Mixtures of Logit and MEV Models

This case study deals with the specification of mixtures of Logit models. The objectives can be summarized as follows:

- Gaining an overview of the different formulations of mixtures of logit and becoming familiar with the concepts of flexible correlation structures and taste heterogeneity.
- Specification and estimation of alternative specific variance models.
- Specification and estimation of error component models.
- Specification and estimation of random coefficients models.
- Specification and estimation of mixtures of MEV models.

For this case study, the *Swissmetro* dataset is considered. Details on the dataset can be found in the Appendix, section A. You are free to use another dataset.

Your Tasks:

- Study the models and estimation results provided to you. (If you are using a different dataset, you can look at the provided models for understanding the implementation of mixture specifications.)
- Following your observations develop your own model.
  - Try different specifications and present at least 3 models that have different mixture notions in their specifications. The models you present should (in total, NOT each model) span at least 3 of the mixture notions among the listed ones (alternative specific variance, error component, random coefficient, and mixtures of MEV).
  - Include necessary discussions to explain why do you think your model is better (or worse) than the others given to you. Draw analogies between nesting structures and mixtures when applicable. (If you are not using the Swissmetro dataset, you can compare the models with your previous good specifications (logit, nested logit, cross-nested logit etc.).)

Note 1: *Mixture of Logit with Panel Data* is kept for your reference. However, you are not expected to provide estimation results with panel data.

Note 2: In this case study you do not need to come up with a comprehensive data analysis. This is to teach you how to specify mixture models, we will not look for detailed data analysis. Briefly explain the idea behind your own specifications.

1 Swissmetro Case

Alternative Specific Variance Model

Files to use with Biogeme:
Model file: *Mixture_SM_Heteroskedastic.py*
Data file: *swissmetro.dat*

In this first model specification, we assume that the ASC's are randomly distributed. We show below the utility expressions:

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

The alternative specific variance for the constants is coded in Biogeme in the following way:

```
# Random parameters
ASC_CAR_random = ASC_CAR_mean + ASC_CAR_std * bioDraws('ASC_CAR_random')
ASC_SM_random= ASC_SM_mean + ASC_SM_std * bioDraws('ASC_SM_random')
```

This model is very simple. The parameters are assumed to be generic over the alternatives, and just a few variables are taken into account. ASC_{car} and ASC_{SM} are now randomly distributed, with mean $\bar{\alpha}_{\text{car}}$ and $\bar{\alpha}_{\text{SM}}$ and standard deviation σ_{car} and σ_{SM}, which are both estimated. We normalize with respect to the train alternative, and the estimation results are shown in Table 1. Note that this is a simplification of the proper estimation process that is needed for alternative specific variance estimation. Recall that the normalization is not arbitrary in that only the minimum variance alternative can be normalized to 0. Therefore, proper estimation requires first that an unidentified model be estimated (with all three variances in this case). Then, the model should be re-estimated with the smallest variance from the unidentified model normalized to 0.

The estimated values of the time, cost and headway coefficients show their negative impact on the utility functions. Time and cost estimated coeffi-

Estimation results				
Parameter number	Parameter name	Parameter estimate	Robust standard error	Robust *t statistic*
1	$\bar{\alpha}_{car}$	-1.23	0.502	-2.46
2	$\bar{\alpha}_{SM}$	0.708	0.479	1.48
3	σ_{car}	-5.02	1.66	-3.03
4	σ_{SM}	4.67	0.919	5.08
5	β_{cost}	-0.0372	0.00695	-5.35
6	β_{he}	-0.0118	0.00291	-4.07
7	β_{time}	-0.0299	0.00618	-4.84

Summary statistics
Number of draws = 1000
Number of observations = 6768
$\mathcal{L}(0) = -6964.663$
$\mathcal{L}(\hat{\beta}) = -5215.226$
$\bar{\rho}^2 = 0.250$

Table 1: Alternative specific variance specification

cients are numerically very close, indicating the same negative impact, which is larger than that of headway. The estimated ASC's show that, all the rest remaining constant, Swissmetro and train are prefered compared to car. Between Swissmetro and train we can not compare, because $\bar{\alpha}_{SM}$ is not significantly different from zero. We can see that the standard deviations both of car and Swissmetro are large compared to the mean, which indicates a great variability. This suggests that it is a good idea to consider the ASC to be a random variable instead of a constant.

