

文章编号:1000-582X(2006)06-0144-05

基于主成分分析的住宅项目特征定价模型*

陈安明

(重庆大学 建设管理与房地产学院, 重庆 400030)

摘要:讨论了房地产特征价格模型在确定新建住宅项目单位销售均价中的应用,针对应用中出现的问题,将主成分分析运用于房地产特征价格模型,建立了基于特征因素主成分分析的住宅项目特征定价模型.通过案例,分析了特征隐含价格的经济意义与作用,利用所建模型较准确地确定新建住宅项目的单位销售均价.

关键词:特征价格模型;主成分分析;计量经济模型;隐含价格

中图分类号:F293.3

文献标识码:A

1 房地产特征价格模型及应用

1.1 房地产特征价格模型与应用

特征价格模型(Hedonic price model 又译为享乐价格模型、隐含价格模型或按质论价模型)是国外用于分析处理异质商品差异特征与商品价格关系的模型,美国学者 Lancaster (1966)提出的消费者理论与美国经济学家 Rosen(1974)提出的供求均衡模型构成了特征价格模型的理论基础^[1].住宅属于典型的异质产品,可将其看成由一系列内在特征(characteristics)构成的集合,人们购置的正是这一能够体验到的特征集合.住宅产品内在特征的质与量,确定了人们使用住宅效用的大小.因此,人们购买住宅产品所愿支付的价格,也应该由与住宅产品特征的质与量相对应的多种价格组成.这些反映住宅产品内在特征质与量的一系列价格,被称之为住宅的特征价格,由于其不可观测性,通常又称为隐含价格.

根据 Lancaster 理论,住宅产品的价格可由住宅所有特征构成的特征价格函数(hedonic price function)确定.假设住宅的特征集为 $x = (x_1, \dots, x_m)$, 价格为 p , 住宅特征价格函数记为: $P = f(x)$, 特征函数有线性方程、半对数方程、双对数方程、对数方程和其他非线性方程形式.

住宅特征 x_i 的隐含价格定义为:在其他因素不变的情况下,特征 x_i 每增加一个单位所导致的价格 p 的改变.即: $\frac{\partial p}{\partial x_i} = \frac{\partial f}{\partial x_i}(x)$.

房地产特征价格模型,是这样一种定价方法,通过采集房地产样本的价格与特征因素数据,对选定的特征价格方程进行多元回归分析(称之为特征回归 hedonic regression),对住宅所有特征进行隐含定价,最终确定住宅产品的特征价格函数,并根据特征价格函数进行定价与分析.

在国外,房地产特征价格模型得到广泛研究与应用,其应用主要体现在如下各方面^[1-2]:1)编制各种房地产价格指数;2)对传统城市经济模型进行改进;3)对环境质量进行测度;4)进行住宅特征隐含价格系数解释;5)对房地产、非市场物品的价值评估;6)进行住房公共提供、住房价格补贴、地方税收等政府公共政策效果评估.

在国内,房地产特征价格模型的研究与应用文献较少,贾生华、温海珍(2004)^[1]对国外房地产特征价格模型理论发展与应用进行了综述性分析,温海珍、贾生华(2004)^[3-4]利用房地产价格模型对存量房地产进行估价分析.

利用特征定价模型确定新建住宅项目的销售均价是一种新尝试,应用中存在一些问题需分析解决.

1.2 应用中的问题分析与解决途径

对于新建住宅项目,开发商首先确定整个项目均价,然后确定各楼栋均价,最后根据调价因素确定各套住宅的价格.因此,项目均价是房地产项目定价的基础.应用房地产特征价格模型测算新建住宅项目销售均价,存在3个方面问题:

* 收稿日期:2006-01-20

作者简介:陈安明(1965-),男,重庆开县人,重庆大学讲师,硕士,主要从事数量经济、系统分析、房地产经济的研究.

