Version: Thursday 4th January, 2024 (h17:14)

Università degli Studi di Napoli "Federico II"
Master in Economics and Finance (MEF)

ECONOMETRICS

Pietro Coretto

SYLLABUS

Prerequisites: basic notion of matrix algebra, probability theory for one-dimensional random variables, statistical inference (point estimation, confidence intervals, hypothesis testing).

Course structure: theory and methodology (28 hours), computing labs (8 hours).

Curse materials: slide sets given before lectures, exercises (not graded), lab session (computing with R language), project work (graded).

Evaluation

- **Option A:** available only for students taking the final exam on March 2022 (first examination date after the end of the course). In this case 40% of the final grade is given based on two take-home projects involving computer programming and real data analysis. The remaining 60% of the final grade is given based on a written examination including both theoretical and practical problems. In case of failure, the take-home project is not considered for the following examinations (see option B below).
- **Option B:** the student takes the final exam after March 2021. 100% of the final grade is given based on a written examination including both theoretical and practical problems.

Program: theory and methodology

Introduction. What is econometrics? – Dependence and causality – Predictive and explanatory modeling – Joint correlation and linear models – Twists in modern big-data – Collecting data and types of data sets

Background tools. Vector and metric spaces – Inner product spaces, linear dependence and orthogonality – Matrix algebra and systems of linear equations – Multivariate random variables, multivariate distributions and moments – Covariance and correlation matrices – Inner product/Orthogonality for random variables – Linear maps and quadratic forms of random variables – Information, conditional expectations, and stochastic conditioning

Linear least squares fitting. Data sets and their matrix representation – Linear approximation problem – Least squares solution – Geometric properties of the least squares projection

Regression models and the CLRM. Regression analysis – Optimal predictor under mean squared loss – Linear conditional mean regression models – Using linear regression models – Marginal/partial effects and elasticities – Dummy variables – Interaction terms – Log-linear models – Semilog-linear models – Classical linear regression model (CLRM)

Inference based on the CLRM. OLS estimator of the coefficients – Estimation of the error variance – Variance decomposition and measures of fit – Finite-sample properties of the estimators – Estimation of standard errors – Gaussian errors – Hypothesis tests for individual coefficients – Confidence intervals – Linear restrictions tests – Prediction – Connections to maximum likelihood – Violations of CLRM assumptions and the GLS estimator

Asymptotic approximations. Convergence of random sequences – Limit theorems for IID sequences – Dependent sequences and times-series data – Stationarity and ergodicity – Martingale difference sequences (MDS), and linear processes – Limit theorems for MDS sequences

Linear model with predetermined regressors. The linear model with predetermined regressors – Large sample properties of the OLS estimator – Hypothesis testing for the individual coefficients (robust t-statistic) – Linear and nonlinear restriction tests (Wald-type tests) – Testing for conditional homoskedasticity – Testing for serial correlation

Specification analysis, indentification and endogeneity. Model selection: predictive-vs-explanatory paradigm – Inclusion of irrelevant variables – Omission of relevant variables – Unobserved heterogeneity – General notion of identification of probability models – Identification and estimation – Identification of simple linear models – Endogeneity bias – Instrumental variables – Identification with instrumental variables – Instrumental variables, proxy variables and control variables

Linear model with both endogenous and exogenous regressors. Specification of linear models via moment conditions – Linear model with endogenous and exogenous regressors – The GMM estimator for just-identified models – IV estimator as a special case of the GMM estimator – Large sample properties of the GMM/IV estimator – Inference with the GMM/IV estimator

Program: computing and applications

Introduction to the R computing language. R language, CRAN website, libraries and packages, IDEs – Data types (integer, floating points, logical, characters, etc) – Data structures: vectors, arrays, matrices, lists and data.frames – Indexing data structures, slicing and data subsetting – Element-wise operations and matrix algebra – Storing R data objects in RData files – Main techniques for data import exports

Statistics and graphics. Sample statistics – Probability distributions – Sampling with/without replacement – XY-plots and the basics of the graphic device – Data visualization (boxplots, ecdf, density, pairs-plot, line plots, etc.)

Elements of computer programming. Statements: if-else, for, while – Loops vs array computations, and apply-type functions – Programming with functions – Writing an R program

Least squares and inference in linear models. Least squares using matrix algebra – lm() function – Estimation of the asymptotic variance – Linear and nonlinear restrictions tests – Confidence intervals

GMM/IV estimation and testing. Example of endogenous variables using real data – Implementing the IV estimator and the related testing

Software

Lab classes are based on the use of the R programming language. The *R base package* is freely available through the *Comprehensive R Archive Network* (CRAN) for most OS

- **Linux/Unix/nix:** follow instructions (link) for specific distros
- **macOS:** obtain and install the R base package binary from CRAN (link). Carefully read the section "*latest release*" before taking any action.
- **MS Window:** obtain and install the R base package binary from CRAN (link)

Moreover, the use of an appropriate IDE can improve your productivity. One of the most popular choice is *RStudio*. You can download the latest free version from the RStudio website (see "Installers for Supported Platforms" section).

References

Suggested books

- Hayashi, F. (2000), "*Econometrics*", Princeton University Press
- Greene, W. H. (2017), " *Econometric Analysis*", Pearson
- Verbeek, M. (2017), "*A Guide to Modern Econometrics*", Wiley

The theoretical and methodological part of the course essentially covers chapters 1–4 of Hayashi (2000). Any recent edition of these books can fit the course. For R language there is a number of free manuals. The following list is suggested

- For this course is sufficient to read (if needed) the official CRAN's introductory manual.
- Zeileis, A. and and C. Kleiber (2008), "*Applied Econometrics with R*", Springer.
- Wickham, H. (2017). "*R for Data Science: Import, Tidy, Transform, Visualize, and Model Data*". O'Reilly Media, Inc.
- Wickham, H. (2019). "*Advanced R*", CRC Press.