

价格聚类视角下商品住宅价格 影响因素实证研究

笪可宁^{1,2}, 杨贺贞¹

(1. 沈阳建筑大学管理学院, 辽宁 沈阳 110168; 2. 沈阳化工大学人文与社会科学学院, 辽宁 沈阳 110142)

摘 要:近年来,居高不下的商品住宅价格在各种调控政策下并没有得到有效控制,且不同城市间的调控结果也各不相同。为探究城市间商品住宅价格影响因素的差异性,从城市等级视角出发,利用全国35个大中城市的相关数据,探究商品住宅开发成本、生产总值和人口密度3个变量对不同等级城市商品住宅价格的影响程度,以期政府因城施策调控房价提供参考。

关键词:商品住宅价格;凝聚聚类;城市等级;影响因素

中图分类号:F293.353 **文献标志码:**A

在房地产行业快速发展带动经济持续增长的同时,商品住宅价格也在不断上涨,这不仅不利于房地产市场的稳定,而且增加了人们的生活压力。住房问题关乎民生,中央政府不断出台相关政策,从金融、土地、限购等方面进行调控。虽然政府相继出台调控政策,但是对商品住宅价格的调控效果却并不乐观。

由于我国地域辽阔、人口众多、城市发展不均衡,商品住宅价格也表现出地域性差异,东部发达地区的商品住宅价格明显高于西部落后地区。而城市间发展水平不同,影响商品住宅价格的主要因素也会存在差异,即使相同的调控政策也会出现不同的调控结果。

探究不同等级城市商品住宅价格的影响因素,能为各地政府更好地调控商品住宅价格、落实因城施策、建立房地产市场调控长效机制提供参考,对维持房地产行业的稳定发展具有重要意义。

吴焕军^[1]通过分析土地政策对房地产

市场的影响,得出了土地政策对房价影响作用较小的结论;任行伟等^[2]从影子银行和房地产价格之间的作用机理出发,认为影子银行规模扩大对商品房价格有正向影响;周小平等^[3]利用杜宾模型,分析2010—2017年我国35个大中城市的地价房价比,指出住宅地价房价比受土地供应、房地产开发规模、人均可支配收入3种因素的影响,区域间绝对值和变化趋势的差异较为明显;李超等^[4]从资源配置的角度进行分析,指出城市资源配置是决定城市房价的关键性因素;李继玲^[5]从人口因素、宏观因素、房地产行业因素、市场需求4个方面进行实证研究,发现城镇化因素对推动商品房价格上涨的影响最显著;唐坚^[6]指出短期商品房价格过快上涨或与城市人口增加和商品房竣工面积不足有关;李成等^[7]指出要根据“因城施策”导向加强不同城市房价的差异化调控,形成动态差别化供给房地产的长效机制。

虽然许多学者都对商品住宅价格的影响因素作过研究,但以商品住宅价格为城市等级划分依据,研究不同等级城市商品住宅价格影响因素的较少。以商品住宅价格来划分城市等级,利用 MATLAB 软件,采用凝聚聚类方法,合并相似簇的原理,把具有相似簇的城市聚集到一起,研究结果更具准确性。

笔者通过文献分析,在已有研究的基础上综合考虑相关因素,结合研究重点,选取 2010—2018 年全国 35 个大中城市数据,探

究城市等级视角下商品住宅价格的主要影响因素。

一、城市等级划分及指标选择

房价调控政策的制定要依据城市等级而有所差异^[8]。城市发展不均衡导致各地商品住宅价格走势也各不相同。由 2011—2018 年全国 35 个大中城市商品住宅价格走势(见图 1),可以看出商品住宅价格的差异性十分明显。

