ENAC
Laboratoire TRANSP-OR

Prof. Michel Bierlaire

Mathematical Modeling of Behavior
(MATH–463)

January 17th 2025

Name

SCIPER

Section

Signature

Question	Points	
1	20	
2	20	
3	20	
4	20	

Total	

Grade	

This exam is written and lasts 3 hours, from 15:00 to 18:00.

The only material that you are allowed to use is the handwritten summary, maximum length of 4 pages (2 double-sided A4 sheets or 4 single-side A4 sheets).

The summaries will be collected at the end along with the exam.

The exam comprises four questions. Please answer each question in the space provided after it. The TAs can provide you with additional sheets if needed.

Make sure that your name and the date are mentioned on every page of the exam, the additional sheets (if used), and on the summary sheets.

You shall answer in English.

All answers have to be carefully justified.

No calculator is allowed.
Make sure to simplify your calculations as much as possible.

Question 1

(20 points)
Consider the Swissmetro (SM) case study, which provides information in order to analyze the impact of the modal innovation in transportation, represented by the Swissmetro (SM), against the usual transport modes represented by car and train. The utility functions for the model are defined as follows:

$$U_{in} = V_{in} + \varepsilon_{in},$$

where i denotes the alternative and n the individual. ε_{in} are error terms that are independently and identically extreme value-distributed; $\varepsilon_{in} \overset{\text{iid}}{\sim} \text{EV}(0,1)$. V_{in} are the deterministic part of utilities for alternative $i \in \{\text{car}, \text{train}, \text{SM}\}$ and individual n. In the first modeling specification (M1), the deterministic utilities are defined as follows:

Model M1:

$$\begin{aligned}
V_{\text{car},n} &= \text{ASC}_{\text{car}} + \beta_{\text{cost}}\,\text{cost}_{\text{car},n} + \beta_{\text{time}}\,\text{time}_{\text{car},n} + \beta_{\text{female,car}}\,\text{female}_n,\\
V_{\text{train},n} &= \text{ASC}_{\text{train}} + \beta_{\text{cost}}\,\text{cost}_{\text{train},n} + \beta_{\text{time}}\,\text{time}_{\text{train},n} + \beta_{\text{female,train}}\,\text{female}_n,\\
V_{\text{SM},n} &= \text{ASC}_{\text{SM}} + \beta_{\text{cost}}\,\text{cost}_{\text{SM},n} + \beta_{\text{time}}\,\text{time}_{\text{SM},n}.
\end{aligned}$$

where $\text{cost}_{i,n}$ is the travel cost in CHF and $\text{time}_{i,n}$ is the travel time in minutes associated with alternative i and individual n. female_n is a binary variable that takes value 1 if individual n is female and 0 if male. Table 1 reports the estimation results for model M1.

Parameter	**Value**	**Std. err**	**t-stat**	**p-value**
ASC_{SM}	0	–	–	–
ASC_{car}	-0.461	0.0973	-4.74	0.00
$\text{ASC}_{\text{train}}$	0.0906	0.0913	0.99	0.32
β_{cost}	-0.0108	0.000670	-16.18	0.00
β_{time}	-0.0125	0.00105	-11.81	0.00
$\beta_{\text{female,car}}$	0.309	0.102	3.04	0.00
$\beta_{\text{female,train}}$	-1.23	0.0792	-15.53	0.00

Summary statistics
Number of observations = 6768
Number of excluded observations = 3960
Number of estimated parameters = 6
$\mathcal{L}(\beta_0) = -6964.663$
$\mathcal{L}(\hat{\beta}) = -5187.983$

Table 1: Estimation results for model M1, normalization $\text{ASC}_{\text{SM}} = 0$

1. **[1 point]** Individual n does not have a driving license. What is her choice set?

2. **[1 point]** Explain the behavioral assumptions that justify the use of alternative specific constants in the model.

3. **[1 point]** Explain why there is no alternative specific constant associated with the Swissmetro (SM) alternative.

4. **[1.5 points]**

   (a) Explain the behavioral assumptions that justify the use of parameters $\beta_{\text{female,car}}$ and $\beta_{\text{female,train}}$ in the model.

   (b) Why is there no $\beta_{\text{female,SM}}$ parameter in the model?

5. **[3 points]** What if instead of capturing the effect of the female attribute in the model M1, we defined an attribute male = (1 − female) and estimated $\beta_{\text{male,car}}$ and $\beta_{\text{male,train}}$? Describe the changes in the specification and give the new values of all the estimated parameters $\beta_{\text{male,car}}$, $\beta_{\text{male,train}}$, ASC_{car}, $\text{ASC}_{\text{train}}$.

6. **[1.5 points]** Consider a model where the alternative specific constant for car is normalized to zero, and the constants for Swissmetro and train are estimated. What would be the estimated value of the parameters for this specification? Complete Table 2.

Name	Value
ASC_{SM}	
ASC_{car}	
$\text{ASC}_{\text{train}}$	
β_{cost}	
β_{time}	
$\beta_{\text{female,car}}$	
$\beta_{\text{female,train}}$	

Table 2: Parameter estimates model M1, normalization $\text{ASC}_{\text{car}} = 0$

Consider now the following model specifications:

Model M2:

$$
\begin{aligned}
V_{\text{car},n} &= \text{ASC}_{\text{car}} + \beta_{\text{cost}} \cdot \text{cost}_{\text{car},n} + \beta_{\text{time, car}} \cdot \frac{{\text{time}_{\text{car},n}}^{\lambda} - 1}{\lambda} + \beta_{\text{female, car}} \cdot \text{female}_n, \\
V_{\text{train},n} &= \text{ASC}_{\text{train}} + \beta_{\text{cost}} \cdot \text{cost}_{\text{train},n} + \beta_{\text{time, train}} \cdot \text{time}_{\text{train},n} + \beta_{\text{female, train}} \cdot \text{female}_n \\
&\quad + \beta_{\text{GA, train}} \cdot \text{GA}_n, \\
V_{\text{SM},n} &= \text{ASC}_{\text{SM}} + \beta_{\text{cost}} \cdot \text{cost}_{\text{SM},n} + \beta_{\text{time, SM}} \cdot \text{time}_{\text{SM},n} + \beta_{\text{GA, SM}} \cdot \text{GA}_n.
\end{aligned}
$$

where GA_n is a variable that takes value 1 if individual n owns a GA travelcard, and 0 otherwise. The estimation results for model M2 are reported in Table 3.