Only 1000 random draws have been used for the estimation. Note that this is not enough. We have chosen few draws in order to decrease the estimation time for the case study. For more theoretical details on this choice, we refer the reader to Train (2003) [1].

Error Component Model

Files to use with Biogeme:
Model files: *Mixture_SM_ErrorComp_01.py, Mixture_SM_ErrorComp_02.py*
Data file: *swissmetro.dat*

This first error component model attempts to capture the correlation between the train and Swissmetro alternatives. They are both rail-based transportation modes, so the hypothesis is that they share unobserved attributes. We show below the systematic utility expressions.

$$\begin{aligned}
V_{car} &= ASC_{car} + \beta_{time} CAR_TT + \beta_{cost} CAR_CO \\
V_{train} &= \beta_{time} TRAIN_TT + \beta_{cost} TRAIN_CO \\
&\quad + \beta_{he} TRAIN_HE + \zeta_{rail} \\
V_{SM} &= ASC_{SM} + \beta_{time} SM_TT + \beta_{cost} SM_CO \\
&\quad + \beta_{he} SM_HE + \zeta_{rail}
\end{aligned}$$

The corresponding code for Biogeme is as follows:

[1] The number of random draws is an important issue in simulated estimations. For reliable values, such a number should theoretically be ∞, as the Simulated Maximum Likelihood estimator is not consistent for a finite number of draws. In practical applications, the trade-off between the reliability of the estimates and a reasonable computational time becomes the most important issue. By default, Biogeme uses pseudo-random draws.

```
# Random parameters
RAIL_random = RAIL_mean + RAIL_std * bioDraws('RAIL_random')

# Utilities
V_Car_SP = ASC_CAR * one + BETA_TIME * CAR_TT + BETA_COST * CAR_CO

V_SBB_SP = ASC_SBB * one + BETA_TIME * TRAIN_TT + BETA_COST * TRAIN_COST +
+ BETA_HE * TRAIN_HE + RAIL_random * one

V_SM_SP = ASC_SM * one + BETA_TIME * SM_TT + BETA_COST * SM_COST +
+ BETA_HE * SM_HE + RAIL_random * one
```

The train and SM modes share the random term ζ_{rail}, which is assumed to be normally distributed $\zeta_{\text{rail}} \sim N(m_{\text{rail}}, \sigma^2_{\text{rail}})$. We estimate the standard deviation σ_{rail} of this error component, while the mean m_{rail} is fixed to zero. The estimation results are shown in Table 2. The interpretation is now a bit different: all else being equal, the Swissmetro mode is the most preferred, followed by the car. Both $\mathsf{ASC}_{\mathsf{CAR}}$ and $\mathsf{ASC}_{\mathsf{SM}}$ are now significantly different from zero. The parameters related to cost, time and headway are still negative, as expected. σ_{rail} has been estimated significantly different from zero, capturing the correlation between the train and the Swissmetro alternatives. This parameter is actually the element of the variance-covariance matrix capturing the correlation between Swissmetro and train.

In the following model, we use a more complex error structure. The idea is that train and SM are correlated, both being rail-based transportation modes, but also that train and car are correlated representing more classical transportation modes with respect to the more innovative Swissmetro. The corresponding utility functions are

$$
\begin{aligned}
V_{\text{car}} &= \mathsf{ASC}_{\text{car}} + \beta_{\text{time}}\mathsf{CAR\_TT} + \beta_{\text{cost}}\mathsf{CAR\_CO} + \zeta_{\text{classic}} \\
V_{\text{train}} &= \beta_{\text{time}}\mathsf{TRAIN\_TT} + \beta_{\text{cost}}\mathsf{TRAIN\_CO} + \beta_{\text{he}}\mathsf{TRAIN\_HE} \\
&\quad + \zeta_{\text{rail}} + \zeta_{\text{classic}} \\
V_{\mathsf{SM}} &= \mathsf{ASC}_{\mathsf{SM}} + \beta_{\text{time}}\mathsf{SM\_TT} + \beta_{\text{cost}}\mathsf{SM\_CO} \\
&\quad + \beta_{\text{he}}\mathsf{SM\_HE} + \zeta_{\text{rail}}
\end{aligned}
$$

