# FGV Sao Paulo School of Business

# Causal Inference in Strategy Research (2017-II)
# Métodos Causais na Pesquisa em Estratégia (2018-II)

Instructor: Rodrigo Bandeira-de-Mello

Release: **February 27, 2018**

**Time & Room**

Classes: From July 30 to August 3
Sessions: 8am-11pm and 12pm-3pm
Room TBD

**Office & Email**

Room 1105, Itapeva 474
Email: rodrigo.mello@fgv.br

## Overview and Course Goals

The major sources of data in strategy research comes from natural observations of sample units in their own settings. This is why empirical research in strategy has increasingly made use of sophisticated methods to overcome the major drawbacks of inferring causality from observational studies. This seminar covers the main designs and inference methods suitable for causal effect identification in observational studies. This content is an extension of the actual courses on quantitative methods in our graduate program. I address the topics of this course from a practical point of view, not from a purely statistical analysis. The statistical notation used here is sufficient to make the researcher more confident when discussing the "tricks of the trade" of various methods. The class is open to all qualified students from other research streams other than strategy.

I expect that, by the end of this course, you will be able:

- to propose creative designs to identify causal effects for major problems in strategy research;
- to critically analyze the existing publications that aims at testing causality;
- to compute estimates for causal effects using R.

## Prerequisites

These are the three prerequisites for this course:

- Research methods: proposing research questions, deriving hypotheses, identifying the basic research designs in quantitative research. These topics are covered in the course "Métodos de Pesquisa", mandatory for all grad students.

- Statistics: correlation, partial correlation, OLS regression, hypothesis testing, probability distributions. These topics are covered in the courses "Análise Multivariada de Dados" and "Metodos Quantitativos de Pesquisa".
- Computation: familiarity with any statistical software. We will use R in this course (more on this below).

In order to help you to decide whether to take this course or not, I prepared the following self-assessment test. Please assign the most probable answer you give to each one of the three questions and then sum up the final score.

Question 1) Look at the equation below and assign your answer:

$$\hat{y} = \alpha + \sum_{i=1}^{n} \beta_i x + \epsilon \tag{1}$$

(0 pt) "I have no idea how to read this and what it implies".
(1 pt) "I can read it, and I guess what it is, but I do not know how to write one by myself".
(2 pts) "I can read it, understand it, and know how to write a new equation like this".

Question 2) Look at the table below and assign your answer:

**Table 2. Effects of resources and exchange conditions on interfirm cooperation**

| | Resources | | Exchange Conditions | |
|---|---|---|---|---|
| Age | 0.09 | 0.82 | 0.05 | 0.43 |
| Size | –0.05 | 0.43 | –0.43** | 3.12 |
| Growth | 0.10 | 0.86 | –0.07 | 0.67 |
| Brand name | –0.25* | 2.39 | | |
| TMT Experience | –0.01 | 0.12 | | |
| Slack Capital | –0.25* | 2.10 | | |
| Asset Specificity | | | 0.20* | 2.06 |
| Specific Knowledge | | | –0.23* | 2.23 |
| Geographic Dispersion | | | 0.39** | 2.85 |
| df | (6,87) | | (6,87) | |
| $R^2$ | 0.16 | | 0.24 | |
| $F$ | 2.83* | | 4.42*** | |

$N$ = 94
*$p$<0.05
**$p$<0.01
***$p$<0.001

(0 pt) "I have no idea how to read this and what it implies".
(1 pt) "I can interpret the main results".
(2 pts) "I can fully understand all tests performed in the table".

Question 3) Can I write a statistical software code to produce the table presented in Question 2?

(0 pt) "I have never used any statistical software before".
(1 pt) "I can produce the table only by using the software menus, but I never wrote a code".
(2 pts) "I can write the code, run it, and present the output".

If your total score is zero, I am afraid this course is not for you this semester. If your total score is between one and three, you are qualified to take the course but keep in mind that you will need additional work on some prerequisites. If you total score is greater or equal to four, then this course is the right one for you.

### Course Requirements

**Paper presentation (40%)**: Course sessions for each topic rely on theory and examples of applications. One important part of this course is to discuss strengths and weakness of the decisions made by the authors of selected applications. During the course, you will provide your own evaluation for one or more papers using, at least, the content of this course. Please prepare a presentation on the following topics: a) question and motivation; b) contribution; c) hypotheses (in a graphical representation, if possible); d) design and estimation methods; e) your personal assessment. One slide per topic is sufficient. Item e) is the most important item for grading purposes.

