Solutions to the
Examination for the Master's Course
Methods of Econometrics, winter semester 2022/23

Problem 1 (35 points)

Given:

(i) a physics-based model for the energy consumption y of electrical vehicles (EV) per kilometer [kWh/km] at constant speed v [km/h] on level roads:

$$y = \beta_0 + \beta_1 \frac{1}{v} + \beta_2 v^2 + \epsilon, \quad \epsilon \sim \text{ i.i.d. } N(0, \sigma^2).$$

(ii) OLS estimate:

$$\hat{\beta}_0 = 0.08, \quad \hat{\beta}_1 = 2.0, \quad \hat{\beta}_2 = 1.1\text{e} - 5.$$

(iii) Variance-covariance matrix

$$\hat{\mathbf{V}}(\hat{\boldsymbol{\beta}}) = \begin{pmatrix} 0.0011 & -0.02 & -1.5\text{e}-8 \\ -0.02 & 1.0 & -1.2\text{e}-7 \\ -1.5\text{e}-8 & -1.2\text{e}-7 & 1\text{e}-11 \end{pmatrix}$$

(a) The factors of the model $y = \sum_{j=0}^{2} \beta_j x_j + \epsilon$ are given by

$$x_0 = 1, \quad x_1 = \frac{1}{v}, \quad x_2 = v^2.$$

Notice that one exogenous variable, the speed v, leads here to three factors

(b) (i) $H_{01} : \beta_1 = 0$:

$$T = \frac{\hat{\beta}_1}{\sqrt{\hat{V}_{11}}} \sim T(10-3), \quad t^{\text{data}} = 2, \quad \mathcal{R} = \{t : |t| > t_{0.975}^{T(7)} = 2.305\} : \text{not rejected}$$

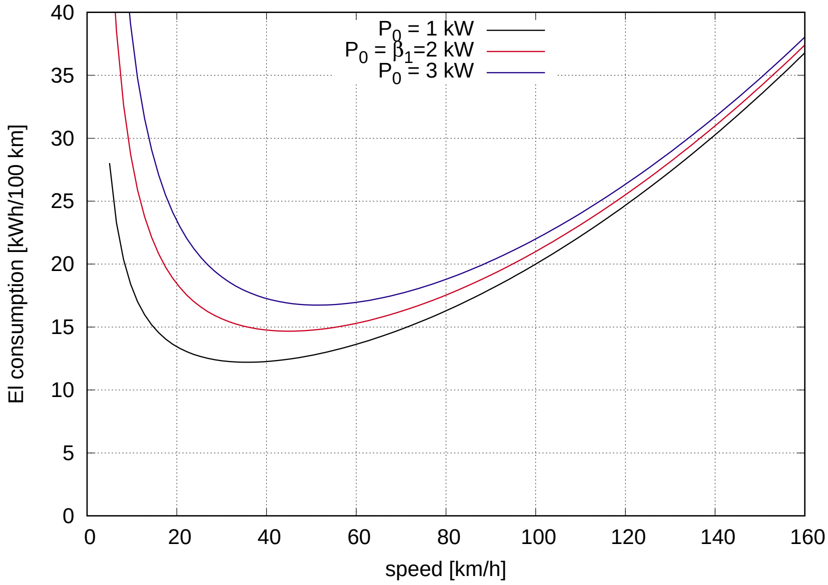
(ii) $H_{01} : \beta_1 \geq 1$:

$$T = \frac{\hat{\beta}_1 - 1}{\sqrt{\hat{V}_{11}}} \sim T(10-3), \quad t^{\text{data}} = 1, \quad \mathcal{R} = \{t : t < -t_{0.95}^{T(7)} = -1.895\} : \text{not rejected}$$

(c) Expected energy consumption per kilometer $\hat{y}(\boldsymbol{x}) = \hat{y}(v)$ as a function of the speed v in km/h (whatch out for the units! Use the one given in the problem setting!)[1]

