

计量经济学

Basic Econometrics 下册

(第三版)

〔美〕古扎拉蒂 / 著

Damodar N. Gujarati

中国人民大学出版社

经济
科学
译丛

1006842

计量经济学

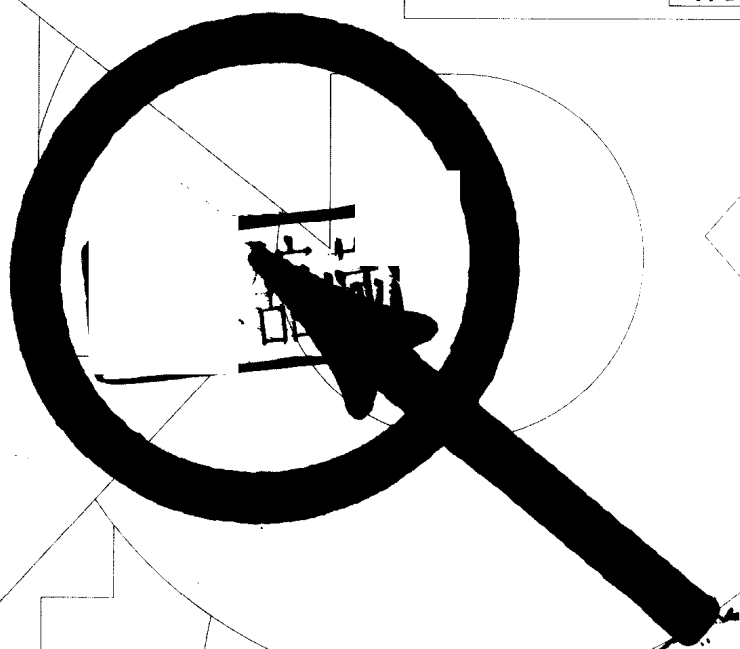
Basic Econometrics 下册

(第三版)

〔美〕古扎拉蒂 / 著

Damodar N. Gujarati

林少宫 / 译



中国人民大学出版社

经济科学译丛

图书在版编目(CIP)数据

计量经济学(第三版)/(美)古扎拉蒂著;林少宫译

北京:中国人民大学出版社,2000

(经济科学译丛)

ISBN 7-300-03044-0/F·1014

I. 计…

II. ①古…②林…

III. 计量经济学

IV. F224.0

中国版本图书馆 CIP 数据核字(2000)第 03448 号

经济科学译丛

计量经济学(第三版) 上、下册

[美] 古扎拉蒂 著

林少宫 译

出版发行:中国人民大学出版社

(北京海淀路 157 号 邮编 100080)

发行部:62514146 门市部:62511369

总编室:62511242 出版部:62511239

E-mail:rendafx@public3.bta.net.cn

经 销:新华书店

印 刷:涿州市星河印刷厂

开本:787×1092 毫米 1/16 印张:55.75 插页 10

2000 年 3 月第 1 版 2000 年 3 月第 1 次印刷

印数:1-10 000

定价:98.00 元

(图书出现印装问题,本社负责调换)

《经济科学译丛》

编辑委员会

学术顾问 高鸿业 王传纶
胡代光 范家骧
朱绍文 吴易风

主 编 陈岱孙

副主编 梁 晶 海 闻

编 委

王逸舟
王利民
贝多广
林毅夫
李 扬
李晓西
刘 伟
朱 玲
金 碚
张 宇
易 纲
徐 宽
高 培
姚 建
盛 洪
樊 纲

目 录

第 12 章 自相关	393
§ 12.1 问题的性质	394
§ 12.2 自相关出现时的 OLS 估计量	399
§ 12.3 自相关出现时的 BLUE 估计量	401
§ 12.4 自相关出现时使用 OLS 的后果	402
考虑到自相关的 OLS 估计	402
忽视自相关的 OLS 估计	403
§ 12.5 侦察自相关	407
图解法	407
游程检验	410
德宾—沃森 d 检验	412
自相关的其他检验	416
§ 12.6 补救措施	417
自相关的结构已知	417
ρ 未知	419
§ 12.7 一个说明性例子: 美国零工招聘指数与失业率的关系方法比较	423
§ 12.8 自回归条件异方差 (ARCH) 模型	427
ARCH 出现时怎么办?	428
为德宾—沃森 d 统计量与 ARCH 效应进一言	429

