s_n = 1$ if n resides in the USA, 0 otherwise.

Specification of the utility

$$\begin{aligned} U_{in} &= V_{in} + \theta_{i1}\ s_n + \theta_{i2}\ (1 - s_n) + \frac{1}{\mu}\varepsilon_{in}, \\ &= V_{in} + (\theta_{i1} - \theta_{i2})\ s_n + \theta_{i2} + \frac{1}{\mu}\varepsilon_{in}, \\ &= V_{in} + \theta_{i,\text{diff}}\ s_n + \theta_{i2} + \frac{1}{\mu}\varepsilon_{in}. \end{aligned}$$

Modeling heterogeneity

Notes

- Including a socio-economic characteristic in the utility captures an heterogenous alternative specific constant.
- The constant normalized to zero must not be segmented. Therefore, no socio-economic characteristic should appear alone in the specification.
- There are two possible specifications:

$$U_{in} = V_{in} + \theta_{i1}\, s_n + \theta_{i2}\,(1 - s_n) + \frac{1}{\mu}\varepsilon_{in},$$

$$U_{in} = V_{in} + \theta_{i,\text{diff}}\, s_n + \theta_{i2} + \frac{1}{\mu}\varepsilon_{in}.$$

The first is easier to interpret. The second is more appropriate to combine several characteristics.

Modeling heterogeneity

Curse of dimensionality

- Several socio-economic characteristics can be used to segment the population.
- However, the number of segments may become too large.
- For example, if K is the number of binary socio-economic characteristics, the number of segments is 2^K.

Practical simplification

- Include a reference parameter.
- Include one parameter for each binary socio-economic characteristic, capturing the difference with the reference one.
- Total number of parameters: $1 + K$.

Modeling heterogeneity

Example

Three characteristics: s_1, s_2, s_3.

Combinatorial specification: $2^3 = 8$ parameters

$$
\begin{aligned}
&\theta_1 s_1 s_2 s_3 + \theta_2 s_1 s_2 (1 - s_3) + \theta_3 s_1 (1 - s_2) s_3 + \theta_4 s_1 (1 - s_2)(1 - s_3) + \\
&\theta_5 (1 - s_1) s_2 s_3 + \theta_6 (1 - s_1) s_2 (1 - s_3) + \\
&\theta_7 (1 - s_1)(1 - s_2) s_3 + \theta_8 (1 - s_1)(1 - s_2)(1 - s_3).
\end{aligned}
$$

Practical simplification: $1 + 3 = 4$ parameters

$$\theta_{\text{ref}} + \theta_1 s_1 + \theta_2 s_2 + \theta_3 s_3.$$

Modeling heterogeneity

Practical simplification: $1 + 3 = 4$ parameters

$$\theta_{\text{ref}} + \theta_1 s_1 + \theta_2 s_2 + \theta_3 s_3.$$

Value per segment

s_1	s_2	s_3	θ	s_1	s_2	s_3	θ
0	0	0	θ_{ref}	1	0	0	$\theta_{\text{ref}} + \theta_1$
0	0	1	$\theta_{\text{ref}} + \theta_3$	1	0	1	$\theta_{\text{ref}} + \theta_1 + \theta_3$
0	1	0	$\theta_{\text{ref}} + \theta_2$	1	1	0	$\theta_{\text{ref}} + \theta_1 + \theta_2$
0	1	1	$\theta_{\text{ref}} + \theta_2 + \theta_3$	1	1	1	$\theta_{\text{ref}} + \theta_1 + \theta_2 + \theta_3$

Extension to all parameters

Alternative Specific Constants

- Qualitative variables.
- Alternative specific coefficient.
- Heterogeneity: coefficients vary across individuals.

This potentially applies to all coefficients

Outline

Qualitative attributes

Examples

- Level of comfort for the train.
- Reliability of the bus.
- Color.
- Shape.
- etc.

Modeling

Identify all possible levels of the variable

- Comfortable,
- Not comfortable.

Introduce a 0/1 attribute for each level k

$$z_k \in \{0,1\}, \forall k = 1, \ldots, K,$$

$$\sum_{k=1}^{K} z_k = 1.$$

Example with $K = 2$

- z_{c} for comfortable.
- z_{nc} for not comfortable.

	z_{c}	z_{nc}
comfortable	1	0
not comfortable	0	1

Modeling

Utility function with ASC

$$
\begin{aligned}
\mu V_{in} &= -\mu(\text{cost}_{in} + \beta_{\text{c}} z_{\text{c}} + \beta_{\text{nc}} z_{\text{nc}} + \beta_i) \\
&= -\mu(\text{cost}_{in} + \beta_{\text{c}} z_{\text{c}} + \beta_{\text{nc}}(1 - z_{\text{c}}) + \beta_i) \\
&= -\mu(\text{cost}_{in} + (\beta_{\text{c}} - \beta_{\text{nc}}) z_{\text{c}} + \beta_{\text{nc}} + \beta_i) \\
&= -\mu(\text{cost}_{in} + \beta'_{\text{c}} z_{\text{c}} + \beta'_i)
\end{aligned}
$$

Comments

- The coefficient of one level is confounded with the coefficient of the constant.
- In the presence of an ASC, only $K - 1$ binary variables can be included.

