



UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

EKT 816

Applied Microeconometrics

Thursdays, 1630-1900, Tukkieurf



Prof SF Koch

steve.koch@up.ac.za

<https://sites.google.com/site/kocheconomics/>

Click-UP

Tukkiewerf 1.10-1

Office Hours: Thursdays, 1330-1630

by appointment: steve.koch@up.ac.za

012 420 5285

Although every effort is made to prepare a syllabus ahead of time and to stick to that syllabus, the reality, of course, is that coverage is subject to change. I will do my best to post updates, as needed on the abovementioned Click-UP page. In the past, I have made extensive use of my website; in general, you can find code and data related to past research papers.

Course Description: Within applied econometrics, there are six major topics; these are: (1) Data and its quality, (2) Measurement, (3) Determinants, (4) Causes, (5) Testing and (6) Presentation. Most of you would likely, also, think ‘techniques’ are a major topic (and the syllabus suggests as much). Although there are plenty of ‘techniques’ out there, I think it is too easy to focus on technique, rather than the question of interest. For that reason, I want us to think of technique or techniques as being an underlying set of skills. One cannot know them all, or teach them all, and, as stated above, we will look at some techniques. However, in providing empirical support or using empirical evidence to refute a theory, technique for the sake of technique is fairly empty of meaning... However, with the proper presentation of information and work to break the results that you have, the ‘artistic’ endeavors in econometrics, empirical evidence can be used to answer (or not) a question of interest.

As this is an applied econometrics course, and I think there are six main components, our objectives are rather simple. We want to learn a bit about data, and its quality (or lack thereof). That data can be used for measurement and analysis, and we want to use it for a bit of both. Once we are comfortable interacting with the data, we would like to use it for some sort of analysis, including determinants, and, especially, differences in determinants that might help us better understand the world around us. Finally, we will investigate, where possible (sometimes we will make unsupported assumptions just for the purpose of practice), the causes of the measures that we observe in the data. Throughout all of this practical engagement, we will pay careful attention to statistical testing (and result sensitivity), and what it can say in that (those) circumstance(s). However, we will avoid focusing our attention on “star” econometrics, i.e., looking for statistically significant results.

Prerequisite(s): Intermediate Econometrics – at the University of Pretoria, EKT 713 and/or EKT 723 would qualify. An equivalent module might also be considered. In particular, students are expected to have learned linear regression, binary response models, and simple systems of equations, especially IV. According to the 2017 syllabus from EKT 713, for example, the module covered bivariate and multiple regression, inference and some asymptotics, binary indicators, heteroskedasticity specification and data issues, pooled cross-sections and instrumental variables.

Note(s): A minimum grade of 40 for the semester test is required to participate in the exam.

Credit Hours: 10

Text(s): At this stage of your academic careers, you should think of textbooks as reference materials. If you really want to learn the concepts; however, you need to try to apply them. The more applications you attempt, the more you will learn. A number of useful textbooks are listed. Clearly, there are many others out there. Also, please note that we will make extensive use of working papers and published articles that you are expected to read *before the scheduled class*.

- Angrist, J. D. & Pischke, J.-S. 2015. *Mastering Metrics: The Path from Cause to Effect*, Princeton University Press, Princeton, NJ, USA
- Verbeek, M. 2012. *A Guide to Modern Econometrics, Fourth Edition*. John Wiley & Sons, Sussex, UK.
- Wooldridge, J. M. 2010. *Econometric Analysis of Cross Section and Panel Data: Second Edition*, The MIT Press, Cambridge, MA, USA
- Angrist, J. D. & Pischke, J.-S. 2009. *Mostly Harmless Econometrics: An Empiricist's Companion*, Princeton University Press, Princeton, NJ, USA
- Cameron, A. C. & Trivedi, P. K. 2005. *Microeconomics: Methods and Applications*, Cambridge University Press, Cambridge, UK
- Deaton, A. 1997. *The Analysis of Household Surveys: A Microeconomic Approach to Development Policy*. The Johns Hopkins University Press (for The World Bank), Baltimore, USA. It is available, free, online through The World Bank.
- O'Donnell, O., E. van Doorslaer, A. Wagstaff and M. Lindelow. 2008. *Analyzing Health Equity Using Household Survey Data: A Guide to Techniques and Their Implementation*. The World Bank, Washington, USA. This is also available, free, online through The World Bank.

Software Reference Text(s): Even more than in the case of standard textbook information, it is quite easy to locate useful software references. I work almost exclusively in R (<https://cran.r-project.org/>). Why? It is entirely free, and it ‘interacts’ with latex (<https://www.ctan.org/?lang=en>), which is also free.¹ Finally, I should note that it is possible to link R with Word, and similarly create decent papers that way, although it is not my preferred software for writing papers (because Latex looks better).

