



《应用统计学》课程论文

经济基本面与宏观政策对房价的影响

——基于 31 个省市的面板数据分析

Impact of Economic Fundamentals and Macro-policies on Housing Prices

院 系	<u>工业工程系</u>
专 业	<u>物流工程</u>
姓 名	<u>刘晓彤、李科、张意舸</u>
学 号	<u>2016210545、2016213696、2016210542</u>
指导老师	<u>杨 朋</u>

2017 年 1 月 5 日

摘 要

近十几年来,我国的国民经济发展迅速,工业化水平和城镇化水平不断提高,其中以房地产市场的发展最为耀眼。现如今,房地产业已成为推动我国国民经济发展的重要因素,俨然已成为政府制定政策以及民众关注的焦点。

本文从房产的消费品与投资品两重属性入手,将影响房地产价格的因素分为需求方因素、供给方因素、宏观经济与政策三个方面。为考察市场本身有哪些因素影响房价,选取 2005-2014 年我国 31 个省市自治区的面板数据,综合运用 R 语言编程、SPSS 和 Eviews 等统计软件对影响房价的因素进行定量分析。结果表明从全国范围内来说,供给和需求是影响房价的重要因素。我们又根据聚类分析结果将全国按照经济发展水平划分为不同区域,分区域建立了多元回归模型。发现影响不同经济发展水平区域房价的因素是不同的。并且本文采用实证研究的方法着重考察了货币政策和限购政策对于房价的影响。发现货币政策对于不同区域的影响程度是有差异的,限购政策对于房价的调控作用是比较有效的。最后本文根据统计分析的结果提出了相关的政策建议。

关键词: 多元回归分析; K-means 聚类 ; 向量自回归模型;

ABSTRACT

As the reform of housing system carried out in 1998, the government of China has canceled the welfare housing and implemented monetization of housing distribution, which promoted the marketization of real estate. Afterwards, real-estate prices rose continually and the real estate industry went through an unprecedented development, which gradually become the pillar industry of national economy.

This paper starts with the dual attributes of consumer goods and investment goods, and explains three factors that influence the price of house, consumer demand, house supply and macroeconomic policy. In order to find out factors affect the price for the market itself, 2005-2014 panel data of China's 31 provinces and autonomous regions are selected. R programming language, SPSS and Eviews statistical software are used for the quantitative analysis of the impact of price. The results show that house supply and consumer demand are important factors affecting the price of the house. In the base of the results of cluster analysis, we divide the whole country into different regions according to the level of economic development. It is found that the factors that affect the regional economic development level are different. In this paper, we use empirical research methods to study the impact of monetary policy and the restriction policy on house prices. And it is found that the impact of monetary policy on different regions is different, and the restriction policy is more effective for the regulation of house prices. Finally, according to the results of statistical analysis, this paper puts forward some policy recommendations.

Key words: Multiple Regression Analysis; K-means Clustering Method;

Vector Autoregression;

目 录

摘 要.....	2
ABSTRACT	3
第一章 绪论.....	6
1.1 选题背景.....	6
1.2 文献综述.....	6
1.3 研究思路与框架结构.....	8
第二章 房价影响因素的定性探索.....	9
2.1 商品住宅销售价格波动.....	9
2.2 人口因素.....	10
2.3 宏观经济因素.....	12
2.3.1 经济增长、居民经济收入与支出.....	12
2.3.2 经济景气指数.....	15
2.4 房地产投资规模.....	16
2.5 宏观经济因素.....	20
2.6 区域差异.....	21
2.6.1 数据预处理.....	21
2.6.2 确定分类数.....	22
2.6.3 K-means 聚类及结果	24
2.6.4 房价区域差异分析.....	26
2.7 本章小结.....	27
第 3 章 房价回归模型.....	27
3.1 变量选取与数据处理.....	27

3.1.1 选择变量建立数据集.....	28
3.2 建立全国房价回归模型.....	31
3.2.1 模型选取.....	31
3.2.2 模型建立.....	34
3.3 建立区域回归模型.....	37
3.4 本章小结.....	42
第四章 宏观政策对房价影响分析.....	42
4.1 货币政策对房价影响分析.....	42
4.1.1 全国范围货币政策影响.....	43
4.1.2 发达地区货币政策影响.....	44
4.1.3 较发达地区货币政策影响.....	46
4.1.4 中等发达地区货币政策影响.....	48
4.2 限购政策对房价的影响分析.....	50
4.3 本章小结.....	56
第五章 结论与政策建议.....	56
5.1 结论.....	56
5.2 政策建议.....	57
参考文献.....	58
附录 1 附图和附表.....	59
附录 2 代码.....	64

第一章 绪论

1.1 选题背景

住房是事关民生的大问题。自 1998 年国家颁布《个人住房贷款管理办法》以来,房地产业的发展如火如荼,房地产价格一路攀升。1998 年商品住宅房价接近 200 元,2016 年初是元。房价飙升和越来越热的“炒房”日益成为社会问题,已经引起全社会的关注和忧虑。伴随房价上涨而来的是国民生活成本大幅增加,导致中低层收入群买不起房,高房价不仅是经济问题也同时成为社会问题。这也引发了我们的思考。到底是什么因素在支撑着这么高的房价?我们小组的成员带着这样的疑问,开始了对房价的影响因素的探索之旅。

本文希望探索影响房价的因素,并建立回归模型。2005 年是国家开始大手笔调控房价的年份,因此本文的时间跨度是从 2005 年年初开始到 2014 年底,部分数据是年至年的年度数据。房产包括商品房、商品住宅、营业用房、办公楼,本文限定只考虑商品住宅。

1.2 文献综述

本文主要从房价影响因素,货币政策对房价的影响,以及调控政策对房价的影响入手进行文献综述

1. 房价影响因素

赵丽丽,焦继文(2007)从供给和需求两方面科学的选取房价影响因素指标体系,引入灰色关联分析方法,对房价影响因素相对重要性进行分析^[3]

梁云芳(2007)关注到了房价与地区有关,他认为市场短期需求量变化对西部地区的房价有负作用;中部地区房价的长期趋势和短期波动与该地区的经济发展水平高度相关^[4]。

杨华(2008)认为商品房供需差、人均、人口密度、地价、居民消费物价、原材料价格、地产商预期利润、利率是影响房价的主要因素。人均 GDP 是影响房价的最重要的因素^[5]。

曹瑞(2010)走的更远,它认为供需差、人均、人口密度、地价是影响房价最重要的因素,并利用多项式回归分别讨论各个因素与房价的关系^[6]。

2. 货币政策对房价的影响

对于货币政策对房价的影响,已有大量文献对此进行实证研究。刘传哲、何凌云(2006)选取货币供应量 M2、利率、汇率、国内生产总值、消费者物价指数作为货币政策变量,通过定性分析和 Granger 因果检验,发现 1998 年后这些变量与房价的相关性趋于上升:货币供应量、国内生产总值、消费者物价指数与房价正相关,利率、汇率与房价负相关^[7]。

蒋益民、陈璋(2009)把中国分为八大经济区域来研究货币政策是否具有区域效应,他们运用结构向量自回归模型(SVAR)和脉冲响应函数,对其进行实证分析,最终得出结论,我国存在明显的货币政策区域差异效应^[8]。

高云峰、阮莉莉(2012)以货币政策传导机制作为理论基础,构建了面板数据模型,定量分析货币政策区域效应。从贷款投放到对经济增长产生刺激作用来看的话,东部地区、东北地区和中部地区的效果相较更明显,但是西部地区效果却比较差^[9]。

3. 调控政策对房价的影响

赵鹤芹(2013)以现在广受热议的房地产“限购令”政策为研究对象,重点观察二、三线城市限购措施实施情况,并针对南昌市房价在限购前后进行定性的图表分析与定量的两配对样本检验 SPSS13.0 实证分析,得到的结论是限购政策在南昌房价调控上起了一定作用,但效果并不显著^[10]。

张德荣(2013)本文按照户籍和限购地域两个标准对限购政策进行了细分,在倍差法的研究框架内,通过动态面板 GMM 模型估计来评估限购政策对抑制住宅价格上涨的效果。结果表明,对非户籍购房人实施限购的政策效果最为明显。限购范围影响政策效果,仅在市区范围内限购难以起到抑制房价上涨的作用^[11]。

综上思路,为了更好的分析房价影响因素,我们从经济基本面和宏观政策两个角度进行分析。前人大多把研究区域划分为东部地区、中部地区和西部地区,而忽略了而同一地区、相邻省市房价差别很大,因此,鉴于传统划分的不足,本文采用聚类分析方法,将全国 31 个省(市、自治区)划分为发达地区、较发达地区、中等发达地区和欠发达地区,建立面板数据模型,实证研究房价的影响因素。

1.3 研究思路与框架结构

综合前人的研究，本文考虑的经济因素包含人均、可支配收入、居民消费价格指数、工业化水平、房地产投资规模，人口因素包含城镇化水平、城市人口密度、人口总量，政策因素主要是利率变动、房产政策。本文希望从这些变量中找出影响房价的因素并建立回归方程。

本文的结构如下：

第一章绪论，简述房地产业的发展 and 房价影响因子的研究现状，以及本文的目标、方法和一些必要的假定。

第二章房价影响因素的定性探索，分别探讨人口因素、宏观经济因素、房地产投资规模、国家政策和利率、区域差异的变动情况以及它们与房价的关系。

第三章房价回归模型，通过第二章的分析，选取对房价有影响的变量建立回归方程。回归诊断是回归分析至关重要的步骤，分别进行回归方程检验、回归系数检验、残差分析以确定回归方程。

第四章宏观政策对房价的影响分析，针对货币政策和调控政策对房价的影响进行分析

第五章结论，并根据实证分析结果提出相应的政策建议

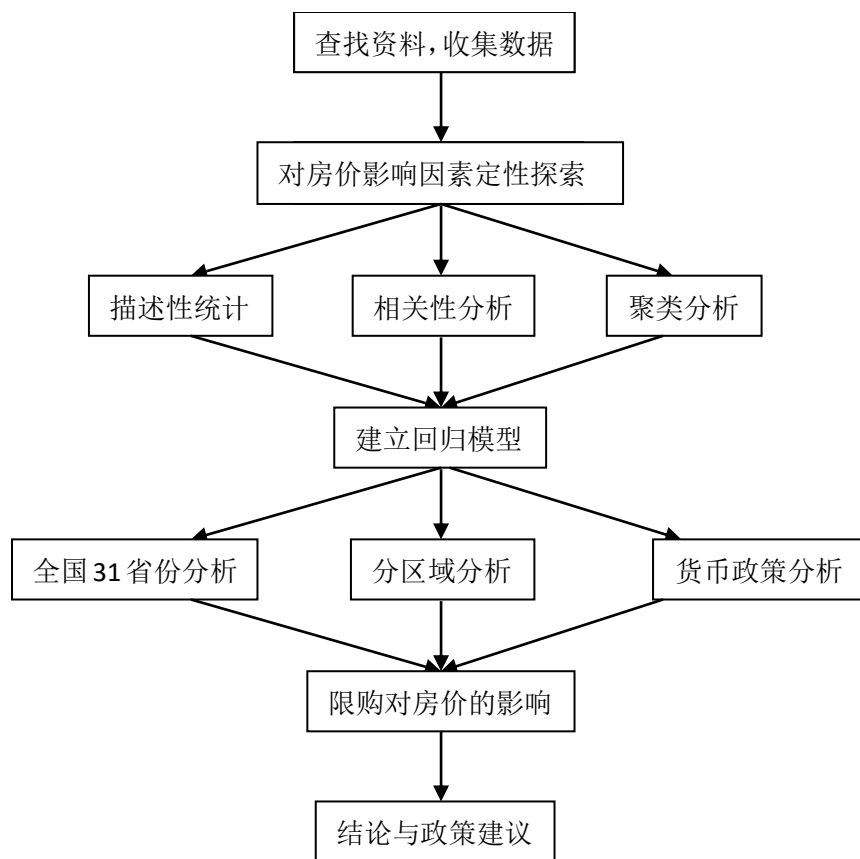


图 1-1 论文结构路线图

第二章 房价影响因素的定性探索

2.1 商品住宅销售价格波动

我们从 Wind 数据库中获取了 2004 年 2 月-2016 年 11 月的全国商品住宅销售价格数据。首先绘制房价随时间波动的趋势图。

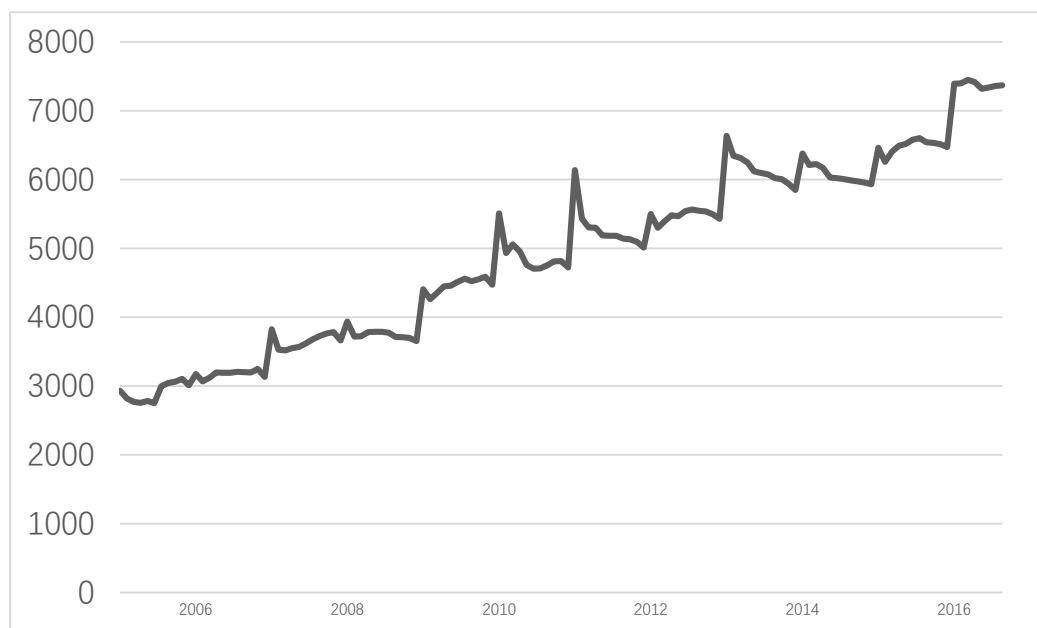


图 2-1 2004-2016 全国房价走势图

可见全国房价变动整体上升，局部波动。在 2010 年和 2011 年都出现过房价大涨的情况。在数据处理中，许多模型都建立在正态性假设的基础之上，因此随机变量的正态性假设往往至关重要。用卡方检验看看全国房价在这一时期内是否符合正态分布：

```
> N_price=read.table("clipboard",header = T)
> chisq.test(N_price)

chi-squared test for given probabilities

data:  N_price
X-squared = 45269, df = 127, p-value < 2.2e-16
```

从结果上看值小于 0.05，拒绝正态性假设，亦即全国的房价不服从正态分布。

2.2 人口因素

衣食住行是人的基本要素，住房是人的必需品，人口总量和城乡结构对房价必然有很大的影响。城市人口的自然出生率将影响未来几十年对房的需求，外来人口在城市的迁入迁出则对房屋当下的需求（主要是消费型需求）有影响。人口总量是影响房屋消费的双刃剑，一方面人口增加加剧了资源稀缺性，使资源人均占有率下降，减缓区域经济增长，人均收入下降，房屋购买力下降。另一方面，人口增长伴随着城镇化进程，城镇人口增加大大激发了房屋需求，从而拉动房地

产及房价的上升。人口的年龄结构对房地产业有一定影响，20-30 岁年龄层的青年面临结婚成家，他们购房主要是结婚所需，中年层的消费者生活基本稳定，没有婚姻的压力，他们购房或者是寻求更加舒适的居住条件或者是将购房作为一种投资方式。数据集数据来源：国家统计局）保存了全国各地区 2005 至 2014 年度人口总量、城乡人口数量、城镇化水平的数据。

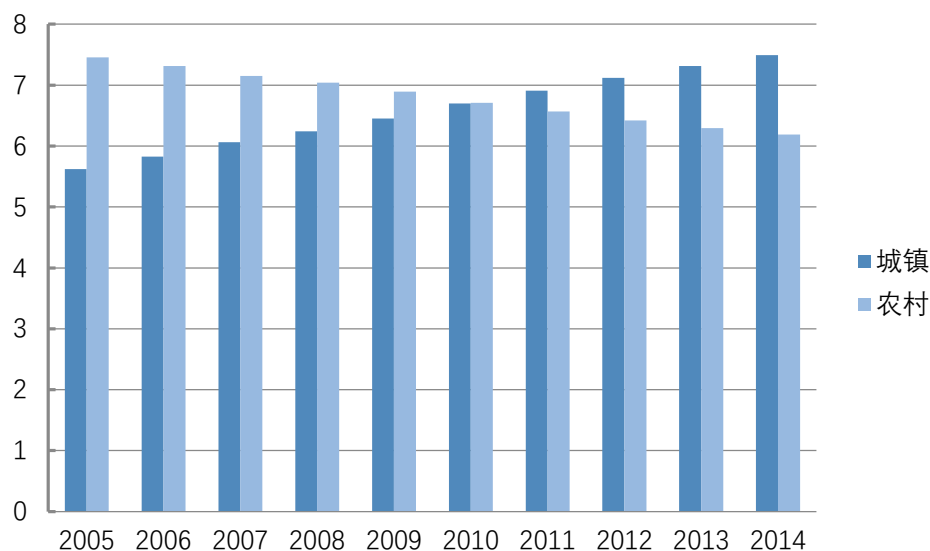


图 2-2 2005-2014 全国城乡人口分布

从图 2-2 可以看出，2005 至 2014 年间，全国人口从 130756 万人增长到 136782 万人，年平均增长率为 0.50%，城镇人口从 56212 万人增加到 74916 万人，年平均增长率为 3.3%，是全国人口平均增长率的 7.22 倍；农村人口逐年下降，呈现负增长，从 2011 年以后全国城镇人口超过农村人口；全国城镇化水平逐年递增，从 2005 年的 0.43 增加到 2014 年的 0.55。

数据集中的数据是年度数据，取 price 中每年 12 月的数据作为该年年度数据。分别比较各地区房价对人口总量、房价对城镇化水平。分析可知，房价与人口总量关系不明显，房价与城镇化水平关系密切，相关系数达到 0.47 左右。由于各地区的城镇化水平是年度数据，每个地区只有 10 个观测值，以此建立的回归模型误差非常大。尽管如此，它们仍反应出房价与城镇化水平的高度相关性。