Parameter	Value	Std. err	t-stat	p-value
ASC_{SM}	0	–	–	–
ASC_{car}	0.549	0.565	0.97	0.33
ASC_{train}	-1.58	0.139	-11.32	0.00
β_{cost}	-0.0108	0.000730	-14.82	0.00
$\beta_{GA,\ SM}$	0.458	0.203	2.25	0.02
$\beta_{GA,\ train}$	2.34	0.213	11.00	0.00
$\beta_{time,\ SM}$	-0.0214	0.00181	-6.73	0.00
$\beta_{time,\ car}$	-0.0648	0.0643	-1.06	0.29
$\beta_{time,\ train}$	-0.0216	0.00115	-10.75	0.00
$\beta_{female,\ car}$	-0.428	0.102	-4.20	0.00
$\beta_{female,\ train}$	1.10	0.0863	12.74	0.00
λ	0.6	0.197	3.29	0.00

Summary statistics
Number of observations = 6768
Number of excluded observations = 3960
Number of estimated parameters = 11
$\mathcal{L}(\beta_0) = -6964.663$
$\mathcal{L}(\hat{\beta}) = -4936.917$

Table 3: Estimation results for model M2, normalization $ASC_{SM} = 0$

7. **[3 points]** What are the behavioral assumptions motivating the model specification M2?

8. **[1 point]** The utility $V_{\text{car,n}}$, is influenced by travel time $\text{time}_{\text{car},n}$, where the functional form involves an estimated parameter λ. Given the estimated λ, what is the implication in terms of the functional relationship (e.g., concavity or convexity) between $\text{time}_{\text{car},n}$ and $V_{\text{car},n}$?
   Hint: Explain your answer based on the sign of $\beta_{\text{time,car}}$.

9. **[3 points]** In model M2, calculate the value of time for train and car. What are the units? What does the values represent?
   Hint: For the car alternative, express the value of time as a function of $\text{time}_{\text{car,n}}$.

10. **[4 points]** Consider models M1 and M2.
    (a) State formally a test that can be used to compare model M1 with M2. Explain why the test can be used.
    (b) Write down explicitly the linear restrictions defining the null hypothesis of such a test. How many degrees of freedom are there for the hypothesis?
    (c) Which model is preferred using a level of significance of 5%? Use Table 4 to justify your answer.

k	$P(X \leq x)$	k	$P(X \leq x)$	k	$P(X \leq x)$
1	3.841	4	9.488	7	14.067
2	5.991	5	11.070	8	15.507
3	7.815	6	12.592	9	16.919

Table 4: Critical values of a chi-square squared distribution with k degrees of freedom for a level of significance of 5%

CORRECTION

Question 1

(20 points)

1. $C = \{\text{train}, \text{SM}\}$ **[1 point]**

2. The ASCs account for the mean of the difference of the error terms/the difference in utility of each alternative ignoring the effect of all other attributes **[1 point]**.

3. As only difference of utilities matters, only the difference of alternative specific constants can be identified. Therefore, one of them must be normalized to zero **[1 point]**.

4. (a) By using parameters $\beta_{\text{female,car}}$ and $\beta_{\text{female,train}}$ in the model, we assume different tastes between male and female in choosing mode of transport (Captures taste heterogeneity between female and male in choice of mode of transport) **[0.75 points]**.

   (b) female_n is not included in one of the utilities because it is a socio-economic characteristic and it the same between alternatives. This term has to be normalized for one alternative to be able to estimate its corresponding parameter. If not the parameter cannot be identified because the effect cancels out (we need to establish a reference level, similar as how we do for the ASCs **[0.75 point]**.

5. The new specification would be as follow. The specification of SM would stay the same.
   [0.5 point]

$$
\begin{aligned}
V_{\text{car},n} &= \text{ASC}^{\text{new}}_{\text{car}} + \beta_{\text{cost}}\,\text{cost}_{\text{car},n} + \beta_{\text{time}}\,\text{time}_{\text{car},n} + \beta_{\text{male,car}}\,\text{male}_n \\
&= \text{ASC}^{\text{new}}_{\text{car}} + \beta_{\text{cost}}\,\text{cost}_{\text{car},n} + \beta_{\text{time}}\,\text{time}_{\text{car},n} + \beta_{\text{male,car}} \cdot (1 - \text{female}_n) \\
&= (\text{ASC}^{\text{new}}_{\text{car}} + \beta_{\text{male,car}}) + \beta_{\text{cost}}\,\text{cost}_{\text{car},n} + \beta_{\text{time}}\,\text{time}_{\text{car},n} + (-\beta_{\text{male,car}}) \cdot \text{female}_n
\end{aligned}
$$

[0.5 point]

$$
\begin{aligned}
V_{\text{train},n} &= \text{ASC}^{\text{new}}_{\text{train}} + \beta_{\text{cost}}\,\text{cost}_{\text{train},n} + \beta_{\text{time}}\,\text{time}_{\text{car},n} + \beta_{\text{male,train}}\,\text{male}_n \\
&= \text{ASC}^{\text{new}}_{\text{train}} + \beta_{\text{cost}}\,\text{cost}_{\text{train},n} + \beta_{\text{time}}\,\text{time}_{\text{train},n} + \beta_{\text{male,train}} \cdot (1 - \text{female}_n) \\
&= (\text{ASC}^{\text{new}}_{\text{train}} + \beta_{\text{male,train}}) + \beta_{\text{cost}}\,\text{cost}_{\text{train},n} + \beta_{\text{time}}\,\text{time}_{\text{train},n} + (-\beta_{\text{male,train}}) \cdot \text{female}_n
\end{aligned}
$$

$$\beta_{\text{male,car}} = -\beta_{\text{female,car}} = -0.309 \ \textbf{[0.5 point]}$$

$$\beta_{\text{male,train}} = -\beta_{\text{female,train}} = 1.23 \textbf{ [0.5 point]}$$

$$\text{ASC}^{\text{new}}_{\text{car}} = \text{ASC}_{\text{car}} - \beta_{\text{male,car}} = \text{ASC}_{\text{car}} + \beta_{\text{female,car}} = -0.461 + 0.309 = -0.152 \textbf{ [0.5 point]}$$

$$\text{ASC}^{\text{new}}_{\text{train}} = \text{ASC}_{\text{train}} - \beta_{\text{male,train}} = \text{ASC}_{\text{train}} + \beta_{\text{female,train}} = 0.0906 - 1.23 = -1.1394 \textbf{ [0.5 point]}$$

6. **[1.5 points]** Table 5: [0.25] per value, except for ASC_{car}.

Name	Value
ASC_{SM}	0.461
ASC_{car}	0
$\text{ASC}_{\text{train}}$	0.5516
β_{cost}	−0.0108
β_{time}	−0.0125
$\beta_{\text{female,car}}$	0.309
$\beta_{\text{female,train}}$	−1.23

Table 5: Parameter estimates normalization $\text{ASC}_{\text{car}} = 0$ solution

7. Travel time is perceived differently in each alternative. **[1 point]**
   Marginal effect of travel time varies with travel time in car alternative. **[1 point]**
   Captures taste heterogeneity between GA holders and non-GA holders in choice of mode of transport **[1 point]**

8. As the estimated λ is less than 1, the second derivatives has the opposite sign of $\beta_{\text{time, car}}$. It means that the relationship is concave if $\beta_{\text{time, car}}$ is positive, and convex otherwise. As $\beta_{\text{time, car}}$ is is negative, the relationship is convex **[1 point]**.

