

# INTRODUCTORY ECONOMETRICS

*Using Monte Carlo Simulation with  
Microsoft Excel®*

**HUMBERTO BARRETO**

*Wabash College*

**FRANK M. HOWLAND**

*Wabash College*



**CAMBRIDGE  
UNIVERSITY PRESS**

# Contents

<i>Preface</i>	<i>page xvii</i>
<b>User Guide</b>	<b>1</b>
0.1. Conventions and Organization of Files	1
0.2. Preparing and Working with Microsoft Excel®	3
<b>1. Introduction</b>	<b>10</b>
1.1. Definition of Econometrics	10
1.2. Regression Analysis	11
<i>Cig.xls</i>	
1.3. Conclusion	28
1.4. Exercises	29
References	30
<b>PART 1. DESCRIPTION</b>	
<b>2. Correlation</b>	<b>33</b>
2.1. Introduction	33
2.2. Correlation Basics	33
<i>Correlation.xls</i>	
2.3. Correlation Dangers	40
<i>Correlation.xls</i>	
<i>IMRGDP.xls</i>	
2.4. Ecological Correlation	44
<i>EcolCorr.xls</i>	
<i>EcolCorrCPS.xls</i>	
2.5. Conclusion	50
2.6. Exercises	51
References	51

<b>3. PivotTables</b>	53
3.1. Introduction	53
3.2. The Basic PivotTable	53
<i>IndianaFTWorkers.xls (in Basic Tools\InternetData\CPS)</i>	
<i>Histogram.xla (Excel add-in)</i>	
3.3. The Crosstab and Conditional Average	59
<i>IndianaFTWorkers.xls (in Basic Tools\InternetData\CPS)</i>	
3.4. PivotTables and the Conditional Mean Function	65
<i>EastNorthCentralFTWorkers.xls</i>	
3.5. Conclusion	69
3.6. Exercises	70
References	71
<b>4. Computing the OLS Regression Line</b>	72
4.1. Introduction	72
4.2. Fitting the Ordinary Least Squares Regression Line	72
<i>Reg.xls</i>	
4.3. Least Squares Formulas	82
<i>Reg.xls</i>	
4.4. Fitting the Regression Line in Practice	84
<i>Reg.xls</i>	
4.5. Conclusion	90
4.6. Exercises	90
References	91
Appendix: Deriving the Least Squares Formulas	91
<b>5. Interpreting OLS Regression</b>	95
5.1. Introduction	95
5.2. Regression as Double Compression	95
<i>DoubleCompression.xls</i>	
<i>EastNorthCentralFTWorkers.xls</i>	
5.3. Galton and Two Regression Lines	104
<i>TwoRegressionLines.xls</i>	
5.4. Properties of the Sample Average and the Regression Line	107
<i>OLSFormula.xls</i>	
5.5. Residuals and the Root-Mean-Square Error	114
<i>ResidualPlot.xls</i>	
<i>RMSE.xls</i>	
5.6. R-Squared ( $R^2$ )	122
<i>RSquared.xls</i>	
5.7. Limitations of Data Description with Regression	126
<i>Anscombe.xls</i>	
<i>IMRGDPReg.xls</i>	

<i>SameRegLineDifferentData.xls</i>	
<i>HourlyEarnings.xls</i>	
5.8. Conclusion	135
5.9. Exercises	135
References	136
Appendix: Proof that the Sample Average is a Least Squares Estimator	136
<b>6. Functional Form of the Regression</b>	138
6.1. Introduction	138
6.2. Understanding Functional Form via an Econometric Fable	139
<i>Galileo.xls</i>	
6.3. Exploring Two Other Functional Forms	144
<i>IMRGDPFunForm.xls</i>	
6.4. The Earnings Function	148
<i>SemiLogEarningsFn.xls</i>	
6.5. Elasticity	155
6.6. Conclusion	159
6.7. Exercises	160
References	160
Appendix: A Catalog of Functional Forms	161
<i>FuncFormCatalog.xls</i>	
<b>7. Multiple Regression</b>	164
7.1. Introduction	164
7.2. Introducing Multiple Regression	165
<i>MultiReg.xls</i>	
7.3. Improving Description via Multiple Regression	174
<i>MultiReg.xls</i>	
7.4. Multicollinearity	184
<i>Multicollinearity.xls</i>	
7.5. Conclusion	191
7.6. Exercises	192
References	194
Appendix: The Multivariate Least Squares Formula and the Omitted Variable Rule	194
<b>8. Dummy Variables</b>	198
8.1. Introduction	198
8.2. Defining and Using Dummy Variables	199
<i>Female.xls</i>	
8.3. Properties of Dummy Variables	202
<i>Female.xls</i>	

