Demand Modeling: 1.202
Spring 2019
Alexandra Berke

Case Study 3 - Report

This report is about the Swissmetro case.

Part 1: Model Development

Best model specification
See *Specification 3* and *model and output files: MNL_SM_traveltime.**

The best model specification from the tests of Part 1 has utility function with expressions defined below.

```
V_car = ASC_car + B_car_time*CAR_TT + B_car_cost*CAR_CO + B_senior*SENIOR

V_train = B_train_time*TRAIN_TT + B_train_cost*TRAIN_COST + B_he*TRAIN_HE +
B_ga*GA + B_firstclass*FIRST

V_sm = ASC_sm + B_sm_time*SM_TT + B_sm_cost*SM_COST + B_he*SM_HE +
B_senior*SENIOR + B_ga*GA + B_firstclass*FIRST
```

The log likelihood value for this specification is -4913.23.
See Table 1 for the estimations of coefficients.

The provided model specification with socio-economic characteristics (MNL_SM_socioec.py) was plausible, with estimated variables supporting intuitive hypotheses. To arrive at this report's "best" model specification, I tested a variety of models that built on top of this provided model specification (MNL_SM_socioec.py).

Specification 1.
*Model and output files: MNL_SM_firstclass.**

I add one new variable to the previous (MNL_SM_socioec.py) model specification to capture the effect of first class travelers. This is a dummy variable (FIRST). This additional variable is motivated by the hypothesis that first class travelers are more likely to prefer the Swissmetro and train alternatives over car. On their train and (potential) Swiss Metro trips, they are better able to enjoy their time traveling in luxury than those who are not first class travelers.

The utility expressions are defined as below.

```
V_car = ASC_car + B_time*CAR_TT + B_car_cost*CAR_CO + B_senior*SENIOR

V_train = B_time*TRAIN_TT + B_train_cost*TRAIN_COST + B_he*TRAIN_HE + B_ga*GA +
B_firstclass*FIRST

V_sm = ASC_sm + B_time*SM_TT + B_sm_cost*SM_COST + B_he*SM_HE + B_senior*SENIOR
+ B_ga*GA + B_firstclass*FIRST
```

The estimation results for this model are shown in the table below.

Name	Value	Std err	t-test	p-value	Rob. Std err	Rob. t-test	Rob. p-value
ASC_CAR	-0.56	0.122	-4.59	4.54e-06	0.142	-3.94	8.05e-05
ASC_SM	-0.145	0.0845	-1.72	0.0857	0.107	-1.35	0.175
B_CAR_COST	-0.00921	0.00091	-10.1	0	0.00118	-7.83	4.88e-15
B_FIRST	0.249	0.0731	3.4	0.000663	0.0778	3.2	0.00139
B_GA	0.529	0.188	2.81	0.00492	0.192	2.75	0.00591
B_HE	-0.00587	0.00105	-5.6	2.18e-08	0.00106	-5.55	2.82e-08
B_SENIOR	-1.88	0.116	-16.2	0	0.109	-17.3	0
B_SM_COST	-0.0112	0.000597	-18.7	0	0.000844	-13.2	0
B_TIME	-0.0112	0.000615	-18.3	0	0.00122	-9.21	0
B_TRAIN_COST	-0.0278	0.0012	-23.1	0	0.00185	-15	0

Table 1

Init log likelihood:	-6958.425
Final log likelihood:	-4921.365
Likelihood ratio test for the init. model:	4074.12
Rho-square for the init. model:	0.293
Rho-square-bar for the init. model:	0.291

The parameter for first class travelers (B_FIRST) is significantly different from zero at a 95% confidence level. The positive value of this parameter was expected. It supports the hypothesis that first class travelers prefer the train and Swissmetro over the car travel alternative because of the luxurious experience they can provide to those traveling first class.

Note that the log likelihood is slightly better (higher) than for the previous model specification that this one was built on (MNL_SM_socioec.py) and and rho-square-bar statistic did not change. The other parameters barely changed. Keeping the first class traveler variable (FIRST) in the model specification therefore seems like a safely preferable model specification from which my subsequent models will iterate from.

