Mathematical Modeling of Behavior (2022-2023)

Michel Bierlaire. TAs: Silvia Varotto, Cloe Cortes, Tom Haering, Nicola Ortelli

Course assignment — December 22, 2022

Group 07

Model 0 [1 point] Start with a model specification that includes alternative specific constants, and cost and travel time of the different alternatives associated with generic parameters. Report both the specification (i.e., the utility functions) and the estimation results (parameter values, t-tests or p-values, null and final log likelihoods). [**0.5 point**]

For each mode choice, we have a different utility function $V_{\text{mode\_choice}}$ *for model 0; these are given by the following expressions:*

$$V_{\text{walk}} = \beta_{\text{time}} \text{dur}_{\text{walk}} \quad (1)$$
$$V_{\text{cycle}} = \text{ASC}_{\text{cycle}} + \beta_{\text{time}} \text{dur}_{\text{cycle}} \quad (2)$$
$$V_{\text{pt}} = \text{ASC}_{\text{pt}} + \beta_{\text{time}} \text{dur}_{\text{pt}} + \beta_{\text{cost}} \text{cost}_{\text{pt}} \quad (3)$$
$$V_{\text{driv}} = \text{ASC}_{\text{driv}} + \beta_{\text{time}} \text{dur}_{\text{driv}} + \beta_{\text{cost}} \text{cost}_{\text{driv}} \quad (4)$$

Where walk *stands for the choice of walking,* cycle *for cycling,* pt *for public transportation and* driv *for driving.* dur *stands for the duration of a trip and* cost *its correspoding cost.* $\text{cost}_{\text{driv}}$ *is taken as the sum of the estimated congestion charge cost and the estimated fuel cost.* dur_{pt} *is taken as the sum of the predicted total access and egress time, the rail in-vehicle time, the bus in-vehicle time and the interchange time.*

Note that all alternatives are available to all individuals of the sample (imposed by the project). If this was not the case, it would have been possible to make the walking option only available to people whose distance to their destination is smaller than a certain threshold. In this case, people who chose alternative "walking" while the distance to their destination were higher than this threshold could simply be removes from the data set (considered as liers). It would also be possible to make the driving option unavailable to individuals with no driving license and no car in the household. In this case, carpooling option would have been treated as well.

The null and final log likelihoods are given by -6931.47 *and* -4668.76. *The results of this model are presented in Table 1.*

Parameter	Value	t-test value
$\text{ASC}_{\text{cycle}}$	-3.671670	-34.217362
ASC_{driv}	-1.104164	-13.275731
ASC_{pt}	-0.427631	-7.654184
β_{cost}	-0.151331	-11.331599
β_{time}	-5.129646	-24.371452

Table 1: Values of the different parameters given by model specification 0 and their respective value of the t-test.

1. Comment on the estimation output (sign and significance of all parameters). [**0.5 point**]

The signs of all three ASC are negative. It means that if the cost and the time where the same for all options, walking is predicted by the model to be preferred to others options. The fact that both β *are negative reflects the intuition that individuals prefer when the time and the cost of a transportation mode is smaller. Since all t-test are, in absolute value, larger than* 1.96 *which is the common practice quantile of* 95% *of the normal distribution, the hypothesis that all the parameters are equal to* 0 *can be rejected. The null log likelihood is simply a benchmark while the final log likelihood gives a mean to measure the goodness of a fit to a model. A high value of this parameter implies a better fit. It will be of used when different model will be compared.*

Model 1 [**2.5 points**] Using Model 0 as the base model, include alternative-specific parameters for one of the attributes of Model 0. Report both the specification and the estimation results (as defined previously). [**0.5 point**]

We choose to include alternative specific parameters for the attribute time of model 0. The utility functions are now given by the following equations:

$$V_{walk} = \beta_{time,walk} dur_{walk} \tag{5}$$
$$V_{cycle} = ASC_{cycle} + \beta_{time,cycle} dur_{cycle} \tag{6}$$
$$V_{pt} = ASC_{pt} + \beta_{time,pt} dur_{pt} + \beta_{cost} cost_{pt} \tag{7}$$
$$V_{driv} = ASC_{driv} + \beta_{time,driv} dur_{driv} + \beta_{cost} cost_{driv} \tag{8}$$

The null and final log likelihoods are given by -6931.47 *and* -4361.57*. The results of this model are presented in Table 2.*

Parameter	Value	t-test value
ASC_{cycle}	-4.743495	-22.522778
ASC_{driv}	-1.865705	-11.864122
ASC_{pt}	-2.311875	-14.700747
β_{cost}	-0.131324	-9.144126
$\beta_{time,cycle}$	-4.880140	-11.483382
$\beta_{time,driv}$	-6.113741	-16.356896
$\beta_{time,pt}$	-3.270473	-13.632966
$\beta_{time,walk}$	-8.208049	-17.115501

Table 2: Values of the different parameters given by model specification 1 and their respective values of the t-test.