Alternative specific coefficients

Generic vs alternative specific

$$
\begin{aligned}
\mu V_{\text{car},n} &= -\mu\text{cost} + \mu\beta_t \text{TT}_{\text{car},n} + \cdots \\
\mu V_{\text{bus},n} &= -\mu\text{cost} + \mu\beta_t \text{TT}_{\text{bus},n} + \cdots
\end{aligned}
$$

or

$$
\begin{aligned}
\mu V_{\text{car},n} &= -\mu\text{cost} + \mu\beta_{\text{car},t} \text{TT}_{\text{car},n} + \cdots \\
\mu V_{\text{bus},n} &= -\mu\text{cost} + \mu\beta_{\text{bus},t} \text{TT}_{\text{bus},n} + \cdots
\end{aligned}
$$

Modeling assumption: a minute has/has not the same value whether it is incurred on the auto or bus mode.

Outline

Modeling heterogeneity

Behavioral assumption

- Individuals have different taste parameters.
- The difference is explained by **continuous** socio-economic characteristics.

Example: one characteristic

$$\mu V_{in} = -\mu \text{cost}_{in} + \mu \beta_n z_{in} + \cdots$$

where

$$\beta_n = \beta_n(\text{income}_n).$$

Modeling heterogeneity

Interaction

Typical definition of β_n:

$$\beta_n = \beta \text{ income}_n.$$

$$\mu V_{in} = -\mu \text{cost}_{in} + \mu\beta \text{ income}_n z_{in} + \cdots$$
$$\mu V_{in} = -\mu \text{cost}_{in} + \mu\beta x_{in} + \cdots$$

where

$$x_{in} = \text{income}_n z_{in}.$$

Modeling heterogeneity

Example: several characteristics

$$\beta_n = \beta_n(\text{income}_n, \text{age}_n).$$

Interaction

$$\beta_n = \beta \text{ income}_n \text{ age}_n.$$

$$\begin{aligned}\mu V_{in} &= -\mu \text{cost}_{in} + \mu\beta_n z_{in} + \cdots \\ &= -\mu \text{cost}_{in} + \mu\beta \text{ income}_n \text{ age}_n z_{in} + \cdots \\ &= -\mu \text{cost}_{in} + \mu\beta x_{in} + \cdots\end{aligned}$$

where $x_{in} = \text{income}_n \text{age}_n z_{in}$.

Qualitative characteristics

Examples

- Sex.
- Education.
- Professional status.
- etc.

Modeling heterogeneity

Behavioral assumption

- Individuals have different taste parameters.
- The difference is explained by **qualitative/discrete** socio-economic characteristics.

Example: one characteristic

$$\mu V_{in} = -\mu \text{cost}_{in} + \mu \beta_n z_{in} + \cdots$$

where

$$\beta_n = \beta_n(\text{education}_n).$$

Modeling heterogeneity

Segmentation

- Assume that there are K levels for the qualitative variable (e.g. education).
- They characterize K segments in the population.
- Define

$$\delta_{kn} = \begin{cases} 1 & \text{if individual } n \text{ is associated with segment } k, \\ 0 & \text{otherwise.} \end{cases}$$

- Introduce a parameter β^k for each level and define

$$\beta_n = \sum_{k=1}^{K} \beta^k \delta_{kn}.$$

Modeling heterogeneity

Segmentation

$$\mu V_{in} = -\mu \text{cost}_{in} + \mu\beta_n z_{in} + \cdots$$

$$\mu V_{in} = -\mu \text{cost}_{in} + \mu \sum_{k=1}^{K} \beta^k \delta_{kn} z_{in} + \cdots$$

$$\mu V_{in} = -\mu \text{cost}_{in} + \mu \sum_{k=1}^{K} \beta^k x_{ink} + \cdots$$

where

$$x_{ink} = \delta_{kn} z_{in}.$$

Segmentation with several variables

Example

- Gender (M, F).
- House location (metro, suburb, perimeter areas).
- 6 segments: (M, m), (M, s), (M, p), (F, m), (F, s), (F, p).

Segmentation

Specification

$$
\begin{gathered}
\beta_{M,m} TT_{M,m} + \beta_{M,s} TT_{M,s} + \beta_{M,p} TT_{M,p} + \\
\beta_{F,m} TT_{F,m} + \beta_{F,s} TT_{F,s} + \beta_{F,p} TT_{F,p} +
\end{gathered}
$$

$TT_i = TT$ if indiv. belongs to segment i, and 0 otherwise.

Remarks

- For a given individual, exactly one of these terms is non zero.
- The number of segments grows exponentially with the number of variables.

Summary

- Specification may involve both continuous/quantitative and discrete/qualitative variables.
- Alternative specific constants are qualitative attributes capturing the name of the alternative.
- Coefficients can be generic or alternative specific.
- Taste heterogeneity: coefficients are function of socio-economic attributes, both continuous/quantitative and discrete/qualitative.

Outline

Relax the i.i.d. assumption

i.i.d. assumption

- ✓ Same η for all alternatives *i*: relaxed in this lecture.
- ✓ Same η for all observations *n*: relaxed in this lecture.
- ▶ Same μ for all alternatives *i*: relaxed in the lecture on mixtures.
- ▶ Same μ for all observations *n*: relaxed in this lecture.
- ▶ Independence across alternatives *i*: relaxed in the lecture on MEV models.
- ▶ Independence across observations *n*: relaxed in the lecture on panel data.

Motivation

Logit is homoscedastic

- Error terms have the same variance across n.
- Specification involves the same μ across n.
- But people may have different levels of knowledge (e.g. taxi drivers).
- Different sources of data:
  - RP / SP
  - Different locations.
  - Different points in time.
- Homoscedastic assumption should be relaxed.

Heteroscedasticity

Data

- K groups of data.
- Each individual n belongs to exactly one group k.
- Characterized by indicators:

$$\delta_{kn} = \begin{cases} 1 & \text{if } n \text{ belongs to group k,} \\ 0 & \text{otherwise,} \end{cases}$$

and $\sum_k \delta_{kn} = 1$, for all n.