Specification 2.

*Model and output files: MNL_SM_age.**

This specification builds off of the observation (in the specification of MNL_SM_socioec.py) that the age characteristic affects the preference of travel alternative. That previous model supported the hypothesis that senior travelers prefer the train, likely due to the safety it provides, and that older travelers may be less likely to try an innovative new form of travel like the Swissmetro. My model specification below takes that notion further, by replacing the dummy variable for seniors (SENIOR) with the discrete ordinal variable for age buckets that is provided in the dataset. My hypothesis is that younger travelers are more likely to prefer the Swissmetro than older travelers. The specification below includes the age variable (AGE) as generic, where the observations where age is unknown are again removed. I include the variable in the utility expressions for Vcar and Vtrain, with an expectation that the coefficient will be positive, to reflect a relative preference for the car and train alternatives over Swissmetro as age increases.

```
V_car = ASC_car + B_time*CAR_TT + B_car_cost*CAR_CO + B_age*AGE

V_train = B_time*TRAIN_TT + B_train_cost*TRAIN_COST + B_he*TRAIN_HE + B_ga*GA +
B_firstclass*FIRST + B_age*AGE

V_sm = ASC_sm + B_time*SM_TT + B_sm_cost*SM_COST + B_he*SM_HE + B_ga*GA +
B_firstclass*FIRST
```

The estimation results for this model are shown in the table below.

Name	Value	Std err	t-test	p-value	Rob. Std err	Rob. t-test	Rob. p-value
ASC_CAR	-0.824	0.119	-6.91	4.77e-12	0.139	-5.95	2.74e-09
ASC_SM	0.133	0.112	1.19	0.234	0.127	1.05	0.294
B_AGE	0.199	0.0278	7.17	7.5e-13	0.0299	6.65	2.9e-11
B_CAR_COST	-0.0094	0.000911	-10.3	0	0.00117	-8	1.11e-15
B_FIRST	0.31	0.0735	4.22	2.41e-05	0.078	3.98	7.01e-05
B_GA	0.487	0.188	2.59	0.00959	0.192	2.54	0.011
B_HE	-0.0055	0.00101	-5.44	5.33e-08	0.00101	-5.42	5.99e-08
B_SM_COST	-0.0113	0.000598	-18.9	0	0.000843	-13.5	0
B_TIME	-0.0113	0.000614	-18.3	0	0.00122	-9.22	0
B_TRAIN_COST	-0.0297	0.0012	-24.7	0	0.00185	-16	0

Table 2.

Init log likelihood:	-6958.425
Final log likelihood:	-5023.722
Likelihood ratio test for the init. model:	3869.406
Rho-square for the init. model:	0.278
Rho-square-bar for the init. model:	0.277

This was not a successful change to the model specification. While the coefficient for age was positive (supporting the hypothesis) and is statistically significant, it has a lower robust t-statistic than the coefficient for the senior dummy variable (SENIOR), which it replaced. Moreover, the log likelihood value is lower.
In the subsequent model specification, I return to the use of the senior dummy variable (SENIOR), without the AGE variable.

Specification 3.
*Model and output files: MNL_SM_traveltime.**

This model specification tests the hypothesis that travel time is alternative specific.
The hypothesis is that longer travel times effect utility more for the car alternative than the train and Swissmetro alternatives because time spent on the train and Swissmetro does not need to feel like time lost - when travelers are not spending time driving, they can instead read or pass the time doing things they enjoy.

The expressions for the model utilities are below, defined with the expectation that the parameter for car travel time (B_car_time) will be more negative than the parameters for train and Swissmetro travel times (B_train_time, B_sm, respectively).
The rest of the model specification builds off of specification 1 (MNL_SM_firstclass).

```
V_car = ASC_car + B_car_time*CAR_TT + B_car_cost*CAR_CO + B_senior*SENIOR

V_train = B_train_time*TRAIN_TT + B_train_cost*TRAIN_COST + B_he*TRAIN_HE +
B_ga*GA + B_firstclass*FIRST

V_sm = ASC_sm + B_sm_time*SM_TT + B_sm_cost*SM_COST + B_he*SM_HE +
B_senior*SENIOR + B_ga*GA + B_firstclass*FIRST
```

The estimation results for this model are shown in the table below.