1. State the underlying assumption of defining alternative-specific parameters in this specific situation. [**0.5 point**]

The assumption is that time is not perceived in the same way depending on the transportation mode. For example, spending one hour in a train is felt differently than spending one hour cycling.

2. Comment on the estimation output (as defined previously, including any changes from the previous model). [**0.5 point**]

It could be observed from Table 2 that the signs of all ASCs are negative. It means that if the cost and the time were not taken into account, walking would be predicted by the model to be preferred over the remaining mode choices. Moreover, the fact that all β*s values are negative reflects the intuition that individuals have a preference for shorter time and lower cost of a transportation mode. Since all t-test are, in absolute value, larger than 1.96, we have evidence to reject the null hypothesis, which states that all the parameters are equal to 0. The fact that each* β_{time} *changes significantly from a given choice mode to the others. This, in turn, helps support the hypothesis that time is perceived differently depending upon what transportation mode is chosen. Time in public transportation is perceived more positively than time spent for the other transportation options. The null log likelihood is again just a benchmark.*

3. Compare Model 0 and Model 1 with an appropriate statistical test. Justify your choice of test. State the null hypothesis and the result of the test. Denote the preferred model as $\text{Model}_{\text{pref}}$. [**1 point**]

Since Model 0 is a restricted version of Model 1 (the restriction is $\beta_{\text{time,walk}} = \beta_{\text{time,cycle}} = \beta_{\text{time,pt}} = \beta_{\text{time,driv}} = \beta_{\text{time}}$*), we can perform a likelihood ratio test to compare the two models. Under such circumstances, we usually refer to Model 0 and Model 1 as restricted and unrestricted model, respectively. Furthermore, it is important to notice that the null hypothesis we plan on either rejecting or accepting states that these restrictions could be supported by the available data. As discussed during the course lectures, we know that under the null hypothesis, the test statistic* $-2(\mathcal{L}_R - \mathcal{L}_U)$ *is asymptotically distributed as a* χ^2_d *distribution with* d *degrees of freedom equal to the number of restrictions. If such a statistic is "large" in the statistical sense, we reject the null hypothesis. We point out that* $d = K_U - K_R$*, where* K_U

and K_R are the number of parameters in the unrestricted and restricted models, respectively. Moreover, the quantity $\mathcal{L}_R$ represents the final log likelihood of the restricted model and $\mathcal{L}_U$ the one of the unrestricted model. In our case, we compute $-2(\mathcal{L}_R - \mathcal{L}_U) = 614.39$. Since there are 3 more parameters in Model 1 than in Model 0, we need to use the χ^2_3 distribution to determine whether this is large enough to reject the null hypothesis. Using a level of significance of 5%, which corresponds to the probability of rejecting the null hypothesis when it is true, the critical value of the χ^2 distribution with 3 degrees of freedom is 7.81 (this value can be obtained in a χ^2 table, which can be found in a statistics book). Because our test statistic value 614.39 is well above this critical value, we have evidence to reject the null hypothesis with at least 95% confidence and we could state that time is, indeed, perceived differently for the four analyzed mode choices.
It is important to emphasize that when comparing two models in this assignment, we always assume that the null hypothesis corresponds to the case where the restricted model, also known as the parsimonious model, is the model we prefer over the unrestricted one since it contains fewer parameter; we should bear in mind the 'keep the model as simple as possible' motto. The intuition behind the previous statistical test is that we examine how much additional information the unrestricted model can explain in contrast to the restricted one, and whether or not it is worth including more parameters. Moreover, we perform the same previous statistical analysis further down to compare models. Hereafter, we refer to Model 1 as the preferred model and it is denoted by Model$_{\text{pref}}$.