8.4. Dummy Variables as Intercept Shifters	205
<i>Female.xls</i>	
8.5. Dummy Variable Interaction Terms	208
<i>Female.xls</i>	
8.6. Conclusion	211
8.7. Exercises	212
References	212
<b>PART 2. INFERENCE</b>	
<b>9. Monte Carlo Simulation</b>	215
9.1. Introduction	215
9.2. Random Number Generation Theory	216
<i>RNGTheory.xls</i>	
9.3. Random Number Generation in Practice	220
<i>RNGPractice.xls</i>	
9.4. Monte Carlo Simulation: An Example	225
<i>MonteCarlo.xls</i>	
9.5. The Monte Carlo Simulation Add-In	232
<i>MonteCarlo.xls</i>	
<i>MCSim.xla (Excel add-in)</i>	
<i>MCSimSolver.xla (Excel add-in)</i>	
9.6. Conclusion	235
9.7. Exercises	236
References	236
<b>10. Review of Statistical Inference</b>	238
10.1. Introduction	238
10.2. Introducing Box Models for Chance Processes	239
10.3. The Coin-Flip Box Model	242
<i>BoxModel.xls</i>	
10.4. The Polling Box Model	251
<i>PresidentialHeights.xls</i>	
10.5. Hypothesis Testing	257
<i>PValue.xla (Excel add-in)</i>	
10.6. Consistent Estimators	262
<i>Consistency.xls</i>	
10.7. The Algebra of Expectations	265
<i>AlgebraofExpectations.xls</i>	
10.8. Conclusion	277
10.9. Exercises	277
References	278
Appendix: The Normal Approximation	278

<b>11. The Measurement Box Model</b>	281
11.1. Introduction	281
11.2. Introducing the Problem	283
11.3. The Measurement Box Model	285
11.4. Monte Carlo Simulation	290
<i>Measure.xls</i>	
11.5. Applying the Box Model	293
<i>Measure.xls</i>	
11.6. Hooke's Law	296
<i>HookesLaw.xls</i>	
11.7. Conclusion	301
11.8. Exercises	301
References	302
<b>12. Comparing Two Populations</b>	303
12.1. Introduction	303
12.2. Two Boxes	303
12.3. Monte Carlo Simulation of a Two Box Model	306
<i>TwoBoxModel.xls</i>	
12.4. A Real Example: Education and Wages	309
<i>CPS90Workers.xls</i>	
12.5. Conclusion	314
12.6. Exercises	315
<i>CPS90ExpWorkers.xls</i>	
References	315
<b>13. The Classical Econometric Model</b>	316
13.1. Introduction	316
13.2. Introducing the CEM via a Skiing Example	316
<i>Skiing.xls</i>	
13.3. Implementing the CEM via a Skiing Example	322
<i>Skiing.xls</i>	
13.4. CEM Requirements	328
13.5. Conclusion	332
13.6. Exercises	333
References	334
<b>14. The Gauss–Markov Theorem</b>	335
14.1. Introduction	335
14.2. Linear Estimators	336
<i>GaussMarkovUnivariate.xls</i>	
14.3. Choosing an Estimator	342
<i>GaussMarkovUnivariate.xls</i>	