Each individual is observed at most once. This assumption will be relaxed in the lecture on panel data.

Heteroscedasticity

Assumption: variance of error terms is different across groups

Consider individual n_1 belonging to group 1, and individual n_2 belonging to group 2.

$$\begin{aligned} U_{in_1} &= V_{in_1} + \tfrac{1}{\mu_1}\varepsilon_{in_1}, \\ U_{in_2} &= V_{in_2} + \tfrac{1}{\mu_2}\varepsilon_{in_2}, \end{aligned}$$

where $\mu_2 \neq \mu_1$.

Equivalent specification

Multiply all utilities for n_2 by μ_2/μ_1 so that the error terms are the same:

$$U_{in_1} = V_{in_1} + \frac{1}{\mu_1}\varepsilon_{in_1},$$

$$\frac{\mu_2}{\mu_1}U_{in_2} = \frac{\mu_2}{\mu_1}V_{in_2} + \frac{1}{\mu_1}\varepsilon_{in_2}.$$

Heteroscedasticity

Moneymetric utility: $\beta_c = -1$

$$\mu_1 V_{in_1} = -\mu_1 \text{cost}_{in_1} + \sum_k \mu_1 \beta_k z_{in_1 k},$$
$$\mu_2 V_{in_2} = -\mu_2 \text{cost}_{in_2} + \sum_k \mu_2 \beta_k z_{in_2 k}.$$

Comment

- Estimate K scale parameters, one per group.

Heteroscedasticity

Linear-in-parameters: $\mu_1 = 1$

$$\mu_1 V_{in_1} = \beta_c \text{cost}_{in_1} + \sum_k \beta_k z_{in_1 k},$$

$$\mu_2 V_{in_2} = -\mu_2 \beta_c \text{cost}_{in_2} + \sum_k \mu_2 \beta_k z_{in_2 k}.$$

Comment

- Estimate $K - 1$ scale parameters.
- The scale parameter for one group is normalized to one.
- Specification not linear-in-parameters anymore.

Outline

Nonlinear transformations of the variables

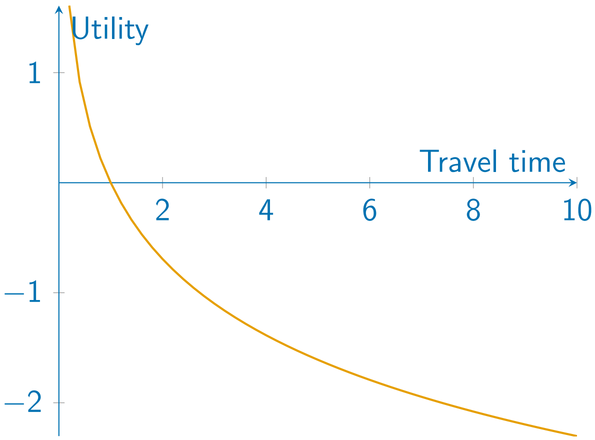
Example with travel time

- Compare a trip of 5 min with a trip of 10 min.
- Compare a trip of 120 min with a trip of 125 min.
- Utility difference: $\beta_T \cdot$ 5 min, in both cases.

Behavioral assumption

One more minute of travel is not perceived the same way for short trips as for long trips.

Nonlinear transformations of the variables



Nonlinear transformations of the variables

Assumption 1: the marginal impact of travel time is constant

$$V_{in} = \beta_T \text{time}_{in} + \cdots$$

Assumption 2: the marginal impact of travel time decreases with travel time

$$V_{in} = \beta_T \ln(\text{time}_{in}) + \cdots$$

Remarks

- It is still a linear-in-parameters specification.
- The unit, the value, and the interpretation of β_T are different.

Nonlinear transformations of the variables

Data can be preprocessed to account for nonlinearities

$$V_{in} = V(h(z_{in}, s_n)) = \sum_k \beta_k (h(z_{in}, s_n))_k$$

It is linear-in-parameter, even with h nonlinear.

Note

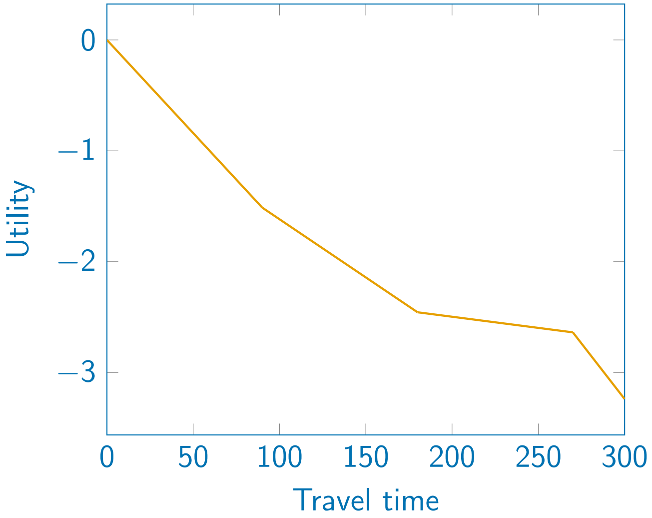
Interactions between attributes and socio-economic characteristics are a special case of h.

Piecewise linear specification

Again: sensitivity to travel time varies with travel time

- Log transform is not the only specification.
- Another possibility: split the range of values of the variable.
- Each category is associated with a different coefficient.

