Solutions to the
Examination for the Masters Course
Methods of Econometrics, winter semester 2019/20

Problem 1 (30 points)

(a) In this context, Leontief's Input-output model (IOM) describes the product and service streams between the car manufacturers, their suppliers, the remaining sectors of this national economy, and the consumers. Specifically, x_1 denotes the value of produced vehicles, x_2 the value of the parts needed by the car producers, x_3 the monetary value of all the other products and services, and y_i what is delivered from sector i to the end consumers. The assumptions are

(i) linearity (double input leads to double output),

(ii) stationarity (no buildup or depletion of storages).

The coefficients A_{ij} of the IO-matrix A denote the amount of material and services from sector i (in €) to produce one €worth of product j. Specifically, $A_{21} = 0.7$ indicates that 70 % of the value of a new car are delivered to the car manufacturers by their suppliers, which is plausible.

(b) The second line of the IOM equation $\boldsymbol{x} = \boldsymbol{y} + \mathsf{A}\,\boldsymbol{x}$ reads

$$x_2 = y_2 + A_{21}x_1 + A_{22}x_2 + A_{23}x_3 = A_{21}x_1 + A_{22}x_2 + A_{23}x_3$$

(since no car parts are delivered to the end consumers, $y_2 = 0$). Hence

$$A_{23} = \frac{x_2 - A_{21}x_1 - A_{22}x_2}{x_3} = 0.005.$$

(c)

$$\boldsymbol{y} = \boldsymbol{x} - \mathsf{A}\,\boldsymbol{x} = (\mathbf{1} - \mathsf{A})\,\boldsymbol{x} = \begin{pmatrix} 0.95 & -0.10 & -0.01 \\ -0.70 & 0.9375 & -0.005 \\ -0.10 & -0.40 & 0.70 \end{pmatrix} \begin{pmatrix} 1 \\ 0.8 \\ 10 \end{pmatrix} = \begin{pmatrix} 0.77 \\ 0 \\ 6.58 \end{pmatrix}.$$

(d) The production vector as a function of the demand vector is given by an inversion of $\mathbf{1} - \mathsf{A}$:

$$\boldsymbol{x} = (\mathbf{1} - \mathsf{A})^{-1}\,\boldsymbol{y} \equiv \mathsf{B}\,\boldsymbol{y}.$$

Because of linearity, this also applies to the changes. With a demand change of $\Delta\,\boldsymbol{y} = (0.1y_1, 0, 0)^{\mathrm{T}} = (0.077, 0, 0)^{\mathrm{T}}$, we have

$$\Delta\,\boldsymbol{x} = \mathsf{B}\,\Delta\,\boldsymbol{y} = \begin{pmatrix} 1.15 & 0.13 & 0.0174 \\ 0.862 & 1.17 & 0.0207 \\ 0.657 & 0.686 & 1.44 \end{pmatrix} \begin{pmatrix} -0.077 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} -0.0886 \\ -0.0664 \\ -0.0506 \end{pmatrix}.$$

The new outputs in the three sectors are

$$\boldsymbol{x}^{\text{ new}} = \boldsymbol{x} + \Delta\,\boldsymbol{x} = \begin{pmatrix} 0.911 \\ 0.734 \\ 9.95 \end{pmatrix}$$

and the relative changes

$$\frac{\Delta x_1}{x_1} = -8.86\,\%, \quad \frac{\Delta x_2}{x_2} = -8.30\,\%, \quad \frac{\Delta x_3}{x_3} = -0.51\,\%,$$

So, a drop in the demand for vehicles also affects the supplyers and, to a lesser extent, all the other sectors: Everything in a national economy is coupled!

Watch out for the following common errors!

(d) Read carefully: A 10 % drop from 0.77 means -0.077, not -0.1

Problem 2 (45 points)

Binary-choice situation for accepting or rejecting gaps when crossing a street at an unsignalized point:

Gap index n	1	2	3	4	5	6	7	8	9	10	11	12
Time gap T [s]	7	5	9	5	4	1	5	4	5	2	6	8
speed v [km/h]	50	15	30	40	15	50	60	20	30	40	60	45
Vehicle type	truck	bike	car	car	bike	car	car	bike	car	car	car	car
Gap accepted ($i=1$)	1	3	3	1	1	0	0	1	1	0	1	2
Gap rejected ($i=2$)	2	0	1	2	1	1	2	2	1	2	2	0

(a) None of the exogenous variables influencing the decision (time gap, speed, vehicle type) is a characteristic of the alternatives: If a truck is approaching, this is so regardless whether the gap is accepted or not. All three influencing variables can be considered as external variables (like the weather in the exercise examples) and therefore need to me modelled in an alternative-specific way with a reference alternative.