14.4. Proving the Gauss–Markov Theorem in the Univariate Case	347
<i>GaussMarkovUnivariate.xls</i>	
14.5. Linear Estimators in Regression Analysis	352
<i>GaussMarkovBivariate.xls</i>	
14.6. OLS is BLUE: The Gauss–Markov Theorem for the Bivariate Case	361
<i>GaussMarkovBivariate.xls</i>	
14.7. Using the Algebra of Expectations	366
<i>GaussMarkovUnivariate.xls</i>	
<i>GaussMarkovBivariate.xls</i>	
14.8. Conclusion	374
14.9. Exercises	375
References	376
<b>15. Understanding the Standard Error</b>	<b>378</b>
15.1. Introduction	378
15.2. SE Intuition	378
<i>SEb1OLS.xls</i>	
15.3. The Estimated SE	383
<i>SEb1OLS.xls</i>	
15.4. The Determinants of the SE of the OLS Sample Slope	387
<i>SEb1OLS.xls</i>	
15.5. Estimating the SD of the Errors	392
<i>EstimatingSDErrors.xls</i>	
15.6. The Standard Error of the Forecast and the Standard Error of the Forecast Error	398
<i>SEForecast.xls</i>	
15.7. Conclusion	408
15.8. Exercises	409
References	410
<b>16. Confidence Intervals and Hypothesis Testing</b>	<b>411</b>
16.1. Introduction	411
16.2. Distributions of OLS Regression Statistics	412
<i>LinestRandomVariables.xls</i>	
16.3. Understanding Confidence Intervals	421
<i>ConfidenceIntervals.xls</i>	
16.4. The Logic of Hypothesis Testing	430
<i>HypothesisTest.xls</i>	
16.5. Z- and T-Tests	434
<i>ConfidenceIntervals.xls</i>	
<i>ZandTTests.xls</i>	

16.6. A Practical Example	443
<i>CigDataInference.xls</i>	
16.7. Conclusion	450
16.8. Exercises	450
<i>SemiLogEarningsFn.xls</i>	
References	451
<b>17. Joint Hypothesis Testing</b>	453
17.1. Introduction	453
17.2. Restricted Regression	456
<i>NoInterceptBug.xls</i>	
17.3. The Chi-Square Distribution	458
<i>ChiSquareDist.xls</i>	
17.4. The $F$ -Distribution	461
<i>FDist.xls</i>	
17.5. An $F$ -Test: The Galileo Example	462
<i>FDistGalileo.xls</i>	
17.6. $F$ - and $T$ -Tests for Equality of Two Parameters	468
<i>FDistFoodStamps.xls</i>	
17.7. $F$ -Test for Multiple Parameters	475
<i>FDistEarningsFn.xls</i>	
17.8. The Consequences of Multicollinearity	478
<i>CorrelatedEstimates.xls</i>	
17.9. Conclusion	487
17.10. Exercises	487
<i>MyMonteCarlo.xls</i>	
References	488
<b>18. Omitted Variable Bias</b>	490
18.1. Introduction	490
18.2. Why Omitted Variable Bias Is Important	491
18.3. Omitted Variable Bias Defined and Demonstrated	493
<i>SkiingOVB.xls</i>	
18.4. A Real Example of Omitted Variable Bias	498
<i>ComputerUse1997.xls</i>	
18.5. Random $X$ 's: A More Realistic Data Generation Process	502
<i>ComputerUse1997.xls</i>	
18.6. Conclusion	506
18.7. Exercises	506
References	507

<b>19. Heteroskedasticity</b>	508
19.1. Introduction	508
19.2. A Univariate Example of Heteroskedasticity	510
<i>Het.xls</i>	
19.3. A Bivariate Example of Heteroskedasticity	518
<i>Het.xls</i>	
19.4. Diagnosing Heteroskedasticity with the B-P Test	527
<i>Het.xls</i>	
<i>BPSampDist.xls</i>	
19.5. Dealing with Heteroskedasticity: Robust Standard Errors	533
<i>HetRobusSE.xls</i>	
<i>OLSRegression.xls (Excel add-in)</i>	
19.6. Correcting for Heteroskedasticity: Generalized Least Squares	542
<i>HetGLS.xls</i>	
19.7. A Real Example of Heteroskedasticity: The Earnings Function	549
<i>WagesOct97.xls</i>	
19.8. Conclusion	555
19.9. Exercises	556
References	557
<b>20. Autocorrelation</b>	558
20.1. Introduction	558
20.2. Understanding Autocorrelation	560
<i>AutoCorr.xls</i>	
20.3. Consequences of Autocorrelation	566
<i>AutoCorr.xls</i>	
20.4. Diagnosing Autocorrelation	576
<i>AutoCorr.xls</i>	
20.5. Correcting Autocorrelation	588
<i>AutoCorr.xls</i>	
20.6. Conclusion	600
<i>CPIMZM.xls</i>	
<i>Luteinizing.xls</i>	
20.7. Exercises	602
<i>Misspecification.xls</i>	
<i>FreeThrowAutoCorr.xls</i>	
References	603
<b>21. Topics in Time Series</b>	604
21.1. Introduction	604
21.2. Trends in Time Series Models	605
<i>IndiaPopulation.xls</i>	
<i>ExpGrowthModel.xls</i>	