Piecewise linear specification



Piecewise linear specification

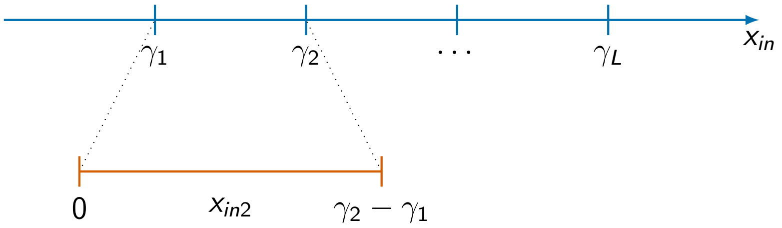
Procedure

- Select breakpoints $\gamma_1 < \gamma_2 < \ldots < \gamma_L$.
- Define new variables.

Piecewise linear specification

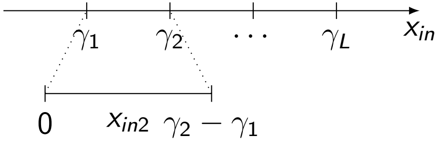
Piecewise linear specification

Piecewise linear specification



Piecewise linear specification

Formulation



$$x_{in1} = \begin{cases} x_{in} & \text{if } x_{in} < \gamma_1, \\ \gamma_1 & \text{otherwise.} \end{cases}$$

$$x_{in\ell} = \begin{cases} 0 & \text{if } x_{in} < \gamma_{\ell-1}, \\ x_{in} - \gamma_{\ell-1} & \text{if } \gamma_{\ell-1} \leq x_{in} < \gamma_\ell, \\ \gamma_\ell - \gamma_{\ell-1} & \text{otherwise.} \end{cases}$$

$$x_{in(L+1)} = \begin{cases} 0 & \text{if } x_{in} < \gamma_L, \\ x_{in} - \gamma_L & \text{otherwise.} \end{cases}$$

Piecewise linear specification

Equivalent formulations

$$x_{in1} = \begin{cases} x_{in} & \text{if } x_{in} < \gamma_1, \\ \gamma_1 & \text{otherwise.} \end{cases} \qquad \min(x_{in}, \gamma_1).$$

$$x_{in\ell} = \begin{cases} 0 & \text{if } x_{in} < \gamma_{\ell-1}, \\ x_{in} - \gamma_{\ell-1} & \text{if } \gamma_{\ell-1} \leq x_{in} < \gamma_\ell, \\ \gamma_\ell - \gamma_{\ell-1} & \text{otherwise.} \end{cases} \qquad \max(0, \min(x_{in} - \gamma_{\ell-1}, \gamma_\ell - \gamma_{\ell-1})).$$

$$x_{in(L+1)} = \begin{cases} 0 & \text{if } x_{in} < \gamma_L \\ x_{in} - \gamma_L & \text{otherwise} \end{cases} \qquad \max(0, x_{in} - \gamma_L).$$

Piecewise linear specification

Examples

$\gamma_1 = 90$, $\gamma_2 = 180$, $\gamma_3 = 270$.

x_{in}	50	100	200	300
x_{in1}	50	90	90	90
x_{in2}	0	10	90	90
x_{in3}	0	0	20	90
x_{in4}	0	0	0	30

$\gamma_1 = 1$, $\gamma_2 = 5$, $\gamma_3 = 10$.

x_{in}	0.5	4	8	12
x_{in1}	0.5	1	1	1
x_{in2}	0	3	4	4
x_{in3}	0	0	3	5
x_{in4}	0	0	0	2

Utility function

$$V_{in} = \sum_{\ell=1}^{L} \beta_\ell x_{in\ell}.$$

Box-Cox transforms

Box and Cox (1964)

$$V_{in} = \beta B(x_{in}; \lambda) + \cdots$$

where

$$B(x_{in}; \lambda) = \begin{cases} \dfrac{x_{in}^{\lambda} - 1}{\lambda} & \text{if } \lambda \neq 0, \\ \ln x_{in} & \text{if } \lambda = 0, \end{cases}$$

and $x_{in} > 0$.

Box-Cox transforms

Continuity

The function is continuous and differentiable in λ, because

$$\lim_{\lambda \to 0} \frac{x_{in}^{\lambda} - 1}{\lambda} = \ln x_{in}.$$

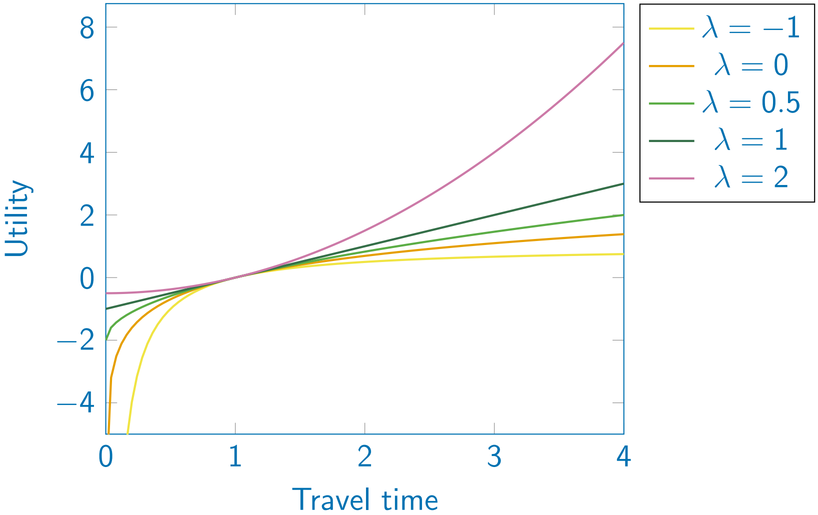
Box-Cox transforms

Convexity / concavity

$$\partial B(x_{in}; \lambda)/\partial x_{in} = x^{\lambda-1},$$
$$\partial^2 B(x_{in}; \lambda)/\partial x_{in}^2 = (\lambda - 1)x^{\lambda-2}.$$

- B is convex if $\lambda \geq 1$.
- B is concave if $\lambda \leq 1$.