Estimation results				
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
6	σ_{rail}	-0.00444	0.00157	-2.82

Summary statistics
Number of draws = 1000
Number of observations = 6768
$\mathcal{L}(0) = -6964.663$
$\mathcal{L}(\hat{\beta}) = -5315.386$
$\bar{\rho}^2 = 0.236$

Table 2: Error component specification. The σ_{rail} coefficient is the standard deviation of the random term capturing the unobserved shared attributes between the train and Swissmetro alternatives.

The corresponding code for Biogeme is as follows:

```
# Random parameters
CLASSIC_random = CLASSIC_mean + CLASSIC_std * bioDraws('CLASSIC_random')
RAIL_random = RAIL_mean + RAIL_std * bioDraws('RAIL_random')

# Utilities
V_Car_SP = ASC_CAR * one + BETA_TIME * CAR_TT + BETA_COST * CAR_CO +
+ CLASSIC_random * one

V_SBB_SP = ASC_SBB * one + BETA_TIME * TRAIN_TT + BETA_COST * TRAIN_COST +
+ BETA_HE * TRAIN_HE + RAIL_random * one + CLASSIC_random * one

V_SM_SP = ASC_SM * one + BETA_TIME * SM_TT + BETA_COST * SM_COST +
+ BETA_HE * SM_HE + RAIL_random * one
```

As before, the random terms are assumed to be normally distributed $\zeta_{rail} \sim N(m_{rail}, \sigma^2_{rail})$ and $\zeta_{classic} \sim N(m_{classic}, \sigma^2_{classic})$. The standard deviations,

Estimation results				
Parameter number	Parameter name	Parameter estimate	Robust standard error	Robust *t statistic*
1	ASC_{car}	-1.12	0.462	-2.43
2	ASC_{SM}	0.665	0.420	1.59
3	β_{cost}	-0.0356	0.00573	-6.21
4	β_{he}	-0.0114	0.00256	-4.44
5	β_{time}	-0.0290	0.00576	-5.04
6	$\sigma_{classic}$	4.48	0.794	5.64
7	σ_{rail}	-4.66	1.47	-3.177

Summary statistics
Number of draws = 1000
Number of observations = 6768
$\mathcal{L}(0) = -6964.663$
$\mathcal{L}(\hat{\beta}) = -5216.971$
$\bar{\rho}^2 = 0.250$

Table 3: Error component specification. Train and car share unobserved attributes through $\zeta_{classic}$ and train and SM through ζ_{rail}.

σ_{rail} and $\sigma_{classic}$, are estimated, while the means m_{rail} and $m_{classic}$ are fixed to zero.

A similar correlation pattern could be specified by means of a Cross-Nested Logit model where the SM alternative belongs to a *rail* nest, the car alternative belongs to a *classic* nest and the train alternative is assigned with certain degrees of membership to both rail and classic nests. In the model, we have normalized with respect to the train alternative. The estimation results are shown in Table 3.

ASC_{CAR} is negative, indicating a preference towards train over car, all the rest being constant. ASC_{SM} is not significantly different from zero. The interpretation of the cost, time and headway coefficients remains the same. Both standard deviations are significantly different from zero [2]

[2]The signs of the estimated standard deviations are always reported as positive. In Biogeme they may be reported as negative. If so, just ignore the sign and consider the absolute value.

