

Regression Analysis of Time-Series Cross-Sections

Robert J. Franzese, Jr., Ph.D.

Professor of Political Science, University of Michigan
franzese@umich.edu <http://www.umich.edu/~franzese>

OBJECTIVES: Participants will build from a brief but thorough review of the classical and generalized linear-regression models (the CLRM and GLRM) an understanding of the typical departures from those models in time-series-cross-section contexts. From this basis, participants will learn how to specify, estimate, test, and interpret empirical models that accurately reflect and powerfully leverage cross-unit and cross-temporal variation and dynamics, and how to present the estimated effects and dynamics from these empirical TSCS analyses effectively. Participants will practice conducting such TSCS analyses in lab sessions using STATA.

ELIGIBILITY: Participants should have previous experience with the CLRM, preferably also the GLRM, and some familiarity with these models and linear-regression estimation of them (i.e., ordinary and feasible generalized least-squares: OLS and FGLS), and preferably in matrix (linear-algebra) notation. The short-course does begin with review of these subjects, which should suffice to prepare participants with previous experience only with the CLRM and OLS in scalar algebra for the rest of the material.

COURSE DESCRIPTION: Time-series-cross-section (TSCS) data harness both cross-temporal and cross-spatial variation to maximize leverage for empirical analysis. However, this powerful data structure also requires careful consideration of temporal and spatial (cross-unit) heterogeneity, temporal and spatial dynamic processes, and potentially complex stochastic-error structures. After brief review of the baseline classical and generalized regression models, this course covers specification, estimation, interpretation, and presentation of empirical models that are appropriate for TSCS data. The course discusses first the nature of pooled data and the ways that they deviate from the assumptions associated with the classic linear regression model. It then addresses a number of issues that are typically associated with TSCS data: fixed or stochastic unit-heterogeneity, complex error-structures, and temporal and spatial correlation and dynamics. The course considers a variety of methodological strategies for confronting these issues in an effective manner, such as: fixed or random-effect models and associated tests; feasible-generalized-least-squares (FGLS); consistent coefficient-estimate variance-covariance (HAC) estimators; and temporal, spatial, and spatiotemporal lag models. Throughout, the emphasis remains firmly on the specification of models that reflect theoretical and substantive understandings and on the substantive interpretation and presentation of the empirical estimates of those models (as well as on statistical testing and evaluation of those models). Participants will have hands-on practice with these sorts of analyses in the brief laboratory sessions, conducted in STATA, that conclude each day. Participants are encouraged to bring their own time-series-cross-section datasets (in STATA format) for additional exploration in these lab sessions.

DAY 1: WEDNESDAY, September 9, 2009 — Introduction & Review

08:15 – 09:00 Introductions; Advantages & Challenges of TSCS Data

09:00 – 10:15 Math Review: Matrix Algebra & Basic Calculus

10:15 – 12:00 Review: Basic Probability & Statistics; Classical (Normal) Linear-Regression Model

- BACKGROUND: Wooldridge, Jeffrey M. 2002. *Econometric Analysis of Cross Section and Panel Data*, MIT Press, Cambridge, Chs. 4-6.

12:00 – 13:00 Lunch

13:00 – 14:00 Review: Generalized (Normal) Linear-Regression Model

- BACKGROUND: Wooldridge, Jeffrey M. 2002. *Econometric Analysis of Cross Section and Panel Data*, MIT Press, Cambridge, Chs. 7-9.

14:00 – 14:45 Review: Maximum-Likelihood Estimation and *QualDep* Models

14:45 – 15:30 C&G(N)LRM and Time-Series-Cross-Section Data

- BACKGROUND: Stimson, J. 1985. “Regression in Space and Time: A Statistical Essay.” *American Journal of Political Science* 29:914-947.

OMITTED: Testing and Sensitivity Analysis

- BACKGROUND: Fox, John. 1997. *Applied Regression Analysis, Linear Models, and Related Methods*, Sage Publications, Chs. 11-13.

15:30 – 16:00 *Lab Session*: OLS & FGLS Regression; Specification Testing and Sensitivity Analysis; Hypothesis Testing and Presentation of Results.

DAY 2: THURSDAY, September 10, 2009 — TSCS: Heterogeneity

08:00 – 09:15 Heterogeneity across units and over time

09:15 – 10:30 Heterogeneity and Least-Squares Dummy-Variable (Fixed-Effect) Models

- Wooldridge, Jeffrey M. 2002. *Econometric Analysis of Cross Section and Panel Data*, MIT Press, Cambridge, Ch. 10.1, 10.5-10.6, or equivalent.

10:30 – 12:00 Heterogeneity and Random-Effect Models

- Wooldridge, Jeffrey M. 2002. *Econometric Analysis of Cross Section and Panel Data*, MIT Press, Cambridge, Ch. 10.2-10.4, 10.7 or equivalent.

12:00 – 13:00 Lunch

13:00 – 14:30 Heterogeneity and Interaction Models

- Kam, Cindy D. & Franzese, Robert J., Jr. 2007. *Modeling & Interpreting Interactive Hypothesis in Regression Analysis*, University of Michigan Press.
- Franzese. 2005. “Empirical Strategies for Various Manifestations of Multilevel Data,” *Political Analysis* 13(4):430-46.

14:30 – 15:00 Random-Coefficient (a.k.a. Hierarchical, Multilevel, Mixed) Models

- Wooldridge, Jeffrey M. 2002. *Econometric Analysis of Cross Section and Panel Data*, MIT Press, Cambridge, Ch. 10.7 & 11.1, 11.3-4, or equivalent.

15:00 – 16:00 *Lab Session*: Fixed Effects; Random Effects; Random Coefficients; Hybrids; Interaction-Model Estimation, Testing, and Presentation

DAY 3: FRIDAY, September 11, 2009 — TSCS: Dynamics

08:00 – 09:30 (Time-)Dynamic Models

- Beck, N. 1991. “Comparing Dynamic Specifications,” *Political Analysis* 3(1):51-87.

09:30 – 11:00 (Time-)Dynamic Panel-Data Models

- Wawro, Gregory. 2002. “Estimating Dynamic Panel Data Models in Political Science,” *Political Analysis* 10(1):25-48.

11:00 – 12:00 *Lab Session*: Estimating and Presenting Temporally Dynamic Models

12:00 – 13:00 Lunch

13:00 – 15:00 Spatial and Spatiotemporal Interdependence in TSCS Data

- Franzese, R., & Jude C. Hays. 2008. “Empirical Models of Spatial Interdependence,” in Box-Steffensmeier, Brady, Collier (eds.): *Oxford Handbook of Political Methodology*.

15:00 – 16:00 *Lab Session*: Estimating and Presenting Spatially & Spatiotemporally Dynamic Models