- v=10 km/h: $\hat{y} = 0.281$ kWh/km = 28.1 kWh/100km
- v=30 km/h: $\hat{y} = 0.157$ kWh/km = 15.7 kWh/100km
- v=50 km/h: $\hat{y} = 0.148$ kWh/km = 14.8 kWh/100km
- v=100 km/h: $\hat{y} = 0.21$ kWh/km = 21 kWh/100km
- v=130 km/h: $\hat{y} = 0.281$ kWh/km = 28.1 kWh/100km

[1] In numerical simulations, particularly physics based ones, it is always a good idea to use SI units (s, m, m/s, Ws, Ws/m) since fatal unit errors are very probable, otherwise. However, this does not apply when applying standard statistical methods rather than simulation.

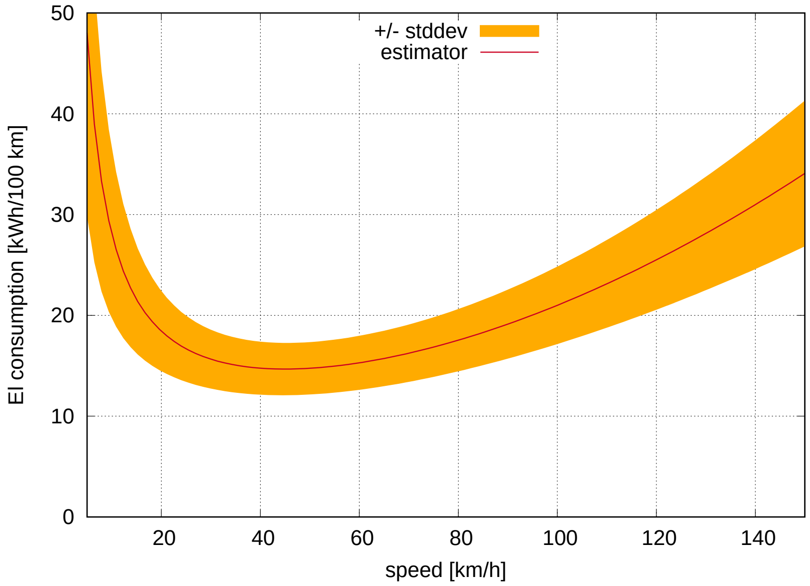


(d) The energy consumption $y(\hat{\boldsymbol{\beta}}, v)$ is a linear function of the elements of $\hat{\boldsymbol{\beta}}$, so both the variances and covariances of the estimation errors are needed:

$$\hat{V}_{\hat{y}}(v) = \hat{V}_{00} + \hat{V}_{11}\frac{1}{v^2} + \hat{V}_{22}v^4 + 2(\hat{V}_{01}\frac{1}{v} + \hat{V}_{02}v^2 + \hat{V}_{12}v)$$

For $v = 50[km/h]$, we obtain $\hat{V}_{\hat{y}}(50) = 0.0545\,(\mathrm{kWh/km})^2$

For illustrative purposes, the $\pm 1\sigma$-band $\left[\hat{y} - \sqrt{\hat{V}_{\hat{y}}}, \hat{y} + \sqrt{\hat{V}_{\hat{y}}}\right]$ is plotted for the standard value $P_0 = \hat{\beta}_1 = 2\,\mathrm{kW}$



(e) As usual, to find an extremal value, you set the derivative equal to zero:

$$\hat{y}'(v) = -\frac{\hat{\beta}_1}{v^2} + 2\hat{\beta}_2 v \stackrel{!}{=} 0, \quad \Rightarrow \quad v_{\min} = \left(\frac{\hat{\beta}_1}{2\hat{\beta}_2}\right)^{1/3} = 45.0\,\mathrm{km/h}$$

Since for very small v the consumption increases as $1/v$ and for very large speeds as v^2, this extremum is also a minimum.

