

基于主成分分析对地产商业价格影响因素的研究

安徽财经大学金融学院 竺天辰

摘要: 为了分析房价影响因素,本文首先借鉴历史文献中的影响因素进行回归分析,其次使用主成分分析方法消除多重共线性对模型的干扰。对数据进行平稳性检验、协整检验,并使用Eviews软件构建回归模型。影响房价因素主要是地区生产总值、全社会固定资产投资、城市人口密度和城镇居民人均可支配收入,最后结合模型对主要影响因素进行经济层面的分析并提出减缓房价上升的建议。

关键词: 房价 回归分析 主成分分析 影响因素

中图分类号: F715.53

文献标识码: A

文章编号: 2096-0298(2017)08(a)-139-04

当前我国社会中面临的突出问题,莫过于高房价。近年来房地产的快速扩张,使房价保持高昂的上升趋势。以北京为例,北京的商品住房均价从2005年的6788元每平方米上涨到2015年的22633元每平方米,平均每年以12.8%的涨幅快速上升。房价持续上涨已经成为影响中国经济安全和社会稳定的重大隐患。党和政府明确提出了住房保障政策目标,享有住房是每个人的基本权利,保障城镇居民的住房权利,是政府义不容辞的责任。而高房价已经成为这一政策目标的最大障碍。那么影响高房价的因素有哪些?又是如何影响房价的?

1 文献回顾

从20世纪90年代开始,学者们就对房地产价格的影响因素投入了大量的研究。Karl E. Case^[1]研究发现房价与人均收入有着较强的正相关性。Clapp J. M. 和 Giaccotto C^[2]研究发现人口对于住宅价格的变化有较好的预测能力。Quigley J.M^[3]的研究认为经济基本面的相关指标可以解释房地产价格的走势。Malpezzi S. A^[4]通过133个都市区域的年度数据,得出住宅价格的变化不是随机的。结合国内经济,袁志刚、樊潇彦^[5]构造了房地产市场局部均衡模型。周京奎^[6]提出了过度金融支持假说,认为过度的金融支持能影响房价的走势。余壮雄和林建浩^[7]在房地产市场局部均衡的框架下,提出了政府支持假说。还有的学者运用灰色预测模型对住宅价格进行分析,对未来房价走势作出了较准确的预测^[8]。

2 房价影响因素模型的设定

2.1 变量的选取、解释及数据处理

2.1.1 变量的选取

房价的高低受多重因素的影响,但并不是所有的都可以量化,本文在借鉴上述学者的研究成果基础上,选取北京2005年~2015年数据作为分析对象,选择可以量化的指标进行模型分析。取地区生产总值、全社会固定资产投资、房地产开发企业土地购置费用、房地产开发企业国内贷款、城市人口密度、城镇居民人均可支配收入作为解释变量。取商品房平均销售价格作为被解释变量。

