## Mathematical Modeling of Behavior (2022-2023)

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

Course assignment — October 11, 2022

**Group 10: Zixuan Fu, Yihan Wang, Wissam Nael Pheng, Germain Quoc-Huan Vu**

**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**]

**Answer** We first define the cost of driving `cost_driving` as the sum of `cost_driving_fuel` and `cost_driving_ccharge`, and the total duration of travelling via public transports `dur_public` as `dur_pt_access` + `dur_pt_rail` + `dur_pt_bus` + `dur_pt_int`.

**Specification**

$$
\begin{aligned}
V_{\text{walk}} &= ASC_{\text{walk}} + \beta_{\text{time}} * \text{dur\_walk} \\
V_{\text{cycling}} &= ASC_{\text{cycling}} + \beta_{\text{time}} * \text{dur\_cycling} \\
V_{\text{public}} &= ASC_{\text{public}} + \beta_{\text{time}} * \text{dur\_public} + \beta_{\text{cost}} * \text{cost\_transit} \\
V_{\text{driving}} &= ASC\_\text{driving} + \beta_{\text{time}} * \text{dur}_{\text{driving}} + \beta_{\text{cost}} * \text{cost\_driving}
\end{aligned}
$$

**Estimation results** See table 1

| | Value | Rob. Std err | Rob. t-test | Rob. p-value |
|---|---|---|---|---|
| asc_cycling | -3.66 | 0.101 | -36.3 | 0.0 |
| asc_driving | -1.34 | 0.0764 | -17.5 | 0.0 |
| asc_public | -0.559 | 0.0526 | -10.6 | 0.0 |
| beta_cost | -0.15 | 0.0138 | -10.9 | 0.0 |
| beta_time | -5.41 | 0.188 | -28.7 | 0.0 |

Table 1: Parameter estimates

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

All the parameters are relevant at the significance level $\alpha = 0.05$. The parameters associated to cost and time are both negative which is logic as people are less likely to choose expansive and time-consuming modes of transportation.

**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]**

**Answer** We decided to include alternative-specific parameters for time attribute of Model 0 since time spent in each transport mean is felt differently. For instance, spending one hour in a car is completely different than spending one hour by walk, for example due to comfort. For that, we thus introduce $\beta_{\text{time,walk}}$, $\beta_{\text{time,cycling}}$, $\beta_{\text{time,driving}}$ and $\beta_{\text{time,public}}$ in our model. The latter thus becomes :

**Specification**

$$\begin{aligned}
V_{\text{walk}} &= \beta_{\text{time,walk}} * \text{dur_walk} \\
V_{\text{cycling}} &= \beta_{\text{time,cycling}} * \text{dur_cycling} \\
V_{\text{public}} &= \beta_{\text{time,public}} * \text{dur_public} + \beta_{\text{cost}} * \text{cost_transit} \\
V_{\text{driving}} &= \beta_{\text{time,driving}} * \text{dur}_{\text{driving}} + \beta_{\text{cost}} * \text{cost_driving}
\end{aligned}$$

**Estimation results** See table 2

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

| | Value | Rob. Std err | Rob. t-test | Rob. p-value |
|---|---|---|---|---|
| asc_cycling | -4.59 | 0.179 | -25.6 | 0.0 |
| asc_driving | -2.07 | 0.12 | -17.2 | 0.0 |
| asc_public | -2.44 | 0.122 | -20.0 | 0.0 |
| beta_cost | -0.142 | 0.0152 | -9.33 | 0.0 |
| beta_time_cycling | -5.2 | 0.424 | -12.3 | 0.0 |
| beta_time_driving | -5.88 | 0.36 | -16.3 | 0.0 |
| beta_time_public | -3.2 | 0.231 | -13.9 | 0.0 |
| beta_time_walk | -8.37 | 0.36 | -23.2 | 0.0 |

Table 2: Parameter estimates

**Answer** Modeling assumption : a minute has/has not the same marginal utility whether it is occurred on the auto or bus mode.

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

**Answer** All the parameters are significant as their p-values are below the $\alpha = 0.05$ threshold. They are still negative which is again logic for the same reason as in Model 0.

