

参赛密码 _____
(由组委会填写)

全国第八届研究生数学建模竞赛

学 校	中南财经政法大学
参赛队号	11943008
队员姓名	1.贾晓惠
	2.宋 迎
	3.姚 琼

参赛密码 _____
(由组委会填写)

全国第八届研究生数学建模竞赛



题 目 我国房地产行业问题研究

摘 要

针对房地产行业问题，本文将其主要归纳为四个方面来进行研究：

问题 1，如何建立房地产行业需求和供给模型来反映该行业市场供需状况；

问题 2，如何建立有效的房地产行业定价模型对未来房价走向进行预测；

问题 3，如何建立模型对房地产行业与其他国民经济行业关联度以及其自身发展态势进行研究；

问题 4，如何建立房地产行业可持续发展模型对其自身发展进行跟踪研究。

为了解决上述问题，本文首先考虑我国房地产市场存在地区差异性，利用地理加权回归方法分别建立需求与供给模型，在其求解基础上，基于 VAR 模型分区域建立了房地产定价模型，用于对房地产市场定价进行合理预测。然后，分别从国民经济核算角度和计量经济学角度对房地产行业关联度和发展态势建立了投入产出模型和固定效应的变系数面板数据模型。最后，本文基于熵值法建立综合评价模型对天津市房地产行业的可持续发展性进行了研究和探讨。

本文的亮点和创新之处在于：第一，充分考虑我国房地产市场存在区域发展不平衡性，建立了基于地理加权回归方法的需求和供给模型，此种空间计量方法较为新颖，且模型结果具有较强的政策参考性；第二，在研究房地产行业关联度和发展态势时，我们采用了国民经济核算和计量经济学两套方法，可以进行对比分析；第三，采用基于熵值法的综合评价模型来评价房地产行业的可持续发展性，以熵值法赋权避免了主观性，模型结果更加可靠。

针对问题 1，首先对影响房地产行业需求和供给的变量进行空间异质性和空间自相关性检验，在验证了空间自相关性存在的基础上，利用地理加权回归（GWR）方法建立了需求模型和供给模型，运用 Matlab 编程求解，此方法分省

份研究房地产市场的需求和供给，充分考虑了房地产市场地区差异性，模型结果显示我国有必要实行差异化的贷款利率政策，进一步完善土地交易市场机制，培养理性投资人等。然后，利用需求和供给模型求解的结果，我们对房地产行业进行分层聚类分析，进而将我国房地产行业划分为东、中、西三类，为下文建立分区域的房地产定价模型做了有力的铺垫。

针对问题 2，我们基于问题 1 对房地产市场的划分结果，采用了计量经济学中的向量自回归模型分东、中、西三个区域进行建模和预测，其中引入了脉冲响应分析和方差分解，不但考虑了各变量整体对房价的综合影响，更通过脉冲图形清晰地反映了单个变量变动对房价的作用，以西部地区的货币供应量对房价的影响为例，其滞后一期的货币供应量发生 1 个单位的正向变动，房价发生 0.85 个单位的正向变动。

针对问题 3，分别采用了投入产出模型和固定效应的变系数面板数据模型来研究房地产行业关联度和发展态势，通过对比分析二者的模型结果，我们发现，房地产行业与金融业、租赁和商务服务业、建筑业存在显著关联度，如根据固定效应的变系数面板数据模型结果，可知房地产行业增加值每提高 1%，将带动金融业增加值的 1.35% 的有利变动。

针对问题 4，本文富有创新性的采用基于熵值法的综合评价方法对天津市的房地产行业的可持续发展进行研究，研究结果表明 2005 以来，由于美国次贷危机对该行业造成的冲击，房地产行业分别于 2006 年出现了一次向下运行的拐点，而在政府积极宏观调控政策的作用下，房地产行业于 2008 年出现了一次向上运行的拐点，这也告诉我们，建立健全房地产行业的对抗外部突发因素干扰的机制对于房地产行业的可持续发展显得十分必要。

关键字 地理加权回归 VAR 模型 投入产出模型 固定效应模型 熵值法

目 录

一、问题的重述.....	4
二、问题的分析.....	4
2.1 问题 1 的分析.....	5
2.2 问题 2 的分析.....	5
2.3 问题 3 的分析.....	5
2.4 问题 4 的分析.....	6
三、模型假设.....	6
四、符号说明.....	6
五、模型和求解.....	7
5.1 房地产行业需求、供给模型和求解.....	7
5.1.1 需求、供给模型的建立.....	7
5.1.2 需求、供给模型的求解.....	9
5.1.3 基于需求、供给模型的房地产市场区域聚类分析.....	25
5.1.4 需求、供给模型结论.....	26
5.2 房价模型与求解.....	27
5.2.1 房价模型的建立.....	27
5.2.2 房价模型的分析与求解（以西部地区为例）.....	30
5.2.3 房价预测模型（以西部地区为例）.....	34
5.2.4 东部和中部地区的房价预测模型.....	35
5.3 房地产行业关联度与发展态势投入产出模型与求解.....	36
5.3.1 投入产出模型理论分析.....	36
5.3.2 投入产出模型求解.....	38
5.4 房地产行业关联度与发展态势空间面板数据模型与求解.....	41
5.4.1 面板数据模型的理论分析.....	41
5.4.2 模型变量选择与数据处理.....	43
5.4.3 基于面板数据模型的房地产行业与其他行业相关性效应实证分析.....	44
5.5 基于熵值法的房地产可持续发展模型.....	49
5.5.1 基于熵值法的房地产行业可持续发展模型理论分析.....	49
5.5.2 基于熵值法的房地产行业可持续发展模型求解.....	52
六、模型的评价与推广.....	54
七、参考文献.....	55

一、问题的重述

房地产行业既是国民经济的支柱产业之一，又是与人民生活密切相关的行业之一，同时自身也是一个庞大的系统，该系统的状态和发展对国民经济的整个态势和全国人民的生活水平影响很大。近年来，我国房地产业发展迅速，不仅为整个国民经济的发展做出了贡献，而且为改善我国百姓居住条件发挥了决定性的作用，但同时房地产业也面临较为严峻的问题和挑战，引起诸多争议，各方都坚持自己的观点，然而多是从政策层面、心理层面和资金层面等因素来考虑，定性分析多于定量分析。显然，从系统的高度认清当前房地产行业的态势，从定量角度把握各指标之间的数量关系，依据较为准确的预见对房地产行业进行有效地调控，深刻认识房地产行业的经济规律进而实现可持续发展是解决问题的有效途径，因此，通过建立数学模型研究我国房地产问题是一个值得探索的方向。

在建立数学模型的基础上进行分析，量化研究该行业当前的态势、未来的趋势，模拟房地产行业经济调控策略的成效，希望在深化认识上取得进步，产生若干结论和观点。由于对房地产问题已经有许多研究成果和讨论材料，本文在引用其他人的成果和数据，尤其对于定量分析的成果时，都将注明参考文献及其出处。

研究房地产问题并不需要很多、很深的专业知识，问题也不难理解，通过对目前房地产行业现状的观察和研究，本文自主地提出自己希望解决的房地产中的问题，并建立相应的数学模型予以解决，所建的每个模型力求系统、深入，自成兼容系统，数据可靠，结论和观点有较多的数据支撑，同时，具有较强的说服力和实际应用价值。

在本文中，我们将利用附录中提供的及自己查找到的资料和数据建立房地产行业的数学模型，主要包括：

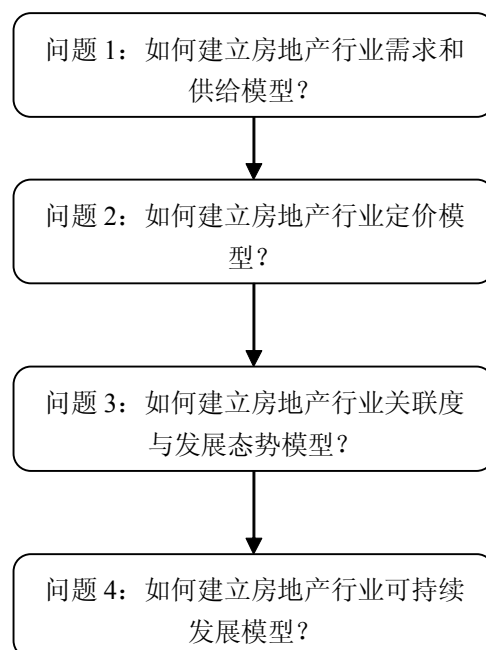
1. 住房需求模型；
2. 住房供给模型；
3. 房价模型；
4. 房地产行业与国民经济其他行业关系模型；
5. 对我国房地产行业态势分析模型；
6. 房地产行业可持续发展模型。

二、问题的分析

房地产业是目前我国大力支持和频繁进行宏观调控的行业，是关系到人民生活水平和质量的重要行业，从其自身发展特点来看，它是基础的、具有带动性和风险性的行业，1998年以来，我国的住房体制日趋成熟且不断完善，银行的信贷业务迅速发展，为开发商和消费者提供了巨大的便利，但是，自从美国次贷危机爆发后，我国对房地产行业的政策放宽，加之各地财政收入对房地产行业的过分依赖，使近年来房价的不断攀升，房地产“泡沫”现象的日益凸出，针对房地产“泡沫”问题，我国政府分别在2010年末和2011年上半年对房地产业的进行了频繁政策调控，如去年年末和今年年初各地的“限购令”和“第三套房贷”政策，以及多次的银行贷款利率和存款准备金率调整等，基于此，越来越多的学者力图有效的对房地产行业的相关问题进行研究，可见建立确定的、较有参考价

值的模型有显著的现实意义。

基于房地产行业问题的日益紧迫感，本文将采用全新的视角和方法来建立有关房地产行业的供给、需求以及定价模型，进而从宏观层面对房地产行业态势进行分析，并研究房地产行业与国民经济其他行业的关联、探讨房地产行业自身的可持续性发展模式，旨在为相关部门制定行业政策提供有效、全新的思路。



2.1 问题 1 的分析

对于问题 1，我们需要针对房地产行业分别建立需求和供给模型，根据经济学知识中的供需定理以及均衡理论，我们能够分别找到可能对房地产行业需求和供给产生影响的变量，然后通过建立有效的计量经济模型来量化需求模型和供给模型，并结合经济学知识对模型结果进行合理的分析。

2.2 问题 2 的分析

房价模型主要涉及定价及预测问题。这里我们先要确定影响房价的主要因素，建立相应的计量模型。再在前面区域划分的基础上，分别确定能够体现地区性差异的房价预测模型。由于我国对房价的调控主要是政策调控，为此要尽可能在模型中引入相关的政策变量。

2.3 问题 3 的分析

在问题 1 和问题 2 对房地产行业自身发展状况有了一定的认知基础上，问题 3 要求我们从宏观层面对房地产业发展态势进行分析，并同时研究该行业与国民经济其他行业的关联性，欲解决此问题，我们考虑从两个方面着手，一方面从国民经济核算角度入手建立投入产出模型，另一方面从计量角度入手建立时间序列或者空间面板数据模型。

2.4 问题 4 的分析

可持续发展模型要能体现房地产行业与经济、社会以及其它行业共同发展、协调发展的特点。为此，我们要对经济和社会中与房地产行业密切相关的因素的发展状况进行综合研究，力求能够建立一个评价体系，根据评价的结果分析不同时间段房地产行业的可持续发展性。

三、模型假设

- 假设一：统计年鉴所得数据真实可靠，能够反映指标所代表的真实含义。
 假设二：国民经济一个部门只生产一种产品，而且只采用一种技术生产，同时，一种产品只由一个部门生产。
 假设三：产品生产中各投入要素间有固定比例，即投入要素增减均呈现统一比例。
 假设四：投入产出模型假设直接消耗系数在一个周期内是不变的。
 假设五：各部门的生产经营活动从生产要素投入到产出的分配与使用均在一个周期内完成。

四、符号说明

模型类别	符号	符号说明（单位）
住房需求模型	<i>SAREA</i>	房屋销售面积（万平方米）
	<i>FPRICE</i>	房屋销售价格（元/平方米）
	<i>INCOME</i>	年可支配收入（元）
	<i>RATE</i>	利率（%）
	<i>STRU</i>	人口结构（%）
	<i>GRATE</i>	经济增长率（%）
住房供给模型	<i>CAREA</i>	房屋竣工面积（万平方米）
	<i>FPRICE</i>	房屋销售价格（元/平方米）
	<i>TPRICE</i>	土地交易价格（元/平方米）
	<i>RATE</i>	利率（%）
	<i>GRATE</i>	经济增长率（%）
房价模型	<i>PRICE</i>	房屋销售价格（元/平方米）
	<i>SAREA</i>	房屋销售面积（万平方米）
	<i>CAREA</i>	房屋竣工面积（万平方米）
	<i>DRR</i>	存款准备金率（%）
	<i>M2</i>	广义货币供应量（亿元）
	<i>LIR</i>	一年期贷款利率（%）
房地产行业关联度模型	<i>FAV</i>	房地产业生产总值增加值（亿元）
	<i>OAV</i>	其他行业生产总值增加值（亿元）

