Econometrics

Bakti Siregar, M.Sc

2023-11-29

Contents

Foreword

Welcome to the exciting and complex world of econometrics! This book is designed to provide an in-depth understanding of the basic concepts and practical applications in econometric analysis. Econometrics is a branch of economics that combines economic theory with mathematical and statistical methods to test economic hypotheses and make predictions about economic phenomena in the real world.

This book was prepared with the aim of providing a comprehensive guide, starting from the basics of econometrics to advanced applications relevant to the contemporary world of economics. Readers will be guided through basic concepts such as linear regression, the assumptions underlying econometric models, to sophisticated data analysis techniques.

Emphasis is placed on the application of econometrics in the context of modern economics, with relevant case studies and applicable examples. The book also integrates the latest developments in the field of econometrics, ensuring that readers gain insight into the latest approaches and methodological developments.

As the author, I hope that this book will not only be a valuable reference source for students, academics, and economic practitioners, but will also inspire further interest and understanding of the complexity of economic phenomena. I would like to thank everyone who has contributed and supported me in the process of writing this book.

Summary

Econometrics is a branch of economics that uses mathematical and statistical methods to analyze and test hypotheses in an economic context. It combines economic concepts with statistical tools to understand and model economic relationships.

Here are some common topics in econometrics:

- **Regression Models:** Econometrics often begins with the study of regression models, in which the relationship between one or more dependent and independent variables is explained. Regression models can be simple or complex depending on the data and problem at hand.

- **Regression Assumptions:** Regression analysis requires several assumptions, such as the homoscedasticity assumption (constant residual variance), the residual independence assumption, and the residual normality assumption.

- **Parameter Estimation:** Using empirical data, we can estimate parameters in a regression model to find out how well the model fits the data.

- **Hypothesis Testing:** Econometrics is used to test statistical hypotheses about model parameters. This helps determine whether the observed relationship is significantly different from zero.

- **Multicollinearity and Heteroscedasticity:** Some common problems in regression analysis involve multicollinearity (high correlation between independent variables) and heteroscedasticity (non-constant residual variance). Econometricians develop methods to deal with these problems.

- **Time Series Model:** Econometrics is also often used to analyze time series data, which includes observations taken sequentially over a certain time.

- **Applied Econometrics:** In addition to basic models, econometrics is also applied in various fields of economics, including finance, human resources, marketing, and others.

Writer

- **Bakti Siregar, M.Sc** is Head of the Study Program at the Department of Statistics, Matana University. Graduate of Masters in Applied Mathematics from National Sun Yat Sen University, Taiwan. He is also a lecturer and Data Scientist consultant at well-known companies such as JNE, Samora Group, Pertamina, and others. He has special enthusiasm in teaching Big Data Analytics, Machine Learning, Optimization, and Time Series Analysis in the fields of finance and investment. His expertise is also evident in the use of Statistical programming languages such as R Studio and Python. He applies the MySQL/NoSQL database system in learning data management, and is proficient in using Big Data tools such as Spark and Hadoop. Some of his projects can be seen at the following links: Rpubs, Github, Website, and Kaggle.

My Gratitude

Thank you to fellow academics and econometric practitioners who have shared their knowledge and insights, provided valuable input, and provided inspiration in preparing this material. I would also like to thank my family and friends who provided moral support and encouragement during the writing process. You are a source of strength and inspiration for me. The entire process of writing this was an incredible journey, and I hope this ebook will be of benefit to readers, both those who are studying econometrics for the first time and those who want to improve their understanding.

Finally, thank you to the publishing team who has worked hard in the production process of this ebook. Hopefully this ebook will be a useful reference source and can contribute to the development of econometric studies.

Feedback & Suggestions

All your input and feedback means a lot to us to improve this template in the future. For readers/users who wish to submit input and responses, please use the contact below!