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**]

**Answer** We proceed to the likelihood ratio test since the latter is applied when one of the two (here Model 0) is a restricted version of the other (Model 1), and the restriction is characterized by linear constraints on the parameters. We first extract the log-likelihood of each model ($\mathcal{L}(0)$ being the one of Model 0, the restricted model, and $\mathcal{L}(1)$ the unrestricted model) from their general statistics. Computing the ratio test yields :

$$-2\left(\mathcal{L}(0) - \mathcal{L}(1)\right) \approx 551.4407578162431.$$

Since one attribute is alternatively specified and there are four utilities, the likelihood ratio follows a $\chi^2$ law with 3 degrees of freedom. The 95% quantile is approximately 7.815, and since the ratio is clearly bigger we can conclude that Model 1 is significantly better (we reject the null hypothesis that all the alternative specific parameters for time are equal).

**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**]

**Specification** Alternative attribute: predicted traffic variability on driving route.
Individual characteristics: Whether the person has a driving license.

$$
\begin{aligned}
V_{\text{walk}} &= ASC_{\text{walk}} + \beta_{\text{time,walk}} * \text{dur\_walk} \\
V_{\text{cycling}} &= ASC_{\text{cycling}} + \beta_{\text{time,cycling}} * \text{dur\_cycling} \\
V_{\text{public}} &= ASC_{\text{public}} + \beta_{\text{time,public}} * \text{dur\_public} + \beta_{\text{cost}} * \text{cost\_transit} \\
V_{\text{driving}} &= \text{ASC\_driving} + \beta_{\text{time,driving}} * \text{dur}_{\text{driving}} + \beta_{\text{cost}} * \text{cost\_driving} \\
&\quad + \beta_{\text{driving\_traffic}} * \text{driving\_traffic\_percentage} \\
&\quad + \beta_{\text{drivinglicens}} * \text{driving\_license}
\end{aligned}
$$

**Estimation results** See table 3.

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

**Answer** We introduce predicted traffic variability on driving route because we believe that the agent makes a travel mode decision before actually spending time on route. Therefore, the expected travel time and its variability plays an important role in decision-making.
We interact driving license with ASC_driving because we believe that whether the agent holds a license affect how easily he can choose cars.

| | Value | Rob. Std err | Rob. t-test | Rob. p-value |
|---|---|---|---|---|
| asc_cycling | -4.66 | 0.186 | -25.0 | 0.0 |
| asc_driving | -2.41 | 0.138 | -17.5 | 0.0 |
| asc_public | -2.62 | 0.128 | -20.4 | 0.0 |
| beta_cost | -0.0899 | 0.0151 | -5.97 | 2.4e-09 |
| beta_driving_traffic_percent | -2.96 | 0.236 | -12.5 | 0.0 |
| beta_drivingliscence | 1.44 | 0.0712 | 20.3 | 0.0 |
| beta_time_cycling | -4.78 | 0.438 | -10.9 | 0.0 |
| beta_time_driving | -4.42 | 0.37 | -11.9 | 0.0 |
| beta_time_public | -2.83 | 0.235 | -12.0 | 0.0 |
| beta_time_walk | -8.37 | 0.37 | -22.6 | 0.0 |

Table 3: Parameter estimates

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

**Answer** The added variables are statistically significant. More specifically, betadriving_traffic_percent negatively affect $V_d riving$, beta_drivinglicense positively affect $V_d riving$. The signs are consistent with our expectation.

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**]

**Answer** Null hypothesis: Model 1 (restricted model) is the true model. Since model 2 is the less restricted version of model 1 (we relax the assumption that the additional parameters are 0), we can apply the likelihood ratio test with 2 degrees of freedom.

$$-2\left(\mathcal{L}(1) - \mathcal{L}(2)\right) \approx 591.7382206371785 > 5.991464547107979(\text{threshold})$$

Thus, the null hypothesis can be rejected at level 5%.

**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**]

**Specification** Box-cox transformation on time parameters.

$$
\begin{aligned}
V_{\text{walk}} &= ASC_{\text{walk}} + \beta_{\text{time,walk}} * \frac{\text{dur\_walk}^{\lambda} - 1}{\lambda} \\
V_{\text{cycling}} &= ASC_{\text{cycling}} + \beta_{\text{time,cycling}} * \frac{\text{dur\_cycling}^{\lambda} - 1}{\lambda} \\
V_{\text{public}} &= ASC_{\text{public}} + \beta_{\text{time,public}} * \frac{\text{dur\_public}^{\lambda} - 1}{\lambda} + \beta_{\text{cost}} * \text{cost\_transit} \\
V_{\text{driving}} &= ASC\_driving + \beta_{\text{time,driving}} * \frac{\text{dur\_driving}^{\lambda} - 1}{\lambda} + \beta_{\text{cost}} * \text{cost\_driving} \\
&\quad + \beta_{\text{driving\_traffic}} * \text{driving\_traffic\_percentage} \\
&\quad + \beta_{\text{drivinglicens}} * \text{driving\_license}
\end{aligned}
$$

