EPFL
ENAC TRANSP-OR
Prof. M. Bierlaire - Dr. de Lapparent - Dr. Sharif

Decision-aid Methodologies in Transportation
Spring 2016

Swissmetro MNL

1 Model Specification with Generic Attributes

Files to use with Biogeme:
Model file: *MNL_SM_generic.mod*
Data file: *swissmetro.dat*

The dataset consists of survey data collected on the trains between St. Gallen and Geneva in Switzerland. The idea is to analyze the impact of modal innovation in transportation, represented by the Swissmetro, against the more classic types of transport modes. The choice variable consists of three alternatives: train, Swissmetro and car (for car owners). In this first model specification, we assume that travel time, cost and headway of public transportation modes influence the utility functions. We also assume that the coefficients of the explanatory variables are generic, that is, they do not vary over the alternatives. The corresponding expressions of the utilities are defined as follows:

$$\begin{aligned} V_{\mathtt{car}} &= \mathrm{ASC}_{\mathrm{car}} + \beta_{\mathrm{time}}\mathrm{CAR\_TT} + \beta_{\mathrm{cost}}\mathrm{CAR\_CO} \\ V_{\mathtt{train}} &= \beta_{\mathrm{time}}\mathrm{TRAIN\_TT} + \beta_{\mathrm{cost}}\mathrm{TRAIN\_COST} + \beta_{\mathrm{he}}\mathrm{TRAIN\_HE} \\ V_{\mathtt{SM}} &= \mathrm{ASC}_{\mathrm{SM}} + \beta_{\mathrm{time}}\mathrm{SM\_TT} + \beta_{\mathrm{cost}}\mathrm{SM\_COST} + \beta_{\mathrm{he}}\mathrm{SM\_HE} \end{aligned}$$

where CAR_TT is the car travel time, CAR_CO is the car cost, TRAIN_TT is the train travel time, TRAIN_COST is the train cost (considering the ownership of Swiss annual season ticket, GA), TRAIN_HE is train headway (in minutes), SM_TT is the Swissmetro travel time, SM_COST is the Swissmetro cost (considering the ownership of GA), and SM_HE is the Swissmetro headway.

The estimation results are shown in Table 1. For estimation purposes, we have normalized the alternative specific constant of train to zero. The estimated values for the alternative specific constants $\mathrm{ASC}_{\mathrm{car}}$ and $\mathrm{ASC}_{\mathrm{SM}}$ show that, all the rest remaining constant, there is a preference in the choice of car and Swissmetro with

Logit model with generic attributes				
Parameter number	Parameter name	Parameter estimate	Robust standard error	Robust *t statistic*
1	ASC_{car}	0.189	0.0798	2.37
2	ASC_{SM}	0.451	0.0932	4.84
3	β_{cost}	-0.0108	0.000682	-15.90
4	β_{he}	-0.00535	0.000983	-5.45
5	β_{time}	-0.0128	0.00104	-12.23
Summary statistics				
Number of observations = 6768				
$\mathcal{L}(0) = -6964.663$				
$\mathcal{L}(\hat{\beta}) = -5315.386$				
$\bar{\rho}^2 = 0.236$				

Table 1: Logit model with generic attributes

respect to train. Moreover, the higher value of ASC_{SM} shows a greater preference for Swissmetro compared to car. As expected, both the travel time and cost coefficients have negative signs. The higher the travel time or the cost of an alternative, the lower the related utility. The negative estimate of the headway coefficient β_{he} indicates that the higher the headway, the lower the frequency of service, and thus the lower the utility.

2 Model Specification with Alternative Specific Attributes

Files to use with Biogeme:
Model file: *MNL_SM_specific.mod*
Data file: *swissmetro.dat*

In this second model, we relax the hypothesis of generic coefficients. To illustrate this idea, we use three different cost coefficients, one for each alternative. The corresponding utility functions are

$$
\begin{aligned}
V_{\texttt{car}} &= \text{ASC}_{\text{car}} + \beta_{\text{time}}\text{CAR\_TT} + \beta_{\text{car\_cost}}\text{CAR\_CO} \\
V_{\texttt{train}} &= \beta_{\text{time}}\text{TRAIN\_TT} + \beta_{\text{train\_cost}}\text{TRAIN\_COST} + \beta_{\text{he}}\text{TRAIN\_HE} \\
V_{\texttt{SM}} &= \text{ASC}_{\text{SM}} + \beta_{\text{time}}\text{SM\_TT} + \beta_{\text{SM\_cost}}\text{SM\_COST} + \beta_{\text{he}}\text{SM\_HE}.
\end{aligned}
$$

Logit model with alternative specific travel cost				
Parameter number	Parameter name	Parameter estimate	Robust standard error	Robust *t statistic*
1	ASC_{car}	-0.971	0.134	-7.22
2	ASC_{SM}	-0.444	0.102	-4.34
3	$\beta_{\text{car\_cost}}$	-0.00949	0.00116	-8.21
4	β_{he}	-0.00542	0.00101	-5.36
5	$\beta_{\text{SM\_cost}}$	-0.0109	0.000703	-15.49
6	β_{time}	-0.0111	0.00120	-9.26
7	$\beta_{\text{train\_cost}}$	-0.0293	0.00169	-17.32
Summary statistics				
Number of observations = 6768				
$\mathcal{L}(0) = -6964.663$				
$\mathcal{L}(\hat{\beta}) = -5068.559$				
$\bar{\rho}^2 = 0.271$				