1) 与以前的应用^[5]相比,考虑的样本对象不同。此应用以住宅项目整体为样本对象,而不是个体住宅。影响定价的特征因素的选取,应以项目总体特征与所处环境外在特征为主。

2) 样本数据局限在目标项目所在地,并以住宅项目为单位进行收集,样本数量十分有限。

3) 部分住宅特征因素间存在较高相关性。

对于后2个问题,如果直接建立价格变量与特征因素变量的计量经济模型,在一定置信度下,不少消费者能明显体验到的,反映住宅特征的解释变量可能不会出现在模型中。为此,需要寻找一种方法,对众多可能相关的住宅特征因素施行变换,由较少的、相互独立的、能集中全部原始变量较多信息的若干综合因素,替代原始变量进行建模分析。主成分分析法是一种客观的较理想的方法^[6],将主成分分析引入房地产特征价格模型,建立基于住宅特征因素主成分分析的住宅项目特征价格模型。

2 基于主成分分析的住宅项目特征价格模型

2.1 住宅项目特征因素的主成分分析

设 n 个样本住宅项目的 m 个特征因素量化值构成的矩阵为:

$$X = \begin{bmatrix} x_{11} & x_{21} & \cdots & x_{m1} \\ x_{12} & x_{22} & \cdots & x_{m2} \\ \vdots & \vdots & \vdots & \vdots \\ x_{1n} & x_{2n} & \cdots & x_{mn} \end{bmatrix}$$

对上述矩阵进行标准化处理,令

$$y_i = \frac{x_i - \bar{x}_i}{s_i}, i = 1, \dots, m. \quad (1)$$

$$y_{ij} = \frac{x_{ij} - \bar{x}_i}{s_i} \quad i = 1, \dots, m; j = 1, \dots, n. \quad (2)$$

其中 $\bar{x}_i = \frac{1}{n} \sum_{j=1}^n x_{ij}$, $s_i^2 = \frac{1}{n-1} \sum_{j=1}^n (x_{ij} - \bar{x}_i)^2$.

变化后的特征指标向量为 y_1, \dots, y_m , 样本值为 y_{1j}, \dots, y_{mj} , $j = 1, \dots, n$.

设原始观测数据矩阵 X 的相关系数矩阵 R 的特征值为 $\lambda_1 > \lambda_2 > \dots > \lambda_m$, λ_i 对应的特征向量为:

$$c_i = (c_{i1}, c_{i2}, \dots, c_{im})^T, i = 1, 2, \dots, m.$$

因素矩阵经标准化处理后矩阵 Y 的主成分为:

$$F_i = c_{i1}y_1 + c_{i2}y_2 + \dots + c_{im}y_m, i = 1, \dots, m. \quad (3)$$

若主成分累计贡献率:

$$\sum_{j=1}^k (\lambda_j / \sum_{i=1}^m \lambda_i) > 85\% (k < m),$$

则选择前 k 个主成分为主分量:

F_1, F_2, \dots, F_k , 这 K 个主分量相互独立,反映了原始变

量中较多信息,可作为进一步建立住宅项目特征定价模型中的解释变量。

2.2 基于主分量的住宅项目特征定价模型

以住宅项目单位销售均价 p 为自变量,以其特征主分量 F_1, F_2, \dots, F_k 为解释变量,采用线性(也可选半对数)特征方程建立住宅项目特征价格模型:

$$p = \beta_0 + \beta_1 F_1 + \beta_2 F_2 + \dots + \beta_k F_k + \varepsilon. \quad (4)$$

对样本数据

$$p_j, y_{1j}, \dots, y_{mj}, j = 1, \dots, n.$$

经计算得主分量模型样本数据

$$p_j, F_{1j}, F_{2j}, \dots, F_{kj}, j = 1, \dots, n.$$

利用此样本数据,采用最小二乘法(OLS)与逐步回归技术,对式(4),在设定的置信度下,选择显著的解释变量,进行参数估计,不失一般性,设经参数估计后的模型为:

$$p = \hat{\beta}_0 + \hat{\beta}_1 F_1 + \hat{\beta}_2 F_2 + \dots + \hat{\beta}_k F_k. \quad (5)$$