Model 2 [3.5 points] Using Model$_{\text{pref}}$ as the base model, include one additional alternative attribute and one interaction of a socioeconomic characteristic with either the ASCs or one of the attributes. Report both the specification and the estimation results (as defined previously). **[0.5 point]**

We choose to include the additional alternative attribute "driving traf-

fic percent" in Model$_{\text{pref}}$. This parameter gives the predicted traffic variability on the driving route. We also add the interaction between travel cost and the socioeconomic characteristic age. Individuals will be separated into two different groups: people that are less active professionally, who correspond to 18-year-old people and 65-year-old (age of retirement in London, and more professionally active people being the rest of the population. We refer to the first group as young_old *and the second as* middle_age. *Then we have the following forms of the deterministic part of utility functions:*

$$
\begin{aligned}
V_{\text{walk}} &= \beta_{\text{time,walk}} \text{dur}_{\text{walk}} && (9)\\
V_{\text{cycle}} &= ASC_{\text{cycle}} + \beta_{\text{time,cycle}} \text{dur}_{\text{cycle}} && (10)\\
V_{\text{pt}} &= ASC_{\text{pt}} + \beta_{\text{time,pt}} \text{dur}_{\text{pt}} + \beta_{\text{cost,middle\_age}} \text{cost}_{\text{pt,middle\_age}} && (11)\\
&\quad + \beta_{\text{cost,young\_old}} \text{cost}_{\text{pt,young\_old}} \\
V_{\text{driv}} &= ASC_{\text{driv}} + \beta_{\text{time,driv}} \text{dur}_{\text{driv}} + \beta_{\text{cost,middle\_age}} \text{cost}_{\text{driv,middle\_age}} && (12)\\
&\quad + \beta_{\text{cost,young\_old}} \text{cost}_{\text{driv,young\_old}} \\
&\quad + \beta_{\text{traffic\_percent}} \text{traffic\_percent}
\end{aligned}
$$

After running our simulations, the null and final log likelihoods are -6931.47 and -4272.92, respectively. The results of this model are presented in Table 3.

1. State the underlying assumptions of the additional attribute and interaction in this specific situation. [**1.0 point**]

The assumption behind the additional attribute "driving traffic percent" is that the traffic along the driving route will change the utility of car driving since it is clear that traffic impacts greatly the driving time and, thus, how this choice mode is perceived. On the other hand, the assumption behind the interaction between age and cost is that different-aged individuals have not the same perception/conception of money when it comes to selecting a transportation mode.

Parameter	Value	t-test value
ASC_{cycle}	−4.769377	−22.351194
ASC_{driv}	−1.357151	−8.393631
ASC_{pt}	−2.469546	−15.327582
$\beta_{cost,middle\_age}$	−0.080101	−1.681746
$\beta_{cost,young\_old}$	−0.102830	−2.734643
$\beta_{time,cycle}$	−4.291779	−10.148372
$\beta_{time,driv}$	−4.109835	−10.974263
$\beta_{time,pt}$	−2.686530	−11.464694
$\beta_{time,walk}$	−8.062313	−16.843198
$\beta_{traffic\_percent}$	−2.858280	−12.477522

Table 3: Values of the different parameters given by model specification 2 and their respective values of the t-test.

2. Comment on the estimation output (as defined previously). [**1.0 point**]

The sign of the ASCs and βs related to time, the t-tests and the differences in β_{time} are interpreted as before. The negative sign in the value of β for traffic percentage reflects, indeed, the fact that when there is much traffic, the utility of the option car decreases. The t-test is again, in absolute value, larger than 1.96. There is then evidence to reject the hypothesis that this parameter is equal to 0. In addition, both βs related to cost are negative, again reflecting the fact that people do not want to spend more money for the same service. We can see that the values of β are slightly different when we look at people in different age groups. Young and old people (younger than 18 and older than 65) seem to be more influenced by the price than the other group because the value β for young and old people is more negative. The t-test of cost for old and young people is, in absolute value, larger than 1.96. The hypothesis that this parameter is equal to 0 can be discarded. On the contrary, this does not happen for middle-aged people. Nevertheless, we know from everyday-experiences that the cost of a service is a relevant parameter so it makes no sense to set its corresponding *parameter to zero. Lack of variability in the data can explain the low value of its t-test.*

3. Compare Model$_{\text{pref}}$ and Model 2 with an appropriate statistical test. Justify your choice of test. State the null hypothesis and the result of the test. Denote the preferred model as Model$_{\text{pref}}$. [**1.0 point**]