五、模型和求解

5.1 房地产行业需求、供给模型和求解

通过对房地产行业相关文献资料的研究,我们发现,相关专家和学者建立房地产行业的需求和供给建模时,大致可以分为如下几类:一,未考虑到地区差异性,将我国房地产市场视为一个整体进行截面或者时序研究,忽略了地区特性;二,基于空间面板数据进行建模,但是大部分都假设截距项系数随着地理空间位置的不同而不同,假定斜率系数都是相同的;三,考虑了地区性发展不平衡和联动性特点,然而在对地区进行区分时,仅主观的根据通常意义上的东部、中部和西部地区的方式进行划分,缺乏量化理论支持。在前人研究和探讨的基础上,本文基于一个全新的视角和方法对房地产的需求和供给建立模型,即主要采用空间计量的分析方法,通过地理加权回归方法(*GWR*)建立基于全国30^①个省市的房地产需求和供给模型,并在此模型结果的基础上,通过聚类分析方法对我国房地产行业进行区域划分,为下文的房地产分区域定价模型做铺垫。

本部分在建立房地产行业需求和供给模型时,首先利用 GeoDa 软件对各变量的空间相关性进行检验,验证空间异质性的存在,在此基础上,采用地理加权回归方法建立模型并进行实证分析,地理加权回归模型的实现是通过 Matlab 软件编程实现的,相关的程序见附录一,之后,利用房地产需求和供给模型的结果对全国房地产进行聚类分析,此过程是利用 SPSS 软件实现的,聚类结果见附录四。

5.1.1 需求、供给模型的建立

1. 地理加权回归模型理论分析

下面,将简要介绍地理加权回归模型的相关理论,地理加权回归模型主要适用于存在空间异质性的截面数据,是普通线性回归模型的扩展,对比全局回归模型和地理加权回归模型(如下),可以发现,在*GWR*模型中,回归系数的估计不是利用全局信息统一进行,而是利用相互临近区域的信息进行局部回归系数的估计,随着空间地理位置的改变,局部回归系数会随之而改变。

$$y_i = \beta + \sum_{j=1}^k \beta_j x_{ij} + \varepsilon_i \quad (1)$$

$$y_i = \beta_0(u_i, v_i) + \sum_{j=1}^k \beta_j(u_i, v_i) x_{ij} + \varepsilon_i \quad (2)$$

(2)式中, $\beta_j(u_i, v_i)$ 表示与各个截面单位相联系的待估参数,是地理位置或者截面单位的 $k+1$ 元函数, *GWR*模型可以对每个观测值估计出一个含有 $k+1$ 个值的参数向量, ε_i 是满足零均值、同方差、相互独立假定的随机误差项。

对于 $\beta_j(u_i, v_i)$ 的估计,本文根据 Tibshirani 和 Hastie 提出的局域求解原理,即

^① 由于西藏地区房地产业并不发达,并且缺乏相关数据支持,故本文在建立模型时没有考虑西藏地区。

对于第 i 个空间位置 $(u_i, v_i), i=1, 2, \dots, n$ ，可以任取一个与其位置向邻近的空间位置 (u_0, v_0) ，从而构造一个回归模型：

$$y_i = \gamma_0 + \sum_{j=1}^k \gamma_j x_{ij} + \varepsilon_i \quad (3)$$

(3) 式中 γ_j 为 GWR 模型中 $\beta_j(u_i, v_i)$ 的近似值，然后通过采用加权最小二乘法对 γ_j 的值进行估计，即使得 $\sum_{i=1}^n W(d_{0i}) \left(y_i - \gamma_0 - \sum_{j=1}^k \gamma_j x_{ij} \right)^2$ ① 最小化，从而得到 $\beta_j(u_i, v_i)$ 的估计值 $\hat{\gamma}_j$ ， $\hat{\gamma}_j = (X'W^2X)^{-1}(X'W^2Y)$ 。

关于权数矩阵 W 的计算，常用的方法有高斯距离权值、指数距离权值以及三次方距离权值，本文采用的是高斯距离权值，即 $W_{ij} = \Phi(d_{ij} / \sigma\theta)$ ， d_{ij} 为第 i 个位置和第 j 个位置的距离，在本文中即第 i 个省与第 j 个省的省会城市的空间地理距离，此空间地理距离可以根据各省会城市的经纬度计算而来， Φ 为标准正态分布密度函数， σ 为距离的标准差， θ 为窗宽。在权数矩阵的计算中， d 和 θ 的选择对模型的估计的准确性具有重要意义， d 的值越大，局域模型估计的结果越趋近于全局估计的解； θ 越大，当其趋于无穷大时，任意两个位置的权重将趋于一致，此时的 *GWR* 模型和 *OLS* 模型一致，反之，当 θ 很小的时候，参数的估计越依赖于邻近的观测值，对于 θ 的选择，可以采用最小二乘法进行，即令

$$CV = \sum_{i=1}^n [y_i - \hat{y}_i(\theta)]^2 \rightarrow 0。$$

2. 模型变量选择 and 数据处理

通过阅读房地产行业需求和供给的相关文献，结合经济学的市场均衡理论知识，考虑题目中所给出的影响房地产行业的一般因素和区域因素，本文最终选择的影响房地产供求的指标涵盖了经济因素、社会因素以及政策因素，因为过多的加入指标可以提高模型的精确度，但数据的收集和模型的显著性等可能会带来复杂的问题，为此我们对每个层面的因素仅选择了具有代表性的变量指标，下面，本文就将影响房地产行业需求或者供给的各变量加以阐述和说明。

(1) 需求模型变量和数据处理

在需求模型中，本文采用房地产开发企业商品房屋销售面积来表示住房的需求量，也即需求模型的被解释变量，解释变量可以分为经济、社会和政策三个层面，经济层面的解释变量为商品住房销售价格、各地区经济增长率以及居民年可支配收入；社会因素为城镇人口数/总人口数，此变量可以反映各地区的城市化水平；政策因素（宏观调控因素）为一年期贷款利率。本文建立模型的时间截点为 2009 年，其中，由于《统计年鉴》中仅有各省销售价格指数，没有提供各省

① 式中， d_{0i} 为位置 (u_0, v_0) 和 (u_i, v_i) 的空间距离， $W(d_{0i})$ 为空间权值。

销售价格数据，为此我们用商品房的销售额除以商品房的销售面积来计算商品住房销售价格；经济增长率采用环比增长率 $= (GDP_{2009} - GDP_{2008}) / GDP_{2008}$ ，并使用 2009 年的消费者价格指数进行调整。

(2) 供给模型变量和数据处理

在供给模型中，本文采用房地产开发企业商品房屋竣工面积来表示住房的供给量，也即供给模型的被解释变量，解释变量包括与影响需求重叠的商品房销售价格、经济增长率以及一年期贷款利率因素，并考虑到建筑成本的问题，添加了土地交易价格这一经济成本指标。

本文各变量数据主要来自《中国统计年鉴 2010》以及巨灵金融平台的宏观行业数据库，对于个别数据的缺失时，采用总体均值替代修匀；同时，为了消除建立模型过程中可能存在的异方差现象，对商品房屋销售面积、商品住房销售价格、居民年可支配收入、商品房屋竣工面积、土地交易价格进行了对数化处理。

综上，本文采用地理加权回归模型建立的需求和供给模型如下：

$$\begin{aligned} \text{需求模型 } sarea_i = & \beta_0(u_i, v_i) + \beta_1(u_i, v_i)rate_{i1} + \beta_2(u_i, v_i)fprice_{i2} + \beta_3(u_i, v_i)grate_{i3} \\ & + \beta_4(u_i, v_i)stru_{i4} + \beta_5(u_i, v_i)income_{i5} + \varepsilon_i \end{aligned} \quad (4)$$

$$\begin{aligned} \text{供给模型 } carea_i = & \beta_0(u_i, v_i) + \beta_1(u_i, v_i)grate_{i1} + \beta_2(u_i, v_i)tprice_{i2} + \beta_3(u_i, v_i)fprice_{i3} \\ & + \beta_4(u_i, v_i)rate_{i4} + \varepsilon_i \end{aligned} \quad (5)$$

5.1.2 需求、供给模型的求解

1. 需求模型的求解

(1) 基于 GeoDa 软件的需求模型各变量空间异质性检验

为了说明建立地理加权回归模型的必要性，在建模之前，我们有必要对各变量进行空间异质性和自相关性的检验，以对这些变量的区域分布情况和各变量的空间自相关性进行一定的初步了解。

下面将首先进行房屋销售面积、经济增长率、人口结构、人年均可支配收入、房屋销售价格五个变量的空间分布描述，在对各个变量的空间分布有了一个直观的了解后，我们分别基于 Moran's I 指数对各个变量进行空间自相关性检验，文中相关的描述统计分析是通过 GeoDa 软件实现的。由于一年期贷款利率为全国统一利率，在这里不对其区域分布情况进行分析。

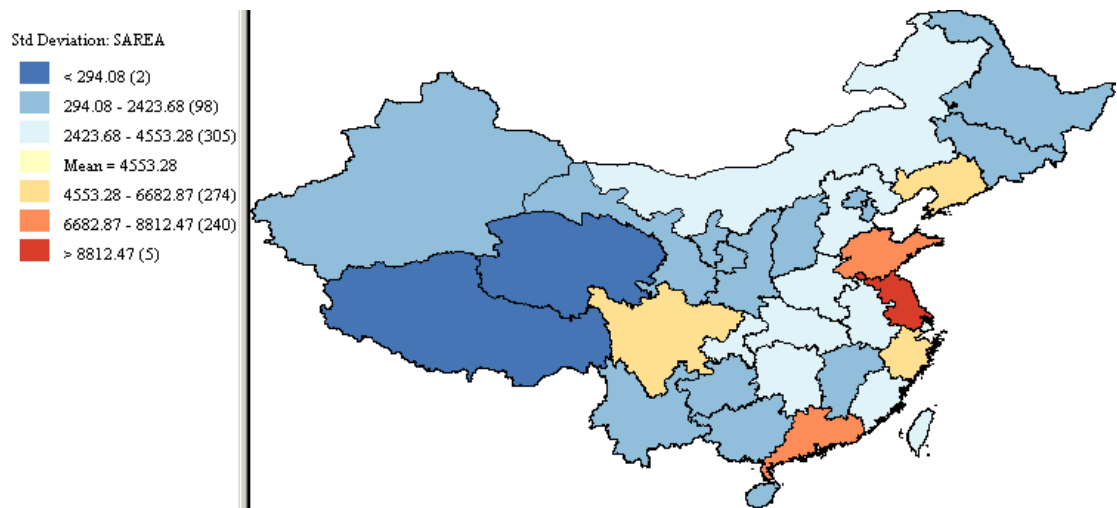


图 1 SAREA（房屋销售面积）分布图

从图 1 中我们可以发现，我国房屋销售面积的平均水平为 4553.28 万平方米，江苏、广东、山东的房屋销售面积相对较大，其中江苏省的房屋销售面积达 10248.2 万平方米，湖南、湖北等中部地区的销售面积处于平均水平，而房屋销售面积较少的城市包括西部地区的青海和宁夏，只有不足一千万平方米，说明了我国各地对房屋的需求量有着一定的差异。

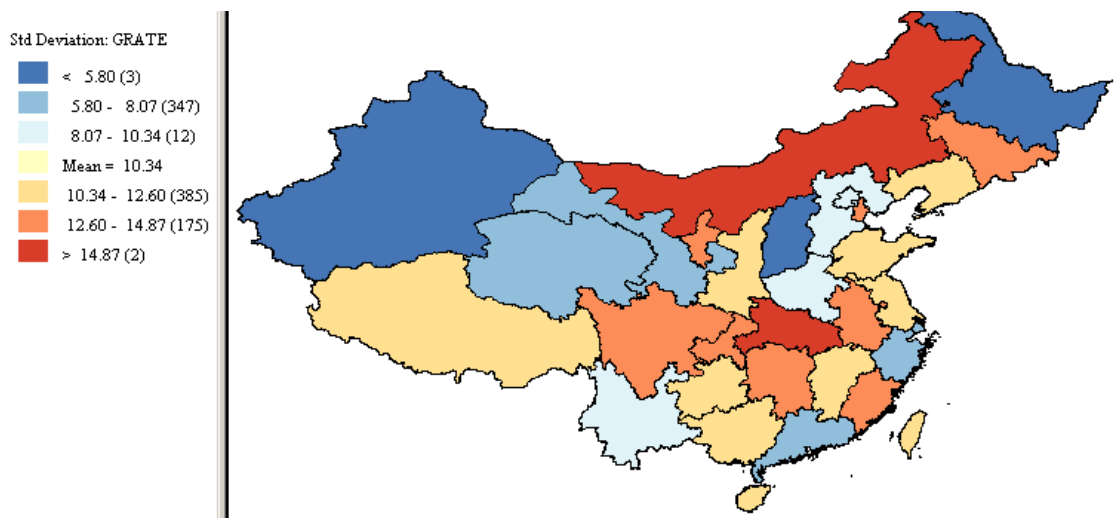


图 2 GRATE（经济增长率）分布

由图 2 可知，我国各省的经济增长率的平均水平为 10.34%，其中内蒙古自治区和湖北省的经济增长率最大，湖南、安徽等中部地区的增长率也在平均水平之上，然而新疆、青海和甘肃省等西部地区的增长率偏低，其中，新疆维吾尔自治区的增长率最低，为 2.97%，这说明了我国地区间经济发展水平的差距较大。