将式(3)与式(1)分别代入模型估计结果式(5),得到基于主成分分析的住宅项目特征定价模型:

$$p = \hat{\alpha}_0 + \hat{\alpha}_1 x_1 + \hat{\alpha}_2 x_2 + \dots + \hat{\alpha}_m x_m. \quad (6)$$

需要说明的是,式(4)表示的模型可以有不同的形式。可根据情况,选择部分特征因素进行主成分分析,由确定的主分量与余下变量建立混合模型。

3 应用案例

选择重庆市某区2005年前后开发的23个住宅项目,利用前21个作为分析样本,建立基于主成分分析的住宅项目特征定价模型。

3.1 模型设立及参数估计

3.1.1 样本项目特征因素的选择及量化

案例所选择的因变量、特征因素自变量及原始数据见表1。

建筑类型: 根据各类建筑成本差异,超高层取值1.2,高层取值1,多层取值0.8,混合型则取相应平均值。这具有一定主观相对性,经试算,改变取值比例,只影响该系数与模型常数项。

交通: 按项目附近经过的公共汽车线路数,根据车站与项目距离远近分别取相应权重加权计数。

交易时间: 以2004年12月取值为0,前后以月为单位计数,在后取正数,反之取负数。

是否装修: 装修房取1,清水房取0。其他因素,根据相关内容综合评价取值,取值范围限定在0~5,精确到0.5,具体说明如下:

户型与设施: 根据项目提供户型的实用性,小区内设施情况综合评价取值;建筑外观风格,根据建筑外观效果,立面使用建材,项目建设规划等因素综合评分

取值;地段,选定研究地区内若干政治经济文化中心,根据其重要度确定各中心点的地段得分,然后根据区位中心辐射原理,视项目所在地与中心点的位置关系,交通情况进行综合评分取值;周边环境,综合考虑加分因素、扣分因素及突出特点后评分取值,考虑的加分因素有:项目临江、依山、邻近大学或公园、小区占地面积及绿化率大小、小区内是否有大型人造景观等,考虑扣分的因素有:靠近高速公路、铁路、城市主干道、坡度较大的道路、有噪声或粉尘的工厂等;教育配套,根据小区内及周边幼儿园、小学、中学、大学和其他教育培训机构设置情况,是否重点学校来综合评分取值,对在销售合同中开发商承诺可免费到附近重点中小学校入学的项目加分;生活配套,根据项目所在地附近一定范围内拥有各种商场、超市、银行、邮政电讯营业部、农贸市场数量多少、规模大小,与本地区重要生活设施场所的距离远近综合评分取值;适合投资,以项目所在位

置、项目户型大小,判断其基于出租、出售的增值潜力,并据此综合评分取值;营运评价,根据项目开发商经营资质,开发业绩,项目设计、施工、物业管理单位资信情况,项目营销广告情况等综合评价取值。

3.1.2 样本特征因素的主成分分析

选择体现项目个别特征与外在性特征的前 10 个因素 $x_1 - x_{10}$, 对前 21 个样本数据,利用 SAS 统计软件进行主成分分析,其相关系数矩阵 R 的特征根及累计贡献率见表 2。前 5 个特征值累计贡献率为 90.3%, 超过 85%, 故对于形如(3)式的主成分,选择前 5 个为主分量,此 5 个主成分系数(即前 5 个特征向量对应的特征向量)见表 2 右半部分。对于原始样本数据及样本均值与方差(见表 1),经过变换(2)代入各主分量表达式中,得到 5 个主分量的样本数据: $F_{1j}, F_{2j}, \dots, F_{5j}$ $j=1, \dots, 21$ 。