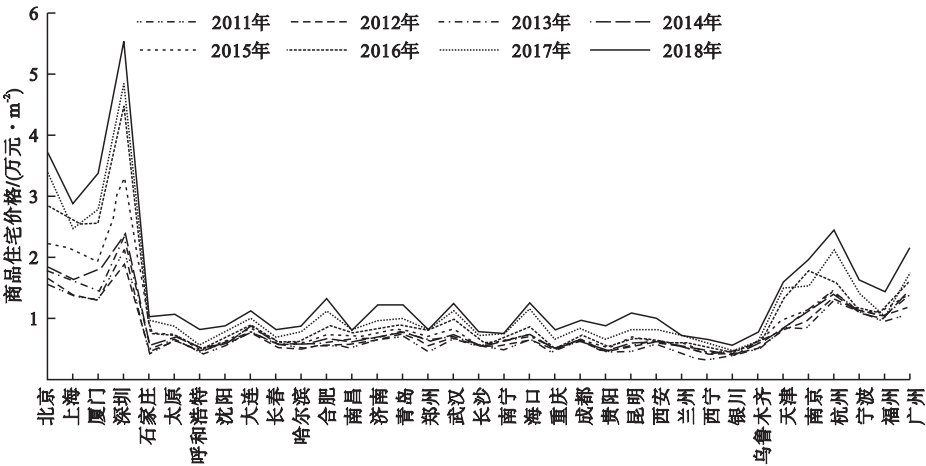


图 1 2011—2018 年我国 35 个大中城市商品住宅价格变化趋势

1. 凝聚聚类

聚类指按照某种相似的度量,把相似的样本(数据点)归为一类^[9]。由聚类所生成的簇是一组数据对象的集合,这些对象与同一个簇中的对象彼此相似,与其他簇中的对象相异。

假设有 N 个待聚类的样本,对于凝聚聚类来说,基本步骤如下:

- ①(初始化)把每个样本归为一类,计算每两个类之间的距离,也就是样本与样本之间的相似度;
- ②寻找各个类之间最近的两个类,把他们归为一类(这样类的总数就少了一个);
- ③重新计算新生成的这个类与剩余类之间的相似度;
- ④重复步骤②和步骤③,直到所有样本点都归为一类结束。

同一种类城市的商品住宅价格在坐标系中具有更加短的空间距离,而且对于城市分

类的个数无法提前确定。因此,采用凝聚分类法,利用商品住宅价格,对城市进行分类,城市的种类个数由算法确定。

2. 城市等级划分

单纯以城市优质资源集聚度或资源禀赋为依据划分城市等级,不能准确反映商品住宅价格的影响因素,而以商品住宅价格为聚类变量,把相似簇的城市聚集到一类,对于政府因城施策更具有指导意义。以 2010—2018 年全国 35 个大中城市商品住宅销售均价为聚类变量,采用凝聚聚类法,运用 MATLAB 软件将城市分成 3 类(见表 1)。

表 1 全国 35 个大中城市等级划分

| 类别 | 城市 | 合计 |
|-----|--|----|
| A 类 | 北京、上海、深圳、厦门 | 4 |
| B 类 | 广州、天津、南京、杭州、宁波、福州 | 6 |
| C 类 | 沈阳、武汉、重庆、合肥、西安、长沙、大连、成都、郑州、长春、太原、石家庄、呼和浩特、哈尔滨、兰州、青岛、西宁、银川、南宁、海口、贵阳、昆明、济南、南昌、乌鲁木齐 | 25 |

根据聚类分层结果,整理2010—2018年3类城市商品住宅销售均价走势(见图2),可以看出3类城市之间商品住宅价格差异较大,A类城市的商品住宅销售价格最高,B类城市处于中间水平,C类城市销售价格相对较低。

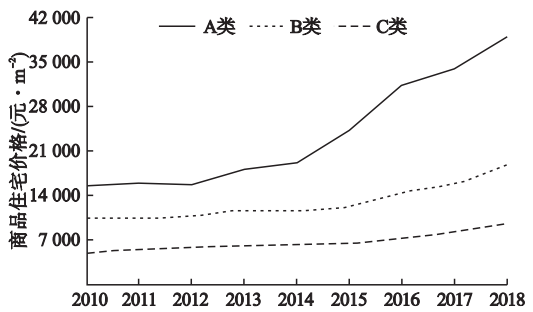


图2 2010—2018年3类城市商品住宅价格走势

3. 指标选择

通过对比分析3类城市,发现商品住宅价格较高的地区经济发展水平较高,商品住宅价格会受当地经济发展水平的影响,因此,选择地区生产总值(GDP)为衡量指标,考察经济水平对商品住宅价格的影响。商品住宅价格由开发成本和开发利润构成,同时受到供求关系的影响。在投资额一定的条件下,开发成本的增加会导致开发数量的下降,即市场供应量降低。考虑到开发成本和市场供给面积之间存在的线性关系,以及市场上商品住宅供给面积数量的可获得性,以开发成本为衡量指标,同时体现供给数量以及开发成本对商品住宅价格的影响。