9.

$$\text{VOT}_{\text{train}} = \frac{\beta_{\text{time,train}}}{\beta_{\text{cost}}} = \frac{-0.0216}{-0.0108} = 2 \quad \textbf{[1 point]} \quad \text{CHF/min} \quad \textbf{[0.5 point]}$$

$$\text{VOT}_{\text{car}} = \frac{\frac{\delta V_{\text{car,n}}}{\delta \text{time}_{\text{car,n}}}}{\frac{\delta V_{\text{car,n}}}{\delta \text{cost}_{\text{car,n}}}} = \frac{\beta_{\text{time, car}} \cdot \text{time}^{\lambda-1}_{\text{car},n}}{\beta_{\text{cost}}} = \frac{-0.0648 \cdot \text{time}^{0.6-1}_{\text{car},n}}{-0.0108} = 6\text{time}^{-0.4}_{\text{car},n} \quad \text{CHF/min}$$

[1 point]

The value of time represents the amount (CHF in this case) that individual n is willing to pay to save one unit (minute in this case) of travel time of alternative train **[0.5 point]**.

10. (a) Likelihood ratio test **[0.5 point]** as M2 is restricted version of M1 **[0.5 point]**.

(b) The null hypothesis is defined as follows:

$$\beta_{\text{GA,SM}} = \beta_{\text{GA,train}} = 0, \ \textbf{[0.5 point]}$$

$$\beta_{\text{time,car}} = \beta_{\text{time,train}} = \beta_{\text{time,SM}}, \ \textbf{[0.5 point]}$$

$$\lambda = 1 \ \textbf{[0.5 point]}$$

The degrees of freedom is $11 - 6 = 5$ **[0.5 point]**.

(c) The restricted model is M_1 and the unrestricted model is M_2. Then,

$-2(\mathcal{L}^R - \mathcal{L}^U) = -2(\mathcal{L}^1 - \mathcal{L}^2) = -2(-5187.983 + 4936.917) = 502.132.$ **[0.5 point]**

Since $\chi^2_{0.95,5} = 11.070 < 502.132$, M_2 is preferred to M_1 with a 95% confidence level **[0.5 point]** .

Question 2

(20 points)
Suppose we have data from a sample of N smartphone owners at EPFL. For each of the three main smartphone brands, we know the number of sales during the last year (see Table 6).

ID	Brand name	Number of observations
1	Apple	n_1
2	Samsung	n_2
3	Xiaomi	n_3

Table 6: Smartphone brand choice data for EPFL sample.

Given this limited dataset, we aim to analyze and model consumer behavior under several assumptions and model specifications.

1. Logit Model with Alternative-Specific Constants.

   The choice of each individual follows the logit model where the systematic utility of each alternative i is given by an alternative-specific constant α_i.

   - **[1 point]** Formulate the log-likelihood function of this specification.
   - **[0.5 points]** Determine the maximum likelihood estimators of the model parameters using the Apple alternative as a reference point.
   - **[0.5 points]** Determine the probability of choosing each alternative for each individual.

2. Logit Model with Alternative-Specific Constants and Two Categories of Individuals (Apple Lovers vs. Others).

   Each individual belongs to a category, depending on the observed choice. Apple Lovers: these individuals consider only Apple products and ignore other alternatives. Non-Apple Lovers: they ignore Apple and consider only non-Apple products.

   The choice of each Non-Apple lover follows the logit model where the systematic utility of each considered alternative i is given by an alternative-specific constant α_i.

   - **[1 point]** Formulate the log-likelihood function of this specification.
   - **[2 points]** Determine the maximum likelihood estimators of the model parameters using the Samsung alternative as a reference point.
   - **[2 points]** Determine the probability of choosing each alternative for each individual.

3. Logit Model with Alternative-Specific Constants and Three Categories of Individuals (Apple Lovers, Non-Apple Lovers, Indifferent) with known proportions.

   Each individual belongs to a category, depending on the observed choice. Apple Lovers: these individuals consider only Apple products and ignore other alternatives. Non-Apple Lovers: they ignore Apple and consider only non-Apple products. Indifferent Individuals: other individuals who consider all available products.

   The choice of each Non-Apple lover and each Indifferent Individual follows the logit model where the systematic utility of each considered alternative i is given by an alternative-specific constant α_i.

   Proportions of categories are defined as follows. p_l equals the proportion of Apple lovers among people who own Apple products. p_h equals the proportion of Non-Apple lovers among people who own Non-Apple products. Proportions p_l and p_h are assumed to be known.

   - **[1 point]** Formulate the log-likelihood function of this specification.
   - **[2.5 points]** Determine the maximum likelihood estimators of the model parameters, using the Apple alternative as a reference point.
   - **[2.5 points]** Determine the probability of choosing each alternative for each individual.
   - **[3 points]** Calculate the choice probabilities for each individual when $p_l = p_h$.
   - **[4 points]** Determine how the choice probabilities change with respect to parameters. First, fix p_l and change p_h. Second, fix p_h and change p_l. Comment on your findings.

CORRECTION

Question 2

(20 points)

1. **[2 point]** The log-likelihood of the model is

$$n_1 \log \frac{e^{\alpha_1}}{e^{\alpha_1}+e^{\alpha_2}+e^{\alpha_3}} + n_2 \log \frac{e^{\alpha_2}}{e^{\alpha_1}+e^{\alpha_2}+e^{\alpha_3}} + n_3 \log \frac{e^{\alpha_3}}{e^{\alpha_1}+e^{\alpha_2}+e^{\alpha_3}}$$

The first-order condition with respect to α_i results in

$$n_i = \frac{N}{e^{\alpha_1}+e^{\alpha_2}+e^{\alpha_3}} e^{\alpha_i}.$$

Therefore, we get that

$$\frac{e^{\alpha_i}}{e^{\alpha_1}+e^{\alpha_2}+e^{\alpha_3}} = \frac{n_i}{N}.$$

The probability of choosing the alternative i equals the proportion of the time the corresponding alternative is chosen in the sample. Additionally, we get

$$e^{\alpha_2} = e^{\alpha_1}\frac{n_2}{n_1}; \qquad e^{\alpha_3} = e^{\alpha_1}\frac{n_3}{n_1}.$$

Normalizing $\alpha_1 = 0$ we obtain that $\alpha_2 = \log \frac{n_2}{n_1}$ and $\alpha_3 = \log \frac{n_3}{n_1}$.

2. **[5 points]** Given the individual's choice, we can identify the type of the individual. If she chooses the Apple product, then she is the Apple lover. If she does not choose the Apple product, she is the Non-Apple Lover.