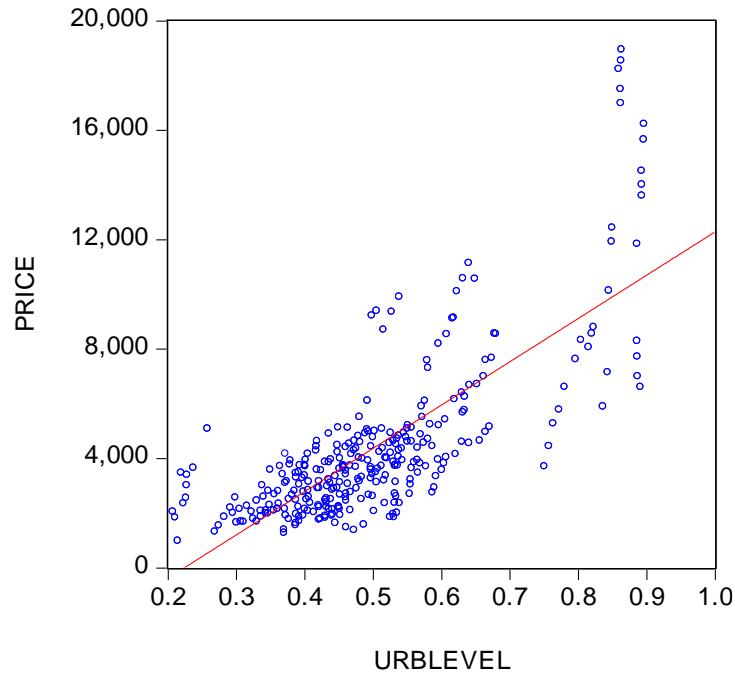


图 2-3 城镇化水平散点图

2.3 宏观经济因素

房地产是具有投资与消费双重属性的特殊商品。在国内宏观经济的大环境中, 房地产的发展必然受到宏观经济因素的制约或促进。经济发展水平、工业化水平、居民收入水平、消费价格指数等等都是宏观经济状态的指标。理论上认为, 经济发展状况良好时, 投资和消费高涨, 必然存在通货膨胀, 物价上涨, 房价自然不例外。供求关系是相互制约相互影响的, 房地产投资增加必然导致房地产供给增加, 给房地产业带来一个阻力, 使之发展放缓。但是房价与投资、物价、收入水平的关系仍然复杂。

本文从以下几个方面探索宏观经济发展状况: 国内生产总值(GDP)、人均国内生产总值、工业化水平; 居民消费价格指数(CPI)、城镇居民可支配收入; 消费者信心指数(CCI)、经济预警指数、国房指数。

2.3.1 经济增长、居民经济收入与支出

数据集数据(来源: 国家统计局)收集了从 2005 年到 2014 年国内生产总值、第二第三产业贡献、工业化水平、城镇居民人均年收入、城镇居民可支配收入指数、居民消费价格指数的年度数据。2005 到 2014 年间, 国内生产总值、工业化

水平、人均 GDP、可支配年收入逐年递增。GDP 从 184937.4 亿元增加到 643974 亿元，增加了 2.4 倍。工业化水平是第二第三产业总产值除以国内总产值得到的数值，它是国家工业化程度的度量，它从 2005 年的 0.88 增加到 0.90。人均 GDP 从 1.4368 万元增加到 4.7203 万元，增长了 2.29 倍；城镇居民可支配年收入从 10493 元增加到 2012 年的 24564.7 元，增长了 1.34 倍。

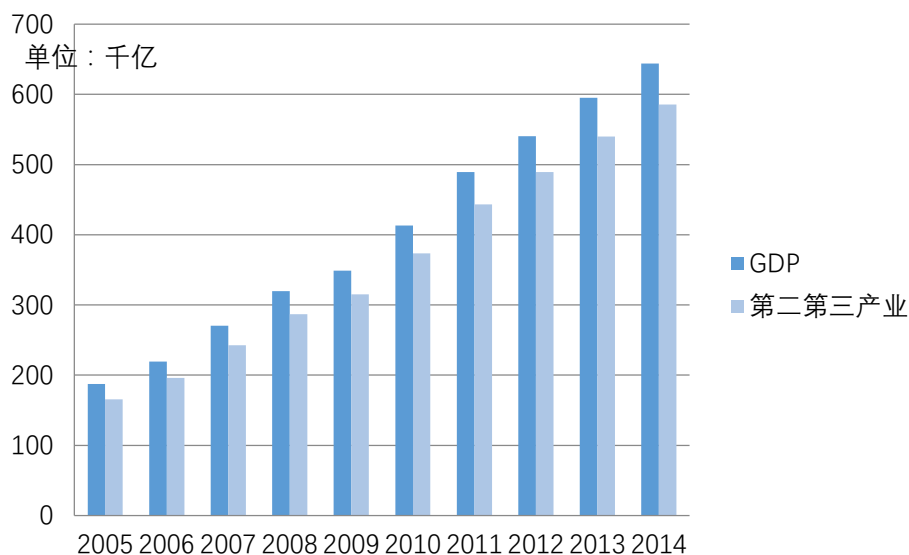


图 2-4 2005-2014 年全国生产总值

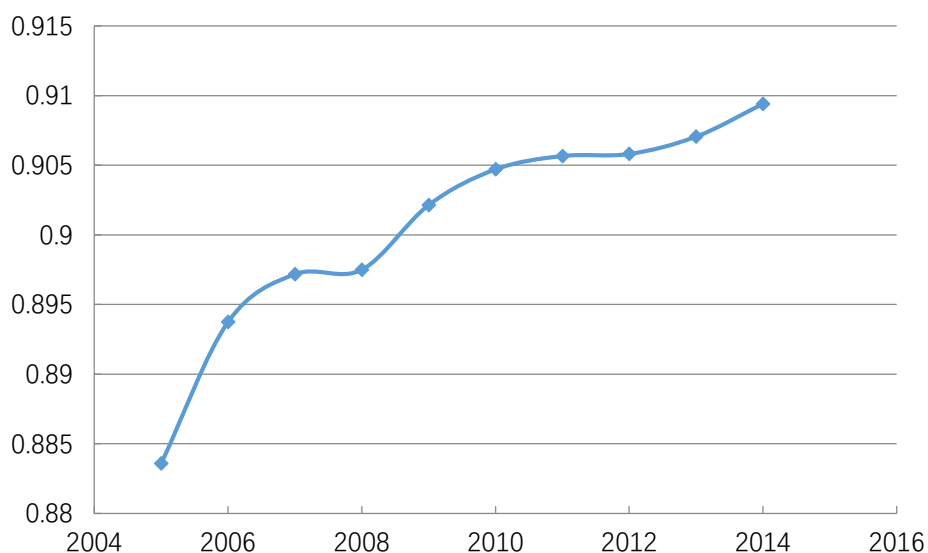


图 2-5 2004-2016 年全国工业化水平

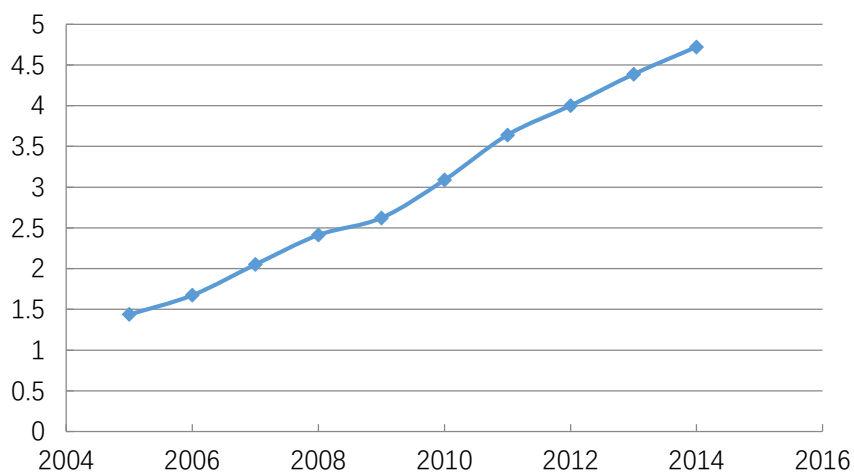


图 2-6 2004-2016 年全国人均 GDP（万元）

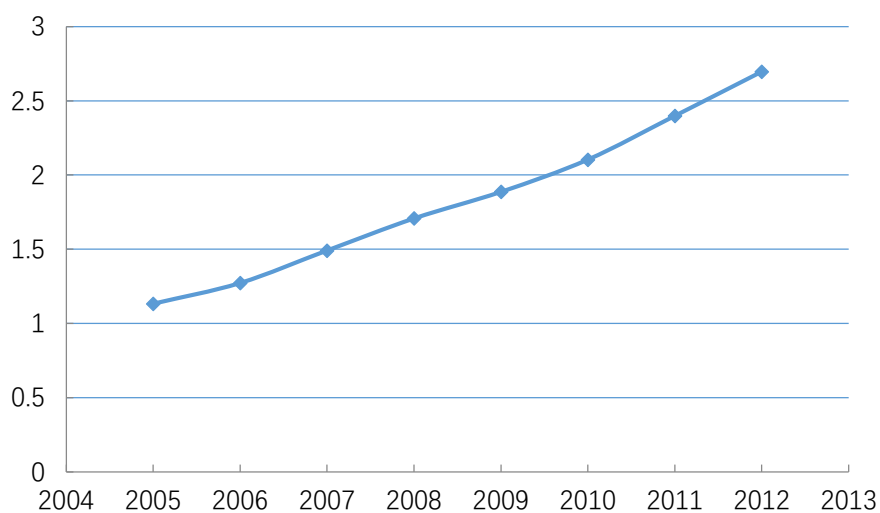


图 2-7 2004-2014 年 城镇居民可支配收入（万元）

观察工业化水平、人均 GDP、可支配人均年收入、CPI 分别对房价的散点图，它们中与房价可能有相关性的是工业化水平、人均 GDP、可支配人均年收入。表 2-1 是它们的相关系数矩阵。

表 2-1 相关系数矩阵

	pri	indus	gdpp	income	CPI	Urblevel
pri	1.0000000	0.9162670	0.9832844	0.9835544	0.1358254	0.9912979
indus	0.9162670	1.0000000	0.8953282	0.8932981	0.2128419	0.9321588
gdpp	0.9832844	0.8953282	1.0000000	0.9984568	0.2632620	0.9943677
income	0.9835544	0.8932981	0.9984568	1.0000000	0.2323480	0.9937818
CPI	0.1358254	0.2128419	0.2632620	0.2323480	1.0000000	0.2283410
Urblevel	0.9912979	0.9321588	0.9943677	0.9937818	0.2283410	1.0000000

可知除了 CPI，房价与其他变量都有明显的正相关性，尤其是人均 GDP、可支配人均年收入和城镇化水平；城镇化水平、工业化水平、人均 GDP 和可支配人均年收入都有很大的相关系数，它们的值越大房价越高。用城镇化水平、工业化水平和人均 GDP 作为地区分类的指标。

2.3.2 经济景气指数

宏观经济景气指数内容有预警指数、一致指数、先行指数和滞后指数。预警指数被称为宏观经济形势的晴雨表，本文用它作为经济景气指数的度量，它在一定程度上反应当前经济形势是冷还是热。下表 2-2 是预警指数对应预警信号。

表 2-2 预警指数与预警信号

预警指数	<60	60-80	80-120	120-140	>140
预警信号	过冷	趋冷	稳定	趋热	过热

消费者信心指数（是消费者结合自身收入状况、收入预期以及对宏观经济的主观判断的一种量化指标，国房指数是房地产业发展状况的度量。（数据集数据来源：新浪财经）保存了预警指数、消费者信心指数和国房指数在 2005 年至 2014 年的月度数据。

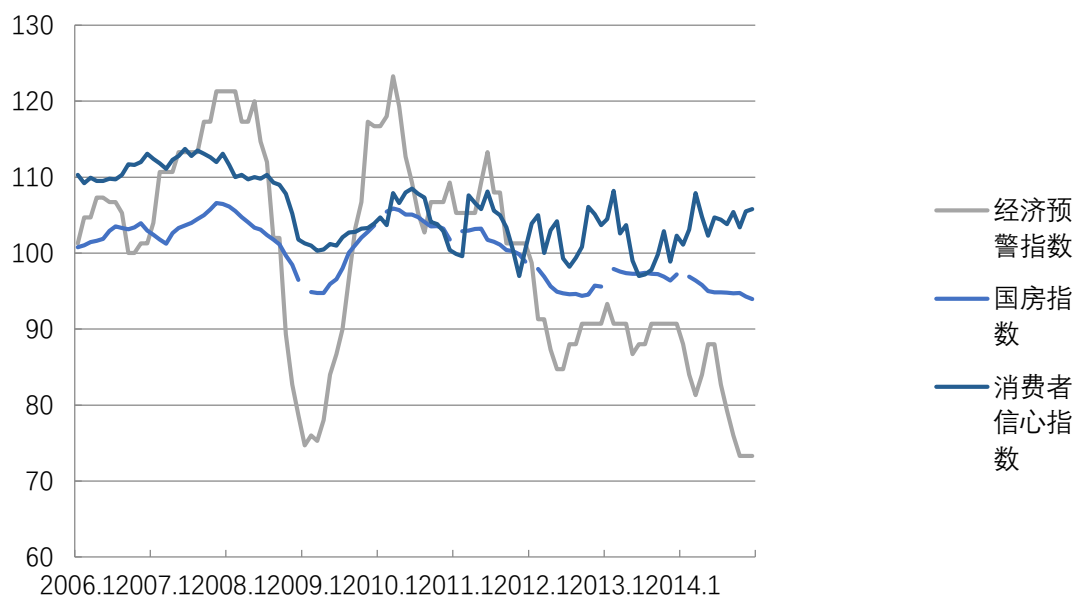


图 2-8 2005-2014 宏观经济指数走势图

从图 2-8 可以看出三个指标都在 2008 年初和 2010 年初达到局部最大值，预警指数超过了 120，经济形势趋热；预警指数和国房指数在 2009 年初达到最小值 74.7，经济形势趋冷；在 2012 年达到局部最小值，这时经济放缓。图 2-8 中经济预警指数和国房指数有相同的走势，消费者信心指数与它们走势相反。它们的相关矩阵为

表 2-3 宏观经济指数相关系数矩阵

	EWI	HI	CCI
EWI	1.000000	0.9010725	0.5808422
HI	0.9010725	1.000000	0.6460435
CCI	0.5808422	0.6460435	1.000000

从表 2-3 可以看出国房指数和经济预警指数相关系数为 0.90，消费者信心指数与国房指数、经济预警指数相关程度不明显。同时结合走势图，当经济形势趋热时，房地产业趋热，消费者信心指数降低。反之成立。

2.4 房地产投资规模

从以下几个方面认识房地产业的发展规模：商品房投资额及其增速、商品房施工和竣工面积、土地购置费及其增速。数据集 `inve_s`（数据来源：中宏）保存了全国各地区商品房投资额月度数据，时间跨度是 2005 年 2 月到 2016 年 9 月，每年一月的数据缺失，单位是亿元。

纵向地，从 2005 年 2 月到 2016 年 9 月，全国商品房投资额从 802.87 亿元跃升到 49930.85 亿元，增长了 62 倍。变化曲线表现为波动中稳步上升，具有明显的季节变化。年初处于低位，此后按月增加，到年底到达最大值，次年一月投资额下降到较低水平，此后进行下一轮循环。全国各地区商品房投资额变化走势与全国的商品房投资额大体相当，在波动中稳步上升。

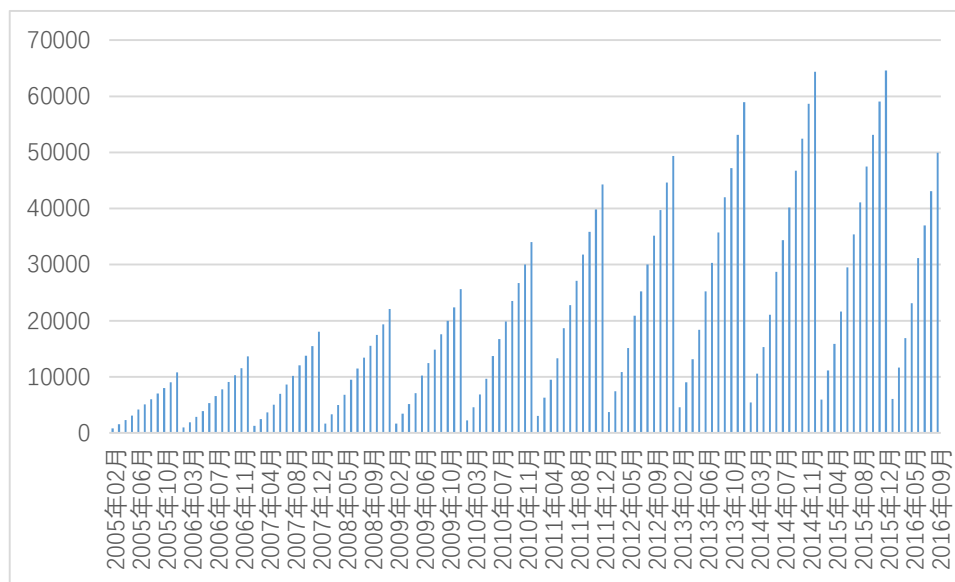


图 2-9 全国商品住宅开发投资额

横向地，全国 31 个省份商品房投资呈现明显的从东部沿海到西部内陆递减的趋势。从 2005 年 2 月到 2016 年 9 月各地区商品房投资额的均值最大的五个省份是江苏、广东、浙江、山东、辽宁，它们的均值分别为 2675.14、2358.73、2106.88、1829.45、1706.29，它们的总和占全国商品房总投资的 37.44%；均值最小的五个省份分别是新疆、甘肃、青海、宁夏、西藏，它们的均值分别为 225.40、195.22、149.74、76.62、8.16，总和之占全国商品房总投资的 2.30%。

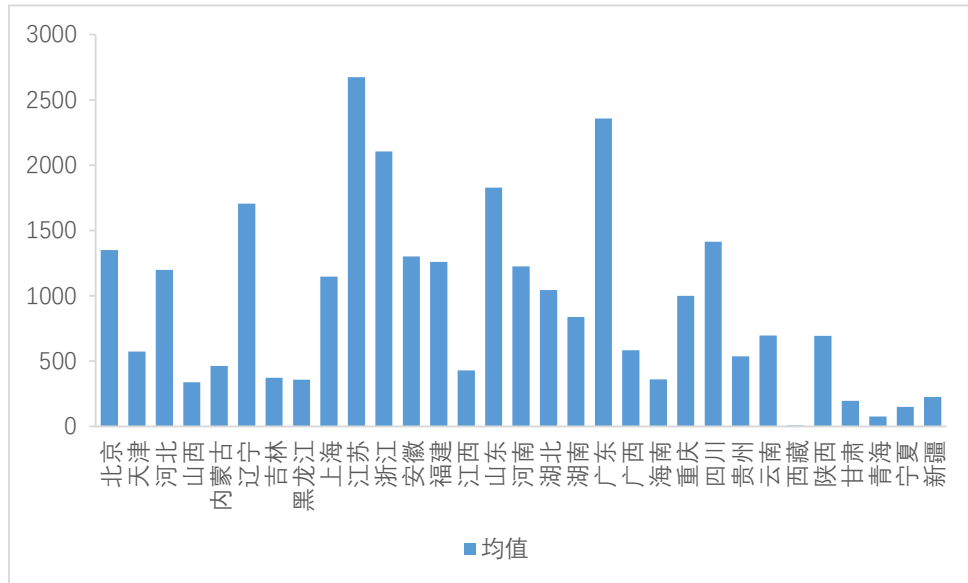


图 2-10 各地区商品住宅投资额均值

数据集 invc_s（数据来源：中宏）是全国平均商品房投资额增速月度数据，它的时序图(图 2-10)形状类似于经济预警指数，随时间波动较大，2008 年、2010 年和 2011 年都保持了 30%左右的增幅，2009 年、2015 年增速很小，甚至接近于 0。



图 2-11 全国商品住宅投资额增速时序图

观察各地区商品房投资额增速的时序图，它们有一个明显的特点是，全国大陆 31 个省市，除宁夏、青海、西藏、陕西、天津这几个省份（直辖市）外，其他所有地区均在 2008 至 2009 年间增长率下降，有半数以上出现负增长；到

了 2009 年以后，它们的商品房投资额增长率都在逐步回升，有些地区（如上海、辽宁）甚至超过了年以前的水平；当时间的年轮转过 2012 年，这些地区的投资增速又出现了一个低潮。这与经济预警指数相一致，预警指数显示，2008 和 2010 年经济形势趋热，2009 经济形势趋冷。