(f) – Small basis power demand of $\beta_1 = 1\,\mathrm{kW}$: $v_{\min} = 35.7\,\mathrm{km/h}$

- Large basis power demand of $\beta_1 = 3\,\mathrm{kW}$: $v_{\mathrm{min}} = 51.5\,\mathrm{km/h}$

Because the basis power demand increases the consumption/km particularly for small speeds (simply because the car takes longer), the optimal speed v_{min} generally increases with the size of the BEV (see the figure). It is also lower by some km/h compared to gasoline/Dieel vehicles because, unlike electrical engines with an essentially constant efficiency, the gas/Diesel engines are more effective at a higher power demand, i.e., higher speeds.

Problem 2 (30 points)

(a) Combined material and fuel repository y^{s} for all three life phases:

$$\boldsymbol{y}^s = \begin{pmatrix} 9\,000\,\mathrm{kg} \\ 250\,\mathrm{kg} \\ 2\,000\,\mathrm{kg} \\ 400\,\mathrm{kg} \\ 800\,\mathrm{kg} \\ 600\,000\,\mathrm{l} \end{pmatrix}$$

(b) The total CO_2 emissions e_1^{mat} for making and recycling the bus (without driving it) are given by

$$e_1^{\mathrm{mat}} = \sum_{i=1}^{5} y_i^s C_i = 32\,950\,\mathrm{kgCO_2}$$

(c) The CO_2 emissions per kilometer of driving are given by $e_1' = \frac{y_6^s C_6}{1.2\,10^6\,\mathrm{km}}$ or also by $e_1' = 0.5\,\mathrm{l/km} * (0.4 + 2.7)\,\mathrm{kg/l} = 1.55\,\mathrm{kg/km}$

(d) Setting the driving emissions $e_1' x$ after kilometrage x equal to e_1^{mat} gives

$$x = \frac{e_1^{\mathrm{mat}}}{e_1'} = 21\,258\,\mathrm{km}$$

The overwhelming majority of the life-time emissions is produced during the operation phase driving 1.2 milllion kilometers. Less than 2 % of the total LCA CO_2 emissions are produced during the production and the wrecking of the bus!

(e) Footprint per passenger kilometer during the driving phase alone:

$$e_{\mathrm{pkm}} = e_1'/12 = \frac{e_1^{\mathrm{drive}}}{14.4\,10^6\,\mathrm{pkm}} = \frac{y_6^s C_6}{14.4\,10^6\,\mathrm{pkm}} = 0.129\,\mathrm{kg/pkm}$$

(e) Footprint per passenger kilometer considering the whole life cycle:

$$e_{\mathrm{pkm}} = \frac{e_1^{\mathrm{mat}} + e_1^{\mathrm{drive}}}{14.4\,10^6\,\mathrm{pkm}} = 0.131\,\mathrm{kg/pkm}$$

Because, per assumption, this bus had so little demand, the footprint of about 130 g/km per passenger is only somewhat lower than typical car footprints at an average occupancy of 1.4 persons

Problem 3 (55 points)

(a) It is a revealed preference (RP) survey because the participants have been asked about hypothetical situations.

(b) Complete: There is no third option such as "Do not buy any car". This, of course, has to be made explicit in the questionaire.

Unique: There is no possibility to buy both cars, only exactly one

(c) Because of the high density of the gas stations and the large range, this is not really a buying criterion for a gasoline car. Regarding fueling time, this is so short with respect to the charging time that it is negligible (you could set, to a good approximation, a time of zero for the ICEV).

(d) AC: δ_{i1}, characteristics of alternatives: C_i, C_i', R and T. While R and T are given for BEVs, only, these are still characteristics and no socio-economic variables although, formally, they are modelled in an alternative-specific way as socio-economic variables (because values are missing for the ICEV).