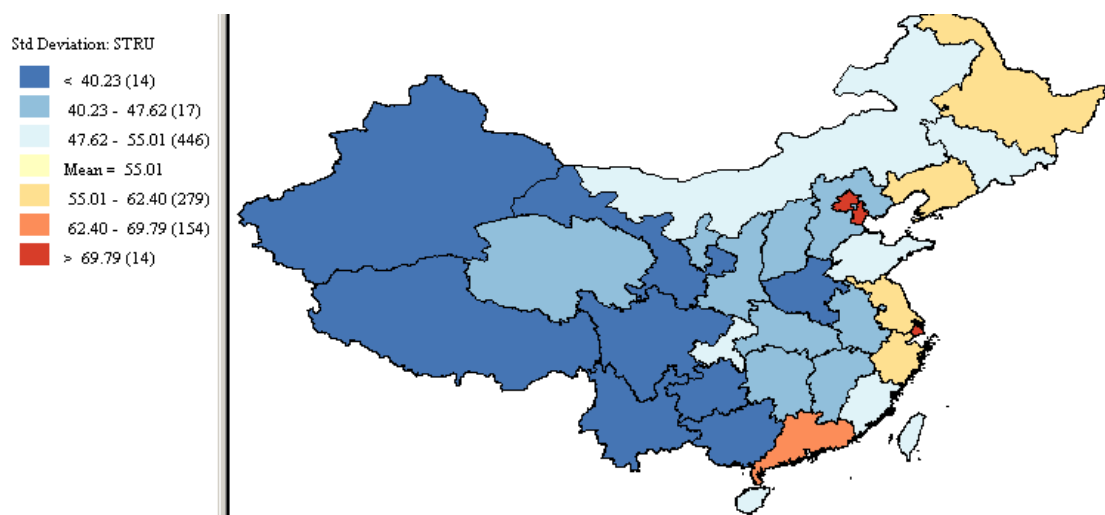


图3 STRU（人口结构）分布

由图3，我们可以看出我国人口结构的平均水平为55.01%，即我国城镇居民人口数占我国总人口数的55.01%。其中北京、上海等一线城市的人口结构偏高，北京的人口结构水平达到了85%，而新疆、甘肃、云南等西部地区的人口结构水平明显偏低，其中甘肃省的人口结构水平仅为32.65%，其农村居民人数为城镇居民人数的两倍，如此明显的地区间城市化差异也说明了我国西部大开发战略的必要性。

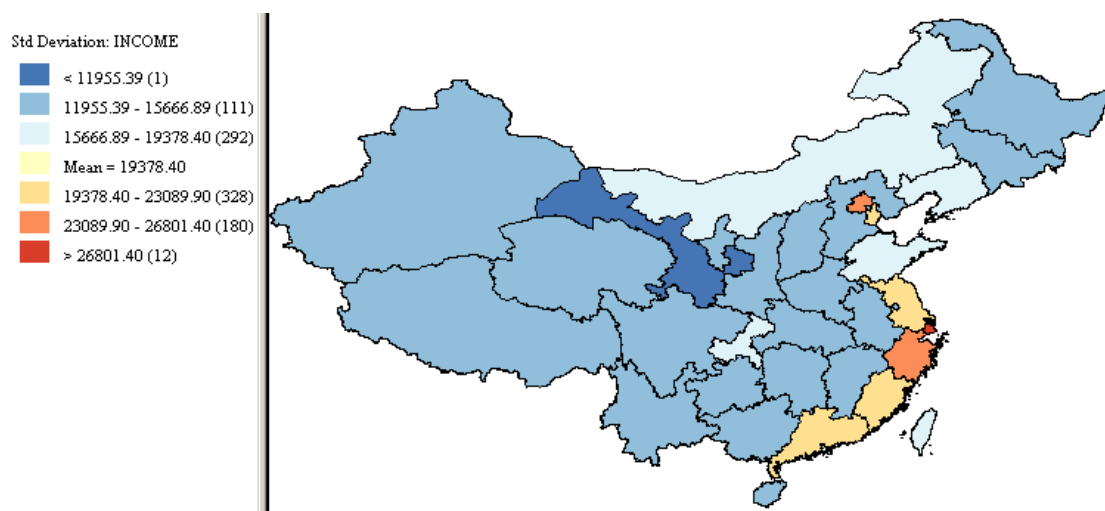


图4 INCOME（年人均可支配收入）分布

由图4知，我国年人均可支配收入的平均水平为19378.40元，其中上海、北京和浙江省的年人均可支配收入较高，最高达到28837.78元，沿海城市的年人均可支配收入水平也相对较高，除甘肃以外的其他城市的水平也处于平均水平，但是甘肃省的年人均可支配收入只有11929.78元，还不到发达城市水平的二分之一，说明了年人均可支配收入的地区差异性也较大。

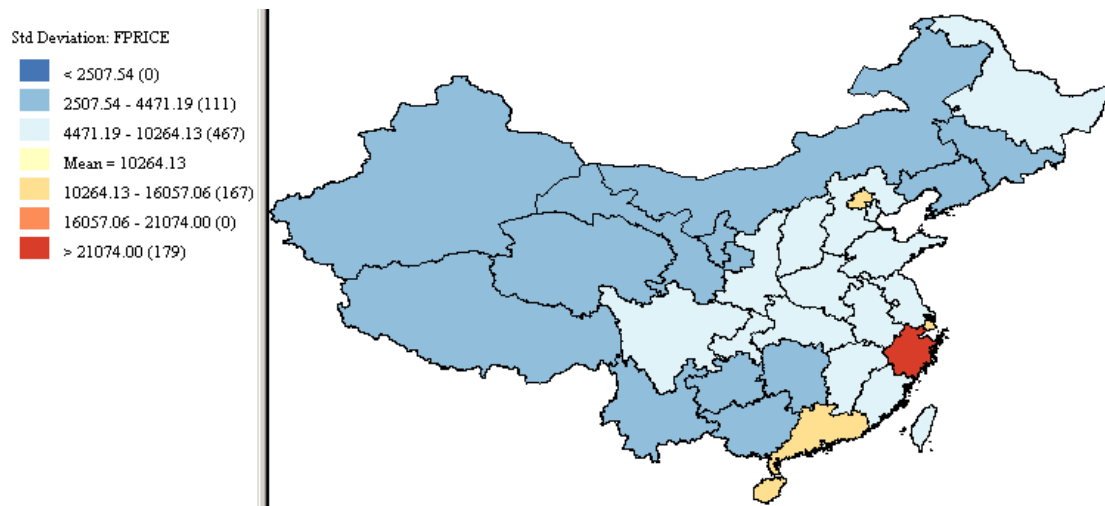
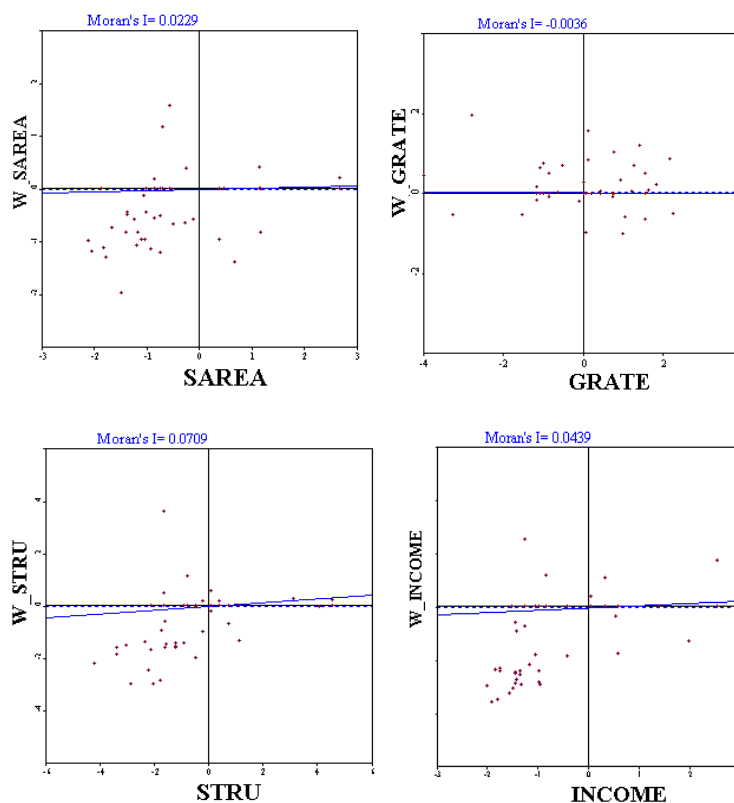


图 5 FPRICE（房屋销售价格）分布

由图 5 知，我国房屋销售价格的平均水平为 10264.13 元/平方米，其中浙江省的房屋销售价格明显高于其他地区，达到 21074 元/平方米，中部地区的销售价格处于平均水平左右，而西部地区的销售价格则明显偏低，其中甘肃省仅为 2507.54 元/平方米，这也说明了我国各地房屋销售价格的差距之大。

在对各变量进行了空间分布描述之后，我们需要从统计角度对各变量在空间上是否存在自相关性进行检验，本文使用 Moran's I 作为空间自相关性检验指标。Moran's I 指标是用来衡量相邻的空间分布对象及其属性取值之间的关系参数，系数取值范围在-1 到 1 之间，正值表示该空间事物的属性分布具有正相关性；负值表示该空间事物的属性分布具有负相关性；0 表示该空间事物的属性分布不存在相关性。

利用 GeoDa 软件对人口结构这个变量进行空间自相关分析，得到结果如下。



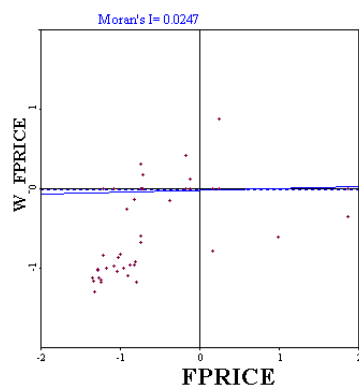


图 6 需求模型各变量空间自相关检验图

图 6 显示，房屋销售面积、经济增长率、人口结构、人均年可支配收入、房屋销售价格的 Moran's I 系数分别为 0.0229、-0.0036、0.0709、0.0439、0.0247，根据 Moran's I 系数的正负可以判定各变量全国 30 省市的空间上呈现的正负相关关系。而 Moran's I 系数是否显著对于我们的判定空间自相关性至观重要，因此，为了检验 Moran's I 系数的显著性，我们利用蒙特卡罗模拟的方法来检验 Moran's I 系数的显著水平，得到的结果如下图所示。