Since Model$_{\text{pref}}$ is a restricted version of Model 2 (the restriction is $\beta_{\text{cost,middle\_age}} = \beta_{\text{cost,young\_old}} = \beta_{\text{cost}}$ and $\beta_{\text{traffic\_percent}} = 0$), we can perform a likelihood ratio test to compare both models. As done previously, the value of the quantity $-2(\mathcal{L}_R - \mathcal{L}_U)$ is 177.08. Since there are 2 more parameters in Model 2 than in Model$_{\text{pref}}$, we use the χ^2 distribution with 2 degrees of freedom to find the threshold value corresponding to a 95% confidence interval; such a value is 5.99. Evidently, $177.29 > 5.99$, thus we have evidence to reject the null hypothesis. This means that the claims that the perception of cost is different for the two groups of people, and that driving traffic percent plays a significant role when users choosing between the mode choices are supported by data and computations. In other words, it is worth keeping the extra parameters of the unrestricted model. As a consequence, Model 2 is better than Model$_{\text{pref}}$. We now set Model$_{\text{pref}}$ equal to Model 2.

Model 3 [2.5 points] Using Model$_{\text{pref}}$ as the base model, include an appropriate non-linear transformation of one of the variables. Report both the specification and the estimation results (as defined previously). [**0.5 point**]

We decide to include a non-linear transformation (Box-Cox transform) to the duration of the trip for the four mode choices, which *leads to the deterministic parts of the utility functions shown below:*

$$
\begin{aligned}
V_{\text{walk}} &= \beta_{\text{time,walk}} \text{dur}_{\text{walk}}(\lambda) && (13)\\
V_{\text{cycle}} &= \text{ASC}_{\text{cycle}} + \beta_{\text{time,cycle}} \text{dur}_{\text{cycle}}(\lambda) && (14)\\
V_{\text{pt}} &= \text{ASC}_{\text{pt}} + \beta_{\text{time,pt}} \text{dur}_{\text{pt}}(\lambda) + \beta_{\text{cost,middle\_age}} \text{cost}_{\text{pt,middle\_age}} && (15)\\
&\quad + \beta_{\text{cost,young\_old}} \text{cost}_{\text{pt,young\_old}} \\
V_{\text{driv}} &= \text{ASC}_{\text{driv}} + \beta_{\text{time,driv}} \text{dur}_{\text{driv}}(\lambda) && (16)\\
&\quad + \beta_{\text{cost,middle\_age}} \text{cost}_{\text{driv,middle\_age}} + \beta_{\text{cost,young\_old}} \text{cost}_{\text{driv,young\_old}} \\
&\quad + \beta_{\text{traffic\_percent}} \text{traffic\_percent},
\end{aligned}
$$

where

$$
\text{dur}_{\text{transportation\_mode}}(\lambda) = \begin{cases} \dfrac{(\text{dur}_{\text{transportation\_mode}})^{\lambda} - 1}{\lambda}, & \text{if } \lambda \neq 0, \\ \log \text{dur}_{\text{transportation\_mode}}, & \text{if } \lambda = 0. \end{cases} \tag{17}
$$

We notice that $\text{dur}_{\text{transportation\_mode}} > 0$ *for all users,* $\text{dur}_{\text{transportation\_mode}}(\lambda)$ *corresponds to the Box-Cox transform of the variable* $\text{dur}_{\text{transportation\_mode}}$, $\text{transportation\_mode}$, *as usual, corresponds to walking, cycling, public transport or driving. The parameter* λ *is estimated from data during our simulations. The null and final log likelihoods of this model are given by* -6931.47 *and* -4224.29*, respectively. After running our simulation, the model parameters are given in Table 4.*

1. State the underlying assumption of the non-linear specification defined in this situation. [**0.5 point**]

The assumption is that the perception of time for the first travel minutes is different to the perception one has when travelling for a longer time; that is, for instance, the first 5 minutes of a 10-minute trip on public transportation, driving, cyclic or walking feel different than the last 5 minutes of a 40-minute trip.

2. Comment on the estimation output (as defined previously). [**0.5 point**]

Parameter	Value	t-test value
$\mathrm{ASC}_{\mathrm{cycle}}$	-1.583609	-6.049883
$\mathrm{ASC}_{\mathrm{driv}}$	2.007417	7.251523
$\mathrm{ASC}_{\mathrm{pt}}$	2.096810	10.870072
$\beta_{\mathrm{cost,middle\_age}}$	-0.080281	-1.728710
$\beta_{\mathrm{cost,young\_old}}$	-0.104890	-2.857933
$\beta_{\mathrm{time,cycle}}$	-2.771042	-9.322635
$\beta_{\mathrm{time,driv}}$	-2.521645	-8.406939
$\beta_{\mathrm{time,pt}}$	-2.155863	-10.600513
$\beta_{\mathrm{time,walk}}$	-5.374291	-14.413584
$\beta_{\mathrm{traffic\_percent}}$	-2.530745	-10.712891
λ	0.379196	-8.827881

Table 4: Values of the different parameters given by model specification 3 and their respective values of the t-test.