数据集 houscons（数据来源：中宏）是全国各地区商品住宅施工面积月度数据，数据集 houscomp（数据来源：中宏）是全国各地区商品住宅竣工面积月度数据。landex（数据来源：中宏）是土地购置费月度数据，landexad（数据来源：中宏）是土地购置费增速月度数据。土地购置费随时间变化也与商品房投资额类似，整体上升，并伴随季节变化。计算以上 6 个房地产投资规模指标对房价的相关系数和散点图。

表 2-4 全国房价与投资规模的相关系数矩阵

	price	inve_s	invc_s	houscons	houscomp	landex	landexad
price	1.00	0.04	-0.10	0.47	0.03	0.03	-0.17
inve_s	0.04	1.00	0.00	-0.05	-0.15	-0.40	0.06
invc_s	-0.10	0.00	1.00	-0.25	-0.21	0.06	-0.08
houscons	0.47	-0.05	-0.25	1.00	0.34	0.05	-0.06
houscomp	0.03	-0.15	-0.21	0.34	1.00	0.09	0.13
landex	0.03	-0.40	0.06	0.05	0.09	1.00	0.01
landexad	-0.17	0.06	-0.08	-0.06	0.13	0.01	1.00

表 2-4 显示，房价与施工面积的相关系数较高，为 0.47，而房价与其他变量的相关系数很小。即房价只与商品住宅施工面积有较好的依赖关系，与其他变量无关。由图 2-11 也可以看出，施工面积与房价存在一定相关关系。因此考虑在房价回归方程中加入商品房施工面积。

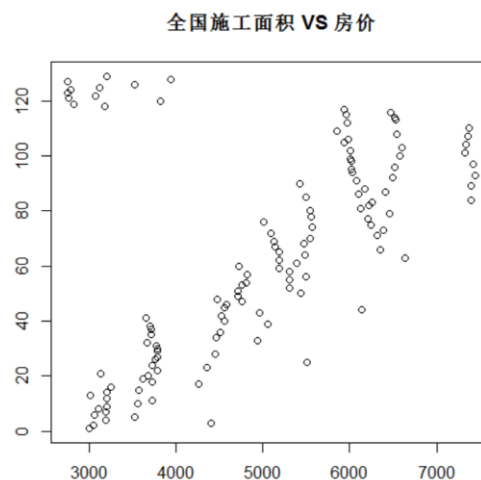


图 2-11 施工面积对房价散点图(全国)

2.5 宏观经济因素

从年以来，政府一直不遗余力地控制房价过高，各项住房政策层出不穷。房地产开发投资层面，政府在土地开发、银行贷款方面施加强有力的控制，消费层面，政府规定第一套住房的首付比例和第二第三套住房的首付比例，以及银行贷款利率。银行利率的调整反应了政府对房价的调控，从国家房产政策可以看出，对房价的宏观调控，利率是必用的手段。

从半年期贷款利率图中看出，2007 年是国家密集提高利率的年份，半年期贷款利率共五次加息，依次提高 0.18,0.18,0.18,0.27,0.09 个百分点。，

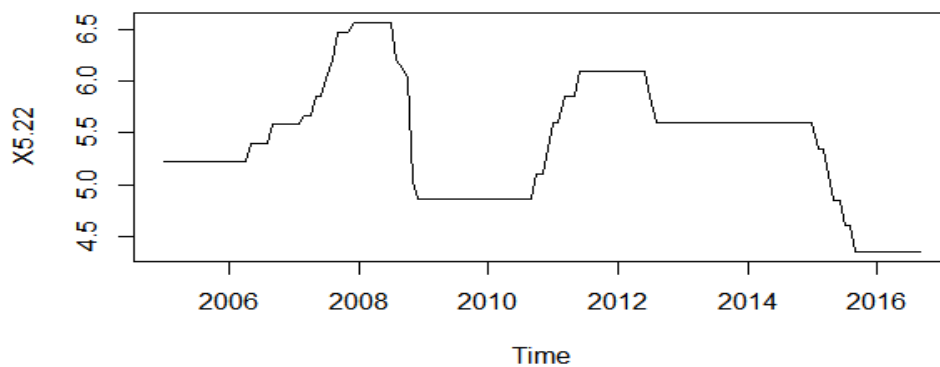


图 2-12 2004-2016 年半年期贷款基准利率走势图

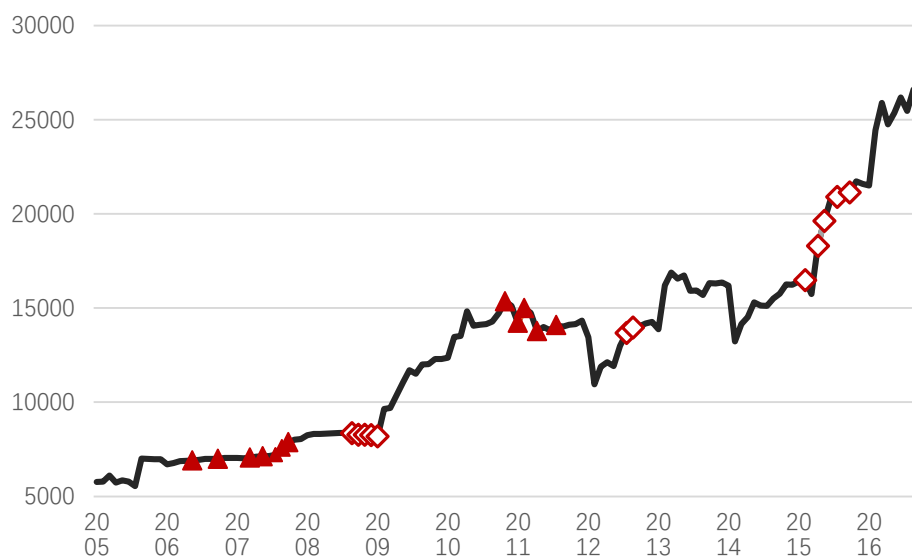


图 2-13 2005-2016 上海房价走势图

为探索利率对房价的影响，作一阶差分，并且差分为正记为 1，差分为负记为 -1，差分为 0 记为 0。则某时刻利率值为 1 表示利率升高，-1 表示利率下降，0 表示利率不变。利率对房价的影响在于，高利率下房产投资开发成本加大，利润空间变小，房产投机者的融资成本增加，持币待购者支付能力下降，多重因素迫使房价下跌。利率是影响房产泡沫的重要因素，当利率上升，预期房产实际价格下降，当降低到一定程度房产泡沫破灭。低利率下的情形相反。利率对房价的作用存在一个滞后效应，年利率处于低位，未来两年内的房价高速增长；年利率处于高位，未来两三年房价仍然很高但是房价增长速度明显下降。因此认为，利率提高对未来两年内的房价增速有反作用力，利率下降则使房价增速变大。

2.6 区域差异

由 3.4 节的结果，可知房价具有明显的地区差异，呈现从东南部沿海地区向中西部递减的趋势。用人均 GDP、城镇化水平、工业化水平这三个变量来衡量一个地区人口结构和经济发展状况，并通过它们用 k-means 聚类，将全国大陆 31 个地区分成若干类，以探讨不同类别的房价差异。

本文采用的工业化水平，是指第二第三产业总产值除以国内总产值得到的比值，它是国家工业化程度的度量。

2.6.1 数据预处理

数据集 region 是从到年按年平均的人均 GDP、城镇化水平和工业化水平的数据，它是一个 31 行 3 列的矩阵，每一行是一个地区的人均 GDP、城镇化水平和工业化水平。后两个变量的取值都在 0-1 之间，因此将数据中的人均 GDP 作极差正规化变换，使之取值在 0-1 之间。极差正变换的公式[1]如下：

$$x_{ij}^* = \frac{x_{ij} - \min_{1 \leq t \leq n} x_{tj}}{R_j} \quad \begin{pmatrix} i = 1, 2, \dots, n \\ j = 1, 2, \dots, m \end{pmatrix}$$

其中， R_j 为第 j 列数据的极差。

2.6.2 确定分类数

k-means 聚类必须事先知道类数，因此我们先采用系统聚类的方法画出谱系图，确定分成几类比较合适。根据图 3.6.1 所示的 Ward 法的聚类结果，以及图 3.6.12 所示的类平均法的聚类结果，将各地区分为四类或者五类比较合适。

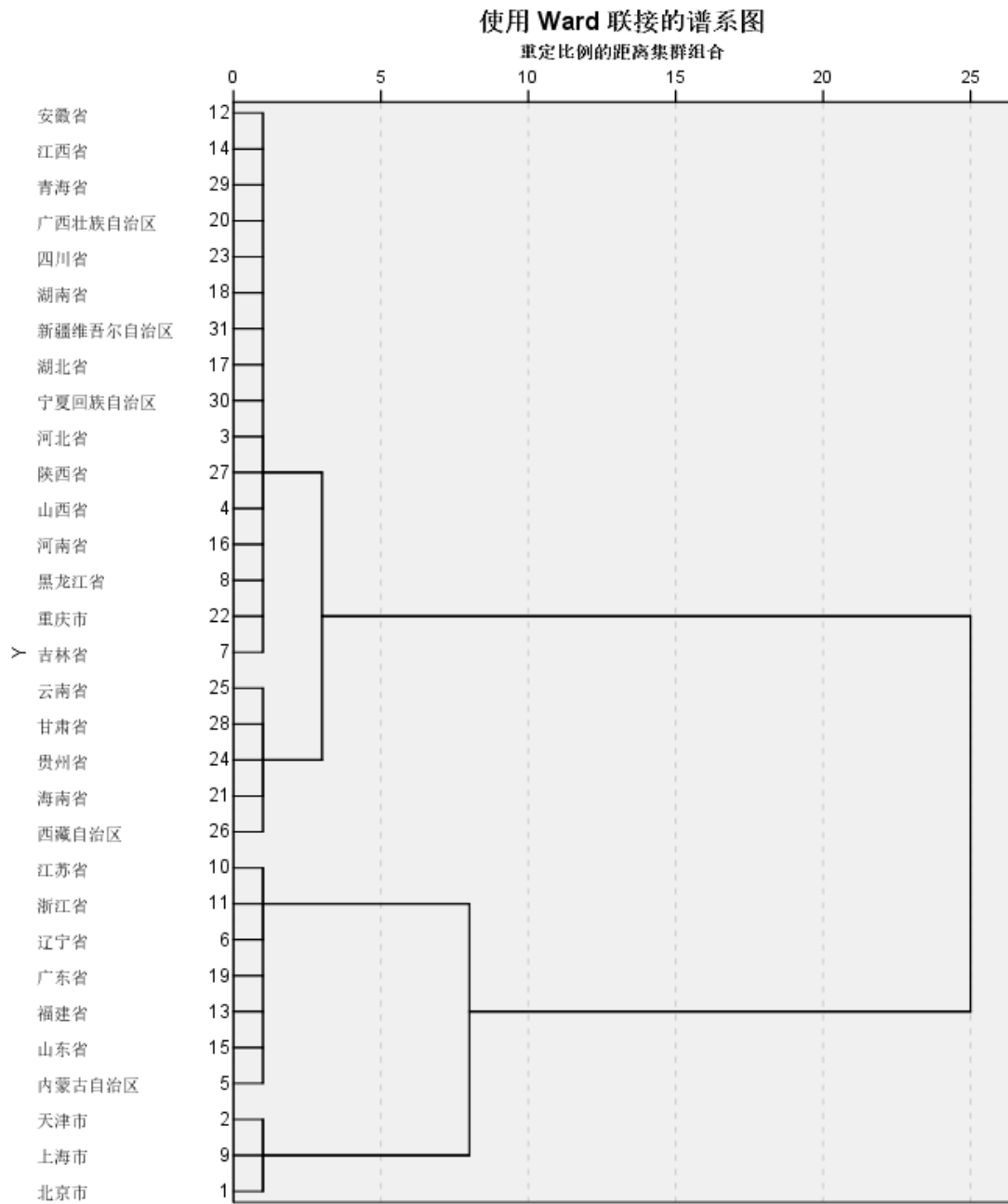


图 2-14 Ward 法聚类谱系图

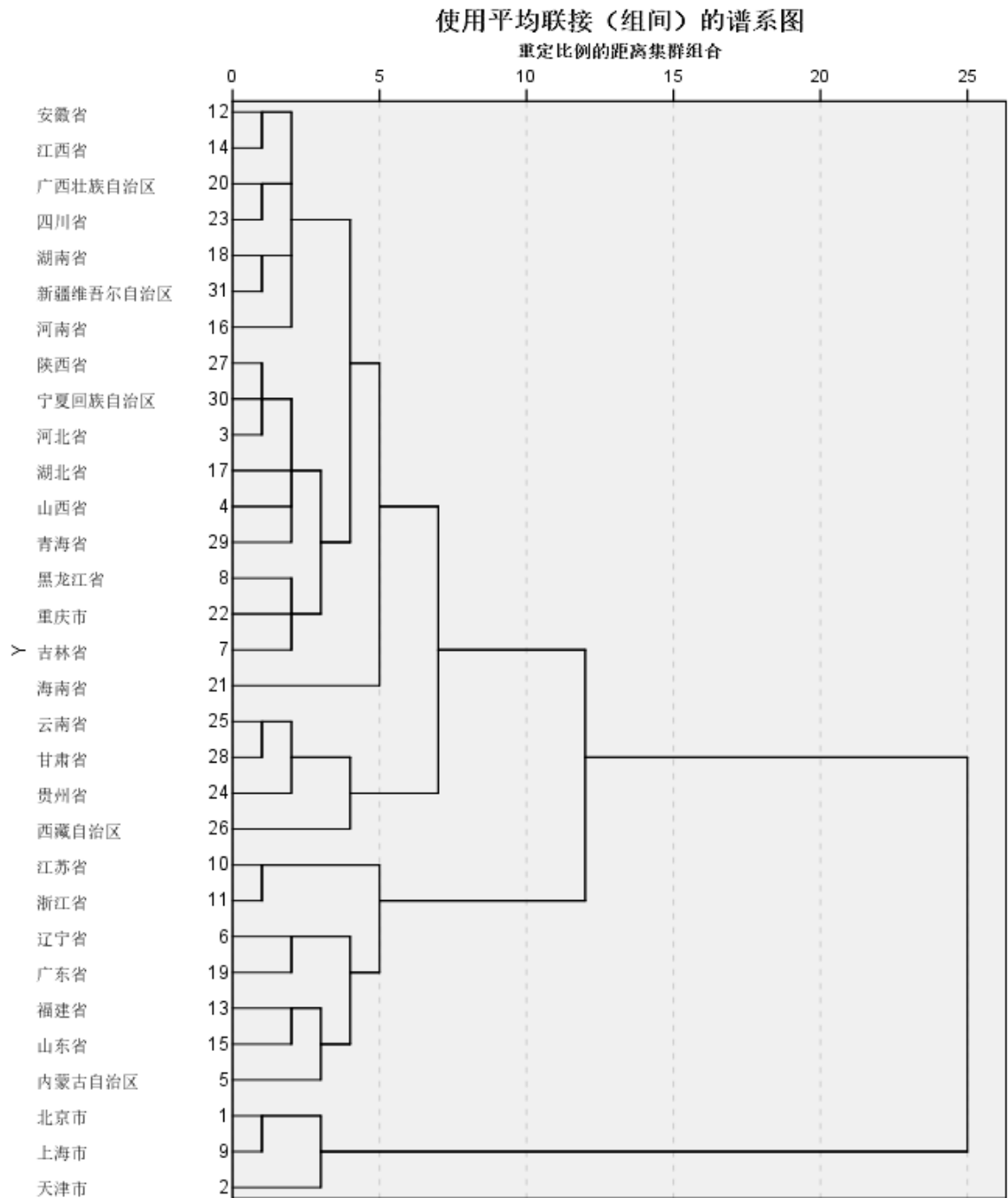


图 2-15 类平均法（组间）聚类谱系图

2.6.3 K-means 聚类及结果

按照 2.6.2 中的结果，设定初始分类数为 4 类，进行 k-means 聚类，得到的结果如下：

表 2-5 聚类的中心

```

> km$centers #各类中心
工业化水平 城镇化率 人均GDP1
1 0.8616567 0.4563025 0.18085458
2 0.9200537 0.5778576 0.50411036
3 0.9885039 0.8440788 0.97125913
4 0.8508952 0.3142413 0.03548721

```

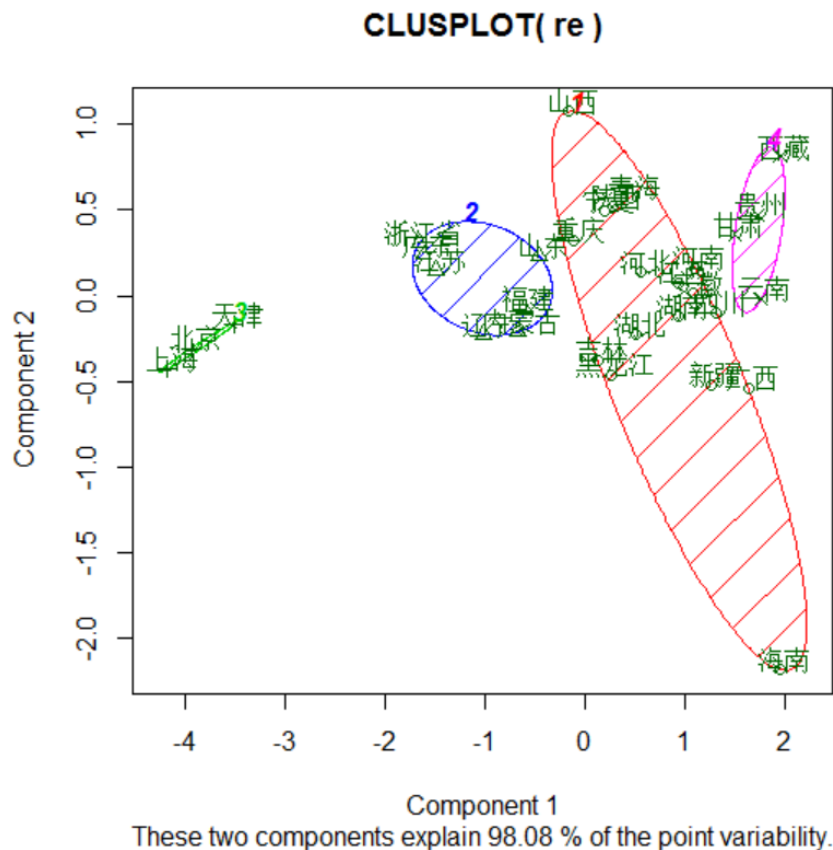


图 2-16 k-means 聚类结果(a)

```

> sort(km$cluster)
河北 山西 吉林 黑龙江 安徽 江西 河南 湖北 湖南 广西 海南
1 1 1 1 1 1 1 1 1 1 1
重庆 四川 陕西 青海 宁夏 新疆 内蒙古 辽宁 江苏 浙江省 福建
1 1 1 1 1 1 2 2 2 2 2
山东 广东 北京 天津 上海 贵州 云南 西藏 甘肃
2 2 3 3 3 4 4 4 4

```

图 2-17 k-means 聚类结果(b)

表 3.6.1 是聚类后各个类的中心，第 3 类的“城镇化率”和“人均 GDP”指标都是最高的，我们称这一类为发达地区；第 2 类各项指标都排在第二位，我们

称它为较发达地区；第 1 类各项指标都排在第三位，不妨称它是中等地区；第 4 类各项指标均在最后，我们称之为落后地区。结合分类的结果，我们根据对各类中心的特点重新排列：

