

Ec228 Course Topics ... with mappings to sections of the Wooldridge text.

Introduction

Unit 1 – Introduction & Getting Started: [Chapter 1]¹ Estimating the relationship between x and y; causality v. correlation; data types; economic v. statistical significance; robust analysis (how many regressions did you run?); art v. science; sample statistics (sample means, variances, standard deviations, covariances, and correlations); standardizing data; OLS as minimization of SSRs (FOCs and SOC's)

Simple Linear Regression (OLS/SLR) Models

Unit 2 – OLS/SLR analytics (single explanatory variable): [Sections 2.1-2.4 & 6.1] *In the beginning* (SLR.1: the data generation process); residuals and sum squared residuals (SSR); OLS, FOCs and SOC's, Sample Regression Function (SRF), predictions and residuals; economic significance/meaningfulness (elasticity and *beta* regressions)

Unit 3 – OLS/SLR assessment: [Section 2.5] Sum Squared Explained (SSE) and Sum Squared Total (SST); $SST = SSE + SSR$ (w/ constant term in the model); Goodness of Fit (GOF) metrics - Coefficient of Determination (R^2), Mean Squared Error (MSE) and Root MSE (RMSE); comparing SLR models using GOF metrics

Multiple Linear Regression (MLR) Models

Unit 4 – OLS/MLR analytics I (adding, and subtracting, explanatory variables): [Sections 3.1-3.3 & 9.2] Comparing SLR and MLR analytics; interpreting coefficients I – *ceteris paribus* (partial effects and the SRF); interpreting coefficients II – partial correlations (*WhatsLeft* and *WhatsNew*); an overview of omitted variable bias (endogeneity)

Unit 5 – OLSMLR assessment: [Section 6.3] Comparing SLR and MLR assessment (GOF metrics); shortcomings of R^2 ; adjusted R^2 ; comparing MLR models using GOF metrics

Unit 6 – OLS/MLR analytics II: [Sections 3.2.-3.3] The *collinearity* regression; multicollinearity, R^2 's and Variance Inflation Factors (VIFs); Omitted variable bias/impact (endogeneity); simple v. partial correlations

Way-Too-Fast Review of Statistics

Unit 7 – Review of Estimation and Inference: [Appendix C] Our focus will be on estimation of the population mean; LUEs (Linear Unbiased Estimators); BLUEs (Best Linear Unbiased Estimators); point and interval estimators; standard errors, t statistics, p-values; confidence levels; critical values; confidence intervals; hypothesis testing; significance levels

Estimation and Inference in Regression Analysis

Unit 8 – SLR Estimation: [Sections 2.5 & 3.2-3.5] Gauss-Markov assumptions (SLR.1 – SLR.5); Population Regression Function (PRF); conditional means; means, variances, standard deviations and standard errors of OLS estimators (intercepts and slopes); unbiasedness (OLS coefficients; MSE); LUEs; homoskedasticity; *BLUE*: The Gauss-Markov Theorem

¹ []'s are for Wooldridge text references... so [Chapter 1] is Wooldridge, Chapter 1

Course Topics: Wooldridge Mappings

Unit 8a – Heteroskedasticity: [Chapter 8] Issues (OLS standard errors no longer correct; LUE but not BLUE); White-corrected standard errors (*robust* inference); working towards BLUE (weighted least squares... but where do those weights come from?)

Unit 9 – SLR Inference: [Sections 4.1-4.3] Add SLR.6 to the mix; normally distributed errors; variances, standard deviations and standard errors; t statistics; t-tests (Null hypotheses); p values; confidence intervals; hypothesis tests; economic v. statistical significance (elasticities v. p-values); *Convergence I* (t stats and R^2)

Unit 10 – MLR Estimation and Inference: [Sections 3.2-3.5 & 4.1-4.3] Compare to SLR; *What's new? ... Not much!*; now MLR.1-MLR.5; $n-k-1$; multicollinearity, standard errors and Variance Inflation Factors (VIFs); MLR.6; heteroskedasticity and *robust* standard errors

Topics

Unit 11 – Dummy Variables and Fixed Effects: [Sections 7.1-7.5 & 14.1] Dummies revisited; on the RHS and on the LHS; uses on the RHS (slope and intercept dummies); quieting the endogeneity critics (fixed effects); Examples (sovereign debt ratings, gender bias in wages, and death penalty deterrence)

Unit 12 – F-tests and Convergence: [Sections 4.4-4.6 & 7.4] Extension of t-tests to more complicated null hypotheses; testing linear restrictions with the $F(q, n-k-1)$ distribution; *Convergence II:* connects Goodness-of-Fit metrics and inference stats (t stats and incremental R^2 , SSR and SSE); reported F stats (for the regression) and associated p values; relation to adjusted R^2 ; *Babies and bathwater*; Chow tests; machine learning

Unit 13 – Linear Probability Models (LPMs) and Functional Forms: [Sections 6.2 & 7.5] Dummies on the LHS; Linear Probability Models (LPMs); exploring functional forms (quantile dummies; linear splines; logarithms and exponentials; polynomials; cubic splines, and fixed effects)

Further Topics

Unit 14 – Further Topics I: [Sections 13.1-13.3 & 15.1-15.3 & Chapter 17] *Differences-in-Differences* (*Deflategate*; NBA Referee Own-Race Bias); *Regression Discontinuity Designs* (Highway Fatalities & Daylight Savings Time); *Instrumental Variables* (The Oregon Health Insurance Experiment (Medicaid)); Maximum Likelihood Estimation (MLE); limited dependent variables; logit and probit models; censored and truncated regression models

Unit 15 – Further Topics II: [Sections 5.1-5.3, 6.4 & 9.3-9.4] OLS asymptotics (large sample properties; consistency; convergence in distribution); misspecified models; proxy variables; missing data; outliers; non-random samples; forecasting and *prediction* intervals