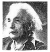
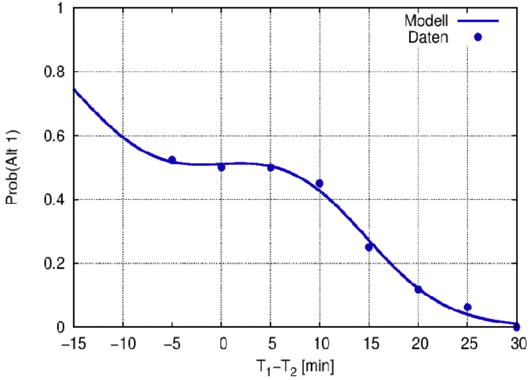
11 Advanced Concepts of Discrete-Choice Theory

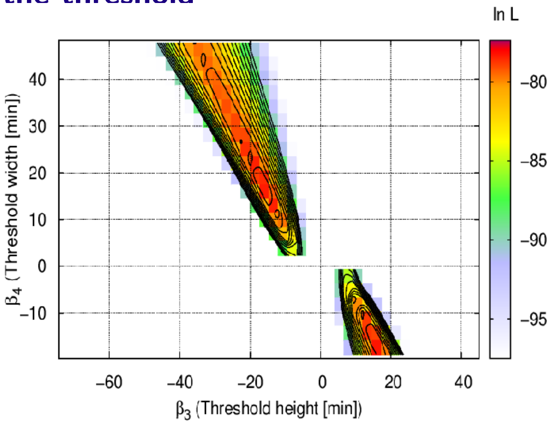
11.1  Parameter Nonlinear Models

Application: Determining subjective thresholds/indifference regions

Person class	Time Alternative 1 [min]	Time Alternative 2 [min]	Choice Alt. 1	Choice Alt. 2
1	25	30	11	10
2	30	30	10	10
3	35	30	10	10
4	40	30	9	11
5	45	30	5	15
6	50	30	2	15
7	55	30	1	15
8	60	30	0	15

Modelling the threshold





$$V_{n1} - V_{n2} = \beta_1 + \beta_2 \left[\Delta T_n + \beta_3 \tanh\left(\frac{\Delta T_n}{\beta_4}\right)\right]$$

$\hat{\beta}_1 = 0.043 \pm 0.236$ AC

$\hat{\beta}_2 = -0.29 \pm 0.38$ asymptotic time sensitivity

$\hat{\beta}_3 = -15 \pm 18$ degree of nonlinearity $\geq -\beta_4$

$\hat{\beta}_4 = 14 \pm 21$ threshold width

!! Generally, $L(\boldsymbol{\beta})$ has no longer a unique maximum, here, because of

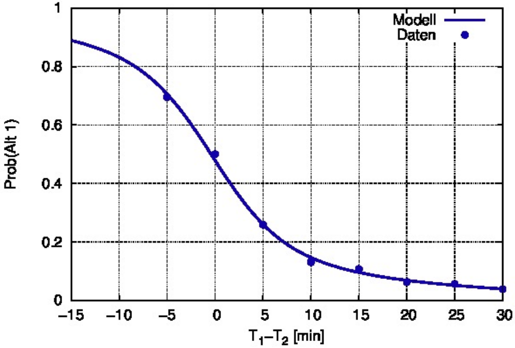
$$\beta_3 \tanh\left(\frac{\Delta T_n}{\beta_4}\right) = -\beta_3 \tanh\left(\frac{\Delta T_n}{-\beta_4}\right)$$

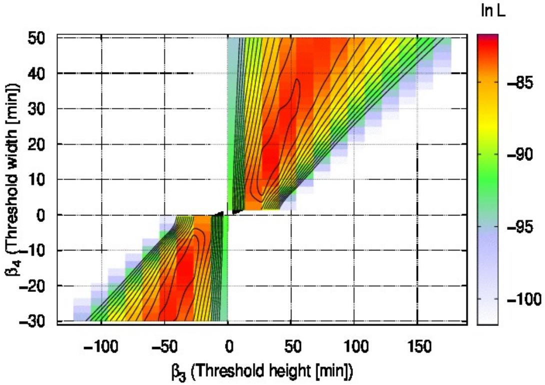
The reverse: Increased sensitivity at reference point

