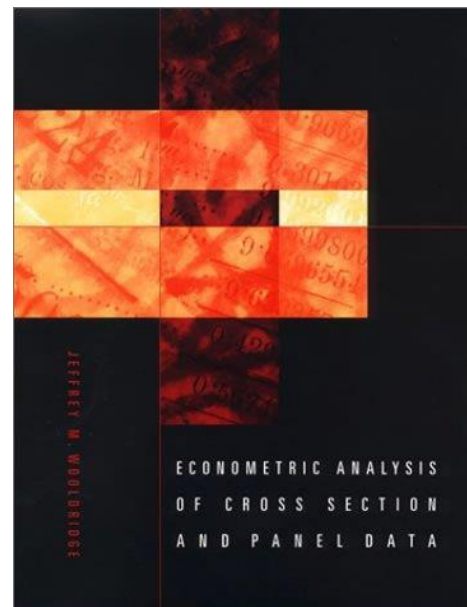
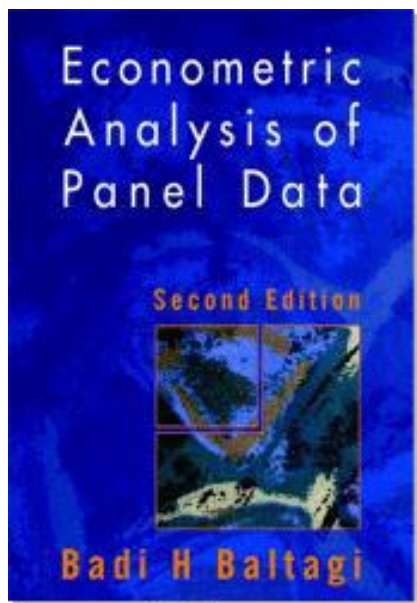


第七章 面板数据模型与应用

从 20 世纪 70 年代末以来, Panel Data 模型的理论方法已经逐步成熟, 形成了现代计量经济学的一个相对独立的分支, 在世界著名的大学如 Harvard、MIT、Berkeley、Stanford 的经济学院系的课程表中, 尤其是研究生课程表中, 已经成为一门独立完整的课程。在应用方面, 无论是宏观经济分析和微观经济分析领域, 都相当普遍。美国南加州大学的 Cheng Hsiao 教授在其有关 Panel Data 的专著中列举了被 SSCI 收录的有关 Panel Data 模型的论文数量, 从 1989 年的 29 篇, 到 1999 年的 650 篇, 说明 Panel Data 模型理论方法的发展和应用的普及。清华大学的 教授查阅了美国和中国的两本综合性经济理论与学术刊物 The American Economic Review(AER)和《经济研究》1984 年至 2004 年的全部论文, 在全部计量经济学应用研究论文中, 采用 Panel Data 模型的论文数量, 在 AER 中处于第二位, 仅次于经典单方程模型, 在《经济研究》中处于第三位, 仅次于经典单方程模型和时间序列分析模型, 而且它的增长速度是最快的。可见学习 Panel Data 模型是非常有必要的了。由于时间的原因, 本章仅仅对 Panel Data 模型做基础性的介绍, 如要进一步系统学习 Panel Data 模型, 请参考有关专著。

参考书目：

1. Badi H. Baltagi, **Econometric Analysis of Panel Data**, John Wiley & Sons, 2005.
2. Jeffrey M. Wooldridge, **Econometric Analysis of Cross Section and Panel Data**, MIT Press, 2002.
3. Cheng Hsiao, **Analysis of Panel Data**, Cambridge University Press, 2003.