2.1.2 变量的解释

(1)地区生产总值(X_1):是指本地区所有常住单位在一定时期内生产活动的最终成果。

(2)全社会固定资产投资(X_2):是以货币表现的建造和购置固

定资产活动的工作量,它是反映固定资产投资规模、速度、比例关系和使用方向的综合性指标。

(3)房地产开发企业土地购置费用(X_3):指房地产开发企业通过各种方式取得土地使用权而支付的费用。

(4)房地产开发企业国内贷款(X_4):是指房地产开发经营企业或土地储备机构发放的与房地产开发经营活动有关的贷款。

(5)城市人口密度(X_5):指生活在城市范围内的人口稀密的程度。

(6)城镇居民人均可支配收入(X_6):是指反映居民家庭全部现金收入能用于安排家庭日常生活的那部分收入。

2.1.3 数据的处理

为了消除通货膨胀因素,在处理各年数据时,以2005年为基准年,查询出各年通货膨胀率,然后将各年的名义数据进行处理,处理数据如表1、表2所示。

2.2 模型的建立

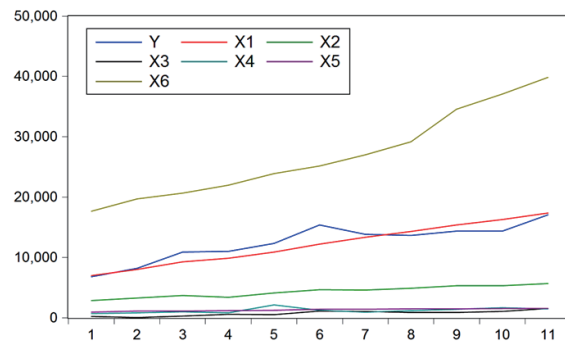


图1 Y 、 X_1 、 X_2 、 X_3 、 X_4 、 X_5 、 X_6 的折线图

从图1看出, Y 随着 X_1 、 X_2 、 X_3 、 X_4 、 X_5 、 X_6 的提高而增加,近似于线性关系,为分析被解释变量与解释变量变动的数量规律性,本文设定模型为:

$$\ln Y = C + \beta_1 \ln X_1 + \beta_2 \ln X_2 + \beta_3 \ln X_3 + \beta_4 \ln X_4 + \beta_5 \ln X_5 + \beta_6 \ln X_6 + \mu$$

首选对数线性模型,是因为自变量的斜率度量了因变量对自变量的弹性,以方便于建立模型后经济意义的解释。

2.3 主成分分析前后的多元线性回归分析

(1)利用Eviews软件对经过处理后的数据进行多元线性回归,并进行多重共线性的检验,判断该模型是否具有多重共线性。基于2005年~2015年的数据,得到的回归方程如下所示。

$$\begin{aligned} \ln Y = & 0.543369130851 - 0.144199114303 \ln X_1 + 0.461291183138 \ln X_2 \\ & + 0.0906385687966 \ln X_3 + 0.112524832775 \ln X_4 \\ & + 1.25787773813 \ln X_5 - 0.395505252152 \ln X_6 \end{aligned}$$

作者简介:竺天辰(1997-),男,浙江宁波人,安徽财经大学在读,主要从事金融工程方面的研究。

表1 未消除通货膨胀因素后北京市2005年~2015年房价相关数据

年份	X1	X2	X3	X4	X5	X6	Y
2005 年	6969.52	2827.23	239.76	680.78	937	17653	6788.09
2006 年	8117.78	3296.38	60.31	842.49	1094	19977.5	8279.51
2007 年	9846.81	3907.2	293.21	1063.21	1132	21988.7	11553.26
2008 年	11115	3814.73	639	889.37	1181	24724.9	12418
2009 年	12153.03	4616.92	587.7	2367.77	1224	26738.5	13799
2010 年	14113.58	5402.95	1292.7	1439.08	1383	29072.9	17782
2011 年	16251.93	5578.93	1301.23	1167.95	1428	32903	16851.95
2012 年	17879.4	6112.4	1102.69	1484.74	1464	36468.8	17021.63
2013 年	19800.81	6847.06	1159.47	1836.95	1498	44563.93	18553
2014 年	21330.83	6924.23	1378.94	2158.03	1525	48531.85	18833
2015 年	23014.59	7495.99	2052.92	1970.97	1541	52859.17	22633

注: 原数据来源于国家统计局。

表2 消除通货膨胀因素后北京市2005年~2015年房价相关数据

年份	x1	x2	x3	x4	x5	x6	y
2005 年	6969.52	2827.23	239.76	680.78	937	17653	6788.09
2006 年	7997.813	3247.665	59.41872	830.0394	1094	19682.27	8157.153
2007 年	9256.957	3673.147	275.6458	999.5206	1132	20671.51	10861.19
2008 年	9867.024	3386.418	567.254	789.5128	1181	21948.82	11023.72
2009 年	10864.56	4127.431	525.3916	2116.737	1224	23903.67	12336.02
2010 年	12214.18	4675.823	1118.729	1245.409	1383	25160.28	15388.91
2011 年	13344.17	4580.759	1068.417	958.9827	1428	27016.06	13836.83
2012 年	14308.43	4891.6	882.4551	1188.2	1464	29185.06	13621.98
2013 年	15354.74	5309.622	899.1228	1424.481	1498	34557.55	14387.11
2014 年	16296.76	5290.113	1053.51	1648.735	1525	37078.34	14388.41
2015 年	17340.38	5647.867	1546.776	1485.031	1541	39826.84	17052.87