Name	Value	Std err	t-test	p-value	Rob. Std err	Rob. t-test	Rob. p-value
ASC_CAR	-0.216	0.149	-1.45	0.148	0.161	-1.34	0.18
ASC_SM	0.123	0.119	1.03	0.301	0.12	1.02	0.306
B_CAR_COST	-0.00751	0.00107	-7.01	2.39e-12	0.00151	-4.99	6.11e-07
B_CAR_TIME	-0.0131	0.000812	-16.1	0	0.00164	-7.95	1.78e-15
B_FIRST	0.262	0.0735	3.56	0.000375	0.0789	3.31	0.000925
B_GA	0.518	0.188	2.75	0.00588	0.194	2.67	0.00764
B_HE	-0.00589	0.00105	-5.63	1.83e-08	0.00105	-5.61	2.01e-08
B_SENIOR	-1.86	0.116	-16.1	0	0.107	-17.4	0
B_SM_COST	-0.0116	0.000617	-18.8	0	0.000931	-12.5	0
B_SM_TIME	-0.011	0.000877	-12.5	0	0.00181	-6.06	1.36e-09
B_TRAIN_COST	-0.0293	0.00128	-23	0	0.0021	-14	0
B_TRAIN_TIME	-0.00881	0.000855	-10.3	0	0.0012	-7.35	2.03e-13

Table 3.

Init log likelihood:	-6958.425
Final log likelihood:	-4913.23
Likelihood ratio test for the init. model:	4090.389
Rho-square for the init. model:	0.294
Rho-square-bar for the init. model:	0.292

The estimated coefficients for the alternative specific travel time variables differ from my expectation that the coefficient for car travel time would be more negative than the others. However, these parameters are statistically significant with a 95% confidence interval. To evaluate whether these parameters should be alternative specific instead of generic, I use the likelihood ratio test.

The null hypothesis is for the generic case, where the parameters for travel time are equivalent.
H0: B_car_time = B_train_time = B_sm_time

The test statistic for the null hypothesis is given by −2(L_r − L_u),
Where L_r is the log likelihood for the restricted model (generic parameters) and L_u is the log likelihood for the unrestricted model (alternative specific parameters).
The critical chi-squared value for the alpha = 0.05 significance level, and 2 degrees of freedom is 5.991

-2(L_r - L_u) = -2(-4921.365 + 4913.23) = 16.27 > 5.991

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

Due to the results of this test, an the fact that the log likelihood of the model specification (-4913.23) is the best value yet attained, I will continue to use this model specification in Part 2.

Part 2: Forecasting

The model used for forecasting was developed in Part 1, as *Specification 3. (see MNL_SM_traveltime.*)* with utility expressions defined as follows:

```
V_car = ASC_car + B_car_time*CAR_TT + B_car_cost*CAR_CO + B_senior*SENIOR

V_train = B_train_time*TRAIN_TT + B_train_cost*TRAIN_COST + B_he*TRAIN_HE +
B_ga*GA + B_firstclass*FIRST

V_sm = ASC_sm + B_sm_time*SM_TT + B_sm_cost*SM_COST + B_he*SM_HE +
B_senior*SENIOR + B_ga*GA + B_firstclass*FIRST
```

Swiss Metro Price Changes by Age

Files:
*MNL_SM_price_base_other.**
*MNL_SM_price_base_youth.**

*MNL_SM_price_base_senior.**
*MNL_SM_price_forecast_other.**
*MNL_SM_price_forecast_youth.**
*MNL_SM_price_forcast_senior.**

In this section I investigate the impact of the proposed policy scenario:
The Swissmetro SA has decided to provide a 20% discount to youths (age < 24) and 50% discount to elderly (age > 65) when using Swissmetro. To compensate for the lost revenue, the company considers increasing the general Swissmetro fare uniformly by 10%.