**Estimation results** See table 4.

| | Value | Rob. Std err | Rob. t-test | Rob. p-value |
|---|---|---|---|---|
| asc_cycling | -1.81 | 0.279 | -6.46 | 1.03e-10 |
| asc_driving | 0.814 | 0.274 | 2.97 | 0.00295 |
| asc_public | 1.97 | 0.174 | 11.4 | 0.0 |
| beta_cost | -0.0915 | 0.0146 | -6.27 | 3.55e-10 |
| beta_driving_traffic_percent | -2.6 | 0.241 | -10.8 | 0.0 |
| beta_drivinglicense | 1.45 | 0.0718 | 20.1 | 0.0 |
| beta_time_cycling | -3.03 | 0.277 | -10.9 | 0.0 |
| beta_time_driving | -2.65 | 0.27 | -9.82 | 0.0 |
| beta_time_public | -2.22 | 0.192 | -11.6 | 0.0 |
| beta_time_walk | -5.38 | 0.299 | -18.0 | 0.0 |
| lambda_boxcox | 0.351 | 0.0583 | 6.01 | 1.87e-09 |

Table 4: Parameter estimates

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

**Answer** We believe that agents do not feel the same about each additional minutes of time spent in transportation. More specifically, they became less sensitive to time when the duration is already long. The $\lambda$

coefficient capture this phenomenon. The smaller the $\lambda$, the less sensitive agents become when time is large. (When $\lambda = 0$, the time variable is log transformed.)

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

**Answer** Notice that the expected $\lambda < 1$ which is consistent with our expectation.
Also, the size of the $\beta$_$time$_$i$ becomes smaller for all 4 alternatives.
All ASCs change by a significant amount but this is due to the $\frac{-1}{\lambda}$ in the box transformation

3. Compare Model$_{\text{pref}}$ 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 Model$_{\text{pref}}$. [**1 point**]

The null hypothesis is $\lambda = 1$.
Under the null hypothesis that $\lambda = 1$, the statistic

$$\frac{\widehat{\lambda} - 1}{\widehat{\sigma}_{\lambda}}$$

follows approximatively a $N(0, 1)$. We perform a t-test on $\lambda$, where the t-statistics is $11.130021870608987 >> 1.96$. Thus, the null hypothesis is rejected.

**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**]

**Specification** First of all, we define a fourth model as a nested version of Model 3 where we regroup motorized (public transports and car), as a nest, and walking and cycling as two other nests. The nesting structure can be reported by the following graph:

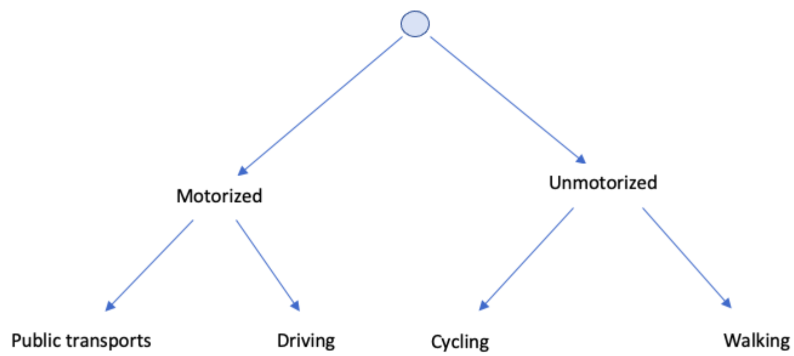


Figure 1: Nested model

| Name | Value | Rob. Std err | Rob. t-test | Rob. p-value |
|---|---|---|---|---|
| MU_motorized | 1.892844 | 0.245826 | 7.699925 | 1.354472e-14 |
| MU_non_motorized | 0.686478 | 0.071778 | 9.563958 | 0.000000e+00 |
| asc_cycling | -0.316613 | 0.294821 | -1.073916 | 2.828603e-01 |
| asc_driving | 2.371441 | 0.301763 | 7.858618 | 3.774758e-15 |
| asc_public | 3.128218 | 0.235327 | 13.293061 | 0.000000e+00 |
| beta_cost | -0.039085 | 0.010655 | -3.668285 | 2.441831e-04 |
| beta_driving_traffic_percent | -1.486760 | 0.240186 -6.190039 | 6.014937e-10 | |
| beta_drivingliscence | 0.895696 | 0.107427 | 8.337745 | 0.000000e+00 |
| beta_time_cycling | -1.323322 | 0.259310 | -5.103240 | 3.338869e-07 |
| beta_time_driving | -1.606233 | 0.247199 | -6.497742 | 8.153433e-11 |
| beta_time_public | -1.340126 | 0.197290 | -6.792657 | 1.100875e-11 |
| beta_time_walk | -4.935801 | 0.347756 | -14.193270 | 0.000000e+00 |
| lambda_boxcox | 0.335751 | 0.059858 | 5.609143 | 2.033313e-08 |