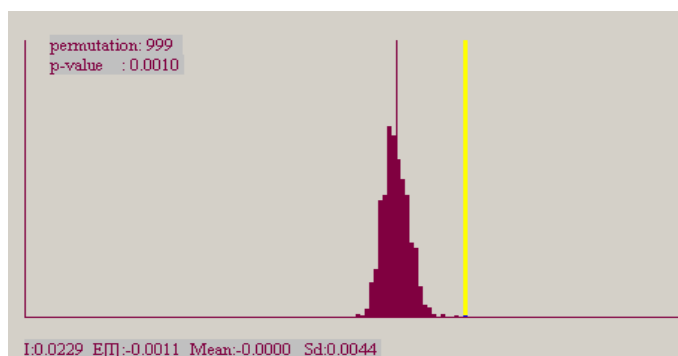


图 7 房屋销售面积 Moran's I 系数检验图

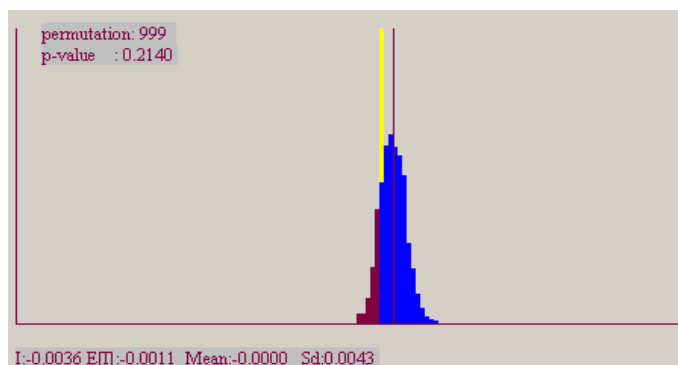


图 8 经济增长率 Moran's I 系数检验图

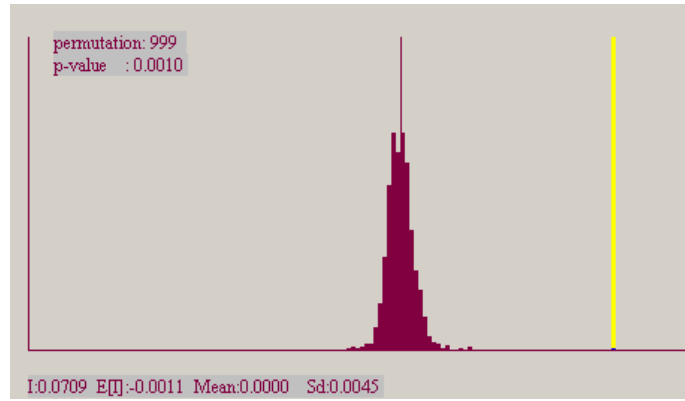


图 9 人口结构 **Moran's I** 系数检验图

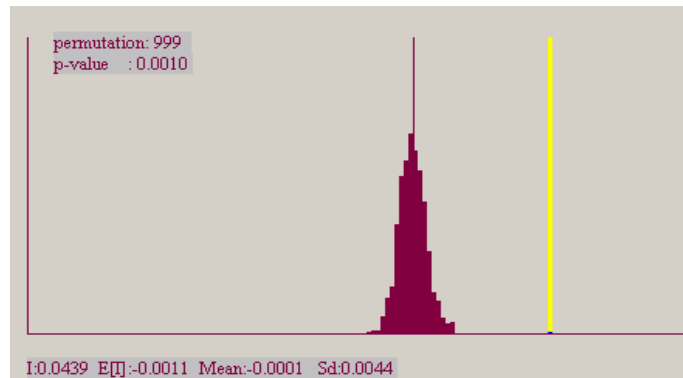


图 10 人均年可支配收入 **Moran's I** 系数检验图

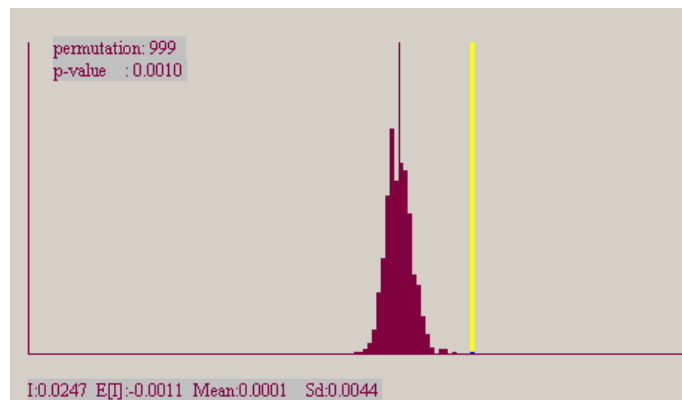


图 11 房屋销售价格 **Moran's I** 系数检验图

图 7-11 显示，房屋销售面积、经济增长率、人口结构、人均年可支配收入、房屋销售价格的 Moran's I 系数显著性检验的 p 值分别为 0.001、0.214、0.001、0.001、0.001，以人口结构为例，其对应 p 值为 0.0010，表明在 99.9% 的显著性水平下人口结构在空间上正的自相关关系。

综上，我们可以得到房屋销售面积、经济增长率、人口结构、人均年可支配收入、房屋销售价格 Moran's I 系数及显著性水平，以列表形式呈现在表 1 中，可知，房屋销售面积、人口结构、年人均可支配收入、房屋销售价格均在空间上呈现正的自相关关系，且显著性非常高，而经济增长率在空间上负自相关性并不明显。由于变量空间自相关性的存在，我们有必要利用空间地理加权回归方法建

立房地产需求模型。

表 1 需求模型各变量 Moran's I 置信度表

	<i>SAREA</i>	<i>GRATE</i>	<i>STRU</i>	<i>INCOME</i>	<i>FPRICE</i>
Moran's I	0.0229	-0.0036	0.0709	0.0439	0.0247
置信度 (%)	99.9	78.6	99.9	99.9	99.9

(2) 地理加权回归需求模型实证结果

基于 2009 年全国 30 个省市的房地产数据建立地理加权回归需求模型，来透视房地产需求市场存在的空间异质性和区域差异性，从而更加有针对性和精准性地分析各地区房地产市场供需状况。

地理加权回归过程通过 Matlab 软件实现，结果整理见表 2、表 3、图 12。

在具体观察每个省份解释变量系数估计值之前，我们首先对 30 个省市的解释变量系数估计值进行了描述统计分析，如表 12，我们可以观察到各个解释变量系数的极大值、极小值、均值和标准差，对解释变量在各省市的分布有一个大致的了解和认识。

表2 需求模型参数估计值描述统计分析表

变量名	极小值	极大值	均值	标准差
<i>c</i>	-1.575	-0.241	-0.910	0.361
<i>rate</i>	-8.365	-1.282	-4.832	1.915
<i>fprice</i>	-0.486	1.666	0.585	0.573
<i>grate</i>	0.022	0.117	0.072	0.029
<i>stru</i>	-0.099	-0.024	-0.054	0.019
<i>income</i>	0.831	4.461	3.232	1.054

图 12 给出了基于高斯距离、指数距离和三次方距离为权值的 *GWR* 模型的参数估计值，可以直观的看出各解释变量回归系数在空间截面上的大小分布和走势，其中，横坐标为与中国统计年鉴相一致的全国 30 个省的编号，纵坐标分别表示常数项、利率、商品住房销售价格、经济增长率、人口结构、人均年可支配收入，通过观察我们发现，基于高斯距离和三次方距离权值的参数估计值在各个省域上基本一致，而基于指数距离权值的参数估计值在各个省域之间存在较大的波动和变化。

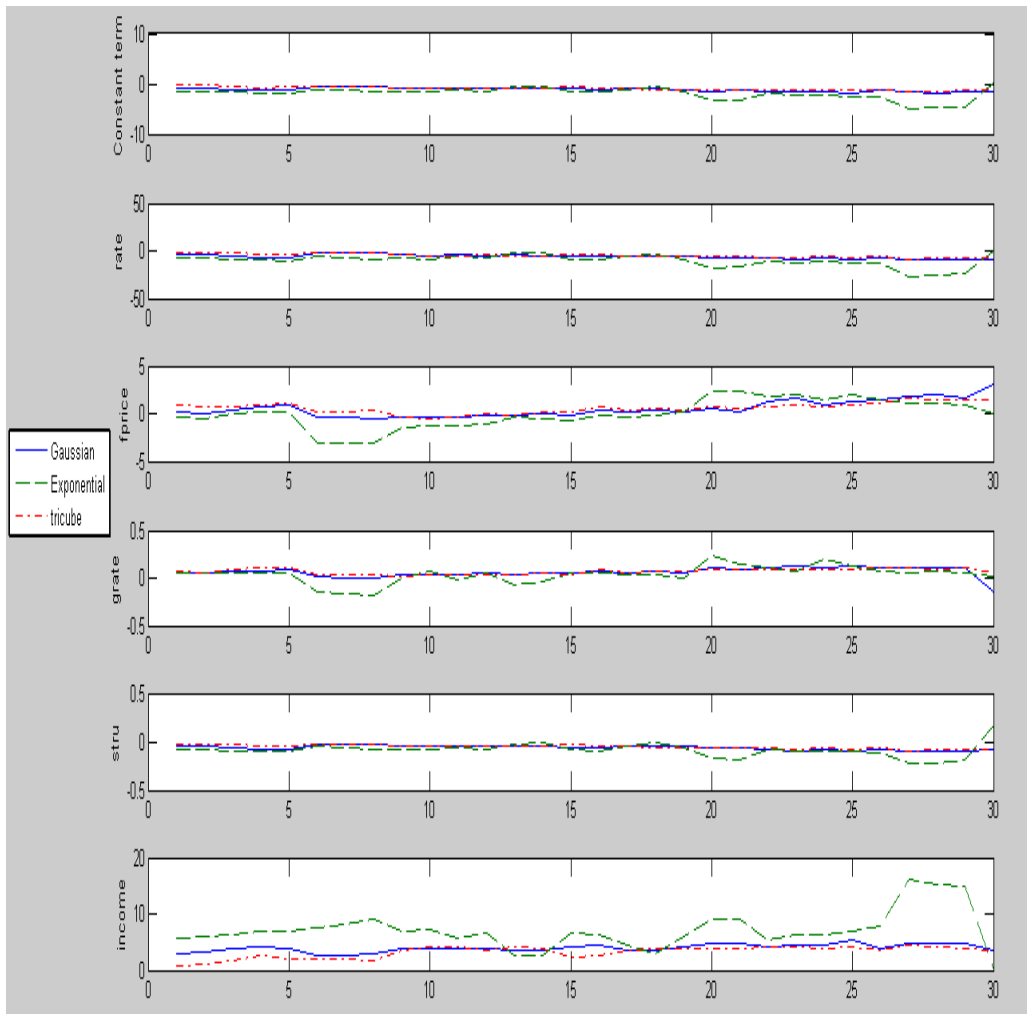


图 12 GWR 模型中各参数的估计值走势图

本文主要根据高斯距离权重矩阵而进行计算，具体的计算而得的结果如表 3，其中分 30 个省市全面展示了房地产需求模型中各解释变量（一年期贷款利率、商品住房销售价格、经济增长率、人口结构、人均年可支配收入）对房屋需求量的影响程度和显著性水平。

表 3 基于地理加权回归房地产需求模型结果表

省份	C	$RATE$	$FPRICE$	$GRATE$	$STRU$	$INCOME$
北 京	-0.241	-1.282	0.856	0.068	-0.024	0.831
t 统计量	-0.686	-0.686	1.699	2.027	-1.590	0.566
天 津	-0.283	-1.500	0.684	0.059	-0.026	1.127
t 统计量	-0.828	-0.828	1.370	1.769	-1.712	0.785
河 北	-0.491	-2.609	0.785	0.085	-0.035	1.681
t 统计量	-1.197	-1.197	1.535	2.522	-2.030	1.011
山 西	-0.818	-4.343	0.852	0.106	-0.051	2.662
t 统计量	-1.791	-1.791	1.701	3.278	-2.721	1.477
内蒙古	-0.644	-3.419	1.126	0.109	-0.046	1.869
t 统计量	-1.437	-1.437	2.192	3.317	-2.508	1.055

辽 宁	-0.390	-2.070	0.226	0.034	-0.030	1.914
t 统计量	-1.274	-1.274	0.465	1.072	-2.196	1.444
吉 林	-0.385	-2.043	0.234	0.035	-0.030	1.891
t 统计量	-1.255	-1.255	0.479	1.092	-2.194	1.422
黑龙江	-0.368	-1.953	0.282	0.037	-0.029	1.791
t 统计量	-1.187	-1.187	0.573	1.161	-2.136	1.334
上 海	-0.746	-3.964	-0.412	0.022	-0.041	3.634
t 统计量	-2.535	-2.535	-0.932	0.748	-3.189	2.928
江 苏	-0.936	-4.970	-0.486	0.032	-0.049	4.302
t 统计量	-2.794	-2.794	-1.085	1.155	-3.540	3.106
浙 江	-0.915	-4.858	-0.430	0.034	-0.046	4.167
t 统计量	-2.902	-2.902	-1.001	1.117	-3.421	3.249
安 徽	-0.834	-4.426	-0.040	0.045	-0.045	3.550
t 统计量	-2.416	-2.416	-0.093	1.449	-3.050	2.579
福 建	-1.014	-5.385	-0.168	0.045	-0.049	4.225
t 统计量	-3.265	-3.265	-0.419	1.323	-3.418	3.535
江 西	-0.979	-5.197	0.158	0.059	-0.050	3.808
t 统计量	-3.168	-3.168	0.422	1.729	-3.260	3.321
山 东	-0.521	-2.766	0.197	0.043	-0.034	2.347
t 统计量	-1.526	-1.526	0.406	1.365	-2.348	1.637
河 南	-0.768	-4.081	0.676	0.082	-0.048	2.679
t 统计量	-1.874	-1.874	1.465	2.648	-2.778	1.654
湖 北	-0.934	-4.959	0.280	0.056	-0.052	3.582
t 统计量	-2.920	-2.920	0.729	1.871	-3.570	2.909
湖 南	-1.080	-5.735	0.517	0.078	-0.058	3.808
t 统计量	-3.258	-3.258	1.368	2.355	-3.662	3.075
广 东	-1.044	-5.543	0.434	0.077	-0.054	3.754
t 统计量	-3.218	-3.218	1.155	2.239	-3.344	3.126
广 西	-1.154	-6.125	0.666	0.092	-0.061	3.895
t 统计量	-3.352	-3.352	1.725	2.671	-3.603	3.035
海 南	-1.110	-5.895	0.576	0.089	-0.057	3.830
t 统计量	-3.286	-3.286	1.499	2.520	-3.390	3.053
重 庆	-1.221	-6.482	0.776	0.095	-0.068	4.028
t 统计量	-3.530	-3.530	2.033	2.912	-4.081	3.128
四 川	-1.308	-6.944	0.914	0.099	-0.074	4.197
t 统计量	-3.843	-3.843	2.453	3.167	-4.445	3.353
贵 州	-1.195	-6.343	0.732	0.094	-0.065	3.976
t 统计量	-3.444	-3.444	1.902	2.800	-3.856	3.073
云 南	-1.302	-6.911	0.899	0.098	-0.072	4.178