§ 12.9	要点与结论	429
	习题	433
附录 12A		443
12A.1	1960—1991 年美国工资(Y)—生产力(X)TSP 输出	443
第 13 章	计量经济建模 I:传统计量经济学方法论	446
§ 13.1	计量经济建模的传统观点:平均经济回归(AER)	446
§ 13.2	设定误差的类型	448
§ 13.3	设定误差的后果	450
	漏掉一个有关变量(模型的过低拟合)	450
	包含一个无关变量(模型的过高拟合)	451
§ 13.4	设定误差的检验	453
	侦察是否含有无关变量	453
	对遗漏变量和不正确函数形式的检验	454
§ 13.5	观测误差	460
	应变量 Y 中的观测误差	460
	解释变量 X 中的观测误差	461
	一个例子	463
	仅应变量 Y 有观测误差	463
	X 中的观测误差	464
§ 13.6	要点与结论	464
	习题	466
附录 13A		472
13A.1	含有无关变量的后果:无偏性质	472
13A.2(13.5.10)	的证明	473
第 14 章	计量经济建模 II:另立计量经济学方法论	475
§ 14.1	利莫尔的模型选择方法	476
§ 14.2	韩德瑞的模型选择方法	479
§ 14.3	诊断性检验选讲:总评	480
§ 14.4	非嵌套假设的检验	481
	判别方法	481
	辨识方法	482
§ 14.5	要点与结论	487
	习题	489
第 3 篇	计量经济学专题	491
第 15 章	关于虚拟变量的回归	493
§ 15.1	虚拟变量的性质	493
	例 15.1:按性别划分的教授薪金	494
§ 15.2	对一个定量变量和一个两分定性变量的回归	496
	例 15.2:库存对利率敏感吗?	498
§ 15.3	对一个定量变量和一个多分定性变量的回归	499

§ 15.4	对一个定量变量和两个定性变量的回归	500
§ 15.5	例 15.3:“兼职”经济学	502
§ 15.6	检验回归模型的结构稳定性	503
	例 15.4:1946—1963 年联合王国储蓄与收入	503
§ 15.7	比较两个回归:虚拟变量法	505
§ 15.8	比较两个回归:进一步的说明	507
	例 15.5:1958—1971 年英国失业和空缺情况	507
§ 15.9	交互作用效应	509
§ 15.10	虚拟变量在季节分析中的应用	510
	例 15.6:美国制造业的利润—销售额行为	511
§ 15.11	分段线性回归	512
	例 15.7:总成本与产出的关系	514
§ 15.12	在时间序列和横截面数据的合并中使用虚拟变量	515
	混合回归:时间序列与横截面数据并用	515
	例 15.8:通用汽车与西屋电气公司的投资函数	517
§ 15.13	虚拟变量方法的一些技术问题	518
	在半对数回归中的虚拟变量的解释	518
	例 15.9:带虚拟变量的半对数回归	518
	避免虚拟变量陷阱的另一种方法	519
	虚拟变量与异方差性	520
	虚拟变量与自相关	520
§ 15.14	进一步的研究问题	521
§ 15.15	要点与结论	522
	习题	524
附录 15A		533
	15A.1 回归(15.8.2)的数据矩阵	533
	15A.2 回归(15.10.2)的数据矩阵	534
第 16 章	关于虚拟应变量的回归:线性概率模型、对数单位、概率单位及托比模型	535
§ 16.1	虚拟应变量	535
§ 16.2	线性概率模型(LPM)	536
§ 16.3	LPM 的估计问题	537
	干扰 u_i 的非正态性	537
	干扰的异方差性	538
	$0 \leq E(Y_i X_i) \leq 1$ 不被满足	539
	可疑的拟合优度: R^2 值	539
§ 16.4	LPM:一个数值例子	540
§ 16.5	LPM 的应用	543
	例 16.1:科恩—雷—勒曼研究	543
	例 16.2:对债券评级的预测	545