Table 5: Parameter estimates

**Estimation results** See table 5.

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

**Answer** We introduce the scale parameters $\mu_{\text{motorized}}$ and $\mu_{\text{non\_motorized}}$ associated within the nests "motorized" and "non motorized".

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

**Answer** As $\mu_{\text{non\_motorized}}$ is $< 1$, we reject the nested model.

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**]

**Answer** Eventually, Model 3 is preferred over the nested model.

**Additional answer** We also thought about a cross-nested model to see if the latter is better than model 4. Let's call it model 5. For that, instead of creating nests on motorized/unmotorized vehicles, we create nests on private/public vehicles.
At the end, we got the following results in table 6:

| Name | Value | Rob. Std err | Rob. t-test | Rob. p-value |
|---|---|---|---|---|
| MU_private | 0.650852 | 0.042759 | 15.221570 | 0.000000e+00 |
| asc_cycling | -1.516913 | 0.388054 | -3.909025 | 9.266953e-05 |
| asc_driving | 2.213199 | 0.406244 | 5.447951 | 5.095350e-08 |
| asc_public | 3.368701 | 0.358377 | 9.399888 | 0.000000e+00 |
| beta_cost | -0.083946 | 0.016215 | -5.176952 | 2.255398e-07 |
| beta_driving_traffic_percent | -3.188256 | 0.288127 | -11.065462 | 0.000000e+00 |
| beta_driving_license | 1.777428 | 0.100008 | 17.772867 | 0.000000e+00 |
| beta_time_cycling | -3.101769 | 0.319821 | -9.698449 | 0.000000e+00v |
| beta_time_driving | -2.767161 | 0.286049 | -9.673735 | 0.000000e+00 |
| beta_time_public | v-2.645343 | 0.228022 | -11.601266 | 0.000000e+00 |
| beta_time_walk | -6.980728 | 0.517760 | -13.482544 | 0.000000e+00 |
| lambda_boxcox | 0.341023 | 0.060617 | 5.625911 | 1.845316e-08 |

Table 6: Parameter estimates

However, Since $\mu_{\text{private}} < 1$, we should also reject this hypothesis. Eventually, model 3 remains the best one.

**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 7 gives the size of each category in the full population.

| | Age $\leq 40$ | Age $> 40$ |
|---|---|---|
| Male | 2'676'249 | 1'633'263 |
| Female | 2'599'058 | 1'765'143 |

Table 7: London population estimates in 2015 (Source: ONS)

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

| | Age $\leq 40$ | Age $> 40$ |
|---|---|---|
| Male | 1256 | 1084 |
| Female | 1456 | 1204 |

Table 8: Sample size of each stratum

**Answer** Let S the size of the sample and N the size of the full population. For any stratum $\sigma$, the corresponding weight is computed by $w_\sigma := \frac{N_\sigma}{N}\frac{S}{S_\sigma}$ where $N_\sigma, S_\sigma$ are the stratum's total and sample sizes respectively. We sum up the weights as following :

| | Age $\leq 40$ | Age $> 40$ |
|---|---|---|
| Male | 1.228 | 0.869 |
| Female | 1.029 | 0.845 |

Table 9: Sample weight of each stratum

2. Using $\text{Model}_{\text{pref}}$ 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**]

**Answer**

- Market share for walk: 18.3% and the confidence interval is: [ 17.4% ; 19.2% ]
- Market share for cycling: 3.6% and the confidence interval is: [ 3.2% ; 4.3% ]
- Market share for public: 36.3% and the confidence interval is: [ 34.5% ; 38.1% ]
- Market share for driving: 41.9% and the confidence interval is: [ 39.7% ; 43.8% ]

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

**Answer** To compute the weighted market shares using the actual choices, we sum the frequency of each transportation mode for each stratum multiplied by the corresponding weight. The results are the following :

- Market share for walk: 18.5
- Market share for cycling: 3.7
- Market share for public: 36.1
- Market share for driving: 41.7

All the values are in the previously computed confidence intervals, which shows consistency with the model.