<i>t</i> 统计量	-3.807	-3.807	2.342	2.974	-3.902	3.339
陕 西	-1.122	-5.956	1.098	0.106	-0.070	3.450
<i>t</i> 统计量	-3.234	-3.234	2.994	3.745	-4.529	2.654
甘 肃	-1.575	-8.365	1.666	0.113	-0.099	4.461
<i>t</i> 统计量	-5.226	-5.226	5.963	5.535	-7.510	4.169
青 海	-1.407	-7.473	1.520	0.086	-0.089	4.068
<i>t</i> 统计量	-3.755	-3.755	4.374	3.508	-5.457	3.052
宁 夏	-1.293	-6.868	1.435	0.117	-0.085	3.737
<i>t</i> 统计量	-3.802	-3.802	4.267	4.942	-5.889	2.967
新 疆	-1.224	-6.499	1.482	0.062	-0.080	3.530
<i>t</i> 统计量	-2.552	-2.552	3.519	2.140	-3.984	2.061

基于上表中的实证结果，同时结合中国房地产市场客观现实，我们分别对各解释变量对需求的作用机制进行分析。

贷款利率——对于全国 30 个省市，贷款利率对商品住房需求量的影响存在着普遍的负相关关系，即贷款利率越高，住房需求量越小，结合现实生活，由于商品住房价格相对于普通家庭收入而言是一笔巨大的支出费用，目前人们购买住房通常采取住房抵押贷款的方式，如果贷款利率升高，则意味着人们的贷款成本增加，这将减弱人们的购房动机和需求。在对贷款利率对商品住房需求量的影响方向有了一个认识之后，我们来分地区考虑其影响程度的大小，影响程度较大的地区是甘肃、青海、宁夏、新疆、云南、陕西、贵州、四川、重庆等地，以甘肃省为例，回归系数为-8.365，说明贷款利率每变动一个百分点，商品住房需求面积将减少 8.365%，这些地区主要集中在西部地区，这里透露了现实的经济背景，即这些地区人们的生活水平和收入水平较低，购买住房主要依靠贷款，如果贷款利率提高，人们宁愿拒绝购买住房，而拒绝购买住房将直接影响这部分地区的经济发展和人民生活水平的提高；影响程度次高的地区主要为浙江、江苏、福建、广东、广西、海南等地区，这些地区主要集中在东部经济发达地段，人均收入水平高，城市化水平进程发展较快，人们购买住房的目的与中西部地区不同，他们并非满足自身居住需求而是为进行房地产投资，因此，当贷款利率提高时，东部地区人们投资热情将减弱，购买住房需求将减少，这一部分需求的减少有利于消除房地产泡沫，从而有助于我国房地产行业的可持续发展。综合考虑我国东、中、西部地区的经济发展特点和房地产行业发展现状，政府对东、中、西部地区实行差别化贷款利率调整是有必要的。

商品住房销售价格——从经济学角度而言，人民购买住房的动机分为两类，一类是将房屋作为正常消费品，通过购买满足自身的住房需求，另一类是将房屋作为投资品，通过投资房地产行业来实现房屋的增值保值。对于第一类人群，商品住房销售价格理应与其需求呈现负相关关系，对于第二类人群，商品住房销售价格的上升将给其带来投资收益，故商品住房销售价格会与需求呈现正相关关系。从实证分析结果来看，我国绝大部分地区的商品住房销售价格与需求呈现正相关关系，即人们都投资于房地产行业谋取资本增值，这也折射了我国房地产行业过热的现状，出现了非理性的投资和需求现象，市场这只看不见的手的调节作用无法正常体现，需要政府进行有力的宏观政策调控。仅有的出现负相关关系的省份分别是上海、江苏、浙江、安徽、福建，这些东部沿海地区人们应划分为第二类人群，房价与需求本应呈现正相关关系，而实证的结果却为负相关，说明，

这些地区的房价被“炒”得过高，出现了非理性定价，即使作为第二类人群的投资品，房价也超过其可支付的范围和能力。

经济增长率——从全国 30 个省市的实证结果可以看出，经济增长率与商品住房需求呈现普遍的正相关现象，即经济发展越快，商品住房的需求量也越大，这是符合客观经济规律的，当经济发展水平提高时，人均收入水平将得到改善，提高生活质量的追求会驱使人们增加商品房屋的需求量，但是经济增长率对住房需求正向影响普遍较小，回归系数均值为 0.072，即经济增长率提高一个百分点，人们的住房需求将增加 0.072%。

人口结构——人口结构即各地区城镇人口数与总人口数的比值，该指标可以反映一个地区的城市化水平，理论上言，城市化水平越高，会产生更多的城市商品住房需求，即城市化水平与商品住房需求呈现正相关关系，而本文实证研究结果却恰恰相反，城市化水平与商品住房需求呈现负相关关系，通过查阅文献资料知道，原因在于我国城市化水平可能存在水分，城市化的衡量指标应为城市常住人口与总人口的比例，而我们现在将近 47.5% 的城市化推进中，有 2 亿左右的农民工没有正式纳入城市化。因此，扎实稳健推进我国城市化进程，抓住城市化发展契机来推动我国房地产行业的发展意义重大。

人均年可支配收入——实证结果表明，全国 30 个省市的人均可支配收入同商品住房需求均呈现正相关关系，即人均可支配收入的提高可以大大提高人们购买住房的需求量，无论是出于自身住房需求还是投资需求。回归系数较大的省市分别是甘肃、青海、云南、四川、重庆、贵州等地，这地西部地区经济发展水平较东部地区落后，人均可支配收入水平的提高将显著地提高人们的住房需求，可见，加快推进西部地区的经济发展，提高人们的收入水平，对于西部地区的房地产行业的发展至关重要。

2. 供给模型的求解

（1）基于 GeoDa 软件的供给模型各变量空间异质性检验

在这部分建立供给模型与建立需求模型的思路相同，首先，为了说明建立地理加权回归供给模型的必要性，在建模之前，我们有必要对各供给模型变量进行空间异质性和自相关性的检验。

在房地产供给模型中，共有商品住房竣工面积、经济增长率、土地交易价格、商品住房销售价格、贷款利率五个变量，鉴于经济增长率、商品住房销售价格的空间分布状况已在需求模型中体现，而一年期贷款利率为全国统一利率，在这里不对其区域分布情况进行分析。故在此，我们仅对商品住房竣工面积和土地交易价格进行空间分布的描述和空间自相关性的检验，空间自相关性的检验同样基于 Moran's I 指数。

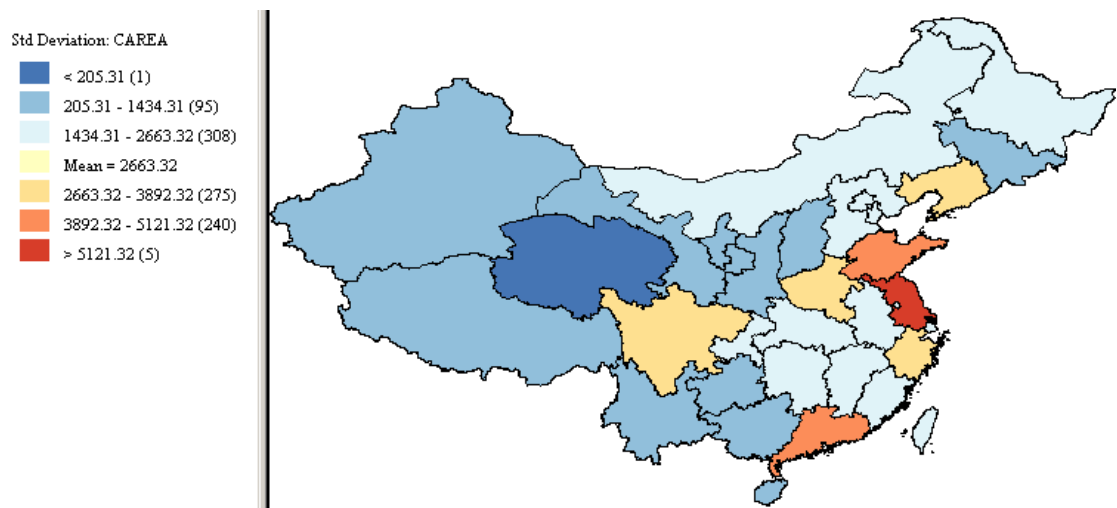


图 13 CAREA（房屋竣工面积）分布

从图 13 中我们可知，我国房屋竣工面积的平均水平为 2663.32 万平方米，其中房屋竣工面积较高的地区有江苏、山东和广东省，房屋供给量最高的江苏省房屋竣工面积为 6731.41 万平方米，四川、河南和辽宁在内的中、东部地区的房屋供给量在平均水平左右，而西部地区的房屋供给量明显较低，特别是甘肃省的房屋竣工面积最低，仅为 461.78 万平方米。

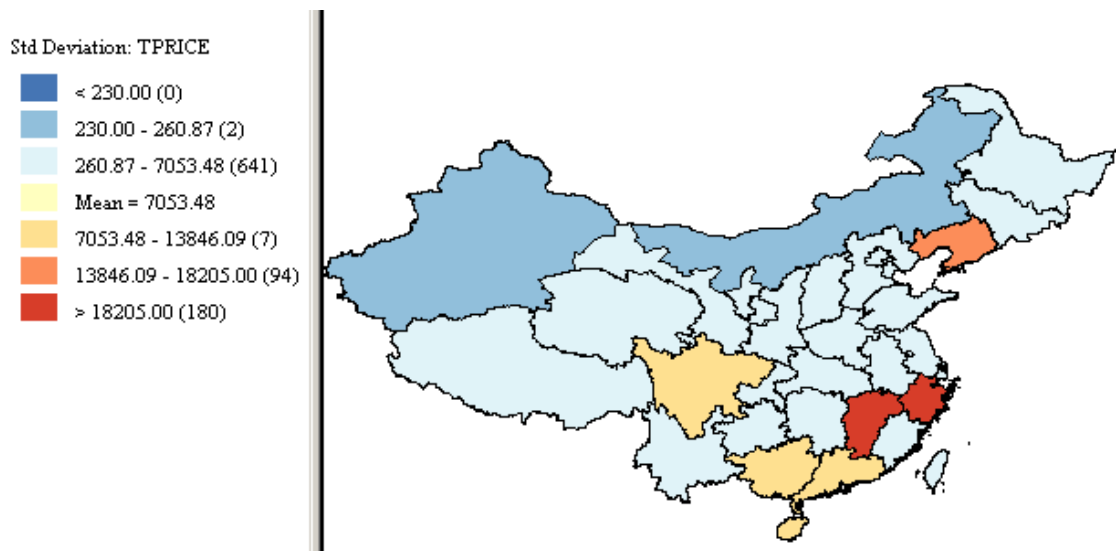


图 14 TPRICE（土地交易价格）分布

由图 14 知，我国土地价格的平均水平为 7053.48 元/平方米，其中浙江、江西和辽宁的土地价格最高，广东等沿海城市的土地价格也在平均水平之上，而新疆、内蒙古等地区的土地价格则相对较低。

在对商品住房竣工面积和土地交易价格进行了空间分布描述之后，我们同样需要从统计角度对各变量在空间上是否存在自相关性进行检验，本文使用 Moran's I 作为空间自相关性检验指标。利用 GeoDa 软件对这两个变量的空间自相关分析结果如下。

