INSEAD

The Business School for the World®

PhD Programme
March/April 2016

Econometrics B for PhDs

Professor: Amine Ouazad
amine.ouazad@insead.edu

Website: ***http://www.ouazad.com/*** *"Teaching Tab" "Econometrics"*

Course Description

Econometrics A is devoted to the estimation of the causal effect of treatments using linear econometrics, i.e. ordinary least squares and instrumental variable estimation. The econometrics we used in Econometrics A were largely "model-free" in the sense that they do not relied on specific assumptions about agents' behavior. In particular we did not try to identify and estimate deep parameters such as preferences for product attributes, risk aversion in lottery choices, or the exact functional form for demand and supply curves.

This second part of the econometrics course will focus on structural econometrics, i.e. econometrics with an underlying economic model. Two main estimation frameworks are employed: the maximum likelihood framework, introduced by Ronald Fisher and widely employed in top publications; and the generalized method of moments, developed by lars Peter Hansen, Nobel Prize 2013 in economics.

We start with maximum likelihood (ML). Maximum likelihood, unlike OLS, assumes specific distributions for all variables (e.g. normal, poisson, logistic). ML allows us to understand the widely-used logit, probit, and multinomial logit models for binary choices. The estimation procedures are analyzed. General ML coding in Stata is introduced. Then applications to censored, truncated models, as well as the sample selection and the treatment selection models (Heckman I and Heckman II) are introduced.

The second estimation framework is the GMM estimation framework. It is shown that all estimators (including ML) introduced in Econometrics A and B fit within the GMM framework. GMM is applied to the case of dynamic panel data.

Materials

The following textbooks will be used. Mandatory textbooks are preceded by a *.

- *William H. Greene's *Econometric Analysis*.
- *A. Colin Cameron, Pravin K. Trivedi, *Microeconometrics Using Stata*, Stata Press, 2009.
- *J. Angrist and S. Pischke, *Mostly Harmless Econometrics: An Empiricist's Companion,* Princeton University Press, 2009.

- A. Colin Cameron and P.K. Trivedi's Microeconometrics, Methods and Applications, at Cambridge University Press.
- Yudi Pawitan, In All Likelihood, Oxford University Press, 2013.

Everything else you need is either handed out in class or posted on the course website at http://www.ouazad.com/PhD-Econometrics/.

Final examination

A 3-hour written examination at the end of econometrics B will be conducted. You will be asked questions of one of three kinds:

1. Comment of the analysis/results of the tables of a paper.
2. Econometric exercise of the same type as in W.H. Greene.
3. Questions on the implementation of a regression analysis in Stata.

Past exams are available either through me (amine.ouazad@insead.edu) or on the website.

Software

We will use Stata 13 or above in the course for ML estimation. Stata has the advantage of providing the right statistics for a paper – thus it is a great tool to learn parts of econometrics on the go.

We will use R or Matlab for GMM estimation. Stata does not provide simple ways of coding general GMM estimation, and the industry standard for GMM estimation is Matlab.

Frequently Asked Questions

What are the course's prerequisites?

- The course will assume that econometrics A has been undertaken. The maximum likelihood and GMM sessions will use: (i) convex optimization (Hessian, Newton Raphson algorithm) (ii) partial derivatives, multivariate analysis, differentials (iii) linear algebra (iv) probability and statistics. Knowledge of a flexible programming language will help (Matlab or R) but is not required. If needed, both the ML and the GMM parts can be understood using Stata.

Is there an econometrics project for Econometrics B?

- There is one econometrics project for both econometrics A and B. The econometrics project is part of econometrics A's grade. As part of my service to the broader INSEAD community, I am available for individual advice on your empirical research.

Session 1: The Maximum Likelihood Framework

- **Key concepts:**
  - The likelihood function.
  - Observational equivalence, identification in the ML framework.
  - The maximum likelihood estimator: consistency and asymptotic normality.
  - The Cramer-Rao lower bound.
  - Hypothesis testing in the Likelihood framework: Wald, Lagrange, Likelihood Ratio.
  - Numerical optimization for ML estimation: the Newton-Raphson algorithm.

- **Chapters**
  - *Econometric Analysis*, Chapter 17.
  - *Cameron and Trivedi,* Chapter 5.
  - *Yudi Pawitan,* In All Likelihood.

Likelihood is a *parametric* framework. OLS was a semi-parametric framework. What that implies for ML estimation is that we need to start by assuming a specific distribution for our observations. The likelihood function, as its name indicates, measures how likely the data are given a vector of parameters to estimate. The *likelihood principle* is that the true parameter vector maximizes the true likelihood of the observations. Thus naturally an empirical analogue of the true parameter vector is the maximum likelihood estimator, which maximizes the measured likelihood of the observations. This session shows that the maximum likelihood is consistent and asymptotically normal under general regularity conditions. We provide some counter examples to convergence of the maximum likelihood estimator (see Jean Le Cam for more counterexamples). The maximum likelihood estimator is show to be efficient, in that its variance achieves the Cramer Rao lower bound.

In practice, in ML as in semi-parametric OLS, we would like to test whether specific constraints are satisfied by the true parameter vector. We thus introduce three types of tests in maximum likelihood: the Wald test, the Lagrange test, and the Likelihood ratio test.

We implement maximum likelihood for the specifications seen in econometrics A: ordinary least squares, panel data, and instrumental variables. The algorithms are coded in Stata.

