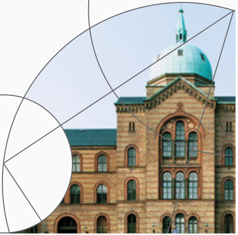
The Last Goodbye

Advanced Microeconometrics

Anders Munk-Nielsen

2021



Plan

1. Example: Last Year

2. Advice

3. From Absalon

Plan for lectures: Helicopter

Part I: Linear methods. ✓

Part II: High-dimensional methods. ✓

Part III: M-estimation, theory ✓

Part IV: M-estimation, examples ✓

Where are we in the course?

Part	Topic	Parameterization non-linear	Estimation non-linear	Dimension $\dim(x)$	Numerical optimization	M-estimation (Part III)	Outcome (y_i)	Panel (c_i)
I	OLS	$\div$	$\div$	low	$\div$	✓	$\mathbb{R}$	✓
II	LASSO	$\div$	✓	high	✓	$\div$	$\mathbb{R}$	$\div$
IV	Probit	✓	✓	low	✓	✓	$\{0,1\}$	$\div$
	Logit	✓	✓	low	✓	✓	$\{1,2,...,J\}$	$\div$
	Tobit	✓	✓	low	✓	✓	$[0;\infty)$	$\div$
	Simulated Likelihood	✓	✓	low	✓	✓	Any	✓
	Sample selection	✓	✓	low	✓	✓	$\mathbb{R}$ and $\{0,1\}$	$\div$
	Quantile Regression	$\div$	✓	(low)	✓	✓	$\mathbb{R}$	$\div$
	~~Non-parametric~~	~~✓~~	~~(✓)~~	$\cancel{\infty}$	$\cancel{\div}$	$\cancel{\div}$	$\cancel{\mathbb{R}}$	$\cancel{\div}$

Course Contents

- **Parameterization – linear vs. non-linear:**
- **Estimation – non-linear?** Whether or not we use `minimize`
- **High-dimensional?** Penalization, LASSO
- **Numerical optimization:** Gradient-based or gradient-free? Flat gradients?
- **M-estimation:**
  - **Thm. 12.1 (consistency):** Typically argued from likelihood or conditional mean assumption.
  - **Thm. 12.3 (normality):** "Smooth" criterion.
- **Outcome:** Continuous, binary, censored, discrete, ... $\Rightarrow$ dictates the model.
- **Panel aspects:** FE vs. RE, efficiency, consistency, (strict) exogeneity, ...

Exam

Part I: Re-submission

- **One project chosen at random** from the three
- The two other projects **will not count** towards the final grade.

Part II: New project

- **New data:** One or two .csv dataset(s)
- **Code:** `getting_started.ipynb` (reads in data)
- **New vs. old:** The model will be similar to *one* of the models you have seen...
  - ... but will be an extension / variation.
  - Start your code from the familiar!

Outline

Model

Model

$$
\begin{aligned}
y_i^* &= \mathbf{x}_i\beta + u_i, \quad u_i|\mathbf{x}_i \sim \mathrm{N}(0, \sigma_i^2), \\
\sigma_i &= \gamma h(\mathbf{x}_i\beta), \quad \gamma > 0, \\
y_i &= \max(y_i^*, 0).
\end{aligned}
$$

- **Cross-sectional** dataset and model
- Many **intuitive** questions and analyzing questions.
- **Part I** dealt with the **pure theory**, which implies:
  - Part II has no questions in this regard,
  - ... and it is a "brief" for a 48h exam.
- **Overall:** An illustration of why Tobit fails under strong heteroscedasticity.

Questions

1. Estimate Tobit and discuss,
2. Show quantiles,
3. Derive loglikelihood function for the model,
4. Estimate parameters for two $h(\cdot)$: $h(z) = \exp(z)$ and $h(z) = \exp(-z)$, and pick the best,
5. Based on 4, exlain findings in 1 and 2.

Questions

1. Estimate Tobit and discuss,
   A gentle start.
2. Show quantiles,
3. Derive loglikelihood function for the model,
4. Estimate parameters for two $h(\cdot)$: $h(z) = \exp(z)$ and $h(z) = \exp(-z)$, and pick the best,
5. Based on 4, exlain findings in 1 and 2.