表3 相关系数表

名称	LnX ₁	LnX ₂	LnX ₃	LnX ₄	LnX ₅	LnX ₆
LnX ₁	1	0.982748	0.844122	0.685715	0.984981	0.974543
LnX ₂	0.982748	1	0.810283	0.741205	0.977359	0.952228
LnX ₃	0.844122	0.810283	1	0.539983	0.81009	0.771158
LnX ₄	0.685715	0.741205	0.539983	1	0.650726	0.67913
LnX ₅	0.984981	0.977359	0.81009	0.650726	1	0.936117
LnX ₆	0.974543	0.952228	0.771158	0.67913	0.936117	1

回归结果显示: 可决系数为0.9515, 说明模型的拟合程度较好, 即被解释变量商品房平均销售价格可以很好地由回归模型来解释。但X₁、X₆经济意义不通过。这表明模型可能存在严重的多重共线性。为了分析模型多重共线性, 根据数据得相关系数检验如表3所示。

由表3可以看出, 解释变量相关系数较高。但是较高的简单相关系数只是多重共线性存在的充分条件, 而不是必要条件。为了确定数据确实存在多重共线性, 对数据进行辅助回归, 并依据辅助回归数据研究变量的方差扩大因子VIF_j。VIF_j的大小反映了解释变量之间是否存在多重共线性。当VIF_j大于等于10时, 说明解释变量与其余解释变量之间有严重的多重共线性, 且这种多重共线性可能会过度地影响OLS估计。

各个解释变量的方差扩大因子值如表4所示。

表4 VIF_j表

VIF1	VIF2	VIF3	VIF4	VIF5	VIF6
434.9717	50.60217	9.069883	3.366131	136.0915	78.5546

根据表4所得数据可以发现该模型的确存在多重共线性。

(2)主成分分析后的多元线性回归分析。主成分分析能够有效消除模型中的多重共线性, 它的基本原理是通过投影的方法, 实现数据的降维^[9]。使用软件对数据进行主成分分析, 分析之后得到数据表5、表6所示。

表5 主成分提取结果

成分	方差贡献率 %	累计方差贡献率 %
主成分 1	85.89	85.89
主成分 2	8.35	94.23
主成分 3	4.4	98.63
主成分 4	1.11	99.74
主成分 5	0.24	99.97
主成分 6	0.03	100

其中第一主成分的方差贡献率为85.9%, 说明该主成分反映了原指标85.9%的信息量。已能较好反映总体变动情况。

表6 主成分特征向量

名称	第一主成分	第二主成分	第三主成分	第四主成分	第五主成分	第六主成分
LnX_1	0.436622	-0.13037	-0.16155	0.029401	0.271694	-0.83162
LnX_2	0.435268	0.005213	-0.17604	-0.2916	-0.83311	-0.02058
LnX_3	0.335566	0.894714	0.266	-0.0012	0.124519	0.024891
LnX_4	0.379567	-0.38946	0.828462	0.089806	-0.01304	0.098316
LnX_5	0.428194	-0.16232	-0.25476	-0.56845	0.46483	0.431513
LnX_6	0.424119	-0.06662	-0.34772	0.763484	0.019165	0.333917

根据表7中已给出对应的特征向量第一成分可表达为:

$$\text{Main} = 0.436622\text{LnX}_1 + 0.435268\text{LnX}_2 + 0.379567\text{LnX}_3 + 0.335566\text{LnX}_4 + 0.428194\text{LnX}_5 + 0.424119\text{LnX}_6$$

多元线性回归分析,运用EViews软件对商品房平均销售价格、第一主成分重新进行多元回归分析,得到的回归方程为:

$$\text{LnY} = 2.95843772803 + 0.324280224461\text{Main}$$

回归结果显示:可决系数为0.9005,拟合优度较高,T检验均通过,F统计量为81.04681,F统计量概率值接近于0,模型通过方差膨胀因子检验,说明进行主成分分析降维后,消除了原模型的多重共线性。

将原来得到的变量数据代入到该回归方程中:

$$\text{LnY} = 2.95843 + 0.1415\text{LnX}_1 + 0.1411\text{LnX}_2 + 0.1230\text{LnX}_3 + 0.1088\text{LnX}_4 + 0.1388\text{LnX}_5 + 0.1375\text{LnX}_6$$

