CENGAGE Learning

计量经济学导论

（第四版）

高等学校经济学类英文版教材

普通高等教育"十一五"国家级规划教材

Jeffrey M. Wooldridge

# Introductory Econometrics

4th Edition

王少平　改编

李子奈　审校

高等教育出版社

CENGAGE Learning 圣智学习

# Introductory Econometrics

CENGAGE Learning

# 计量经济学导论

（第四版）

高等学校经济学类英文版教材

普通高等教育“十一五”国家级规划教材

Jeffrey M. Wooldridge

# Introductory Econometrics

## 4th Edition

王少平　改编
李子奈　审校

JILIANG JINGJIXUE DAOLUN

高等教育出版社·北京
HIGHER EDUCATION PRESS　BEIJING

**图字：01-2013-9060号**

Introductory Econometrics, Fourth Edition
Wooldridge



ISBN: 978-7-04-039594-5

Cengage Learning Asia Pte. Ltd
151 Lorong Chuan, #02-08 New Tech Park, Singapore 556741



**图书在版编目（CIP）数据**

计量经济学导论 ：第4版 = Introductory econometrics, 4e ：英文 /（美）伍德里奇(Wooldridge, J. M.) 著 ；王少平改编. -- 北京 ：高等教育出版社，2014.5
ISBN 978-7-04-039594-5

Ⅰ. ①计… Ⅱ. ①伍… ②王… Ⅲ. ①计量经济学-高等学校-教材-英文 Ⅳ. ①F224.0

中国版本图书馆CIP数据核字(2014)第067566号

策划编辑 施春花　　责任编辑 施春花　　封面设计 张　楠　　版式设计 王艳红
插图绘制 李大鹏　　责任印制 刘思涵

| | | | |
|---|---|---|---|
| 出版发行 | 高等教育出版社 | 咨询电话 | 400-810-0598 |
| 社　　址 | 北京市西城区德外大街4号 | 网　　址 | http://www.hep.edu.cn |
| 邮政编码 | 100120 | | http://www.hep.com.cn |
| 印　　刷 | 北京明月印务有限责任公司 | 网上订购 | http://www.landraco.com |
| 开　　本 | 787mm×1092mm　1/16 | | http://www.landraco.com.cn |
| 印　　张 | 31 | 版　　次 | 2014年5月第1版 |
| 字　　数 | 780千字 | 印　　次 | 2014年5月第1次印刷 |
| 购书热线 | 010-58581118 | 定　　价 | 59.00元 |


物 料 号　39594-00

# 内容简介

本书为 Wooldridge 所著的 *Introductory Econometrics—A Modern Approach, Fourth Edition* 的英文改编版教材。改编后的教材内容简洁、逻辑清晰、篇幅与深度适当，并且具有比较完整的知识体系，符合我国高等学校计量经济学的本科教学需求。

改编后的教材集中于计量经济学的主流框架，加强了基础性理论，适当弱化了应用。具体分为四个部分：一是基于横截面数据的模型、最小二乘估计（OLS）和假设检验及其应用；二是时间序列数据的模型设定、估计和检验理论与应用；三是面板数据模型的理论和应用；四是离散选择模型或者微观计量经济学，用于研究个体选择的决定因素。

