

# MBR: Quantitative Methods

Harm Schuett

Winter 2016/2017

E-mail: [schuett@bwl.lmu.de](mailto:schuett@bwl.lmu.de)

Office Hours: tba

Office: Ludwigstraße 28RG / Room 423

Web: [svmiller.com/teaching](http://svmiller.com/teaching)

Class Hours: 9:00 am - 5:00 p.m.

Class Room: Ludwigstraße 28RG / Room 416

---

**This Syllabus will be updated continuously till January 10th 2017**

## Course Dates

Day	Date	Room	Time
1	Tue, 9. Jan, 2017	416	9 am - 5 pm
2	Thu, 11. Jan, 2017	416	11:30 am - 7 pm
3	Wed, 17. Jan, 2017	B-103	9 am - 5 pm
4	Fri, 19. Jan, 2017	416	9 am - 5 pm
5	Mon, 22. Jan, 2017	416	9 am - 5 pm
6	Mon, 29. Jan, 2017	416	9 am - 5 pm

## Course Description

This course is intended to teach you the basics and intermediate elements of the current state of empirical methods in economics. We will start with linear regression and work our way up to extensions such as fixed effects estimators, Difference-in-Difference estimators. We'll switch to matching estimators and highlight its link to normal regression analysis, then spend quite some time on instrumental variables approaches and similar approaches to causal inference. We will expand the linear regression framework to so called "general linear models" in order to analyze ordinal and nominal variables (e.g., logistic regression). The course will end with a glance at the advanced topics of GMM, Bayesian Statistics, and Shrinkage. Throughout the course, the focus, especially at the beginning, will be to develop a solid understanding of the key statistical concepts and tools, underlying modern empirical thinking: Identification, potential outcomes, causal graphs, inference in the presence of uncertainty. My hope is that if you grasp these you will:

1. have a clear idea how to setup up any empirical test (e.g., what estimator to choose, which variables to include, which **not** to include),
2. see the connections between all discussed approaches: how they are different ways of exploiting variation in the data based on what you assume about the underlying data,
3. be able to teach yourself any of the more advanced topics we did not have time to cover.

## Required Software

All of the course will be taught in R. The course handouts will be R-Notebooks (In fact the course syllabus is compiled from a R-Markdown document). So please, bring a laptop with the newest [R version](#) and [RStudio](#) installed. Both is free software, which is part of the reason why I chose to hold the course in R. The other is that R and Python have taken the Data community [by storm](#) during the last few years and you will likely need to know both in the future. Try to play around with R a bit before the course starts. Part of the course will be to teach you R. But it helps immensely if you try to find your way around before. In particular, figure out how to install additional packages from within RStudio (check for instance [R for Data Science](#)) and install the following packages:

- tidyverse
- lfe
- haven
- *(list will be updated)*

## Recommended Reading for Empiricists

You don't have to read these. But I encourage you to have access to at lease the first two. And if you will later do empirical work, you definitely want to have the first three ready at hand. (*Note, that Gelman and Hill are significantly revising their book at the Moment. A new version is expected some time in 2018*)

### Great Textbooks

Angrist, Joshua D and Jörn-Steffen Pischke (2008). *Mostly harmless econometrics: An empiricist's companion*. Princeton University Press.

Gelman, Andrew and Jennifer Hill (2007). *Data analysis using regression and multilevel hierarchical models*. Cambridge University Press.

Wooldridge, Jeffrey M (2010). *Econometric analysis of cross section and panel data*. MIT Press.

### Great Additional Reading

McElreath, Richard (2016). *Statistical rethinking: A Bayesian course with examples in R and Stan*. CRC Press.

Tufte, Edward R (1990). *Envisioning information*. Graphics press.

## Course Policy

### Attendance Policy

*Showing up is 80 percent of life* – Woody Allen, [via Marshall Brickman](#)

Class participation is necessary to understand the presented methods and to use them properly. MBR participants are required to attend at least 75% of the course hours to receive a passing grade. At the end of this course there will be an exam which is mandatory for all MBR participants.