<i>AnnualGDP.xls</i>	
<i>Spurious.xls</i>	
21.3. Dummy Variables in Time Series Models	613
<i>TimeSeriesDummyVariables.xls</i>	
<i>CoalMining.xls</i>	
21.4. Seasonal Adjustment	617
<i>SeasonalTheory.xls</i>	
<i>SeasonalPractice.xls</i>	
21.5. Stationarity	624
<i>Stationarity.xls</i>	
21.6. Weak Dependence	633
<i>Stationarity.xls</i>	
<i>Spurious.xls</i>	
21.7. Lagged Dependent Variables	638
<i>PartialAdjustment.xls</i>	
21.8. Money Demand	645
<i>MoneyDemand.xls</i>	
<i>LaggedDepVar.xls</i>	
21.9. Comparing Forecasts Using Different Models of the DGP	652
<i>AnnualGDP.xls</i>	
<i>ForecastingGDP.xls</i>	
21.10. Conclusion	658
21.11. Exercises	659
References	661
<b>22. Dummy Dependent Variable Models</b>	<b>663</b>
22.1. Introduction	663
22.2. Developing Intuition about Dummy Dependent Variable Models	666
<i>Raid.xls</i>	
22.3. The Campaign Contributions Example	669
<i>CampCont.xls</i>	
22.4. A DDV Box Model	671
<i>Raid.xls</i>	
<i>CampCont.xls</i>	
22.5. The Linear Probability Model (OLS with a Dummy Dependent Variable)	674
<i>CampCont.xls</i>	
<i>LPMMonteCarlo.xls</i>	
22.6. Nonlinear Least Squares Applied to Dummy Dependent Variable Models	680
<i>NLLSFit.xls</i>	
<i>NLLSMCSim.xls</i>	

22.7. Interpreting NLLS Estimates	690
<i>NLLSFit.xls</i>	
22.8. Is There Mortgage Discrimination?	695
<i>MortDisc.xls</i>	
<i>MortDiscMCSim.xls</i>	
<i>DDV.xls</i> (Excel add-in)	
<i>DDVGN.xls</i> (Excel add-in)	
22.9. Conclusion	706
22.10. Exercises	707
References	707
<b>23. Bootstrap</b>	<b>709</b>
23.1. Introduction	709
23.2. Bootstrapping the Sample Percentage	710
<i>PercentageBootstrap.xls</i>	
23.3. Paired XY Bootstrap	713
<i>PairedXYBootstrap.xls</i>	
23.4. The Bootstrap Add-in	718
<i>PairedXYBootstrap.xls</i>	
<i>Bootstrap.xls</i> (Excel add-in)	
23.5. Bootstrapping $R^2$	721
<i>BootstrapR2.xls</i>	
23.6. Conclusion	726
23.7. Exercises	728
References	728
<b>24. Simultaneous Equations</b>	<b>730</b>
24.1. Introduction	730
24.2. Simultaneous Equations Model Example	731
24.3. Simultaneity Bias with OLS	735
<i>SimEq.xls</i>	
24.4. Two Stage Least Squares	741
<i>SimEq.xls</i>	
24.5. Conclusion	745
24.6. Exercises	746
References	747
<i>Glossary</i>	749
<i>Index</i>	761