TRAINING COURSE | ONLINE

ANALYSING MICRO DATA IN STATA

31 August - 4 September 2020

TStat's *Analysing Micro Data in Stata* course offers participants a comprehensive introduction to the principle methodologies used in the analysis of micro data. Micro data, data which contains information at the level of a specific unit (such as individuals, firms or entities), has by its very nature become an increasingly important source of information offering researchers and policy makers an effective *tool* with which to obtain a more in-depth understanding of an array of political, socio-economic and public health phenomena. As such the collection and subsequent analysis of micro data over recent years has proved to be the key to policy formulation, the targeting of interventions and the subsequent monitoring and measurement of the impact of such interventions and policies. Whilst these techniques have been traditionally more applied in the field of economics, the increasing availability of micro data has over recent years resulted in a steady increase in the analysis of micro data by researchers working in Political and Social Sciences, Biostatistics, Epidemiology and Public health.

TStat's introduction to micro data analysis course focuses from both a theoretical and applied point of view, on the following methodologies: linear models, count models, binary dependent variable models, multinomial models, Tobit and Interval Regression models, models with treatment variables and models with Sample Selection. The concluding sessions focus on: i) the Control Function approach for the estimation of non-linear models with endogenous continuous variables; and ii) the latest *Stata* commands for the estimation of *Extended Regression Models* (ERM), which implement Maximum Likelihood estimators capable of simultaneously treating issues of Sample Selection and the presence of both endogenous regressors and treatment variables.

In common with TStat's training philosophy, each individual session is composed of both a theoretical component (in which the techniques and underlying principles behind them are explained), and an applied (hands-on) segment, during which participants have the opportunity to implement the techniques using real data under the watchful eye of the course tutor. Throughout the course, theoretical sessions are reinforced by case study examples, in which the course tutor discusses and highlights potential pitfalls and the advantages of individual techniques. The intuition behind the choice and implementation of a specific technique is of the utmost importance. In this manner, the course leader is able to bridge the "often difficult" gap between abstract theoretical methodologies, and the practical issues one encounters when dealing with real data. At the end of the course, participants are expected to be able to autonomously implement the theories and methodologies discussed during the course.

COURSE CODE

D-EF13/3-OL

DATE AND LOCATION

Due to the current COVID-19 situation, the 2020 edition of this Training Course will now be offered ONLINE, on a part-time basis from the 31st August to 4th September 2020. To this end, this year's programme has been transformed into a series of module based on 5 sessions from 10.00 am to 1.30 pm Central European Summer Time (CEST).

In addition, an informal hour long evening Study Group session will take place on Tuesday the 1st and Thursday the 3rd September.

ANALYSING MICRO DATA IN STATA

TARGET AUDIENCE

Researchers and professionals working in biostatistics, economics, epidemiology, finance, psychology, social and political sciences needing to acquire the necessary statistical requisites required to independently conduct empirical analysis using micro data.

PREREQUISITES

It is assumed that course participants have at some point followed a basic course in econometrics or statistics. Previous exposure to Stata or other statistical software packages would also be an advantage.

PROGRAM

SESSION I: THE LINEAR MODEL

1. Identification, estimation and specification tests for linear models with exogenous regressors/variables: ***regress, estat imtest, estat hettest, estat bgodfrey, actest***
2. Identification, estimation and specification tests for linear models with exogenous regressors/variables: Instrumental Variables (IV), Limited Information Maximum Likelihood (LIML) and Generalized Method of Moments (GMM) estimators: ***ivregress, gmm, ivhettest, actest, estat overid, estat endogenous, estat firststage, weakivtest***

SESSION II: COUNT MODELS

1. Count Model Estimators in Stata: The Poisson Model
   - Non-Linear Least Squares and GMM Estimators, Maximum Likelihood Estimators in Stata: ***nl, gmm, poisson***
   - Models with endogenous regressors: ***gmm*** and ***ivpoisson***
2. Estimation and Specification tests in the presence of *overdispersion*: the *Generalized Negative Binomial Model*: ***nbreg, gnbreg***
3. Estimation and interpretation of marginal effects using the Stata post estimation command ***margins***

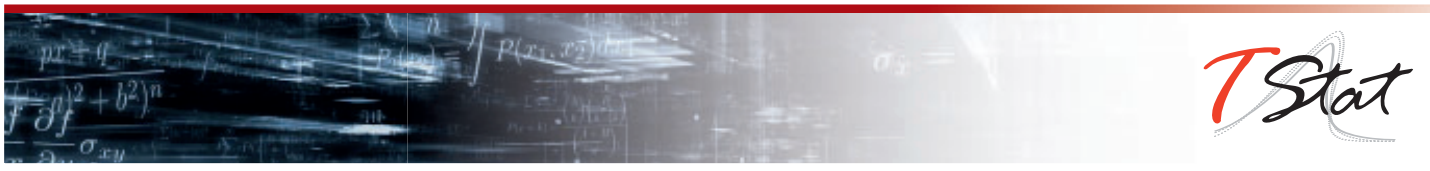
SESSION III: DISCRETE DEPENDENT VARIABLE MODELS