If an individual is the Apple lover, she always chooses the Apple product. That is, $P(1|\text{Apple lover}) = 1$. Therefore, no model is required to predict her choice. The analysis for the Non-Apple lovers is identical to the previous case with $n_2 + n_3$ individuals. That is, the probability that the Non-Apple lover chooses Samsung equals to $\frac{e^{\alpha_2}}{e^{\alpha_2}+e^{\alpha_3}}$, and she chooses Xiaomi with probability $\frac{e^{\alpha_3}}{e^{\alpha_2}+e^{\alpha_3}}$. Normalizing $\alpha_2 = 0$ we get that $\alpha_3 = \log \frac{n_3}{n_2}$.

To sum up, we introduced two types of individuals. The choice of the first group is deterministic; the choice of the second group follows the same pattern as previously.

3. **[13 points]** The log-likelihood of the model is

$$\begin{aligned}
&(1-p_l)n_1 \log \frac{e^{\alpha_1}}{e^{\alpha_1}+e^{\alpha_2}+e^{\alpha_3}} + \\
&+ (1-p_h)n_2 \log \frac{e^{\alpha_2}}{e^{\alpha_1}+e^{\alpha_2}+e^{\alpha_3}} + p_h n_2 \log \frac{e^{\alpha_2}}{e^{\alpha_2}+e^{\alpha_3}} \\
&+ (1-p_h)n_3 \log \frac{e^{\alpha_3}}{e^{\alpha_1}+e^{\alpha_2}+e^{\alpha_3}} + +p_h n_3 \log \frac{e^{\alpha_3}}{e^{\alpha_2}+e^{\alpha_3}}.
\end{aligned}$$

The first-order condition with respect to α_i results in

$$(1-p_l)n_1 = \Big((1-p_l)n_1 + (1-p_h)(n_2+n_3)\Big)\frac{e^{\alpha_1}}{e^{\alpha_1}+e^{\alpha_2}+e^{\alpha_3}};$$
$$n_2 = \Big((1-p_l)n_1 + (1-p_h)(n_2+n_3)\Big)\frac{e^{\alpha_2}}{e^{\alpha_1}+e^{\alpha_2}+e^{\alpha_3}} + p_h(n_2+n_3)\frac{e^{\alpha_2}}{e^{\alpha_2}+e^{\alpha_3}};$$
$$n_3 = \Big((1-p_l)n_1 + (1-p_h)(n_2+n_3)\Big)\frac{e^{\alpha_3}}{e^{\alpha_1}+e^{\alpha_2}+e^{\alpha_3}} + p_h(n_2+n_3)\frac{e^{\alpha_3}}{e^{\alpha_2}+e^{\alpha_3}}.$$

Summing up and rearranging the first-order conditions with respect to α_2 and α_3 gives

$$e^{\alpha_2}+e^{\alpha_3} = \Big(e^{\alpha_1}+e^{\alpha_2}+e^{\alpha_3}\Big)\frac{(1-p_h)(n_2+n_3)}{(1-p_l)n_1+(1-p_h)(n_2+n_3)}.$$

From the first-order condition with respect to α_1 using the normalization $\alpha_1 = 0$ we get that

$$e^{\alpha_2}+e^{\alpha_3} = \frac{(1-p_h)(n_2+n_3)}{(1-p_l)n_1}.$$

Substituting the value of $e^{\alpha_2}+e^{\alpha_3}$ back into the first-order condition with respect to α_2 results in

$$n_2 = (1-p_l)n_1 e^{\alpha_2} + \frac{p_h(1-p_l)n_1}{1-p_h}e^{\alpha_2}.$$

Rearranging we get that $\alpha_2 = \log\frac{(1-p_h)n_2}{(1-p_l)n_1}$ and $\alpha_3 = \log\frac{(1-p_h)n_3}{(1-p_l)n_1}$.

Therefore, if $p_l = p_h$ we get that the solution is identical to the first model. $\alpha_1 = 0, \alpha_2 = \log\frac{n_2}{n_1}, \alpha_3 = \log\frac{n_3}{n_1}$.

If an individual is indifferent, her probability of choice equals the probability of choice from the first model. If an individual is Apple lover or Non-Apple lover, her probability of choice equals to the probability of choice from the second model.

If $p_l \neq p_h$ we observe that if an individual is the Apple lover or the Non-Apple lover, her probability of choice equals the probability of choice from the second model. That is, $P(1|\text{Apple lover}) = 1, P(2|\text{Non-Apple lover}) = \frac{n_2}{n_2+n_3}, P(3|\text{Non-Apple lover}) = \frac{n_3}{n_2+n_3}$. If an individual is indifferent to the Apple products, her probability of choice equals to

$$P(1|\text{IND}) = \frac{(1-p_l)n_1}{(1-p_l)n_1+(1-p_h)(n_2+n_3)};$$
$$P(2|\text{IND}) = \frac{(1-p_h)n_2}{(1-p_l)n_1+(1-p_h)(n_2+n_3)};$$
$$P(3|\text{IND}) = \frac{(1-p_h)n_3}{(1-p_l)n_1+(1-p_h)(n_2+n_3)}.$$

In the sample, there are $(1 - p_l)n_1 + (1 - p_h)(n_2 + n_3)$ individuals who consider all alternatives when they make a decision. Intuitively, if p_h is large, then individuals who consider all alternatives mostly buy the Apple product. Therefore, if p_h is large, then the values of α_2 and α_3 should be relatively small. The opposite argument applies when p_h is small. If p_l is large, then individuals who consider all alternatives mostly buy Samsung or Xiaomi. Therefore, if p_l is large, then the values of α_2 and α_3 should be relatevily large. The opposite argument applies when p_l is small.

We confirm our intuition, observing that α_2 and α_3 are monotone with respect to p_l and p_h. In particular, both parameters decrease with respect to p_h and increase with respect to p_l.

Question 3

(20 points)
The City of Lausanne conducted a study to better understand the commuting mobility patterns of the employees and students of EPFL. To this end, a team of researchers conducted a data collection with the help of an application which the participants installed on their smartphones. The app was tracking them for two weeks and collected information about the mode choice and travel time. Participants also provided information about the travel cost of each trip. After some data processing, the team of researchers used the data of N participants. For each of the participants, T observations were available. The researchers decided to estimate a cross-nested logit (CNL) model considering five alternatives (car, bus, metro, bicycle, walking) adopting the specification shown in Figure 1.

