

# **Lecture Notes in Econometrics**

Erlend Dancke Sandorf

2024-07-03

# Table of contents

<b>Acknowledgements</b>	<b>3</b>
<b>Preface</b>	<b>4</b>
What is econometrics? . . . . .	4
<b>I Introduction to econometrics and economic data</b>	<b>5</b>
<b>1 Economic data</b>	<b>6</b>
1.1 Cross-sectional data . . . . .	6
1.2 Time-series data . . . . .	6
1.3 Panel data . . . . .	6
<b>Heteroskedasticity</b>	<b>7</b>
1.4 How to detect heteroskedasticity? . . . . .	7
<b>2 Graphical representation of data</b>	<b>8</b>
<b>References</b>	<b>9</b>
<b>II The Simple Linear Regression (SLR) model</b>	<b>10</b>
<b>III References</b>	<b>12</b>
<b>References</b>	<b>13</b>
<b>Appendices</b>	<b>14</b>
<b>A Installing R and RStudio</b>	<b>14</b>
A.1 Installing R . . . . .	14

# Acknowledgements

This book is the result of input and revisions based on input from collaborators, colleagues, and students taking my classes in econometrics.

# Preface

This book is a living document of how to do econometrics in R. It is the result of many years doing applied econometric research and teaching econometrics to students. This book contains many practical examples that I use extensively in my teaching to illustrate theoretical points as well as practical implications. In my teaching, I rely on the excellent introductory textbook by Jeffrey Wooldridge (Wooldridge 2019), and use and reproduce many of the examples from that book.

Lecture Notes in Econometrics assumes that you are familiar with R. If you are not familiar with R, you are strongly encouraged to get familiar before proceeding. Excellent resources include Grolemond (2014) and Wickham, Çetinkaya-Rundel, and Grolemond (2023).

## What is econometrics?

## **Part I**

# **Introduction to econometrics and economic data**

# **1 Economic data**

Different types of econometric models are needed for different types of economic data.

## **1.1 Cross-sectional data**

## **1.2 Time-series data**

## **1.3 Panel data**

# Heteroskedasticity

In this chapter, we will use the following packages:

```
library(tidyverse)
library(wooldridge)
library(lmtest)
```

We the constant variance assumption.

$$\text{Var}(u_i|X) = \sigma^2$$

Heteroskedasticity is a violation of the constant variance assumption and instead we have:

$$\text{Var}(u_i|X) = \sigma_i^2$$

## 1.4 How to detect heteroskedasticity?

We can test for heteroskedasticity using the Breusch-Pagan test. The null hypothesis is that there is no heteroskedasticity. The alternative hypothesis is that there is heteroskedasticity.

## 2 Graphical representation of data

In this chapter, we will use the following packages:

```
library(tidyverse)
library(wooldridge)
```

The importance of plotting your data cannot be understated. A good graph can reveal patterns in the data that guide model building or extreme observations that point to potential data problems. It is hard to build a model

```
model_1 <- lm(wage ~ exper, data = wage1)
```

```
rlang::f_lhs(model_1$call$formula)
```

wage



## References

## **Part II**

# **The Simple Linear Regression (SLR) model**

In this chapter, we will use the following packages:

```
library(tidyverse)
library(wooldridge)
```

```
model_1 <- lm(wage ~ exper, data = wage1)
```

```
rlang::f_lhs(model_1$call$formula)
```

wage

## **Part III**

# **References**

# References

- Grolemund, Garrett. 2014. *Hands-on Programming with r: Write Your Own Functions and Simulations*. O'Reilly Media, Inc.
- Wickham, Hadley, Mine Çetinkaya-Rundel, and Garrett Grolemund. 2023. *R for Data Science*. O'Reilly Media, Inc.
- Wooldridge, Jeffrey. 2019. *Introductory Econometrics - a Modern Approach*. 7th ed. South-Western College Publishing.

# **A Installing R and RStudio**

## **A.1 Installing R**

R is an open source and freely available language and environment for statistical computing and graphics.