Person class	Time Alternative 1 [min]	Time Alternative 2 [min]	Choice Alt. 1	Choice Alt. 2
1	25	30	16	7
2	30	30	10	10
3	35	30	7	20
4	40	30	3	20
5	45	30	3	25
6	50	30	2	30
7	55	30	1	17
8	60	30	2	50

Such increased sensitivity at the reference (here: equal trip times) is proposed by the **Prospect Theory** of Kahneman/Twersky in certain situations

Modelling the increased sensitivity

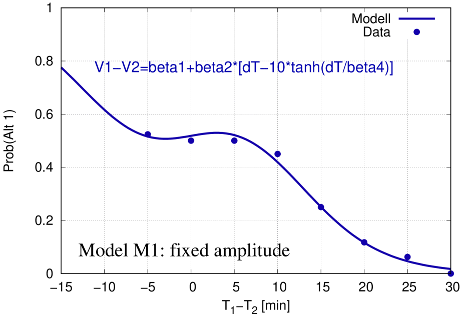


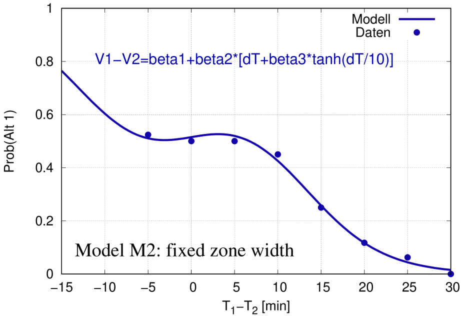


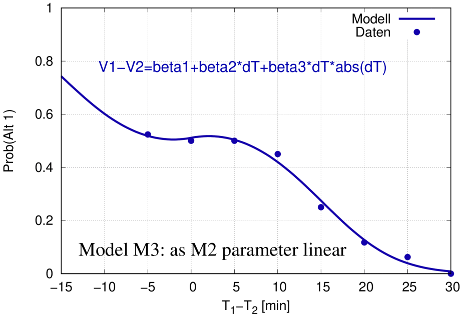
$$V_{n1} - V_{n2} = \beta_1 + \beta_2 \left[\Delta T_n + \beta_3 \tanh\left(\frac{\Delta T_n}{\beta_4}\right)\right]$$

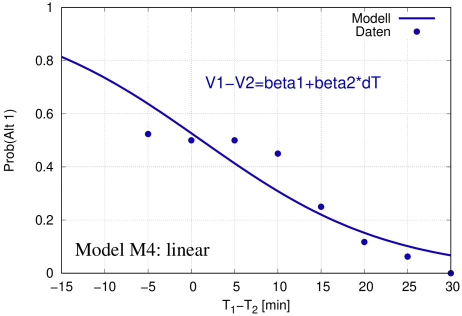
$$\hat{\beta}_1 = -0.08 \pm 0.25,$$
$$\hat{\beta}_2 = -0.05 \pm 0.10,$$
$$\hat{\beta}_3 = 27 \pm 101,$$
$$\hat{\beta}_4 = 10 \pm 16$$

Four further models applied to the threshold data









11.2 GEV and Nested Logit Models

Motivation

When taking decisions, the available options are often coupled in a way that i.i.d. random utilities are not applicable:

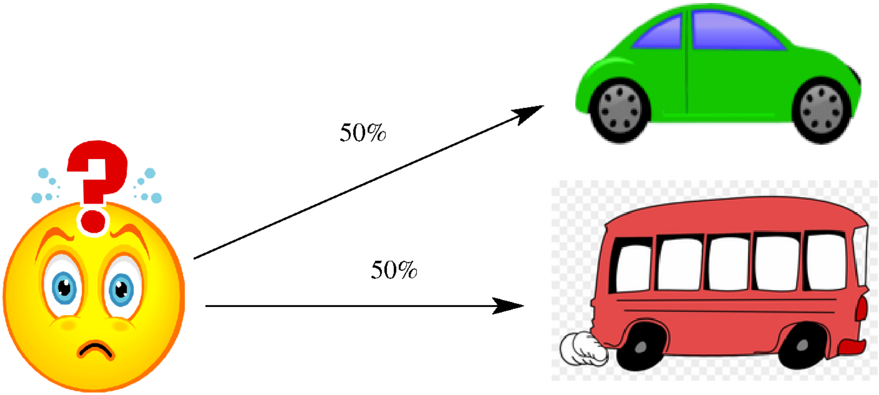
- Destination and mode choice
- Destination city and job offers when about to moving
- Expansion of a company: Creating a new branch office and if so, where?