人作为需求市场的承担者,人口越多,市场需求量就越大,单纯以人口数量来衡量市

场需求,会忽略城市面积对商品住宅开发的限制,因此,以每平方千米人口的数量为指标,考察人口密度对商品住宅价格的影响,既考虑需求量对商品住宅价格的影响,也考虑城市可开发面积对价格带来的影响。

最终选取生产总值、商品住宅开发成本和人口密度3个因素为解释变量,选取商品住宅平均销售价格为被解释变量,研究城市等级视角下,3个因素对商品住宅价格的影响程度,变量说明如表2所示。

表2 变量说明

| 变量 | 单位 | 符号 | 性质 |
|----------|-------------------|-----------|-------|
| 商品住宅销售均价 | 元/m ² | <i>P</i> | 被解释变量 |
| 商品住宅开发成本 | 元/m ² | <i>HI</i> | |
| 生产总值 | 亿元 | GDP | |
| 人口密度 | 人/km ² | <i>PD</i> | 解释变量 |

二、回归模型建立

1. 数据说明

利用2010—2018年全国35个大中城市的数据进行实证研究。数据来源于国家数据网站和2010—2018年《中国城市统计年鉴》。

2. 单位根检验

为避免出现伪回归现象,根据凝聚聚类的分类结果,运用Eviews 7.2软件,对商品住宅销售均价、商品住宅开发成本、生产总值和人口密度4个指标进行单位根检验,当*P*值小于5%时,表示序列平稳,否则不平稳,检验结果如表3所示。其中ln为对原数据取对数处理。

表3 3类城市指标单位根检验结果

| 类别 | 性质 | 原数据 | | | 处理后数据 | | |
|----|-------|-----------|------------|-----|-----------------|------------|-----|
| | | 符号 | <i>P</i> 值 | 平稳性 | 符号 | <i>P</i> 值 | 平稳性 |
| A类 | 被解释变量 | <i>P</i> | 0.288 3 | 不平稳 | ln(<i>P</i>) | 0.021 2 | 平稳 |
| | | <i>HI</i> | 0.000 0 | 平稳 | ln(<i>HI</i>) | 0.013 7 | 平稳 |
| | 解释变量 | GDP | 0.455 1 | 不平稳 | ln(GDP) | 0.000 8 | 平稳 |
| | | <i>PD</i> | 0.272 0 | 不平稳 | ln(<i>PD</i>) | 0.025 1 | 平稳 |
| B类 | 被解释变量 | <i>P</i> | 0.069 9 | 不平稳 | ln(<i>P</i>) | 0.010 0 | 平稳 |
| | | <i>HI</i> | 0.002 5 | 平稳 | ln(<i>HI</i>) | 0.008 5 | 平稳 |
| | 解释变量 | GDP | 0.202 0 | 不平稳 | ln(GDP) | 0.038 8 | 平稳 |
| | | <i>PD</i> | 0.976 5 | 不平稳 | ln(<i>PD</i>) | 0.008 8 | 平稳 |
| C类 | 被解释变量 | <i>P</i> | 0.014 4 | 平稳 | ln(<i>P</i>) | 0.000 7 | 平稳 |
| | | <i>HI</i> | 0.045 3 | 平稳 | ln(<i>HI</i>) | 0.000 0 | 平稳 |
| | 解释变量 | GDP | 0.007 5 | 平稳 | ln(GDP) | 0.000 0 | 平稳 |
| | | <i>PD</i> | 0.000 0 | 平稳 | ln(<i>PD</i>) | 0.000 0 | 平稳 |

由表 3 可知,3 类城市中,部分变量显示非平稳,再对所有变量数据进行取对数处理,并对处理结果进行单位根检验,检验结果显示均平稳。

3. 多重共线性检验

为检验所选择变量之间是否存在共线性,运用综合检验法进行检测,检测结果为:A 类城市最终选定 GDP 和商品住宅开发成本作为自变量;B 类城市选定商品住宅开发成本和 GDP 为自变量;C 类城市选定 GDP、

商品住宅开发成本和人口密度为自变量。

4. 协整检验

在通过单位根检验和多重共线性检验之后,为验证选定的解释变量与被解释变量之间是否存在长期稳定关系,还需进行协整检验,检验结果如表 4 所示。由表 4 可知,3 类城市的协整检验结果对应的 P 值均小于 5%,即解释变量与被解释变量存在长期稳定关系。其中,ADF 为检验方法; t -Statistic 为 T 统计量值;Prob 为检验 P 值。