表 1 住宅项目特征因素量化与模型预测值

变量	p	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}	x_{11}	x_{12}	\hat{p}	
内容与项目	销售均价 /元·米 ⁻²	户型与设施	建筑类型	建筑外观风格	地段	交通	周边环境	教育配套	生活配套	适合投资	营运评价	交易时间	是否装修	预测值	相对误差/%
1	3 300	3.5	1.1	4.0	3.0	10	4.0	4.0	3.5	2.5	3.5	5.0	0	3 365	1.97
2	3 400	4.0	1.2	4.0	3.5	12	4.0	3.0	3.5	3.0	5.0	5	0	3 566	4.89
3	4 400	4.0	1.2	3.5	5.0	20	4.0	4.0	4.5	4.5	4.0	4.0	1.0	4 400	0.00
4	3 200	3.5	1.2	3.5	3.5	10	3.5	3.0	3.0	2.5	3.0	4	0	3 118	-2.57
5	3 280	3.0	1.2	3.0	4.0	16	3.5	3.5	3.5	4.0	3.0	7	0	3 338	1.77
6	3 450	3.5	1.0	4.0	3.0	13	4.5	2.5	3.5	2.0	4.0	6	0	3 347	-2.98
7	3 500	3.5	1.2	4.0	4.0	24	4.0	3.5	4.0	3.0	3.5	7	0	3 672	4.92
8	3 180	3.0	1.1	3.0	4.0	15	3.5	3.0	3.5	3.5	3.5	3	0	3 171	-0.29
9	3 000	3.0	1.2	3.0	3.0	19	3.5	3.0	3.0	3.5	3.0	3	0	3 046	1.53
10	2 680	3.0	0.9	3.0	2.0	8	3.0	3.0	2.5	1.5	3.0	5	0	2 630	-1.83
11	3 600	4.0	1.2	3.5	3.5	11	3.5	4.0	3.5	4.5	3.0	3	0	3 376	-6.23
12	2 750	2.5	1.0	2.5	3.5	15	3.0	3.5	3.0	3.0	3.0	3	0	2 840	3.28
13	3 200	3.0	0.8	3.0	4.0	14	4.0	3.0	3.5	1.5	3.5	5	0	3 098	-3.19
14	3 800	3.0	1.2	4.0	4.5	16	4.0	4.5	4.5	4.5	3.0	3	0	3 667	-3.51
15	2 900	3.0	1.2	3.0	3.5	17	4.0	3.0	3.0	3.5	3.0	-7	0	2 873	-0.92
16	2 800	2.5	1.0	3.5	3.0	15	3.5	2.5	3.0	3.5	2.5	1	0	2 825	0.88
17	3 400	3.5	1.2	4.0	3.5	11	4.0	4.5	3.5	4.5	4.0	0	0	3 475	2.22
18	3 000	3.0	1.2	3.5	3.5	16	3.5	4.0	3.0	3.5	3.0	-5	0	2 983	-0.56
19	3 350	3.5	1.0	3.5	4.0	14	4.0	3.5	3.5	2.5	3.5	6	0	3 377	0.81
20	2 320	2.5	0.8	2.5	3.0	15	3.0	3.0	3.0	3.0	2.5	-7	0	2 390	3.03
21	3 700	3.0	1.0	4.0	4.0	24	4.0	4.0	4.0	3.5	4.0	7	0	3 651	-1.31
22	3 500	3.5	1.1	4.0	3.5	12	4.0	3.0	3.5	3.0	4.0	7	0	3 457	-1.22
23	3 200	3.0	1.2	3.0	3.5	17	4.0	3.0	3.0	3.5	3.0	6	0	3 220	0.62
新项目	-	4.0	1.2	3.5	4.5	20	3.5	4.0	3.5	3.0	4.5	8	0	-	-
\bar{x}_i	-	3.214 2	1.090 5	3.428 6	3.571 4	15	3.714 3	3.428 6	3.428 6	3.214 3	3.357 1	-	-	-	-
s_i	-	0.462 9	0.137 5	0.507 1	0.638 1	4.219 0	0.405 3	0.596 1	0.507 1	0.916 1	0.594 6	-	-	-	-