In these cases, a decision involves taking two or more sub-decisions with nearly fixed random utilities in the associated alternative sets, so the total random utility is correlated

$\Rightarrow$ **Red-Bus-Blue-Bus problem.**

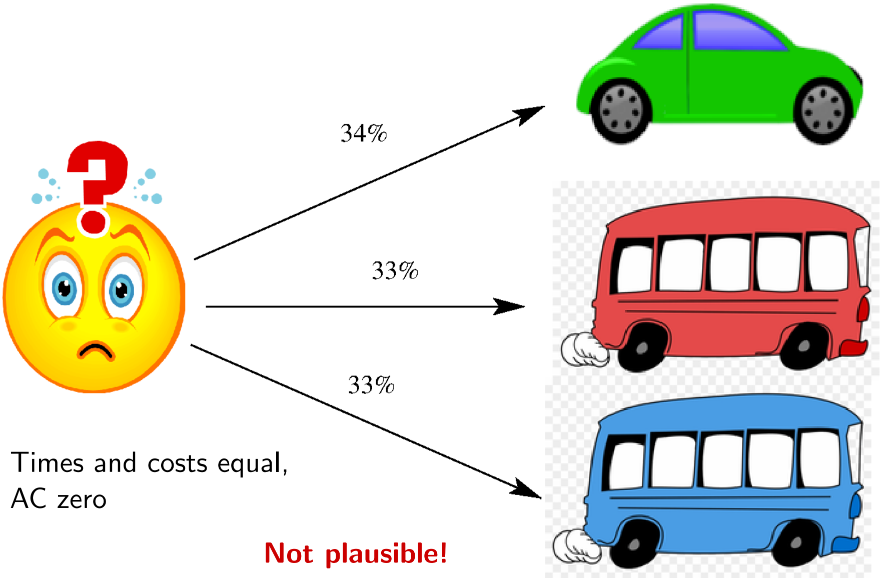
$\Rightarrow$ How to model this while retaining explicit expressions for the choice probabilities?