¹There are a number of free programs and packages that offer viable alternatives to MS Office, if you are interested, including OpenOffice (<https://www.openoffice.org/>).

I do, however, work in STATA, too, partly because of its capabilities, when it comes to manipulating data, and the fact that many datasets are available in STATA. Having said that, R is at least as capable, and is capable of reading data from a wide range of sources: STATA, SAS, SQL, and more (without the added cost of additional software). Thus, my notes will be built mostly in R, and I will include code for you. I will try to put together additional material, where plausible, in STATA. Most importantly, I will not stop you from using other programs, and the lab includes all of the required material. Please, note that there are also many resources available online to help you with ‘coding’, when you are busy with any project (now and into the future). However, as with anything... If you want to be ‘good’, you have to put in the effort!

For your information, I do list a few useful texts for working with STATA and R, but do not forget the internet...

- Cameron, A. C. & Trivedi, P. K. 2009. *Microeconomics Using Stata*, Revised Edition, STATA Press, College Station, USA; many people have been able to access this book for free from the internet.
- Kleiber, C. & Zeileis, A. 2008. *Applied Econometrics with R*, Springer Science and Business Media, Amsterdam, the Netherlands
- Gelman, A. & Hill, J. 2007. *Data Analysis Using Regression and Multilevel/Hierarchical Models*, Cambridge University Press, Cambridge, UK

Course Objectives: At the completion of this course, students will be able to:

1. Find, source and cite data
2. Understand a data set and its limitations
3. Undertake computational analysis in a range of areas
4. Examine robustness of analysis
5. Present and illustrate analytical results

Grade Distribution:

Assignments	30%
Presentation	30%
Exam (Possibly Project)	40%

Course Grading: I generally follow a *fuzzy logic* approach to actual grading, because there are two thresholds that matter: (i) 50% and (ii) 75%. Therefore, I start the marking process by assigning ordered factor values, $\psi = \{0, 1, 2, 3, 4\}$. Zero ought to be obvious. In principle, one corresponds to “extremely poor performance, and unacceptable for even an honours student”; two corresponds to “poor performance, which means it is not acceptable at the M level”; three corresponds to “acceptable performance” and 4 corresponds to “distinctive performance”. Each milestone/assignment is, first, placed within these relative categories, and, second, rankings are made within each category. Third, following the rankings, marks befitting the category and the marks weight of the milestone are ascribed to the actual assignment.

All typewritten assignments should be submitted to me, via email, using, “studentsurname_studentnumber_deliverable#.pdf”.² Sometimes, email servers go down, so I

²You are always welcome to submit written assignments as scanned copies, as well, using this convention. If it is a group assignment, I only want one to be sent, so be sure that all names are included; I will not ‘accept’ additional names after the time. Thus, it is also up to group members to be sure that they get the credit they deserve.

would suggest not waiting for the last minute to email the file. Beforehand is always acceptable. Serious penalties are invoked for late deliverables: $m = \max\{0, \psi - t\}$, where m is the final category for your mark, t is the number of days the deliverable is late and ψ is the category, wherein your mark would have fallen (see above) without the delay. In other words, it is necessary to receive an initial mark beyond the number of days late; otherwise, a zero is your likely score! Importantly, at 1631 on the day the deliverable is due, $t \rightarrow 1$.

Academic Honesty Policy Summary: In addition to skills and knowledge, the University of Pretoria aims to teach students appropriate Ethical and Professional Standards of Conduct. The Academic Honesty Policy exists to inform students and Faculty of their obligations in upholding the highest standards of professional and ethical integrity. All student work is subject to the Academic Honesty Policy. Professional and Academic practice provides guidance about how to properly cite, reference, and attribute the intellectual property of others. Any attempt to deceive a faculty member or to help another student to do so will be considered a violation of this standard.

Instructor's Intended Purpose The student's work must match the instructor's intended purpose for an assignment. While the instructor will establish the intent of an assignment, each student must clarify outstanding questions of that intent for a given assignment.

Unauthorized/Excessive Assistance The student may not give or get any unauthorized or excessive assistance in the preparation of any work.

Authorship The student must clearly establish authorship of a work. Referenced work must be clearly documented, cited, and attributed, regardless of media or distribution. Even in the case of work licensed as public domain or Copyleft, (See: <http://creativecommons.org/>) the student must provide attribution of that work in order to uphold the standards of intent and authorship.