说明:表中 \bar{x}_i, s_i 分别为前 21 个样本数据的样本均值与方差,表中数据根据楼盘调查资料综合评定确定。

表2 相关系数矩阵特征值与主分量系数

λ	特征值	累计贡献	变量	主分量				
				F_1	F_2	F_3	F_4	F_5
λ_1	4.497	0.450	γ_1	0.321	-0.341	-0.279	0.075	-0.425
λ_2	1.985	0.648	γ_2	0.295	0.140	-0.479	0.569	-0.010
λ_3	1.308	0.779	γ_3	0.351	-0.301	-0.087	-0.031	0.610
λ_4	0.678	0.847	γ_4	0.347	0.283	0.280	-0.105	-0.453
λ_5	0.564	0.903	γ_5	0.181	0.360	0.527	0.404	0.210
λ_6	0.349	0.938	γ_6	0.347	-0.317	0.254	0.105	0.259
λ_7	0.243	0.962	γ_7	0.295	0.258	-0.285	-0.637	0.193
λ_8	0.192	0.982	γ_8	0.414	0.110	0.263	-0.252	-0.054
λ_9	0.130	0.995	γ_9	0.260	0.465	-0.314	0.119	0.011
λ_{10}	0.055	1.000	γ_{10}	0.297	-0.410	0.127	0.024	-0.301

3.1.3 基于主成分特征定价模型及参数估计

以主分量 F_1, \dots, F_5 与因素 x_{11}, x_{12} 为特征定价模型的解释变量,以项目销售均价为因变量,选择线性形式特征方程建立如下模型:

$$p = \beta_0 + \beta_1 F_1 + \beta_2 F_2 + \beta_3 F_3 + \beta_4 F_4 + \beta_5 F_5 + \beta_6 x_{10} + \beta_7 x_{12} + \varepsilon \quad (7)$$

用最小二乘法与逐步回归技术进行参数估计,得到式(7)对应模型的估计式(8).

$$p = 3148.413 + 153.099F_1 + 26.659x_{11} + 108.41 (<0.0001) + 11.78 (<0.0001) + 4.51 (0.0003) + 547.099x_{12}, \quad R^2 = 0.9529, \quad \bar{R}^2 = 0.9446, \\ (4.57 (0.0003) \quad F = 114.59 (<0.0001)) \quad (8)$$

其中,括号内数据为参数的 t 统计量及对应概率,说明各回归系数是显著的,模型 F 检验值及拟合优度及修正拟合优度说明模型拟合效果好.

将主分量 F_1 关于 γ_1 形如式(3)的表达式代入式(8),再将由表1最后两行数据确定的变换式(1)代入,最后得到由21个样本项目数据建立的基于主成分分析的住宅项目特征价格模型:

$$p = 120.6 + 106.05x_1 + 328.06x_2 + 106.06x_3 + 83.15x_4 + 6.55x_5 + 130.94x_6 + 75.56x_7 + 124.92x_8 + 43.42x_9 + 76.55x_{10} + 26.66x_{11} + 547.10x_{12} \quad (9)$$

3.2 模型的经济意义及应用

式(9)中,特征因素前的系数为其隐含价格,反映了特征因素对项目销售均价的影响的大小.交易时间每延后1个月,价格将上调26.66元,装修单价为547.1元,与该地区2005年的实际情况相符.