Box-Cox transforms



Nonlinear interactions

Taste heterogeneity

$$V_{in} = \beta_{in} z_{in} + \cdots$$

Linear interaction

$$\beta_{in} = \beta \text{ income}_n.$$

Nonlinear interaction

$$\beta_{in} = \beta \text{ income}_n^{\lambda}, \text{ where } \lambda = \frac{\partial \beta_{in}}{\partial \text{income}_n} \frac{\text{income}_n}{\beta_{in}}.$$

Nonlinear interactions

Remarks

- λ must be estimated.
- Use a reference value for the socio-economic characteristic:

$$\beta_{in} = \beta \left(\frac{\text{income}_n}{\text{refIncome}} \right)^{\lambda}.$$

- Reference value is arbitrary.
- Several (continuous) characteristics can be combined:

$$\beta_{in} = \beta \left(\frac{\text{income}_n}{\text{refIncome}} \right)^{\lambda_1} \left(\frac{\text{age}_n}{\text{refAge}} \right)^{\lambda_2}.$$

Example: interurban mode choice in Switzerland

Sample

- Revealed preference data.
- Survey conducted between 2009 and 2010 for PostBus.
- Questionnaires sent to people living in rural areas.
- Each observation corresponds to a sequence of trips from home to home..
- Sample size: 1693.

Model: 3 alternatives

- Car,
- public transportation (PT),
- slow modes.

Attributes

Travel cost

- Alternatives: car and public transportation.
- Units: CHF.
- Specification: linear.
- Coefficient: generic.

Travel time

- Alternatives: car and public transportation.
- Units: minutes.
- Specification: Box-Cox.
- Coefficient: alternative specific, interacted with gender.

Attributes

Waiting time

- Alternative: public transportation.
- Units: minutes
- Specification: linear.
- Coefficient: interacted with income and socio-professional category.

Distance

- Alternative: slow mode.
- Units: kilometers.
- Specification: linear.
- Coefficient: piecewise linear interaction with age.

Parameters of the error terms

Alternative specific constants

- Alternatives: car and public transportation.
- Units: none.
- Coefficient: interacted with public transportation subscription.

Scale parameter

Interacted with language region.

Utility functions

Public transportation

$$\mu_n V_{\text{pt},n} = \mu_n(\text{asc}_{\text{pt},n} + \beta_c \text{tc}_{\text{pt},n} + \beta_{t,\text{pt},n} \text{boxcox\_tt}_{\text{pt},n} + \beta_{w,n} \text{wt}_n),$$

where

$$\begin{aligned}
\text{asc}_{\text{pt},n} =& \text{asc}_{\text{pt}} + \text{asc}_{\text{pt,GA}}\ \mathbb{1}(n \in \text{GA}), \\
\mu_n =& \mu_{\text{french}}\ \mathbb{1}(n \in \text{french}) + \mu_{\text{german}}\ \mathbb{1}(n \in \text{german}), \\
\beta_{t,\text{pt},n} =& \beta_{t,\text{pt}} + \beta_{t,\text{pt,female}}\ \mathbb{1}(n \in \text{female}) + \beta_{t,\text{pt,unknown}}\ \mathbb{1}(n \in \text{gender unknown}), \\
\beta_{w,n} =& \beta_w + \beta_{w,\text{craftman}}\ \mathbb{1}(n \in \text{craftman}) + \beta_{w,\text{manager}}\ \mathbb{1}(n \in \text{manager}) + \\
& \beta_{w,\text{intellectual}}\ \mathbb{1}(n \in \text{intellectual}) + \beta_{w,\text{high income}}\ \mathbb{1}(n \in \text{high income}), \\
\text{boxcox_tt}_{\text{pt},n} =& \frac{\text{tt}_{\text{pt},n}^{\lambda_{tt}} - 1}{\lambda_{tt}}.
\end{aligned}$$

Utility functions

Car

$$\mu_n V_{\text{car},n} = \mu_n(\text{asc}_{\text{car},n} + \beta_c \text{tc}_{\text{car},n} + \beta_{t,\text{car},n}\text{tt}_{\text{car},n}),$$

where

$$\begin{aligned}
\text{asc}_{\text{car},n} =& \text{asc}_{\text{car}} + \text{asc}_{\text{car,GA}} \ \mathbb{1}(n \in \text{GA}), \\
\mu_n =& \mu_{\text{french}} \ \mathbb{1}(n \in \text{french}) + \mu_{\text{german}} \ \mathbb{1}(n \in \text{german}), \\
\beta_{t,\text{car},n} =& \beta_{t,\text{car}} + \beta_{t,\text{car,female}} \ \mathbb{1}(n \in \text{female}) + \beta_{t,\text{car,unknown}} \ \mathbb{1}(n \in \text{gender unknown}) \\
\text{boxcox_tt}_{\text{car},n} =& \frac{\text{tt}_{\text{car},n}^{\lambda_{tt}} - 1}{\lambda_{tt}}.
\end{aligned}$$

Utility functions

Slow modes

$$\mu_n V_{\mathrm{sm},n} = \mu_n(\beta_{d,n} \mathrm{distance}_n),$$

where

$$\begin{aligned}
\mu_n =& \mu_{\mathrm{french}} \; \mathbb{1}(n \in \mathrm{french}) + \mu_{\mathrm{german}} \; \mathbb{1}(n \in \mathrm{german}), \\
\beta_{d,n} =& \beta_d(\mathrm{age0\text{-}40}_n + \beta_{age40-60}\mathrm{age40\text{-}60}_n + \beta_{age60-100}\mathrm{age60\text{-}100}_n), \\
\mathrm{age0\text{-}40}_n =& \min(\mathrm{age}_n, 40), \\
\mathrm{age40\text{-}60}_n =& \max(0, \min(\mathrm{age}_n - 40, 20)), \\
\mathrm{age60\text{-}100}_n =& \max(0, \min(\mathrm{age}_n - 60), 40).
\end{aligned}$$