MNL: The Red-Bus-Blue-Bus Problem



Times and costs equal,
AC zero

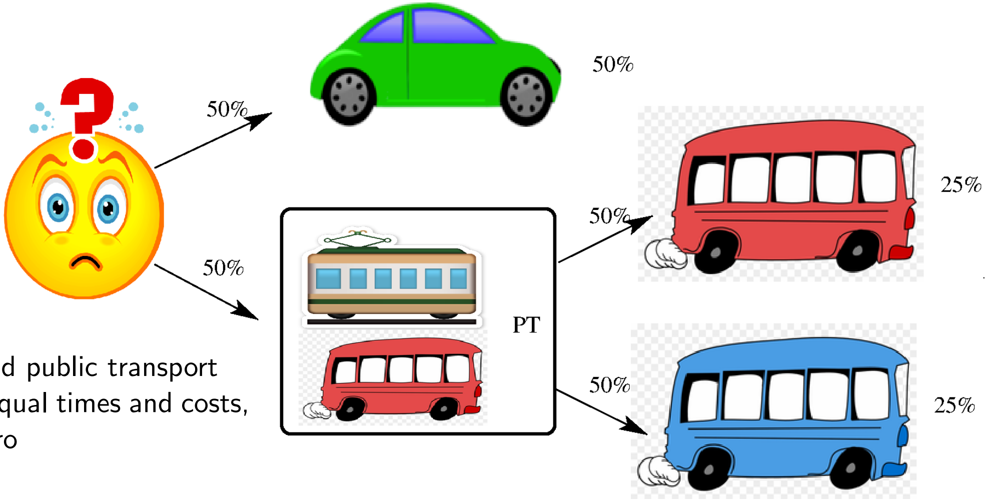
MNL: The Red-Bus-Blue-Bus Problem



Times and costs equal,
AC zero

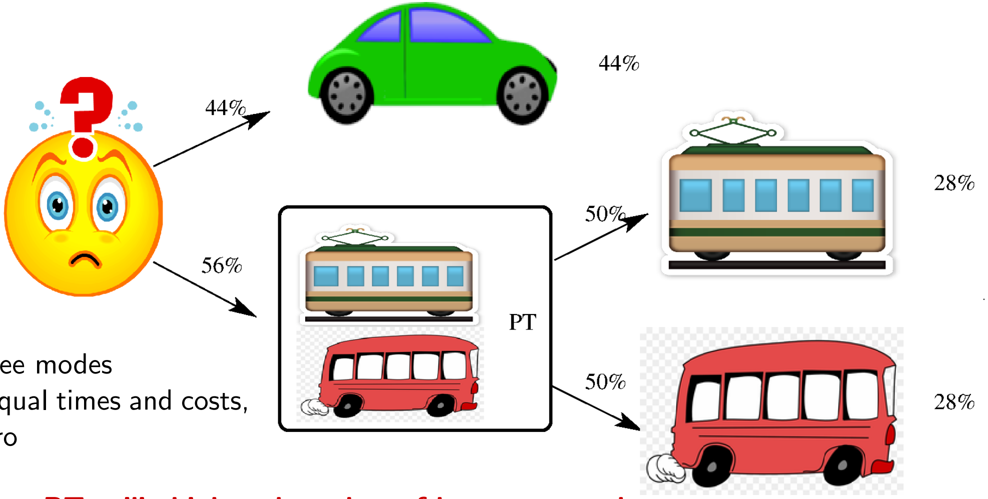
Not plausible!

100% correlated random utilities: Problem solved!



Car and public transport have equal times and costs, AC zero

Nontrivial nested decision: partial correlations



All three modes
have equal times and costs,
AC zero

Average PT utiliy higher than that of bus or tram alone because some prefer tram, some bus

The general GEV generating function

All the GEV models are defined via a **Generating function $G(y) = G(y_1, ..., y_I)$** satisfying following formal conditions:

- Not negative: $G(\boldsymbol{y}) \geq 0$ for all $\boldsymbol{y}$,
- Asymptotics: $G \to \infty$ if any $y_i \to \infty$,
- Sign of derivatives:

$$\begin{aligned} G_i &\equiv \frac{\partial G}{\partial y_i} \geq 0, \\ G_{ij} &\equiv \frac{\partial^2 G}{\partial y_i \, \partial y_j} \leq 0 \text{ if } i \neq j, \\ G_{ijk} &\geq 0 \text{ and so on,} \end{aligned}$$

- Homogeneity of degree 1: $G(\alpha \boldsymbol{y}) = \alpha G(\boldsymbol{y})$

The Nobel-Price winning result of McFadden et. al.

Any GEV function $G(\boldsymbol{y})$ satisfying the above four conditions

- generates a random vector ϵ satisfying a generalized extreme-value distribution with the distribution function
$$F(\boldsymbol{e}) = P(\epsilon_1 \le e_1, ..., \epsilon_I \le e_I) = e^{-G(\boldsymbol{y})} \text{ with } y_i = e^{-e_i}$$
- has analytic choice probabilities when maximizing the total utilities $U_i = V_i + \epsilon_i$:
$$P_i = \frac{y_i G_i(\boldsymbol{y})}{G(\boldsymbol{y})} \text{ with } G_i = \frac{\partial G}{\partial y_i},\ y_i = e^{+V_i}$$

? Check why the above conditions for $G(\boldsymbol{y})$ must be true

Question: Check the conditions for $G(\boldsymbol{y})$

? Why $G(\boldsymbol{y}) \geq 0$ for all $\boldsymbol{y}$?