利用模型对21个样本数据进行测算,测算值及与样本原始值相对误差列入表1,除少数几个样本数据外,多数预测误差在3%以内,样本平均绝对误差为73.7,相对平均绝对误差为2.25%,预测误差的标准差为93.5,泰尔不等系数为0.0143,模型预测效果好.利用两个未参与模型参测的销售均价与实际均价

误差在2%以内,说明模型用于新建住宅项目均价的定价的效果良好.利用模型对新建项目进行定价,某项目,拟在2005年8月开盘,相关特征为超高层清水房,其它特征见表1新项目行.将其代入式(9)得均价预测值为3700元,要体现出消费者对能体验到的项目特征的认可,应围绕每平方米3700元确定项目销售均价.

4 结论

1)与住宅项目均价的传统定价法相比,利用特征定价模型确定新建项目均价更能体现供求双方对房地产这一异质商品内在特征的质与量的共同认可度.开发商也可将特征定价结果作为项目开发成本的控制依据;

2)引入主成分分析法建立模型,既克服了直接利用特征因素建立模型所面临的样本数据少、解释变量间存在一定相关性的困难,提高了模型模拟效果,又能使较多的主要特征因素进入最终模型,定价时可供考虑的因素较多,定价较准确;

3)对于模型中体现的特征因素隐含价格而言,虽然对其进行经济意义解释时要慎重,但也在一定程度上说明了消费者对自己能体验到的项目特征在价格上的认可程度,房地产开发商可以参考这些特征隐含价格,了解本地区消费者对住宅项目的消费偏好,开发出适销对路,具有高附加值的住宅项目产品.

应注意的是,模型的应用要受到地域与时间的限制.

参考文献:

- [1] 贾生华,温海珍. 房地产特征价格模型的理论发展及其应用[J]. 外国经济与管理, 2004, 26(5): 42-44.
- [2] MALPEZZI S. Hedonic Pricing Models: A Selective and Applied Review [EB/OL]. <http://www.bus.wisc.edu/wcre/paper.asp>, 2005-08-20.
- [3] 温海珍,贾生华. 基于特征价格的房地产评估新方法[J]. 外国经济与管理, 2004, 26(6): 31-35.
- [4] 温海珍,贾生华. 住宅的特征与特征价格——基于特征价格模型的分析[J]. 浙江大学学报(工学版), 2004, 38(10): 1338-1442.
- [5] STEPHANIE E VANDERFORD, YOKO MINURA, ANNE L SWEANEY. A Hedonic Price Comparison of Manufactured and Site-Built Homes in the Non-MSA U.S. [J]. The Journal of Real Estate Research, 2005, 27(1): 83-104.
- [6] MARIUS THERIAULT, FRANCOIS DES ROSIERS, PAUL VILLENEUVE, et al. Modeling Interactions of Location with Specific Value of Housing Attributes [J]. Property Management, 2003, 21(1): 25-48.

Housing Project Hedonic Price Model Based on Principal Components Analysis

CHEN An-ming

(College of Construction Management and Real Estate, Chongqing University, Chongqing 400030, China)

Abstract: Real estate hedonic price model is applied to determine the average sale unit price of newly built housing project. In order to solve the problems of the lack of sample and the high correlation of some of the housing characteristic factors with each other in this application, a housing project hedonic pricing model based on principal components analysis of housing characteristic factors has been established. Through a case, the author has analyzed the economical significance of the implicit prices. With the model, the average sale unit price of a newly built housing project can be determined by rule and line.

Key words: hedonic price model; principal components analysis; econometric model; implicit price

(编辑 姚 飞)

(上接第 134 页)

Empirical Research on Value Strategies in China Security Market

(LU Da-yin, LIN Cheng-dong, YANG Chao-jun)

(Management School, Shanghai Jiaotong University, Shanghai 200052, China)

Abstract: The empirical research shows that the value strategies that buys historically low price, high book-to market ratios, or high sale-price of A stocks, can outperform the glamour strategies that buys historically high price, low book-to market ratios, or low sale-price of A stocks after formation-period in China Security Market. CAPM can not explain the phenomena, but FF's three factors pricing model can.

Key words: value strategies; price; book-to-market ratios; sale-price ratios; earnings-price ratios

(编辑 姚 飞)