Mathematical Modeling of Behavior (2024-2025)

Michel Bierlaire.

TAs: Cloe Cortes, Pavel Ilinov, Evangelos Paschalidis, Anne-Valérie Preto, Negar Rezvany

Mid-term wrap-up — November 12, 2024

Question 1: Specification testing and forecasting

Consider the following transportation mode choice model (M_1) for the *Swissmetro* (SM) case study, which involves three alternatives: car, rail and Swissmetro. The utility functions are defined as

$$U_{in} = V_{in} + \varepsilon_{in},$$

where i denotes the alternative, n the individual, ε_{in} are the error terms, i.i.d. EV(0, 1), and V_{in} are the deterministic part of the utility functions, specified as follows:

$$\begin{aligned}
V_{car,n} &= ASC_{car} + \beta_c cost_{car,n} + \beta_t time_{car,n} + \beta_{f,car} female_n, \\
V_{rail,n} &= ASC_{rail} + \beta_c cost_{rail,n} + \beta_t time_{rail,n} + \beta_{f,rail} female_n, \\
V_{SM,n} &= \phantom{ASC_{rail} +\ } \beta_c cost_{SM,n} + \beta_t time_{SM,n},
\end{aligned}$$

where $cost_{i,n}$ is the travel cost in CHF associated with alternative i and individual n, $time_{i,n}$ is the travel time in minutes of alternative i and individual n, and $female_n$ is a binary variable that takes value 1 if individual n is female and 0 otherwise. The estimates for the parameters ASC_{car}, ASC_{rail}, β_c, β_t, $\beta_{f,car}$ and $\beta_{f,rail}$ can be found in Table 1 on the following page.

Interpretation of the specification

1. What are the parameters ASC_{car} and ASC_{rail}, and what are they capturing?
2. The model does not involve a parameter ASC_{SM}. Why?
3. What is the behavioral interpretation of the sign of β_{cost} and β_{time}?
4. The model does not involve a term $\beta_{f,SM} female_n$. Why?
5. What is the value of the alternative specific constants for females?
6. What is the value of the alternative specific constants for males?

Parameter number	Description	Coeff. estimate	Robust Asympt. std. error	t-stat	p-value
1	ASC_{car}	-0.461	0.0973	-4.74	0.00
2	ASC_{rail}	0.0906	0.0913	0.99	0.32
3	β_c	-0.0108	0.000670	-16.18	0.00
4	β_t	-0.0125	0.00105	-11.81	0.00
5	$\beta_{f,car}$	0.309	0.102	3.04	0.00
6	$\beta_{f,rail}$	-1.23	0.0792	-15.53	0.00

Summary statistics

Number of observations = 6768

Number of excluded observations = 3960

Number of estimated parameters = 6

$$\begin{aligned}
\mathcal{L}(\beta_0) &= -6964.663 \\
\mathcal{L}(\hat{\beta}) &= -5187.983 \\
-2[\mathcal{L}(\beta_0) - \mathcal{L}(\hat{\beta})] &= 3553.359 \\
\rho^2 &= 0.255 \\
\bar{\rho}^2 &= 0.254
\end{aligned}$$

Table 1: Estimation results for M_1

7. Specify an equivalent model (M_2) that involves the variable male_n instead of female_n. That variable is 0 if individual n is a female, and 1 otherwise. Provide the estimated value of the parameters for this new specification.

8. Specify an equivalent model (M_3) that explicitly includes an alternative specific constant for females, and an alternative specific constant for others. Provide the estimated value of the parameters for this new specification.

Testing

We now characterize a new model (M_4) as follows:

$$
\begin{aligned}
V_{\text{car},n} &= \text{ASC}_{\text{car}} + \beta_c \text{cost}_{\text{car},n} + \beta_{t,\text{car}} \frac{\text{time}_{\text{car},n}^{\lambda} - 1}{\lambda} + \beta_{f,\text{car}} \text{female}_n\,, \\
V_{\text{rail},n} &= \text{ASC}_{\text{rail}} + \beta_c \text{cost}_{\text{rail},n} + \beta_{t,\text{rail}} \text{time}_{\text{rail},n} + \beta_{f,\text{rail}} \text{female}_n + \beta_{\text{GA},\text{rail}} \text{GA}_n, \\
V_{\text{SM},n} &= \beta_c \text{cost}_{\text{SM},n} + \beta_{t,\text{SM}} \text{time}_{\text{SM},n} + \beta_{\text{GA},\text{SM}} \text{GA}_n,
\end{aligned}
$$

where all the elements are defined in the same way as for model M_1, GA_n is a variable that takes value 1 if individual n owns a yearly subscription, and 0 otherwise, and λ is the parameter of the Box-Cox transform, to be estimated. The estimates for the parameters can be found in Table 2 on the next page.

1. What does it mean for a coefficient to be significant? Is there any parameter in Table 2 on the following page that is not significant? What are the implications in terms of model specification?

2. We want to use a likelihood ratio test to compare the two specifications M_1 and M_4. What is the null hypothesis for such a test? What are the corresponding linear restrictions?

3. Which model is preferred using a level of significance of 5%? Use Table 3 on the next page to justify your answer.

4. Specify a new linear-in-parameter model **with the exact same number of parameters** as M_4, and call it M_5. It shall capture the fact that the marginal effect of travel cost in the utility varies with travel cost.

5. What statistical tests can be used to compare M_5 against model M_4? Explain how they are used and what are the possible outcomes.

6. What is the expression for the value of time associated with each alternative and individual according to model M_5?

Parameter number	Description	Coeff. estimate	Robust Asympt. std. error	t-stat	p-value
1	ASC_{car}	0.549	0.565	0.97	0.33
2	ASC_{rail}	-1.58	0.139	-11.32	0.00
3	β_c	-0.0108	0.000730	-14.82	0.00
4	$\beta_{GA,SM}$	0.458	0.203	2.25	0.02
5	$\beta_{GA,rail}$	2.34	0.213	11.00	0.00
6	$\beta_{t,SM}$	-0.0122	0.00181	-6.73	0.00
7	$\beta_{t,car}$	-0.0684	0.0643	-1.06	0.29
8	$\beta_{t,rail}$	-0.0123	0.00115	-10.75	0.00
9	$\beta_{f,car}$	-0.428	0.102	-4.20	0.00
10	$\beta_{f,rail}$	1.10	0.0863	12.74	0.00
11	λ	0.647	0.197	3.29	0.00

Summary statistics

Number of observations $= 6768$

Number of excluded observations $= 3960$

Number of estimated parameters $= 11$

$$\begin{aligned}
\mathcal{L}(\beta_0) &= -6964.663 \\
\mathcal{L}(\hat{\beta}) &= -4936.917 \\
-2[\mathcal{L}(\beta_0) - \mathcal{L}(\hat{\beta})] &= 4055.492 \\
\rho^2 &= 0.291 \\
\bar{\rho}^2 &= 0.290
\end{aligned}$$

Table 2: Estimation results for M_4

k	$P(X \geq x) = 5\%$	k	$P(X \geq x) = 5\%$	k	$P(X \geq x) = 5\%$
1	3.841	4	9.488	7	14.07
2	5.991	5	11.07	8	15.51
3	7.815	6	12.59	9	16.92

Table 3: Critical values of a χ^2 distribution with k degrees of freedom for a 5% level of significance