! Because a distribution function $F = e^{-G}$ must be ≤ 1 (the condition $F \geq 0$ is satisfied automatically)

? Why $G \to \infty$ if any $y_i \to \infty$?

! If $y_i \to \infty$ then the argument $e_i = -\ln y_i$ of the distribution function tends to $-\infty$. Since the corresponding random variable ϵ_i is always $> -\infty$, we have $F = e^{-G} = 0$, hence $G \to \infty$

? Sign of derivatives of G?

! We check only the first derivative $G_i = \frac{\partial G}{\partial y_i}$. We have $P_i = y_i G_i / G$ with P_i, $y_i = e^{-e_i}$ and G because of the first requirement all ≥ 0. Hence $G_i \geq 0$. The other conditions follow from the non-negativity of the distribution functions

? Homogeneity $G(\alpha\boldsymbol{y}) = \alpha G(\boldsymbol{y})$ for any $\alpha > 0$?

! Because of $P_i = y_i G_i / G$ and the scaling invariance $P(\epsilon_1 < e_1) = P(\lambda\epsilon_1 < \lambda e_1)$ with $\alpha = e^{\lambda}$

Special Case I: Multinomial-Logit

- Generating function:

$$G(\boldsymbol{y})^{\mathsf{MNL}} = \sum_{j=1}^{I} y_j$$

- Distribution function of the random utilities (RUs):

$$\begin{aligned} F(\boldsymbol{e}) &= \exp\left[-G\left(e^{-e_1}, \ldots\right)\right] = \exp\left(-\sum_j e^{-e_j}\right) \\ &= \prod_j \exp\left(-e^{-e_j}\right) \Rightarrow \epsilon_i \sim \text{ i.i.d. Gumbel} \end{aligned}$$

- Choice probabilities:

$$\begin{aligned} G_i &= \frac{\partial G}{\partial y_i} = 1, \\ P_i &= \frac{y_i}{\sum_{j=1}^{I} y_j} = \frac{\exp(V_i)}{\sum_{j=1}^{I} \exp(V_j)} \end{aligned}$$

Special Case II: Two-level Nested Logit model

- Hierarchical decision: $i = (l, m)$, l: top-level alternatives, m second-level alternatives depending on l

- Generating GEV function:

$$G^{\text{NL}}(\boldsymbol{y}) = \sum_{l=1}^{L} \left(\sum_{m=1}^{M_l} y_{lm}^{1/\lambda_l} \right)^{\lambda_l}$$

  where $\lambda_l \in [0, 1]$ determine the correlations of the RUs in "nest" l:

  - $\lambda_l \to 1$: Limit of MNL, zero correlation $\Rightarrow$ check it!
  - $\lambda_l \to 0$: no RUs inside the nests, correlation=1: **sequential model: blue and red buses**

- Distribution of the RUs:

$$\begin{aligned} F(\boldsymbol{e}) &= \exp\left[-\sum_{l} \left(\sum_{m} e^{-e_{lm}/\lambda_l} \right)^{\lambda_l} \right] = \prod_{l} \exp\left[-\left(\sum_{m} e^{-e_{lm}/\lambda_l} \right)^{\lambda_l} \right] \\ &= \prod_{l} F_l(\boldsymbol{e}_l) \Rightarrow \text{independent at top level} \end{aligned}$$

Nested Logit choice probabilities

Insert $G^{\mathsf{NL}}(\boldsymbol{y})$ into the general expression $P_i = y_i G_i / G$:

$$P_i = P_{lm} = P_l P_{m|l} = \frac{e^{V_{lm}/\lambda_l}\left(\sum_{m'} e^{V_{lm'}/\lambda_l}\right)^{\lambda_l - 1}}{\sum_{l'}\left(\sum_{m'} e^{V_{l'm'}/\lambda_{l'}}\right)^{\lambda_{l'}}}$$

$\Rightarrow$ complicated and non-intuitive!