Questions

1. Estimate Tobit and discuss,
2. Show quantiles,
3. Derive loglikelihood function for the model,
4. Estimate parameters for two $h(\cdot)$: $h(z) = \exp(z)$ and $h(z) = \exp(-z)$, and pick the best,
5. Based on 4, exlain findings in 1 and 2.

Questions

1. Estimate Tobit and discuss,
2. Quantiles,
   Introduces another estimation method,
   ... and quantiles reveal that the *mean* and *P90* are moving in *opposite directions*
   (consistent with strong heteroscedasticity)
3. Derive loglikelihood function for the model,
4. Estimate parameters for two $h(\cdot)$: $h(z) = \exp(z)$ and $h(z) = \exp(-z)$, and pick the best,
5. Based on 4, exlain findings in 1 and 2.

Questions

1. Estimate Tobit and discuss,
2. Show quantiles,
3. Derive loglikelihood function for the model,
   Standard derivation that follows the book/slides with a tiny difference.
4. Estimate parameters for two $h(\cdot)$: $h(z) = \exp(z)$ and $h(z) = \exp(-z)$, and pick the best,
5. Based on 4, exlain findings in 1 and 2.

Questions

1. Estimate Tobit and discuss,
2. Show quantiles,
3. Derive loglikelihood function for the model,
4. Estimate parameters for two $h(\cdot)$: $h(z) = \exp(z)$ and $h(z) = \exp(-z)$, and pick the best,
   Likelihood criterion (the two models are **nested**)
   Intuitive criterion (only $h(z) = \exp(-z)$ is consistent with findings in 2)
5. Based on 4, exlain findings in 1 and 2.

Questions

1. Estimate Tobit and discuss,
2. Show quantiles,
3. Derive loglikelihood function for the model,
4. Estimate parameters for two $h(\cdot)$: $h(z) = \exp(z)$ and $h(z) = \exp(-z)$, and pick the best,
5. Based on 4, exlain findings in 1 and 2.
   Low x_{i2}: σ_i is large, $\mathbf{x}_i\beta < 0$, so $y_i > 0$ occur due to strong heteroscedasticity (i.e. large draws of u_i).
   High x_{i2}: σ_i is small, but $\mathbf{x}_i\beta > 0$, so $y_i > 0$ occur because $\mathbf{x}_i\beta$ becomes positive (and y_i becomes less dispersed around $\mathbf{x}_i\beta$)
   P90: driven by σ_i.
   Heteroscedasticity causes *inconsistency* of Tobit!
   Even OLS gets the wrong sign (due to the large outliers for low x_{i2}).

Outline

General Advice

- Read the full problem set!
- Write something for *each* question!
  - E.g. write what you were *intending* to do.
- Use the definitions and names from the course wherever you can.
  - E.g. consistency, inference, strict exogeneity, Wald test, ...
- For each topic, find the *key concepts*
  - Search for good places to enter them.

add these to the problem sets?

How much to write

- **General rule:** a fellow student must be able to reproduce your results based on your pdf alone.
- **Assumptions:** Named and stated.
- **Derivations:**
  - **From slides or Wooldridge:** entirely satisfactory to just give a reference (page) and state the result.
  - **Otherwise:** we must at least be able to see that the correct approach is taken
    - I.e. make sure the *starting point* in particular is clear: is it a FOC, are you isolating something, etc.

Common mistakes last year

- **Question 1 (2020 Exam):** State the assumptions for consistent estimation.
  - **Wrong answer:** θ_o must be the unique minimizer of Q.
    (this is just the definition of identification)
  - **Correct answer:** Consistency rests on the model specification, including the error term distribution, since then $\ell_i(\theta)$ is the correct likelihood function, which guarantees consistency of the maximizer.
- **Sufficient vs. necessary conditions:** What happens to the maximizer under model misspecification
  - **Wrong**: if the error term is not homoskedastic, then $\hat{\theta}$ is inconosistent.
  - **Correct:** If the error term is homoskedastic, then $\hat{\theta}$ is consistent. If it is heteroskedastic, we don't know; it could be.
- **Do** use course language
  - **Incomplete:** The estimate could be inconsistent if there is learning.
  - **Better:** Learning about severity can generate sample selection problems, where unobserved needs is an omitted variable correlated with time.