**First-week exam (30%)**: You will be asked to provide your interpretation of the R output tables for problems whose analyses used any of the methods covered in the first week. This evaluation will take place on the Friday of the first week during the class time. The use of the textbook will be allowed but not the use of notebooks.

**Final exam (30%)**: This final evaluation will take place on the last day of our course during class time. You will be asked to write, and handle to me, your R code that solve a practical problem assigned to you, as well as the interpretation of the results (the use of Latex is a plus). You will also analyze the empirical strategy of an application.

### Computation

I will teach the course using R software. You can download it for free here. This is a open-source software with great tutorials and resources available on line. Just google it. You need to use R with an integrated development environment (IDE), such as RStudio. You can downloaded RStudio for free here. A good suggestion are the tutorials provided by Dan Goldstein (tutorial 1 and tutorial 2) and DataCamp.

I will integrate R with Latex. Latex is a free typesetting software that produces high-quality, professional-looking manuscript. The integration of Latex with R, for instance, increases the productivity when writing-up the research paper. You can download the Texmaker 4.5 to use Latex in your computer `here`. You will find on youtube several tutorials on Texmaker. Some of them are `here`.

**Books**

Angrist J, Pischke JS. 2008. *Mostly Harmless Econometrics: An empiricist's companion.* Princeton University Press

**Course Sessions and Schedule**

**Session (1): Course Introduction**

- Course overview, requirements, and outline
- Introduction to R and Latex

*Readings:*

- Watch R and Latex tutorials before coming to class

**Session (2): Causality, Endogeneity, and Quasi-Experiments**

- The selection problem
- The potential outcome model
- Randomization and quasi-randomization

*Readings:*

- Angrist and Pischke (2008, chapter 1-2)
- Hamilton and Nickerson (2003)
- Chatterji *et al.* (2016)
- Suggested:
  - Sekhon and Titiunik (2012)
  - Bettis, Gambardella, Helfat, and Mitchell (2014)
  - Bettis *et al.* (2016)

**Session (3): Regression and Matching**

- Selection on observables
- The propensity score

*Readings:*

- Angrist and Pischke (2008, chapter 3)
- Application:
  - Paulo: Galilea, Gama, Bandeira-de Mello, and Marcon (2017)
  - Renato: Valentini (2012)
- Suggested:
  - Rubin (2001)
  - Caliendo and Kopenig (2005)
  - Imbens (2014)

**Session (4): Instrumental Variables**

- Local average treatment effects (LATE)
- The exclusion restriction and the “good" instrument
- Two-stage least squares

*Readings:*

- Angrist and Pischke (2008, chapter 4)
- Semadeni, Withers, and Certo (2014)
- Application:
  - Marina: Arreola and Bandeira-de Mello (2017)
  - Otavio: Castaner and Kavadiz (2013)
- Suggested:
  - Bascle (2008)
  - Acemoglu, Johnson, and Robinson (2000)

**Session (5): First-week exam**

**Session (6): Fixed Effects**

- Time-invariant (fixed) unobservables
- Practical considerations

*Readings:*

- Angrist and Pischke (2008, chapter 5)
- Certo, Withers, and Semadeni (2017)
- Application:
  - Renato: Chakrabarti, Singh, and Mahmood (2007)
  - Karina: Costa, Bandeira-de Mello, and Marcon (2013)
- Suggested:
  - Seamans (2013)

**Session (7): Differences-in-Differences**

- Exogenous "shocks" and interaction with treatment
- Practical considerations

*Readings:*

- Angrist and Pischke (2008, chapter 5)
- Application:
  - Karina: Lazzarini, Musacchio, Bandeira-de Mello, and Marcon (2015)
  - Paulo: Chatterji and Toffel (2010)
- Suggested:
  - Bertrand, Duflo, and Mullainathan (2004)

**Session (8): Regression-Discontinuity Design (Sharp)**

- Identification
- Assumptions
- Estimation
- Practical considerations

*Readings:*

- Angrist and Pischke (2008, chapter 6)
- Application:

  - Marina: Flamer and Bansal (2017)
  - Otavio: Bandeira-de Mello (2017)

- Suggested:
  - Boas, Hidalgo, and Richardson (2014)
  - Imbens and Lemieux (2008)
  - Hahn, Todd, and Klaauw (2001)

**Session (9): Wrap-up!**

**Session (10): Final Exam**

## References


Acemoglu D, Johnson S, Robinson J. 2000. The colonial origins of comparative development: an empirical investigation. *American Economic Review* **91**(5): 1369–1401.