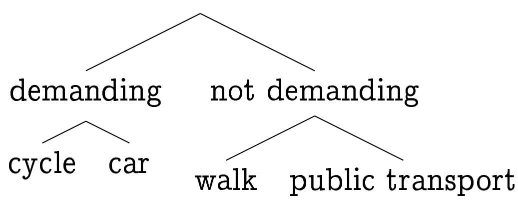
All comments from the same question of Model 2 remain true for this new model too and we obtain $\lambda = 0.379196$. *Since the t-test of the lambda parameter is greater that* 1.96 *in absolute value, the null hypothesis that the true value of* λ *is equal to 1 can be rejected with at least 95% confidence. Note that the values of the alternative specific constants are different than for the previous model. This is explained by the fact that the Box-Cox transform absorbs a part of the ASC.*

3. Compare $\mathrm{Model}_{\mathrm{pref}}$ and Model 3 with an appropriate statistical test. Justify your choice of test. State the null hypothesis and the result of the test. Denote the preferred model as $\mathrm{Model}_{\mathrm{pref}}$. [**1 point**]

Since $\mathrm{Model}_{\mathrm{pref}}$ *is a restricted version of Model 3 (with the restriction given by* $\lambda = 1$*), we can perform a likelihood ratio test to compare both models. The value of the quantity* $-2(\mathcal{L}_R - \mathcal{L}_U)$ *is 97.26. Since there is 1 more parameter in Model 3 than in* $\mathrm{Model}_{\mathrm{pref}}$*, we use the* χ^2 *distribution with 1 degree of freedom to find the threshold value corresponding* *to a 95% confidence interval; this value is 3.84. Since $97.26 > 3.84$, we can reject the null hypothesis that Model$_{\text{pref}}$ is better than Model 3 to model individuals' choice with the available data. Thus, we now set Model$_{\text{pref}}$ equal to Model 3.*

Model 4 [3 points] Using Model$_{\text{pref}}$ as the base model, propose and test a nested or cross-nested structure. Report the nesting structure by means of a graph, together with the specification and the estimation results (as defined previously). [**1 point**]

For the nesting structure, we differentiate between demanding modes of transport that require a high concentration and are more dangerous due to traffic, and not demanding modes that require no concentration and are safer. This is illustrated below:



The deterministic part of the utility functions, V, are given as in Model$_{\text{pref}}$, but now some random factors ϵ in the utilities U are correlated. The utilities are thus given by:

$$U_{\text{walk}} = V_{\text{walk}} + \epsilon_{\text{not demanding}} + \epsilon_{\text{walk}} \tag{18}$$

$$U_{\text{cycle}} = V_{\text{cycle}} + \epsilon_{\text{demanding}} + \epsilon_{\text{cycle}} \tag{19}$$

$$U_{\text{pt}} = V_{\text{pt}} + \epsilon_{\text{not demanding}} + \epsilon_{\text{pt}} \tag{20}$$

$$U_{\text{driv}} = V_{\text{driv}} + \epsilon_{\text{demanding}} + \epsilon_{\text{driv}}, \tag{21}$$

where the random terms follow extreme value distributions accordingly *to their nests:*

$$\epsilon_{\text{demanding}}, \epsilon_{\text{not demanding}} \sim EV(0, 1) \quad (22)$$

$$\epsilon_{\text{walk}}, \epsilon_{\text{pt}} \sim EV(0, \mu_{\text{not demanding}}) \quad (23)$$

$$\epsilon_{\text{cycle}}, \epsilon_{\text{driv}} \sim EV(0, \mu_{\text{demanding}}). \quad (24)$$

The null and final log likelihoods are -6931.47 *and* -4220.682*, respectively. The results of the model are given in Table 5.*

Parameter	Value	t-test value
ASC_{cycle}	-1.627140	-4.394982
ASC_{driv}	1.622294	5.402689
ASC_{pt}	1.647707	6.711181
$\beta_{\text{cost,middle\_age}}$	-0.075355	-1.740688
$\beta_{\text{cost,young\_old}}$	-0.107561	-3.074511
$\beta_{\text{time,cycle}}$	-2.678963	-9.219496
$\beta_{\text{time,driv}}$	-2.409871	-7.885460
$\beta_{\text{time,pt}}$	-2.166825	-10.181926
$\beta_{\text{time,walk}}$	-4.873209	-12.933348
$\beta_{\text{traffic\_percent}}$	-2.462150	-10.555405
λ_{cox}	0.382546	-8.3580
$\mu_{\text{not demanding}}$	1.208714	1.7424
$\mu_{\text{demanding}}$	1.172845	1.5194

Table 5: Values of the different parameters given by model specification 4 and their respective values of the t-test.

