Chapter 1

An Introduction to Econometrics

1.1 Why Study Econometrics?

- Econometrics is a set of research tools also employed in the business disciplines of accounting, finance, marketing and management. It is also used by social scientists, specifically researchers in history, political science and sociology. Econometrics plays an important role in such diverse fields as forestry, and in agricultural economics.
- Studying econometrics fills a gap between being "a student of economics" and being "a practicing economist."
- By taking this introduction to econometrics you will gain an overview of what econometrics is about, and develop some "intuition" about how things work.

1.2 What is Econometrics?

- In economics we express our ideas about relationships between economic variables using the mathematical concept of a function.
- For example, to express a relationship between income i and consumption c, we may write

$$c = f(i)$$

- The demand for an individual commodity, say the Honda Accord, might be expressed as

$$q^d = f(p, p^s, p^c, i)$$

- The quantity of Honda Accords demanded, q^d, is a function $f(p, p^s, p^c, i)$ of the price of Honda Accords p, the price of cars that are substitutes p^s, the price of items that are complements p^c, like gasoline, and the level of income i.

- The supply of an agricultural commodity such as beef might be written as
-

$$q^s = f(p, p^c, p^f)$$

- q^s is the quantity supplied, p is the price of beef, p^c is the price of competitive products in production (for example, the price of hogs), and p^f is the price of factors or inputs (for example, the price of corn) used in the production process.

Econometrics is about how we can use economic, business or social science theory and data, along with tools from statistics, to answer "how much" type questions.

1.2.1 Some Examples

- A question facing Alan Greenspan is "*How much* should we increase the discount rate to slow inflation, and yet maintain a stable and growing economy?" The answer will depend on the responsiveness of firms and individuals to increases in the interest rates and to the effects of reduced investment on Gross National Product. The key elasticities and multipliers are called **parameters**. The values of economic parameters are unknown and must be estimated using a sample of economic data when formulating economic policies.
- Econometrics is about how to best estimate economic parameters given the data we have. "Good" econometrics is important, since errors in the estimates used by policy makers such as the FRB may lead to interest rate corrections that are too large or too small, which has consequences for all of us. Other examples include:

- A city council ponders the question of how much violent crime will be reduced if an additional million dollars is spent putting uniformed police on the street.
- U.S. Presidential candidate Gore questions how many additional California voters will support him if he spends an additional million dollars in advertising in that state.
- The owner of a local Pizza Hut franchise must decide how much advertising space to purchase in the local newspaper, and thus must estimate the relationship between advertising and sales.
- Louisiana State University must estimate how much enrollment will fall if tuition is raised by $100 per semester, and thus whether its revenue from tuition will rise or fall.
- The CEO of Proctor & Gamble must estimate how much demand there will be in ten years for the detergent Tide, as she decides how much to invest in new plant and equipment.

- A real estate developer must predict by how much population and income will increase to the south of Baton Rouge, Louisiana, over the next few years, and if it will be profitable to begin construction of a new strip-mall.
- You must decide how much of your savings will go into a stock fund and how much into the money market. This requires you to make predictions of the level of economic activity, the rate of inflation and interest rates over your planning horizon.

1.3 The Econometric Model

- When studying Honda car sales we recognize that the *actual* number of Hondas sold is the sum of this systematic part and a random and unpredictable component, *e*, that we will call a **random error**.
- An **econometric model** representing the sales of Honda Accords is

$$q^d = f(p, p^s, p^c, i) + e$$

- The random error *e* accounts for the many factors that affect sales that we have omitted from this simple model, and it also reflects the intrinsic uncertainty in economic activity.
- To complete the specification of the econometric model, we must also say something about the form of the algebraic relationship among our economic variables. For

example, in your first economics courses quantity demanded was depicted as a *linear* function of price. We assume that the systematic part of the demand relation is linear

$$f(p,\ p^s,\ p^c, i) = \beta_1 + \beta_2 p + \beta_3 p^s + \beta_4 p^c + \beta_5 i$$