In particular, we can set $V_2 = 0$ if we assume alternative $i = 2$ (waiting) to be the reference for all the alternative-specific factors and also for the AC.

Comment: Generally, we can *always* set the determinsitic utility of one alternative equal to zero since only utility differences matter (translation invariance). However, if we have characteristics assuming nonzero values for all alternatives (which we do not have here), this will be a counterintuitive formulation, in most cases.

(b)
- β_2 denotes the increase of the preference for walking instead of waiting per additional second of the time gap. Since longer gaps are good, $\beta_2 > 0$ expected.
- β_3 denotes the increase of the utility for accepting the gap (crossing right now) per additional km/h of the approaching vehicle. Since higher speeds are bad for the pedestrians (even if the gap is measured in seconds rather than meters), $\beta_3 < 0$ is expected.
- β_4: Increase in utility for crossing right now if a bicycle instead of a car is approaching. Since a bike is less dangerous/intimitading than a car, β_4 should be positive.
- β_5: The same for approaching trucks. Obviously, $\beta_5 < 0$.

Comment: β_1 denotes the preference for walking instead of waiting if the gap is zero and a car is approaching at zero speed: Since this is beyond the applicability limits, nothing can be said about the expected sign (not required in the examination).

(c) In order to calculate the model property sum for $\boldsymbol{\beta} = \mathbf{0}$, we observe that always $P_{ni} = 1/2$ and the total number of taken decisions is $N = 30$.

- Property sum X_1 for the AC: Total number of accepted gaps

$$X_1^{\text{data}} = \sum_n \sum_i \delta_{i1} y_{ni} = \sum_n y_{n1} = 14, \quad \hat{X}_1(\boldsymbol{\beta} = \mathbf{0}) = N/2 = 15$$

- Property sum X_4: Number of accepted gaps if a bicycle is approaching. The total number of decisions with an approaching bike is $N_{\text{ bike}} = 8$:

$$X_4^{\text{data}} = \sum_n y_{n1} \begin{cases} 1 & \text{bike approaching} \\ 0 & \text{otherwise} \end{cases} = 5, \quad \hat{X}_1(\boldsymbol{\beta} = \mathbf{0}) = \frac{N_{\text{ bike}}}{2} = 4.$$

– Property sum X_5: Number of accepted gaps if a truck is approaching ($N_{\text{truck}} = 3$):

$$X_5^{\text{data}} = \sum_n y_{n1} \begin{cases} 1 & \text{truck approaching} \\ 0 & \text{otherwise} \end{cases} = 1, \quad \hat{X}_1(\boldsymbol{\beta} = \mathbf{0}) = \frac{N_{\text{truck}}}{2} = 1.5.$$

$\boldsymbol{\beta} = \mathbf{0}$ is not a ML calibration of the Logit model since the realized and modelled property sums are different from each other.

(d) For the first gap $n = 1$, we have

$$x_2 = T = 7\,\text{s}, \quad x_3 = v = 50\,\text{km/h}, \quad x_4 = 0, \quad x_5 = 1 \quad \text{(a truck is approaching)},$$

so

$$V_1 = \hat{\beta}_1 + 7\beta_2 + 50\beta_3 + \beta_5 = -0.693$$

and

$$P_1 = \frac{e^{V_1}}{e^{V_1} + 1} = 0.333$$

This coincides with the observed percentaged frequencies of 1/3 but this is by chance.

(e) A factor is "relevant" if its associated model parameter is significantly different from zero, hence we have to test the null hypotheses $\beta_2 = 0$ and $\beta_3 = 0$, respectively.