例 16.3: 预测债券违约	546
§ 16.6 LPM 以外的其他方法	547
§ 16.7 对数单位模型	548
§ 16.8 对数单位模型的估计	550
§ 16.9 对数单位模型: 一个数值例子	552
§ 16.10 对数单位模型: 说明性例子	555
例 16.4: “对数单位模型在预测兼并对象中的一个应用”	555
例 16.5: 债券评级的预测	556
§ 16.11 概率单位模型	557
§ 16.12 概率单位模型: 一个数值例子	560
对数单位与概率单位	561
比较对数单位与概率单位的估计值	562
回归元单位变化的边际效应	562
§ 16.13 概率单位模型: 例 16.5	562
§ 16.14 托比模型	565
§ 16.15 要点与结论	568
习题	572
第 17 章 动态计量经济模型: 自回归与分布滞后模型	581
§ 17.1 “时间”或“滞后”在经济学中的作用	582
§ 17.2 滞后的原因	586
§ 17.3 分布滞后模型的估计	587
分布滞后模型的现式估计法	587
§ 17.4 分布滞后模型的考伊克方法	588
中位滞后	591
平均滞后	591
§ 17.5 考伊克模型的理性化: 适应性期望模型	592
§ 17.6 考伊克模型理性化的另一形式: 存量调整或部分调整模型	594
§ 17.7 适应性期望与部分调整模型的组合	596
§ 17.8 自回归模型的估计	597
§ 17.9 工具变量(IV)法	598
§ 17.10 在自回归模型中侦察自相关: 德宾 h 检验	599
§ 17.11 一个数值例子: 印度的货币需求	601
§ 17.12 说明性例题	603
例 17.7: 联邦储备与实际利率	603
例 17.8: 1946—1972 年美国短期与长期总消费函数	604
§ 17.13 分布滞后模型的阿尔蒙方法: 阿尔蒙或多项式分布滞后(PDL)	606
§ 17.14 经济学中的因果关系: 葛兰杰检验	613
葛兰杰检验	613
经验结果	614
§ 17.15 要点与结论	616

第4篇

习题	620
联立方程模型	629
第18章 联立方程模型	631
§ 18.1 联立方程模型的性质	631
§ 18.2 联立方程模型举例	632
例 18.1: 需求与供给模型	632
例 18.2: 凯恩斯收入决定模型	633
例 18.3: 工资—价格模型	634
例 18.4: 宏观经济学中的 IS 模型	635
例 18.5: LM 模型	636
例 18.6: 计量经济模型	637
§ 18.3 联立方程偏误: OLS 估计量的非一致性	638
§ 18.4 联立方程偏误: 一个数值例子	641
§ 18.5 要点与结论	643
习题	644
第19章 识别问题	649
§ 19.1 符号与定义	649
§ 19.2 识别问题	652
不足识别情形	652
恰好或恰可识别情形	654
过度识别情形	658
§ 19.3 识别规则	659
可识别性的阶条件	659
可识别性的秩条件	661
§ 19.4 联立性检验	664
豪斯曼设定检验	664
例 19.5: 平狄克—鲁宾费尔德公共支出模型	666
§ 19.5 外生性检验	667
关于因果律与外生性	667
§ 19.6 要点与结论	668
习题	669
第20章 联立方程方法	673
§ 20.1 估计的方法	673
§ 20.2 递归模型与普通最小二乘法	674
§ 20.3 恰可识别方程的估计: 间接最小二乘 (ILS) 法	676
一个说明性例子	677
ILS 估计量的性质	680
§ 20.4 过度识别方程的估计: 二阶最小二乘法 (2SLS)	680
§ 20.5 2SLS: 一个数值例子	683
§ 20.6 说明性例子	686