Random Coefficients

Files to use with Biogeme:
Model file: *Mixture_SM_RandomCoeff.py*
Data file: *swissmetro.dat*

In this specification, the unknown parameters are assumed to be randomly distributed over the population. They capture the so called *taste variation* of individuals. The utility expressions and the related Biogeme code are shown below.

$$
\begin{aligned}
V_{car} &= ASC_{car} + \beta_{time}CAR\_TT + \beta_{car\_cost}CAR\_CO \\
V_{train} &= \beta_{time}TRAIN\_TT + \beta_{train\_cost}TRAIN\_CO + \beta_{he}TRAIN\_HE \\
V_{SM} &= ASC_{SM} + \beta_{time}SM\_TT + \beta_{SM\_cost}SM\_CO + \beta_{he}SM\_HE
\end{aligned}
$$

```
# Random parameters
BETA_CAR_COST_random = BETA_CAR_COST_mean + BETA_CAR_COST_std * bioDraws('BETA_CAR_COST_random')
BETA_HE_random = BETA_HE_mean + BETA_HE_std * bioDraws('BETA_HE_random')
BETA_SM_COST_random = BETA_SM_COST_mean + BETA_SM_COST_std * bioDraws('BETA_SM_COST_random')
BETA_TRAIN_COST_random = BETA_TRAIN_COST_mean +
 + BETA_TRAIN_COST_std * bioDraws('BETA_TRAIN_COST_random')

# Utilities
V_Car_SP = ASC_CAR * one + BETA_TIME * CAR_TT + BETA_CAR_COST_random * CAR_CO

V_SBB_SP = ASC_SBB * one + BETA_TIME * TRAIN_TT + BETA_TRAIN_COST_random * TRAIN_COST +
+ BETA_HE_random * TRAIN_HE

V_SM_SP = ASC_SM * one + BETA_TIME * SM_TT + BETA_SM_COST_random * SM_COST +
+ BETA_HE_random * SM_HE
```

We have three alternative-specific coefficients for the cost variable which are normally distributed with means m_{car_cost}, m_{train_cost}, and m_{SM_cost} and standard deviations σ_{car_cost}, σ_{train_cost}, and σ_{SM_cost}, respectively. The coefficient related to headway is also assumed to be randomly distributed over the population, with mean m_{he} and standard deviation σ_{he}.

The estimation results are shown in Table 4. The ASC's have negative signs, and their values show a preference, all the rest remaining constant, for the train with respect to both the car and Swissmetro alternatives. The mean for the car cost coefficient is negative, as expected, and the standard deviation

Estimation results				
Parameter number	Parameter name	Parameter estimate	Robust standard error	Robust *t statistic*
1	$\mathsf{ASC}_{\mathsf{car}}$	-1.78	0.262	-6.82
2	$\mathsf{ASC}_{\mathsf{SM}}$	-1.11	0.193	-5.74
3	$m_{\mathsf{car\_cost}}$	-0.0185	0.00375	-4.94
4	$\sigma_{\mathsf{car\_cost}}$	-0.00963	0.00326	-2.96
5	$m_{\mathsf{train\_cost}}$	-0.0691	0.00790	-8.74
6	$\sigma_{\mathsf{train\_cost}}$	0.0263	0.00360	7.31
7	$m_{\mathsf{SM\_cost}}$	-0.0185	0.00273	-6.77
8	$\sigma_{\mathsf{SM\_cost}}$	0.0114	0.00281	4.05
9	m_{he}	-0.00912	0.00221	-4.12
10	σ_{he}	0.0122	0.00427	2.86
11	β_{time}	-0.0144	0.00239	-6.04

Summary statistics
Number of draws = 1000
Number of observations = 6768
$\mathcal{L}(0) = -6964.663$
$\mathcal{L}(\hat{\beta}) = -4972.915$
$\bar{\rho}^2 = 0.284$

Table 4: Random coefficient specification assuming normal distributions.

```
B_TIME_RND = -exp(B_TIME + B_TIME_S * bioDraws('B_TIME_RND'))
```

Figure 1: The Biogeme Log-Normal specification.

σ_{car_cost} is significantly different from zero. Its numerical value indicates that the probability that the parameter has a negative value is 1. The assumed Normal distribution allows for non-zero probabilities of having a positive car cost coefficient. Similar considerations can be made for the other random coefficients. The mean for the train cost coefficient is negative, as expected, and both the mean and the standard deviation are significant. Computing the cumulative distribution function (cdf) for the Normal distribution with these parameters, we observe that the cumulative probability of having a train cost coefficient less than zero is also 1. For the SM cost parameter (both mean and standard deviation are significant), we have the cdf for negative values equal to 1 also. The mean of the headway parameter is negative as expected, and its standard deviation has been estimated significantly different from zero.

