

Causal Inference for the Social Sciences  
Master in Economics: Empirical Applications and  
Policies

Universidad del País Vasco - Euskal Herriko  
Unibertsitatea

Javier Gardeazabal

January 5, 2017

Office: 2C45, third floor.

Phone: +34 946013780

E-mail: javier.gardeazabal@ehu.eus

Office hours: Monday and Tuesday, 8:30 - 9:30 and 11:30 - 12:00, Wednesday,  
Thursday and Friday 10:00-12:00 or by appointment.

## Description

Causal inference for the Social Sciences covers methods to establish causal relationships between a treatment, policy or intervention and an outcome or endogenous variable using different types of data: experimental and observational data. A particularly important application of causal inference is the evaluation of public programs or policies. Sometimes, people refer to the methods described in this course as econometric policy evaluation or econometric program evaluation. These methods allow the researcher to determine whether a policy or program has the intended effect in a quantitatively sound manner.

## Pre-requisites

In my lectures, I will assume anyone attending this course is familiar with descriptive statistics, probability theory and regression methods.

## After taking this course you will know how to:

- perform program evaluation in a quantitatively sound manner

- determine whether a causal relationship exists using observational data
- be able to apply the core methods used for statistical inference in economics and other social sciences
- apply the core causal inference methods to real data using STATA
- address a wide array of research questions empirically

## Outline of the course

The scientific method (point 1 below) relies on experimental methods (point 2 below) to validate hypotheses. In economics and other social sciences, it turns out that applying the experimental methods is often not feasible, unethical or unlawful. However, social scientists do have access to rich data sets which are collected by recording magnitudes in the real world, the so called observational data. This course covers a variety of methods (points 3 to 10 below) to establish a causal relationship using observational data.

1. The scientific method:  
An outline of the scientific method. Sampling methods. External and internal Validity. Construct validity. Reliability. Levels of measurement. Research design. Types of experiments.
2. Randomized experiments:  
Subjects. Treatments. Outcomes. Potential Outcomes. Treatment effects. Random assignment. Estimation. Testing. Regression interpretation. Examples.
3. Regression methods:  
Non-random assignment. Selection bias. Conditional Independence. Regression formulation. Propensity score. Estimation and testing. Examples.
4. Matching methods:  
Matching at the cell level. Common support. Matching on the score. Nearest neighbor matching. Combining matching and regression. Examples.
5. Inverse Probability Weighting:  
Missing data analog. Treatment effects as weighted means. Estimation. Combining inverse probability weighting and regression. Examples.
6. Regression discontinuity design:  
Treatment under discontinuity. Treatment effect at the margin. Local regression. Sharp and fuzzy regression discontinuity designs. Estimation. Examples.
7. Instrumental Variables:  
Endogenous treatment status. Instrumental variables: relevance and exclusion restrictions. IV estimation. Binary instruments. Local average treatment effects. Examples.

8. Difference-in-differences:  
Regression interpretation. Pre- versus post-treatment differences. Treatment versus control differences. Difference-in-differences. Parallel trends. Examples.
9. Panel data methods:  
Fixed effects. First differences. Difference-in-differences interpretation. Treatment histories. Propensity score weighting. Dynamic treatment effects. Examples.
10. Comparative case studies:  
Case studies. Comparative case studies. The synthetic control method. Placebo analysis and inference. Examples.

## References

The main text for this course is:

- Angrist, J. D. and J. S. Pischke, 2009. Mostly harmless econometrics; An empiricist's companion. Princeton University Press.  
You can borrow this book from the library: [library's catalog reference](#)  
A pre-print version of this textbook is available [here](#)  
An undergrad-rated textbook from the same authors is:  
Angrist, J. D. and J. S. Pischke, 2015. Mastering Metrics. Princeton University Press. Available at the library shortly.

Most algebra details can be found in chapter 21st of the textbook:

- Wooldridge, J. M. 2010. Econometric Analysis of Cross Section and Panel Data. Cambridge, MA: MIT Press.  
You can borrow this book from the library: [library's catalog reference](#)  
The 2002 edition of this book is freely available [here](#) .

Other superb textbooks:

- Hernán, M. and J. Robins, 2015. Causal Inference. Available online [here](#)
- Rosenbaum, Paul R. 2009. Design of Observational Studies. Springer-Verlag.  
Under the terms of the contract between Springer and UPV/EHU, you can read on-line and download this textbook from the university premises [clicking here](#)

A survey on methods:

- Imbens, Guido W. and Jeffrey M. Wooldridge. 2009. Recent developments in the econometrics of program evaluation. *Journal of Economic Literature* 47, no. 1: 5-86. Available online [here](#)

References

Articles (most of these are available from the university premises or using VPN)