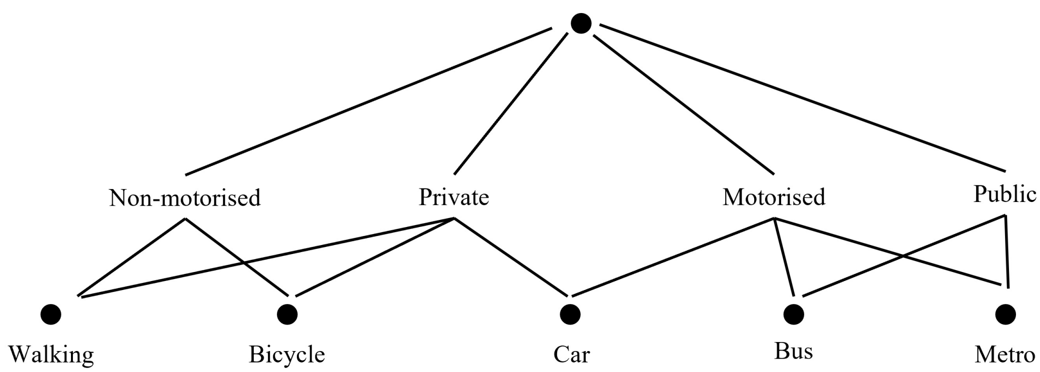

Figure 1: The CNL model specification

The attributes used in the model were:

- Travel time (for all alternatives)
- Travel cost (not for cycling and walking)

1. **(5 points)** Specify an **error component mixture logit model** that captures the cross-nested structure depicted in Figure 1. Provide the complete mathematical formulation of the utility functions for each alternative, explicitly including all terms and notation. Ensure the formulation accounts for the cross-nested structure described in the figure. Consider the follwoing instructions:

    - Assume that a linear-in-parameter specification, where the scale parameter μ is normalized to 1.
    - Do not apply any parameter normalization at this stage, except for the scale parameter.

- Include generic coefficients for travel time and travel cost across alternatives.
- Clearly define and explain all terms and parameters used in the utility functions.
- State all assumptions regarding the random terms used in the model, including their distributions and dependencies.
- Do not omit any error terms from the specification.

Note: To save time, group terms that can be described in general. For example, instead of repeating explanations for each alternative, you may use phrasing like "ASC_i represents the ... of alternative i."

2. **(3 points)** After estimating both the CNL and the error component mixture logit model, the researchers noticed a difference in the cost parameter, between the two models. In particular, they found that:

   β_{costCNL} = -1.52
   β_{costMIXT} = - 4.25

   Based on this outcome, the researchers concluded that the CNL model underestimated the sensitivity of individuals to cost. Explain why this conclusion is incorrect and suggest a way to compare the sensitivity to cost between the two models.

3. **(12 points)** Even though the researchers had repeated observations over time from each individual (panel data), the specification of the error component mixture logit model in question 3.1 is static:

   i. What is serial correlation, and what is is capturing in the context of the tracking data described in the question? (3 points)
   ii. Apart from serial correlation, name one additional limitation of a static model specification when panel data is used (2 point)
   iii. For each of the two limitations, propose a way to resolve the issue (2 points)
   iv. What is the initial condition problem? What may cause the problem in this specific mode choice model example? What is the impact on the parameter estimates? (3 points)
   v. Propose a solution to the initial condition problem. (2 points)

CORRECTION

Question 3

(20 points)

1. **(5 points)** The utilities of the fully specified model are:

   $U_{\text{car,n,t}} = \text{ASC}_{\text{car}} + \beta_c \text{Cost}_{\text{car,n,t}} + \beta_{tt} \text{Time}_{\text{car,n,t}} + \sigma_{\text{mot}} \xi_{\text{mot}} + \sigma_{\text{priv}} \xi_{\text{priv}} + \epsilon_{\text{car,n,t}}$
   $U_{\text{bus,n,t}} = \text{ASC}_{\text{bus}} + \beta_c \text{Cost}_{\text{bus,n,t}} + \beta_{tt} \text{Time}_{\text{bus,n,t}} + \sigma_{\text{mot}} \xi_{\text{mot}} + \sigma_{\text{pub}} \xi_{\text{pub}} + \epsilon_{\text{bus,n,t}}$
   $U_{\text{metro,n,t}} = \text{ASC}_{\text{metro}} + \beta_c \text{Cost}_{\text{metro,n,t}} + \beta_{tt} \text{Time}_{\text{metro,n,t}} + \sigma_{\text{mot}} \xi_{\text{mot}} + \sigma_{\text{pub}} \xi_{\text{pub}} + \epsilon_{\text{metro,n,t}}$
   $U_{\text{bike,n,t}} = \text{ASC}_{\text{bike,n,t}} + \beta_{tt} \text{Time}_{\text{bike,n,t}} + \sigma_{\text{nonmot}} \xi_{\text{nonmot}} + \sigma_{\text{priv}} \xi_{\text{priv}} + \epsilon_{\text{bike,n,t}}$
   $U_{\text{walk,n,t}} = \text{ASC}_{\text{walk}} + \beta_{tt} \text{Time}_{\text{walk,n,t}} + \sigma_{\text{nonmot}} \xi_{\text{nonmot}} + \sigma_{\text{priv}} \xi_{\text{priv}} + \epsilon_{\text{walk,n,t}}$

   where:

   ASC_i are the alternative specific constants
   β_c is the travel cost parameter
   β_{tt} is the travel time parameter
   $\text{Cost}_{i,n,t}$ is the cost of alternative i for the t^{th} observation of individual n
   $\text{Time}_{i,n,t}$ is the travel time of alternative i for the t^{th} observation of individual n
   σ_* are error term related parameters that capture the correlation between alternatives (for nest *)
   $\xi \sim N(0,1)$ are normally distributed error terms
   $\epsilon \sim EV(0,1)$ are Gumbel distributed error terms

2. **(3 points)** The linear in parameters specification of the logit mixtures is not directly comparable to the CNL specification, models have different scales (1.5 point). To address the issue we can either re-estimate using money metric utilities or divide all parameters by β_c (1.5 point).

3. **(12 points)**

   i. Serial correlation refers to the dependence between the error terms of the model across repeated observations for the same individual (1.5 points). Simply put, it is related to the persistence of unobserved factors over time, for the choices of an individual (1.5 points).

   ii. Another limitation of static models is related to the dynamics, in particular the impact of past choices on the current choice. This effect is not captured in a static model specification (2 point).

iii.
- For serial correlation: addition of error components (typically normally distributed) that capture unobserved heterogeneity across individuals (1 point).
- For dynamics: include past choices as an explanatory variable of the current choice (1 point).

iv. The initial condition problem arises when dynamic Markov model is combined with random effects (1 point). By introducing the past choices as explanatory variables there is an endogeneity problem due the the presence of the error term that captures persistence (since the same error term was used as an explanatory variable in the previous time step to explain the choice). Going backwards, this problem starts from the correlation between the first observed choice and the error term (1 point). This problem (the initial condition problem) leads to incosisntent parameter estimates (1 point).

v. (2 points) We can discard the first observation of each individual and express the persistence error term as a function of the first choice (Wooldridge approach). Hence, instead of

$\alpha_{i,n=}\nu_{\text{i}}\zeta_{\text{i,n}}$ (where i is an alternative)

we have to do

$\alpha_{i,n} = a_i + b_i y_{n,0} + c_i x_n + \nu_i \zeta_{i,n}$

where:
a_i is a constant for alternative i
b_i is the parameter related to the fist choice of the individual n
x_n are time invariant variables related to individual n
c_i are the parameters associated to x_n
ζ_{in} a draw for which we have $\xi_n \sim$ N(0,1)
ν_i parameter to be estimated for alternative i

Question 4

(20 points)
The city of Chicago experiences severe winter weather, including heavy snowstorms and extreme cold, leading to frequent high-risk meteorological alerts. As part of a research team focused on understanding behavioral responses to these conditions, your task is to build a model that explains why people will take the choice i to Work From Home (WFH) instead of commuting to the office and Work From the Office (WFO) during such high-risk situations. To this end, your team has collected survey data from 300 employees in various roles with various professional backgrounds. Each respondent n has indicated their level of agreement with a set of 30 statements, I_n^k where $k = 1, \cdots, 30$, related to safety concerns and their remote work environment. We also collected their commute time (commute_{in}) to work and the sector they are working in (sector_{in}). Finally, the survey includes each respondent's age (age_n). Finally, we define R_n^* as the commuting risk perception latent variable.