1. State the underlying assumption of the proposed nesting structure in this specific situation. [**0.5 point**]

We assumed that the amount of concentration and risk people must be willing to accept for a specific transportation mode are significant variables that correlate the travel modes inside a nest.

2. Comment on the estimation output (as defined previously). [**0.5 point**]

Both μs are greater than one, meaning that there is indeed a correlation of the error terms inside the nests. However, t-test values corresponding to the μs parameters are not greater than 1.96 as previously discussed with the other models. We then cannot reject the hypothesis that these values are equal to 1. Nevertheless, since the t-test values are not that far from the value 1.96, we think that the model is comparable with the previous one. The comments about all other values and their respective t-test are the same as for Model 3.

3. Compare $\text{Model}_{\text{pref}}$ and Model 4 with an appropriate statistical test. Justify your choice of test. State the null hypothesis and the result of the test. Denote the preferred model as $\text{Model}_{\text{pref}}$. [**1 point**]

Model 3 is just Model 4 with the restriction $\mu_{\text{not demanding}} = \mu_{\text{demanding}} = 1$. To test if Model 4 is really better than Model 3, we state, as before, the null hypothesis "Model 3 is actually the true model", equivalent to $\mu_{\text{not demanding}} = \mu_{\text{demanding}} = 1$. Under this hypothesis, the likelihood ratio test $-2(\mathcal{L}_R - \mathcal{L}_U) = 7.31$ follows the χ^2 distribution with two degrees of freedom, since Model 4 has two additional parameters than Model 3. However, since the 0.95 quantile of this distribution is 5.991, the realization $7.31 > 5.991$ is either a very rare (rarer than 0.05 occurence) event or, more likely, Model 3 is not the true model. Consequently, we reject Model 3 and choose $\text{Model}_{\text{pref}}$ to be Model 4. The small difference between 7.31 and 5.991 is probably explained by the fact that the t-test of the μs are not larger than 1.96.

Market shares [2.5 points] Assume that stratified random sampling was used to produce your sample. We consider the following strata:

- **S1:** females aged 40 years or younger;
- **S2:** females aged 41 years or older;

- **S3:** males aged 40 years or younger;
- **S4:** males aged 41 years or older.

Table 6 gives the size of each category in the full population.

	Age ≤ 40	Age > 40
Male	2'676'249	1'633'263
Female	2'599'058	1'765'143

Table 6: London population estimates in 2015 (Source: ONS)

1. Report the size and weight of each stratum in your sample [**1 point**].

Table 7 reports the size of each stratum in our sample. Concerning the weights, they are reported in Table 8.

	Age ≤ 40	Age > 40
Male	1271	1075
Female	1457	1197

Table 7: Size of each stratum in our sample.

	Age ≤ 40	Age > 40
Male	1.2138	0.8758
Female	1.0283	0.8501

Table 8: Weights of each stratum in our sample.

2. Using Model$_{\text{pref}}$ and the weights of your strata, compute the predicted market share of each mode and their confidence intervals. Do the obtained results match your expectations? [**1 point**]

Mode	market shares	confidence intervals
Walk	16.6%	[15.6%, 17.8%]
Cycle	3%	[2.5%, 3.6%]
Public transportation	35.8%	[33.6%, 37.9%]
Drive	44.6%	[42.4%, 47.0%]

Table 9: Market shares with their confidence intervals for the four choice modes.

Market shares are computed and reported in Table 9 for the preferred model (Model 4).

We remark that the majority of people prefer car over other choices. The second highest choice is public transportation. This seems reasonable knowing that many people choose public transportation for practical reasons in cities; one reason behind this could be, as an example, problematic circulation of car traffic. Moreover, some people also choose public transportation because of environmentally-friendly reasons. Walk and cycle are two less successful choice modes. For instance, this may be explained by the fact that the commuting distance from their workplace can be too large. The market shares for walking can seem bigger than expected but as there are also children in the sample. This is important since they could commute to school by this mean of transportation since schools are usually close to their home. We conclude then that results from Table 9 are coherent.