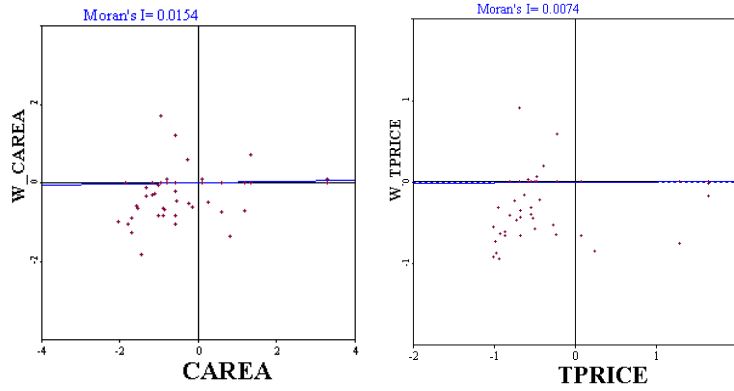


图 15 供给模型变量空间自相关检验图

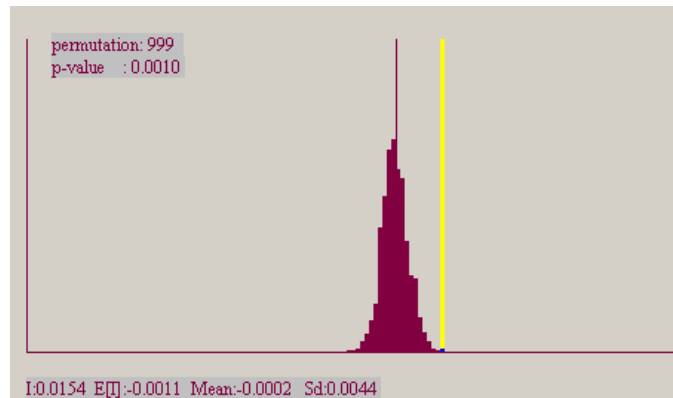


图 16 商品住房竣工面积 **Moran's I** 系数检验图

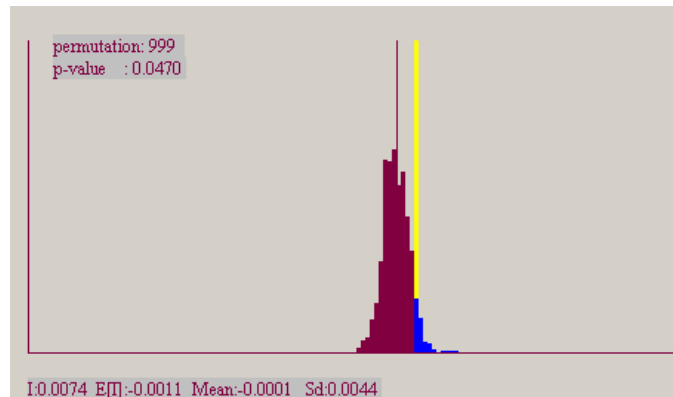


图 17 土地交易价格 **Moran's I** 系数检验图

图 16、17 显示,房屋竣工面积、土地交易价格的 Moran's I 系数分别为 0.0154、0.0074,其显著性水平分别为 0.001、0.047,说明房屋竣工面积和土地交易价格在空间上存在着显著的正相关性。结合经济增长率、商品住房销售价格的 Moran's I 系数及其显著性水平,我们将供给模型中各变量的 Moran's I 系数及其置信度水平列于表 4,分析表中数据我们可以知道空间自相关性的显著存在性和建立地理加权回归模型的必要性。

表 4 供给模型各变量 **Moran's I** 置信度表

	<i>CAREA</i>	<i>GRATE</i>	<i>FPRICE</i>	<i>TPRICE</i>
Moran's I	0.0154	-0.0036	0.0247	0.0074
置信度 (%)	99.9	78.6	99.9	95.3

(2) 地理加权回归供给模型实证结果

与需求模型采用相同的方法，我们基于 2009 年全国 30 个省市的房地产数据建立地理加权回归供给模型，来透视房地产供给市场存在的空间异质性和地区差异性，希望通过实证结果的分析可以为相关部门制定政策提供一个全新的思路。

地理加权回归过程通过 Matlab 软件实现，结果整理见表 5、表 6 和图 18。

表 5 首先对 30 个省市的解释变量系数估计值进行了描述统计分析，我们可以观察到各个解释变量系数的极大值、极小值、均值和标准差，对解释变量在各省市的分布有一个初步的了解和认识。

表5 供给模型参数估计值描述统计分析表

变量名	极小值	极大值	均值	标准差
c	-0.156	0.170	0.034	0.102
$grate$	0.041	0.105	0.068	0.020
$tprice$	0.043	0.340	0.233	0.067
$fprice$	0.104	1.061	0.453	0.305
$rate$	-0.828	0.903	0.181	0.540

图 18 给出了基于高斯距离、指数距离和三次方距离为权值的 GWR 模型的参数估计值，可以直观的看出各解释变量回归系数在空间截面上的大小分布和走势，其中，横坐标为与中国统计年鉴相一致的全国 30 个省的编号，纵坐标分别表示常数项、经济增长率、土地交易价格、商品住房销售价格、利率，通过观察我们发现，基于高斯距离和三次方距离权值的参数估计值在各个省域上基本一致，而基于指数距离权值的参数估计值在各个省域之间存在较大的波动和变化。

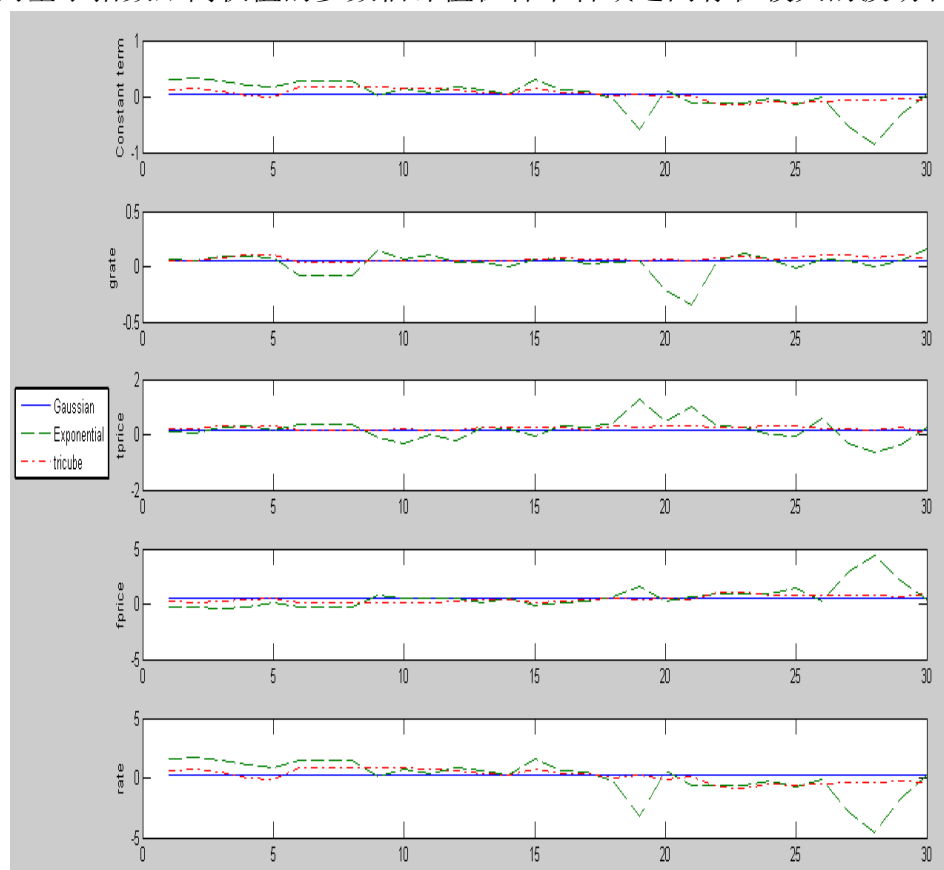


图 18 GWR 模型中各参数的估计值走势图

本文主要根据高斯距离权数矩阵而进行计算，具体的计算而得的结果如表 6，其中分 30 个省市全面展示了房地产供给模型中各解释变量（经济增长率、土地交易价格、商品住房销售价格、利率）对房屋供给量的影响程度和显著性水平。

表 6 基于地理加权回归房地产供给模型结果表

省份	<i>C</i>	<i>GRATE</i>	<i>TPRICE</i>	<i>FPRICE</i>	<i>RATE</i>
北 京	0.110	0.053	0.211	0.239	0.583
<i>t</i> 统计量	1.544	1.742	1.713	0.885	1.544
天 津	0.133	0.048	0.205	0.175	0.705
<i>t</i> 统计量	1.905	1.617	1.721	0.671	1.905
河 北	0.082	0.079	0.294	0.220	0.437
<i>t</i> 统计量	1.150	2.420	2.174	0.797	1.150
山 西	0.011	0.101	0.291	0.430	0.060
<i>t</i> 统计量	0.158	2.994	1.933	1.427	0.158
内 蒙 古	-0.016	0.105	0.295	0.513	-0.087
<i>t</i> 统计量	-0.216	3.022	2.000	1.655	-0.216
辽 宁	0.165	0.041	0.167	0.115	0.874
<i>t</i> 统计量	2.442	1.447	1.482	0.462	2.442
吉 林	0.169	0.041	0.162	0.106	0.899
<i>t</i> 统计量	2.472	1.420	1.420	0.419	2.472
黑 龙 江	0.170	0.041	0.160	0.105	0.903
<i>t</i> 统计量	2.540	1.450	1.443	0.425	2.540
上 海	0.161	0.047	0.177	0.116	0.854
<i>t</i> 统计量	2.228	1.472	1.251	0.492	2.228
江 苏	0.155	0.056	0.190	0.113	0.820
<i>t</i> 统计量	2.137	1.739	1.305	0.458	2.137
浙 江	0.139	0.051	0.185	0.172	0.740
<i>t</i> 统计量	1.792	1.476	1.140	0.686	1.792
安 徽	0.128	0.057	0.175	0.212	0.679
<i>t</i> 统计量	1.790	1.755	1.148	0.850	1.790
福 建	0.061	0.057	0.243	0.363	0.323
<i>t</i> 统计量	0.696	1.348	1.202	1.336	0.696
江 西	0.041	0.058	0.258	0.414	0.218
<i>t</i> 统计量	0.489	1.347	1.331	1.521	0.489
山 东	0.139	0.060	0.242	0.104	0.738
<i>t</i> 统计量	2.014	1.952	1.957	0.422	2.014
河 南	0.077	0.078	0.272	0.261	0.407
<i>t</i> 统计量	1.113	2.455	1.786	0.917	1.113
湖 北	0.071	0.062	0.181	0.385	0.379
<i>t</i> 统计量	0.966	1.808	1.069	1.498	0.966
湖 南	0.011	0.060	0.302	0.474	0.056

t 统计量	0.127	1.498	1.610	1.771	0.127
广 东	0.047	0.053	0.259	0.398	0.251
t 统计量	0.607	1.202	1.488	1.488	0.607
广 西	-0.018	0.065	0.336	0.526	-0.094
t 统计量	-0.241	1.482	2.045	1.882	-0.241
海 南	0.026	0.055	0.297	0.429	0.137
t 统计量	0.343	1.193	1.764	1.577	0.343
重 庆	-0.146	0.080	0.282	1.004	-0.775
t 统计量	-1.679	2.176	1.913	2.925	-1.679
四 川	-0.156	0.089	0.245	1.061	-0.828
t 统计量	-1.736	2.533	1.759	2.968	-1.736
贵 州	-0.083	0.072	0.340	0.737	-0.443
t 统计量	-1.059	1.760	2.135	2.381	-1.059
云 南	-0.107	0.078	0.326	0.824	-0.569
t 统计量	-1.297	1.853	2.084	2.532	-1.297
陕 西	-0.085	0.101	0.234	0.811	-0.450
t 统计量	-1.062	3.076	1.561	2.420	-1.062
甘 肃	-0.076	0.101	0.212	0.801	-0.404
t 统计量	-0.916	3.047	1.474	2.304	-0.916
青 海	-0.066	0.085	0.148	0.847	-0.351
t 统计量	-0.753	2.569	1.035	2.303	-0.753
宁 夏	-0.047	0.105	0.250	0.659	-0.248
t 统计量	-0.586	3.032	1.562	1.920	-0.586
新 疆	-0.073	0.066	0.043	0.986	-0.387
t 统计量	-0.743	1.933	0.284	2.346	-0.743