本书可作为高等学校经济学类、管理学类本科的计量经济学教材，也可以作为研究生的参考教材。本书配套的数据文件等教学资源可通过书后的教辅材料申请表索取。

# About the Author

Jeffrey M. Wooldridge is University Distinguished Professor of Economics at Michigan State University, where he has taught since 1991. From 1986 to 1991, Dr. Wooldridge was an assistant professor of economics at the Massachusetts Institute of Technology. He received his bachelor of arts, with majors in computer science and economics, from the University of California, Berkeley, in 1982 and received his doctorate in economics in 1986 from the University of California, San Diego. Dr. Wooldridge has published more than three dozen articles in internationally recognized journals, as well as several book chapters. He is also the author of *Econometric Analysis of Cross Section and Panel Data*. His awards include an Alfred P. Sloan Research Fellowship, the Plura Scripsit award from *Econometric Theory*, the Sir Richard Stone prize from the *Journal of Applied Econometrics*, and three graduate teacher-of-the-year awards from MIT. He is a fellow of the *Econometric Society* and of the *Journal of Econometrics*. Dr. Wooldridge has been editor of the *Journal of Business and Economic Statistics* and econometrics coeditor of *Economics Letters*, and he has served on the editorial boards of *Econometric Theory*, the *Journal of Economic Literature*, the *Journal of Econometrics*, the *Review of Economics and Statistics*, and the *Stata Journal*. He has also acted as an occasional econometrics consultant for Arthur Andersen, Charles River Associates, and the Washington State Institute for Public Policy.

# 导读——兼改编说明

随着我国计量经济学教学和研究水平的不断提高，以及经济学人才培养国际合作项目的不断加深，特别是大批经济学海外人才的迅速引进，对计量经济学英文教材的需求快速增长。应高等教育出版社之约，我对Wooldridge所著的英文原版教材*Introductory Econometrics—A Modern Approach, Fourth Edition*进行改编。这是计量经济学专业一本著名的教材。我深知，要改编好一本著名的英文教材，实在是一件不容易的事情，其困难在于：改编既要体现原书的精华，又要保留原书章与章之间的逻辑联系，从而保证前后内容的衔接，使教学内容能够有序展开。坦率地讲，英文原版教材内容有些冗杂，某些内容的讲述稍显累赘，使得学生通常感觉理论虽新，但却雾里看花、抓不住重点。为了便于学生在学习计量经济学的时候抓住要领，这就要求改编的时候必须删除或者重写部分知识点。然而，发现和删除某些累赘的内容是简单的，但要使删除后的知识点相互贯通则有比较大的难度。改编后的内容应基本符合我国大学生计量经济学的教学及其课时要求，内容相互衔接，篇幅和难度基本适当。然而据我所知，几乎所有的英文原版教材，无一例外地都是越来越厚。从上述角度来说，改编工作无疑是困难的，但要使改编后的英文教材内容简洁、逻辑清晰、篇幅和难度适当，也是一件有意义的学术工作。

首先说明改编后教材的基本特征。概括地说，改编后的内容置于主流的计量经济学框架内，并且具有比较完整的知识体系。不仅如此，计量经济学的理论由浅入深、前后递进、理论与应用相结合。改编后的内容大致可以分为四个部分：其一，基于横截面数据的模型、最小二乘估计（OLS）和假设检验及其应用。这是本书最清晰的部分。我们知道，回归分析是计量经济学的基本工具和基础性理论。现有的教材，基本没有基于数据类型讲述回归分析，导致基于横截面数据的模型、估计和估计量的性质及假设检验理论与基于时间序列数据的性质时常相互混淆，增加了学生学习的难度。通读本书可以看出，基于横截面数据讲述OLS及其性质，从理论上是非常适合的。其二，时间序列数据的模型设定、估计和检验理论与应用。从一般的意义来说，时间序列数据模型估计、检验及其应用，既与横截面数据有相同之处，也有其独特的性质。因此，将时间序列数据的模型区别于横截面数据的模型，不仅有助于介绍计量经济学的基本理论，而且使各种概念之间的内在联系更加清晰。其三，面板数据模型的理论和应用。从横截面数据、时间序列数据，到面板数据模型，这很显然是一种具有内在逻辑关系的知识发展和演进。由此而言，改编的内容安排体现了原书的精华。其四，离散选择模型或者微观计量经济学，用于研究个体选择的决定因素，如是否继续读研究生、是否购买住房。从理论和方法看，这一类模型由于不可能简单使用OLS估计与检验，

从而与前述模型相区别。而且，微观计量经济学是计量经济学非常重要的一个分支，因此所有的教材都必须介绍这部分内容。以上四个方面，形成了计量经济学的主流框架，也是每一本计量经济学教材的主要内容。改编后的内容集中于上述主流框架，加强了基础性理论，适当弱化了应用。原书的特点之一是，基于大量的例子，说明回归分析的应用。但是，这些例子都是早期的例子，而现有的文献表明，仅仅使用回归分析几乎不可能形成高水平的经济学研究。因此，我们删除了若干例子，以减少篇幅。