1. Estimating linear models with binary dependent variables – Logit, Probit and the Linear Probability Model: ***probit, logit, regress***
2. The Heteroskedastic Probit Model and tests of heteroskadicity: ***hetprobit***
3. Measures of Goodness of Fit and Specification Tests: ***tabulate, estat classification, estat gof***
4. Independent Latent Heterogeneity in Probit Models
5. Estimating marginal effects: ***margins***
6. Numerical problems with Logit and Probit

SESSION IV: PROBIT MODELS WITH ENDOGENOUS REGRESSORS

1. The *Control Function* (CF) in the presence of continuous endogenous regressors
2. Testing for exogeneity in the CF framework
3. *Bootstrap* standard error estimation in the CF approach
4. Maximum likelihood estimation in the presence of continuous endogenous regressors: ***ivprobit***
5. The multivariate recursive Probit estimator as a solution to the problem of the presence of binary endogenous regressors: ***biprobit, mvprobit, cmp***
6. Measures of Goodness of Fit: ***tabulate, estat classification, estat correlation***
7. Estimating marginal effects: ***margins***

https://www.tstattraining.eu/training/analysing-micro-data-stata-ol/



SESSION V: MULTINOMIAL MODELS

1. Ordered categorical variable models (the Ordered Probit and Ordered Logit Estimators): ***oprobit*** and ***ologit***
2. The Heteroskedastic Probit Model and tests of heteroskadicity: ***hetoprobit***
3. Models with categorical (but unordered) variables - Multinomial Logit and Multinomial Probit estimators: ***mlogit, mprobit***
4. MacFadden's Choice Model - categorical variable models with alternative specific regressors: ***cmclogit, cmcprobit***
5. Measures of Goodness of Fit and Specification Tests
6. Estimation and interpretation of marginal effects using the Stata post estimation command ***margins***

SESSION VI: THE TOBIT MODEL, INTERVAL REGRESSION E *SAMPLE SELECTION*

1. The Tobit Model - ML and Two-Step Least Squares: ***tobit, heckman***
2. The *Control Function* (CF) approach in the presence of continuous endogenous regressors, exogeneity tests and Bootstrap *standard errors*
3. The Maximum Likelihood estimator for Tobit models with endogenous regressors: ***ivtobit***
4. Interval Regression: a generalization of the Tobit Model: ***intreg***
5. Estimators for Sample Selection Models: ***heckman***
6. Estimation and interpretation of marginal effects using the Stata post estimation command ***margins***

COURSE REFERENCES

- Wooldridge, (2010) Econometric Analysis of Cross Section and Panel Data, Second Edition MIT Press
- Cameron e Trivedi, (2010) Microeconometrics Using Stata, Revised Edition StataPress
- Cameron e Trivedi, (2005) Microeconometrics: Methods and Applications, Cameron e Trivedi, Cambridge University Press

ANALYSING MICRO DATA IN STATA

REGISTRATION FEES

Full-time Students*: € 1065.00
Academic: € 1515.00
Commercial: € 2020.00

*To be eligible for student prices, participants must provide proof of their full-time student status for the current academic year.

Fees are subject to VAT (applied at the current Italian rate of 22%). Under current EU fiscal regulations, VAT will not however applied to companies, Institutions or Universities providing a valid tax registration number.

The number of participants is limited to 8. Places will be allocated on a first come, first serve basis. The course will only be confirmed when at least 5 people have enrolled.

Course fees cover: teaching materials (handouts, Stata *do files* and datasets to used during the course), a temporary licence of Stata valid for 30 days from the beginning of the course.

Individuals interested in attending this workshop must return their completed registration forms by email (training@tstat.eu) to TStat by the 10th August 2020.

CONTACTS

Monica Gianni
TStat Training | Kleebergstraße, 8
D-60322 Frankfurt am Main

TStat S.r.l. | Via Rettangolo, 12-14
I-67039 Sulmona (AQ)
T. +39 0864 210101

training@tstat.eu
www.tstattraining.eu - www.tstat.eu

Further details regarding our registration procedures, including our commercial terms and conditions, can be found at https://www.tstattraining.eu/training/analysing-micro-data-stata-ol/