Parameters

22 parameters

asc_{pt}	$\beta_{t,\text{car,unknown}}$
$\text{asc}_{\text{pt,GA}}$	λ_{tt}
asc_{car}	β_w
$\text{asc}_{\text{car,GA}}$	$\beta_{w,\text{craftman}}$
μ_{french}	$\beta_{w,\text{manager}}$
μ_{german}	$\beta_{w,\text{intellectual}}$
$\beta_{t,\text{pt}}$	$\beta_{w,\text{high inc}}$
$\beta_{t,\text{pt,female}}$	β_d
$\beta_{t,\text{pt,unknown}}$	$\beta_{age40-60}$
$\beta_{t,\text{car}}$	$\beta_{age60-100}$
$\beta_{t,\text{car,female}}$	β_c

Normalizations: moneymetric: $\beta_c = -1$, linear-in-parameters: $\mu_{\text{french}} = 1$.

Estimation results

General statistics

	Moneymetric	Linear
Number of estimated parameters	21	21
Sample size	1785	1785
Null log likelihood	-1922.909	-1922.909
Final log likelihood	-1007.661	-1007.644
Likelihood ratio test for the null model	1830.497	1830.531
Rho-square for the null model	0.476	0.476
Rho-square-bar for the null model	0.465	0.465
Akaike Information Criterion	2057.321	2057.287
Bayesian Information Criterion	2172.5552	2172.518
Number of iterations	74	42

Estimation results: moneymetric

Parameter number	Description	Coeff. estimate	Robust Asympt. std. error	t-stat	p-value
1	asc_{pt}	-8.87	7.8	-1.14	0.255
2	$\text{asc}_{\text{pt,GA}}$	19.9	13.2	1.51	0.131
3	$\beta_{t,\text{pt}}$	-0.771	0.559	-1.38	0.168
4	$\beta_{t,\text{pt},\text{female}}$	0.407	0.317	1.28	0.2
5	$\beta_{t,\text{pt},\text{unknown}}$	-4.31	2.09	-2.06	0.0392
6	λ_{tt}	0.758	0.122	6.22	4.82e-10
7	β_w	-0.584	0.264	-2.21	0.0271
8	$\beta_{w,\text{craftman}}$	0.415	0.328	1.27	0.206
9	$\beta_{w,\text{intellectual}}$	0.314	0.3	1.05	0.296
10	$\beta_{w,\text{manager}}$	-0.661	0.445	-1.49	0.137
11	$\beta_{w,\text{high inc}}$	-0.658	0.342	-1.93	0.0542
12	asc_{car}	16.3	8.04	2.03	0.0428
13	$\text{asc}_{\text{car,GA}}$	-15.4	13.9	-1.1	0.271
14	$\beta_{t,\text{car}}$	-2.16	1.31	-1.64	0.101
15	$\beta_{t,\text{car},\text{female}}$	0.847	0.591	1.43	0.152
16	$\beta_{t,\text{car},\text{unknown}}$	6.08	4.01	1.52	0.13
17	β_d	-3.04	0.575	-5.28	1.26e-07
18	$\beta_{age40-60}$	4.39	1.04	4.22	2.44e-05
19	$\beta_{age60-100}$	-1.74	1.56	-1.11	0.265
20	β_c	-1.0			
21	μ_{french}	0.0725	0.0195	3.71	0.000203
22	μ_{german}	0.0373	0.00975	3.83	0.000129

Estimation results: linear-in-parameters

Parameter number	Description	Coeff. estimate	Robust Asympt. std. error	t-stat	p-value
1	asc_{pt}	-0.639	0.529	-1.21	0.227
2	$\text{asc}_{\text{pt},\text{GA}}$	1.48	0.874	1.7	0.09
3	$\beta_{t,\text{pt}}$	-0.0561	0.0354	-1.59	0.113
4	$\beta_{t,\text{pt},\text{female}}$	0.0296	0.0212	1.39	0.163
5	$\beta_{t,\text{pt},\text{unknown}}$	-0.415	0.145	-2.86	0.00429
6	λ_{tt}	0.757	0.122	6.22	4.91e-10
7	β_w	-0.0423	0.0153	-2.76	0.00569
8	$\beta_{w,\text{craftman}}$	0.03	0.0231	1.3	0.194
9	$\beta_{w,\text{intellectual}}$	0.0227	0.0208	1.09	0.274
10	$\beta_{w,\text{manager}}$	-0.0479	0.0297	-1.61	0.107
11	$\beta_{w,\text{high inc}}$	-0.0477	0.023	-2.08	0.0377
12	asc_{car}	1.18	0.483	2.45	0.0143
13	$\text{asc}_{\text{car},\text{GA}}$	-1.07	0.929	-1.16	0.247
14	$\beta_{t,\text{car}}$	-0.157	0.0797	-1.97	0.0492
15	$\beta_{t,\text{car},\text{female}}$	0.0615	0.0392	1.57	0.117
16	$\beta_{t,\text{car},\text{unknown}}$	0.972	0.302	3.21	0.00131
17	β_d	-0.899	0.295	-3.04	0.00234
18	$\beta_{age40-60}$	0.319	0.421	0.758	0.448
19	$\beta_{age60-100}$	-0.425	0.235	-1.81	0.0702
20	β_c	-0.0724	0.0195	-3.71	0.000209
21	μ_{french}	1.0			
22	μ_{german}	0.515	0.0508	10.1	0.0

