# Introduction to econometrics *

Jean-Marie Dufour †
McGill University

First version: September 1981
Revised: July 2011, August 2011
This version: August 2011
Compiled: September 7, 2011, 4:09

* This work was supported by the William Dow Chair in Political Economy (McGill University), the Canada Research Chair Program (Chair in Econometrics, Université de Montréal), the Bank of Canada (Research Fellowship), a Guggenheim Fellowship, a Konrad-Adenauer Fellowship (Alexander-von-Humboldt Foundation, Germany), the Institut de finance mathématique de Montréal (IFM2), the Canadian Network of Centres of Excellence [program on *Mathematics of Information Technology and Complex Systems* (MITACS)], the Natural Sciences and Engineering Research Council of Canada, the Social Sciences and Humanities Research Council of Canada, and the Fonds de recherche sur la société et la culture (Québec).

† William Dow Professor of Economics, McGill University, Centre interuniversitaire de recherche en analyse des organisations (CIRANO), and Centre interuniversitaire de recherche en économie quantitative (CIREQ). Mailing address: Department of Economics, McGill University, Leacock Building, Room 519, 855 Sherbrooke Street West, Montréal, Québec H3A 2T7, Canada. TEL: (1) 514 398 8879; FAX: (1) 514 398 4938; e-mail: jean-marie.dufour@mcgill.ca . Web page: http://www.jeanmariedufour.com

# Contents

# 1. Definition and role of econometrics

## 1.1. Distinction with related disciplines

- Mathematical economics – Mathematical economics is economic theory formulated in mathematical language.
  Mathematical economics proposes relations between economic variables.

- Statistical economics – Statistical economics deals with problems associated with collecting and building economic data.

  - How to measure economic phenomena (*e.g.*, price indices, national accounts)?
  - How to collect data (*sample survey theory*)?

## 1.2. The role of econometrics

The role of econometrics is to link economic theory (especially, theory formulated in a mathematical language) with observation.

To do this, econometrics uses the methods of mathematical statistics.

1. Descriptive statistics: methods for describing, representing, summarizing data (exploration).

2. Statistical inference: methods for making generalizations from data, usually based on probability theory.

   (a) Estimation
   (b) Hypothesis tests
   (c) Prediction

### 1.3. Definition of econometrics

Econometrics is the discipline whose purpose is to empirically determine economic laws.
It is the discipline which allows us

starting from

- data,
- a model, which reflects
  - economic theory,
  - the way data were generated,

to pursue various scientific operations, such as:

- to estimate a model;
- to test hypotheses;
- to evaluate the adequacy of the model;
- to make predictions.

Laws studied in this way take two important forms:

1. a relation between two or several variables;
2. the distribution of one or several variables.

We can also distinguish:

1. theoretical econometrics,
2. applied econometrics.

## 2. Descriptive analysis

Two basic types of data are used in econometrics:

1. cross-sectional data;
2. time series.

Some data combine bots features: panel data.
Preliminary steps in the analysis of data.
Its is always **very important**, before a statistical analysis is performed,

**to look at the data.**

The investigator should have an idea of the behavior of the data, and check for the presence of surprising features.

Descriptive statistics supplies instruments for doing this.

Suppose we have $N$ observations on $k$ variables:

$$X_{1i}, X_{2i}, \ldots, X_{1i}, \quad i = 1, \ldots, N$$

Different observations may correspond to different individuals or units (cross sections), or to different time periods (time series).

## 2.1. Examination of individual data sets

Useful descriptive operations.

1. Histogram: tables and graphs

2. Maximum

3. Minimum

4. Measures of location:
$$\bar{X} = \frac{1}{N}\sum_{i=1}^{N} X_i \quad \text{(arithmetic mean)} \tag{2.1}$$

5. Dispersion measures

   (a) Range:
   $$R = \text{Max}(X_i) - \text{Min}(X_i)\,. \tag{2.2}$$

   (b) Sample variance:
   $$S_X^2 = \sum_{i=1}^{N}(X_i - \bar{X})^2/N\,. \tag{2.3}$$

   (c) Sample standard error:
   $$S_X = \sqrt{S_X^2}\,. \tag{2.4}$$

## 2.2. Graphs

1. Graph data as functions of time (for time series)

   (a) Types of behavior: trend (linear, exponential, etc.), cycles, random fluctuations?

   (b) Discontinuities?

   (c) Outlying observations? (data errors)

   (d) Comparisons between series: do different series behave in the same way?

2. If we consider that a variable $X_1$ may be explained by other variables $(X_2, X_3, \dots)$, graph $X_1$ as functions of $X_2, X_3, \dots$

   (a) Are the relationships linear? non-linear?

   (b) Discontinuities

   (c) Etc.