Different distributions We show here two examples of Biogeme code to specify a random coefficient model where the parameters are log-normally and Johnson SB distributed. The Biogeme code is shown in Figures 1 and 2, respectively. Recall that a variable X is log-normally distributed if $y = \ln(X)$ is normally distributed. We can easily define in Biogeme such a distribution by assuming a generic time coefficient to be log-normally distributed.

In the case of Johnson SB distribution, the functional form is derived using a Logit-like transformation of a Normal distribution, as defined in the following equation

$$\xi = a + (b - a)\frac{e^{\zeta}}{e^{\zeta} + 1}$$

where $\zeta \sim N(\mu, \sigma^2)$. This distribution is very flexible; it is bounded between a and b and its shape can change from a very flat one to a bimodal, changing the parameters of the normal variable. It requires the estimation of four parameters (a, b, μ and σ) and a nonlinear specification, assuming as before, a generic time coefficient following such a distribution.

The topic of the functional form for random coefficient distributions is treated in more detail in, for example, Train (2003) and Walker et al. (2007).

```
B_TIME_RND = ( A + ( ( B - A ) * ( exp (B_TIME + B_TIME_S * bioDraws('B_TIME_RND'))
/ ( exp( B_TIME + B_TIME_S * bioDraws('B_TIME_RND') ) + 1 ) ) ) ) * TRAIN_TT
```

Figure 2: The Biogeme SB specification.

Mixture of MEV Models

Files to use with Biogeme:
Model file: *Mixture_SM_MEV.py*
Data file: *swissmetro.dat*

In this example, we capture the substitution patterns using a Nested Logit model, and we allow for some parameters to be randomly distributed over the population.

$$
\begin{aligned}
V_{car} &= ASC_{car} + \beta_{car\_time} CAR\_TT + \beta_{cost} CAR\_CO \\
V_{train} &= \beta_{train\_time} TRAIN\_TT + \beta_{cost} TRAIN\_CO + \beta_{he} TRAIN\_HE \\
&\quad + \beta_{ga} GA + \beta_{senior} SENIOR \\
V_{SM} &= ASC_{SM} + \beta_{SM\_time} SM\_TT + \beta_{cost} SM\_CO + \beta_{he} SM\_HE \\
&\quad + \beta_{ga} GA + \beta_{seats} SM\_SEATS
\end{aligned}
$$

We have added the socio-economic characteristics *senior* (a dummy variable for senior people, i.e. age above 65), *ga* and *SM_seats*. A few observations have been removed where the variable *Age* was missing. We specify a nest composed of alternatives *car* and *train* representing standard transportation modes, while the Swissmetro alternative represents the technological innovation. We further assume a generic cost parameter and three randomly distributed alternative-specific time parameters. Normal distributions are used for the random coefficients, that is,

$$
\begin{aligned}
\beta_{car\_time} &\sim N(m_{car\_time}, \sigma^2_{car\_time}) \\
\beta_{train\_time} &\sim N(m_{train\_time}, \sigma^2_{train\_time}) \\
\beta_{SM\_time} &\sim N(m_{SM\_time}, \sigma^2_{SM\_time}).
\end{aligned}
$$

Estimation results					
Parameter number	Parameter name	Parameter estimate	Robust standard error	Robust *t stat. 0*	Robust *t stat. 1*
1	$\mathrm{ASC}_{\mathrm{car}}$	-0.145	0.122	-1.18	
2	$\mathrm{ASC}_{\mathrm{SM}}$	0.195	0.114	1.71	
3	β_{senior}	1.54	0.132	11.64	
4	$m_{\mathrm{car\_time}}$	-0.0137	0.00102	-13.42	
5	β_{cost}	-0.00973	0.000843	-11.54	
6	$\sigma_{\mathrm{car\_time}}$	0.00443	0.000948	4.67	
7	β_{ga}	1.03	0.152	6.73	
8	β_{he}	-0.00470	0.000888	-5.29	
9	β_{seats}	-0.256	0.102	-2.51	
10	$m_{\mathrm{SM\_time}}$	-0.0159	0.00136	-11.69	
11	$\sigma_{\mathrm{SM\_time}}$	-0.00952	0.00147	-6.48	
12	$m_{\mathrm{train\_time}}$	-0.0162	0.00112	-14.41	
13	$\sigma_{\mathrm{train\_time}}$	-0.000122	0.000147	-0.83	
14	μ_{classic}	1.81	0	144	5.62