1. $H_0 : \beta_2 = 0$,
2. Test statistics $T = (\hat{\beta}_2 - 0)/\sqrt{\hat{V}(\hat{\beta}_2)} \sim N(0, 1)$ if H_0,
3. Data value: $t_{\text{data}} = 0.747/0.369 = 2.02$ (notice that the standard deviations, not the variances, are given!),
4. Decision: H_0 rejected if $|t_{\text{data}}| > z_{0.975} = 1.96$: Yes

For $H_0 : \beta_3 = 0$, we obtain similarly

$$t_{\text{data}} = -0.0369/0.0508 = -0.727, \quad |t_{\text{data}}| < z_{0.975} = 1.96 \;\Rightarrow\; \text{not rejected}$$

Comment 1: The p-values $p = 2(1 - \Phi(|z_{\text{data}}|)$ (not required) are $p_2 = 4.3\,\%$ and $p_3 = 46.8\,\%$, respectively.

Comment 2: For tests of discrete-choice model apraмeters, we always assume the asymptotic expansion which requires a minimum number N of decisions. Here, $N = 30$ is just sufficient.

(f) Since $\hat{\beta}_4$ denotes the utility difference between approaching bikes and cars and $\hat{\beta}_5$ that between trucks and cars, $\hat{\beta}_4 - \hat{\beta}_5$ denotes the difference between bicycles and trucks.

Test of $\beta_4 - \beta_5 = 0$:

– $H_0 : \gamma = \beta_4 - \beta_5 = 0$

– Test statistics: $T = \hat{\gamma}/\sqrt{V(\hat{\gamma})} \sim N(0, 1)$ where the error variance is given by

$$V(\hat{\gamma}) = V(\hat{\beta}_4 - \hat{\beta}_5) = V(\hat{\beta}_4) + V(\hat{\beta}_5) - 2\,\text{Cov}(\hat{\beta}_4, \hat{\beta}_5) = 0.63$$

– Data value (notice the squares at 1.09 and 1.46 since these numbers denote the standard deviations!):

$$t_{\text{data}} = \frac{1.04 - (-0.88)}{\sqrt{1.09^2 + 1.46^2 + 2 * 0.355}} = 0.78$$

Decision: $|t_{\text{data}}| < z_{0.975} = 1.96 \;\Rightarrow\; H_0$ cannot be rejected.

Comment: Notice that the p-value of 0.43 is of the same order as that for β_4 and β_5 separately. It would have been smaller if there was no correlation/covariance.

(g) There are too few data points for a five-parameter model: We have just 12 situations with 30 decisions.

Watch out for the following common errors!

(a) V_i does *not* mean that a utilility of an alternative is zero since the utility is not defined: Only the utility *difference* is a useful concept! So, $V = 0$ can serve as a reference: If another alternative is better, it has a positive utility.

(b) Don't copy past problems, tutorials, or solutions to old exams! In contrast to all these, the time sensitivity here is expected to be *positive* since T is here the time gap, where more is better, and not the travel time where less is better!

(c) Don's solve extra problems such as calculating property sums for the factors of β_0, β_3, and β_4. This only costs time and brings no extra points.

(e) In the maximum-likelihood estimation, we always use the asymptotic expansion where parameter estimation errors are Gaussian, i.e., the test statistic is standard-normal instead of student-t distributed. The exact distribution for small samples is much more complicated than a simple Student-t distribution.

(e) as well: If a parameter is "significant", this always means that the hypothesis that it is *equal* to zero can be rejected (if nothing else is stated, at $\alpha = 5\,\%$). So no $\geq$ or $\leq$ null hypotheses to check if a parameter is significant!

Problem 3 (45 points)

Analysis of data from bike computers:

Person/section	1	2	3	4	5	6	7	8	9	10
Speed[m/s]	8	9	2.8	11.5	6	4.5	2.5	6	12.5	11
Acceleration[m/s^2]	0	0	0	0.2	0.4	0	-0.3	0	0	0
Gradient [%]	0	0	5	-3	0	2	6	0	-3	0
Mass[kg]	78	78	78	69	69	69	92	92	92	92
Power [W]	134	166	135	212	240	100	89	87	35	280

(a) The wind drag depends on the wind speed, only, while all other contributions (rolling resistance, gradient, acceleration) are proportional to the mass m. Hence, a pure analysis of the influencing variables leaves only $\beta_1 v_i^3$ to describe the contribution of the wind drag.

Remark: Of course, you may know that the wind drag increases drastically with the speed also leading you to the $\beta_1 v_i^3$ term (power is force times speed, and the force due to the wind drag increases quadratically with the relative wind speed). This explanation will also give full marks.