7.1 面板数据定义

Panel Data 中译：面板数据、平行数据、纵列数据（Longitudinal Data）、时间序列截面数据（Time Series and Cross Section Data）、混合数据（Pool Data）等。

时间序列数据与截面数据都是一维数据，即时间序列数据是变量按时间得到的数据；截面数据是变量在截面空间上的数据。而面板数据是在时间和截面空间上同时取得的二维数据。面板数据示意图见图 1。面板数据从横截面（cross section）上看，是由若干个体（entity, unit, individual）在某一时刻构成的截面观测值，从纵剖面（longitudinal section）上看是一个时间序列。

面板数据特点：

- （1）面板数据定义为相同截面上的个体在不同时点的重复观测数据。
- （2）称为纵向变量序列（个体）的多次测量。

平衡面板数据（balanced panel data）：数据是完整的，每一个时期的观测个体相等。即每期的 N 相等，样本数为 NT 。

非平衡面板数据：每一个时期的观测个体不再相等，有些个体消失，没有数据观测，即 $\sum_{t=1}^T n_t < NT$ 。

例如，在 20 年中，有些厂商倒闭了等等。

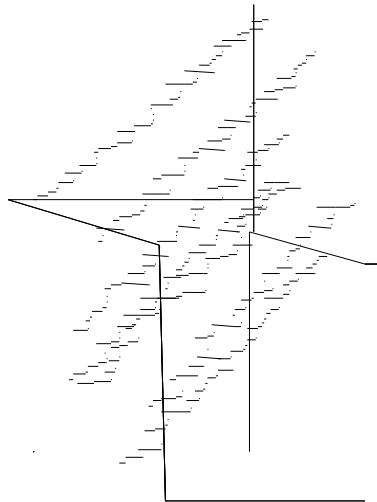


图 1 $N=7, T=50$ 的面板数据示意图

面板数据用双下标变量表示。例如

$$y_{it}, i=1, 2, \dots, N; t=1, 2, \dots, T$$

i 对应面板数据中不同个体， N 表示面板数据中含有 N 个个体。 t 对应面板数据中不同时点， T 表示时间序列的最大长度。若固定 t 不变， $y_{i.}$ ($i=1, 2, \dots, N$) 是横截面上的 N 个随机变量；若固定 i 不变， $y_{i.}$ ($t=1, 2, \dots, T$) 是纵剖面上的一个时间序列（个体）。

本章所讨论的面板数据主要指时期短（ T 小）而截面上包括的个体多（ N 大）的数据。

案例 1：1996-2002 年中国东北、华北、华东 15 个地区的居民家庭人均消费（CP）（不变价格）和人均收入（IP）数据见表 1 和表 2（file:panel02）。数据是 7 年的，每一年都有 15 个数据，共 105 组观测值。

人均消费和收入两个面板数据都是平衡面板数据，各有 15 个个体。人均消费和收入的面板数据从纵剖面观察分别见图 2 和图 3。从横截面观察分别见图 4 和图 5。横截面数据散点图的表现与观测值顺序有关。图 4 和图 5 中人均消费和收入观测值顺序是按地区名的汉语拼音字母顺序排序的。

表 1 1999-2002 年中国东北、华北、华东 15 个省级地区的居民家庭人均消费数据（不变价格）

地区人均消费	1996	1997	1998	1999	2000	2001	2002
CP-AH（安徽）	3282.466	3646.150	3777.410	3989.581	4203.555	4495.174	4784.364
CP-BJ（北京）	5133.978	6203.048	6807.451	7453.757	8206.271	8654.433	10473.12
CP-FJ（福建）	4011.775	4853.441	5197.041	5314.521	5522.762	6094.336	6665.005
CP-HB（河北）	3197.339	3868.319	3896.778	4104.281	4361.555	4457.463	5120.485
CP-HLJ（黑龙江）	2904.687	3077.989	3289.990	3596.839	3890.580	4159.087	4493.535
CP-JL（吉林）	2833.321	3286.432	3477.560	3736.408	4077.961	4281.560	4998.874
CP-JS（江苏）	3712.260	4457.788	4918.944	5076.910	5317.862	5488.829	6091.331
CP-JX（江西）	2714.124	3136.873	3234.465	3531.775	3612.722	3914.080	4544.775
CP-LN（辽宁）	3237.275	3608.060	3918.167	4046.582	4360.420	4654.420	5402.063
CP-NMG（内蒙古）	2572.342	2901.722	3127.633	3475.942	3877.345	4170.596	4850.180
CP-SD（山东）	3440.684	3930.574	4168.974	4546.878	5011.976	5159.538	5635.770
CP-SH（上海）	6193.333	6634.183	6866.410	8125.803	8651.893	9336.100	10411.94
CP-SX（山西）	2813.336	3131.629	3314.097	3507.008	3793.908	4131.273	4787.561
CP-TJ（天津）	4293.220	5047.672	5498.503	5916.613	6145.622	6904.368	7220.843
CP-ZJ（浙江）	5342.234	6002.082	6236.640	6600.749	6950.713	7968.327	8792.210

