CHAPTER ONE

Introduction

In this chapter I provide an introductory discussion of what econometrics is and what econometricians do. Next, I consider a more detailed motivation for writing this book. Finally, I give an outline of the other chapters of the book.

What is econometrics?

Econometric techniques are usually developed and employed for answering practical questions. As the first five letters of the word "econometrics" indicate, these questions tend to deal with economic issues, although applications to other disciplines are widespread. The economic issues can concern macroeconomics, international economics, and microeconomics, but also finance, marketing, and accounting. The questions usually aim at a better understanding of an actually observed phenomenon and sometimes also at providing forecasts for future situations. Often it is hoped that these insights can be used to modify current policies or to

put forward new strategies. For example, one may wonder about the causes of economic crises, and if these are identified, one can think of trying to reduce the effects of crises in the future. Or, it may be interesting to know what motivates people to donate to charity, and use this in order to better address prospective donors. One can also try to understand how stock markets go up – and, particularly, how they go down – in order to adjust investment decisions.

The whole range of econometric methods is usually simply called "econometrics," and this will also be done in this book. And anyone who either invents new econometric techniques, or applies old or new techniques, is called an "econometrician." One might also think of an econometrician as being a statistician who investigates the properties particular to economic data. Econometrics can be divided into *econometric theory* and *applied econometrics*. Econometric theory usually involves the development of new methods and the study of their properties. Applied econometrics concerns the development and application of tools to solve relevant practical questions.

In order to answer practical questions, econometric techniques are applied to actually observed data. These data can concern (1) observations over time, like a country's GDP when measured annually, (2) observations across individuals, like donations to charity, or (3) observations over time and over individuals. Perhaps "individuals" would be better phrased as "individual cases," to indicate that these observations can also concern countries, firms, or households, to

mention just a few. Additionally, when one thinks about observations over time, these can concern seconds, days, or years.

Sometimes the relevant data are easy to access. Financial data concerning, for example, stock markets, can be found in daily newspapers or on the internet. Macroeconomic data on imports, exports, consumption, and income are often available on a monthly basis. In both cases one may need to pay a statistical agency in order to be able to download macroeconomic and financial indicators. Data in marketing are less easy to obtain, and this can be owing to issues of confidentiality. In general, data on individual behavior are not easy and usually are costly to obtain, and often one has to survey individuals oneself.

As one might expect, the type of question that one intends to answer using an econometric method is closely linked to the availability of actual data. When one can obtain purchase behavior of various households, one can try to answer questions about this behavior. If there are almost no data, there is usually not much to say. For example, a question like **"how many households will use this new product within 10 years from now?"** seems rather difficult to answer. And, **"what would the stock market do next year?"** is complicated, too. Of course, one can always come up with an answer, but whether one would have great confidence in this answer is rather doubtful. This touches upon a key aspect of the application of econometric techniques, which is that one aims at answering questions with *some*

degree of confidence. In other words, econometricians do not provide answers like "yes" or "no," but instead one will hear something like "with great confidence we believe that poor countries will not catch up with rich countries within the next 25 years." Usually, the size of "great" in "great confidence" is a choice, although a typical phrase would be something like "with 95 per cent confidence." What that means will become clear in chapter 2 below.

The econometrician uses an *econometric model*. This model usually amounts to one or more equations. In words, these equations can be like "the probability that an individual donates to charity is 0.6 when the same individual donated last time and 0.2 when s/he did not," or "on average, today's stock market return on the Amsterdam Exchange is equal to yesterday's return on the New York Stock Exchange," or "the upward trend in Nigeria's *per capita* GDP is half the size of that of Kenya." Even though these three examples are hypothetical, the verbal expressions come close to the outcomes of actual econometric models.

The key activities of econometricians can now be illustrated. First, an econometrician needs to *translate a practical question* like, for example, **"what can explain today's stock market returns in Amsterdam?"** *into a model*. This usually amounts to thinking about the economic issue at stake, and also about the availability and quality of the data. Fluctuations in the Dow Jones may lead to similar fluctuations in Amsterdam, and this is perhaps not much of a surprise. However, it is by no means certain that this is best

observed for daily data. Indeed, perhaps one should focus only on the first few minutes of a trading day, or perhaps even look at monthly data to get rid of erratic and irrelevant fluctuations, thereby obtaining a better overall picture. In sum, a key activity is to translate a practical question into an econometric model, where this model also somehow matches with the available data. For this translation, econometricians tend to rely on mathematics, as a sort of language. Econometricians are by no means mathematicians, but mathematical tools usually serve to condense notation and simplify certain technical matters. First, it comes in handy to know a little bit about matrix algebra before taking econometrics courses. Note that in this book I will not use any such algebra as I will just stick to simple examples. Second, it is relevant to know some of the basics of calculus, in particular, differential and integral calculus. To become an econometrician, one needs to have some knowledge of these tools.

The second key activity of an econometrician concerns the *match of the model with the data*. In the examples above, one could note numerical statements such as "equal" or "half the size." How does one get these numbers? There are various methods to get them, and these are collected under the header "estimation." More precisely, these numbers are often associated with unknown parameters. The notion "parameter estimation" already indicates that econometricians are never certain about these numbers. However, what econometricians can do is to provide a certain degree of

confidence around these numbers. For example, one could say that **"it is very likely that growth in *per capita* GDP in Nigeria is smaller than that of Kenya"** or that **"it is unlikely that an individual donates to charity again if s/he did last time."** To make such statements, econometricians use statistical techniques.