Table 2: Logit model with alternative-specific cost attributes

The estimation results for this model specification are shown in Table 2. The results show the significance of the alternative-specific cost coefficients. The influence of the cost is different, showing a larger negative impact on the train alternative with respect to car and Swissmetro. In this model, the ASC's are negative implying a preference, with all the rest constant, for the train alternative. These results are different from those of the previous model where $\text{ASC}_{\texttt{car}}$ and $\text{ASC}_{\texttt{SM}}$ were positive and significant. The larger negative value of $\text{ASC}_{\texttt{car}}$ implies that this alternative is more negatively perceived with respect to train than the Swissmetro alternative. Considering that the deterministic utilities are very simple, only including three explanatory variables, the alternative specific constants can capture various effects. Their signs and magnitudes should therefore be further investigated.

2.1 Generic vs. Specific Test

To test whether a coefficient should be generic or alternative-specific, we use the likelihood ratio test (see pages 28 and 164-167 in [?]). We compare the log likelihood functions of the restricted and unrestricted models of interest. The restricted model includes generic travel cost coefficients over the three alternatives, and the unrestricted model includes alternative-specific travel cost coefficients. Hence, the null hypothesis is

$$H_0 : \beta_{\text{car\_cost}} = \beta_{\text{train\_cost}} = \beta_{\text{SM\_cost}}$$

and the test statistic for the null hypothesis is given by

$$-2(\mathcal{L}_R - \mathcal{L}_U)$$

which is asymptotically distributed as χ^2 with $df = K_U - K_R$ degrees of freedom, where K_U and K_R are the numbers of estimated parameters in the unrestricted and restricted models, respectively. We reject the null hypothesis that the restrictions are true if

$$-2(\mathcal{L}_R - \mathcal{L}_U) > \chi^2_{((1-\alpha),df)}$$

where α is the level of significance. In this specific case, using $\alpha = 0.05$ yields

$$-2(-5315.386 + 5068.559) = 493.654 > 5.991$$

We can therefore reject the null hypothesis and conclude that the travel cost coefficient should be alternative-specific.

Model Specification with Socio-Economic Characteristics

Files to use with Biogeme:
Model file: *MNL_SM_socioec.mod*
Data file: *swissmetro.dat*

To capture the average of the differences between the individuals in the sample, we make use of socio-economic characteristics. These types of variables do not change over the choice set and are individual specific. In this example, we add two variables to the model: a dummy variable (SENIOR) for senior people (age above 65) and a dummy variable that captures the effect of the Swiss annual season ticket for train (GA). A few observations, where the variable AGE is unknown (coded as 6), are removed from the estimation. The deterministic utilities are:

$$
\begin{aligned}
V_{\text{car}} &= \text{ASC}_{\text{car}} + \beta_{\text{time}}\text{CAR\_TT} + \beta_{\text{car\_cost}}\text{CAR\_CO} + \beta_{\text{senior}}\text{SENIOR} \\
V_{\text{train}} &= \beta_{\text{time}}\text{TRAIN\_TT} + \beta_{\text{train\_cost}}\text{TRAIN\_COST} + \beta_{\text{he}}\text{TRAIN\_HE} + \\
&\quad \beta_{\text{ga}}\text{GA} \\
V_{\text{SM}} &= \text{ASC}_{\text{SM}} + \beta_{\text{time}}\text{SM\_TT} + \beta_{\text{SM\_cost}}\text{SM\_COST} + \beta_{\text{he}}\text{SM\_HE} + \\
&\quad \beta_{\text{senior}}\text{SENIOR} + \beta_{\text{ga}}\text{GA}
\end{aligned}
$$

Logit model with socio-economic variables

Parameter number	Parameter name	Parameter estimate	Robust standard error	Robust *t statistic*
1	ASC_{car}	-0.608	0.143	-4.24
2	ASC_{SM}	-0.135	0.106	-1.26
3	$\beta_{\text{car\_cost}}$	-0.00936	0.00117	-8.02
4	β_{he}	-0.00586	0.00106	-5.55
5	$\beta_{\text{SM\_cost}}$	-0.0104	0.000744	-14.02
6	β_{time}	-0.0111	0.00121	-9.20
7	$\beta_{\text{train\_cost}}$	-0.0268	0.00176	-15.24
8	β_{senior}	-1.88	0.109	-17.31
9	β_{ga}	0.557	0.191	2.91

Summary statistics
Number of observations = 6759
$\mathcal{L}(0) = -6958.425$
$\mathcal{L}(\hat{\beta}) = -4927.167$
$\bar{\rho}^2 = 0.291$

Table 3: Logit model with socio-economic variables

The estimation results for this model are shown in Table 3. The coefficients of the socio-economic variables have been estimated and are significantly different from zero at a 95% confidence level. The negative sign of the age coefficient (referring to SENIOR dummy variable) reflects a preference of older individuals for the train alternative. It seems a reasonable conclusion, dictated probably by safety reasons with respect to the car choice and a kind of "inertia" with respect to the modal innovation represented by the Swissmetro alternative. The coefficient related to the ownership of the Swiss annual season ticket (GA) is positive, as expected. It reflects a preference for the SM and train alternative with respect to car, given that the traveler possesses a season ticket. Finally, the interpretation of the alternative specific constants is similar to that of the previous model specification.