This investigation looks at the effect on following market segments:

- Youth (who enjoy 20% discount)
- Seniors (who enjoy 50% discount)
- Other people who are not youth or seniors (24 <= age <= 65)

It also investigates the effect on the overall population, and overall revenue, by using a weighted average from the market segment estimates.

Since the Swissmetro has already decided to provide the youth and senior discounts, the base case is such that these discounts are already provided.
The case to forecast is the effect of the 10% uniform fare increase.

Results:

	Base case				**Forecast**			
	Youth	**Senior**	**Other**	**Weighted Avg**	**Youth**	**Senior**	**Other**	**Weighted Avg**
CAR	7	8	29	26	7	8	31	28
TRAIN	21	48	10	13	22	49	10	13
SM	72	44	61	61	71	43	59	59

Table 4: Market shares (percent) for Swissmetro price changes.

The estimated results are consistent with intuitive expectations. The proposed Swissmetro fare increase is forecasted to cause an average decrease in the use of Swissmetro across all market segments. The fare increase has a smaller effect on youth and seniors, which makes sense because they would still enjoy their discounts.

The forecasted estimates are useful to evaluate whether the proposed 10% uniform fare increase would help compensate for otherwise lost revenue. Policy makers might be concerned that the fare increase would be so great as to deter riders and have an overall negative effect on revenue. However, this model's results support the 10% fare increase as a means to increase revenue.

From the weighted average results, we can see that the forecasted decrease in Swissmetro use is small enough that the proposed 10% fare increase would have a positive effect on revenue. (61 * 1 < 59 * 1.1).

Methods:

For the base case,
*MNL_SM_price_base_other.** is used to estimate the preferences for people neither youth nor senior.
In this model, all observations where AGE is 1, 5, or 6 are excluded. The SM_COST variable is not changed.

*MNL_SM_price_base_youth.** is used to estimate the preferences for youth.
In this model, all observations where AGE is 2, 3, 4, 5, or 6 are excluded. The SM_COST variable changed to reflect the 20% discount. *(SM_COST = DefineVariable('SM_COST', 0.8 * SM_CO * (GA == 0),database).*
*MNL_SM_price_base_senior.** is used to estimate the preferences for youth.
In this model, all observations where AGE is 1, 2, 3, 4, or 6 are excluded. The SM_COST variable changed to reflect the 50% discount. *(SM_COST = DefineVariable('SM_COST', 0.5 * SM_CO * (GA == 0),database).*

For the forecasts of the effect of the fare increase,
*MNL_SM_price_forecast_other.**
*MNL_SM_price_forecast_youth.**
*MNL_SM_price_forcast_senior.**
Are used similarly to the base case. The only change is that each model file includes the fare increase by redefining the *SM_COST* variable with a 1.1 scalar multiplier.

I compiled the results in a spreadsheet and calculated the market segment group weights to compute the weighted average.

	Group Weights			
	Youth	**Senior**	**Other**	**Total**
Count	423	432	5904	6759
Weight	0.06258322237	0.06391478029	0.8735019973	1

Table 5: Market segment group weights calculation (by age groups)

Weight is calculated as group (e.g. Youth) count divided by total count.

Swissmetro and regular train frequency changes

Files:
*MNL_SM_headway_base_all.**
*MNL_SM_headway_base_commuter.**
*MNL_SM_headway_base_business.**
*MNL_SM_headway_base_leisure.**
*MNL_SM_headway_base_other.**
*MNL_SM_headway_forecast_all.**
*MNL_SM_headway_forecast_commuter.**
*MNL_SM_headway_forecast_business.**
*MNL_SM_headway_forecast_leisure.**
*MNL_SM_headway_forecast_other.**

In this section I investigate the impact of the proposed policy scenario:
The Swissmetro SA is considering an alternative option of making incremental investment in Swissmetro and initially starting with half the maglev trains they originally planned to purchase. To meet the growing demand, they are also considering doubling the frequency of the regular trains.