第 5 篇

例 20.1: 广告费、集中度与价差	686
例 20.2: 克莱因的模型 I	688
例 20.3: 表达为递归方程组的资本性资产定价模型	689
例 20.4: 圣路易斯模型的修订版本	691
§ 20.7 要点与结论	693
习题	696
附录 20A	700
20A.1 间接最小二乘估计量的偏误	700
20A.2 2SLS 估计量的标准误的估计	701
时间序列计量经济学	703
第 21 章 时间序列计量经济学 I: 平稳性、单位根与协积	705
§ 21.1 选看一些美国经济的时间序列	706
§ 21.2 平稳随机过程	709
§ 21.3 根据相关图的平稳性检验	711
§ 21.4 平稳性的单位根检验	713
美国 GDP 时间序列是平稳的吗?	715
GDP 序列的一阶差分是平稳的吗?	716
§ 21.5 趋势平稳(TS)与差分平稳(DS)随机过程	717
§ 21.6 谬误回归	719
§ 21.7 协积	721
恩格尔—葛兰杰(EG)或扩充恩格尔—葛兰杰(AEG)检验	721
协积回归德宾—沃森(CRDW)检验	722
§ 21.8 协积与误差纠正机制(ECM)	722
§ 21.9 要点与结论	723
习题	727
附录 21A	729
21A.1 随机步游模型	729
第 22 章 时间序列计量经济学 II: 用于预测的 ARIMA 与 VAR 模型	730
§ 22.1 经济预测方法	730
§ 22.2 时间序列数据的 AR、MA 和 ARIMA 建模	731
自回归(AR)过程	732
移动平均(MA)过程	732
自回归与移动平均(ARMA)过程	733
自回归求积移动平均(ARIMA)过程	733
§ 22.3 博克斯—詹金斯(BJ)方法论	734
§ 22.4 识别	735
§ 22.5 ARIMA 模型的估计	739
§ 22.6 诊断	739
§ 22.7 预测	739
§ 22.8 BJ 方法论的其他方面	741

§ 22.9 向量自回归 (VAR)	741
VAR 的估计	742
VAR 用于预测	744
VAR 建模的一些问题	744
VAR 的一个应用:得克萨斯州经济的一个 VAR 模型	745
§ 22.10 要点与结论	747
习题	749
附录 A 统计学中的若干概念复习	751
§ A.1 总和与乘积运算符	751
§ A.2 样本空间,样本点与事件	752
§ A.3 概率与随机变量概率	753
§ A.4 概率密度函数 (PDF)	754
§ A.5 概率分布的特征值	759
§ A.6 若干重要的理论概率分布	768
§ A.7 统计推断	774
§ A.8 统计推断:假设检验	782
附录 B 矩阵代数初步	788
§ B.1 定义	788
§ B.2 矩阵的类型	790
§ B.3 矩阵运算	792
§ B.4 行列式	796
§ B.5 求一个方阵的逆阵	799
§ B.6 矩阵微分法	801
附录 C 计量经济软件包选编	803
附录 D 统计学用表	805
表 D.1 标准化正态分布下的面积	806
表 D.2 t 分布的百分点	807
表 D.3 F 分布的上端百分点	808
表 D.4 χ^2 分布的上端百分点	814
表 D.5 德宾-沃森 d 统计量:在 0.05 和 0.01 显著性水平上 d_L 和 d_U 的显著点	816
表 D.6 游程检验中的游程临界值	820
参考书目	822
人名索引中译	825
标题索引	829
后 记	853

第 12 章 自相关

400

没有普遍有效的方法能防止把回归函数的错误设定误解为序列相关误差的出现。^[1]

第 1 篇阐述的经典线性模型有一重要假定,即在进入总体回归函数的干扰 u_i 中无自相关或序列相关。本章中我们从严分析这一假定,以期得到对下列问题的答案:

1. 自相关的性质是什么?
2. 自相关会导致什么理论上和实际上的后果?
3. 由于非自相关假定涉及不可观测的干扰 u_i ,怎样能知道在一个任给的情况中是否有自相关?
4. 怎样补救自相关的问题?

