Course Title: Econometric Modeling for Public Policy
Course Code: GP-EA-EMPP-7290
Program: MAPP 2025 elective + Cross elective
Semester: Fall 2026

Course Information

Course Duration: 1st week of August till last week of September - About 8 teaching weeks
Meetings: Weekly one and half hour office hour will be provided
Location: Campus
Prerequisites:
1. Students must have done at least one semester of Statistics (including hypothesis testing)
2. No prior knowledge of Econometrics or any Statistical Software needed.

Instructor Information

Instructor: Sunetra Ghatak (Ph.D.)

Biography: Dr. Sunetra Ghatak is an Associate Professor of Economics at O. P. Jindal Global University, India. She is also a Visiting Fellow at Asian Confluence and a HUC Fellow at Asian Mountain Academic Alliance, ICIMOD. She serves as Topic Editor, Mountain Research and Development (MRD) Journal (ISSN: 1994-7151, SCOPUS-indexed), published by the University of Bern, Switzerland. She holds a Ph.D. in Economics from Jawaharlal Nehru University, India. With more than seventeen years of academic and research experience, her work spans international trade, migration, gender, and economic development.

Email: sunetra.ghatak@jgu.edu.in
Office Hours: TBA
Homepage: Here

1. Course Description

This is a course on Applied Regression Methods. It focuses on the Linear Regression Model and Ordinary Least Squares (OLS), the foundation of quantitative policy analysis. The course will start with the basics - correlation and scatter plots followed by bivariate regression as an extension of the scatter plots. Within bivariate regression, the emphasis will be on the important differences in interpretation when the independent variable is an interval (or continuous) level variable and when it is a binary or dichotomous variable. Most of the course is dedicated to multiple regression using different types of independent variables – all interval level, all dichotomous, mixture of interval and dichotomous, interaction terms (moderator analysis) involving different types of variables, and finally all types together. Later, the course will cover hypothesis testing and residual analysis, and violations of assumptions like heteroskedasticity, feasible generalized least squares, and robust standard errors. The course will finish with some discussion of the causes, implications and remedies for some of these failures.

The idea of developing this course is to provide students with application-based exposure for eliciting useful information and identifying patterns and relationships in data. This will enable students to gain the confidence and ability to deal with raw data, from cleaning to estimating models, testing hypotheses, assessing strengths and weaknesses, generating useful conclusions, and communicating the results effectively, both verbally and in writing.

2. Course Intended Learning Objectives (Aim)

Course Intended Learning Outcomes	Teaching and Learning Activities	Assessments/ Activities
1. Apply the Single Equation Linear Regression Model (LRM) to real-world cross-sectional datasets.	- Interactive lectures deriving OLS mechanics with real-world examples (e.g., housing prices). - In-class "live coding" sessions (using Stata) using public datasets (e.g., CPS, ACS). - Lab worksheets focused on loading, visualizing, and running bivariate regressions.	- Weekly Quiz. - Problem Set 1: Manual calculation of OLS estimates from a small dataset to bridge theory to application.
2. Read, interpret, and critique published research articles that use LRM.	- Journal club sessions: Students bring a printed methodology section to class. - Guided critique templates (identifying IV, model specification, omitted	- Article Critique Memo: Evaluate one empirical paper's regression table and data limitations. - Peer review activity: Students anonymously critique a classmate's

