Netherland Mode Choice Case

Model Specification with Generic Attributes

Files to use with BIOGEME:
Model file: *binary_generic_netherlands.mod*
Data file: *netherlands05.dat*

In this first model we assume that the total travel time (in-vehicle and out-of-vehicle) and travel cost of the modes are the only factors influencing the mode choice. We also assume that the coefficients of the explanatory variables are generic, i.e. they do not vary among alternatives. The expression of utility for this simple model can be expressed as:

$$\begin{aligned} V_{\text{auto}} &= \text{ASC}_{\text{auto}} + \beta_{\text{tt}} \text{tt}_{\text{auto}} + \beta_{\text{tc}} \text{cost}_{\text{auto}} \\ V_{\text{rail}} &= \beta_{\text{tt}} \text{tt}_{\text{rail}} + \beta_{\text{tc}} \text{cost}_{\text{rail}} \end{aligned}$$

Estimation results				
Variable number	Variable name	Coefficient estimate	Robust standard error	Robust *t statistic*
1	ASC_{auto}	-0.798	0.275	-2.895
2	β_{tc}	-0.050	0.010	-4.669
3	β_{tt}	-1.326	0.354	-3.745

Summary statistics
Number of observations = 228
$\mathcal{L}(0) = -158.038$
$\mathcal{L}(\hat{\beta}) = -123.133$
$\bar{\rho}^2 = 0.221$

Table 1: Estimation results with generic attributes

The estimation results are reported in Table 1. All the estimated coefficients are statistically significant different from zero. Looking at the alternative specific constant, the negative sign indicates that the rest of the utilities being equal, auto is less preferred than the rail. However, this may be due to the fact that the model is too simple and there are important variables left out from the model. The negative signs for the generic coefficients for cost and travel time indicates as expected, that the utility perceived by the decision maker for any of the two alternatives decreases with increase in cost and travel time.

Model Specification with Alternative Specific Attributes

Files to use with BIOGEME:
Model file: *binary_specific_netherlands.mod*
Data file: *netherlands05.dat*

In the second specification, we relax the hypothesis of generic travel time coefficients. The alternative specific coefficients are more relevant if people perceive a minute spent in one mode to be different than, a minute spent on the other. To illustrate this idea two different travel time coefficients are introduced for auto and rail. The corresponding utility function is given below:

$$\begin{aligned} V_{auto} &= ASC_{auto} + \beta_{tt\_auto} tt_{auto} + \beta_{tc} cost_{auto} \\ V_{rail} &= \beta_{tt\_rail} tt_{rail} + \beta_{tc} cost_{rail} \end{aligned}$$

The estimation results are reported in table 2. In general, this model has a better likelihood value than the model with generic travel time coefficients. However the coefficient for the travel time of the rail alternative is not statistically significant different from zero. The coefficient for the travel time of the car alternative is negative and significant as expected, and is also bigger in magnitude with respect to the generic one presented in the previous table (-0.022 vs. -0.037). As in the previous example the negative sign indicates that the utility perceived by the decision maker for

Estimation results				
Variable number	Variable name	Coefficient estimate	Robust standard error	Robust *t statistic*
1	ASC_{auto}	2.430	0.973	2.497
2	β_{tc}	-0.054	0.011	-4.785
3	β_{tt_auto}	-2.262	0.485	-4.662
4	β_{tt_rail}	-0.543	0.396	-1.372
Summary statistics				
Number of observations = 228				
$\mathcal{L}(0) = -158.03$				
$\mathcal{L}(\hat{\beta}) = -118.023$				
$\bar{\rho}^2 = 0.253$				

Table 2: Estimation results with alternative-specific attributes

the car alternative decreases with the increase of travel time. However it appears that travel time does not affect the car and rail alternatives in the same way. The results indicate that people have less negative utility for travel time in rail compared to car. This may be due to the fact that people can better utilize their time when travelling on rail. The alternative specific constant for the auto alternative has now the reversed sign denoting increased preference for auto (given equal total time and total cost) which is more intuitive. A likelihood ratio test can be performed to test whether or not there is a significant improvement in the goodness-of-fit in the modified specification with alternative specific coefficients for travel times (see case study).

Generic vs Specific Test

The likelihood ratio test can be used to test the generic vs the alternate-specific specification. The likelihood ratio test statistic for the null hypothesis of generic attributes is

$$-2(L(\beta_G) - L(\beta_{AS}))$$

where G and AS denote the generic and alternate-specific models, respectively. It is distributed χ^2 with the number of degrees of freedom equal to the number of restrictions $(K_{AS}-K_G)$. In this case, $-2(-123.133+118.02) = 10.22$. Since $\chi^2_{0.95,1} = 3.841$ at 95% level of confidence, we can conclude that the model with the alternate-specific constant has a significant improvement in fit.

Model Specification with Socio-Economic Variables

Files to use with BIOGEME:
Model file: *binary_socio_econ_netherlands.mod*
Data file: *netherlands05.dat*

The previous two models only included variables that were alternative specific. We now introduce a socioeconomic variable 'sex' which indicates the respondent gender. The variable is categorical and is equal to 1 if the gender is female and zero if male. Since the variable sex does vary on the alternative (recall that only difference in utility matters), we have normalized the alternative auto to zero. As it is showen in the utility function below, the gender variable only enters the utility of the rail alternative. However this is an arbitrary normalization, as we could also have normalized the rail alternative.

$$\begin{aligned} V_{auto} &= ASC_{auto} + \beta_{tt} tt_{auto} + \beta_{tc} cost_{auto} \\ V_{rail} &= \beta_{tt} tt_{rail} + \beta_{tc} cost_{rail} + \beta_{sex} sex \end{aligned}$$

The estimation results are reported in table 3. The results show that there is a slight improvement in the likelihood value. The coefficient of the gender variable is positive and statistically significant, which indicates that females have higher 'propensity' than males in choosing the rail alternative with respect to the auto alternative. The reader can verify that if we had included the gender variable in the utility of the auto alternative instead of the rail alternative, the conclusion would remain unchanged. In fact the results would be exactly the same. The only difference is that the coefficient would show the opposite sign. In our case it would become negative. The interpretation would be that females would have lower propensity (or utility) than males for using the car alternative with respect to the train alternative, which is exactly the same result we had before. Regarding the coefficients of the other explanatory variables they are almost unchanged with respect to the previous model (reported in table 4.8) and therefore we do not comment on them.

Estimation results				
Variable number	Variable name	Coefficient estimate	Robust standard error	Robust *t statistic*
1	ASC_{auto}	2.852	1.017	2.802
2	β_{sex}	0.675	0.329	2.050
3	β_{tc}	-0.06	0.012	-4.893
4	β_{ttauto}	-2.338	0.495	-4.726
5	β_{ttrail}	-0.529	0.414	-1.280

Summary statistics

Number of observations = 228

$\mathcal{L}(0) = -158.038$

$\mathcal{L}(\hat{\beta}) = -115.88$

$\bar{\rho}^2 = 0.267$

Table 3: Estimation results with socioeconomic characteristics