其次，我们谨借此机会，说明计量经济学的教学安排与体会。上述改编，适用于本科生计量经济学40或者54（60）学时的教学。对于本科生教学而言，必须强调对计量经济学基础性理论的正确理解，通过例子解释模型的经济学意义，必要时应基于学生的正确理解进行适度延伸。例如，无偏性、外生性的基本理解，从无偏性到渐近无偏性逐步延伸。从现在的经济学研究可以看出，任何一篇现代经济学实证论文，都具有理论和方法论的难度，绝不是通过回归就可以产生结论这么简单。指望使用回归分析就可以实现对经济学高水平的研究，显然已经不可能了。原书试图通过大量的例子，将计量经济学的基本理论与应用结合起来。但是，这些例子几乎都是早期（20世纪70年代以前）的论文，与现在的计量经济学应用性论文的计量理论要求，有着非常明显的差别。因此，适当弱化原书中的应用性论文，不仅必要，也为减轻教学内容、压缩篇幅提供了途径。另外，计量经济学理论和方法的新发展，无不建立在对基础性理论的正确理解和延伸之上。从这个意义来说，我们建议将重点放到基础性理论。例如，基于经济学理论设定的模型，主要讲授OLS的原理、估计量的分布、假设检验及其经济学含义的识别和解释、模型系数的经济学意义等。如果学生有兴趣，可以将以上理论和方法延伸至渐进理论。与此同时，可以简单介绍Monte-Carlo仿真技术，并且引导学生实现对诸如无偏性等性质的Monte-Carlo验证。

再次，我们知道，计量经济学教材越来越厚，如果直接使用任何一本计量经济学英文教材，任何大学给予计量经济学的学时都是远远不够的。从这个意义来说，任何教材，实际上都是参考读物，教师必须合理安排适当的内容。本书的英文原版也是这样。“厚”源于两个方面：其一，为了强化对基础性内容的理解和应用，汇集了大量应用型例子。我们强调的是，过多的例子，是教材越来越厚的原因之一。因此，我们出于压缩篇幅而删除了大量的实际例子。其二，随着计量经济学和经济学的发展，知识的积累越来越多，原版教材试图面面俱到，尽可能多地介绍计量经济学的不同分支或者主题，导致教材越来越厚。而且主流计量经济学也随之发生深刻的变化，一个典型的例子是联立方程模型，由于其没有预测到20世纪70年代的石油危机带来的经济衰退，而逐渐退出经济学和计量经济学的文献，VAR模型取而代之，逐渐演变为实证研究的主要工具。因此，简化或者删除部分新的但非主流的计量经济学方法，也是我们改编工作的内容。其三，为了介绍一些新发展起来的方法，原书不惜使用较多篇幅介绍估计量的方差和协方差的稳健（Robust）估计，但其介绍的稳健估计主要是基于回归而实现的估计和相关检验。然而，采用这种方式