对该式适当简化:

$$\text{LnY} = 2.96 + 0.14(\text{LnX}_1 + \text{LnX}_2 + \text{LnX}_5 + \text{LnX}_6) + 0.12\text{LnX}_3 + 0.11\text{LnX}_4$$

即

$$Y = e^{2.96} X_1^{0.14} X_2^{0.14} X_3^{0.14} X_4^{0.14} X_5^{0.12} X_6^{0.11}$$

3 模型检验

3.1 平稳性检验

对于时间序列数据,如果建立的模型中存在非平稳变量,极有可能出现伪回归现象。因此,为真正反映模型中变量之间的联系,避免伪回归问题的发生,要对商品房平均销售价格、第一主成分2个变量进行平稳性检验,检验如表7、表8所示。

表7 Main的平稳性检验

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-3.99022	0.018
Test critical values:	1% level	-4.4206	
	5% level	-3.25981	
	10% level	-2.77113	

表8 LnY的平稳性检验

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-2.34124	0.1786
Test critical values:	1% level	-4.29707	
	5% level	-3.2127	
	10% level	-2.74768	

从检验结果看,在1%、5%、10%三个显著性水平下,单位根检验的Mackinnon临界值分别为-4.29707、-3.2127、-2.74768,t检验值-2.34124大于相应临界值,表明LnY序列存在单位根,是非平稳序列。

3.2 协整检验

协整理论是指非平稳变量的线性组合可能是平稳的,基于此理论对回归模型各变量进行协整检验。检验回归残差的平稳性,由于残差序列的均值为0,所以选择无截距项、无趋势项的ADF检验,估计结果如表9所示。

表9 残差的平稳性检验

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-3.97767	0.0011
Test critical values:	1% level	-2.81674	
	5% level	-1.98234	
	10% level	-1.60114	

在5%的显著性水平下,t检验统计量为-3.97767,小于相应临界值,表明残差序列不存在单位根,是平稳序列,说明LnY和Main之间存在协整关系。

4 模型的分析和相关对策建议

4.1 模型分析

根据给出的模型:

$$Y = e^{2.96} X_1^{0.14} X_2^{0.14} X_3^{0.14} X_4^{0.14} X_5^{0.12} X_6^{0.11}$$

说明了影响2005年~2015年住房价格的最主要的因素是:地区生产总值、全社会固定资产投资、城市人口密度、城镇居民人均可支配收入。该模型主要因素对房地产市场的影响具体表现在:(1)地区总值的增加预示着良好的经济环境,经济的繁荣带来房地产市场的繁荣;(2)人口的增加会拉动对住房消费的刚性需求,会使均衡价格向上增加;(3)全社会固定资产投资、可支配收入的增加加大了住房市场的需求。上述分析从经济意义上较好解释了房地产市场的需求原理,而需求会影响住房价格。本模型较好表达了各个因素对房价的影响,因此本模型的建立是成功的。

4.2 相关对策建议

(1)这些年来,改革开放不断深入和发展,人均收入不断提高,经济的繁荣带来房地产市场的繁荣,因此,房价上涨是人均收入,地区生产总值上涨所导致的不可避免的结果,房价的调控必须首先尊重这一客观规律。(2)各地区经济发展的不平衡导致了人员流动的不均匀,过多的人涌入城市会快速的抬高房价。因此,调节地

快递行业最后一公里问题的现状与模式创新^①

河北农业大学 潘玉焕 马思玮 回增赫 黄子钦

摘要: 随着我国电子商务的爆发式增长,快递业随之进入快速发展期。资料显示,2016年中国快递行业实现了207亿件快递的业务服务规模,同比增长48%。然而作为联系商家与顾客最密切的环节——“最后一公里”,其配送费用却始终制约着整个物流成本的降低,一度占到整个配送成本的30%,因此“最后一公里”价值的挖掘潜力巨大。而本文提到的“互联网+同城快递”模式将社会资源优化整合利用,解决“最后一公里”的痛点。本文将围绕“最后一公里”问题,对现有第三方物流与“互联网+同城快递”模式进行对比分析,用“互联网+”的新观念寻求物流发展的新思路。