Angrist J, Pischke JS. 2008. *Mostly Harmless Econometrics: An empiricist's companion.* Princeton University Press.

Arreola F, Bandeira-de Mello R. 2017. The effect of political ties heterogeneity on the internationalization of multinationals from emerging countries.

Bandeira-de Mello R. 2017. Leveraging the Winner: Corporate political action under resource-dependence heterogeneity.

Bascle G. 2008. Controlling for endogeneity with instrumental variables in strategic management research. *Strategic Organization* **6**(3): 285–327.

Bertrand M, Duflo E, Mullainathan S. 2004. How much should we trust differences-in-differences estimates? (February): 249–275.

Bettis R, Ethiraj S, Gambardella A, Helfat C, Mitchell W. 2016. Creating repeatable cumulative knowledge in strategic management: A call for broad and deep conversation among authors, referees, and editors. *Strategic Management Journal* **37**: 257–261.

Bettis R, Gambardella A, Helfat C, Mitchell W. 2014. Quantitive empirical analysis in strategic management. *Strategic Management Journal* **35**: 949–953.

Boas TC, Hidalgo FD, Richardson NP. 2014. The spoils of victory: Campaign donations and government contracts in Brazil. *The Journal of Politics* **76**(2): 415–429.

Caliendo M, Kopenig S. 2005. Some practical guidance for the implementation of propensity score matching. *IZA Discussion Paper No. 1588* .

Castaner X, Kavadiz N. 2013. Does good governance prevent bad strategy? A study of corporate governance, financial diversification, and value creation by French corporations, 2000-2006. *Strategic Management Journal* **34**: 863–876.

Certo ST, Withers M, Semadeni M. 2017. A tale of two effects: Using longitudinal data to compare within- and between-firm effects. *Strategic Management Journal* **38**: 1536–1556.

Chakrabarti A, Singh K, Mahmood I. 2007. Diversification and performance: Evidence from east asian firms. *Strategic Management Journal* **28**: 101–120.

Chatterji A, Findley M, Jensen N, Meier S, Nielson D. 2016. Field experiments in strategy research. *Strategic Management Journal* **37**: 116–132.

Chatterji A, Toffel M. 2010. How firms respond to being rated. *Strategic Management Journal* **31**: 917–945.

Costa M, Bandeira-de Mello R, Marcon R. 2013. Influência da conexão política na diversificação dos grupos empresariais brasileiros. *Revista de Administração de Empresas* **53**(4): 376–387.

Flamer C, Bansal P. 2017. Does a long-term orientation create value? Evidence from a regression discontinuity. *Strategic Management Journal* (Online EarlyView).

Galilea G, Gama M, Bandeira-de Mello R, Marcon R. 2017. Allocation of resources and earnings quality: does expropriation of minority shareholders exist in politically connected firms?

Hahn J, Todd P, Klaauw W. 2001. Identification and estimation of treatment effects with a regression-discontinuity design. *Econometrica* **69**(1): 201–209.

Hamilton B, Nickerson J. 2003. Correcting for endogeneity in strategic management research. *Strategic Organization* **1**(1): 51–78.

Imbens G. 2014. Matching methods in practice: Three examples.

Imbens G, Lemieux T. 2008. Regression discontinuity designs: A guide to practice. *Journal of Econometrics* **142**(2): 615–635.

Lazzarini S, Musacchio A, Bandeira-de Mello R, Marcon R. 2015. What do state-owned development banks do? Evidence from BNDES, 2002–09. *World Development* **66**(February): 237–253.

Rubin DB. 2001. Using propensity scores to help design observational studies: Application to the tobacco litigation. *Health Services and Outcomes Research Methodology* **2**(3-4): 169–188.

Seamans R. 2013. Threat of entry, asymmetric information, and pricing. *Strategic Management Journal* **34**: 426–444.

Sekhon JS, Titiunik R. 2012. When Natural Experiments Are Neither Natural nor Experiments. *American Political Science Review* **106**(01): 35–57.

Semadeni M, Withers M, Certo ST. 2014. The perils of endogeneity and instrumental variables in strategy research: Understanding through simulations. *Strategic Management Journal* **35**: 1070–1079.

Valentini G. 2012. Measuring the effect of M\&A on patenting quantity and quality. *Strategic Management Journal* **33**: 336–346.