第一类发达地区 3 个：北京、天津、上海；

第二类较发达地区 7 个：内蒙古、辽宁、江苏、浙江、福建、山东、广东；

第三类中等地区 17 个：河北、山西、吉林、黑龙江、安徽、江西、河南、湖北、湖南、广西、海南、重庆、四川、陕西、青海、宁夏、新疆；

第四类落后地区 4 个：贵州、云南、西藏、甘肃。

发达地区为京津沪，较发达地区是东部沿海地区及内蒙、辽宁，中等地区是中西部地区，落后地区是部分西部地区和西南部地区。国家的经济发达水平呈现从东部往西部递减的趋势，从常识上看，分类效果还可以。后续在数据中分别用 vrich、rich、med、poor 标记四类地区。

2.6.4 房价区域差异分析

数据集 region_new 中包含各地区从 2005 年到 2015 年房价的年度数据以及地区分类信息。计算可知，发达地区的房价均值为，较发达地区为，中等地区为，落后地区为，房价均值从发达地区到落后地区递减，相应的分位数也呈现递减的规律。

由于房价总体分布未知，为比较四类地区的房价是否有差异，采用 Kruskal-Wallis 检验和 Wilcoxon 秩和检验。Kruskal-Wallis 检验的结果如下：

```
> kruskal.test(data$平均,data$分类)

Kruskal-Wallis rank sum test

data: data$平均 and data$分类
Kruskal-Wallis chi-squared = 16.798, df = 3, p-value = 0.0007775
```

结果显示 p 值非常小，因此强烈否定零假设，故各类地区的房价不相同。那么各地区房价是否从发达地区到落后地区递减呢？pr1、pr2、pr3 和 pr4 分别是四个地区的房价数据，只需分别检验零假设：

H01:pr1 大于 pr2

H02:pr2 大于 pr3

H03:pr3 大于 pr4

采用 Wilcoxon 秩和检验，结果如下：

```
> wilcox.test(pr1,pr2,alternative="less")

Wilcoxon rank sum test

data:  pr1 and pr2
W = 20, p-value = 0.9917
alternative hypothesis: true location shift is less than 0
```

结果显示，秩和为 20，p 值为 0.9917，强烈接受零假设，即发达地区的房价高于较发达地区。同样地我们可以得出，较发达地区房价高于中等地区，中等地区高于落后区。

2.7 本章小结

本文考虑的可能影响房价影响有：宏观经济形势、国家政策、地区差异、城镇居民收入水平、人均、工业化水平、城镇化水平、房地产投资规模。它们对房价都有一定的影响，影响程度最大的是居民收入水平和地区差异，在一定范围内房地产投资规模是影响房价最小的因素。利率水平影响房产泡沫，利率对房价存在一个滞后的影响，影响程度比较微弱。宏观经济形势、国家政策与利率变动相互关联，当经济形势过热时，国家实行经济紧缩政策，通过提高利率收紧投资和消费；当经济形势低迷时，国家经济政策转为降低利率刺激投资和消费。消费者信心指数（反映消费者对经济形势的主观判断和对收入水平和物价的衡量，当增加时房价增加，降低时房价下跌。各地区的房价存在明显差异，京津沪三个地区的房价是最高的，东部地区次之，中部地区再次之，西部地区房价最低。

第 3 章 房价回归模型

3.1 变量选取与数据处理

第 3 章讨论了房价影响因素极其它们的影响程度，本章用多元回归分析建立房价与这些因素的回归方程，并利用回归方程定量的探讨房价影响因素。回归方程的建立依赖于变量选取、回归诊断，本章的主要内容是如何选取或变换变量建

立回归方程，以及如何判断一个回归模型是否合适。

3.1.1 选择变量建立数据集

由第 3 章的讨论，与房价有较大相关性的变量有城镇化水平、工业化水平、人均 GDP、城镇居民可支配收入、居民消费价格指数、施工面积。消费者信息指数、利率和区域差异对房价的影响已经分别在第二章讨论过，它们虽然对房价有影响，但是由于数据的结构不同于城镇化水平、消费价格指数等变量（前者是全国性的数据，后者是地区性地数据；利率和地区是定型变量，后者均为数值型变量），因此房价回归模型不考虑消费者信心指数、地区和利率的影响。

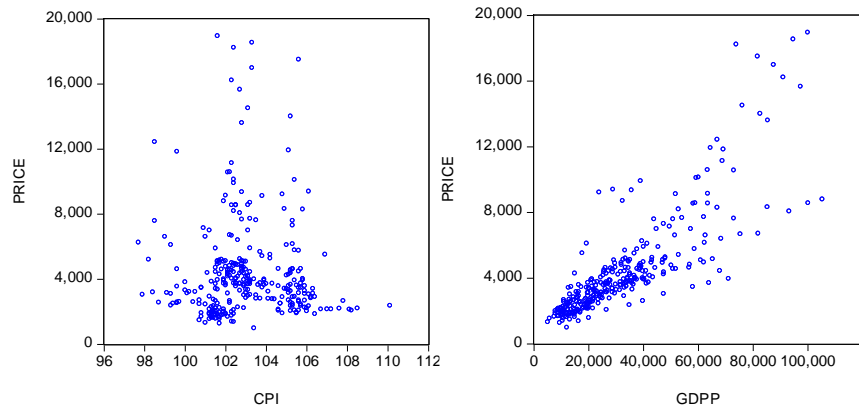
数据集中数据标签 price 代表房价，cpi 代表居民消费价格指数，gdpp 代表人均 GDP，hcon 代表施工面积，income 代表城镇居民可支配收入，indus 代表工业化水平，urblevel 代表城镇化水平。invest 代表投资额，m2 代表货币供应量。代表共计 310 个观测值，来自 31 省份，2005-2014 十年的数据，变量选取如下表 3-1 所示。

用 Eviews 进行操作，看房价与居民消费指数、人均 GDP、施工面积、城镇居民可支配收入、工业化水平与城镇化水平的相关性，得相关表，其中房价与 CPI 的相关系数为-0.038035，相关性非常弱，同时用 Eviews 做房价与 CPI 的散点图得，数据点分布分散，关系非常不明显。因此建立回归方程时，不考虑居民消费指数对房价的影响。但是我们在查阅文献后发现，对于货币供应量的分析应该去除通货膨胀的影响，因此我们用 m2 除以 CPI 来去除通货膨胀的影响

表 3-1 变量选取表

解释变量类型	解释变量	符号	单位
供应方	施工面积	hcon	万平方米
	已完成投资额	invest	万元
需求方	人均 GDP	gdpp	元
	城镇居民收入水平	income	元
	工业化水平	indus	
	城镇化率	urblevel	
货币政策	货币供给量	m2	亿元

注：工业化水平和城镇化率皆没有单位



用 Eviews 分别做房价与人均 GDP、施工面积、城镇居民可支配收入、工业化水平与城镇化水平的散点图，房价与各变量呈现较强的相关性。

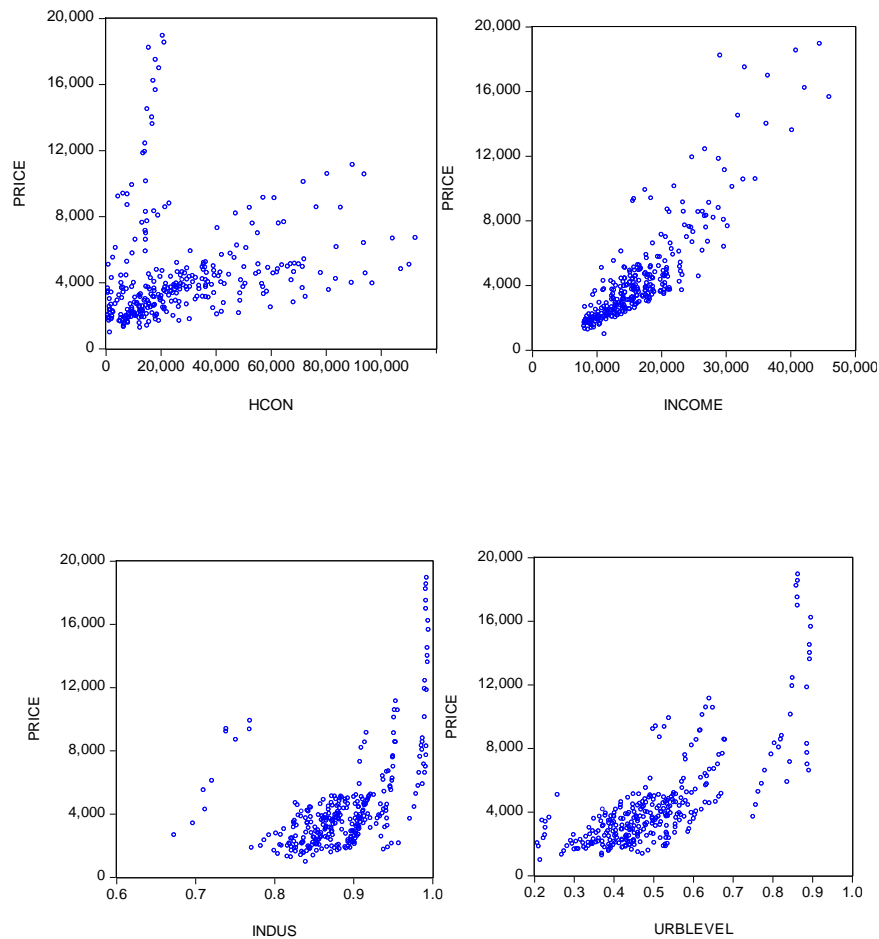


图 3-1 变量与房价的散点图

图反映出房价和城镇化水平可能存在二次项相关，和工业化水平可能与房价

是指数的关系。

4.2 异常值处理

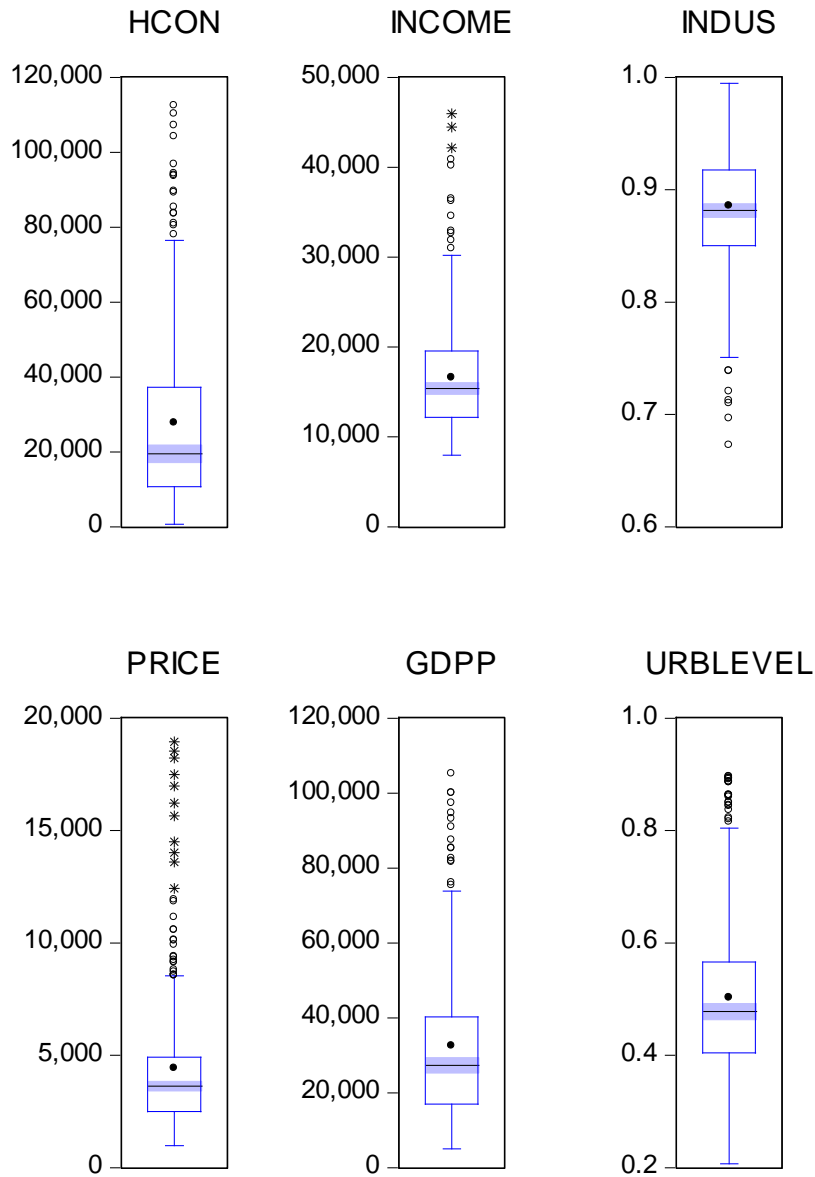


图 3-2 变量箱型图

异常值是一组数据中与平均值的偏差超过两倍标准差的那些数据，偏差超过三倍标准差的数据称为高度异常值。为了检验数据集中各个数值型变量是否有异常值，我们看它们的线箱图，发现前六个变量中有少量的异常值。由于数据观测值只有 310 个，而且数据集中定义的异常值是恒存在的，因此保留这些异常值。在回归诊断时如果发现异常值再删除不晚。

3.2 建立全国房价回归模型

3.2.1 模型选取

利用已有的数据集建立方程

$$\text{price} = f(\text{hcon}, \text{income}, \text{indus}, \text{price}, \text{gdpp}, \text{urblevel} + \text{m2} + \text{invest}) + \varepsilon$$

首先建立单纯多元线性回归模型，1 是混合模型，2 是个体固定效应回归模型，3 是时点固定效应回归模型，4 是个体随机效应回归模型，进行模型效果比较。

1. 混合模型

相应的表达式是：

$$\begin{aligned} \text{price} = & -0.018\ 779 * \text{hcon} + 0.332\ 711 * \text{income} + 0.020\ 639 * \text{gdpp} \\ & + 5784.671 * \text{urblevel} - 464.684 * \text{indus} + 0.000013 * \text{m2} \\ & + 0.000057 * \text{invest} + 4221.167 \\ & (-5.254\ 693)(13.44\ 881)(2.001\ 489)(5.665\ 969)(-4.940\ 303)(0.44)(4.36) \\ & (2.794\ 019) \end{aligned}$$

$$R^2 = 0.825\ 572 \quad SSE_r = 507\ 000\ 000$$

2. 个体固定效应回归模型的估计方法

在 Eviews 将截距项选择区选 Fixed effects（固定效应）

相应的表达式为：

$$\begin{aligned} \text{price} = & -0.016\ 937\text{hcon} + 0.148\ 849\text{income} + 0.080\ 283\text{gdpp} \\ & - 2148.638\text{urblevel} + 6492.412\text{indus} - 4860.399 + 411.367D_1 \\ & - 1789.363D_2 \cdots - 523.030\ 1D_{31} \\ & (-2.518\ 484)(7.155\ 037)(7.972\ 696)(-0.754804)(-1.512\ 825) \\ & (1.447\ 672) \end{aligned}$$

$$R^2 = 0.934\ 291 \quad SSE_u = 191\ 000\ 000$$

其中虚拟变量 $D_1, D_2 \cdots D_{31}$ 的定义是：

$$D_i = \begin{cases} 1, & \text{如果属于第 } i \text{ 个个体, } i = 1, 2, \dots, 31 \\ 0, & \text{其他} \end{cases}$$

接下来用 F 统计量检验是应该建立混合回归模型，还是个体固定效应回归模型。

$H_0: \alpha_i = \alpha$ 。模型中不同个体的截距相同（真实模型为混合回归模型）。

H_1 : 模型中不同个体的截距项 α_i 不同（真实模型为个体固定效应回归模型）。

F 统计量定义为：

$$F = \frac{(SSE_r - SSE_u) / [(NT - k - 1) - (NT - N - k)]}{SSE_u / (NT - N - k)} = \frac{(SSE_r - SSE_u) / (N - 1)}{SSE_u / (NT - N - k)}$$

其中 SSE_r 表示约束模型，即混合估计模型的残差平方和， SSE_u 表示非约束模型，即个体固定效应回归模型的残差平方和。非约束模型比约束模型多了 $N - 1$ 个被估参数。

所以本例中：

$$F = \frac{(507\ 000\ 000 - 191\ 000\ 000) / (31 - 1)}{191\ 000\ 000 / (310 - 31 - 1)} = 15.331 > F_{0.05}(30, 300) = 1.5$$

所以推翻原假设，建立个体固定效应回归模型更合理。

3. 时点固定效应回归模型的估计方法

在 Eviews 中将时间选择为固定效应。相应的表达式为：

$$\begin{aligned} \text{price} = & -0.022\ 974\text{hcon} + 0.497\ 865\text{income} - 0.027\ 729\text{gdpp} \\ & + 5936.791\text{urblevel} - 8971.160\text{indus} + 2654.514 + 460.254\ 5D_1 \\ & + 313.354\ 2D_2 \cdots + 807.939\ 5D_{10} \\ & (-7.234\ 910)(18.314\ 08)(2.362\ 203)(5.778\ 293)(-5.249\ 426) \\ & (1.937\ 968) \end{aligned}$$

$$R^2 = 0.872\ 454 \quad SSE_u = 371\ 000\ 000$$

其中虚拟变量 $D_1, D_2 \dots D_{10}$ 的定义是：

$$D_t = \begin{cases} 1, \text{如果属于第 } t \text{ 个截面, } t = 2005, 2006 \dots, 2010 \\ 0, \text{其他} \end{cases}$$

4. 个体随机效应回归模型估计

在 Eviews 中选择 Random effects（个体随机效应）相应的表达式为：

$$\begin{aligned} \text{price} = & -0.019 \ 501 \text{hcon} + 0.182 \ 106 \text{income} + 0.067 \ 038 \text{gdpp} \\ & + 3211.710 \text{urblevel} - 3061.709 \text{indus} + 855.7715 + 2967.398 D_1 \\ & - 1892.538 D_2 \dots - 585.5701 D_{31} \\ & (-4.334 \ 021)(9.104 \ 869)(7.339 \ 129)(2.278 \ 726)(-1.059 \ 754) \\ & (0.380 \ 952) \end{aligned}$$

$$R^2 = 0.771 \ 874 \quad SSE_u = 238 \ 000 \ 000$$

其中虚拟变量 $D_1, D_2 \dots D_{31}$ 的定义是：

$$D_i = \begin{cases} 1, \text{如果属于第 } i \text{ 个个体, } i = 1, 2 \dots, 31 \\ 0, \text{其他} \end{cases}$$

接下来利用 Hausman 统计量检验应该建立个体随机效应回归模型还是个体固定效应回归模型。

H_0 ：个体效应与回归变量（ IP_{it} ）无关（个体随机效应回归模型）

H_1 ：个体效应与回归变量（ IP_{it} ）相关（个体固定效应回归模型）

得到如下检验结果：

Correlated Random Effects - Hausman Test				
Pool: Untitled				
Test cross-section random effects				
<hr/>				
Test Summary	Chi-Sq. Statistic	Chi-Sq. d.f.	Prob.	
<hr/>				
Cross-section random	41.836589	5	0.0000	
<hr/>				
Cross-section random effects test comparisons:				
Variable	Fixed	Random	Var(Diff.)	Prob.
<hr/>				
HCON?	-0.016937	-0.019501	0.000025	0.6080
INCOME?	0.148849	0.182106	0.000033	0.0000
GDPP?	0.080283	0.067038	0.000018	0.0018
URBLEVEL?	-2148.638...	3211.710079	6116724.2...	0.0302
INDUS?	6492.412...	-3061.7091...	10070912....	0.0026