Summary statistics

Number of draws = 1000

Number of observations = 6759

$\mathcal{L}(0) = -6958.425$

$\mathcal{L}(\hat{\beta}) = -4954.314$

$\bar{\rho}^2 = 0.286$

Table 5: Mixture of Nested Logit estimation results

The estimation results are shown in Table 5. The nest parameter has been estimated significantly different from 1, showing a correlation between the train and car alternatives, as expected. The three mean parameters for the time coefficients have been estimated with negative signs (as expected) and are significantly different from zero. Their numerical values are only slightly different, suggesting that probably a generic specification would have been acceptable. For the car and Swissmetro time coefficients, the estimated standard deviations are significant and one magnitude order less than the mean value. It means that their distribution over the population is very peaked, indicating that the way different individuals perceive the negative impact of travel time on the alternatives' utilities is not so different.

Mixture of Logit with Panel Data

Files to use with Biogeme:
Model file: *Mixture_SM_panel.py*
Data file: *swissmetro.dat*

In this example, we take into account the fact that we have panel data in the sample file. Indeed, the sample file is composed of nine observations per individual. These nine observations correspond to the choices made by a single respondent in nine hypothetical mode choice situations described in the questionnaire of the Swissmetro survey. The idea is thus to specify a model which is able to deal with sequences of observed choices and the intrinsic correlation among the choices of a sequence.

The specification file *Mixture_SM_panel.py* is based on the model *MNL_SM_-specific.py* with alternative-specific cost coefficients which has been analyzed in the Case Study dealing with logit models.

In the dataset, each individual is characterized by an `ID`. This is the name of the variable in the data file identifying the observations belonging to a given individual. We define `ZERO_SIGMA_PANEL` to be the random coefficient which will not vary across observations from the same individual.

The way we deal with panel data is therefore to use a Mixture of Logit model with random coefficients specification. More precisely, we add individual specific error terms (specified in Biogeme by `ZERO_SIGMA_PANEL * one` where `ZERO_SIGMA_PANEL` equals

Estimation results				
Parameter number	Parameter name	Parameter estimate	Robust standard error	Robust *t statistic*
1	$\mathsf{ASC}_{\mathsf{car}}$	-0.988	0.390	-2.53
2	$\mathsf{ASC}_{\mathsf{SM}}$	-0.291	0.531	-0.55
3	$\beta_{\mathsf{car\_cost}}$	-0.0132	0.00324	-4.07
4	$\beta_{\mathsf{train\_cost}}$	-0.0323	0.00574	-5.63
5	$\beta_{\mathsf{SM\_cost}}$	-0.0163	0.00262	-6.22
6	β_{he}	-0.00757	0.00127	-5.96
7	β_{time}	-0.0190	0.00616	-3.09
8	σ_{panel}	2.39	0.216	11.06

Summary statistics
Number of draws = 100
Number of individuals = 752
$\mathcal{L}(0) = -6964.663$
$\mathcal{L}(\hat{\beta}) = -4235.440$
$\bar{\rho}^2 = 0.391$

Table 6: Mixture of logit model with panel data.

`ZERO + SIGMA_PANEL * bioDraws('ZERO_SIGMA_PANEL')`) in two alternatives (we need to normalize one alternative), where the standard deviation (`SIGMA_PANEL`) needs to be estimated while the mean (`ZERO`) is fixed to zero. The utility functions for this model can therefore be specified in Biogeme as follows:

```
# Utilities
V_Car_SP = ASC_CAR * one + BETA_TIME * CAR_TT + BETA_CAR_COST *
* CAR_CO + ZERO_SIGMA_PANEL * one

V_SBB_SP = ASC_SBB * one + BETA_TIME * TRAIN_TT + BETA_TRAIN_COST *
* TRAIN_COST + BETA_HE * TRAIN_HE + ZERO_SIGMA_PANEL * one

V_SM_SP = ASC_SM * one + BETA_TIME * SM_TT + BETA_SM_COST *
* SM_COST + BETA_HE * SM_HE
```