## Grading Policy

The grade will be determined by an exam at the end of the course. *Details on the exam will follow in December*

## E-mail Policy

You can always write me an e-mail. I am usually quick to respond to student e-mails. If you have longer questions however, I would prefer if you would come to my office hours; it's a more efficient use of your and my time in that case. As a general rule, if you have an administrative question, please look at the syllabus first. I will sometimes choose not to answer emails with questions that should be clear from just looking at the final syllabus at the end of December.

## Class Schedule

### Tue, 9. Jan, 2017 - Day 1: Inference in the Presence of Uncertainty

#### *Part 1: Statistical Basics*

- Introduction to econometric and statistical thinking.
- The goal of most of scientific analysis: inference (vs. say prediction)

*Excercise part: Intro to R + simulation and real data excercises to understand inference under uncertainty*

#### *Part 2: Regression Analysis I*

- Introduction to linear regression models.
- Assumptions and properties of linear regression models.

*Excercise part: Paper case 1: replicating a study, reporting, interpreting, and analyzing regression output.*

### Thu, 11. Jan, 2017 - Day 2: Identification

**ATTENTION: Because of a conflict with another MBR course, we will start at 11:30 am and finish at 7:00 pm on Tuesday the 11th.**

#### *Part 1: Causality, Identification, Potential Outcomes*

- How to structure an empirical test using the potential outcomes and causal diagram frameworks
- Confounders, endogeneity, and correlated omitted variable bias

*Excercise part: Simulation and real data excercises to understand causes of consequences of confounding variables*

## *Part 2: Regression Analysis II*

- Regression model properties in the presence of confounders.
- Posttreatment and Pretreatment variables.
- The problem with controlling for posttreatment variables.
- How to design a regression test. How not to test for mediation.
- What variables to include and which **not** to include.

*Excercise part: simulations + replicating a study.*

### **Relevant literature**

Acharya, Avidit, Matthew Blackwell and Maya Sen (2016). “Explaining causal findings without bias: Detecting and assessing direct effects”. In: *American Political Science Review* 110.3, pp. 512–529.

## **Wed, 17. Jan, 2017 - Day 3: Matching Estimators and Panel Models**

**ATTENTION: This day will be held in a different room: Edmund-Rumpler-Str. 13, B-103. Sorry for the inconvenience.**

### *Part 1: Panel Fixed-Effecs Models*

- Introducing advantages and disadvantages of panel models
- Using panel models to control for unobserved confounders.

*Excercise part: Replicating a study and computig fixed-effects models.*

### *Part 2: Potential Outcomes and Control Groups*

- Revisiting the control group framework
- Selection-on-obervables via matching
- Modern matching procedures
- Comparison to linear regression
- Synthetic control groups

*Excercise part: Replicating a study and comparing matching and regression estimates.*

## **Fr, 19. Jan, 2017 - Day 4: Instrumental Variables**

### *Part 1: Instrumental Variables - The Claim to Fame of Econometrics*

- Instrumental Variables as a way to circumvent confounding variables
- The logic of instrumental variables via the causal path framework
- Estimation issued with instrumental variables.
- Why you should always look for instrumental variables and why you shouldn't

*Excercise part: tba.*

*Part 2: Stretch Block.*

- Reserved if some topics need more time
- Question session. Chance to review key parts of the course. Chance for you to ask questions.
- Bigger paper replication.

*Excercise part: tba.*

## **Mon, 22. Jan, 2017 - Day 5: Natural Experiments**

*Part 1: Diff-in-Diff*

- The ideal case: finding a natural experiment. Allmost as good as randomly assigned.
- Pre-/post treatment for treatment and control group in a natural experiment.
- The all-important common trends assumption.

*Excercise part: tba.*

*Part 2: Regression Discontinuity*

- Searching for distribution kinks and defendable comparisons
- Problems with RD designs.
- Examples.

*Excercise part: tba.*

## **Mon, 29. Jan, 2017 - Day 6: General Linear Models and Advanced Topics**

*Part 1: General linear models*

- Extending the regression framework to include ordinal and binary data.
- Logit and Probit models.
- Complications when interpreting Logit and Probit coefficients.
- Poisson Regression.
- Overdispersion.

*Excercise part: tba.*

*Part 2: GMM and Bayesian Statistics*

- What to do if one equation is not enough
- Intro to structural equation modelling

*Excercise part: tba.*