图 3-3 Hausman 检验结果

由检验输出结果 3-3 的上半部分可以看出, Hausman 统计量的值是 41.836589, 相对应的概率是 0.000 0, 即拒接原假设, 应该建立个体固定效应模型。综上所述, 我们将建立个体固定效应模型来对于回归模型进行分析

3.2.2 模型建立

首先将全部变量建立多元回归模型, 得到的系数表如下表 3-1 所示:

表 3-2 未删除异常值的回归结果

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-21443.31	4022.678	-5.330606	0.0000
INVEST	6.89E-05	1.28E-05	5.369378	0.0000
INDUS	155.4849	3983.687	0.039030	0.9689
INCOME	0.118526	0.018476	6.415068	0.0000
HCON	-0.060790	0.010055	-6.045796	0.0000
GDPP	0.027354	0.010870	2.516513	0.0124
URBLEVEL	-6417.756	2581.576	-2.485984	0.0135
LM2	1691.781	316.9012	5.338513	0.0000

我们在此基础上进行了残差分析, 绘制了残差图 3-4

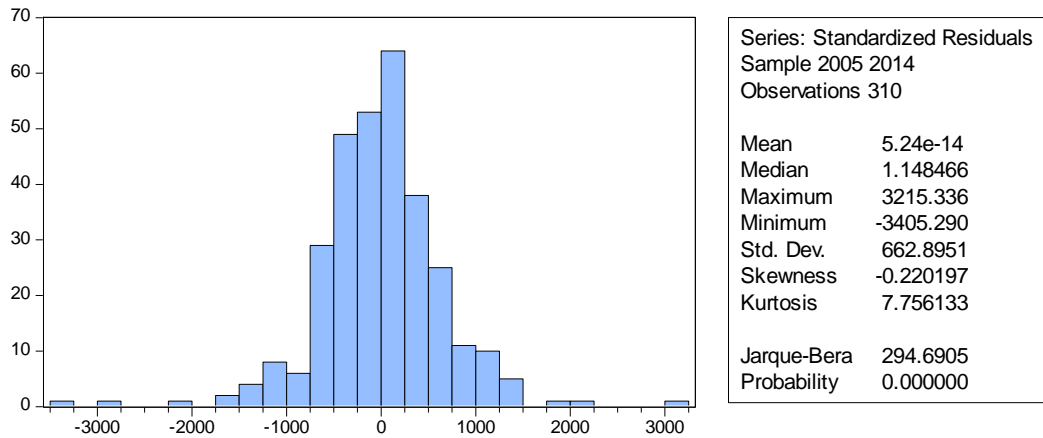


图 3-4 去除异常值的残差图

我们可以从以上的残差图中看出，残差的正态性略差，并且有异常值，所以我们进一步绘制残差表来确定异常值的编号，从残差表见附录 1，我们取超过 2 倍标准差的为异常值。删除北京 2005 年，2006 年，2011 年还有上海 2008 年，2010 年的数据。再次进行回归，得到的结果如下表 3-3 所示，我们在此基础上绘制了残差图 3-5：

表 3-3 删除异常值后的回归结果

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-952.4599	2541.469	-0.374767	0.7081
HCON	-0.038765	0.008127	-4.770026	0.0000
GDPP	0.010910	0.009614	1.134767	0.2575
INCOME	0.143195	0.014469	9.896539	0.0000
INDUS	4149.708	3095.506	1.340559	0.1812
INVEST	4.49E-05	1.03E-05	4.344974	0.0000
M2	0.000209	3.11E-05	6.716709	0.0000
URBLEVEL	-4820.853	2069.139	-2.329884	0.0206

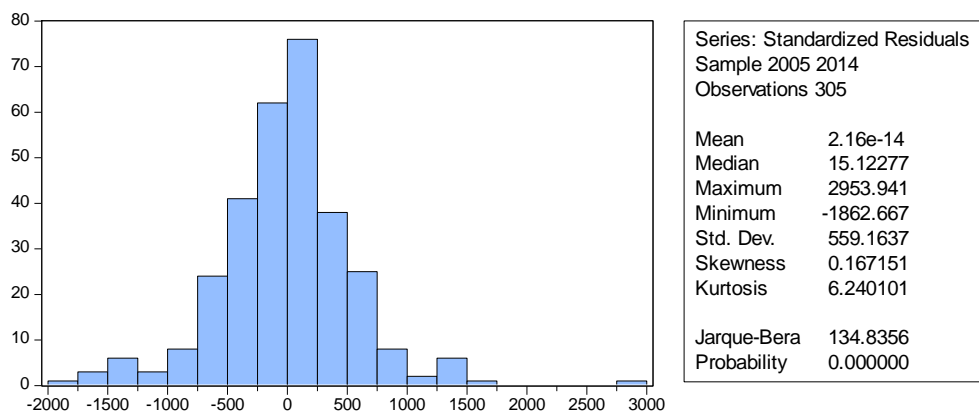


图 3-5 去除异常值后的残差图

我们从残差图中可以看到模型残差的更具有正态性，并且残差的标准差也有由 662.289 降到 559.1637 但是我们从上表中看出有的变量并不显著，于是我们进行了相关性分析，：

表 3-4 变量相关性矩阵

	PRICE	HCON	GDP	INCOME	INDUS	INVEST	M2	URBLEVEL
PRICE	1	0.250690037	0.830197991	0.88287	0.526099	0.511438	0.458526	0.766851
HCON	0.25069004	1	0.419167314	0.434426	0.282673	0.899949	0.44838	0.227023
GDP	0.83019799	0.419167314	1	0.886128	0.692803	0.630261	0.591442	0.842086
INCOME	0.88287018	0.434425814	0.886128486	1	0.616204	0.634984	0.538554	0.735035
INDUS	0.5260989	0.282673411	0.692803471	0.616204	1	0.409045	0.216545	0.702834
INVEST	0.51143841	0.899949419	0.630261372	0.634984	0.409045	1	0.557384	0.45513
M2	0.45852594	0.44837984	0.591442263	0.538554	0.216545	0.557384	1	0.250249
URBLEVEL	0.76685089	0.227022636	0.842086184	0.735035	0.702834	0.45513	0.250249	1

我们从上表 3-4 中可以看出变量之间存在着一定的相关性，所以我们将所有变量都加进去容易出现多重共线性的影响。因此，我们采取逐步回归的方式进行回归。得到的最终模型如下表 3-5：

表 3-5 房价影响因素回归结果

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	550.7241	156.9397	3.509145	0.0005
RM2	0.021420	0.001902	11.26414	0.0000
INVEST	4.68E-05	1.00E-05	4.674391	0.0000
INCOME	0.156279	0.012158	12.85351	0.0000
HCON	-0.043134	0.007879	-5.474453	0.0000

建立的回归模型为：

$$\text{price} = 550.724 + 0.021 * \text{rm2} + 0.00004 * \text{invest} + 0.156 * \text{income} - 0.043 * \text{hcon}$$

$$R^2 = 0.963$$

$$\overline{R^2} = 0.959$$

$$DW = 1.162$$

其中拟合系数达到 0.959，说明模型的拟合效果较好，并且由于采用的是逐步回归的方法，已经不存在多重共线性，下面我们绘制了残差图，可以看到残差的正态性还是比较好的。

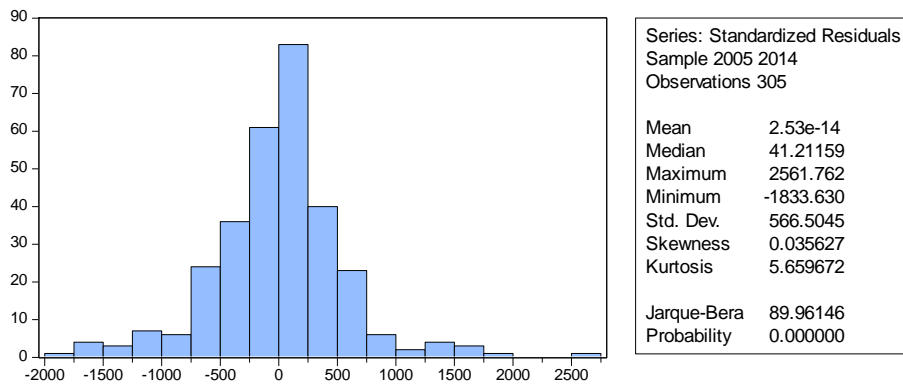


图 3-6 最终模型残差图

最终进入回归模型的变量为本年房地产完成投资、住宅施工面积和城镇居民可支配收入还有货币供给量。供给方因素最终进入回归模型的变量为本年房地产完成投资（投资因素）与住宅施工面积（供给意愿因素）。房价与本年房地产完成投资正相关，这是由于房地产完成投资增加意味着对房地产市场预期的看好，因此房价会随之上升。本年开发土地面积与房价负相关，这是因为开发土地面积增加后，市场上的住房供给随之增加，进而使得房价下跌。

3.3 建立区域回归模型

根据第二章聚类分析的结果，我们分别对四个区域建立多元回归模型

1. 发达地区

$$\text{price} = 1396.44 + 0.212 \cdot \text{income} - 0.219 \cdot \text{hcon} + 0.00037 \cdot \text{invest}$$

$$(645.9) \quad (0.048) \quad (0.045) \quad (0.00007)$$

$$t = (2.16) \quad (-4.33) \quad (-4.89) \quad (5.37)$$

$$R^2 = 0.975 \quad \overline{R}^2 = 0.969 \quad DW = 1.33$$

回归方程修正可决系数为 0.969，方程拟合优度较高；各个解释变量回归系数的显著性水平均小于 0.05，因此，模型中各变量对房价都有显著影响，模型解释能力较强。进一步对模型进行计量经济学检验，分别使用了逐步回归法、white 检验和 DW 检验发现：模型不存在多重共线性，也无异方差和自相关性，因此模型设定与回归结果无误。

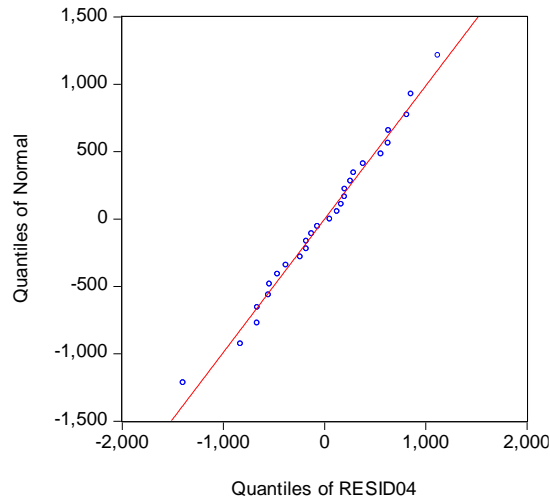


图 3-7 发达地区回归模型残差 QQ 图

从回归结果可以看出,发达地区的房价与本年的城镇居民收入和房地产完成投资额正相关,与本年商品住宅施工面积成负相关。本年房地产完成投资每增加 1 亿元,每平方米住房价格上升约 3.7 元。房地产完成投资增加意味着对市场预期的看好,因此,当房地产完成投资增加时房价会随之上升。人均可支配收入每增加 1 元,每平方米住房价格上升 0.212 元。当收入水平上升时,人们购买住房的能力上升,同时也希望扩大住房面积,所以会购买更多的住房,对住房的需求上升,拉动房价的上升。本年商品住宅面积每增加 1 万平方米,每平方米住房价格下降约 0.219 元。这是因为,施工土地面积增加,意味着市场上住房供给增加,因而使得住房价格下降。各变量回归系数的符号均符合经济含义。

总的来说,对于发达地区,所受供给的影响是比较显著的,这与我们的现实情况是相符的,比如在北京,上海和天津。北京和上海分别作为我国的政治中心和经济中心,因经济发达、工业化进程较快、人口流动大、金融业发展水平较高和房地产升值潜力较大等因素的影响,使得这两个城市的房地产商投资、居民购房更多考虑其政治、经济地位和获得更多的就业、求学机会,该地区投资、消费意愿强烈。所以这些城市的房住目前来说还是供不应求的,所以房价收到供给的影响是比较大的。

2. 较发达地区

$$\text{price} = 898.977 + 0.0235 \cdot \text{rm2} + 0.121 \cdot \text{income}$$

$$(385.17) \quad (0.004) \quad (0.0431)$$

$$t = (2.33) \quad (-5.41) \quad (3.89)$$

$$R^2 = 0.932 \quad \overline{R^2} = 0.923 \quad DW = 1.58$$

回归方程修正可决系数为 0.923，方程拟合优度较高；各个解释变量回归系数的显著性水平都大于 95%，因此，模型中各变量对房价都有显著影响，模型解释能力较强。进一步对模型进行计量经济学检验，分别使用了逐步回归法、white 检验和 DW 检验发现：模型不存在多重共线性，也无异方差和自相关性，因此模型设定与回归结果无误。我们从 QQ 图可以看出，该回归模型的正态性是比较好的

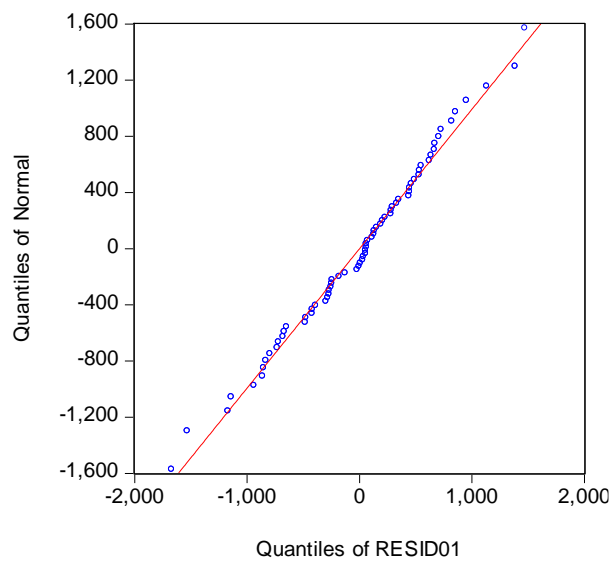


图 3-8 较发达地区回归模型 QQ 图

从回归结果可以看出，较发达地区的房价货币供给量与城镇居民可支配收入呈正相关关系。货币供给量每增加 1 亿元，每平方米住房价格上升约 0.02 元。城镇居民可支配收入每增加 1 元，每平方米住房价格上升 0.121 元。当收入水平上升时，人们购买住房的能力上升，同时也希望扩大住房面积，所以会购买更多的住房，对住房的需求上升，拉动房价的上升。所以对于较发达地区，货币供给量以及需求的影响比较大。

3. 中等发达地区

$$\text{price} = -518.561 + 2826.4 * \text{ubrelevel} + 0.039 * \text{income} + 0.078 * \text{gdpp}$$

$$(498.92) \quad (793.49) \quad (0.016) \quad (0.006)$$

$$t = (-1.039) \quad (2.17) \quad (-2.76) \quad (12.722)$$

$$R^2=0.943 \quad \overline{R^2}=0.936 \quad DW=1.692$$

回归方程修正可决系数为 0.936，方程拟合优度较高；各个解释变量回归系数的显著性水平都大于 95%，因此，模型中各变量对房价都有显著影响，模型解释能力较强。进一步对模型进行计量经济学检验，分别使用了逐步回归法、white 检验和 DW 检验发现：模型不存在多重共线性，也无异方差和自相关性，因此模型设定与回归结果无误。从残差图来看，残差还是比较平稳的。

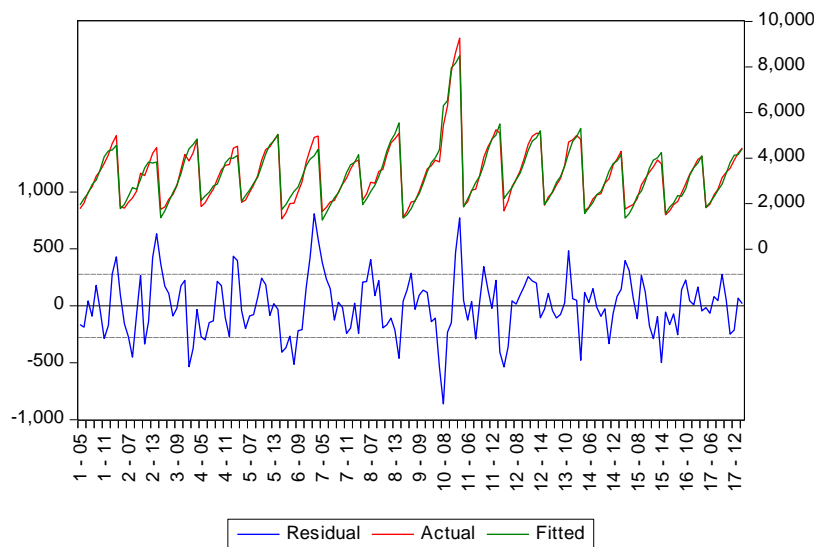


图 3-9 中等发达地区回归模型残差图

从回归结果可以看出，房价与地区城市化水平和地区人均 GDP 以及城镇居民人均可支配收入正相关。人均 GDP 每增加 1 元，每平方米住房价格上升 0.078 元。城镇居民人均可支配收入每增加 1 元，每平方米住房价格上升 0.039 元。当收入水平上升时，人们购买住房的能力上升，同时也希望扩大住房面积，所以会购买更多的住房，对住房的需求上升，拉动房价的上升。城市化率每提高 1%，每平方米住房价格上升约 28.26 元。这是因为，开发土地面积增加，意味着市场上住房供给增加，因而使得住房价格下降。各变量回归系数的符号均符合经济含义。

所以总的来说，对于中等发达地区，影响房价的最重要的因素是需求，这是与现实情况比较相符的，因为对于中等发达地区，住房主要的作用是消费品，人们对于住房是刚性的需求，投资价值不大。所以随着人们收入水平的提高和城市

化水平的提高。需求推动房价上升

4. 欠发达地区

$$\begin{array}{cccc}
 \text{price} = & 68.55 & +2048.012 * \text{urblevel} & + 0.038 * \text{income} & + 0.101 * \text{gdpp} \\
 & (259.76) & (0.048) & (0.045) & (0.008) \\
 t = & (0.262) & (2.58) & (2.34) & (11.32) \\
 R^2 = & 0.90 & \overline{R^2} = 0.89 & DW = 1.415 &
 \end{array}$$

回归方程修正可决系数为 0.936，方程拟合优度较高；各个解释变量回归系数的显著性水平都大于 95%，因此，模型中各变量对房价都有显著影响，模型解释能力较强。进一步对模型进行计量经济学检验，分别使用了逐步回归法、white 检验和 DW 检验发现：模型不存在多重共线性，也无异方差和自相关性，因此模型设定与回归结果无误。因为欠发达地区的数据量比较少，残差分布图太直观，所以我们绘制了残差与残差正态分布的 QQ 图，从 QQ 图来看，残差的正态性还是比较好的

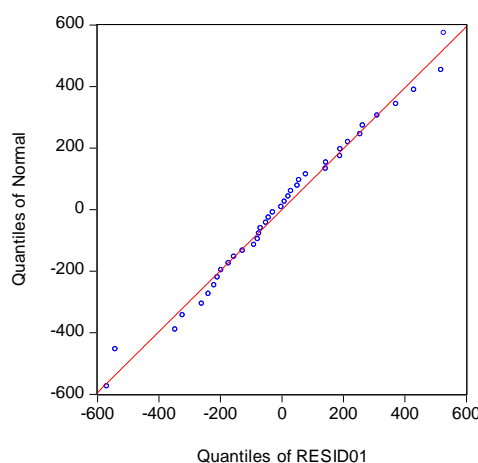


图 3-10 欠发达地区回归模型残差 QQ 图

欠发达地区进入模型的变量与中等发达地区是相同的。从回归结果可以看出，房价与地区城市化水平和地区人均 GDP 以及城镇居民人均可支配收入正相关。人均 GDP 每增加 1 元，每平方米住房价格上升 0.101 元。城镇居民人均可支配收入每增加 1 元，每平方米住房价格上升 0.038 元。当收入水平上升时，人们购

买住房的能力上升，同时也希望扩大住房面积，所以会购买更多的住房，对住房的需求上升，拉动房价的上升。城市化率每提高 1%，每平方米住房价格上升约 20.48 元。这是因为，开发土地面积增加，意味着市场上住房供给增加，因而使得住房价格下降。各变量回归系数的符号均符合经济含义。所以总的来说，对于欠发达地区，影响房价的最重要的因素也是需求，这是与现实情况比较相符的，因为对于欠发达地区，住房主要的作用是消费品，人们对于住房是刚性的需求，投资价值不大。所以随着人们收入水平的提高和城市化水平的提高。需求推动房价上升。

3.4 本章小结

对于全国范围内来说，对面板数据建立了多元回归模型。最终进入回归模型的变量为本年房地产完成投资、住宅施工面积和城镇居民可支配收入还有货币供给量。供给方因素最终进入回归模型的变量为本年房地产完成投资（投资因素）与住宅施工面积（供给意愿因素）。房价与本年房地产完成投资正相关，这是由于房地产完成投资增加意味着对房地产市场预期的看好，因此房价会随之上升。本年住宅施工与房价负相关，这是因为住宅施工面积增加后，市场上的住房供给随之增加，进而使得房价下跌。

总的来说四个地区的房价都受到需求因素的影响，但是每个地区的经济发展水平不同，房价影响因素也有区别。对于发达地区，所受供给的影响是比较显著的，说明在发达地区，房地产显示出了较强的投资品的属性。对于较发达地区的房价所受货币供给量的影响显著。而在中等发达和欠发达地区，房价受到需求也就是人们的收入水平和经济发展水平的影响比较大。

第四章 宏观政策对房价影响分析

4.1 货币政策对房价影响分析

为了进一步研究货币政策对于房价的影响，我们采用计量经济学中向量自回归模型（VAR）进行分析，首先对于房价和货币统计量 $m2$ 取对数，采用 ADF 单位根检验原始序列的平稳性。然后利用格兰杰因果检验（Granger Causality Test）分析两个变量间的因果关系，最后为了更好、更直观地研究货币供给量对房价的影响，本文将采用脉冲响应函数来描述货币供给量对于房价的影响

4.1.1 全国范围货币政策影响

1. 单位根检验

我们首先对全国 31 个省市自治区进行分析。以下为对所有变量及其一阶差分进行的 ADF 单位根检验结果。对房价 Price 和货币供给量 M2 的检验包含趋势项和截距。由 4-1 表，原始序列中房价为平稳序列，但是货币供给量为非平稳序列。但一阶差分后，变量都在 1%水平显著，趋于平稳。

表 4-1 ADF 检验

变量	检验类型	ADF统计量	P值	结论
lnhp	(C,T,0)	-1.6545	0.049**	平稳
d.lnhp	(C,0)	-0.1679	0.4333	非平稳
lnm2	(C,T,0)	5.1826	1.0000	非平稳
d.lnm2	(C,0)	-5.9873	0.000***	平稳

注：上表中，d. 变量名表示对于变量进行一阶差分，检验类型中的C表示包含截距项，T表示包含趋势项，数字0表示滞后期数为0；*，**，***分别表示统计量在10%，5%和1%的显著性水平上显著

2. 格兰杰因果检验

根据 Granger 因果关系的定义，具有因果关系的含义是一个事件 X 如果是另一个事件 Y 的原因，则事件 X 应当领先事件 Y。下表为货币供应量与房价之间，以及变量一阶差分之间的 Granger 因果检验统计结果。

表 4-2 格兰杰因果检验

滞后期数	因果关系	Chi2统计量	P-value
Lag=1	d.lnm2→lnprice	17.3789	0.000***
	lnprice→d.lnm2	3.1278	0.078
Lag=2	d.lnm2→lnprice	12.0991	0.000***
	lnprice→d.lnm2	7.2821	0.0009***
Lag=3	d.lnm2→lnprice	4.4353	0.0049***
	lnprice→d.lnm2	24.2453	0.000***

由表 4-2，房价与 M2 一阶差分后因果关系较为显著。再滞后一期，房价对

货币供给量在 10%水平上显著，在后面的滞后期，房价均在 1%水平上是货币供给量的显著 Granger 原因。这说明，房价上涨对货币供给的影响可能需要较长的时间才能显现。而对于表 4-2 中的任何滞后期，货币供应量 M2 均在 1%水平上是房价的显著 Granger 原因。这一结果说明，在我国，货币供应量对房价的领先作用十分明显，M2 水平的上升会导致房价的上涨，且传导的速度很可能是比较迅速的。

3. 脉冲响应函数

脉冲响应函数反映了来自随机扰动项的一个标准差的新信息冲击对内生变量当前和未来取值的影响，以及其影响路径的变化。图 2 为基于 VAR(1)模拟的脉冲响应函数曲线，横轴代表滞后阶数，将滞后阶数设为 5 期，纵轴表示房价对货币供应量 M2 的响应程度，图中的实线为实际计算值，阴影部分为 95%置信区间。根据图 2 的脉冲响应函数曲线，房屋价格在受到 M2 的一个正向冲击后，会在之后的 2 年内持续上升。也就是说，货币供应量增加对房地产价格有长期的持续正向影响，随后在第 2 期达到了脉冲响应的峰值，之后就一直下降在第 5 期之后响应基本趋于稳定，响应值趋于零。

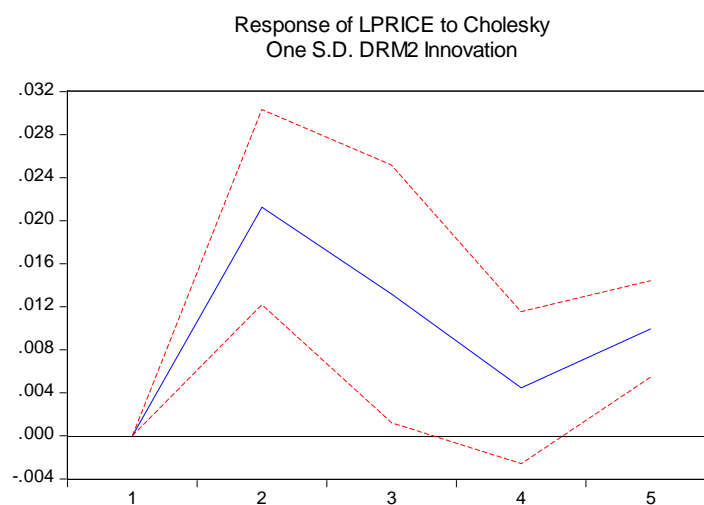


图 4-1 房价对 m2 的脉冲分析

也就是说，从全国范围内来看，货币供给量对于房价的影响还是比较显著的，当一个货币政策实施后，会在 2 年内对房价的影响，这种影响在大约 2 年后开始减弱，直至 4 年后趋于消失。

4.1.2 发达地区货币政策影响

1. 单位根检验

以下为对所有变量及其一阶差分进行的 ADF 单位根检验结果。对发达地区的房价 Price 和货币供给量 M2 的检验包含趋势项和截距，由下表 4-3，原始序列均为非平稳序列，但一阶差分后，变量都在 1%水平显著，趋于平稳

表 4-3 ADF 检验

变量	检验类型	ADF统计量	P值	结论
lnhp	(C,T,0)	-1.1129	0.1329	非平稳
d.lnhp	(C,0)	-2.1115	0.0174**	平稳
lnm2	(C,T,0)	1.5856	0.9436	非平稳
d.lnm2	(C,0)	-1.7633	0.0389**	平稳

注：上表中，d. 变量名表示对于变量进行一阶差分，检验类型中的C表示包含截距项，T表示包含趋势项，数字0表示滞后期数为0；*，**，***分别表示统计量在10%，5%和1%的显著性水平上显著

2. 格兰杰因果检验

对于发达地区房价的一阶滞后值和货币供给量的一阶滞后值进行格兰杰因果检验，检验结果如下表 4-4：

表 4-4 格兰杰因果检验

滞后期数	因果关系	Chi2统计量	P-value
Lag=1	d.lnm2 → d.lnprice	7.0471	0.0161**
	d.lnprice → d.lnm2	3.7141	0.0699*
Lag=2	d.lnm2 → d.lnprice	1.3242	0.2996
	d.lnprice → d.lnm2	0.7408	0.4958
Lag=3	d.lnm2 → d.lnprice	1.66841	0.2308
	d.lnprice → d.lnm2	0.66841	0.5646

由表 4-4, 房价与 M2 因果不太显著。只有在滞后 1 期时, 统计检验量为 0.069, 房价在 10% 水平上是 M2 的 Granger 原因, 但这一因果关系较为微弱。M2 在 5% 水平上是房价的 Granger 原因。在之后的 3 期中, 二者的因果关系不太显著。这说明对于发达地区, 货币政策对于房价的影响并不是很显著。

3. 脉冲响应函数与方差分解

接下来建立发达地区房价对于货币供给量的脉冲响应函数

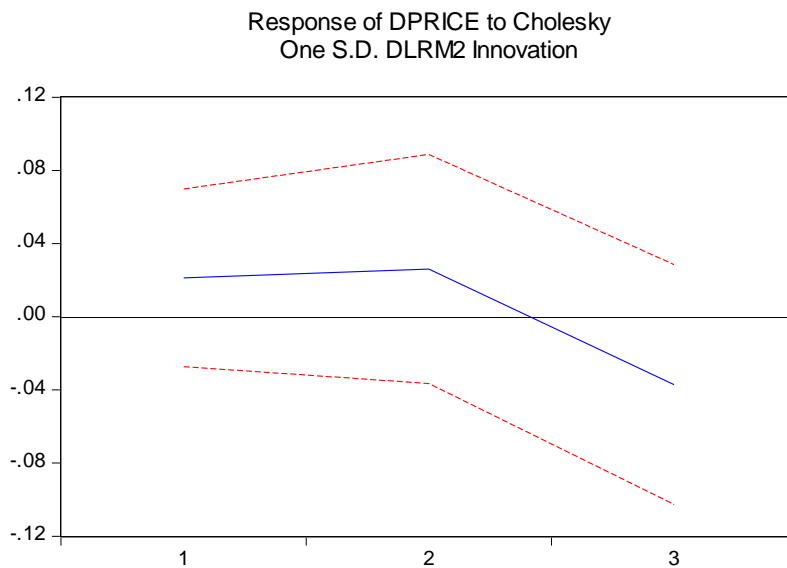


图 4-2 房价对 m2 的响应函数

根据图 4-2 的脉冲响应函数曲线, 房屋价格在受到 M2 的一个正向冲击后, 会在之后的 2 年内有正向的相应, 但是响应程度很弱。之后就一直下降甚至为负, 最后趋近于 0。

4.1.3 较发达地区货币政策影响

1. 单位根检验

我们接着对较发达地区进行分析。以下为对所有变量及其一阶差分进行的 ADF 单位根检验结果。对房价 Price 和货币供给量 M2 的检验包含趋势项和截距。由表 4-5, 原始序列中房价为平稳序列, 但是货币供给量为非平稳序列。但一阶差分后, 变量都在 1% 水平显著, 趋于平稳。

表 4-5 ADF 检验

变量	检验类型	ADF 统计量	P 值	结论
lnprice	(C,T,0)	-2.1803	0.0146**	平稳
d.lnprice	(C,0)	0.9035	0.8169	非平稳
lnm2	(C,T,0)	2.3169=2	0.9897	非平稳
d.lnm2	(C,0)	-2.8881	0.0019***	平稳

注：上表中，d. 变量名表示对于变量进行一阶差分，检验类型中的 C 表示包含截距项，T 表示包含趋势项，数字 0 表示滞后期数为 0；*，**，***分别表示统计量在 10%，5%和 1%的显著性水平上显著

2. 格兰杰因果检验

对于较发达地区房价和货币供给量的一阶滞后值进行格兰杰因果检验，检验结果如下表 4-6 所示：

表 4-6 格兰杰因果检验

滞后期数	因果关系	Chi2统计量	P-value
Lag=1	d.lnm2→lnprice	3.3644	0.0723*
	lnprice→d.lnm2	1.1495	0.2886
Lag=2	d.lnm2→lnprice	2.1309	0.1311
	lnprice→d.lnm2	3.4068	0.0423**
Lag=3	d.lnm2→lnprice	1.8802	0.1515
	lnprice→d.lnm2	10.0216	0.000***

由表 4-6，较发达地区房价与 M2 因果关系较为显著。在滞后 2 期时，统计检验量为 0.0432，房价在 5%水平上是 M2 的 Granger 原因，但这一因果关系较为微弱。这说明，房价上涨对货币供给的影响可能需要较长的时间才能显现。而对于表 3 中的 1 阶滞后期，货币供应量 M2 均在 10%水平上是房价的显著 Granger 原因。这说明较发达地区对于货币供应量的响应是比较迅速的。但是在长期二者的因果关系就不明显了。

3. 脉冲响应函数

接下来建立较发达地区房价对于货币供给量的脉冲响应函数

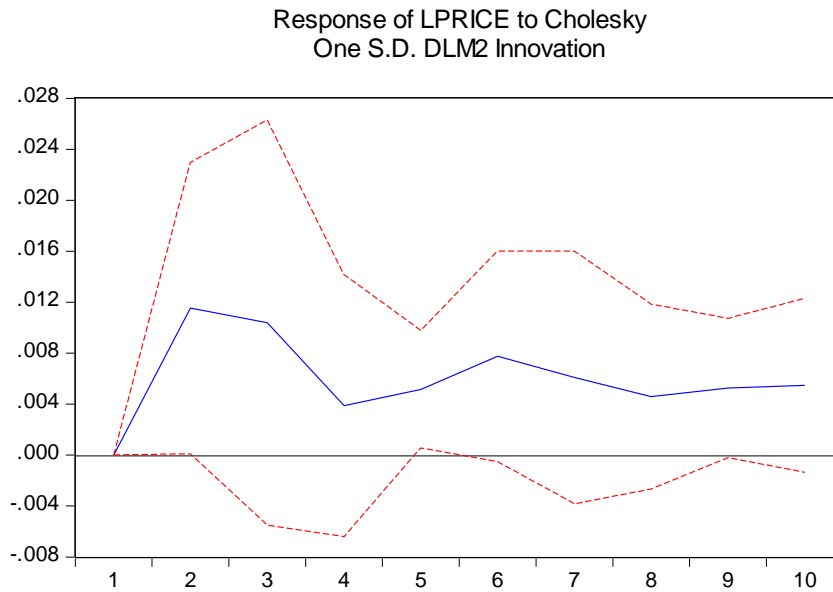


图 4-3 房价对 m2 的响应函数

较发达地区商品房价格对的脉冲响应几乎一直表现为正向冲击。其响应值从一开始逐渐上升，在第 2 期达到了正的响应峰值，响应比较深刻，之后就逐渐趋于下降趋势，变化比较平稳，在第 10 期之后响应值就趋于零了。

4.1.4 中等发达地区货币政策影响

1. 单位根检验

因为在之前建立回归模型的时候，两个地区中货币供应量都没有对货币供应量有显著性影响，所以我们主要对中等发达地区进行分析。以下为对所有变量及其一阶差分进行的 ADF 单位根检验结果。对房价 Price 和货币供给量 M2 的检验包含趋势项和截距，由上表，原始序列均为非平稳序列，但一阶差分后，M2 在 1% 水平显著，而房价却在二阶差分后才显著。

表 4-7 ADF 检验

变量	检验类型	ADF 统计量	P 值	结论
lnprice	(C,T,0)	3.2407	0.9994	非平稳
d.lnprice	(C,0)	-0.3468	0.3644	非平稳
d2.lnprice	(C,0)	-3.0302	0.0012***	平稳
lnm2	(C,T,0)	0.2657	0.6048	非平稳
d.lnm2	(C,0)	-4.4034	0.0000***	平稳

注：上表中，d. 变量名表示对于变量进行一阶差分，d2. 变量名表示对于变量进行二阶差分，检验类型中的 C 表示包含截距项，T 表示包含趋势项，数字 0 表示滞后期数为 0；*，**，*** 分别表示统计量在 10%，5% 和 1% 的显著性水平上显著

2. 格兰杰因果检验

对于较发达地区房价的二阶滞后值和货币供给量的一阶滞后值进行格兰杰因果检验，检验结果如下表 4-8 所示

表 4-8 格兰杰因果检验

滞后期数	因果关系	Chi2 统计量	P-value
Lag=1	d.lnm2 → d.lnprice	14.8348	0.0002***
	d.lnprice → d.lnm2	13.6326	0.0003***
Lag=2	d.lnm2 → d.lnprice	4.0771	0.0201**
	d.lnprice → d.lnm2	6.75825	0.0018***
Lag=3	d.lnm2 → d.lnprice	6.34044	0.0007***
	d.lnprice → d.lnm2	0.689	0.5617

由表 4-8，房价与 M2 因果关系较为显著。只有在滞后 4 期时，统计检验量为 0.5617，房价在 10% 水平上是 M2 的 Granger 原因，但这一因果关系较为微弱。这说明，房价上涨对货币供给的影响可能需要较长的时间才能显现。而对于表 4-8 中的任何滞后期，货币供应量 M2 均在 1% 水平上是房价的显著 Granger 原因。这一结果说明，在我国，货币供应量对房价的领先作用十分明显，M2 水平的上升会导致房价的上涨，且传导的速度很可能是比较迅速的。

3. 脉冲响应函数

接下来建立较发达地区房价对于货币供给量的脉冲响应函数

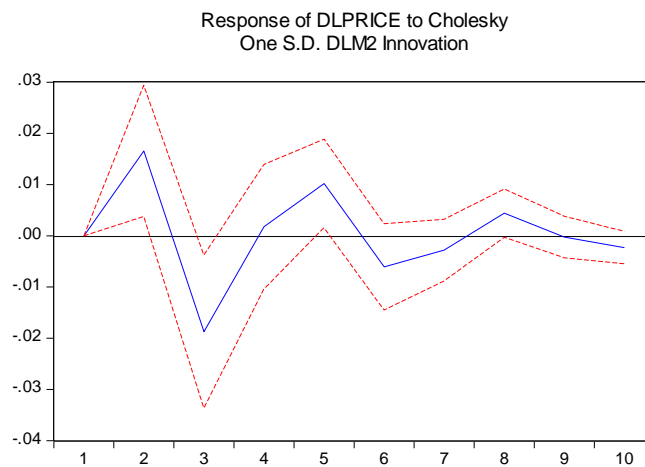


图 4-4 房价对 m2 的脉冲响应函数

中部地区商品房价格对的脉冲响应也出现了波动，表现为正负交替。随后在第 2 期达到了脉冲响应的峰值，之后就一直下降甚至为负，然后在第 3 期响应值又升为正。在第 5 期之后响应基本趋于稳定，响应值趋于零。

4.2 限购政策对房价的影响分析

1. 限购政策

2008 年金融危机之下，中国楼市出现 10 年来大拐点，量价齐跌。2008 年下半年宏观调控转向，管理层启动大规模刺激楼市消费政策，其他金融政策全线松绑。2008 年底国务院正式确定“四万亿”救市计划，降准、降息、降首付。2009 年楼市复苏，房价持续攀升，且其增长速度明显高于经济、居民收入和消费的增长速度。其中，由于房地产过度投机导致房地产泡沫的形成。

为调节房地产业狂热的发展态势，国家采取了一系列措施对其进行宏观调控，2010 年 1 月 10 日国务院出台国十一条，要求二套房贷款首付不得低于 40%。同年 4 月 15 日，国务院要求对贷款购买第二套住房的家庭，贷款首付款不得低于 50%，贷款利率不得低于基准利率的 1.1 倍。4 月 27 日，国务院发布了《国务院关于坚决遏制部分城市房价过快上涨的通知》，（简称“国十条”），商品住房价格过高、上涨过快、供应紧张的地区，商业银行可根据风险状况，暂停发放购买第三套及以上住房贷。被称为“史上最严厉的调控政策”。随后，各个城市陆续出台

了实施细则，包括北京、上海等直辖市，厦门等计划单列市，省会城市和部分房价上涨过快的二、三线城市。到 2011 年 10 月 31 日，已有 46 个城市推行限购政策。限购城市的名单如下表 4-9：

表 4-9 限购城市名单

限购令城市名单			
华南地区		东北地区	
广东	广州、深圳、佛山、珠海	黑龙江	哈尔滨
福建	福州、厦门	吉林	长春
海南	海口、三亚	辽宁	大连、沈阳
华北地区		中南地区	
北京		湖南	长沙
天津		湖北	武汉
山东	济南、青岛	贵州	贵阳
河北	石家庄	广西	南宁
山西	太原	江西	南昌
河南	郑州	西北地区	
华东地区		陕西	西安
上海		甘肃	兰州
江苏	南京、无锡、苏州、徐州	新疆	乌鲁木齐
安徽	合肥	青海	西宁
浙江	宁波、杭州、温州、金华、绍兴、舟山、台州、衢州	宁夏	银川
西南地区		内蒙	呼和浩特
四川	成都		
云南	昆明		

2011 年，中央政府坚持调控不动摇，要求各地不折不扣地落实房地产调控工作。年初，国务院“新国八条”重拳出击，紧接着，楼市限购、限贷、限价接踵而至。

综上所述，2009 年 12 月至 2011 年 5 月是一个政策密集、房价波动显著的阶段，政府大量使用货币、财政、土地和行政手段干预市场，包括史上最严厉的“限购政策”、“国十一条”、“国十条”、多达 11 次的利率调整和保障房政策等。2012 年 2 月至 2013 年 3 月，财政和货币政策等发挥了作用，房价趋于平稳波动，2013 年 1 月涨至历史新高后，基本保持了相对稳定的状态。上述房地产限购政策一直持续了两三年，直到 2014 年部分城市才逐步放开限购。

要估计限购政策的效应，从政策的因果关系来看，估计方法的选取一方面必须把限购政策效应与其他宏观政策效应区分开来，另一方面必须解决政策变量的

内生性问题。

从逻辑上讲，宏观调控中货币政策的松紧以及房地产市场的其他政策变化，如购房的首付比例的变化以及第二套房贷款利率上浮比例等政策变化都会影响房价。因而，如果利用时间序列数据研究单一城市限购与房价的关系，从逻辑上很难厘清限购政策效应和其他政策效应。

第二，政策变量的内生性问题。实施限购政策的城市多为近些年来房价上涨较快的一、二线城市，一个城市是否限购本身可能与房价上涨之间存在系统相关性，由此可能导致解释变量的内生性问题，这要求我们选择合适的计量方法来解决这一内生性问题。

2. 倍差法

倍差法，又称双重差分模型(difference-in-difference, DID)，近年来多用于计量经济学中对于公共政策或项目实施效果的定量评估。

通常大范围的公共政策有别于普通科研性研究，难以保证对于政策实施组和对照组在样本分配上的完全随机。非随机分配政策实施组和对照组的试验称为自然试验(natural trial)，此类试验存在较显著的特点，即不同组间样本在政策实施前可能存在事前差异，仅通过单一前后对比或横向对比的分析方法会忽略这种差异，继而导致对政策实施效果的有偏估计。DID 模型正是基于自然试验得到的数据，通过建模来有效控制研究对象间的事前差异，将政策影响的真正结果有效分离出来。

在倍差法下，我们将城市分为两组，一组是实施限购的城市，另外一组是未实施限购的城市，限购的政策效果是限购城市和非限购城市房价上涨率均值在实施限购政策前后的差异，即下式中的 $\hat{\beta}^{DID}$ 。

$$\begin{aligned}\hat{\beta}^{DID} &= (\bar{Y}^{treatment, after} - \bar{Y}^{treatment, before}) - (\bar{Y}^{control, after} - \bar{Y}^{control, before}) \\ &= \Delta \bar{Y}^{treatment} - \Delta \bar{Y}^{control}\end{aligned}$$

若限购政策有效，绘出限购城市和非限购城市的房价走势图，则得到的结果应如下图 4-5 所示：

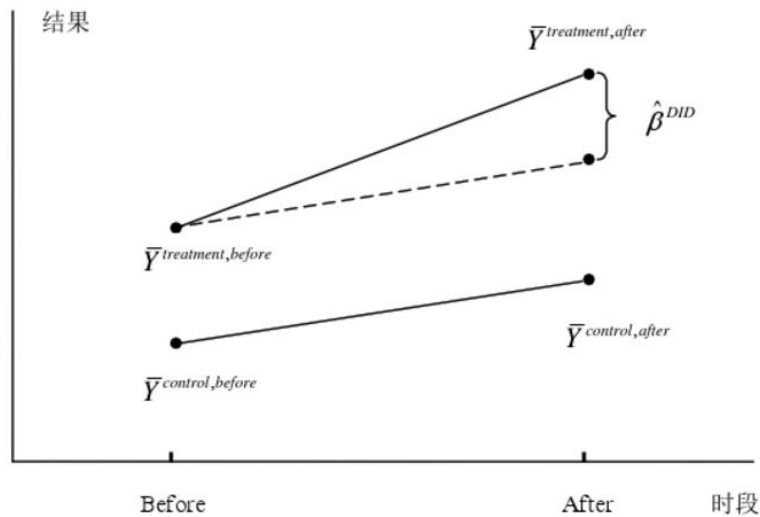


图 4-5 使用倍差法估计限购效应的说明

3. 数据来源

住宅价格数据来自国家统计局。在国家统计局公布月度房价指数的 70 个大中城市中，有 40 个城市实施了限购政策，限购城市包括多数一线、二线城市。限购和未限购城市名单如下表 4-10：

表 4-10 限购城市与未限购城市列表

未限购城市（30 个）	限购城市（40 个）
九江、惠州、牡丹江、赣州、湛江、烟台、韶关、扬州、济宁、桂林、洛阳、北海、平顶山、宜昌、泸州、蚌埠、襄樊、南充、安庆、岳阳、遵义、唐山、秦皇岛、包头、丹东、锦州	上海、南京、武汉、北京、杭州、长沙、天津、宁波、广州、石家庄、合肥、深圳、太原、福州、南宁、呼和浩特、厦门、海口、沈阳、长春、南昌、重庆、大连、济南、成都、青岛、贵阳、西安、兰州、西宁、银川、乌鲁木齐、金华、温州、三亚、无锡、徐州

数据来源：个人整理

我们使用的住宅价格上涨率是月度数据，包括新建住宅同比价格上涨率和二

手住宅同比价格上涨率。70 个大中城市中限购城市政策出台时间跨度是 2010 年 5 月—2011 年 5 月，因而我们选择的样本时间范围是 2010 年 1 月—2012 年 12 月，我们有足够时间跨度的数据来考察限购前与限购后房价变化情况。限购政策细则的数据来自个城市在其官方网站上发布的与限购相关的房地产调控文件，其中部分城市前后发布了两次与限购相关的房地产调控文件。

自 2010 年 5 月北京开始实施限购政策后，我国房地产市场总体上开始降温。图 4-6 是限购城市和非限购城市新建住宅价格上涨率均值的在限购前后的变化趋势，图 4-7 是限购城市和非限购城市二手住宅价格上涨率均值的在限购前后的变化趋势。

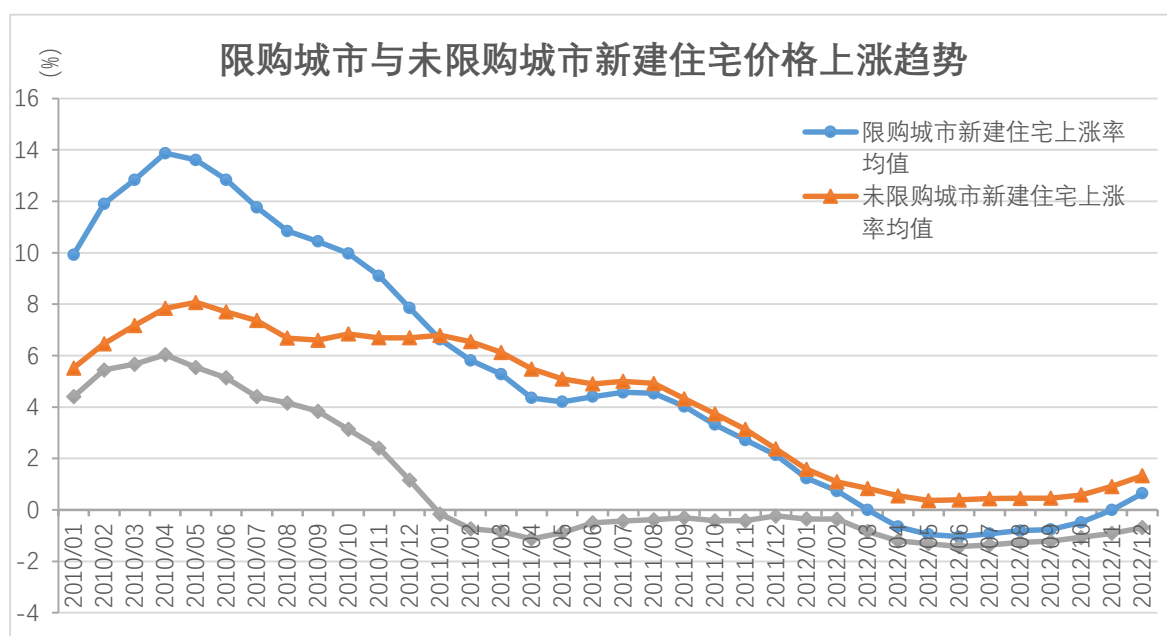


图 4-6 限购城市和未限购城市新建住宅价格上涨率

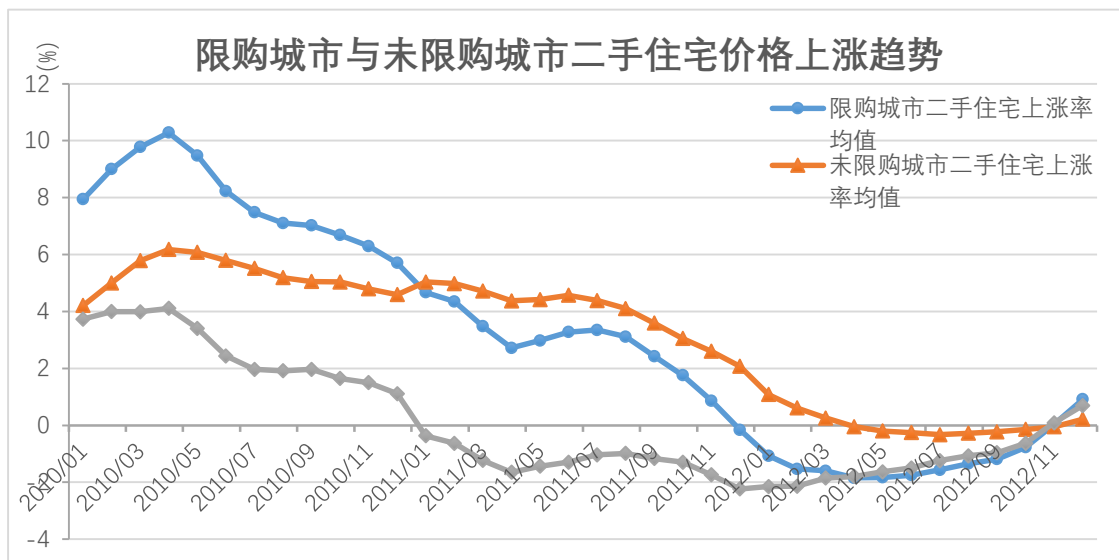


图 4-7 限购城市和未限购城市二手住宅价格上涨率

图 4-6 和图 4-7 的基本趋势是一致的，2010 年 4 月限购城市和非限购城市住宅价格上涨率均值都达到最大，与此同时，两类城市住宅价格上涨率均值差异达到最大。此后，无论是限购城市还是非限购城市，住宅价格上涨率均趋于下降，到 2012 年底，两类城市住宅价格上涨率接近 0%。根据上述的分析可知，这种两类城市相似的住宅价格变动趋势是宏观调控中的非限购政策的效应。

由于每个城市正式实施限购并非同时，从 2010 年 5 月至 2011 年 5 月，限购城市的数目逐渐增多。我们可以明显地观察到，无论是新建住宅还是二手住宅，限购城市和非限购城市住宅价格上涨率均值差异开始缩小。通过图 2，可观察到，在新建住宅中，这一差异在 2011 年 1 月消失，随后限购城市的新建住宅价格上涨率均值开始低于非限购城市。而在图 4-7 中，这一趋势更加明显，在 2011 年 1 月以后，限购城市二手住宅价格上涨率均值明显低于非限购城市，并且随着限购时间的持续这一趋势愈益加强，直至 2012 年初。

2012 年底，趋势开始改变，限购城市新建住宅和二手住宅价格的上涨率均值均重新超过非限购城市均值。而对于二手住宅来说，两类城市价格上涨率变化趋势的改变得更早，在 2012 年 6 月份的时候，限购城市二手住宅价格上涨率与非限购城市二手住宅价格上涨率之差开始减小，并且在 2012 年反超，显示出限购政策对二手住宅价调控的效果开始削弱。这可能也是引发自 2010 年以来第五次调控的原因，即 2013 年 2 月新“国五条”政策的出台，要求限购城市继续严格执行商品住房限购措施。并且从图 4-6 和图 4-7 中，我们可以直观地看到，限购政策似乎取得了图 4-5 的理想效果。表明了限购政策的有效性。

4.3 本章小结

从各个地区来看,对于货币政策与房价之间存在着因果关系,其中较发达地区对于货币政策的相应比较迅速,这主要由于较发达地区,都属于东部沿海地区,经济较为发达,金融系统发展较为完善,房地产商投资和居民购房都较为理性,利用货币政策调控房价具有一定的效果。但是发达地区对于货币政策的响应虽然也是正向的,但是比较微弱。而中等发达地区和欠发达地区对于货币政策的响应则是正负交替的,货币政策对于其影响比较复杂。

对于限购政策来看,通过比较限购城市和非限购城市的房价波动率我们可以看出调控政策是有效果的。

第五章 结论与政策建议

5.1 结论

本文从房产的消费品与投资品两重属性入手,将影响房地产价格的因素分为需求方因素、供给方因素、宏观经济与政策三个方面。

为考察市场本身有哪些因素影响房价,选取 2005-2014 年我国 31 个省市自治区的面板数据,综合运用 R 语言编程、SPSS 和 Eviews 等统计软件对影响房价的因素进行定量分析。结果表明从全国范围内来说,供给和需求是影响房价的重要因素。我们又根据聚类分析结果将全国按照经济发展水平划分为不同区域,分区域建立了多元回归模型。发现影响不同经济发展水平区域房价的因素是不同的,总的来说四个地区的房价都受到需求因素的影响,但是每个地区的经济发展水平不同,房价影响因素也有区别。对于发达地区,所受供给的影响是比较显著的,说明在发达地区,房地产显示出了较强的投资品的属性。对于较发达地区的房价所受货币供给量的影响显著。而在中等发达和欠发达地区,房价受到需求也就是人们的收入水平和经济发展水平的影响比较大。

并且本文采用实证研究的方法着重考察了货币政策和限购政策对于房价的影响。各个地区来看,对于货币政策与房价之间存在着因果关系,其中较发达地区对于货币政策的相应比较迅速,这主要由于较发达地区,都属于东部沿海地区,经济较为发达,金融系统发展较为完善,房地产商投资和居民购房都较为理性,利用货币政策调控房价具有一定的效果。但是发达地区对于货币政策的响应虽然

也是正向的，但是比较微弱。而中等发达地区和欠发达地区对于货币政策的响应则是正负交替的，货币政策对于其影响比较复杂。对于限购政策来看，通过比较限购城市和非限购城市的房价波动率我们可以看出调控政策是有效果的。

5.2 政策建议

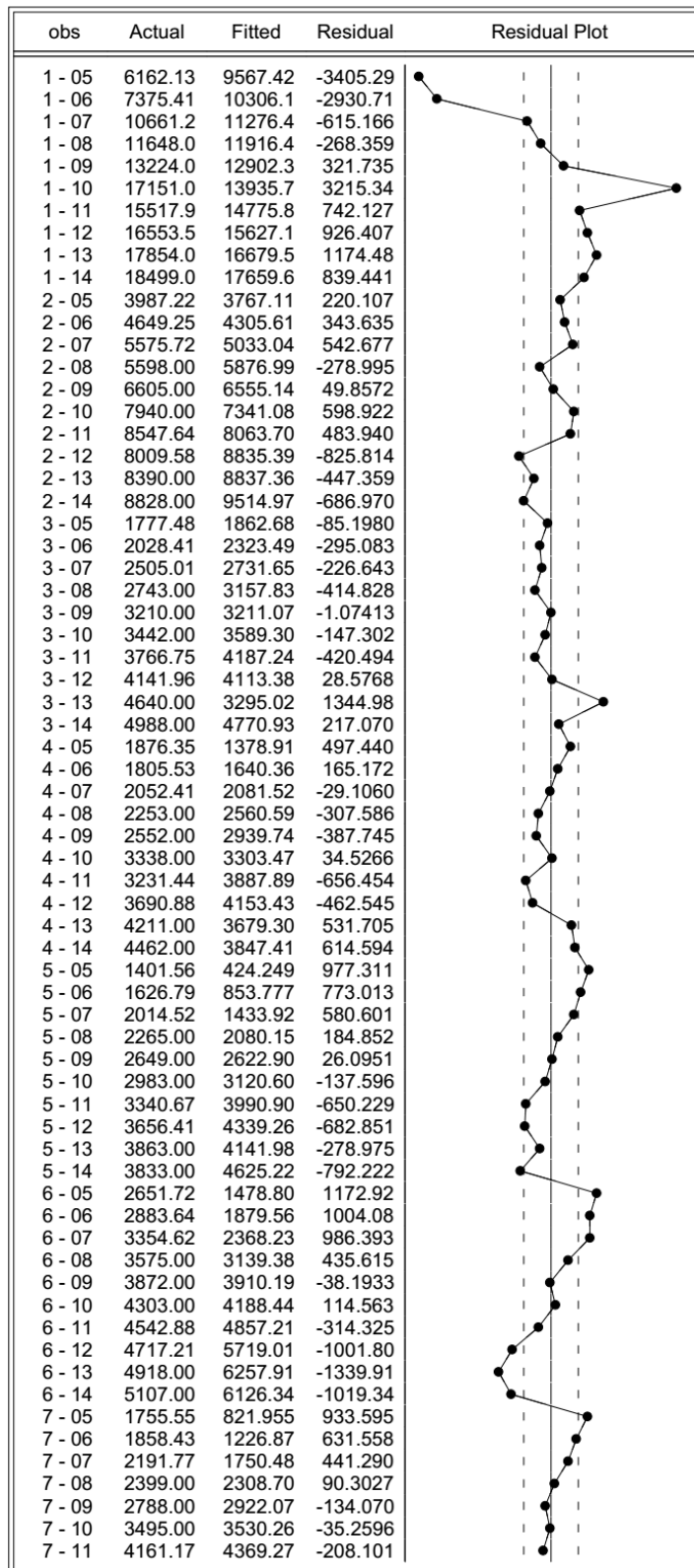
1) 加大房地产的供给量，推进保障性住房及廉租房的建设房价持续上涨的重要原因之一便是，在城镇化的进程阶段，房地产一直处于供不应求的状况。对房价进行调控，最有效的方法不是压制需求，而是增加供给。因此，进一步推进保障性住房及廉租房的建设对于解决中低层人民的住房需求，控制房价的过快上涨具有重要的作用。

2) 货币政策对房价具有一定宏观调控作用，但存在区域差异较大，货币政策传导效果不明显等问题。这源于我国市场机制还不够健全、经济发展不平衡、金融体系不发达，导致货币政策传导机制不够完善，使得货币政策对房价的调控效果不能完全发挥。因此，应根据我国区域发展状况，逐步健全市场机制，针对各地区房价波动差异较大、影响因素不同的特点制定相应政策，避免全国范围内一刀切的情况，以实现有效调控各地区房价。

参考文献

- [1] 张晓峒. EViews 使用指南与案例[M]. 机械工业出版社, 2007.
- [2] 薛毅, 陈立萍. 统计建模与 R 软件[M]. 清华大学出版社, 2007.
- [3] 赵丽丽, 焦继文. 房价影响因素的灰色关联度分析[J]. 统计与决策, 2007(23):74-75.
- [4] 梁云芳, 高铁梅. 中国房地产价格波动区域差异的实证分析[J]. 经济研究, 2007(8):133-142.
- [5] 杨华, 房价分析模型及对策[J]. 武汉轻工大学学报, 2008, 27(1):89-93.
- [6] 曹瑞, 周锋, 欧阳广帅, 等. 基于多项式回归的房价模型分析[J]. 科协论坛, 2010(10):137-138
- [7] 刘传哲, 何凌云. 我国货币政策房地产渠道传导效率检验[J]. 南方金融, 2006(7):5-7.
- [8] 蒋益民, 陈璋. SVAR 模型框架下货币政策区域效应的实证研究:1978~2006[J]. 金融研究, 2009(4):180-195.
- [9] 高云峰, 阮莉莉. 中国货币政策区域效应研究[J]. 投资研究, 2012(4):8-16.
- [10] 赵鹤芹. 从二、三线城市看限购令的房价影响效应——基于南昌市的实证分析[J]. 科技广场, 2013(1):224-227.
- [11] 张德荣, 郑晓婷. “限购令”是抑制房价上涨的有效政策工具吗?——基于 70 个大中城市的实证研究[J]. 数量经济技术经济研究, 2013(11):56-72.

附录 1 附图和附表



obs	Actual	Fitted	Residual	Residual Plot
7 - 12	3875.10	5001.67	-1126.57	
7 - 13	4228.00	4642.85	-414.847	
7 - 14	4810.00	4987.90	-177.899	
8 - 05	1872.80	1269.43	603.374	
8 - 06	2035.17	1669.63	365.536	
8 - 07	2354.24	2126.12	228.120	
8 - 08	2642.00	2508.73	133.269	
8 - 09	3067.00	2932.22	134.779	
8 - 10	3492.00	3378.38	113.624	
8 - 11	3682.78	3953.40	-270.619	
8 - 12	3725.51	4520.19	-794.678	
8 - 13	4435.00	4391.09	43.9051	
8 - 14	4517.00	5074.31	-557.310	
9 - 05	6698.00	7855.73	-1157.73	
9 - 06	7039.00	8554.68	-1515.68	
9 - 07	8253.00	9307.26	-1054.26	
9 - 08	8115.00	10186.5	-2071.46	
9 - 09	12364.0	10978.5	1385.51	
9 - 10	14290.0	12064.1	2225.90	
9 - 11	13565.8	13088.6	477.230	
9 - 12	13869.9	13981.1	-111.178	
9 - 13	16192.0	14870.2	1321.81	
9 - 14	16415.0	15915.1	499.864	
10 - 05	3145.58	2598.33	547.248	
10 - 06	3375.17	3161.34	213.830	
10 - 07	3834.20	3539.39	294.807	
10 - 08	3802.00	4299.08	-497.079	
10 - 09	4805.00	4643.81	161.188	
10 - 10	5592.00	5222.57	369.427	
10 - 11	6145.20	6208.94	-63.7408	
10 - 12	6422.85	6800.99	-378.142	
10 - 13	6650.00	6649.34	0.65804	
10 - 14	6783.00	7431.20	-648.196	
11 - 05	3973.04	5390.14	-1417.10	
11 - 06	4510.12	5768.00	-1257.88	
11 - 07	5623.27	6227.37	-604.096	
11 - 08	6144.00	6727.49	-583.492	
11 - 09	7890.00	7418.08	471.917	
11 - 10	9332.00	8085.96	1246.04	
11 - 11	9801.15	9044.89	756.263	
11 - 12	10679.7	10018.5	661.205	
11 - 13	11016.0	9906.74	1109.26	
11 - 14	10586.0	10968.1	-382.122	
12 - 05	2064.64	2033.18	31.4639	
12 - 06	2152.78	2253.07	-100.286	
12 - 07	2504.78	2580.59	-75.8149	
12 - 08	2808.00	2967.73	-159.729	
12 - 09	3235.00	3269.24	-34.2350	
12 - 10	3899.00	3933.14	-34.1449	
12 - 11	4371.24	4397.56	-26.3166	
12 - 12	4495.12	4706.41	-211.286	
12 - 13	4776.00	4312.82	463.175	
12 - 14	5017.00	4869.83	147.173	
13 - 05	2800.86	3654.64	-853.783	
13 - 06	3655.74	4270.39	-614.653	
13 - 07	4475.95	4681.55	-205.604	
13 - 08	4498.00	5099.22	-601.219	
13 - 09	5366.00	5670.77	-304.771	
13 - 10	6077.00	6472.89	-395.894	
13 - 11	7451.95	7040.04	411.913	
13 - 12	8365.92	7613.21	752.714	
13 - 13	8618.00	7411.25	1206.75	
13 - 14	8843.00	8238.45	604.551	
14 - 05	1335.77	1573.59	-237.823	
14 - 06	1590.60	1918.39	-327.794	
14 - 07	1998.26	2373.32	-375.056	
14 - 08	2022.00	2576.21	-554.207	

obs	Actual	Fitted	Residual	Residual Plot
14 - 09	2517.00	2924.36	-407.362	
14 - 10	2959.00	3353.97	-394.973	
14 - 11	3822.00	4061.80	-239.803	
14 - 12	4381.18	4164.50	216.684	
14 - 13	4905.00	3626.70	1278.30	
14 - 14	4971.00	3928.97	1042.03	
15 - 05	2294.63	1424.12	870.513	
15 - 06	2399.94	2213.12	186.821	
15 - 07	2799.25	2797.61	1.63889	
15 - 08	2851.00	3381.84	-530.845	
15 - 09	3390.00	3745.44	-355.442	
15 - 10	3809.00	4079.01	-270.008	
15 - 11	4298.78	4604.91	-306.133	
15 - 12	4556.63	5010.64	-454.010	
15 - 13	4797.00	4222.45	574.546	
15 - 14	5029.00	4746.08	282.919	
16 - 05	1658.68	1911.11	-252.426	
16 - 06	1842.63	2152.20	-309.573	
16 - 07	2081.22	2150.22	-68.9986	
16 - 08	2138.00	2368.17	-230.169	
16 - 09	2501.00	2369.69	131.314	
16 - 10	2856.00	2808.68	47.3241	
16 - 11	3123.15	3390.89	-267.745	
16 - 12	3511.26	3650.94	-139.680	
16 - 13	3835.00	2730.53	1104.47	
16 - 14	3909.00	3923.51	-14.5133	
17 - 05	2163.65	1587.06	576.585	
17 - 06	2422.14	2081.67	340.468	
17 - 07	2937.45	2610.12	327.334	
17 - 08	2898.00	3000.55	-102.554	
17 - 09	3413.00	3571.51	-158.505	
17 - 10	3506.00	4039.98	-533.983	
17 - 11	4142.13	4507.73	-365.604	
17 - 12	4668.00	4941.65	-273.648	
17 - 13	4847.00	4555.21	291.790	
17 - 14	5085.00	5186.88	-101.882	
18 - 05	1404.59	972.389	432.201	
18 - 06	1655.15	1298.71	356.435	
18 - 07	2067.96	1716.52	351.442	
18 - 08	2113.00	2059.32	53.6841	
18 - 09	2532.00	2435.36	96.6444	
18 - 10	3014.00	2931.72	82.2810	
18 - 11	3523.65	3716.47	-192.821	
18 - 12	3669.63	4318.36	-648.726	
18 - 13	3908.00	3883.74	24.2622	
18 - 14	3830.00	4385.40	-555.403	
19 - 05	4149.26	3653.88	495.382	
19 - 06	4588.62	4212.84	375.780	
19 - 07	5682.14	4953.08	729.058	
19 - 08	5723.00	5661.04	61.9636	
19 - 09	6360.00	6190.00	170.004	
19 - 10	7004.00	6770.43	233.572	
19 - 11	7560.85	7847.73	-286.884	
19 - 12	7667.89	8802.65	-1134.76	
19 - 13	8466.00	8399.11	66.8880	
19 - 14	8526.00	9237.00	-711.004	
20 - 05	1824.51	1374.14	450.371	
20 - 06	1972.86	1807.26	165.603	
20 - 07	2385.52	2275.90	109.615	
20 - 08	2634.00	2631.81	2.18887	
20 - 09	3133.00	3045.02	87.9775	
20 - 10	3382.00	3606.50	-224.497	
20 - 11	3553.96	4143.30	-589.338	
20 - 12	3909.83	4611.46	-701.626	
20 - 13	4219.00	3612.66	606.340	
20 - 14	4442.00	4348.64	93.3648	
21 - 05	2855.27	4360.11	-1504.84	

obs	Actual	Fitted	Residual	Residual Plot
21 - 06	3734.67	4776.33	-1041.66	
21 - 07	4094.58	5247.46	-1152.88	
21 - 08	5441.00	5742.99	-301.987	
21 - 09	6291.00	6250.42	40.5799	
21 - 10	8800.00	6935.58	1864.42	
21 - 11	9083.14	7624.67	1458.47	
21 - 12	7811.26	8295.91	-484.647	
21 - 13	8633.00	8142.36	490.643	
21 - 14	9262.00	8630.10	631.897	
22 - 05	1900.66	1570.17	330.486	
22 - 06	2081.31	1904.69	176.622	
22 - 07	2588.22	2372.13	216.086	
22 - 08	2640.00	2855.60	-215.599	
22 - 09	3266.00	3401.30	-135.298	
22 - 10	4040.00	3945.05	94.9519	
22 - 11	4492.30	3824.61	667.692	
22 - 12	4804.80	5487.05	-682.252	
22 - 13	5239.00	4929.31	309.688	
22 - 14	5094.00	5856.38	-762.377	
23 - 05	1687.87	2451.17	-763.296	
23 - 06	2122.56	2612.29	-489.735	
23 - 07	2752.81	3065.74	-312.933	
23 - 08	3067.00	3191.21	-124.207	
23 - 09	3434.00	2385.00	1049.00	
23 - 10	3985.00	3795.46	189.539	
23 - 11	4595.13	4490.28	104.852	
23 - 12	4959.19	5138.28	-179.088	
23 - 13	5086.00	4552.12	533.876	
23 - 14	5092.00	5100.01	-8.01181	
24 - 05	1307.56	1115.78	191.781	
24 - 06	1583.52	1462.50	121.016	
24 - 07	1899.36	1906.40	-7.03853	
24 - 08	2122.00	2258.77	-136.765	
24 - 09	2642.00	2642.77	-0.76787	
24 - 10	3142.00	2915.33	226.670	
24 - 11	3489.71	3573.01	-83.2974	
24 - 12	3695.36	4172.42	-477.061	
24 - 13	3735.00	3506.85	228.154	
24 - 14	3694.00	3756.69	-62.6903	
25 - 05	2000.84	1357.05	643.786	
25 - 06	2190.65	1775.70	414.947	
25 - 07	2295.74	2047.09	248.649	
25 - 08	2441.00	2452.53	-11.5292	
25 - 09	2723.00	2958.76	-235.763	
25 - 10	2893.00	3380.26	-487.259	
25 - 11	3388.41	4078.70	-690.286	
25 - 12	3861.01	4468.16	-607.150	
25 - 13	4176.00	3766.34	409.659	
25 - 14	4451.00	4136.05	314.946	
26 - 05	1506.08	1206.54	299.540	
26 - 06	1686.85	1398.28	288.566	
26 - 07	2662.32	1941.54	720.775	
26 - 08	3103.00	2412.50	690.499	
26 - 09	2392.00	2922.27	-530.273	
26 - 10	2761.00	3441.95	-680.950	
26 - 11	3311.55	3938.44	-626.894	
26 - 12	2982.19	4472.19	-1490.00	
26 - 13	3883.00	3798.68	84.3163	
26 - 14	5323.00	4078.58	1244.42	
27 - 05	1930.33	1834.36	95.9748	
27 - 06	2297.07	2180.70	116.368	
27 - 07	2487.38	2576.58	-89.1956	
27 - 08	2821.00	3079.56	-258.563	
27 - 09	3113.00	3259.31	-146.311	
27 - 10	3668.00	3863.68	-195.678	
27 - 11	4705.28	4410.05	295.225	
27 - 12	4803.05	4760.70	42.3550	

obs	Actual	Fitted	Residual	Residual Plot
27 - 13	4991.00	4560.66	430.344	
27 - 14	4823.00	5113.52	-290.518	
28 - 05	1739.07	1242.02	497.054	
28 - 06	1703.09	1601.35	101.739	
28 - 07	2146.03	1968.15	177.882	
28 - 08	1851.00	2146.24	-295.243	
28 - 09	2396.00	2388.45	7.54905	
28 - 10	2938.00	3011.37	-73.3700	
28 - 11	3130.04	3560.47	-430.427	
28 - 12	3376.08	4061.02	-684.942	
28 - 13	3684.00	3485.64	198.361	
28 - 14	4234.00	3732.60	501.397	
29 - 05	1681.03	1025.72	655.313	
29 - 06	1840.50	1443.94	396.559	
29 - 07	2205.75	1881.68	324.066	
29 - 08	2384.00	2356.76	27.2449	
29 - 09	2442.00	2794.45	-352.445	
29 - 10	2894.00	3123.34	-229.341	
29 - 11	3090.01	3550.69	-460.677	
29 - 12	3692.21	4112.96	-420.746	
29 - 13	3957.00	3890.04	66.9556	
29 - 14	4294.00	4300.93	-6.92891	
30 - 05	1764.60	840.004	924.596	
30 - 06	1869.27	1214.70	654.573	
30 - 07	1957.71	1714.24	243.471	
30 - 08	2215.00	2278.87	-63.8670	
30 - 09	2824.00	2785.27	38.7341	
30 - 10	3107.00	3261.33	-154.331	
30 - 11	3389.46	3773.86	-384.396	
30 - 12	3620.77	4341.63	-720.858	
30 - 13	3917.00	3890.48	26.5177	
30 - 14	3747.00	4311.44	-564.439	
31 - 05	1509.01	545.206	963.804	
31 - 06	1684.06	1002.71	681.350	
31 - 07	1959.99	1558.89	401.095	
31 - 08	2100.00	2081.19	18.8120	
31 - 09	2466.00	2546.68	-80.6819	
31 - 10	2872.00	3332.48	-460.484	
31 - 11	3286.95	3758.04	-471.088	
31 - 12	3593.82	4212.40	-618.579	
31 - 13	3949.00	3917.96	31.0380	
31 - 14	4057.00	4522.27	-465.266	

附录 2 代码

```
#更改工作空间
setwd("E:\\Tsinghua\\应用统计学\\大作业\\应用统计学数据\\3.4 投资额")
#读入数据
price <- read.csv("全国房价数据.csv",header=T)
colnames(price)<-c("date",names)
inve_s <- read.csv("全国商品住宅开发投资额.csv",header= T)
colnames(inve_s)<-c("date",names)
invc_s <- read.csv("全国商品住宅开发投资额增速.csv",header= T)
colnames(invc_s)<-c("date",names)
houscons <- read.csv("全国商品住宅施工面积.csv",header= T)
colnames(houscons)<-c("date",names)
houscomp <- read.csv("全国商品住宅竣工面积.csv",header= T)
colnames(houscomp)<-c("date",names)
landex <- read.csv("全国土地购置费.csv",header= T)
colnames(landex)<-c("date",names)
landexad <- read.csv("全国土地购置费增速.csv",header= T)
colnames(landexad)<-c("date",names)
#根据列合并数据
con<-
cbind(price[2:129,2],inve_s[4:131,2],invc_s[4:131,2],houscons[4:131,2],houscomp[4:
131,2],landex[4:131,2],landexad[4:131,2])
colnames(con)<-
c("price","inve_s","invc_s","houscons","houscomp","landex","landexad")
#计算相关系数矩阵
round(cor(con),digits=2)
#绘制散点图,房价与施工面积
#jpeg("施工面积 VS 房价.jpeg",width=150,height=150)
plot(price[2:129,2],houscons[4:131,2],main="全国施工面积 VS 房价")
#更改工作空间
setwd("E:\\Tsinghua\\应用统计学\\大作业\\应用统计学数据\\3.6 区域差异")
```



```

#读入数据
region <- read.csv("3.6 区域差异-正规化.csv",header=T)
attach(region)
re <- cbind(工业化水平,城镇化率,人均 GDP1)
rownames(re)<-c("北京","天津","河北","山西","内蒙古","辽宁","吉林","黑龙江","
上海","江苏","浙江省",
"安徽","福建","江西","山东","河南","湖北","湖南","广东","广西","海南","重庆","
四川","贵州",
"云南","西藏","陕西","甘肃","青海","宁夏","新疆")
library(cluster)
km <- kmeans(re,centers=4,algorithm="Forgy")    #建立 k-means 聚类
#algorithm="MacQueen","Forgy","Hartigan-Wong","Lloyd"
clusplot(re,km$cluster,color=T,shade=T,labels=2,lines=0)    #聚类结果,画图显示
km$centers    #各类中心
sort(km$cluster)    #对分类情况排序
#更改工作空间
setwd("E:\\Tsinghua\\应用统计学\\大作业\\应用统计学数据\\3.1 商品住宅销售价格波动")
#读入数据
data <- read.csv("2005-2015 各地区房价.csv",header=T)
kruskal.test(data$平均,data$分类)
pr1<-c(14523.22935,6981.880782,12394.76236)
pr2<-
c(2971.472825,4131.897959,5209.736386,8152.885087,6180.722775,3761.487966,6
824.496886)
pr3<-
c(3495.449748,3082.966385,3520.9761,3438.57998,3691.186311,3187.771403,2919.
351587,3846.89392,
2929.437859,3272.034923,3745.729318,3796.561637,3736.782555,2978.303367,296
2.652004,2853.660954,7127.777941)
pr4<-c(2816.066166,3161.565962,2895.772916,2881.115415)
wilcox.test(pr1,pr2,alternative="less")

```