We see from the estimation results presented in Table 6 that the coefficient σ_{panel} is highly significant, which means that this model allows for capturing intrinsic correlations among the observations of the same individual. Moreover, the final log-likelihood value is -4235.440, which is much smaller (in absolute value) than the value -5068.560 obtained with the model *MNL_SM_specific.py* without a panel term. The interpretation of other coefficients remains the same as that for the coefficients of *MNL_SM_specific.py*, except that $\mathsf{ASC}_{\mathsf{SM}}$ is no longer significantly different from 0.

Note that for this example we have used only 100 draws which is not enough, it is only for illustration purposes. For more theoretical details on this choice, we refer the reader to Train (2003).

References

Bierlaire, M., Axhausen, K. and Abay, G. (2001). The acceptance of modal innovation: The case of swissmetro, *Proceedings of the 1st Swiss Transportation Research Conference*, Ascona, Switzerland. www.strc.ch/bierlaire.pdf.

Train, K. (2003). *Discrete Choice Methods with Simulation*, Cambridge University Press, University of California, Berkeley.

Walker, J., Ben-Akiva, M. and Bolduc, D. (2007). Identification of parameters in normal error component logit-mixture (neclm) models, *Journal of Applied Econometrics* **22**(6): 1095–1125.

A Swissmetro Data Set

This dataset consists of survey data collected on the trains between St. Gallen and Geneva, Switzerland, during March 1998. The respondents provided information in order to analyze the impact of the modal innovation in transportation, represented by the Swissmetro, a revolutionary mag-lev underground system, against the usual transport modes represented by car and train.

Context

Innovation in the market for intercity passenger transportation is a difficult enterprise as the existing modes: private car, coach, rail as well as regional and long-distance air services continue to innovate in their own right by offering new combinations of speeds, services, prices and technologies. Consider for example high-speed rail links between the major centers or direct regional jet services between smaller countries. The Swissmetro SA in Geneva is promoting such an innovation: a mag-lev underground system operating at speeds up to 500 km/h in partial vacuum connecting the major Swiss conurbations, in particular along the Mittelland corridor (St. Gallen, Zurich, Bern, Lausanne and Geneva).

Data Collection

The Swissmetro is a true innovation. It is therefore not appropriate to base forecasts of its impact on observations of existing revealed preferences (RP) data. It is necessary to obtain data from surveys of hypothetical markets/situations, which include the innovation, to assess the impact. Survey data were collected on rail-based travels, interviewing 470 respondents. Due to data problems, only 441 are used here. Nine stated choice situations were generated for each of 441 respondents, offering three alternatives: rail, Swissmetro and car (only for car owners).

A similar method for relevant car trips with a household or telephone survey was deemed impractical. The sample was therefore constructed using license plate observations on the motorways in the corridor by means of video recorders. A total of 10529 relevant license plates were recorded during September 1997. The central Swiss car license agency had agreed to send up to 10000 owners of these cars a survey-pack. Until April 1998, 9658 letters were mailed, of which 1758 were returned. A total of 1070 persons filled in the survey completely and were willing to participate in the second SP survey, which was generated using the same approach used for the rail interviews. 750 usable SP surveys were returned, from the license-plate based survey.

Variables and Descriptive Statistics

The variables of the dataset are described in Tables 7 and 8, and the descriptive statistics are summarized in Table 9. A more detailed description of the data set as well as the data collection procedure is given in Bierlaire et al. (2001).