- The corresponding econometric model is

$$q^d = \beta_1 + \beta_2 p + \beta_3 p^s + \beta_4 p^c + \beta_5 i + e$$

1.4 How Do We Obtain Data?

1.4.1 Experimental Data

- In an ideal world, from a researcher's point of view, an economic model would describe how we might design an experiment that could be used to obtain economic observations or sample information, that then could be used to provide insights about the unknown economic parameters. Repeating the experiment T times would create a sample of T sample observations

- In an ideal world for research, controlled experiments could be conducted to investigate the relationship between these "explanatory" variables and the "dependent" variable q^d. For example, we might set

  p = price of Accords = \$25,000

  p^s = price of Maximas = \$25,000

  p^c = price of gasoline per gallon = \$1.35

  i = income of individuals in sample = \$42,000

- At the end of a month we *observe* the number of Honda Accords sold at one dealership to be $q^d = 37$.
- By repeating this process a number of times, a *sample* of economic data is created.

- The experimental outcome we observe is the sum of a systematic component that depends on the controlled explanatory variables and this random noise. Consequently, the experimental outcome is random too.
- If we repeat the experiment a few times, we obtain observations such as those in Table 1.1.

Table 1.1 Monthly Sales of Honda Accords					
month	q^d	p	p^s	p^c	i
1	37	25	25	1.35	42
2	28	23	25	1.35	32
3	30	23	25	1.35	32
4	35	20	25	1.35	35
5	40	25	27	1.35	42
6	46	25	30	1.35	52
7	52	25	32	1.35	55
8	55	25	25	1.40	60
9	60	25	25	1.45	70
10	65	25	25	1.55	100

- We, the econometricians, wish to determine the relationship between each of the explanatory variables (p, p^c, p^s, i) and the dependent variable q^d using a sample of T observations on values of the explanatory variables and the observed quantities sold.

1.4.2 Nonexperimental Data

Most economic data are collected for administrative rather than research purposes, often by government agencies. The data may be collected in a:

- *time series form*—data collected over discrete intervals of time.
- *cross section form*—data collected over sample units in a particular time period
- *panel data form*—data that follow individual micro-units over time.

These data may be collected at various levels of aggregation:

- *micro*—data collected on individual economic decision making units such as individuals, households or firms.
- *macro*—data resulting from a pooling or aggregating over individuals, households or firms at the local, state or national levels.

The data collected may also represent a flow or a stock:

- *flow*—outcome measures over a period of time, such as the consumption of gasoline during the last quarter of 1999.
- *stock*—outcome measured at a particular point in time, such as the quantity of crude oil held by Chevron in its US storage tanks April 1, 1999

The data collected may be quantitative or qualitative:

- *quantitative*—outcomes such as prices or income that may be expressed as numbers or some transformation of them, such as real prices or per capita income.
- *qualitative*—outcomes that are of an "either-or" situation. For example, a consumer either did or did not make a purchase of a particular good, or a person either is or is not married.

1.5 Statistical Inference

The phrase **statistical inference** will appear often in this book. By this we mean we want to "infer" or learn something about the real world by analyzing a sample of data. The ways in which statistical inference are out carried include:

- Estimating economic parameters, such as elasticities, using econometric methods.

- Predicting economic outcomes, such as the enrollment in 2-year colleges in the U.S. for the next 10 years.
- Testing economic hypotheses, such as the question of whether newspaper advertising is better than store displays for increasing sales.

1.6 A Research Format

1. It all starts with a problem or question
2. Economic theory gives us a way of thinking about the problem: What economic variables are involved and what is the possible direction of the relationship(s)?
3. The working economic model leads to an econometric model. We must choose a functional form and make some assumptions about the nature of the error term.
4. Sample data are obtained, and a desirable method of statistical analysis chosen, based on our initial assumptions, and our understanding of how the data were collected.
5. Estimates of the unknown parameters are obtained with the help of a statistical software package, predictions are made and hypothesis tests are performed.
6. Model diagnostics are performed to check the validity of assumptions we've made. For example, were all of the right-hand-side explanatory variables relevant? Was the correct functional form used?

7. The economic consequences and the implications of the empirical results are analyzed and evaluated. What economic resource allocation and distribution results are implied, and what are their policy-choice implications? What remaining questions might be answered with further study or new and better data?