Estimation results: moneymetric

$$\beta_c = -0.0724$$

Parameter number	Description	Coeff. estimate	Robust Asympt. std. error	t-stat	p-value
1	asc_{pt}	-0.644	0.565	-1.14	0.255
2	$\text{asc}_{\text{pt,GA}}$	1.48	0.978	1.51	0.13
3	$\beta_{t,\text{pt}}$	-0.0561	0.0406	-1.38	0.168
4	$\beta_{t,\text{pt,female}}$	0.0296	0.023	1.28	0.2
5	$\beta_{t,\text{pt,unknown}}$	-0.444	0.214	-2.07	0.0381
6	λ_{tt}	0.757	0.122	6.23	4.77e-10
7	β_w	-0.0423	0.0191	-2.21	0.0272
8	$\beta_{w,\text{craftman}}$	0.03	0.0238	1.26	0.207
9	$\beta_{w,\text{intellectual}}$	0.0227	0.0218	1.04	0.297
10	$\beta_{w,\text{manager}}$	-0.0479	0.0322	-1.49	0.137
11	$\beta_{w,\text{high inc}}$	-0.0477	0.0248	-1.92	0.0542
12	asc_{car}	1.18	0.583	2.02	0.043
13	$\text{asc}_{\text{car,GA}}$	-1.08	1.01	-1.06	0.288
14	$\beta_{t,\text{car}}$	-0.157	0.0954	-1.64	0.101
15	$\beta_{t,\text{car,female}}$	0.0615	0.0429	1.43	0.152
16	$\beta_{t,\text{car,unknown}}$	1.41	0.489	2.89	0.00381
17	β_d	-0.596	0.171	-3.49	0.00048
18	$\beta_{age40-60}$	0.994	0.594	1.67	0.0945
19	$\beta_{age60-100}$	-0.645	0.69	-0.935	0.35
20	β_c	-0.0724			
21	μ_{french}	1.0	0.27	3.71	0.000208
22	μ_{german}	0.515	0.135	3.82	0.000134

Comments

Moneymetric vs linear-in-parameters

- The results are identical, up to numerical precision.
- Moneymetric is more challenging for the optimization algorithm.

Alternative specific constants

Moneymetric model

Public transportation vs. Slow modes

- Individuals without a GA: $\text{asc}_{\text{pt}} = -8.87$.
- Individuals with a GA: $\text{asc}_{\text{pt}} + \text{asc}_{\text{pt,GA}} = -8.87 + 19.9 = 11.03$.

Car vs. Slow modes

- Individuals without a GA: $\text{asc}_{\text{car}} = 16.3$.
- Individuals with a GA: $\text{asc}_{\text{car}} + \text{asc}_{\text{car,GA}} = 16.3 - 15.4 = 0.9$.

Public transportation vs. Car

- Individuals without a GA: $\text{asc}_{\text{pt}} - \text{asc}_{\text{car}} = -8.87 - 16.3 = -25.2$.
- Individuals with a GA:
  $\text{asc}_{\text{pt}} + \text{asc}_{\text{pt,GA}} - \text{asc}_{\text{car}} - \text{asc}_{\text{car,GA}} = 11.03 - 0.9 = 10.1$.

Travel time coefficient

Moneymetric model

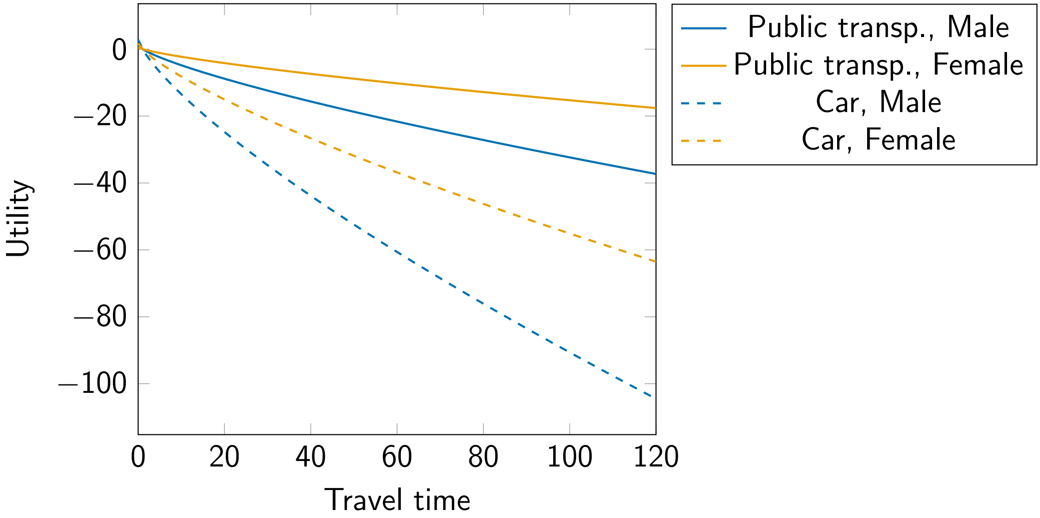
Public transportation

- Males: -0.771.
- Females: $-0.771 + 0.407 = -0.364$
- Unknown gender: irrelevant, never used for applications.

Car

- Males: -2.16.
- Females: $-2.16 + 0.847 = -1.31$
- Unknown gender: irrelevant, never used for applications.