介绍稳健估计，学生不仅不理解估计的基本思想和估计结果的应用和意义，还会陷入其中而不得其解。因此，我们也删除了这样的内容。

上述的改编和删除，是否降低了原书的难度？改编后的教材能否作为研究生的教材？我想从两个方面说明这个问题：其一，为强化计量经济学的基础性理论，我们保留了诸如渐近性理论、某些定理的证明和公式的推导。这些内容，根据学生的兴趣和学习计量经济学的积极性，既可以作为本科生的教学内容（但没必要作为教学重点），也可以作为自学的内容。之所以保留这些具有一定难度的内容，目的是通过这些内容的学习，增强学生阅读现代经济学文献的能力，培养学生的研究能力。但是，大多数本科生的计量经济学教学，都没有将上述内容作为教学内容。另外，从现代计量经济和经济学文献可以看出，大多数计量经济学方法的论文，都可能涉及上述内容，对于继续学习或者读研究生的读者来说，如果不学习这部分内容，计量经济学的学习和研究就无法深入。因此，这一部分的内容作为自学或者有选择的讲授是比较合适的安排。无疑的，保留这些内容体现了相应的难度。其二，关于难度。计量经济学是国内外公认的难度较大的学科。其难度在于，计量经济学对统计知识的要求非常高，几乎任何新的统计方法，都可以应用到计量经济学之中。从这个角度看，计量经济学的难度，就是不断更新、不断发展的数学和统计方法在计量经济学理论之中的应用。事实上，经济学研究也是如此。这些事实意味着，回避计量经济学的难度是不可能的，在教科书中体现基本的难度，有助于学生对计量经济学的正确理解。将计量经济学看做“简单”的回归，认为学习了计量经济学，就可以应用计量经济学研究经济问题，这样的想法也许是天真的。计量经济学的难度在不断增强，我们在改编的过程中，在保留基本难度的同时，删除了前面提到的含糊不清的协方差的稳健估计等内容，适当降低了难度。尽管如此，将本书中的难点（如渐近性理论）作为教学重点，适当增加诸如VAR等主流计量经济模型的介绍，本书亦可以作为研究生的参考教材。

最后，我要特别感谢本书的审校、清华大学李子奈教授。李老师的审校意见，首先是体现了他对计量经济学整个学科的洞悉及其与教学内容的融合与衔接，也体现了他一以贯之的认真。例如，他建议保留原书中计量经济模型的矩阵理论和部分渐近理论。这一建议突然使我意识到，现代计量经济学的研究，大多是基于矩阵进行的，其难点之一是渐近理论。因此，保留这部分内容，不仅为进一步学习和研究计量经济学提供了基础和桥梁，也衔接了计量经济学的发展与本科生的教学内容。再如，联立方程模型，从文献和研究的角度看，这是计量经济学早期的主流模型，由于这一类模型都没有预测到石油危机及其所引发的世界经济的衰退，导致对联立模型的批判。英文版教材通常使用较大篇幅介绍联立模型和它的应用。李老师认为，鉴于联立模型的理论和应用研究基本消失，这部分内容可以从本科生教学内容中删除。李老师的审稿意见简洁轻松，但一语中的。如果没有长期的教学积累，没有对计量经济学前沿准确的把握，就不可能像李老师那样“轻松”地提出建议。其次，李老师的审校，也体现了他对计量经济学教学的热忱和对计量经济学人才培养难以割舍的情结。

我无以表达谢意，谨致深深的敬意！

由于本人的水平有限，本书不可避免地存在这样或那样的问题，欢迎读者批评指正。

王少平

2013年12月于华中科技大学

# Contents

**Chapter 13**

## Instrumental Variables Estimation and Two Stage Least Squares 360



**Chapter 14**

## Limited Dependent Variable Models 395



**Chapter 15**

## Advanced Time Series Topics 422



**Appendix A**

## The Normal and Related Distributions 455

**Appendix B**

## Answers to Chapter Questions 463

**Appendix C**

## Statistical Tables 470

## References 477

Chapter 1

# The Nature of Econometrics and Economic Data

Chapter 1 discusses the scope of econometrics and raises general issues that arise in the application of econometric methods. Section 1.3 examines the kinds of data sets that are used in business, economics, and other social sciences. Section 1.4 provides an intuitive discussion of the difficulties associated with the inference of causality in the social sciences.

## 1.1 What Is Econometrics?

Econometrics is based upon the development of statistical methods for estimating economic relationships, testing economic theories, and evaluating and implementing government and business policy. The most common application of econometrics is the forecasting of such important macroeconomic variables as interest rates, inflation rates, and gross domestic product. Whereas forecasts of economic indicators are highly visible and often widely published, econometric methods can be used in economic areas that have nothing to do with macroeconomic forecasting. For example, we will study the effects of political campaign expenditures on voting outcomes. We will consider the effect of school spending on student performance in the field of education. In addition, we will learn how to use econometric methods for forecasting economic time series.