A more intuitive form of the NL choice probabilities

- Set/assume $V_{lm} = W_l + \tilde{V}_{lm}$
  - W_l: top-level contributions not appearing inside the nests
  - $\tilde{V}_{lm}$: inner contributions of alternative m in nest l
- Then, the NL choice probabilities can be formulated as

$$P_{lm} = P_l P_{m|l}, \quad P_l = \frac{e^{W_l+\lambda_l I_l}}{\sum_{l'} e^{W_{l'}+\lambda_{l'} I_{l'}}}, \quad P_{m|l} = \frac{e^{\tilde{V}_{lm}/\lambda_l}}{\sum_{m'} e^{\tilde{V}_{lm'}/\lambda_l}}$$

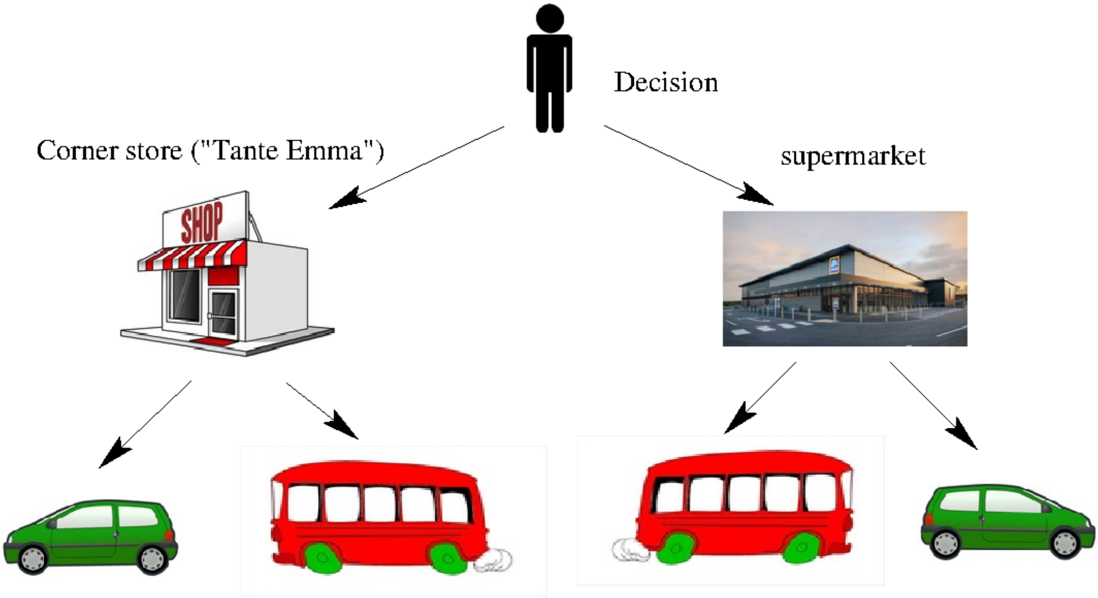
with the **inclusion values** $I_l = \ln\left(\sum_m e^{\tilde{V}_{lm}/\lambda_l}\right)$ (calibrate first $e^{\tilde{V}_{lm}/\lambda_l}$, then determine λ_l with fixed I_l in the outer MNL calibration)

? Argue that the outer nest decision is a normal MNL with the *effective nest utilities* given by $\lambda_l I_l$. Because for these assumptions P_l has the normal MNL form

? Show that $\lambda_l I_l$ is at least as high as the utility $\tilde{V}_{lm_l^*}$ of the best alternative within the nest and that $\lambda_l I_l = \tilde{V}_{lm_l^*}$ for $\lambda_l \to 0$. All contributions of the sum inside the log are exponentials and thus positive. Furthermore, the $\ln$ function is strictly monotonously increasing. Hence, $\lambda_l I_l$ is larger than any single $\tilde{V}_{lm}$ including the maximum. For $\lambda_l \to 0$, only the maximum contributes to the sum

? Argue that the (potential) selection within a nest is independent from the outer decision and obeys a normal MNL Independent because $P_{lm} = P_l P_{m|l}$, MNL for the utilities $\tilde{V}_{lm}/\lambda_l$ for fixed l