1. **[3 points]** Due to the extensive set of indicators available, we have preselected 9 key indicators and conducted a factor analysis to identify underlying correlations between indicators. Table 7 presents the results of this analysis, detailing each indicator along with its definition and measurement method. Based on this information, select the set of indicators from the factor you believe represents best the latent variable Risk Perception and explain your reasoning for this choice and why is better than the other choices.

   ☐ Factor 1.
   ☐ Factor 2.
   ☐ Factor 3.

Factor	Notation	Statement	Likert Scale (1-5)
1	I_n^1	I feel my health is at risk when traveling to work during extreme weather.	1 - Strongly Disagree, 2 - Disagree, 3 - Neutral, 4 - Agree, 5 - Strongly Agree
	I_n^2	I believe commuting in severe weather can lead to accidents or injury.	1 - Strongly Disagree, 2 - Disagree, 3 - Neutral, 4 - Agree, 5 - Strongly Agree
	I_n^3	I am concerned about the potential hazards of commuting in poor weather conditions.	1 - Strongly Disagree, 2 - Disagree, 3 - Neutral, 4 - Agree, 5 - Strongly Agree
2	I_n^4	I have access to all necessary tools and technology to work from home effectively.	1 - Strongly Disagree, 2 - Disagree, 3 - Neutral, 4 - Agree, 5 - Strongly Agree
	I_n^5	I feel confident in my ability to perform my job remotely.	1 - Strongly Disagree, 2 - Disagree, 3 - Neutral, 4 - Agree, 5 - Strongly Agree
	I_n^6	My home environment allows me to work without distractions.	1 - Strongly Disagree, 2 - Disagree, 3 - Neutral, 4 - Agree, 5 - Strongly Agree
3	I_n^7	I feel stressed thinking about severe weather.	1 - Strongly Disagree, 2 - Disagree, 3 - Neutral, 4 - Agree, 5 - Strongly Agree
	I_n^8	I worry about my well-being in severe weather.	1 - Strongly Disagree, 2 - Disagree, 3 - Neutral, 4 - Agree, 5 - Strongly Agree
	I_n^9	I experience anxiety when my kid goes to school in poor weather conditions.	1 - Strongly Disagree, 2 - Disagree, 3 - Neutral, 4 - Agree, 5 - Strongly Agree

Table 7: Factor analysis of the indicators.

2. **[2 points]** Write down the structural equations. Consider the latent variable model, the specification of which is depicted in Figure 2. Note that $\varepsilon_n^* \sim \mathcal{N}(0, \sigma^*)$. Change a, b, c with the superscripts corresponding to the indicators you selected from Table 7 in the previous exercise. Ensure that any new notation added is properly defined.

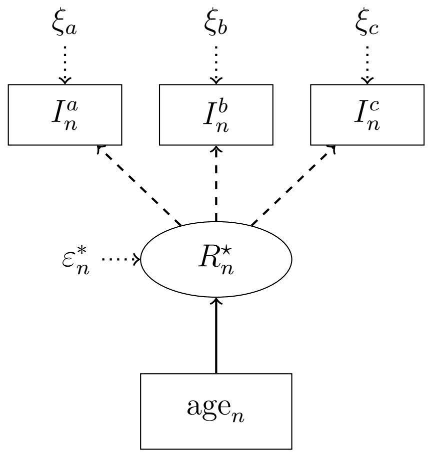

Figure 2: Latent variable model diagram.

3. **[3 points]** Write down the measurement equations. Consider again the latent variable model, the specification of which is depicted in Figure 2. Note that $\xi_a \sim \mathcal{N}(0, \sigma^a), \xi_b \sim \mathcal{N}(0, \sigma^b), \xi_c \sim \mathcal{N}(0, \sigma^c)$.

4. **[2 points]** Write down the structural equations for the choice model from Figure 3.

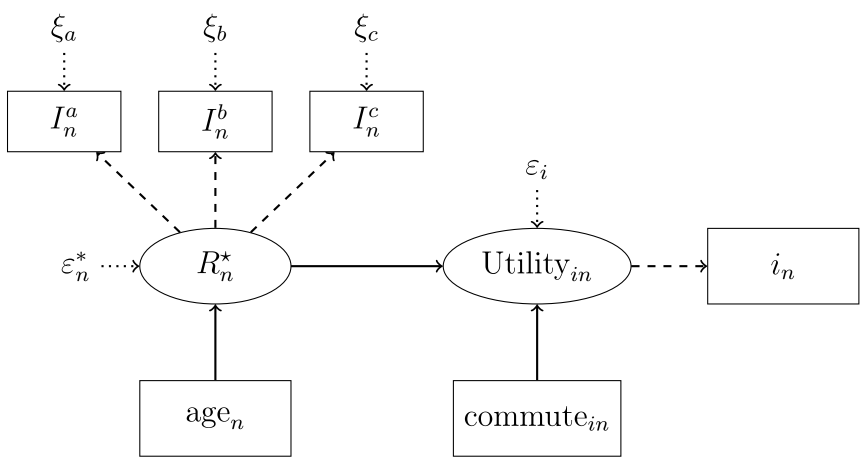

Figure 3: Hybrid choice model diagram.

5. **[4 points]** After estimating the latent variable model, we are trying to fit the choice model using sequential estimation. We assume to use commuting time as an explanatory variable of the binary choice model. However, the parameter estimates that we obtained are non-significant. Which of the following 3 alternative specifications would be appropriate to attempt to re-estimate the parameters? Why? (You may select any number of options, including all,

none, one, or two.)

$$\text{U}_{\text{WFO}} = V_{\text{WFO}} + \varepsilon_{\text{WFO}} = \text{ASC}_{\text{WFO}} + \beta_{\text{WFO}}^{\text{commute}} \text{commute}_{\text{WFO}} + \varepsilon_{\text{WFO}}$$
$$U_{\text{WFH}} = 0.$$

☐ Specification 1. $\beta_{\text{WFO}}^{\text{commute}}$ is defined as:

$$\beta_{\text{WFO}}^{\text{commute}} = \beta_{\text{WFO}}^{\prime\text{commute}} R^*.$$

☐ Specification 2. $\beta_{\text{WFO}}^{\text{commute}}$ is defined as:

$$\beta_{\text{WFO}}^{\text{commute}} = \exp\left(\beta_{\text{WFO}}^{\prime\text{commute}} R^*\right).$$

☐ Specification 3. $\beta_{\text{WFO}}^{\text{commute}}$ is defined as:

$$\beta_{\text{WFO}}^{\text{commute}} = \exp\left(\beta_{\text{WFO}}^{\prime\text{commute}} R^*\right),$$

and ASC_{WFO} is defined as:

$$\text{ASC}_{\text{WFO}} = \beta_{\text{WFO}}^{\prime\text{R}^*} R^*.$$

Table 8 provides the different specification estimated results.