Declaration Online submission of, or placing one's name on an exam, assignment, or any course document is a statement of academic honor that the student has not received or given inappropriate assistance in completing it and that the student has complied with the Academic Honesty Policy in that work.

Consequences An instructor may impose a sanction on the student that varies depending upon the instructor's evaluation of the nature and gravity of the offense. Possible sanctions include but are not limited to, the following: (1) Require the student to redo the assignment; (2) Require the student to complete another assignment; (3) Assign a grade of zero to the assignment or exam; (4) Submit the offending student and conduct to the appropriate oversight committee on academic dishonesty.

Attendance: Although there is no official attendance policy, missing classes is not condoned, especially without informing the instructor. Being late, although often understandable, is also to be avoided; again, curtesy suggests an attempt to inform the instructor beforehand.

Data for Research Disclosure: Any and all results of in-class and out-of-class assignments and examinations are data sources for research and may be used in published research. All such use will always be anonymous.

Tentative Course Outline: As with all course outlines, the following is a plan, and not necessarily a perfect reflection of my ability to predict the pace of coverage. Also, the topics could very easily change, depending upon needs.

Part I: Unconfoundedness

- Feb 8 *Introduction, Data, Some Measures and the Lab*
 - The Syllabus
 - GHS 2012
 - Poverty and Inequality

- Feb 15 *OLS and Issues*
 - Deaton: Chapter 1
 - O’Donnell: Chapters 1 & 2
 - Angrist: Chapter 8
 - Black, SE. 1999. Do Better Schools Matter? Parental Evaluation of Elementary Education. *Quarterly Journal of Economics* 114(2): 577-599.

- Feb 22 *Regression: Redux and Decomposition*
 - Deaton: Chapter 2.1–2.2
 - Angrist: Chapter 3.2–3.2
 - O’Donnell: Chapters 10 & 12
 - Hlavac, M. 2016. `oaxaca`: Blinder-Oaxaca Decomposition in R
<https://cran.r-project.org/web/packages/oaxaca/>

- March 1 *MLE and Categorical Outcome Models*
 - Verbeek: Chapters 6 & 7.1-7.2
 - O’Donnell: Chapter 11

Part II: Endogeneity

- March 8 *Propensity Matching*
 - Angrist: Chapter 3.3
 - Seekhon. 2011. Multivariate and propensity score matching software with automatic balance optimization: The ‘Matching’ package for R. *Journal of Statistical Software* 42(7): 1-52.
 - DiNardo, Fortin, Lemieux. 1996. Labor Market Institutions and the Distribution of Wages, 1973-1992: A Semiparametric Approach. *Econometrica* 65(4): 1001-1044.
 - Recent paper on standard errors in matching problems

- March 15 *Instrumental Variables and Control Functions*

- Deaton: Chapters 2.4 & 2.6
- Angrist: Chapter 4
- Angrist: Lifetime earnings and the Vietnam era draft lottery: Evidence from Social Security Administrative Records. *American Economic Review* 80:313-336.
- Finkelstein et al. 2012. The Oregon health insurance experiment: Evidence from the first year. *Quarterly Journal of Economics* 127(3): 1057-1106.
- Add a control function paper or two
- March 22 *Regression Discontinuity and Differences-in-Differences*
 - Angrist: Chapter 6
 - Carpenter and Dobkin. 2009. The effect of alcohol consumption on mortality: Regression discontinuity evidence from the minimum drinking age. *American Economic Journal: Applied Economics* 1(1): 164-182.
 - Koch and Racine. (2016). Healthcare facility choice and user fee abolition: regression discontinuity in a multinomial choice setting. *Journal of the Royal Statistical Society: Series A* DOI: 10.1111/rssa.12161.
- April 12 *Clusters and Panels*
 - Deaton: Chapter 2.5
 - Angrist: Chapter 5
 - Racine: Chapter 6
 - Bertrand, Duflo and Mullainathan. 2004. How much should we trust differences-in-differences estimates. *Quarterly Journal of Economics* 119(1): 249-275.
- April 19 *Repeated Cross-Sections and Peer Effects*
 - Deaton: Chapter 2.7
 - Verbeek: Chapter 11³
 - Bertrand, Luttmer and Mullainathan. 2000. Peer effects and welfare cultures. *Quarterly Journal of Economics* 115(3): 1019-1055.
- April 19 **Presentations**
- April 26 **Presentations**

³This chapter 11 from a different book than noted above, but it is available from his website. See <http://www.uio.no/studier/emner/sv/oekonomi/ECON5103/v10/undervisningsmateriale/PDAppl17.pdf>