表 4 协整检验结果

| 城市类别 | 变量 | 检验方法 | t - Statistic | Prob |
|------|---|------|---------------|---------|
| A | $\ln(P)$ 、 $\ln(HI)$ 、 $\ln(GDP)$ | ADF | -2.81 | 0.002 5 |
| B | $\ln(P)$ 、 $\ln(HI)$ 、 $\ln(GDP)$ 、 | | -2.13 | 0.016 5 |
| C | $\ln(P)$ 、 $\ln(HI)$ 、 $\ln(PD)$ 、 $\ln(GDP)$ | | -2.28 | 0.011 2 |

三、实证分析

通过单位根、多重共线性和协整检验后,进行实证分析。首先对 3 类城市进行类型假设,3 类城市的面板数据回归模型^[10]设定为

$$\ln(P) = \alpha_i + \beta_1 \ln(HI_{i,t}) + \beta_2 \ln(GDP_{i,t}) + \beta_3 \ln(PD_{i,t}) + U_{i,t}$$

式中: i 为城市; t 为年份; α_i 为常数项; β_i 为系数向量; $U_{i,t}$ 为误差项。

3 类城市计算结果和类型判断如表 5 所示。

表 5 3 类城市 F 值计算结果

| 类别 | S_1 | S_2 | S_3 | N | T | K | F_1 | F_2 | F^1 | F^2 | 类型 |
|-----|-------|-------|-------|-----|-----|-----|-------|--------|---------|---------|-----|
| A 类 | 0.40 | 0.44 | 1.94 | 4 | 9 | 2 | 0.40 | 10.27 | 2.508 2 | 2.300 2 | 变截距 |
| B 类 | 0.40 | 0.60 | 1.50 | 6 | 9 | 2 | 1.80 | 6.60 | 2.106 0 | 1.954 3 | 变截距 |
| C 类 | 1.48 | 2.98 | 78.74 | 25 | 9 | 3 | 3.18 | 108.76 | 1.442 5 | 1.382 5 | 变系数 |

注: S_1 为变系数模型的残差平方和; S_2 为变截距模型的残差平方和; S_3 为混合模型的残差平方和; F^1 、 F^2 为相应自由度下的临界值; N 为城市个数; T 为统计时间; K 为自变量个数。

若 $F_2 > F^2$, $F_1 < F^1$, 则为变截距类型;若 $F_2 > F^2$, $F_1 > F^1$, 则为变系数类型。通过 Hausman 检验,若 Prob < 5%, 则为固定型,否则为随机型,3 类城市检验结果如表 6 所示。利用 Eviews 7.2 软件,进行回归分析(见表 7~表 9)。其中,Chi - Sq. Statistic 为统计量;Chi - Sq. d. f. 为自变量个数;Prob 为 P 值; C 为常数项;Variable 为变量;Coefficient 为系数;Adjusted R - squared 为修正决定系数,该值小于 1,且越大越好。其中,A 类城市的修正决定系数为 0.905 9, F 统计量为 68.376 5;B 类城市的修正决定系

表 6 Hausman 检验结果

| 类别 | Chi - Sq. Statistic | Chi - Sq. d. f. | Prob | 类型 |
|-----|---------------------|-----------------|---------|-----|
| A 类 | 24.120 6 | 2 | 0.000 0 | 固定型 |
| B 类 | 4.129 8 | 2 | 0.124 9 | 随机型 |
| C 类 | 124.140 6 | 3 | 0.000 0 | 固定型 |

表 7 A 类城市分析结果

| Variable | Coefficient | t - Statistic | Prob |
|------------|-------------|---------------|---------|
| C | -2.57 | -3.202 4 | 0.003 2 |
| $\ln(HI)$ | 0.38 | 1.724 5 | 0.094 9 |
| $\ln(GDP)$ | 1.02 | 5.837 6 | 0.000 0 |

表 8 B 类城市回归结果

| Variable | Coefficient | t - Statistic | Prob |
|------------|-------------|---------------|---------|
| C | 2.16 | 3.383 6 | 0.001 4 |
| $\ln(HI)$ | 0.62 | 5.088 9 | 0.000 0 |
| $\ln(GDP)$ | 0.28 | 3.332 0 | 0.001 6 |

数为 0.713 2, F 统计量为 63.491 2。

A 类城市中,北京、上海、深圳的经济发展水平都较高,厦门虽然按照商品住宅价格进行分类时,属于 A 类城市,但 GDP 远低于其他 3 个城市,单从 GDP 与商品住宅价格的适应性看,厦门的商品住宅价格超过了人们的承担能力。由表 7 可知, $\ln(GDP)$ 系数为