	Parameter	**Value**	**Rob. Std err**	**Rob. t-test**	**Rob. p-value**
Specification 1	ASC_{WFO}	0.809	0.133	6.08	$1.19\text{e}-09$
	$\beta_{\text{WFO}}^{\prime\text{commute}}$	-0.0154	0.004	-3.86	$1.14\text{e}-04$
	σ_s	0.416	0.195	2.13	$3.28\text{e}-02$
Specification 2	ASC_{WFO}	0.755	0.106	7.09	$1.35\text{e}-12$
	$\beta_{\text{WFO}}^{\prime\text{commute}}$	-0.0138	0.00331	-4.18	$2.95\text{e}-05$
	σ_s	0.387	0.168	2.31	$2.09\text{e}-02$
Specification 3	ASC_{WFO}	3.64	0.469	7.76	$8.66\text{e}-15$
	$\beta_{\text{WFO}}^{\prime\text{commute}}$	-0.00485	0.00126	-3.84	$1.24\text{e}-04$
	$\beta_{\text{WFO}}^{\prime\text{R}^*}$	-1.44	0.299	-4.82	$1.43\text{e}-06$
	σ_s	-0.210	0.0461	-4.56	$1.07\text{e}-06$

Table 8: Combined results for specifications 1, 2, and 3.

6. **[2 points]** Can you comment on the sign of parameter $\beta_{\text{WFO}}^{\prime\text{commute}}$ from Table 8?

7. **[4 points]** Some companies in Chicago are considering providing subsidies to their employees to help them purchase additional hardware or equipment needed to create an optimal home office setup. These companies believe that ensuring employees have a proper remote setup might encourage employees to work from home, particularly during periods of extreme weather conditions.

   To this end, we define the latent variable Need for Remote Setup (N^*) to investigate its role in reinforcing the decision to work from home when there are extreme weather conditions. Research has shown that when R^* is high, individuals are more likely to feel the need for a proper remote setup, which could, in turn, influence their likelihood of choosing to work remotely. Provide an updated diagram where both R^* and N^* influence the choice. Choose the set of indicators from Table 7 you will use to define N^*.

CORRECTION

Question 4

(20 points)

1. **[3 points]**

   ☒ Factor 1 **[1 point]**.
   ☐ Factor 2.
   ☐ Factor 3.

   Factor 1 effectively captures risk perception in the context of commuting during adverse weather because it directly addresses health and safety concerns associated with travel risks. Statements in this factor, such as "I feel my health is at risk when traveling to work during extreme weather," specifically target the respondent's perception of physical danger and vulnerability, aligning closely with the concept of risk perception. In contrast, Factor 2 focuses on the convenience and readiness to work from home, while Factor 3 reflects general commuting stress rather than an immediate sense of risk. Consequently, Factor 1 includes the physical and emotional concerns that would prompt someone to avoid commuting due to perceived threats to personal safety, making it the most accurate indicator of risk perception for this choice model.

2. **[2 points]** The latent variable for each individual, denoted as R_n^*, represents the high perceived risk and is estimated by the following structural equation:

   $$R_n^* = \beta^* + \beta_n^{\text{age}} \text{age}_n + \sigma^* \varepsilon^*$$

   where β^* is the intercept, β_n^{age} is the coefficients for age_n for each individual n, σ^* is the standard deviation of the error term and ε^* represents the error term associated with the latent variable.

3. **[3 points]**

   The indicators, measured on a Likert scale from 1 to 5, indicate risk perception towards bad weather by an individual n. These indicators are associated with the latent variable through the following measurement equations:

   $$I_n^a = \alpha_0^a + \alpha^a R_n^* + \sigma_a \xi_a$$
   $$I_n^b = \alpha_0^b + \alpha^b R_n^* + \sigma_b \xi_b$$
   $$I_n^c = \alpha_0^c + \alpha^c R_n^* + \sigma_c \xi_c$$

   where $\alpha_0^a, \alpha_0^b, \alpha_0^c$ are the intercepts for the a, b, c indicators, $\alpha^a, \alpha^b, \alpha^c$ are the coefficient relating the latent variable to the a, b, c indicators, $\sigma_a, \sigma_b, \sigma_c$ are the standard deviations of the error term for the a, b, c indicators, ξ_a, ξ_b, ξ_c are the error terms for the a, b, c indicators, and $\gamma_1, \gamma_2, \gamma_3, \gamma_4$ are the thresholds that define the categories of the Likert scale.

$$I_n^k = \begin{cases} 1 & \text{if } I_n^k < \gamma_1, \\ 2 & \text{if } \gamma_1 \leq I_n^k < \gamma_2, \\ 3 & \text{if } \gamma_2 \leq I_n^k < \gamma_3, \\ 4 & \text{if } \gamma_3 \leq I_n^k < \gamma_4, \\ 5 & \text{if } \gamma_4 \leq I_n^k. \end{cases}$$

As the measurements are using a Likert scale with $M = 5$ levels, we define 4 parameters γ_m. The thresholds γ_m are defined symmetrically around zero to facilitate interpretation. For this reason, we define two positive parameters δ_1 and δ_2 as:

$$\begin{aligned} \gamma_1 &= -\delta_1 - \delta_2, \\ \gamma_2 &= -\delta_1, \\ \gamma_3 &= \delta_1, \\ \gamma_4 &= \delta_1 + \delta_2. \end{aligned}$$

4. **[2 points** The structural equation for the choice model for each alternative i is:

$$\begin{aligned} &\mathrm{U_{WFO}} = V_{\mathrm{WFO}} + \varepsilon_{\mathrm{WFO}} = \mathrm{ASC_{WFO}} + \beta_{\mathrm{WFO}}^{\mathrm{commute}} \mathrm{commute}_{ni} + \beta_{\mathrm{WFO}}^{\mathrm{R}^*} R^* + \varepsilon_{\mathrm{WFO}} \\ &\mathrm{U_{WFH}} = 0 \end{aligned}$$

5. **[4 points]**

☒ Specification 1 **[2 points]** .
☐ Specification 2.
☐ Specification 3.