(b) We have $n = 10$ data points and $J + 1 = 5$ parameters:

- $\boldsymbol{y}$ is a $n = 10$ column vector (10×1 matrix)
- $\mathbf{X}$ is a $n \times (J + 1) = 10 \times 5$ matrix
- $\boldsymbol{\beta}$ is a $J + 1 = 5$ column vector (5×1 matrix)
- $\boldsymbol{\epsilon}$ is a $n = 10$ column vector
- $\mathrm{Cov}(\boldsymbol{\epsilon})$ is a 10×10 matrix (with identical diagonal elements and zeroes, otherwise).

(c) The first line (row) of $\mathbf{X}$ relating to the first person/section is given by

$$(\mathbf{X})_1 = (1, v_1^3, m_1 v_1, m_1 \alpha_1 v_1, m_1 a_1 v_1) = (1, 512, 624, 0, 0)$$

Remark: Notice the first entry 1 relating to the constant β_0 (first element of the first column $(1, 1, ..., 1)^{\mathrm{T}}$ of $\mathbf{X}$)

(d)
- For a standing cyclist, we have $y = \beta_0 = 0$. Hence, $\beta_0 = 0$ (and also $\epsilon_1 = 0$ contradicting the i.i.d. $N(0, \sigma^2)$ assumption for ϵ_i).
- Uphill power:

$$P_i^{\text{ up}} = m_i g \alpha_i v_i \stackrel{!}{=} \beta_3 m_i \alpha_i v_i \;\Rightarrow\; \beta_3 = g = 9.81\,\mathrm{m/s^2}$$

- Power to overcome the inertia when accelerating:

$$P_i^{\text{ acc}} = m_i a_i v_i \stackrel{!}{=} \beta_4 m_i a_i v_i \;\Rightarrow\; \beta_4 = 1$$

(e) Expected power (in W) for the first person/section (you might reuse the results from (c), here):

$$\hat{y}_1 = \hat{\beta}_0 + 8^3\hat{\beta}_1 + 78 * 8\hat{\beta}_2 = \hat{\beta}_0 + 512\hat{\beta}_1 + 624\hat{\beta}_2 = 132.4\ [W].$$

If this person drives at $v = 10\,\mathrm{m/s}$ instead of $8\,\mathrm{m/s}$:

$$\hat{y}_1(v = 10\,\mathrm{m/s}) = \hat{\beta}_0 + 10^3\hat{\beta}_1 + 78 * 10\hat{\beta}_2 = \hat{\beta}_0 + 1000\hat{\beta}_1 + 780\hat{\beta}_2 = 212.4\ [W].$$

Remark: This demonstrates how a rather small increase in speed leads to much more needed power ;-)

(f) Standard deviations:

$$\sqrt{\hat{V}(\hat{\beta}_1)} = \sqrt{5.7 * 10^{-5}} = 0.00752, \quad \sqrt{\hat{V}(\hat{\beta}_2)} = \sqrt{0.00016} = 0.01275.$$

Tests:

- $H_0 : \beta_1 = 0$

$$T = \frac{\hat{\beta}_1}{\sqrt{\hat{V}(\hat{\beta}_1)}} \sim T(10-5)|H_0 = T(5)|H_0, \quad t_{\text{ data}} = 0.136/0.00752 = 18.07$$

Decision: $|t_{\text{ data}}| > t_{0.975}^{(5)} = 2.571 \;\Rightarrow\; H_0$ rejected, i.e., β_1 is significantly different from zero and the associated factor v^3 (air drag) is significant.

- $H_0 : \beta_2 = 0$

$$T = \frac{\hat{\beta}_2}{\sqrt{\hat{V}(\hat{\beta}_2)}} \sim t(5)|H_0, \quad t_{\text{ data}} = 0.0876/0.01275 = 6.877 \;\Rightarrow\; H_0 \;\text{ rejected}$$

- $H_0 : \beta_0 = 0$

$$T = \frac{\hat{\beta}_0}{\sqrt{\hat{V}(\hat{\beta}_0)}} \sim t(5)|H_0, \quad t_{\text{ data}} = 8.14/4.636 = 1.756 \;\Rightarrow\; H_0 \;\text{ not rejected}$$

- $H_0 : \beta_3 = g = 9.81\,\mathrm{m/s^2}$

$$T = \frac{\hat{\beta}_3 - 9.81}{\sqrt{\hat{V}(\hat{\beta}_3)}} \sim t(5)|H_0, \quad t_{\text{ data}} = -0.144 \;\Rightarrow\; H_0 \text{ not rejected}$$