表 9 C 类城市分析结果

| 变量 | β | 变量 | β | 变量 | β | 变量 | β |
|----------------|---------|---------------|---------|---------------|---------|---------------|---------|
| C | -6.07 | GY - ln(GDP) | 0.15 | WH - ln(PD) | 3.74 | HEB - ln(HI) | 0.29 |
| SJZ - ln(GDP) | 1.49 | KM - ln(GDP) | 1.10 | CS - ln(PD) | 2.10 | HF - ln(HI) | 0.79 |
| TY - ln(GDP) | 0.45 | XA - ln(GDP) | 0.35 | NJ - ln(PD) | 2.08 | NC - ln(HI) | 0.06 |
| HH - ln(GDP) | 0.29 | LZ - ln(GDP) | 0.66 | HK - ln(PD) | 8.07 | JN - ln(HI) | 0.05 |
| SY - ln(GDP) | -0.12 | XN - ln(GDP) | 0.84 | CQ - ln(PD) | -6.95 | QD - ln(HI) | 0.29 |
| DL - ln(GDP) | 0.66 | YC - ln(GDP) | 0.54 | CD - ln(PD) | -1.22 | ZZ - ln(HI) | 0.55 |
| CC - ln(GDP) | 0.44 | WL - ln(GDP) | 0.58 | GY - ln(PD) | 2.04 | WH - ln(HI) | 0.55 |
| HEB - ln(GDP) | 0.62 | SJZ - ln(PD) | 0.29 | KM - ln(PD) | 0.19 | CS - ln(HI) | 0.52 |
| HF - ln(GDP) | 0.95 | TY - ln(PD) | 5.39 | XA - ln(PD) | 1.52 | NJ - ln(HI) | 0.49 |
| NC - ln(GDP) | 0.75 | HH - ln(PD) | 6.10 | LZ - ln(PD) | 0.03 | HK - ln(HI) | 0.72 |
| JN - ln(GDP) | 0.42 | SY - ln(PD) | 9.40 | XN - ln(PD) | -0.20 | CQ - ln(HI) | 1.49 |
| QD - ln(GDP) | 0.56 | DL - ln(PD) | 3.92 | YC - ln(PD) | -0.65 | CD - ln(HI) | 1.45 |
| ZZ - ln(GDP) | 0.63 | CC - ln(PD) | 2.17 | WL - ln(PD) | -0.35 | GY - ln(HI) | 0.51 |
| WH - ln(GDP) | 0.7 | HEB - ln(PD) | 2.04 | SJZ - ln(HI) | 2.69 | KM - ln(HI) | 0.49 |
| CS - ln(GDP) | 0.28 | HF - ln(PD) | 3.05 | TY - ln(HI) | 0.38 | XA - ln(HI) | 0.11 |
| NJ - ln(GDP) | 0.36 | NC - ln(PD) | -0.30 | HH - ln(HI) | 0.65 | LZ - ln(HI) | 0.02 |
| HK - ln(GDP) | -0.16 | JN - ln(PD) | 2.75 | SY - ln(HI) | 0.01 | XN - ln(HI) | 0.25 |
| CQ - ln(GDP) | 0.85 | QD - ln(PD) | 3.66 | DL - ln(HI) | 0.44 | YC - ln(HI) | 0.5 |
| CD - ln(GDP) | 0.51 | ZZ - ln(PD) | -0.52 | CC - ln(HI) | 0.34 | WL - ln(HI) | 0.26 |

注:C 类城市回归模型为固定变系数;“-”前为城市字母缩写;“-”后为变量。

1.02,ln(HI) 的系数为 0.38,GDP 对商品住宅价格的影响高于商品住宅开发成本带来的影响,Adjusted R - squared 为 0.905 9,分析可靠性较高。GDP 代表经济的发展水平,也反映了人们的生活水平,若商品住宅价格超过人们的承担能力,不断升高的价格会导致房地产泡沫,不利于房地产行业健康稳定发展。

B 类城市商品住宅价格受到商品住宅开发成本和城市生产总值的影响。由表 8 可知,ln(GDP) 系数为 0.28,ln(HI) 的系数为 0.62,商品住宅开发成本对商品住宅价格的影响高于生产总值带来的影响,Adjusted R - squared 为 0.713 2,结果可靠性较高。B 类城市不断增加的商品住宅开发成本最终转移到商品住宅价格上,当商品住宅价格的上涨幅度超过人们的经济承受能力时,会不利于房地产市场的稳定,最终影响经济的发展。