Variable	Description
GROUP	Different groups in the population
SURVEY	Survey performed in train (0) or car (1)
SP	It is fixed to 1 (stated preference survey)
ID	Respondent identifier
PURPOSE	Travel purpose. 1: Commuter, 2: Shopping, 3: Business, 4: Leisure, 5: Return from work, 6: Return from shopping, 7: Return from business, 8: Return from leisure, 9: other
FIRST	First class traveler (0 = no, 1 = yes)
TICKET	Travel ticket. 0: None, 1: Two way with half price card, 2: One way with half price card, 3: Two way normal price, 4: One way normal price, 5: Half day, 6: Annual season ticket, 7: Annual season ticket Junior or Senior, 8: Free travel after 7pm card, 9: Group ticket, 10: Other
WHO	Who pays (0: unknown, 1: self, 2: employer, 3: half-half)
LUGGAGE	0: none, 1: one piece, 3: several pieces
AGE	It captures the age class of individuals. The age-class coding scheme is of the type: 1: age$\leq$24, 2: 24$<$age$\leq$39, 3: 39$<$age$\leq$54, 4: 54$<$age$\leq$ 65, 5: 65 $<$age, 6: not known
MALE	Traveler's Gender 0: female, 1: male
INCOME	Traveler's income per year [thousand CHF] 0 or 1: under 50, 2: between 50 and 100, 3: over 100, 4: unknown
GA	Variable capturing the effect of the Swiss annual season ticket for the rail system and most local public transport. It is 1 if the individual owns a GA, zero otherwise.
ORIGIN	Travel origin (a number corresponding to a Canton, see Table 10)

Table 7: Description of variables

Variable	Description
DEST	Travel destination (a number corresponding to a Canton, see Table 10)
TRAIN_AV	Train availability dummy
CAR_AV	Car availability dummy
SM_AV	SM availability dummy
TRAIN_TT	Train travel time [minutes]. Travel times are door-to-door making assumptions about car-based distances (1.25*crow-flight distance)
TRAIN_CO	Train cost [CHF]. If the traveler has a GA, this cost equals the cost of the annual ticket.
TRAIN_HE	Train headway [minutes] Example: If there are two trains per hour, the value of TRAIN_HE is 30.
SM_TT	SM travel time [minutes] considering the future Swissmetro speed of 500 km/h
SM_CO	SM cost [CHF] calculated at the current relevant rail fare, without considering GA, multiplied by a fixed factor (1.2) to reflect the higher speed.
SM_HE	SM headway [minutes] Example: If there are two Swissmetros per hour, the value of SM_HE is 30.
SM_SEATS	Seats configuration in the Swissmetro (dummy). Airline seats (1) or not (0).
CAR_TT	Car travel time [minutes]
CAR_CO	Car cost [CHF] considering a fixed average cost per kilometer (1.20 CHF/km)
CHOICE	Choice indicator. 0: unknown, 1: Train, 2: SM, 3: Car

Table 8: Description of variables

Variable	**Min**	**Max**	**Mean**	**St. Dev.**
GROUP	2	3	2.63	0.48
SURVEY	0	1	0.63	0.48
SP	1	1	1.00	0.00
ID	1	1192	596.50	344.12
PURPOSE	1	9	2.91	1.15
FIRST	0	1	0.47	0.50
TICKET	1	10	2.89	2.19
WHO	0	3	1.49	0.71
LUGGAGE	0	3	0.68	0.60
AGE	1	6	2.90	1.03
MALE	0	1	0.75	0.43
INCOME	0	4	2.33	0.94
GA	0	1	0.14	0.35
ORIGIN	1	25	13.32	10.14
DEST	1	26	10.80	9.75
TRAIN_AV	1	1	1.00	0.00
CAR_AV	0	1	0.84	0.36
SM_AV	1	1	1.00	0.00
TRAIN_TT	31	1049	166.63	77.35
TRAIN_CO	4	5040	514.34	1088.93
TRAIN_HE	30	120	70.10	37.43
SM_TT	8	796	87.47	53.55
SM_CO	6	6720	670.34	1441.59
SM_HE	10	30	20.02	8.16
SM_SEATS	0	1	0.12	0.32
CAR_TT	0	1560	123.80	88.71
CAR_CO	0	520	78.74	55.26
CHOICE	1	3	2.15	0.63

Table 9: Descriptive statistics

Number	**Canton**
1	ZH
2	BE
3	LU
4	UR
5	SZ
6	OW
7	NW
8	GL
9	ZG
10	FR
11	SO
12	BS
13	BL
14	Schaffhausen
15	AR
16	AI
17	SG
18	GR
19	AG
20	TH
21	TI
22	VD
23	VS
24	NE
25	GE
26	JU

Table 10: Coding of Cantons