基于上表中的实证结果，同时结合中国房地产市场客观现实，我们分别对各解释变量对供给的作用机制进行分析。

经济增长率——观察全国 30 个省市的经济增长率回归系数，可以发现，经济增长率与商品住房供给均呈现正相关关系，即经济的良好平稳运行对于房地产商进行投资有积极促进作用，它提供了一个良好的外部环境，但是从影响力系数大小而言，经济增长对房地产供给的促进作用并不大，回归系数均值为 0.068，即经济增长率提高一个百分点，商品住房供给量增加 0.068%；从影响力系数显著性程度而言，重庆、四川、陕西、甘肃、青海、宁夏等西部地区影响力系数显著，这提供给我们一个有用的信息，经济增长对于西部地区的房地产行业促进作用颇大，而对中东部地区的房地产行业的边际促进力较小。

土地交易价格——从理论上而言，土地作为房地产商的成本，当土地价格上涨时，其生产成本提高，理性的房地产商应该缩减投资规模，减少房屋的供给量，即土地交易价格与商品住房供给量呈现负相关关系，而从实证研究的结果来看，全国 30 个省市土地交易价格对商品住房供给量呈现正相关关系，与经济供需理论相违背，这说明我国土地交易市场对供给量的作用机制不能很好地体现，我国土地交易市场有待进一步完善。众所周知，我国土地供应目前实行“双轨制”，

一种方法是政府无偿划拨国有土地给使用单位，另一种方法是通过市场化手段有偿出让国有土地，为了使得土地定价能够反映资源的稀缺性，取消土地双轨制是我国政府部门目前应该努力的方向。

商品住房销售价格——从理论上而言，商品住房销售价格应该与住房供给量呈现正相关关系，而实证分析的结果也同样验证了理论的正确性，即随着商品住房销售价格的上涨，我国房地产投资商的利润空间变大，因此，他们会通过提高供给的方式来进一步谋取利润。

贷款利率——从理论上讲，房地产商追求利润最大化，贷款利率的提高会加大其投资成本，故贷款利率应该与商品住房供给量呈现负相关关系，但是，实证分析的结果却与理论有所出入。考虑贷款利率影响方向，我国中东部地区的贷款利率与商品住房供给量呈现正相关关系，而唯有重庆、贵州、四川、云南、陕西、甘肃、青海、宁夏、新疆等西部地区的贷款利率与商品住房呈现负相关关系，分析可知，中东部地区的房地产供给市场发展理论与理论相背，而西部地区与理论相同，充分验证了我国中东部地区的房地产市场过热，出现非理性投资和巨大的“泡沫”，而西部地区的房地产市场发展并没有出现过度非理性投资现象。

5.1.3 基于需求、供给模型的房地产市场区域聚类分析

在基于地理加权回归方法建立了房地产市场需求和供给模型之后，考虑到我国疆域广阔，经济发展不平衡的客观现实，我们有必要在需求、供给模型结果的基础上对我国房地产市场进行划分，这也是本文采用地理加权回归方法建模的落脚点——对我国房地产市场进行划分，使得政府在制定政策时，充分考虑地区差异性的存在，避免政策“一刀切”，从而更加有针对性和精准性的对房地产市场进行差别化的宏观调控。

首先，我们基于需求模型的结果对房地产市场进行分层聚类分析，即根据全国 30 个省市的常数项、利率回归系数、商品住房销售价格回归系数、经济增长率回归系数、人口结构回归系数、人均年可支配收入回归系数作为聚类分析变量，采用类间平均距离连接法作为聚类方法，欧式距离平方作为距离测度方法，聚类的结果见表 7，SPSS 输出结果见附录四。

与基于需求模型的聚类方法相同，我们对房地产市场进行基于供给模型和基于需求和供给模型的聚类分析。其中，基于供给模型的聚类分析变量是常数项、商品住房竣工面积回归系数、经济增长率回归系数、土地交易价格回归系数、商品住房销售价格回归系数、贷款利率回归系数，基于需求和供给模型的聚类分析变量是需求聚类变量和供给聚类变量的综合，聚类结果见表 8、9。

综合观察表 7、8、9，并的传统行政划分的中国东、中、西部相对比（见表 10），可以发现，聚类分析的结果与传统行政划分的中国东、中、西部地区非常的接近，这为我们将中国房地产市场划分为东、中、西部地区提供了量化支持，同时，也为政府进行分区域差别化宏观调控提供了支持。在之后的房价模型中，我们将以此聚类分析划分的东、中、西部地区为依据，进行分区域研究和探讨。

表 7 基于需求模型的房地产市场区域聚类分析结果表

类一	辽宁、吉林、黑龙江、河北、山东、北京、天津、山西、河南、内蒙古
类二	上海、安徽、江西、湖北、江苏、浙江、福建

类三	湖南、海南、广东、重庆、贵州、广西、陕西、宁夏、新疆、四川、云南、青海、甘肃
----	--

表 8 基于供给模型的房地产市场区域聚类分析结果表

类一	吉林、黑龙江、辽宁、上海、江苏、天津、浙江、安徽、山东、北京
类二	内蒙古、广西、山西、湖南、海南、河北、河南、江西、广东、福建、湖北
类三	重庆、四川、陕西、甘肃、青海、贵州、云南、宁夏、新疆

表 9 基于需求、供给模型的房地产市场区域聚类分析结果表

类一	辽宁、吉林、黑龙江、山东、河北、北京、天津
类二	甘肃、青海、四川、云南、重庆、贵州、宁夏、新疆、陕西
类三	山西、河南、内蒙古、湖南、海南、广东、广西、湖北、福建、江苏、浙江、上海、安徽

表 10 中国东、中、西部省市划分表

东部	辽宁、北京、天津、河北、山东、江苏、上海、浙江、福建、广东、广西、海南
中部	山西、内蒙古、吉林、黑龙江、安徽、江西、河南、湖北、湖南
西部	陕西、甘肃、青海、宁夏、新疆、四川、重庆、云南、贵州、西藏

5.1.4 需求、供给模型结论

本文在对房地产供给和需求模型结果进行系统分析的基础上，主要以中国房地产市场供需理论机制与实证结论相违背作为切入点，得出如下有突破性和创新性的结论。

贷款利率——从供给实证结果可以知道，我国东部经济发达地区的房地产市场供给存在着非理性投资和过多的房地产“泡沫”，贷款利率调整对于房地产商投资行为的调整作用甚微，有必要继续加大东部地区的贷款利率调整幅度和时间跨度，实行持续性的宏观调控；而对于西部地区，从实证结果可知，贷款利率的调整对于人们购买住房需求存在着显著的负相关作用，考虑到西部地区经济尚不发达，如果频繁提高西部地区的贷款利率将不利于西部地区的房地产行业乃至经济的发展。综合考虑东、中、西部地区的贷款利率调整效应，对东、中西部地区实行差别化的住房抵押贷款政策显得很有必要。

商品住房销售价格——从需求实证结果来看，我国绝大部分地区的商品住房销售价格与需求呈现正相关关系，即人们购买房屋的动机是出于投机，将闲置的钱投资于房地产行业谋取资本增值，这也折射了我国房地产行业过热的现状，出现了非理性的投资和需求现象，市场这只看不见的手的调节作用无法正常体现，需要政府进行有力的宏观政策调控。

城市化率——本文实证研究结果与理论恰恰相反，实证结果表明城市化水平

与商品住房需求呈现负相关关系，它潜在折射出我国城市化水平可能存在水分，城市化的衡量指标应为城市常住人口与总人口的比例。正如我国的现实，将近47.5%的城市化推进中，有2亿左右的农民工没有正式纳入城市化，因此，扎实稳健推进我国城市化进程，抓住城市化发展契机来推动我国房地产行业的发展意义重大。

土地价格——在我国房地产供给市场上，土地价格对住房供给量的影响与理论相违背，全国30个省市土地交易价格对商品住房供给量呈现正相关关系，这说明我国土地交易市场对供给量的作用机制不能很好地体现，我国土地交易市场有待进一步完善，为了使得土地定价能够反映资源的稀缺性，取消土地双轨制是我国政府部门目前应该努力的方向。

5.2 房价模型与求解

文献中已有的大量关于房价的模型多数采用时间序列的相关知识，如 *ARMA* 或 *ARIMA* 模型，且回归分析结果较好，但却不能反映出影响价格变动的因素。而其他一些文献，研究了少数几个变量对房价的影响，但采用的模型多为简单的一元或多元分析，而这些模型所需的条件较为严格，且预测的效果有一定的局限性。特别是涉及到政策类变量时，多数模型无法量化的进行分析。

本部分在注重考虑房价自身的变动特点的基础上，结合前面的供给、需求因素和区域性差异，以及在我国有突出作用的国家宏观调控的相关政策分别建立东部、中部和西部的房价的向量自回归 (*VAR*) 的预测模型。同时，我们还加入了脉冲响应函数和方差分解，在得到各因素的总体预测情况的基础上，明晰各个变量对房价的作用效果以及贡献率。

5.2.1 房价模型的建立

1. *VAR* 模型及脉冲响应函数的理论分析

建立脉冲响应前一般先要对数据进行向量自回归的建模。当所有检验序列均服从同阶单整，可以构造 *VAR* 模型。所谓 *VAR* 模型是采用多方程联立的形式，它不以经济理论为基础，在模型的每一个方程中，内生变量对模型的全部内生变量的滞后值进行回归，从而估计全部内生变量的动态关系，而不带任何约束条件。*VAR* 模型对于相互联系的时间序列变量系统是有效的预测模型，同时，向量自回归模型也被频繁地用于分析不同类型的随机误差项对系统变量的动态影响。

含有 N 个变量的滞后 k 期的 *VAR* 模型表述为：

$$Y_t = c + \Pi_1 Y_{t-1} + \Pi_2 Y_{t-2} + \cdots + \Pi_k Y_{t-k} + \mu_t, \quad \mu_t \sim N(0, \Omega) \quad (6)$$

其中，

$$Y_t = (y_{1t}, y_{2t}, \cdots, y_{Nt})'$$

$$c = (c_1, c_2, \cdots, c_N)'$$

$$\Pi_j = \begin{bmatrix} \pi_{11,j} & \pi_{12,j} & \cdots & \pi_{1N,j} \\ \pi_{21,j} & \pi_{22,j} & \cdots & \pi_{2N,j} \\ \vdots & \vdots & \ddots & \vdots \\ \pi_{N1,j} & \pi_{N2,j} & \cdots & \pi_{NN,j} \end{bmatrix}, j = 1, 2, \cdots, k$$

$$\mu_t = (\mu_{1t}, \mu_{2t} \cdots \mu_{Nt})'$$

Y_t 为 $N \times 1$ 阶时间序列列向量, c 为 $N \times 1$ 阶常数项列向量, Π_j 为 $N \times N$ 阶参数矩阵, $\mu_t \sim N(0, \Omega)$ 为 $N \times 1$ 阶随机误差列向量, 其中每一个元素都是非自相关的, 但不同方程对应的随机误差项之间可能存在相关性。

而脉冲响应函数是描述一个内生变量对误差冲击的反应, 即在随机误差项上施加一个标准差大小的冲击后对内生变量的当期值和未来值所带来的影响。由于在 VAR 模型中, 所有的变量之间都有一定的相关性。因此, 其中任何一个变量的冲击不仅仅影响自身的变化, 还会影响其他变量的变化。

脉冲响应函数的推导较为复杂。一般地, 用 q 的函数来描述冲击情况:

$$c_{ij}^{(q)} = \frac{\partial y_{i,t+q}}{\partial u_{jt}}, q = 0, 1, 2, \dots, t = 1, 2, \dots, T \quad (7)$$

上式描述了在时期 t 下, 其他变量和早期变量不变的情况下 $y_{i,t+q}$ 对 y_{jt} 的一个冲击反映, 称其为一个脉冲函数。

用矩阵可以表示为

$$C_q = \frac{\partial Y_{t+q}}{\partial u_t} \quad (8)$$

即 C_q 的第 i 行和第 j 列元素等于时期 t 第 j 个变量的扰动项增加一个单位, 而其它时期的扰动项为常数时, 对时期 $t+q$ 的第 i 个变量的影响情况。

2. 模型变量选择和数据处理

根据西方经济学中的市场均衡理论, 供给和需求的均衡点即均衡价格, 于是通过得出的供求和需求模型确定房价。但在我国特定的经济背景下, 国家的宏观调控的干预, 我国的房价还受相关政策的影响。特别是近年来, 在美国次贷危机引发的全球性金融危机的背景下, 我国的房地产相关政策对房价有很强的引导作用。显然, 因变量选择商品房销售价格, 而自变量选择的变量主要分为两大类: 一是经济变量, 在前面模型的基础上选择分别代表需求量和供给量的房地产开发企业商品房屋销售面积和房地产开发企业的房屋竣工面积变量; 二是政策变量, 主要包括一年期贷款利率、广义货币供应量、存款准备金率和税收。