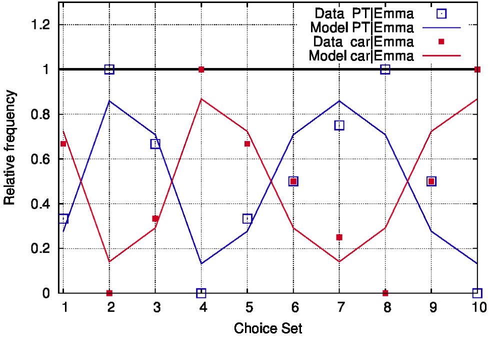
11.2.3 Example: Combined Destination and Mode Choice



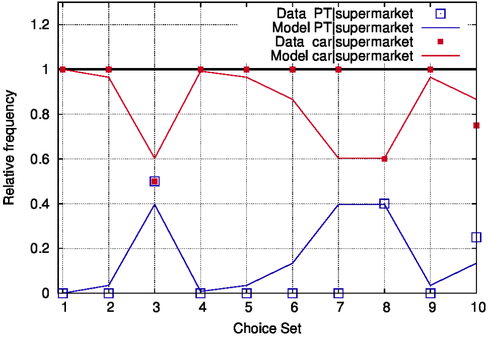
Combined destination and mode choice: the data

Person group	T [min] Emma, PT	T [min] Emma, car	T [min] superm, PT	T [min] superm, car	Fridge fill level F	y_{11}	y_{12}	y_{21}	y_{22}
1	25	15	25	20	0.9	1	2	0	0
2	25	30	40	30	0.8	3	0	0	1
3	20	20	30	30	0.7	2	1	1	1
4	25	10	25	10	0.6	0	3	0	2
5	15	5	30	20	0.5	1	2	0	2
6	15	15	25	20	0.4	1	1	0	1
7	15	20	45	45	0.3	3	1	0	1
8	15	15	15	15	0.2	1	0	2	3
9	25	15	40	30	0.1	1	1	0	1
10	25	10	25	20	0.0	0	1	1	3

Conditional modal splits



Observed and modelled modal split when driving to "Aunt Emma"

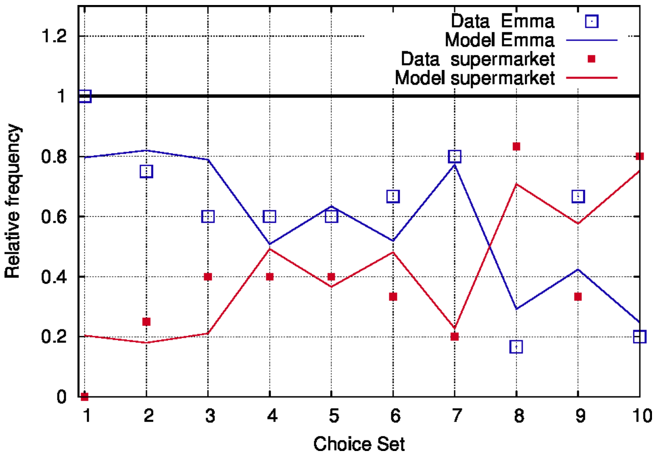


Observed and modelled modal split when driving to the supermarket

$$\begin{aligned}
P_{m|n1} &= \frac{\exp(\tilde{V}_{n1m}/\lambda_1)}{\sum_{m'} \exp(\tilde{V}_{n1m'}/\lambda_1)}, \\
\tilde{V}_{n1m}/\lambda_1 &= \beta_1 T_{n1m} + \beta_2 \delta_{m1}, \\
\hat{\beta}_1 &= -0.18,\ \hat{\beta}_2 = +0.88
\end{aligned}$$

$$\begin{aligned}
P_{m|n2} &= \frac{\exp(\tilde{V}_{n2m}/\lambda_2)}{\sum_{m'} \exp(\tilde{V}_{n2m'}/\lambda_2)}, \\
\tilde{V}_{n2m}/\lambda_2 &= \beta_3 T_{n2m} + \beta_4 \delta_{m1}, \\
\hat{\beta}_3 &= -0.29,\ \hat{\beta}_4 = -0.42
\end{aligned}$$

Top-level choice of the type of shop



Choice of the type of shop: "Aunt Emma" vs supermarket:

$$P_{nl} = \frac{\exp(W_{nl} + \lambda_l I_{nl})}{\sum_{l'} \exp(W_{nl'} + \lambda'_l I_{nl'})}$$