Econometrics has evolved as a separate discipline from mathematical statistics because the former focuses on the problems inherent in collecting and analyzing nonexperimental economic data. **Nonexperimental data** are not accumulated through controlled experiments on individuals, firms, or segments of the economy. (Nonexperimental data are sometimes called **observational data**, or **retrospective data**, to emphasize the fact that the researcher is a passive collector of the data.) **Experimental data** are often collected in laboratory environments in the natural sciences, but they are much more difficult to obtain in the social sciences. Although some social experiments can be devised, it is often impossible, prohibitively expensive, or morally repugnant to conduct the kinds of controlled experiments that would be needed to address economic issues. We give some specific examples of the differences between experimental and nonexperimental data in Section 1.4.

Naturally, econometricians have borrowed from mathematical statisticians whenever possible. The method of multiple regression analysis is the mainstay in both fields, but its focus and interpretation can differ markedly. In addition, economists have devised new techniques to deal with the complexities of economic data and to test the predictions of economic theories.

## 1.2 Steps in Empirical Economic Analysis

Econometric methods are relevant in virtually every branch of applied economics. They come into play either when we have an economic theory to test or when we have a relationship in mind that has some importance for business decisions or policy analysis. An **empirical analysis** uses data to test a theory or to estimate a relationship.

How does one go about structuring an empirical economic analysis? It may seem obvious, but it is worth emphasizing that the first step in any empirical analysis is the careful formulation of the question of interest. The question might deal with testing a certain aspect of an economic theory, or it might pertain to testing the effects of a government policy. In principle, econometric methods can be used to answer a wide range of questions.

In some cases, especially those that involve the testing of economic theories, a formal **economic model** is constructed. An economic model consists of mathematical equations that describe various relationships. Economists are well known for their building of models to describe a vast array of behaviors. For example, in intermediate microeconomics, individual consumption decisions, subject to a budget constraint, are described by mathematical models. The basic premise underlying these models is *utility maximization*. The assumption that individuals make choices to maximize their well-being, subject to resource constraints, gives us a very powerful framework for creating tractable economic models and making clear predictions. In the context of consumption decisions, utility maximization leads to a set of *demand equations*. In a demand equation, the quantity demanded of each commodity depends on the price of the goods, the price of substitute and complementary goods, the consumer's income, and the individual's characteristics that affect taste. These equations can form the basis of an econometric analysis of consumer demand.

Economists have used basic economic tools, such as the utility maximization framework, to explain behaviors that at first glance may appear to be noneconomic in nature. A classic example is Becker's (1968) economic model of criminal behavior.

### Example 1.1

**[Economic Model of Crime]**

In a seminal article, Nobel Prize winner Gary Becker postulated a utility maximization framework to describe an individual's participation in crime. Certain crimes have clear economic rewards, but most criminal behaviors have costs. The opportunity costs of crime prevent the criminal from participating in other activities such as legal employment. In addition, there are costs associated with the possibility of being caught and then, if convicted, the costs associated with incarceration. From Becker's perspective, the decision to undertake illegal activity is one of resource allocation, with the benefits and costs of competing activities taken into account.

Under general assumptions, we can derive an equation describing the amount of time spent in criminal activity as a function of various factors. We might represent such a function as

$$y = f(x_1, x_2, x_3, x_4, x_5, x_6, x_7), \quad \textbf{1.1}$$

where

$y$ = hours spent in criminal activities,
$x_1$ = "wage" for an hour spent in criminal activity,
$x_2$ = hourly wage in legal employment,
$x_3$ = income other than from crime or employment,
$x_4$ = probability of getting caught,
$x_5$ = probability of being convicted if caught,
$x_6$ = expected sentence if convicted, and
$x_7$ = age.