读者将发现本章在许多方面类似于前章对异方差性的讨论。因为,在自相关情况下,和在异方差情况下一样,平常的 OLS 估计量虽然无偏,但不再是所有线性无偏估计量中的最小方差者。简言之,OLS 估计量不再是 BLUE。

§ 12.1 问题的性质

401

自相关(autocorrelation)一词可定义为按时间(如在时间序列数据中)或空间(如在横截面数据中)排序的观测值序列的成员之间的相关。^[2]在回归的论述中,经典线性回归模型假定在干扰 u_i 之间不存在有自相关,用符号表示:

$$E(u_i u_j) = 0 \quad i \neq j \quad (3.2.5)$$

简单地说,经典模型假定,任一次观测的干扰项都不受任何其他观测的干扰项影响。例如,在做产出对劳力和资本投入的回归中,我们用了季度时间序列数据。如果某一季度的产出受到罢工的影响,却没有理由认为这一生产中断会持续到下一季度,就是说,即令本季度产出下降,却没有理由预期下一季度的产出也下降。同理,如果我们在家庭消费支出对家庭收入的回归中,用了横截面数据,那么,一个家庭的收入增加对其消费支出的影响,预料不会波及到另一个家庭的消费支出。

然而,如果存在有这种相依性,我们就有了自相关。用符号表示:

$$E(u_i u_j) \neq 0 \quad i \neq j \quad (12.1.1)$$

这时,由本季罢工引起的生产中断很可能影响下季的产出;或者,一个家庭的消费支出增加,很可能促使另一个家庭为了同邻居攀比也随之增大其消费支出。

在我们寻求自相关存在的原因前,澄清一些名词术语方面的问题是必要的。虽然现在把名词**自相关**(autocorrelation)和**序列相关**(serial correlation)看作同义语已成为习惯,但一些作者比较喜欢把两者区分开来。例如,廷特纳(Tintner)定义自相关为“一给定序列同它自身滞后若干期的序列的滞后相关。而与此同时,他把序列相关一词保留作“两个不同序列的滞后相关。”^[3]例如,时间序列 $u_1, u_2, \dots, u_{10}$ 和它自身滞后一期的序列 $u_2, u_3, \dots, u_{11}$ 之间的相关叫做**自相关**,而两个不同时间序列 $u_1, u_2, \dots, u_{10}$ 和 $v_2, v_3, \dots, v_{11}$ 之间的相关叫做**序列相关**。尽管两名词的区分有一定用处,本书中将把它们看作同义语。

让我们想像自相关和无自相关的一些可能的模式,如图 12.1 所示。图 12.1a 到 d 各显示 u_i 中的一个可分辨的模式。图 12.1a 显示一个周期模样;图 12.1b 和 c 分别表明干扰中的一个上升或下降线性趋势;而图 12.1d 表明干扰中兼有线性和二次趋势项。惟有图 12.1e 表示无系统性模样,符合于经典线性回归模型的无相关假定。

402

一个自然的问题是:为什么会出现序列相关?有种种理由,其中的一些如下:

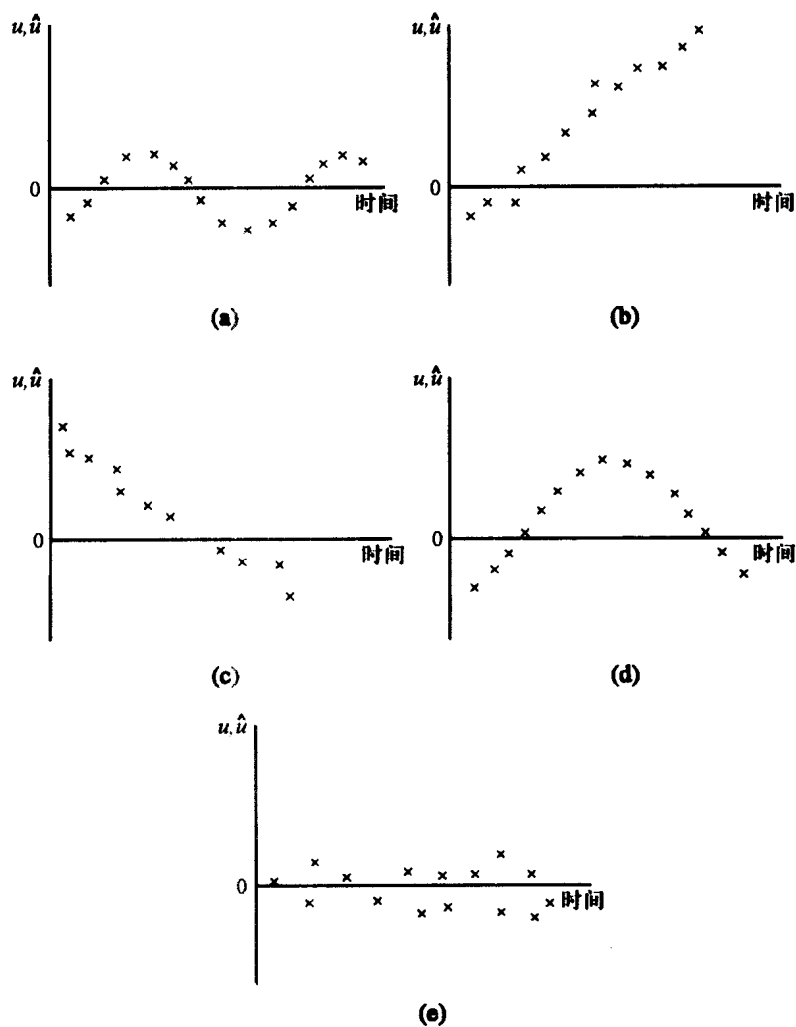


图 12.1 自相关模式

408 **惯性。**大多数经济时间序列都有一个明显的特点,就是它的惯性或粘滞性。众所周知,GNP,价格指数、生产、就业和失业等时间序列都呈现(商业)循环。从衰退的谷底开始,当经济开始复苏时,大多数经济序列开始上升。在此上升期间,序列在每一时刻的值都高于前一时刻的值。看来有一种内在的动力驱使这一势头继续下去,直至某些情况(如利率或课税或两者同时升高)出现才把它拖慢下来。因此,在涉及时间序列的回归中,相继的观测值很可能是相依赖的。

设定偏误:应含而未含变量的情形。在经验分析中,研究者常从一个较好但不一定“最好”的回归模型开始。经过回归分析,研究者做事后检查,看结果是否与事先预期的相一致。如不一致,便要开始动手术。例如,研究者可能将拟合回归的残差描图,并可能观察到类似于图 12a 到 d 那样的模式。这些残差(为 u_i 的替代变量)也许表明,某些原来待选的、却未被包含的变量,由于种种

理由应被包含到模型中来。这就是应含而未含变量(excluded variables)的设定偏误。但这些变量包含进来常常能消除干扰中所观察到的相关模式。例如,假如我们有以下的需求模型:

$$Y_t = \beta_1 + \beta_2 X_{2t} + \beta_3 X_{3t} + \beta_4 X_{4t} + u_t \quad (12.1.2)$$

其中 Y_t = 牛肉需求量, X_2 = 牛肉价格, X_3 = 消费者收入, X_4 = 猪肉价格及 t = 时间。^[4]然而,为了某种理由,我们做了下述回归:

$$Y_t = \beta_1 + \beta_2 X_{2t} + \beta_3 X_{3t} + v_t \quad (12.1.3)$$

现在,如果(12.1.2)是“正确”的模型或“真理”或真实的关系式,则做回归(12.1.3)就无异于令 $v_t = \beta_4 X_{4t} + u_t$ 。于是在猪肉价格影响牛肉消费的限度内,误差或干扰项 v 将反映出一种系统性模式,从而造成(错误的)自相关。真的有自相关吗?一个简单的检验是,同时做回归(12.1.2)和(12.1.3),看看在做(12.1.2)时,原先在模型(12.1.3)中观测到的自相关(如果观测有的话)是否消失。^[5]