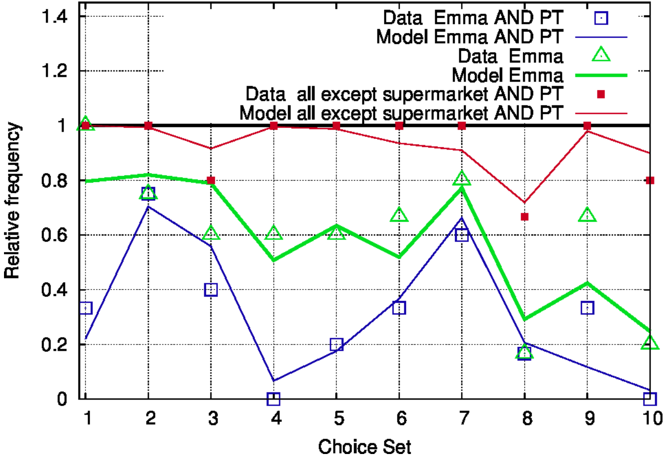
$$W_{nl} = \beta_5 F_n \delta_{l1} + \beta_6 \delta_{l1}$$

$$I_{n1} = \ln\left[\sum_m \exp\left(\hat{\beta}_1 T_{n1m} + \hat{\beta}_2 \delta_{m1}\right)\right]$$

$$I_{n2} = \ln\left[\sum_m \exp\left(\hat{\beta}_3 T_{n2m} + \hat{\beta}_4 \delta_{m1}\right)\right]$$

$$\hat{\beta}_5 = 2.9,\ \hat{\beta}_6 = -2.0,\ \hat{\lambda}_1 = 0.17,\ \hat{\lambda}_2 = 0.21.$$

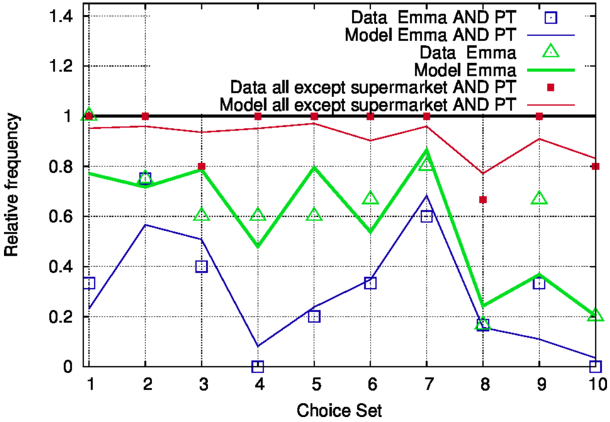
Final combined probabilities



Combined nested choice of shop type and transport mode

$$
\begin{aligned}
P_{ni} &= P_{nl} P_{m|nl} \\
&= \text{Prob(destination)*Prob(mode|destination)}
\end{aligned}
$$

Counter check: normal MNL



$$P_{ni} = \frac{\exp(V_{ni})}{\sum_{i'=1}^{4} \exp(V_{ni'})}$$

$$\begin{aligned}
V_1 &= \beta_1 T_1 + \beta_2 + \beta_5 F + \beta_6 & (l,m) = (1,1) \text{ Emma+PT} \\
V_2 &= \beta_1 T_2 + \beta_6 + \beta_5 F & (l,m) = (1,2) \text{ Emma+car} \\
V_3 &= \beta_3 T_3 + \beta_4 & (l,m) = (2,1) \text{ supermarket+PT} \\
V_4 &= \beta_3 T_4 & (l,m) = (2,2) \text{ supermarket+car}
\end{aligned}$$

$$\hat{\beta}_1 = -0.15,\ \hat{\beta}_2 = 0.60,\ \hat{\beta}_3 = -0.09,\ \hat{\beta}_4 = -0.84,\ \hat{\beta}_5 = 3.49,\ \hat{\beta}_6 = -1.76$$

11.3 Advanced I: Mixed-Logit Models

if time allows, see German script, Sec. 4.14