Other factors generally affect a person's decision to participate in crime, but the list above is representative of what might result from a formal economic analysis. As is common in economic theory, we have not been specific about the function $f(\cdot)$ in (1.1). This function depends on an underlying utility function, which is rarely known. Nevertheless, we can use economic theory—or introspection—to predict the effect that each variable would have on criminal activity. This is the basis for an econometric analysis of individual criminal activity.

Formal economic modeling is sometimes the starting point for empirical analysis, but it is more common to use economic theory less formally, or even to rely entirely on intuition. You may agree that the determinants of criminal behavior appearing in equation (1.1) are reasonable based on common sense; we might arrive at such an equation directly, without starting from utility maximization. This view has some merit, although there are cases in which formal derivations provide insights that intuition can overlook.

Next is an example of an equation that we can derive through somewhat informal reasoning.

### Example 1.2

**[Job Training and Worker Productivity]**

Consider the problem posed at the beginning of Section 1.1. A labor economist would like to examine the effects of job training on worker productivity. In this case, there is little need for formal economic theory. Basic economic understanding is sufficient for realizing that factors such as education, experience, and training affect worker productivity. Also, economists are well aware that workers are paid commensurate with their productivity. This simple reasoning leads to a model such as

$$wage = f(educ, exper, training), \quad \textbf{1.2}$$

where

*wage* = hourly wage,
*educ* = years of formal education,
*exper* = years of workforce experience, and
*training* = weeks spent in job training.

Again, other factors generally affect the wage rate, but equation (1.2) captures the essence of the problem.

After we specify an economic model, we need to turn it into what we call an **econometric model**. Because we will deal with econometric models throughout this text, it is important to know how an econometric model relates to an economic model. Take equation (1.1) as an example. The form of the function $f(\cdot)$ must be specified before we can undertake an econometric analysis. A second issue concerning (1.1) is how to deal with variables that cannot reasonably be observed. For example, consider the wage that a person can earn in criminal activity. In principle, such a quantity is well defined, but it would be difficult if not impossible to observe this wage for a given individual. Even variables such as the probability of being arrested cannot realistically be obtained for a given individual, but at least we can observe relevant arrest statistics and derive a variable that approximates the probability of arrest. Many other factors affect criminal behavior that we cannot even list, let alone observe, but we must somehow account for them.

The ambiguities inherent in the economic model of crime are resolved by specifying a particular econometric model:

$$crime = \beta_0 + \beta_1 wage_m + \beta_2 othinc + \beta_3 freqarr + \beta_4 freqconv + \beta_5 avgsen + \beta_6 age + u, \tag{1.3}$$

where

*crime* = some measure of the frequency of criminal activity,
$wage_m$ = the wage that can be earned in legal employment,
*othinc* = the income from other sources (assets, inheritance, and so on),
*freqarr* = the frequency of arrests for prior infractions (to approximate the probability of arrest),
*freqconv* = the frequency of conviction, and
*avgsen* = the average sentence length after conviction.

The choice of these variables is determined by the economic theory as well as data considerations. The term $u$ contains unobserved factors, such as the wage for criminal activity, moral character, family background, and errors in measuring things like criminal activity and the probability of arrest. We could add family background variables to the model, such as number of siblings, parents' education, and so on, but we can never eliminate $u$ entirely. In fact, dealing with this *error term* or *disturbance term* is perhaps the most important component of any econometric analysis.

The constants $\beta_0$, $\beta_1$, $\cdots$, $\beta_6$ are the *parameters* of the econometric model, and they describe the directions and strengths of the relationship between *crime* and the factors used to determine *crime* in the model.

A complete econometric model for Example 1.2 might be

$$wage = \beta_0 + \beta_1 educ + \beta_2 exper + \beta_3 training + u, \tag{1.4}$$

where the term $u$ contains factors such as "innate ability," quality of education, family background, and the myriad other factors that can influence a person's wage. If we are specifically concerned about the effects of job training, then $\beta_3$ is the parameter of interest.

For the most part, econometric analysis begins by specifying an econometric model, without consideration of the details of the model's creation. We generally follow this approach, largely because careful derivation of something like the economic model of crime is time-consuming and can take us into some specialized and often difficult areas