3. Compare the market shares predicted by $\text{Model}_{\text{pref}}$ with the weighted market shares computed using the actual choices. **[0.5 point]**

Comparison of the market shares from the preferred model and weighted market shares using the actual choices is made in Table 10. We observe that the values are similar (a maximum change of 0.3% for public transportation). This means that $Model_{pref}$ is capable of explaining what people choose.

Mode	ms using $\text{Model}_{\text{pref}}$	ms using actual choices
Walk	16.6%	16.7%
Cycle	3%	3%
Public transportation	35.8%	35.5%
Drive	44.6%	44.8%

Table 10: Market shares (ms) predicted by $\text{Model}_{\text{pref}}$ and market shares using choices of people for the four choice modes.

Forecasting [5 points] Consider the following scenarios: (i) an increase of car cost by 15%; and (ii) a decrease of public transport cost by 15%.

1. Report the market shares predicted by $\text{Model}_{\text{pref}}$ for each scenario. Do they match your expectations? Compare those with the original market shares. [**1 point**]

Let us consider two scenarios from $Model_{pref}$. The first one (i) an increase by 15% of car cost and the second (ii) a decrease by 15% in the cost of public transportation. Table 11 shows the obtained market shares for both scenarios and the ones obtained with the preferred model. The scenario with the increase by 15% of car cost implies a

Mode	ms for scenario (i)	ms for scenario (ii)	ms of $\text{Model}_{\text{pref}}$
Walk	16.6%	16.6%	16.6%
Cycle	3%	3%	3%
PT	36.1%	36.2%	35.8%
Drive	44.2%	44.3%	44.6%

Table 11: Market shares (ms) for the increase of car cost by 15%, decrease of public transport cost by 15% and for $\text{Model}_{\text{pref}}$ for the four choice modes.

decrease of the market shares of car by 0.4%. This makes sense as increasing the cost could discourage some people from using their cars. To compensate this, more people use public transportation (increase

of its market share by 0.3%). Scenario (ii) with the decrease of public transportation cost implies an increase of its market share by 0.4% and a decrease of market share of the car choice mode by 0.3%. The same reasoning applies. If public transportation is cheaper, then more people may be tempted to use it instead of their cars. It can also encourage people to using public transport over their bicycles.
Note that in both scenarios, market shares for walking and cycling did not change in Table 11 because of numerical approximations (only one digit). A larger difference of the cost would also have changed their market shares significantly.

2. Which scenario is the most effective policy if the goal is to decrease the share of car? Explain why. [**0.5 point**]

The first scenario because the change is made directly on the car cost. It means that those people who quit using their cars will have options to choose from, which are the three other transportation alternatives. The market shares will then split among these three options and to quantify this, we shall look at the choice probability of each of these.

3. Which scenario reports the highest public transportation total revenue? Explain why. Is it higher than the total revenue obtained without any of the policies? Can you explain why? [**0.5 point**]

Total revenue for public transportation is 3557.2 GBD *using* $Model_{pref}$, 3594.2 GBD *with scenario (i) and* 3073.4 GBD *with scenario (ii). Scenario (i) has the highest total revenue and is higher than total revenue without any policies. It is coherent since the cost for public transportation is similar but market shares have increased with the change made in scenario (i). Now if we compare scenario (i) with (ii), market shares of the second one are greater. However scenario (ii) does not have a higher total revenue because its cost has been reduced by 15% and it has a more important effect than the increase of its market* *share.*

4. Calculate the average value of time for car and public transportation. Comment on the obtained results. [**1 point**]

On simple linear models, the value of time is just given by $\beta_{\text{time}}/\beta_{\text{cost}}$. *Our model however contains a Box Cox transformation, so we need to dive deeper into the mathematics. The value of time when using such a transformation is given by*

$$\text{VOT} = \frac{\partial V/\partial \text{dur}}{\partial V/\partial \text{cost}} = \frac{\beta_{\text{dur}} \text{dur}^{\lambda-1}}{\beta_{\text{cost}}}, \tag{25}$$

which depends on the duration of the trip. This is why the value of time has to be calculated separately for each individual, and only then can the average be taken. The resulting values of time in GBP per hour are shown in Table 12.

Population	Car	Public Transport
young_old	71.15 GBD/hour	42.00 GBD/hour
middle_age	88.25 GBD/hour	54.52 GBD/hour
mean	79.70 GBD/hour	48.26 GBD/hour

Table 12: Value of time for car and public transport in the age groups young_old and middle_age defined in Model 2. Note that the mean is weighted by group sizes.