侦察自相关的实际步骤将在 12.5 节中讨论。届时我们将看到,用回归(12.1.2)的残差来描图,常常会在很大程度上有助于看清楚自相关的问题。

设定偏误:不正确的函数形式。假使在成本—产出研究中“真实”或“正确”的模型如下:

$$\text{边际成本}_i = \beta_1 + \beta_2 \text{产出}_i + \beta_3 \text{产出}_i^2 + u_i \quad (12.1.4)$$

404 但我们拟合了以下的模型:

$$\text{边际成本}_i = \alpha_1 + \alpha_2 \text{产出}_i + v_i \quad (12.1.5)$$

图 12.2,连同“不正确”的线性成本曲线一起,展示了“真实”模型的边际成本曲线。

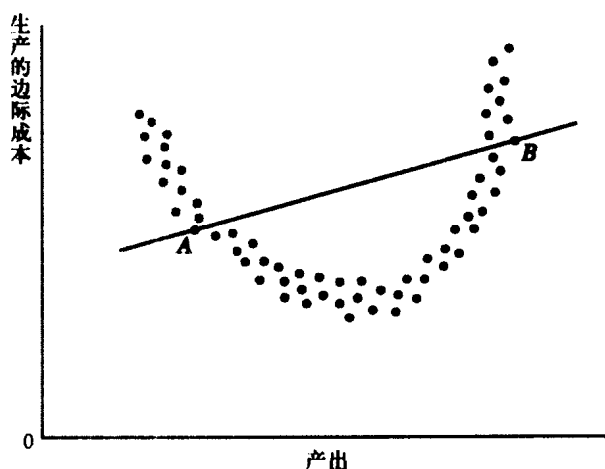


图 12.2 设定偏误:不正确的函数形式

如图 12.2 所示,在点 A 与 B 之间的线性边际成本曲线一直在高估真实的边际成本,而在这两点之外则一致地低估真实的边际成本。这种结果是可以预料到的,因为干扰项 v_i 事实上等于产出² + u_i 从而包含了(产出)² 对边际成本的系统影响。在这种情形中,由于错误函数形式的使用, v_i 将要反映出自相关。在第 13 章中,我们将考虑侦察设立偏误的若干方法。

蛛网现象。许多农产品的供给反映一种所谓的蛛网现象(Cobweb phenomenon)。供给对价格的反应要滞后一个时期,是因为供给需要经过一定的时间才能实现(有一孕育期)。例如,今年年初的作物种植是受去年流行的价格影响的。因此,相关的供给函数是:

$$\text{供给}_t = \beta_1 + \beta_2 \text{价格}_{t-1} + u_t \quad (12.1.6)$$

假使时期 t 的期末价格 P_t 低于 P_{t-1} ,农民就很可能决定在时期 $t+1$ 生产比时期 t 更少的东西。显然,在这种情形中,农民由于在年度 t 的过量生产很可能在年度 $t+1$ 削减他们的产量。诸如此类的现象,就不能期望干扰 u_t 是随机的,从而导致一种蛛网模式。

滞后效应。在一个消费支出对收入的时间序列回归中,人们常常发现当前时期的消费支出除了依赖于其他变量外,还依赖于前期的消费支出,就是:

$$\text{消费}_t = \beta_1 + \beta_2 \text{收入}_t + \beta_3 \text{消费}_{t-1} + u_t \quad (12.1.7)$$

405

像(12.1.7)这样的回归,由于解释变量之一是应变量的滞后值而被称为自回归(autoregression)。(我们将在第 17 章中研究这类类型。)关于(12.1.7)这样一类模型的合理性是简单的。由于心理上、技术上以及制度上的原因,消费者是不轻易改变他们的消费习惯的。如果现在我们忽略了(12.1.7)中的滞后项,所造成的误差项将由于滞后消费对当前消费的影响而反映出一种系统性模式。