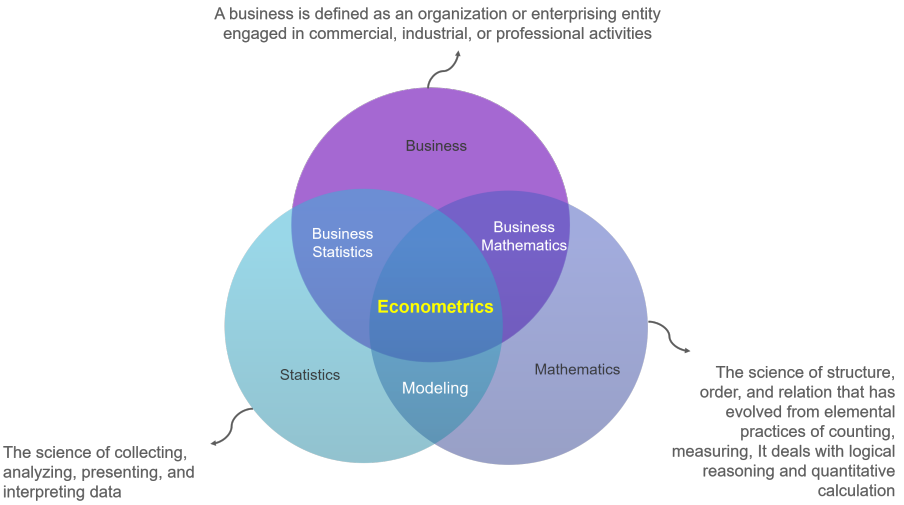
Email: dsciencelabs@outlook.com

Chapter 1

Introduction to Econometrics

1.1 What is Econometrics?

Econometrics is a branch of economics that combines economic concepts, mathematics, and statistics to analyze and measure empirical relationships in the economy. The term "econometrics" comes from combining the words "economics" and "metrics" (which means measurement). The main goal of econometrics is to develop mathematical models that can be used to explain and predict economic behavior, as well as test economic hypotheses and policies.



In practice, econometrics involves applying statistical models to examine relationships between economic variables. By using empirical data, econometricians (econometric practitioners) try to formulate mathematical models that can explain or predict economic behavior in the real world. The main method in econometrics involves statistical regression, in which economic variables are explained and related to each other.

Some basic concepts in econometrics involve:

- **Regression Model:** Create a mathematical model that describes the relationship between the dependent variable and the independent variable.
- **Hypothesis and Statistical Test:** Develop and test hypotheses about relationships between variables, and determine the extent to which analysis results are statistically reliable.
- **Prediction:** Uses an econometric model to predict the value of the dependent variable based on the value of the given independent variable.
- **Economic Policy Evaluation:** Analyze the effects of economic policies using econometric models.

Econometrics has an important role in economic and policy analysis. This allows economists to use empirical data and statistical methods to make more informational and evidence-based decisions in an economic context.

1.2 Goals of Econometrics

The main goal of econometrics is to develop mathematical models that can describe and explain the relationships between economic variables, as well as test economic hypotheses using empirical data. Some specific goals of econometrics involve:

- **Description and Measurement of Economic Relationships:** Econometrics aims to describe and measure the relationships between economic variables. This involves developing a mathematical model that reflects these relationships.
- **Economic Hypothesis Testing:** Econometrics is used to test economic hypotheses, such as whether there is a causal relationship between two variables or whether the effects of an economic policy are significant.
- **Forecasting and Prediction:** Econometrics provides tools for making forecasts and predictions about future economic behavior. By using the developed model, econometricians can estimate the value of the dependent variable based on the value of the given independent variable.
- **Economic Policy Evaluation:** Econometrics is used to evaluate the impact of economic policies. This allows economists to understand how changes in certain economic variables can affect other variables, thereby helping in policy decision making.

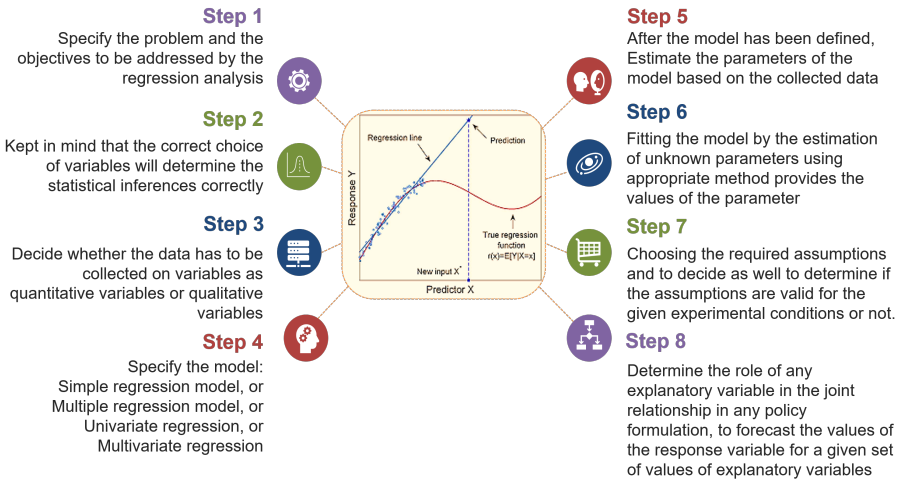
- **Understanding of Economic Processes:** Through statistical analysis, econometrics can help in deep understanding of complex economic processes. This includes identifying factors that influence the behavior of consumers, producers, and the market as a whole.
- **Empirical Research:** Econometrics provides tools for economic researchers to test economic theories using empirical data. This can help fill the gap between economic theory and the economic reality that occurs on the ground.
- **Evidence-Based Decision Making:** One of the main goals of econometrics is to provide an empirical basis for decision making in an economic context. By using data and statistical methods, econometrics helps make more informed and evidence-based decisions.

Thus, econometrics has an important role in providing powerful analytical tools for economists to understand and interpret economic phenomena in the real world.