Table 12 reveals two facts. Firstly, middle-aged people would on average pay more money to save time than young or old people. The two possible causes are that they value their time more because they have less of it due to work, and that they can afford to pay more for it thanks to having a salary. It is likely a combination of the two. Secondly, people who take the car would on average spend more money to decrease travel time than people who take public transports. This *can be because one's time can be used productively in public transportation compared to in a car, so spending more time on a train does not penalize utility as much as spending more time in a car. Another reason could be that people who value their time more tend to choose the car over public transportation because it is often faster and more comfortable. Again, the cause is probably a combination of the two.*

5. Compute the direct and cross aggregate elasticities of car cost and public transport cost and comment on the obtained results. Report the normalization factors. [**2 points**]

From the lecture notes, the disaggregate direct elasticity of public transportation cost is given by (26)*:*

$$E^{P_n(PT)}_{costPT} = \frac{\partial P_n(PT)}{\partial costPT}\frac{costPT}{P_n(PT)}, \tag{26}$$

whereas, the direct aggregate elasticity of public transportation is given by (27)

$$E^{\widehat{W}(PT)}_{costPT} = \frac{\partial \widehat{W}(PT)}{\partial costPT}\frac{costPT}{\widehat{W}(PT)} = \frac{1}{\sum_{\ell=1}^{N}\omega_\ell P_\ell(PT)}\sum_{n=1}^{N}\omega_n P_n(PT)E^{P_n(PT)}_{costPT}. \tag{27}$$

Similarly, the disaggregate cross elasticity of public transportation is given by (28)

$$E^{P_n(PT)}_{costCar} = \frac{\partial P_n(PT)}{\partial costCar}\frac{costCar}{P_n(PT)}. \tag{28}$$

Then, the cross aggregate elasticity of public transportation is given by (29)

$$E^{\widehat{W}(PT)}_{costCar} = \frac{\partial \widehat{W}(PT)}{\partial costCar}\frac{costCar}{\widehat{W}(PT)} = \frac{1}{\sum_{\ell=1}^{N}\omega_\ell P_\ell(PT)}\sum_{n=1}^{N}\omega_n P_n(PT)E^{P_n(PT)}_{costCar}. \tag{29}$$

In (27) *and* (29)*, we notice that ω_n's correspond to the weights of the strata defined in question* Market shares*, and $\widehat{W}(PT)$ corresponds to the market shares of public transportation computed in point 1 of question*

Market shares. Besides, we refer to the quantity $\sum_{\ell=1}^{N} \omega_\ell P_\ell(\mathrm{PT})$ *in* (27) *and* (29) *as normalization factor; we compute one for public transportation and another for car. To compute the elasticity of car cost, we accordingly modify expressions* (26), (27), (28), *and* (29). *After running our simulations, we have the elasticity values shown in Table 13.*

We observe that direct aggregate elasticities are quite low. In fact, we have $|E^{\widehat{W}(\mathrm{PT})}_{\mathrm{costPT}}| < 1$ *and* $|E^{\widehat{W}(\mathrm{Car})}_{\mathrm{costCar}}| < 1$. *When the value of the elasticity is less than one, it is common to say that the demand is inelastic. In other words, if the cost of either mode of transportation increases by a small amount, the probability to choose that mode would decrease by a small quantity too.*

Mode	Norm. factors	Direct agg. elasticity	Cross agg. elasticity
PT	1788.65	$E^{\widehat{W}(\mathrm{PT})}_{\mathrm{costPT}} = -0.0722$	$E^{\widehat{W}(\mathrm{PT})}_{\mathrm{costCar}} = 0.0546$
Car	2229.07	$E^{\widehat{W}(\mathrm{Car})}_{\mathrm{costCar}} = -0.0518$	$E^{\widehat{W}(\mathrm{Car})}_{\mathrm{costPT}} = 0.0468$

Table 13: Elasticities for public transportation and car.Norm. factors stands for normalization factors, and agg. for aggregate.

Similarly, we observe that for both modes of transportation $|E^{\widehat{W}(\mathrm{PT})}_{\mathrm{costCar}}| < 1$ *and* $|E^{\widehat{W}(\mathrm{Car})}_{\mathrm{costPT}}| < 1$, *which means that when the price of car is modified, this has no significant effect on the demand for public transportation usage; likewise, when the price of public transportation changes, this has no significant effect on the demand for car usage.*

In Table 13, it is possible to see that the values of the direct aggregate elasticities are negative while the values of the cross aggregate elasticities are positive which is expected; when the price of an alternatives is raised, this alternatives will be chosen less often while the others will be chosen more frequently.