数据的“编造”。在经验分析中,原始数据(raw data)往往是经过“编造的(manipulated)”。例如,在用到季度数据的时间序列回归中,这些数据通常来自每月数据,不过是把 3 个月的观测值加在一起除以 3 罢了。这种平均的计算减弱了每月数据的波动而引进了数据中的匀滑性。因此,用季度数据描绘的图形要比用月度数据看来匀滑得多。这种匀滑性本身就能使干扰项中出现有系统性样式,从而导入自相关。数据编造的另一来源是数据的内插(interpolation)或外推(extrapolation)。例如在美国每 10 年举行一次人口普查。最近的一次在 1990 年,而此前的一次在 1980 年。假如现在需要两个普查年间 1980 - 1990 年的某年数据,通常的做法就是,根据某些特殊的假定进行内插。所有这些“揉合”技术,都会给数据带来原始数据所没有的系统性样式。^[6]

在本节结束之前,应注意到,虽然自相关的问题可能也确实出现在横截面数据之中,但一般更常出现于时间序列数据中。时间序列数据的观测是按日历顺序排列的。因此,它们在相继的观测中很可能交叉相关,尤其是当相继

观测的时间间隔较短,如每日,每周或每月而不是每年时。一般地说,横截面数据就没有这种日历顺序,虽则有时可能存在某种类似的顺序。因此,在一个消费支出对收入的横截面回归中,用美国的 50 个州作为观测单元,数据就有可能按其所属的南部、西南部、北部等等地区而分组排列的。由于消费模式尽管在同一地区内基本相似,但很可能一个地区不同于另一个地区。这样,由回归估计得的残差可能展现一种与地区差异有联系的系统性模式。问题是要看到,虽然自相关的出现,绝大多数情形都缘于时间序列数据。但它可以出现在横截面数据中。一些作者把横截面数据中的自相关叫做空间自相关(spatial autocorrelation),以表示自相关出现在空间而非时间上。然而,须知在横截面分析中,为要判别是否有自相关,数据的排列必须有某种逻辑上或经济上的旨趣,方有意义。记住这一点是重要的。

406 还应指出,虽然多数经济时间序列都因在一个较长时期内或者上升或者下降而表现出正的自相关,而不像上页图 12.3b 那样表现为一上一下的恒常运动,但是自相关可以是正的或负的。

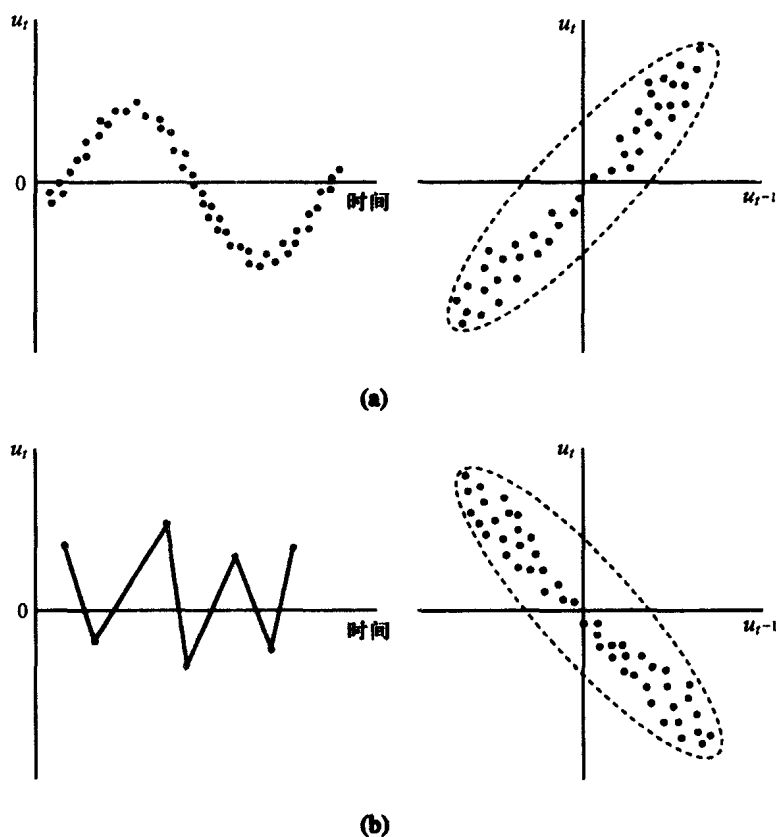


图 12.3 (a)正的自相关和(b)负的自相关