Finally, a third key activity concerns the *implementation of the model outcomes*. This may mean the construction of *forecasts*. It can also be possible to simulate the properties of the model and thereby examine the effects of various policy rules.

To summarize, econometricians use economic insights and mathematical language to construct their econometric model, and they use statistical techniques to analyze its properties. This combination of three input disciplines ensures that courses in econometrics are not the easiest ones to study.

In this book I try to introduce the essentials of econometrics to novices, keeping the mathematical and statistical level at a minimum, but without being condescending. This book can be used prior to any textbook on econometrics, but it should certainly not replace it! The intention is that this book should be used as introductory and supplementary reading. For textbooks on econometrics, one can choose from Verbeek (2000), Koop (2000), Gujarati (1999), Kennedy (1998), Ramanathan (1997), Johnston and Dinardo (1996), Griffiths, Hill and Judge (1993), and Goldberger (1991) at the introductory level, from Heij *et al.* (2002), Ruud (2000),

Greene (1999), Wooldridge (1999), and Poirier (1995), at the intermediate level, and from White (2000), Davidson and MacKinnon (1993), and Amemiya (1985), at the advanced level. For more specific analysis of time series, one can consider Franses (1998), Hamilton (1994), and Hendry (1995), and for financial econometrics, see Campbell, Lo and MacKinlay (1997).

So, do you have any interest in reading more about econometrics? If you are really a novice, then you can perhaps better skip the next section as this is mainly written for colleagues and more experienced econometricians. The final section is helpful, though, as it provides an outline of subsequent chapters.

Why this book?

Fellow econometricians may now wonder why I decided to write this book in the first place. Well, the motivation was based on my teaching experience at the Econometric Institute of the Erasmus University Rotterdam, where we teach econometrics at undergraduate level. My experience mainly concerns the empirical projects that undergraduate students have to do in their final year before graduation. For these projects, many students work as an intern, for example, with a bank or a consultancy firm, and they are supposed to answer a practical question which the supervising manager may have. Typically, this manager knows that econometricians can handle empirical data, and usually

they claim to have available abundant data. Once the student starts working on the project, the following scenario is quite common. The manager appears not to have an exact question in mind, and the student ends up not only constructing an econometric model, but also precisely formulating the question. It is this combination that students find difficult, and indeed, a typical question I get is "how do I start?"

Observing this phenomenon, I became aware that many econometric textbooks behave as if the model is already given from the outset, and it seems to be suggested that the only thing an econometrician needs to do is to estimate the unknown parameters. Of course, there are many different models for different types of data, but this usually implies that textbooks contain a range of chapters treating parameter estimation in different models (see also Granger, 1994). Note that more recent textbooks also address the possibility that the model may be inappropriate and therefore these books contain discussions about diagnostic checks.

Of course, to address in a single textbook all the practical steps that one can take seems like an impossible enterprise. However, it should be possible to indicate various issues other than parameter estimation that arise when one wants to arrive at a useful econometric model. Therefore, in chapter 3 I will go through various concerns that econometricians have when they aim to answer a practical question. This is not to say that parameter estimation is unimportant.

I merely aim to convey that in practice there is usually no model to begin with!

Without wishing to go into philosophical discussions about econometrics, it seems fair to state that the notion of "a model given from the outset" dates back to the first developments in econometrics. In the old days (like, say, fifty years ago), econometricians were supposed to match (mainly macro-) economic theories to data, often with an explicit goal to substantiate the theory. In the unlucky event that the econometric model failed to provide evidence in favor of the theory, it was usually perceived that perhaps the data were wrong or the estimation method was incorrect, implying that the econometrician could start all over again.

A format of a typical econometrics textbook has its origin in this traditional view of econometrics. This view assumes that most aspects of a model, like the relevant variables, the way they are measured, the data themselves, and the functional form, are already available to the econometrician, and the only thing s/he needs to do is to fit the model to the data. The model components are usually assumed to originate from an (often macro-) economic theory, and there is great confidence in its validity. A consequence of this confidence is that if the data cannot be summarized by this model, the econometric textbook first advises us to consider alternative estimation techniques. Finally, and conditional upon a successful result, the resultant empirical econometric model is used to confirm (and perhaps in some cases,

to disconfirm) the thoughts summarized in the economic theory. See Morgan (1990, 2002) for a detailed analysis of the development of econometric ideas.

There are several reasons why this traditional view is losing territory. The first is that there is a decreasing confidence in the usefulness of econometric models to confirm or disconfirm economic theories. Summers (1991) convincingly argues that important new macroeconomic insights can also be obtained from applying rather simple statistical techniques, and that the benefit of considering more complicated models is small. Granger (1999) gives a lucid illustration of the fact that the implications of even a simple economic theory are hard to verify.

With an increased application of econometric methods in finance and marketing, there also seems to be a need for teaching econometrics differently. The main reason for this need is that it is usually impossible to have strong prior thoughts about the model. Also, these modern application areas require new models, which are suggested by the data more than by a theory – see Engle (1995), Wansbeek and Wedel (1999), for example. Hence, an econometrician nowadays uses the data and other sources of information to construct the econometric model. With this stronger emphasis on the data, it becomes important to address in more detail the specification of a model, the evaluation of a model, and its implementation. The evaluation part is relevant for obtaining confidence in the outcomes. It