



## Workshop on Causal Inference

Organised by Nina Liljeqvist & Pedro Riera

Monitoring professor: Fabrizio Bernardi

Instructor: Elias Dinas (Oxford University)

28 May - 1 June 2012, Emeroteca, Badia Fiesolana

Please register with [Monika.Rzemieniecka@eui.eu](mailto:Monika.Rzemieniecka@eui.eu)

### Objectives and Learning Outcomes

*“Whenever, looking at my watch, I see the hand has reached the figure X, I hear the bells beginning to ring in the church close by. But from the fact that the watch hands point to ten whenever the bells begin to ring, I have not the right to infer that the position of the hands of my watch is the cause of the vibration of the bells.”*

[Leo Tolstoy, *War and Peace*, trans. Constance Garnett (New York: Modern Library Classics, 2002), p. 939]

Do hospitals make people healthier? Is it a problem that more people die in hospitals than in bars? Does an additional year of schooling increase future earnings? Do parties that enter the parliament enjoy vote gains in subsequent elections? The answers to these questions (and many others which affect our daily life) involve the identification and measurement of causal links: an old problem in philosophy and statistics. To address this problem we either use experiments or try to mimic them by collecting information on potential factors that may affect both treatment assignment and potential outcomes. Customary ways of doing this in the past entailed the specification of sophisticated versions of multivariate regressions. However, it is by now well understood that causality can only be dealt with during the design, not during the estimation process. The goal of this workshop is to familiarize participants with the logic of causal inference, the underlying theory behind it and introduce research methods that help us approach experimental benchmarks with observational data. Hence, this will be a much applied workshop where we aim to provide researchers with ideas for better research designs in their own work. Following the identification strategies discussed in the workshop, questions about how many Zs to include as controls will henceforth seem irrelevant.



### **Logistics: Format, software and data**

Classes will meet five days for one two-hour sessions each day, from 11-13:00. In addition, on Days 2 through 5 a lab session will be held from 10:00-11:00 before the course. The homework clinic will be optional, but researchers will have to attend all the theoretical sessions and hand in the assigned homework in order to obtain the credits.

The class will employ the CITRIX statistical package, focusing on ivregress, ivprobit, gmm, rd, atnd, qqplot, sensatt, mahapick, rdob (add-on package).

Researchers are encouraged to bring their own data to the class. As far as possible homework will be organized around researchers' own datasets so that the class can be maximally beneficial in helping researchers to overcome data analysis problems relevant to their own research.

### **Schedule**

Day 1: Introduction to the fundamental problem of causal inference and the experimental ideal. Introduction to the Neyman-Rubin potential outcomes framework.

Day 2: Instrumental Variables

Day 3: Regression Discontinuity Design

Day 4: Matching and selection models

Day 5: Wrap-up, miscellaneous, and researchers' practical problems with their own data.