Coding

- Re-start / re-run your notebook often
  - Ensure old variables are not lying around
- Tables: Consider using `dataframe.to_latex()`
- Use functions for repeated tasks
  - If you do something twice, put it in a function
  - ... avoid mistakes where a variable sticks around
- Code library
  - Make sure you can use all models on a new (appropriate) dataset.

Coding: M-estimators

If the exam is an M-estimator

- Copy the nearest `model.py` file
- Make the relevant changes
- If you want a check: write `sim_data` and check that you can estimate back parameters.

Troubleshooting

- Always start by evaluating `model.q(theta0,y,x)`.
- `... lack of precision...`: Probably an error in your code
- `... not defined at initial piont`: Always call your criterion first.
- **Starting values**: If you have problems, think of several options.
  - E.g. $\beta^0 = \hat{\beta}^{\text{OLS}}$ or $\beta^0 = \mathbf{0}_{K\times 1}$.
  - For dispersions (e.g. σ in Tobit): tricky to choose. `Try y.std()`
    - $\sigma \to 0$ leads to zero likelihoods,
    - $\sigma \to \infty$ leads to a "flat" likelihood function.

  Consider holding σ fixed and estimating β to get started.
- If **standard errors** cause trouble (e.g. `linalg` error in inverting `A` or `B`):
  - Split `estimation.estimate()` into two functions (maybe with `try ... except`)
    - First optimize; then compute standard errors.

Standard errors

- **Numerical precision** can impact estimates of A and B
- **For ML:** We know that $A = B$ should hold...
  - Large deviaitons should be noted as "strange" and "unexpected."
  - If one leads to insane standard errors, this is a good reason to pick another.
    (one reason could be numerical problems, though it is typically a sign of a coding error)
- **For non-ML:** Only "Sandwich" can be used...
  - ... hence, it is generally considered mor robust.
- **For functions of parameters:** Either *delta method* or *bootstrap*.
- **Delta method:** Write out the formulas explicitly. Don't just write "we use the delta method".
- **Bootstrap**: remember to evaluate function at the same $\mathbf{x}^0$ across bootstrap replications.
  - If computational constraints apply: better to report with $R = 10$ replications than none at all.

Specific topics

- **Simulated Maximum Likelihood**: intended to follow `sml.py` and `sml.ipynb` very closely.
  - Try extending probit/tobit/etc.
- **Quantile Regression**: standard errors not covered
  - Bootstrap can be used.
- **Bootstrap**: If computational constraints $\Rightarrow$ take 10 iterations and describe your method
  (your code will show that you have done it correctly)

Attached code

- **Requirement:** Attach code
- Serves as the **intermediate steps** (same as derivations)
  - If your estimates are wrong, we can verify that the correct method was used.
- **Make sure** that your entire code runs before uploading.
  (e.g. use the "restart kernel and run all" button.)
  (e.g. include all attachehd files in the same directory and not in obscure paths specific to your computer)
  - Do not submit a notebook full of error messages in its saved form.

Prettifying

- **Figures:** Double check the axis labels!
  - `plt.xlabel('hi')` or `ax.set_xlabel('hi')`
- **Tables**: Put different versions in columns for quick comparison
  - Report N and loglikelihood/R^2 at the bottom

Outline

Formalities

- **Group members:** as exam numbers
  - Separate responsibilities by subsection. Study-wide requirement!
- **Character count** on the front page
  - E.g. use `https://charcounter.com/` copied from the pdf.
  - Lyx: doesn't count math $\Rightarrow$ lower bound.
- **No page / character constraint** on the new part of the exam!
  - ... but be brief!
- **Citing other groups:** E.g. *"In the following, we rely on code developed by Group 13 in the course."*
  - *Always cite sources!!*
  - Exception: core course code material (e.g. estimation.py)

That's all