The model from specification 3 is incorrect since it defines $\mathrm{ASC_{WFO}} = \beta_{\mathrm{WFO}}^{\prime \mathrm{R}^*} R^*$, which makes the ASC depend on the latent variable R^*. ASCs are generally constants that capture unobserved preferences for a particular choice relative to other choices. Making it depend on R^* could complicate interpretation because it introduces additional variability tied to a latent construct. Typically, latent variables should interact with explanatory variables rather than directly modifying constants. The model from specification 2 is incorrect since the commute time is positive, and its parameter should be negative since a higher commute time with a high-risk perception should have a negative impact on the utility.

6. **[2 points]** To analyze the sign of $\beta_{\mathrm{WFO}}^{\prime \mathrm{commute}}$ across the specifications and understand its implications, we compute the derivatives for each specification with respect to R^*.

- **Specification 1:**

$$\beta_{\text{WFO}}^{\text{commute}} = \beta_{\text{WFO}}'^{\text{commute}} R^*$$

Differentiating with respect to R^*:

$$\frac{\partial \beta_{\text{WFO}}^{\text{commute}}}{\partial R^*} = \beta_{\text{WFO}}'^{\text{commute}}$$

Thus, the sign of $\beta_{\text{WFO}}'^{\text{commute}}$ directly determines the sign of $\beta_{\text{WFO}}^{\text{commute}}$ for all $R^* > 0$.

From Table 8, $\beta_{\text{WFO}}'^{\text{commute}} = -0.0154$, which is negative. This means that $\beta_{\text{WFO}}^{\text{commute}}$ decreases as R^* increases.

- **Specification 2:**

$$\beta_{\text{WFO}}^{\text{commute}} = \exp\left(\beta_{\text{WFO}}'^{\text{commute}} R^*\right)$$

Differentiating with respect to R^*:

$$\frac{\partial \beta_{\text{WFO}}^{\text{commute}}}{\partial R^*} = \exp\left(\beta_{\text{WFO}}'^{\text{commute}} R^*\right) \cdot \beta_{\text{WFO}}'^{\text{commute}}$$

Since $\exp\left(\beta_{\text{WFO}}'^{\text{commute}} R^*\right) > 0$, the sign of $\frac{\partial \beta_{\text{WFO}}^{\text{commute}}}{\partial R^*}$ depends entirely on $\beta_{\text{WFO}}'^{\text{commute}}$.

From Table 8, $\beta_{\text{WFO}}'^{\text{commute}} = -0.0138$, which is negative, indicating that $\beta_{\text{WFO}}^{\text{commute}}$ decreases as R^* increases.

Across all specifications, $\beta_{\text{WFO}}'^{\text{commute}}$ is consistently negative, indicating that the utility for working from the office (WFO) decreases as the commute becomes more burdensome. This aligns with the intuitive notion that longer commutes reduce the attractiveness of WFO.

7. **[4 points]** The set of indicators to define N^* are the ones from Factor 2, I_n^4, I_n^5 and I_n^6

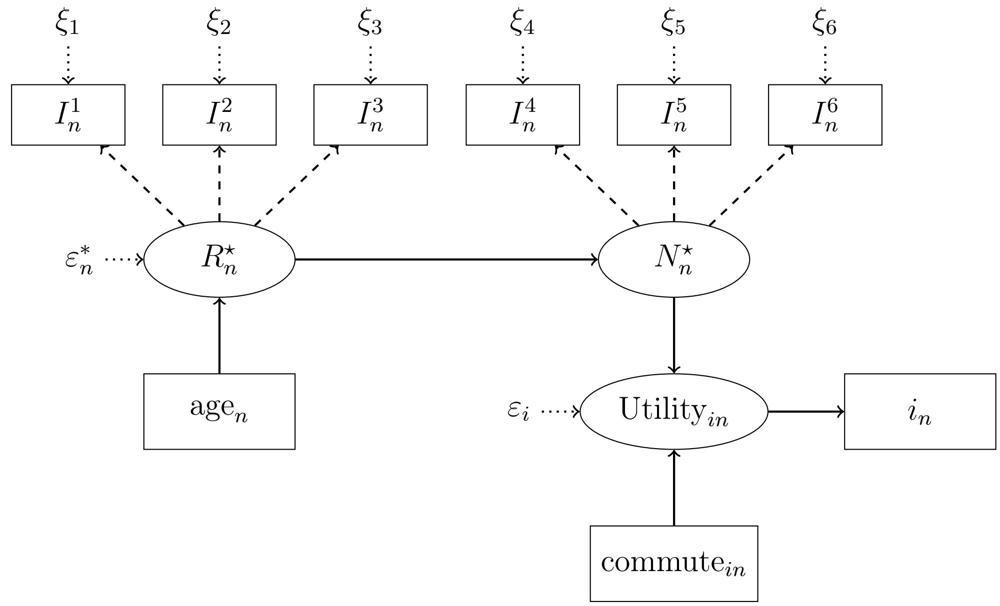

Figure 4: Hybrid choice model diagram with two latent variables (Option 1).

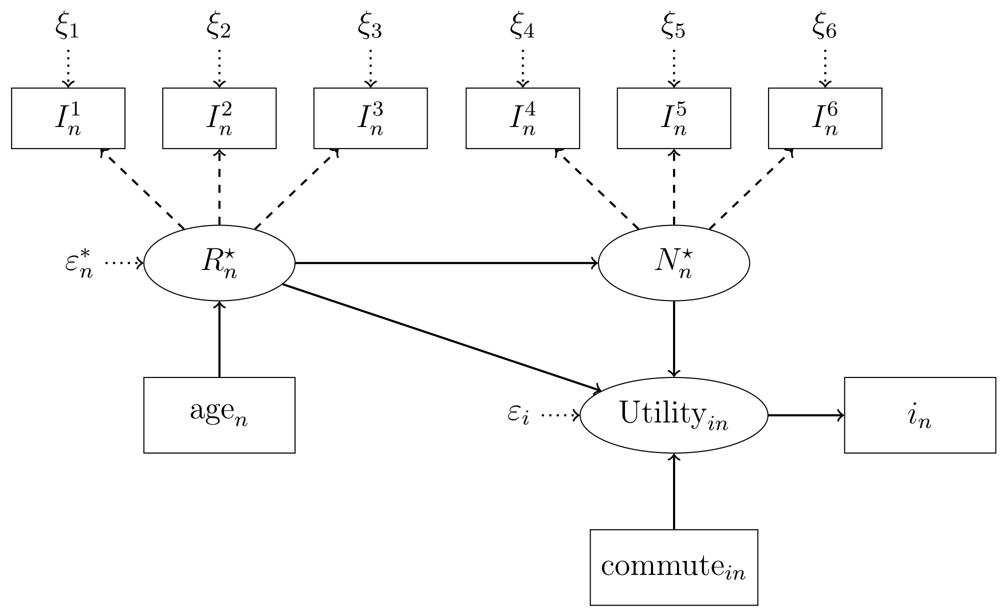

Figure 5: Hybrid choice model diagram with two latent variables (Option 2).