Travel time



Public transp., Male
Public transp., Female
Car, Male
Car, Female
Utility
0
−20
−40
−60
−80
−100
0
20
40
60
80
100
120
Travel time

Waiting time coefficient

Moneymetric model

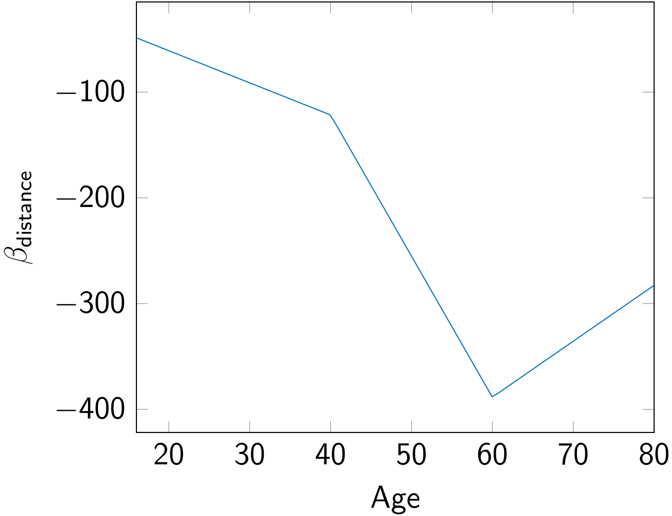
Individuals with low income

- Craftman: $-0.584 + 0.415 = -0.169$.
- Intellectual: $-0.584 + 0.314 = -0.27$.
- Manager: $-0.584 - 0.661 = -1.245$.
- Others: -0.584.

Individuals with high income

- Craftman: $-0.584 + 0.415 - 0.658 = -0.827$.
- Intellectual: $-0.584 + 0.314 - 0.658 = -0.928$.
- Manager: $-0.584 - 0.661 - 0.658 = -1.903$.
- Others: $-0.584 - 0.658 = -1.242$.

Distance coefficient



Choice probabilities

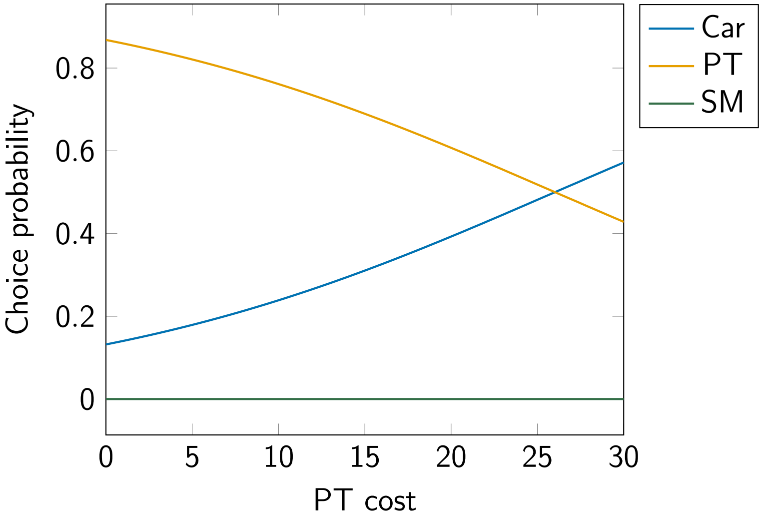
Attributes

- Cost of gasoline: 4.54 CHF.
- Travel time by car: 32 min.
- Cost PT: 12.4 CHF.
- Travel time PT: 45 min.
- Waiting time: 10 min.
- Distance: 30 km.

Socio-eco. characteristics

- French speaking.
- Age: 27.
- Gender: female.
- Subscription: GA.
- Socio-prof. category: intellectual.
- Income: low.
- Car availability: yes.

Choice probabilities



Choice probabilities

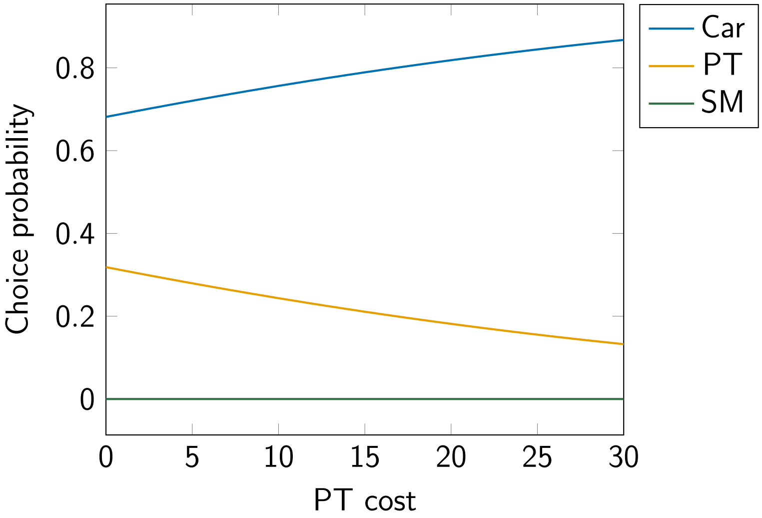
Attributes (same as before)

- Cost of gasoline: 4.54 CHF.
- Travel time by car: 32 min.
- Cost PT: 12.4 CHF.
- Travel time PT: 45 min.
- Waiting time: 10 min.
- Distance: 30 km.

Socio-eco. characteristics

- German speaking.
- Age: 55.
- Gender: male.
- Subscription: no GA.
- Socio-prof. category: manager.
- Income: high.
- Car availability: yes.

Choice probabilities



Summary

- Alternative Specific Constants.
- Continuous/quantitative vs discrete/qualitative variables.
- Taste heterogeneity and interactions.
- Heteroscedasticity.
- Nonlinear specifications.