- β_1: No really sensible meaning: Preference BEV over ICEV if the car and kilometer costs are the same, the charging time is zero but also the range is zero.
- β_2: Price sensitivity; should be < 0
- β_3: Sensitivity to the kilometer costs; should be < 0
- β_4: Appraisal of a higher range; should be positive
- β_5: Sensitivity to the charging time; should be < 0

Realized property somes for the factors x_1 and x_5:

- $X_1^{\text{data}} = \sum_n y_{n1} = 24$: Number of choices for the BEV
- $X_5^{\text{data}} = \sum_n y_{n1} T_{n1} = 3960$: Sum of the charging times [minutes] if everybody opting for a BEV in any choice set charges his/her BEV once

The expected property sums for the null model $\boldsymbol{\beta} = 0$ can be calculated with the choice probabilities of this model $P_{ni} = P_0 = 1/2$:

- $X_1^{\text{MNL}} = 5 \sum_n \sum_i y_{ni} = 40$: Half of the total number of decisions
- $X_5^{\text{MNL}} = 5 \sum_n T_1 = 8\,700$: Five times the sum of all charging times in the ten situations

The null model is not the correctly calibrated model since not all realized and expected property sums are the same

(f) Binomial logit model:

$$P_{n1} = \frac{\exp V_{n1}}{\exp V_{n1} + \exp V_{n2}}$$

For the first choice set, we have

$$V_{11} = \hat{\beta}_1 + 30\hat{\beta}_2 + 20\hat{\beta}_3 + 300\hat{\beta}_4 + 180\hat{\beta}_5 = -5.51,$$
$$V_{12} = 30\hat{\beta}_2 + 20\hat{\beta}_2 = -5.21$$

we obtain

$$N_1 = e^{V_{11}} + e^{V_{12}} = 0.00954$$

and

$$P_{11} = e^{V_{11}}/N = 0.426, \quad P_{12} = 1 - P_{11} = 0.574$$

The same with the second choice set:

$$V_{21} = -7.52, \quad V_{22} = -5.21, \quad P_{21} = 0.090, \quad P_{22} = 0.910$$

(g) In discrete-choice modelling, we assume the asymptotic limit, i.e., the test functions $T_m = \hat{\beta}_m/\sqrt{V_{mm}}$ are standardnormal. The parameter $\hat{\beta}_m$ is significantly if the null hypothesis $H_{0m}: \beta_m = 0$ can be rejected at $\alpha = 5\,\%$:

$$H_{0m} \text{ rejected } \Leftrightarrow |t_m^{\text{data}}| > z_{0.975}$$

We have

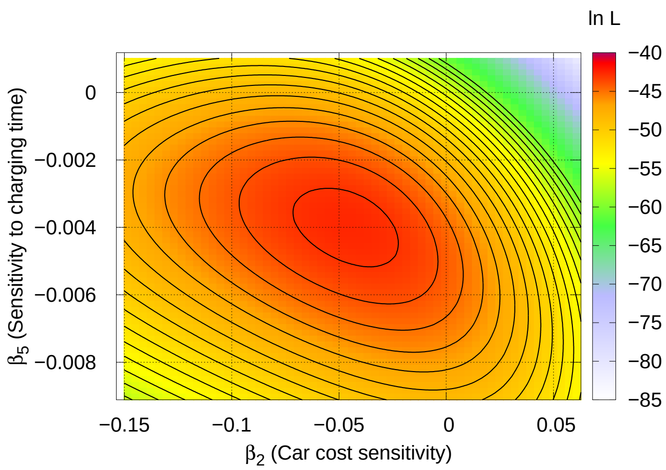
$$t_1^{\text{data}} = 0.13, \quad t_2^{\text{data}} = -2.33, \quad t_3^{\text{data}} = -1.10, \quad t_4^{\text{data}} = 1.16, \quad t_5^{\text{data}} = -2.15$$

while $z_{0.975} = 1.96$. Therefore, only β_2 (price sensitivity) and β_5 (sensitivity to charging time) are significant

(h) Likelihood ratio test between the restricted model with #dof=2 and the full model with #dof=5:

1. H_0: The restrained model is equally predictive as the full model
2. Test statistic $T = 2(\tilde{L} - \tilde{L}^{\text{restr}}) \sim \chi^2(5-2)$
3. Data realisation: $t^{\text{data}} = 25$
4. Decision: H_0 rejected if $t^{\text{data}} > \chi^2_{3,0.95} = 7.815 \Rightarrow$ rejected.

Contour lines of the log-likelihood of the restrained model:



(i) Too little data