1.3 Econometric Relations and Regression

Regression is a very important analytical tool in econometrics, helping researchers and analysts to measure and understand economic relationships based on empirical data.

- **Use of Regression in Econometrics:** Regression is one of the main tools in econometric analysis. Econometricians use regression techniques to estimate parameters and measure relationships between economic variables.
- **Regression Models in Econometrics:** In the context of econometrics, regression models are often used to describe and understand causal relationships between economic variables. This model can include economic variables such as income, prices, production, and others.
- **Hypothesis Testing in Regression Context:** Econometricians use regression to test economic hypotheses. For example, does an increase in variable A cause an increase or decrease in variable B.
- **Prediction and Forecasting:** Regression analysis is used in econometrics to make predictions and forecasting. By using a regression model, econometricians can estimate the value of the dependent variable based on the value of the independent variable.



1.4 Lab Practicum

1.4.1 Simple Regression

Suppose you have data on product sales (Y) and advertising costs (X) for several months. Perform a simple regression to understand the relationship between advertising costs and sales. Use the following dataset:

```
# Data
sales <- c(100, 120, 130, 140, 160)
advert <- c(50, 70, 90, 110, 130)

# Simple Regression Model
model1 <- lm(sales ~ advert)

# results
summary(model1)
```

```
##
## Call:
## lm(formula = sales ~ advert)
##
## Residuals:
##         1         2         3         4         5
## -2.00e+00  4.00e+00 -1.27e-14 -4.00e+00  2.00e+00
##
## Coefficients:
```

```
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  67.0000     5.4467    12.3   0.0012 **
## advert        0.7000     0.0577    12.1   0.0012 **
## ---
## Signif. codes:
## 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3.65 on 3 degrees of freedom
## Multiple R-squared:  0.98,   Adjusted R-squared:  0.973
## F-statistic:  147 on 1 and 3 DF,  p-value: 0.00121
```

Based on the regression results above, interpret the regression coefficient, R-squared value, and significance test to see whether advertising costs have a significant influence on sales.

1.4.2 Multiple Regression

Suppose you have a production dataset that includes production (Y), number of hours worked (X1), and number of machines (X2). Perform multiple regression to model the production relationship with these independent variables.

```r
# Data
product <- c(200, 250, 300, 280, 320)
work_hours <- c(40, 50, 60, 55, 65)
number_of_machines <- c(2, 3, 4, 3, 5)

# Multiple Regression model
model2 <- lm(product  ~ work_hours + number_of_machines)

# result
summary(model2)
```

```
## Warning in summary.lm(model2): essentially perfect
## fit: summary may be unreliable
```

```
##
## Call:
## lm(formula = product ~ work_hours + number_of_machines)
##
## Residuals:
##         1         2         3         4         5
##  2.13e-14 -4.25e-14 -1.43e-17  1.06e-14  1.06e-14
##
## Coefficients:
```

```
##                      Estimate Std. Error   t value
## (Intercept)        -2.00e+01   1.77e-13 -1.13e+14
## work_hours          6.00e+00   6.36e-15  9.44e+14
## number_of_machines -1.00e+01   5.36e-14 -1.87e+14
##                     Pr(>|t|)
## (Intercept)           <2e-16 ***
## work_hours            <2e-16 ***
## number_of_machines    <2e-16 ***
## ---
## Signif. codes:
## 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3.53e-14 on 2 degrees of freedom
## Multiple R-squared:     1,   Adjusted R-squared:     1
## F-statistic: 3.54e+30 on 2 and 2 DF,  p-value: <2e-16
```

Berdasarkan hasil regresi berganda di atas, identifikasi variabel yang memiliki pengaruh signifikan terhadap produksi. Interpretasikan koefisien regresi dan uji signifikan masing-masing variabel.

Catatan

- Pastikan untuk memberikan penjelasan yang jelas tentang interpretasi hasil, termasuk signifikansi statistik dan koefisien determinasi (R-squared).
- Anda dapat menambahkan visualisasi grafik, seperti plot regresi atau residu, untuk memperjelas analisis.
- Selalu periksa dan pastikan bahwa dataset yang digunakan sesuai dengan konteks soal dan telah diimport dengan benar ke dalam lingkungan R.

Chapter 2

Simple Regression

2.1 What is Regression?

Regression is a set of statistical methods used for the estimation of relationships between a dependent variable and one or more independent variables. It can be utilized to assess the strength of the relationship between variables by fitting a line to the observed data. Regression allows you to estimate how the dependent variable changes as the independent variable(s) change. There are several variations of Regression, such as linear, multiple linear, and nonlinear.

The most common models are simple linear and multiple linear. Nonlinear regression analysis is commonly used for more complicated data sets in which the dependent and independent variables show a nonlinear relationship.

2.2 Measures Of The Relationship

There are two quantitative measures of such relationships:

2.2.1 Covariance

In probability theory and statistics, covariance is a measure of the joint variability of two random variables.

$$cov(x, y) = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{n - 1}$$

Covariance between two variables: