NEW ECONOMIC SCHOOL
Master of Arts in Economics

Econometrics IV
Module 6, 2017–2018

Professor: Stanislav Khrapov
New Economic School
skhrapov@nes.ru

Course information

Course Website: my.nes.ru

Instructor's Office Hours: write me an email

Class Time: TBA

Auditory: TBA

TA: TBA

Course description

The goal of this course is primarily to give to any researcher in quantitative economics a very firm understanding of the most widely based estimation techniques for non-linear models, namely Maximum Likelihood (MLE) and Generalized Method of Moments (GMM). The material assumes that a student has a working knowledge of inference tools that were taught in Econometrics III. The data we have in mind is still a large cross-section or a long time series. We will also consider how to deal with a panel when a cross section of subjects is observed over time. After completing the class students should be able to work independently with new non-linear econometric models, analyze the properties of estimators and statistics, and finally interpret any economic theory in terms of statistical hypothesis which is testable using econometric methodology.

Course requirements, grading, and attendance policies

Course prerequisites include Probability Theory, and Econometrics I, II, III.

The grading of student's performance is based on

- 5 weekly homework assignments that account for 20% of the final grade.

Homework assignments will contain analytical problems, as well as practical computational exercises to be written in Python[1]. They will be posted on my.nes.ru. Your individual work must be uploaded on my.nes.ru before the specified deadline. Late submissions have zero weight. Sloppy formatting is heavily penalized. The homework should contain at most two separate files:

  - Pdf with the text. No jpg, doc, txt, etc. are allowed. If your work is handwritten, scan it with one of the available copiers in the school.
  - Working bug-free code. If the code uses external data, its source should be mentioned. All data transformations should be part of your code.

- Final exam will be two-sided A4 format and account for 80% of the final grade.

  Depending on the difficulty of the exam, there may be a minimum passing threshold based on the exam grade regardless of the homework grade.

- The format of the make-up exam is identical to the final exam.

Course materials

Required textbooks and materials

- **[W]** Wooldridge, J. M. (2010) *Econometric Analysis of Cross Section and Panel Data*. 2nd ed. The MIT Press
- **[HA]** Hall, A. R. (2005) *Generalized Method of Moments*. Oxford University Press
- **[C]** Cameron, A. C. & Trivedi, P. K. (2005) *Microeconometrics. Methods and Applications*. Cambridge University Press.

Additional materials

- **[HB]** Hansen, B. E. (2016) *Econometrics*. Unpublished manuscript, University of Wisconsin http://www.ssc.wisc.edu/~bhansen/econometrics/Econometrics.pdf
- **[HF]** Hayashi, F. (2000) *Econometrics*. Princeton University Press
- Sheppard, K. (2014). Python for Econometrics. http://www.kevinsheppard.com/Python_for_Econometrics

[1]For local installation it is recommended to download Anaconda distribution available at https://store.continuum.io/cshop/anaconda.

Academic integrity policy

Cheating, plagiarism, and any other violations of academic ethics at NES are not tolerated.

Course contents

The following is a tentative list of topics which may change depending on the pace of the lectures.

1 Maximum Likelihood [W : 13], [HF : 7,8], [C : 5]

- Numerical optimization
- Likelihood function and likelihood principle
- Conditional, joint and marginal ML
- Consistency and asymptotic normality
- Asymptotic efficiency, Cramer-Rao inequality
- Hypothesis testing: Wald, Likelihood Ratio, Lagrange Multiplier
- MLE with time series data
- Quasi-MLE
- Examples and applications of MLE

2 Method of Moments [HB : 13,15], [W : 14], [HF : 3,7], [HA : 1-3], [C : 6]

- Endogeneity, instrumental variables, 2SLS
- Moment restrictions and moment functions
- Exact identification and overidentification
- Consistency and asymptotic normality
- Efficient GMM
- Test for overidentifying restrictions
- Hypothesis testing: Wald, Distance Difference, Lagrange Multiplier

- Hausman specification test
- GMM with time series data
- Examples and applications of GMM

3 Panel Data [HB : 20], [W : 10], [HF : 5], [C : V]

- Error component models (ECM)
- OLS, GLS, LSDV, Within and Between estimators in one-way ECM
- Dynamic panel
- Panel binary choice models: random effects probit and fixed effects logit

Sample tasks for course evaluation

Final exam 2012

True of False?

1. **(5 points)** For T large, N fixed and endogenous regressors in a standard ECM model, we have to eliminate individual effects to estimate β consistently.
2. **(5 points)** Under standard regularity assumptions GMM estimator in a non-linear model could be more efficient than MLE.
3. **(5 points)** Any additional overidentifying moment restrictions in GMM framework are only useful in testing model misspecification.

Big Problem

Suppose we want to model the publishing activity of NES professors over the last 10 years. Since the number of papers published in a year is an integer, the first choice for the model may be the Poisson distribution. So, denote a number of papers published by a professor $i = 1, \ldots, N$ in a year $t = 1, \ldots, 10$ as M_{it}. We assume that

$$\mathbb{P}\left[M_{it} = m \middle| X_i\right] = \frac{\exp\left(-\lambda_{it}\right)\lambda_{it}^m}{m!}, \quad \lambda_{it} = c_i \exp\left(X_{it}\beta\right).$$

Here λ_{it} is a parameter of Poisson distribution, c_i is an individual effect, X_{it} are K-vectors of (non-negative) explanatory variables, X_i is a $T \times K$ matrix, and β is a K-vector of unknown parameters. Also assume that $M_{i1}, \ldots, M_{iT}$ are mutually independent conditionally on X_i and c_i.

First, assume that $c_i = 1$.

1. Estimate the model using CMLE.

   (a) **(5 points)** Write down the log-likelihood function.

   (b) **(5 points)** Derive the first order conditions.

   (c) **(5 points)** Calculate the asymptotic variance of the estimator.

2. **(10 points)** How would you test the hypothesis that publishing a paper is a pure luck? Write down the test statistic, its asymptotic distribution, and your decision given all of the above.

3. **(5 points)** You are not sure in correct specification of the distribution above, but with no better alternative you still proceed as if the data came from Poisson distribution. How would it change you testing strategy?

4. **(10 points)** Propose a test for correct specification of the model.

Now suppose that

$$c_i | X_i \sim \text{Exp}\left(\psi + \overline{X}_i \delta\right),$$

where ψ and δ are unknown parameters, and $\overline{X}_i = T^{-1} \sum_{t=1}^{T} X_{it}$, and $\text{Exp}(\lambda)$ is an exponential distribution with density $f(u, \lambda) = \lambda e^{-\lambda u}$, $u \geq 0$.

5. **(10 points)** Derive the appropriate log-likelihood function (simplifying it as you can) or show how to compute it numerically.

Next, we want to reduce the risk of misspecification as much as possible. In particular, we only assume that

$$\mathbb{E}\left[M_{it} | c_i, X_{is}, s \leq t\right] = c_i \exp\left(X_{it}\beta\right),$$

and allow for unspecified dependence between X_i and c_i.

6. Estimate the model using GMM

   (a) **(10 points)** Write down conditional moment restrictions that do not include unobservable effect c_i.

   (b) **(5 points)** Write down unconditional moment restrictions using the appropriate set of instruments. How many overidentifying restrictions do you have?

(c) **(10 points)** Estimate β as efficiently as possible given a finite sample of the data and assuming that X_{it} contains $M_{i,t-1}$. Be specific of all the steps you take.

7. **(10 points)** Explain, how would you test whether each explanatory variable in the vector X_{it} has any effect on publishing activity.