This change would mean an increase in the headway of Swissmetro by a factor of 2, and a decrease in the headway of regular trains by a factor of 2.

This report investigates the effect of this change on both the overall population, and the market segmented by travel purpose (PURPOSE):

- Commuter (1: commuter and 5: return from work)
- Business (3: business, 7: return from business)
- Leisure (4: leasure, 8: return from leisure)
- Other (2: shopping, 6: return from shopping, 9: other)

This market segmentation is motivated by the intuition that individuals' preferences will be affected by changes in SM/train frequency differently depending on the reason they are traveling. For example, people traveling for leisure may be less sensitive to changes in SM/train frequency, while commuters making a choice about which transit alternative to regularly take them to and from work are expected to be more sensitive to changes in SM/train frequency.

I hypothesize that all market segments will have an increased preference for the regular train and a decreased preference in the Swissmetro due to the frequency changes. I expect this effect to be more pronounced among the commuter and business market segments, and least pronounced for the leisure market segment.

Results
The base case in this investigation is without the change in Swissmetro and regular train frequency (headway). The forecasted estimates are for the case where Swissmetro headway increases by a factor of 2, and regular train headway decreases by a factor of 2.

	Base Case					Forecast				
	Commuter	Business	Leisure	Other	All	Commuter	Business	Leisure	Other	All
CAR	25	26	33	23	27	26	27	33	23	28
TRAIN	15	13	11	19	14	18	16	11	23	17
SM	60	61	56	58	59	55	57	56	54	55

Table 6: Market shares (percent) for Swissmetro and Train frequency changes.

The forecasted results from the model generally match expectations.
There is an overall decrease in Swissmetro as a preferred travel alternative, as Swissmetro headway is doubled, and there is an overall increase in regular train as a preferred travel alternative, as train headway is halved.

There is no change in the estimated preferences for the leisure traveler market segment. This result is more exaggerated than the hypothesis that the leisure market segment would be least affected by the changes, given that the leisure market segment seems unaffected.

The effect of the changes on the commuter and business market segments are not more pronounced than the overall population.

Methods

For the base case:
*MNL_SM_headway_base_all.** is used to estimate the base case preferences for the entire population. All traveler PURPOSE values are included.

*MNL_SM_headway_base_commuter.** is used to estimate the base case preferences for commuters. Travelers are excluded with PURPOSE values 2, 3, 4, 6, 7, 8, 9.

*MNL_SM_headway_base_business.** is used to estimate the base case preferences for business travelers. Travelers are excluded with PURPOSE values 1, 2, 4, 5, 6, 8, 9.

*MNL_SM_headway_base_leisure.** is used to estimate the base case preferences for leisure travelers. Travelers are excluded with PURPOSE values 1, 2, 3, 5, 6, 7, 9.

*MNL_SM_headway_base_other.** is used to estimate the base case preferences for "other" (and shopping) travelers. Travelers are excluded with PURPOSE values 1, 3, 4, 5, 7, 8.

Forecast case:
*MNL_SM_headway_forecast_all.**
*MNL_SM_headway_forecast_commuter.**
*MNL_SM_headway_forecast_business.**
*MNL_SM_headway_forecast_leisure.**
*MNL_SM_headway_forecast_other.**

These files are used to make similar exclusions as the base case files, but they also modify the swissmetro headway (SM_HE) and train headway (TRAIN_HE) variables to incorporate the proposed changes:

```
TRAIN_HEADWAY = DefineVariable('TRAIN_HEADWAY', TRAIN_HE * 0.5,database)
SM_HEADWAY = DefineVariable('SM_HEADWAY', SM_HE * 2,database)
```

The new TRAIN_HEADWAY and SM_HEADWAY variables are used in the model utility definitions.

Code

See the attached files for model specification code and spreadsheet used for the estimations results. The spreadsheet is also available at this link:
https://docs.google.com/spreadsheets/d/1Trynn9UCfx2zXMHEINRc10fu8kSixLT1-7urp4Rh2RY

Supplemental Problems

See attached.