**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**]

**Answer** The predicted market shares for different scenarios are summarized as the following table:

| **Scenario** | **Walking** | **Cycling** | **Public Transport** | **Driving** |
|---|---|---|---|---|
| **Original** | 18.37% | 3.59% | 36.26% | 41.78% |
| **Scenario I** | 18.40% | 3.61% | 36.52% | 41.47% |
| **Scenario II** | 18.33% | 3.56% | 36.60% | 41.51% |

For both of the scenarios, the market shares of public transport increase compared to the original scenario, while those of driving decrease. Market shares of walking and cycling only increase and decrease slightly, in scenario I and II, respectively. This is in line with the expectations, since in the model, the coefficient of cost is negative, thus in scenario I, increasing the car cost means reducing the utility of driving, the probability of choosing other alternatives will grow. Meanwhile, in scenario II, decreasing the public transport cost means increasing the utility of public transport, the probability of choosing other alternatives will drop.

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

**Answer** Scenario I is the most effective policy in order to decrease the share of car, from 41.78% to 41.47%, because it directly changes the attribute of corresponding alternative, which is the cost of driving. In this scenario, the only alternative with a reduced probability of choice is driving.

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**]

**Answer** The predicted public transportation total revenue (in GBP) for different scenario is summarized as the following table:

| Original | Scenario I | Scenario II |
|---|---|---|
| 3495.0 | 3532.2 | 3016.6 |

Table 10: PT revenue

Scenario I reports the highest public transportation total revenue. Because the market share of public transport in this scenario increases without a reduction in public transport fee. It is higher that the total revenue obtained without any of the policies, because in this scenario, the number of people choosing public transportation becomes larger in comparison, while the cost per person using public transportation remains the same.

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

**Answer** The formula for calculating VOT is as following:

$$\mathrm{VOT}_{in} = \frac{\delta^c_{in}}{\delta^t_{in}} = \frac{(\frac{\partial V_{in}}{\partial t_{in}})(c_{in}, t_{in})}{(\frac{\partial V_{in}}{\partial c_{in}})(c_{in}, t_{in})}$$

$$\mathrm{VOT}_i = \sum(\omega_n \mathrm{VOT}_{in})$$

The predicted VOT for car is 97 GBP/hour, 50% higher than that for public transportation, which is 61.5 GBP/hour. People are more willing to pay for saving time when driving, which further reflects the difference between one more minute spent in public transportation(where people can spare to do other things) and that spent in driving.
However, we also recognize that the magnitude of the VOT reported is

higher than usual. We have investigated the issue but didn't come to a conclusion. In all our models (from 0 to 4), the $\beta_{cost}$ is very small while VOT is too high.

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**]

**Answer** Direct disaggregate elasticity:

$$E^{p_n(i)}_{x_{ink}} = \frac{\partial P_n(i)}{\partial x_{ink}} \frac{x_{ink}}{P_n(i)}$$

Cross disaggregate elasticity:

$$E^{p_n(i)}_{x_{jnk}} = \frac{\partial P_n(i)}{\partial x_{jnk}} \frac{x_{jnk}}{P_n(i)}$$

Direct aggregate elasticity:

$$E^{\hat{W}(i)}_{x_{ik}} = \frac{1}{\sum_{l=1}^{S} \omega_l P_l(i)} \sum \omega_n P_n(i) E^{p_n(i)}_{x_{ink}}$$

Cross aggregate elasticity:

$$E^{\hat{W}(i)}_{x_{jk}} = \frac{1}{\sum_{l=1}^{S} \omega_l P_l(i)} \sum \omega_n P_n(i) E^{p_n(i)}_{x_{jnk}}$$

The predicted aggregate direct elasticity, cross elasticity and normalization factors for car cost and public transport cost are summarized as the following table:

| Alternative | Direct aggregate | Cross aggregate | Normalization Factor |
|---|---|---|---|
| **Public Transportation** | -0.064 | 0.051 | 1812.84 |
| **Driving** | -0.051 | 0.044 | 2089.20 |

Note that the aggregate elasticity of the cost for both alternatives is lower than one (in absolute value), thus we can say that the demand is inelastic for cost, which means that changes in cost of the two modes would not result in significant demand. Also, note that the values for direct aggregate elasticity are less than 0, which is the opposite to cross aggregate elasticity. This is in line with expectancy, for increasing one alternative's cost will discourage the choice of this alternative and increase that of other alternatives.