	variable bias, heteroskedasticity). - Discussion of common reporting flaws (e.g., confusing correlation with causation).	model output.
3. Identify appropriate applications of LRM vs. other methods.	- Decision making whether to choose LRM vs. logit/probit. - Flawed model detection exercises: Instructor shows a regression where LRM is inappropriate (e.g., binary outcome, time series with trends). - Concept mapping activity.	- Multiple-choice quiz with scenario-based questions. - Research design brief: 1-page proposal justifying why LRM is suitable for a specific research question.
4. Clean data, estimate models, test hypotheses, and make inferences using software.	- Data clinic labs: Guided exercises on handling missing values, recoding categorical variables, detecting outliers (leverage/influence). - Hypothesis testing drills (t-tests, F-tests) using software output. - Walkthrough of prediction intervals and confidence intervals for mean response.	- Practice file: Students are given a CSV file (messy) and must produce a clean regression table with tests for heteroskedasticity (Breusch-Pagan) and normality of residuals. - Replication exercise: Replicate one table from a published paper.
5. Interpret results of inference and prediction and present them compellingly in writing and verbally.	- Peer review workshops: Students exchange written results sections and give structured feedback.	- Final sit-down Stata-based exam: Includes a regression table, interpretation of coefficients, prediction for a novel observation, and limitations. - 5-minute oral presentation using exactly 3 slides: research question, key result, real-world takeaway.

3. Scheme of Evaluation and Grading

Evaluation will be based on continuous assessment. Assessment components will include in-class quizzes and sit-down Stata-based exam. 10% of the total marks will be based on attendance.

4. Academic Integrity

Academic Honesty, Cheating, and Plagiarism – All cases will be heavily penalized.
Participation/Attendance Policy – Attendance will be strictly according to the university norms.
Use of phone/ texting/ laptop – Only laptop use is allowed.

5. Keyword Syllabus

Linear Model, Regression, Partial Effect, Interaction, Moderation Analysis, Log Model, Variable Transformation, Heteroskedasticity, Breusch-Pagan Test, Feasible Generalized Least Squares, Huber-White Sandwich Standard Error, Stata software.

6. Course Material

Main Reading:
Jeffrey M. Woolridge - Introductory Econometrics: A Modern Approach, Latest Edition.
For this course we will read the chapter 19, 1, 2, 3, 4, and 6 completely; 7,8, and 13 partly.

Other Readings:
Rachel A. Gordon, (Latest Edition), Regression Analysis for the Social Sciences by Routledge.
Christopher Dougherty (Latest Edition), Introduction to Econometrics by Oxford University Press.

Statistical Software (Stata):
Contact eresources@jgu.edu.in (JGU library) for installing latest Stata version in your computers.

7. Session Plan

Session		General Topic
Week 1	Day 1	1. Test of Statistics 2. Introduction to the Course 3. Motivating the course 4. Basic Bivariate Model of Mean 5. Bivariate Regression a. Estimation in Stata* b. Interpreting Results i. Continuous Predictor ii. Dummy Predictor
	Day 2	6. Bivariate Regression a. Interpretation of Results b. Dummy Predictor c. Predicted Values ($\hat{y}, \hat{u}$)
Week 2	Day 3	7. What OLS regression 8. Multiple Regression a. Interpretation of results i. All continuous predictors ii. All dummy predictors iii. 1 Cont. and 1 Dummy
	Day 4	9. Multiple Regression a. Change of scale b. Change of origin
Quiz 1		
Week 3	Day 5	10. Log-Log Model
	Day 6	11. Log-Lin Model 12. Quadratic Model
Quiz 2		
Week 4	Day 7	13. Interaction Terms/Moderator Analysis in Multiple Regression
	Day 8	14. Factor Notation 15. TSS, RSS, ESS, R2, Adjusted R2
Quiz 3		
Week 5	Day 9	16. Fitted Values from Log Model
	Day 10	17. Standardized Coefficients 18. Algebraic Properties
Quiz 4		
Week 6	Day 11	19. Assumptions of the Linear Model and their breakdown 20. OLS Estimators are BLUE 21. SE of Coefficients
	Day 12	22. Heteroskedasticity 23. BP Tests 24. Solutions
Quiz 5		

Week 7	Day 13	25. Multicollinearity and VIF
	Day 14	26. t test and lincom 27. F Test
Quiz 6		
Week 8	Day 15	Revision
	Day 16	Sit-down Stata-based Exam