关键词: 第三方物流 “互联网+” “最后一公里”

中图分类号: F719

文献标识码: A

文章编号: 2096-0298(2017)08(a)-142-02

1 我国第三方物流现状

第三方物流是指通过签订合同或契约,工商企业将其全部或部分物流功能在一段时间内外协给专业的物流公司管理,相对于基本服务,契约物流服务提供复杂、多功能的物流服务,以长期互益的关系为特征。第三方物流之所以受到青睐,在于其独特的价值,主要体现在成本价值、服务价值、风险分散价值和竞争力提升价值。比较国内外第三方物流的发展情况,各自呈现出不同的特点。第三方物流在我国起步较晚,发展过程中难免存在许多不尽人意的地方。这些问题的存在不仅影响了第三方物流自身健康发展,也给相关行业的发展带来了不利影响。需求不足是物流外包乏力的根源。“大而全、小而全”的思想和对“第三利润源”的误解,是工商企业观念上的障碍;自有物流设备的处理和人员的分流也令许多企业为难;由于担心商业机密外泄和对第三方物流商缺乏信心,风险规避型的经理人宁可选择自营物流等。从第三方物流企业本身来看,传统的经营理念、信息技术落后等原因使其难以提供优质的服务。第三方物流市场潜力巨大,可待开发的价值无限如图1所示。

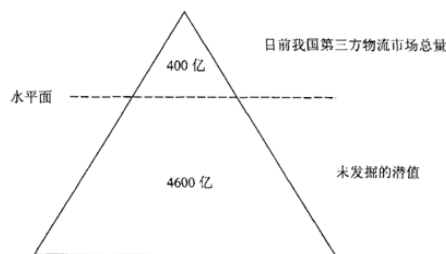


图1 中国第三方物流潜力

现在各大电商企业争相在各地建立仓储中心,物流集散中心,集中精力多拿一二线城市的土地,2013年开始,菜鸟网络先后在北京、天津、上海、广州、杭州、武汉、郑州、重庆、成都等15个城市布点仓储中心,仅在天津就拿地1500亩。相比菜鸟网络,京东的步伐明显更快,据京东在2014年5月19日向美国证券交易所更新递交的F-1招股书显示,京东已在全国34个城市拥有82间仓库,总建筑面积超过130万平方米。此外,京东宣称还将进一步在三线和一线城市扩大版图,加大建设步伐,开发物流市场,挖掘市场价值。

①基金项目:河北农业大学大学生创新创业训练计划资助项目(201710086001)。

2 “最后一公里”配送体系瓶颈

(1)每个快递公司在每个城市都有自己的仓库,各个快递公司

区、城乡间的经济发展差异,是解决人口相对均匀分布问题、解决房价攀高的一个重要手段。(3)投资因素是房价快速上升的又一问题。为了在一定程度上起到抑制房价上涨,让住房使用效率达到最大化,应当在合适的时机推出房地产税以抑制快速上涨的房价。

参考文献

- [1] Karl E. Case and Robert J. Shiller. The Efficiency of the Market for Single-Family Homes[J]. The American Economic Review, 1989, 79(1).
- [2] Clapp J. M. and Giaccotto C. The Influence of Economic Variables on Local House Price Dynamics[J]. Journal of Urban Economics, 1994(36).
- [3] Quigley J. M. Real Estate Prices and Economic Cycles[J]. International Real Estate Review, 1999(2).

- [4] Malpezzi S. A Simple Error Correction Model of Housing prices[J]. Journal of Housing Economics, 1999(8).
- [5] 袁志刚, 樊潇彦. 房地产市场理性泡沫分析[J]. 经济研究, 2003(3).
- [6] 周京奎. 房地产泡沫生成与演化——基于金融支持过度假说的一种解释[J]. 财贸经济, 2006(5).
- [7] 余壮雄, 林建浩. 谁推高了房价? 开发商、置业者还是地方政府[J]. 经济学家, 2010(5).
- [8] 杨东. 成都市住宅房地产价格变动研究及预测[D]. 四川师范大学, 2006(6).
- [9] 高铁梅. 计量经济分析方法与建模[M]. 北京: 清华大学出版社, 2009.