– $H_0 : \beta_4 = 1$

$$T = \frac{\hat{\beta}_4 - 1}{\sqrt{\hat{V}(\hat{\beta}_4)}} \sim t(5)|H_0, \quad t_{\text{ data}} = 1.01 \ \Rightarrow \ H_0 \ \text{ not rejected}$$

Remark 1 (not required): As expected, the null hypotheses $\beta_0 = 0$ (no power at no speed), $\beta_3 = g$ (physics: Power to overcome uphill=force*speed= $mg\alpha v$), and $\beta_4 = 1$ (physics, Newton's law: Power to overcome intertia=force*speed $= mav$) cannot be rejected.

Remark 2 (not required): You might have heard that the wind drag is proportional to v^2, not v^3. This is true. However, the *power* to overcome the wind drag is=force*speed resulting in another factor of v.

Remark 3 (not required): The nontrivial coefficients $\hat{\beta}_1$ and $\hat{\beta}_2$ allow for a reverse engineering of the air-drag and friction coefficients that have been assumed by the app makers. The friction coefficient will be treated in (g), and the air-drag coefficient here:

Equating the wind-drag force $F_w = 1/2 c_d \rho_L A v^2$ (where c_d is the dimensionless air-drag coefficient, $\rho_L = 1.2\,\mathrm{kg/m^3}$ is the density of air, and A the front area of the cyclist) with $\hat{\beta}_1 v^2$ gives a $c_d A$ value of

$$c_d A = \frac{\hat{\beta}_1}{2\rho_L} = 0.227\,\mathrm{m}^2.$$

Assuming a front area $A = 0.5\,\mathrm{m}^2$, this corresponds to a cd value of about 0.45 (for comparison, a modern car has a cd value of about 0.3).

(g) Equating the physical rolling resistance force $F_R = \mu m g$ with $\hat{\beta}_2 m$ gives for the assumed value of the friction coefficient the value

$$\mu = \frac{\hat{\beta}_2}{g} \approx 0.009$$

(Probably, the app makers have assumed $\mu = 0.01$ which is within estimation errors).

(h) Comparing two nested models (in the sense that one model is a restrained special case of another "full" model) is performed by the F-test (quite in analogy to the LR test for discrete-choice models, only the test statistic is different, there):

- H_0: The reduced model describes the observation equally well as the full model.
- The test statistic is Fisher-F distributed provided H_0 is true. The two parameters of the Fisher-F distribution are
  - the numerator degrees of freedom df=3 (the full model has 5 parameters and the restrained one 2),
  - the denominator df=5 ($n = 10$ data points minus 5 parameters of the full model)
- The rejection region is above a critical value, namely the $(1 - \alpha)$ quantile.
- Since physics requires the five-parameter model to be reduced to the two-parameter model, the expected outcome is that H_0 is *not* rejected.

Remark 1: Instead of the F-test, you could also argue with overall model quality metrics for model selection such as $\bar{R}^2$ (adjusted R^2), AIC, or BIC. This gives full marks if it is stated that the selected model is the one with higher $\bar{R}^2$, or lower AIC or BIC (incidentally, all these criteria would erroneously select the full model, so the F-test which is more in favour of Occam's Razor, as shown below, is to be preferred)

Remark 2: (not required) The reduced model takes into account all physical laws while it does not ignore any nontrivial factors of the original model. As expected, the null hypothesis ("the reduced model describes the observation equally well as the full model") cannot be rejected: the data value $f_{\text{data}} = 2.19$ while the rejection region for an α-error of 0.05 is given by $f > f_{0.95}^{(3,5)} = 5.41$.

Watch out for the following common errors!

(c) The system matrix elements x_{ij} include all *prefactors* of the parameters according to $y_i = \sum_j x_{ij}\beta_j + \epsilon_i$ So, for the first line (first cyclist/segment), the prefactor of $\beta_0 = 1$ (as for all segments), that of β_1 is $v_1^3 = 8^3$ and so on.

(d) Look at the problem statement: The expected parameter *values*, not only *signs* should be given.

(h) As the name implies, the LR (Likelihood-ratio) test refers to the maximum-likelihood calibration. With OLS estimation as we have used for regression, the equivalent method is the F-test, so this would be the right answer.