C 类城市由于城市数量多,影响因素和影响程度存在差异,所以模型为固定变系数类型,即常数项不变,而解释变量的系数随着城市不同而变化。C 类城市的商品住宅价格受到生产总值、开发成本和人口密度的影响,由表 9 可知,大部分城市受到人口密度的影响最大,其次是生产总值带来的影响,而开发

成本对商品住宅价格的影响在不同城市间差异较大。

四、结 语

笔者通过对生产总值、商品住宅开发成本和人口密度进行分析,研究 3 类因素对不同等级城市的影响程度。其中,A 类城市商品住宅价格受到生产总值的影响最大;B 类城市在商品住宅开发成本和生产总值两个因素中,开发成本对商品住宅价格的影响更大;C 类城市随着城镇化水平的提高,城市人口不断增多,为房地产市场带来了大量的市场需求,C 类城市的商品住宅价格受人口密度影响最大。

A 类城市在保障商品住宅价格与中高收入人群生活水平和经济承受能力相适应的同时,可以通过增加低档住房或建设长租公寓,满足中低收入人群的需求;商品住宅开发成本对 B 类城市影响最大,主要体现在土地费用不断提高,政府可以逐步征收房产税,增加地方政府财政来源,从而减少对土地财政的依赖;C 类城市政府可以通过调整供给结构,适当增加保障住房,满足不同类型住房需求,使城镇化的发展速度与住房建设类型和规模相适应。

笔者根据商品住宅销售均价进行城市等级划分,仅考虑生产总值、商品住宅开发成本和人口密度 3 个因素,存在一定的片面性,需要进一步增加其他影响因素的论证分析,以便提高实证分析的有效性。

参考文献:

[1] 吴焕军. 土地政策在房地产调控中的政策效果评价[J]. 中南财经政法大学学报, 2011 (6): 23 - 27.

[2] 任行伟,邢天才,张鑫. 影子银行、货币政策与房地产价格[J]. 经济与管理, 2019, 33 (4): 58 - 64.

[3] 周小平,秦振扬,赵松,等. 中国住宅地价房价比的空间格局、演变特征及影响因素:基于 35 个大中城市的空间计量分析[J]. 中国土地科学, 2019, 33 (1): 40 - 48.

[4] 李超,张超. 城市资源与人口集聚:房价的中

介与调节效应[J]. 华南师范大学学报(社会科学版), 2018 (5): 125 - 133.

[5] 李继玲. 房价波动影响因素研究:基于 2005—2015 年数据的实证分析[J]. 经济问题探索, 2017 (9): 30 - 37.

[6] 唐坚. 供需均衡视角下城市房价上涨动因分析[J]. 统计与决策, 2019, 35 (14): 146 - 149.

[7] 李成,李一帆,于海东,等. 城市人口、货币政策与房地产价格:内在机理与实证检验[J]. 当代经济科学, 2020, 42 (1): 108 - 119.

[8] 藏波,杨庆媛,周滔. 不同等级城市房价与地价关系的规律初探[J]. 中国土地科学, 2011, 25 (5): 79 - 84.

[9] 王宇,李晓利. 核 k -凝聚聚类算法[J]. 大连理工大学学报, 2007 (5): 763 - 766.

[10] 罗良文,潘雅茹. 区域城镇化差异对房地产价格影响的实证分析[J]. 统计与决策, 2015 (5): 131 - 134.

An Empirical Study on the Influencing Factors of Commercial Housing Prices from the Perspective of Price Cluster

DA Kening^{1,2}, YANG Hezhen¹

(1. School of Management, Shenyang Jianzhu University, Shenyang 110168, China; 2. School of Economics and Management, Shenyang University of Chemical Technology, Shenyang 110142, China)

Abstract: In recent years, the high prices of commercial housing have not been effectively controlled under various regulation policies, and the results of regulation in different cities are also different. In order to explore the differences in the factors affecting the price of commercial housing between cities, from the perspective of urban hierarchies, using relevant data from 35 large and medium - sized cities across the country, the three variables of commercial housing development cost, gross production value and population density are explored. The influence degree of commodity housing prices is expected to provide a reference for the government to adjust housing prices due to urban policies.

Key words: commercial housing prices; agglomerative clustering; urban hierarchy; influencing factors

(责任编辑:郝 雪 英文审校:林 昊)