资料来源：《中国统计年鉴》1997-2003。

表 2 1999-2002 年中国东北、华北、华东 15 个省级地区的居民家庭人均收入数据（不变价格）

地区人均收入	1996	1997	1998	1999	2000	2001	2002
IP-AH（安徽）	4106.251	4540.247	4770.470	5178.528	5256.753	5640.597	6093.333
IP-BJ（北京）	6569.901	7419.905	8273.418	9127.992	9999.700	11229.66	12692.38
IP-FJ（福建）	4884.731	6040.944	6505.145	6922.109	7279.393	8422.573	9235.538
IP-HB（河北）	4148.282	4790.986	5167.317	5468.940	5678.195	5955.045	6747.152
IP-HLJ（黑龙江）	3518.497	3918.314	4251.494	4747.045	4997.843	5382.808	6143.565
IP-JL（吉林）	3549.935	4041.061	4240.565	4571.439	4878.296	5271.925	6291.618
IP-JS（江苏）	4744.547	5668.830	6054.175	6624.316	6793.437	7316.567	8243.589
IP-JX（江西）	3487.269	3991.490	4209.327	4787.606	5088.315	5533.688	6329.311
IP-LN（辽宁）	3899.194	4382.250	4649.789	4968.164	5363.153	5797.010	6597.088
IP-NMG（内蒙古）	3189.414	3774.804	4383.706	4780.090	5063.228	5502.873	6038.922
IP-SD（山东）	4461.934	5049.407	5412.555	5849.909	6477.016	6975.521	7668.036
IP-SH（上海）	7489.451	8209.037	8773.100	10770.09	11432.20	12883.46	13183.88
IP-SX（山西）	3431.594	3869.952	4156.927	4360.050	4546.785	5401.854	6335.732
IP-TJ（天津）	5474.963	6409.690	7146.271	7734.914	8173.193	8852.470	9375.060
IP-ZJ（浙江）	6446.515	7158.288	7860.341	8530.314	9187.287	10485.64	11822.00