Session 2: Binary Discrete Choice Models, Logit and Probit

- **Key concepts:**
  - Utility maximization with discrete choice: smoking, migration.
  - The logistic and probit distributions: pdf and cdf.
  - Logit and probit likelihood functions.
  - Single-peakedness of the likelihood function with probit and logit: the Hessian.
  - Interpretation of logit and probit output: the marginal effects.
  - Hypothesis testing with logit and probit.

- **Chapters**
  - *Econometric Analysis*. Chapter 21.
  - *Microeconometrics Using Stata*. Chapter 14.

- **Papers**
  - F.Lafontaine, K.Shaw, "Franchising Growth and Franchisor Entry and Exit in the U.S. Market: Myth and Reality," Journal of Business Venturing, Vol.13, pp.95-112, 1998.

While we showed in Econometrics A that a wide range of data sets with binary outcomes can be estimated using OLS, there are a number of issues with the so-called linear probability model. First, the linear probability does not have a structural interpretation: it does not reveal individuals' preferences for either choice. Second, the predictions of the linear probability model can fall outside the [0,1] segment, which is incorrect as predictions should be probabilities. Third, the errors are typically heteroskedastic when the outcome variable is binary.

We thus introduce the logit and the probit models, prove the identification of the models, and present the estimation technique for such models. Output is interpreted and the major caveats are explained: coefficients cannot be directly interpreted, only their ratios. We then show that marginal effects are identified. Finally, we show how to test hypothesis in the logit and probit frameworks. We compare the relative properties of logit and probit.

Session 3: The Heckman Sample Selection and Treatment Selection Models

- **Key concepts**:
  - Sample selection: the case of women's wage regressions and decision to work.
  - The two-equation sample selection model.
  - The maximum likelihood estimator.
  - The control function estimator.
  - Treatment selection: first step with a binary treatment.

- **Chapters**
  - *Econometric Analysis*. Chapter 24.
  - *Cameron and Trivedi*. Chapter 16.

- **Readings**
    - J.M. Shaver, “Accounting for Endogeneity When Assessing Strategy Performance: Does Entry Mode Choice Affect FDI Survival?,” Management Science, Vol. 44, No. 4, April 1998.

Two separate estimators fall under the name of “Heckman model”. The first Heckman model deals with sample selection: when estimating the wage returns of education, individuals out of the labor force are not part of the estimation, and thus the true impact of education on wages may not be consistently estimated. The selection “Heckman model” deals with the binary nature of treatments. The first-stage typically has a binary outcome variable. Such framework is described in Shaver (1998) and is useful when understanding instrumental variable estimation. Shaver (1998) has a substantial flaw that we will discover in class.

Session 4: Truncation and Censoring

- **Key concepts**
    - Tobit.
    - Truncated outcome variable in OLS.
    - Censored outcome variable in OLS.

- **Chapters**
    - *Econometric Analysis*. Chapter 22.
    - *Cameron and Trivedi,* Chapter 16.

A researcher would like to estimate the demand for a stadium’s tickets, but the demand is upper censored: the stadium has a maximum capacity that is often reached. How will the researcher consistently estimate the parameters of the demand function? A second example is as follows: most census data have the unfortunate feature of top coding income. Individuals with an annual income at 1 million dollars and above will see their income reported as 999,999 dollars. Such top coding biases OLS estimates.

This session introduces the truncated and the censored models, and shows how to carefully interpret results of these models.

Session 5: The Generalized Method of Moments

- **Key concepts**
    - Moment conditions in OLS, IV, ML, and consumption models.
    - The moment estimator.
    - Asymptotic normality of the empirical moments.
    - The Generalized Method of Moments estimator for overidentified models.
    - Two-step GMM.
    - Variance of the GMM estimator.
    - Hypothesis testing in the GMM estimator.

- Numerical computation issues in GMM.

- **Reading**
  - Hansen, Lars Peter. "Large sample properties of generalized method of moments estimators." *Econometrica: Journal of the Econometric Society*(1982): 1029-1054.
  - *Econometric Analysis,* Chapter 18.
  - *Microeconometrics*, Chapter 6.

The Generalized Method of Moments framework encompasses all estimators seen in Econometrics A and B. We show how to write your own GMM estimator. We prove consistency and asymptotic normality. The GMM estimator is used in the structural econometrics of differentiated products models. We develop algorithm for the estimation of GMM models, and implement them in Matlab or R.

Session 6: Dynamic Panel Data

- **Key concepts**
  - The incidental parameter problem.
  - The Hsiao method.
  - The Arellano Bond GMM model.

- **Readings**
  - Nickell, Stephen. "Biases in dynamic models with fixed effects."*Econometrica: Journal of the Econometric Society* (1981): 1417-1426.
  - Stephen Bond, Dynamic Panel Data Models: A Guide to Micro Data Methods and Practice, Cemmap, 2002.
  - Arellano, Manuel, and Stephen Bond. "Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations."*The review of economic studies* 58.2 (1991): 277-297.

Linear regression models with a fixed effect and with a lagged dependent variable included in the covariates suffer from a bias described in Nickell (1981). We will show that the bias is severe for small fixed T and for large N. We develop in class the Arellano Bond GMM estimator, which is unbiased under a simple orthogonality condition on the error term. We then estimate such dynamic panel data model in Stata and R.