- Abadie, Alberto & Diamond, Alexis & Hainmueller, Jens, 2010. "Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program," *Journal of the American Statistical Association* 105(490), 493-505. [available here](#)
- Abadie, A. and J. Gardeazabal. 2003. The Economic Costs of Conflict: A Case Study of the Basque Country. *American Economic Review* 93, 113-132. [Available online here](#)
- Angrist, J. (1990). Lifetime Earnings and the Vietnam Era Draft Lottery: Evidence from Social Security Administrative Records. *American Economic Review* 80, 313-336.
- Angrist, J. and A. Krueger (1991) "Does Compulsory School Attendance Affect Schooling and Earnings?" *Quarterly Journal of Economics* 106-4 979–1014. [Available online here](#)
- Brand, J. E. and Halaby, C. N., (2006). Regression and matching estimates of the effects of elite college attendance on educational and career achievement. *Journal Social Science Research* 35, 749-770. [Download this article from the university premises by clicking Available online here](#)
- Card, D. and A. B. Krueger (1994). Minimum Wages and Employment: A Case Study of the Fast-Food Industry in New Jersey and Pennsylvania. *American Economic Review* 84, 772-793. [Available online here](#)
- Dehejia, R. and S. Wahba (1999) "Causal Effects in Non-Experimental Studies: Re-Evaluating the Evaluation of Training Programs," *Journal of the American Statistical Association*, 1053–1062. [Available online here](#)
- Lalonde, R. (1986), "Evaluating the Econometric Evaluations of Training Programs," *American Economic Review*, 76, 604-620.
- Krueger, A. B. 1999. Experimental estimates of education production functions. *Quarterly Journal of Economics* 114, 497–532. [Available online here](#)
- Lee, D. (2008) "Randomized Experiments from Non-Random Selection in U.S. House Elections," *Journal of Econometrics* 142, 675–697. [Available online here](#)
- Ludwig, J. and D. L. Miller, (2007) "Does Head Start Improve Children's Life Chances? Evidence from a Regression Discontinuity Design" *The Quarterly Journal of Economics* 122 (1): 159-208. [Available online here](#)
- Lalive, R. (2008), How do extended benefits affect unemployment duration? A regression discontinuity approach, *Journal of Econometrics* 142, 785-806. [Available online here](#)

- Battistin, E., A. Brugiavini, E. Rettore and G. Weber, (2009), The Retirement Consumption Puzzle: Evidence from a Regression Discontinuity Approach, *American Economic Review* 99, 2209-26.

A web site on the Scientific Method I recommend to visit:

- Professor William M.K. Trochim, Cornell University: web site

A complete course on causal inference on-line:

- Professors Joshua Angrist and Victor Chernozhukov, Massachusetts Institute of Technology: web site

Easy to read policy evaluation manuals:

- Paul J. Gertler, Sebastian Martinez, Patrick Premand, Laura B. Rawlings, Christel M. J. Vermeersch (2011). “Impact Evaluation in Practice” The International Bank for Reconstruction and Development / The World Bank. Available online here
- Shahidur R. Khandker Gayatri B. Koolwal Hussain A. Samad (2010). “Handbook on Impact Evaluation Quantitative Methods and Practices” The World Bank. Available online here

## Web site

The web site for this course is hosted at eGela. You must be registered for the course to get access, I can however grant access to students and scholars sitting in the course. The web site contains this syllabus, the slides of the presentations used for lecturing, the assignments, data sets for you to use with the assignments, chats for interaction among participants, announcements and other materials we will use during the course.

## Lectures and learning methods

We will meet twice a week for two hours. In class, I will use presentations for lecturing, pdf versions of which will be available in eGela. Class participation is greatly encouraged, do not hesitate to stop me whenever you need a clarification or simply do not understand. You are supposed to read the textbook and other readings. The examples covered in class are borrowed from the assigned readings, so the discussion in class will guide you through the articles. In addition to the standard lecture method, we will strongly rely on the peer learning method. With that aim, you are expected to form study groups and discuss the materials covered in class regularly. All members of the study group will take turns in leading study sessions, that is, explaining the materials covered in class or rising questions to the other group members. If you need help, I will be available to gladly answer any question you might have. We will meet at least a couple of times at the computer lab (a detailed schedule will be posted in eGela). There, I will guide you through the computing requirements for your assignments and stand ready to answer your questions.

## Assignments

There will be two assignments. The assignments are meant to train you in the empirical methods providing a “hands on” experience to complement the theoretical methods covered in class. Both assignments will have the same structure. You will be asked to estimate some models and treatment effects using data I will provide. I expect you to use `stata` statistical package, but anyone able and willing to use other statistical packages (SPSS, SAS or the like) or programming languages (R, S-plus, Matlab,..) are welcome to do so. However, you must be aware of the extra difficulty this might imply, as some of the specific routines we will be using might not be available in other packages or languages. You can work in groups for the assignments, but everyone must submit a written version of the assignments.

## Exam

There will be one closed-book two-hour final exam. The date, time and room of the final exam will be announced by the program coordinator. The last day of classes I will hand out a list of about a hundred questions/problems of which I will select a handful for the exam.

## Grades

The grade of this course will be a weighted average of the grades obtained in the assignments (15% each) and the final exam (70%).