资料来源：《中国统计年鉴》1997-2003。

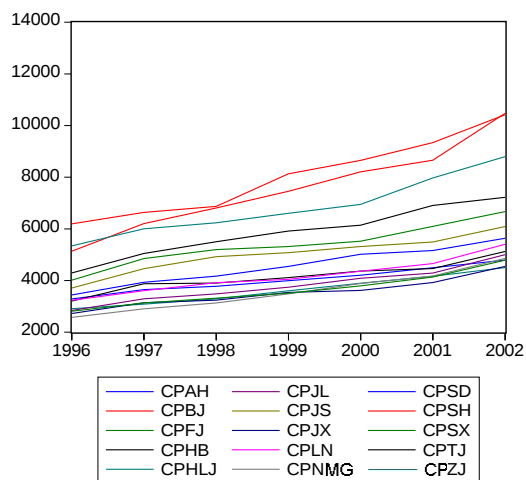


图 2 15 个省级地区的人均消费序列（纵剖面）

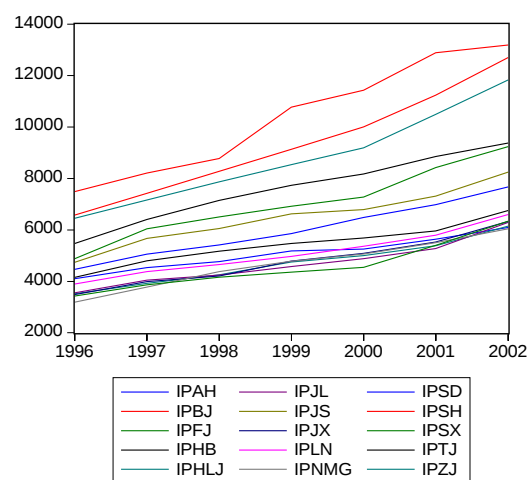


图 3 15 个省级地区的人均收入序列

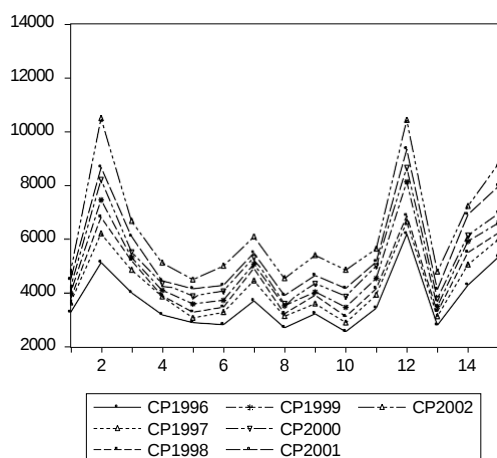


图 4 7 个时点人均消费横截面数据（含 15 个地区）
（每条连线数据表示同一年度 15 个地区的消费值）

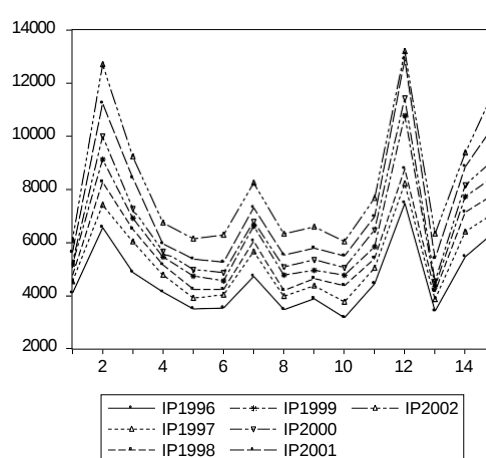


图 5 7 个时点人均收入横截面数据（含 15 个地区）
（每条连线数据表示同一年度 15 个地区的收入值）

15 个地区 7 年人均消费对收入的面板数据散点图见图 6 和图 7。图 6 中每一种符号代表一个省级地区的 7 个观测点组成的时间序列。相当于观察 15 个时间序列。图 7 中每一种符号代表一个年度的截面散点图（共 7 个截面）。相当于观察 7 个截面散点图的叠加。

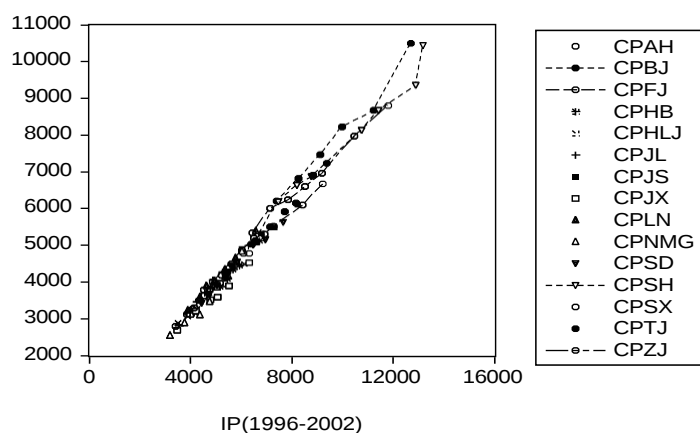


图 6 人均消费对收入的面板数据散点图（15 个时间序列叠加）

以上内容仅为本文档的试下载部分，为可阅读页数的一半内容。如要下载或阅读全文，请访问：<https://d.book118.com/128024003013006042>