本部分以 2005 年 1 月到 2009 年 12 月 60 个月的各变量数据为基础, 数据主要来自国家统计局网站和《统计年鉴》以及巨星金融平台。随后, 根据供给需求部分对区域进行的划分, 分别收集了东部、中部和西部地区的各变量数据。其中, 部分指标缺少 1 月份的对数据, 用 2、3 月的数据平均进行替代。

为了确保符合 VAR 模型的基础数据特征, 先观察原始数据的统计特征, 由于变量的基础量纲不同, 分别绘制线图描述:

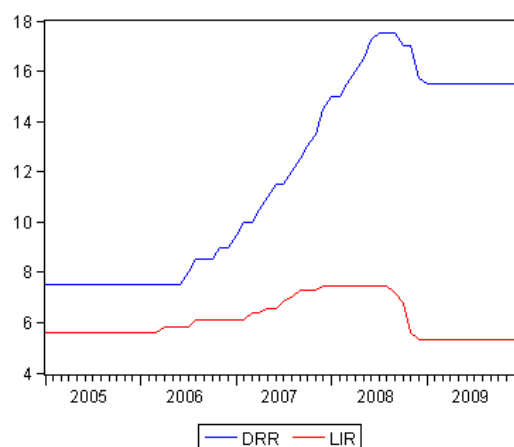


图 19 贷款和存款准备金率

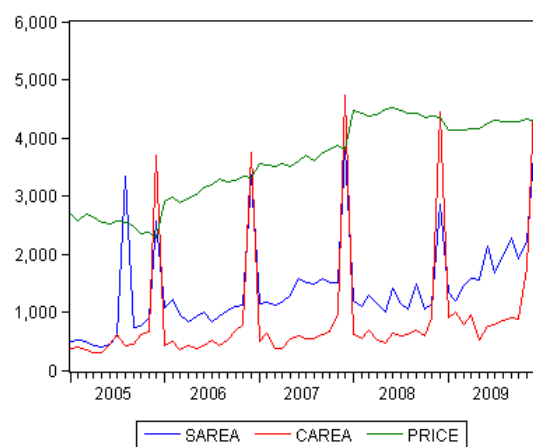


图 20 销售价格、销售面积以及竣工面积

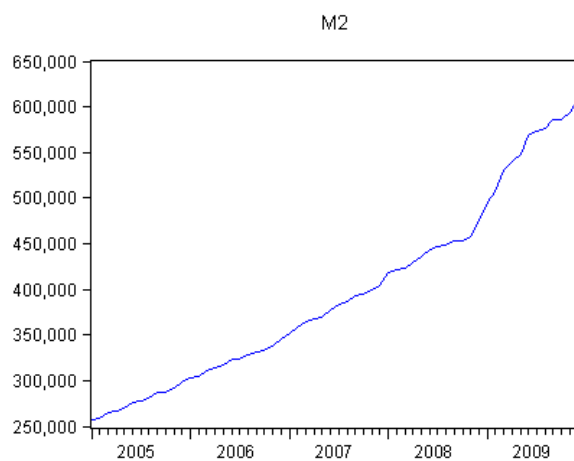


图 21 广义货币供应量

由图 19、20、21 可知销售面积和竣工面积有一定的季节影响，为此用 *Eviews* 作 $X12$ 季节调整，之后为了减少变量数据的波动程度，再对数据波动较大的销售面积、竣工面积、广义货币供应量和商品房销售价格取对数处理。处理后的数据为：

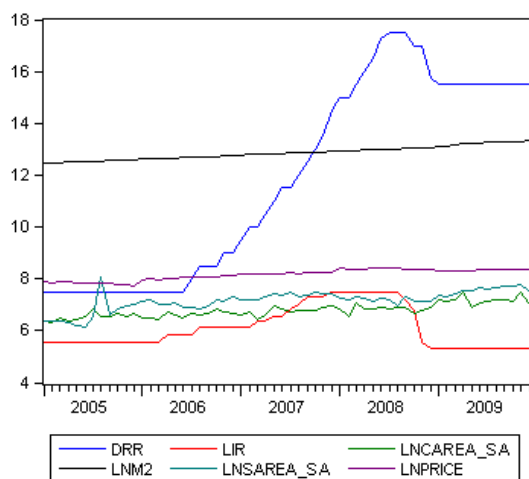


图 22 处理后数据

5.2.2 房价模型的分析 and 求解（以西部地区为例）

1. VAR 模型的建立

(1) 平稳性检验

VAR 模型建立在变量平稳或存在协整关系的基础上，否则会出现伪回归现象，因此需先用单位根检验法对变量进行平稳性检验。

表 11 平稳性检验结果

变量名称	ADF 统计量值	1%临界值	5%临界值	10%临界值	P 值	Lag Length	结论
LNPRICE	-1.0307	-3.5461	-2.9117	-2.5936	0.7367	0	非平稳
DLNPRICE	-8.6666	-3.5482	-2.9126	-2.5940	0.0000	0	平稳
LNSAREA	0.6335	-2.6054	-1.9465	-1.6132	0.8503	1	非平稳
DLNSAREA	-11.1736	-2.6054	-1.9465	-1.6132	0.0000	0	平稳
LNCAREA	-1.2870	-3.5504	-2.9135	-2.5945	0.6298	2	非平稳
DLNCAREA	-9.0835	-3.5504	-2.9135	-2.5945	0.0000	1	平稳
DRR	-0.8088	-3.5482	-2.9126	-2.5940	0.8090	1	非平稳
DDRR	-4.4797	-3.5482	-2.9126	-2.5940	0.0001	0	平稳
LNM2	1.3887	-3.5461	-2.9117	-2.5936	0.9988	0	非平稳
DLNM2	-6.6569	-3.5482	-2.9126	-2.5940	0.0000	0	平稳
LIR	-1.4495	-3.5482	-2.9126	-2.5940	0.5518	1	非平稳
DLIR	-4.5915	-3.5482	-2.9126	-2.5940	0.0004	0	平稳

由表 11 中的结果可以看出：所有变量的 ADF 统计量均显著大于 1%、5% 及 10% 时的临界值，所以接受存在单位根的原假设，原序列是非平稳的。因此对数据进行一阶差分，得到一阶差分变量的 ADF 统计量的均小于显著性水平 1%、5% 及 10% 时的临界值，说明至少可以在 99% 的显著性水平下拒绝存在单位根的原假设，及各序列都是平稳的，符合 VAR 模型的进入该系统的内生变量都是平稳的时间序列的要求。

(2) 滞后阶数的确定

滞后阶数的确定在 $VAR(p)$ 模型中很重要，一方面要使滞后阶数足够大，以便完整的反应出模型的动态特征；但另一方面随着滞后期的增大，待估计参数数

量增多，模型的自由度减少。为此，这里使用实际研究中广泛应用的 AIC 和 SC 准则来进行确定，如表 12：

表 12 VAR 滞后期的选择性检验

Lag	LogL	LR	FPE	AIC	SC	HQ
0	257.2885	NA	3.66e-12	-9.306983	-9.085985*	-9.221753
1	314.0489	98.80502	1.71e-12*	-10.07588	-8.528897	-9.479272*
2	340.7339	40.52169	2.54e-12	-9.730885	-6.857908	-8.622890
3	375.4689	45.02681	3.03e-12	-9.684032	-5.485065	-8.064655
4	423.8100	51.92201*	2.51e-12	-10.14111	-4.616157	-8.010354
5	461.3050	31.94015	3.90e-12	-10.19648*	-3.345536	-7.554340

其中 AIC 和 SC 准则要求它们的值越小越好，由于 SC 是逐渐变小的，为此参看 AIC 准则，滞后 3 期时 $AIC = -9.684032$ 是最小的，所以建立 $VAR(3)$ 模型为：

$$\begin{pmatrix} \Delta \ln price_t \\ \Delta \ln sarea_t \\ \Delta \ln carea_t \\ \Delta drr_t \\ \Delta \ln m2_t \\ \Delta lir_t \end{pmatrix} = A_0 + A_1 \begin{pmatrix} \Delta \ln price_{t-1} \\ \Delta \ln sarea_{t-1} \\ \Delta \ln carea_{t-1} \\ \Delta drr_{t-1} \\ \Delta \ln m2_{t-1} \\ \Delta lir_{t-1} \end{pmatrix} + A_2 \begin{pmatrix} \Delta \ln price_{t-2} \\ \Delta \ln sarea_{t-2} \\ \Delta \ln carea_{t-2} \\ \Delta drr_{t-2} \\ \Delta \ln m2_{t-2} \\ \Delta lir_{t-2} \end{pmatrix} + A_3 \begin{pmatrix} \Delta \ln price_{t-3} \\ \Delta \ln sarea_{t-3} \\ \Delta \ln carea_{t-3} \\ \Delta drr_{t-3} \\ \Delta \ln m2_{t-3} \\ \Delta lir_{t-3} \end{pmatrix} + \begin{pmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \\ \mu_4 \\ \mu_5 \\ \mu_6 \end{pmatrix} \quad (9)$$

其中 A_1, \dots, A_5 为 6×6 的系数矩阵， $t=1, 2, \dots, T$ 。

(3) 模型稳定性检验

对于 VAR 模型而言，稳定性事指把一个脉冲冲击施加在 VAR 模型中某一个方程的过程上，随着时间的推移，分析这个冲击是否消失。如果消失说明系统是稳定的，否则系统不稳定。于是有系统检验的单位根的图为：

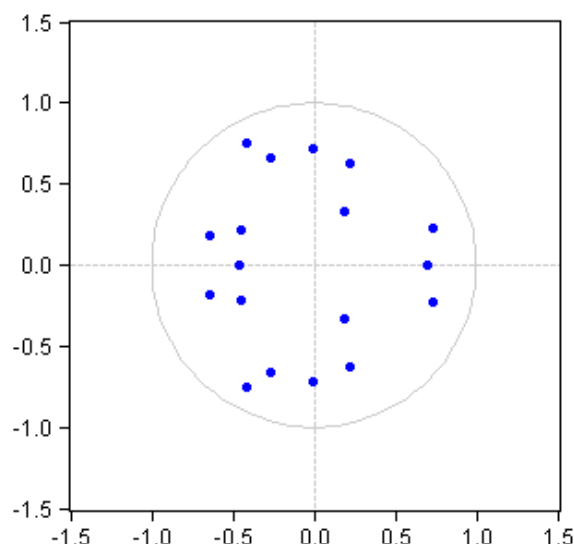


图 23 VAR 模型单位根示意图

由图 23 看出 VAR 模型对应的特征方程的所有根均在单位圆以内，所以建立的 $VAR(3)$ 模型是稳定的，即表示该系统稳定。此时，对变量间的关系建立 VAR 模

型才有意义，并且为进一步的脉冲响应和方差分解奠定了基础。

(4) *VAR* 模型的参数估计

在选择合适的滞后期后且进行了稳定性检验后，可对 *VAR* 模型进行参数估计，并得出相应的参数估计表，见附录五和表 13：

表 13 *VAR* 各方程的检验结果

R-squared	0.655333	0.430197	0.455263	0.611964	0.234066	0.471522
Adj. R-squared	0.487657	0.152995	0.190256	0.423190	-0.138550	0.214425
Sum sq. resids	2.191097	1.284678	1.142616	0.002194	0.082890	2.946521
S.E. equation	0.243349	0.186336	0.175731	0.007701	0.047331	0.282198
F-statistic	3.908333	1.551927	1.717929	3.241775	0.628169	1.834021
Log likelihood	11.28602	26.23507	29.51631	204.6639	102.9761	2.991785
Akaike AIC	0.275499	-0.258395	-0.375582	-6.630852	-2.999148	0.571722
Schwarz SC	0.962672	0.428778	0.311590	-5.943679	-2.311975	1.258895
Mean dependent	0.142857	-0.004821	0.012371	0.014761	0.008574	0.021157
S.D. dependent	0.339977	0.202467	0.195288	0.010140	0.044358	0.318390
Determinant resid covariance (dof adj.)		4.16E-13				
Determinant resid covariance		3.46E-14				
Log likelihood		391.1106				
Akaike information criterion		-9.896807				
Schwarz criterion		-5.773770				

由结果可以看出，贷款利率和广义货币供应量拟合方程拟合效果较好，而价格方程的解释度一般，且各方程的检验统计量通过检验，总体拟合程度较好。

2. 脉冲响应及方差分解分析

(1) 脉冲响应函数

为了更好的考察需求量、供给量以及各政策变量之间的动态关系，我们可以从脉冲响应函数和方程分析上进行考察。通过格兰杰因果关系检验发现，各变量因果分析结果并不是很明显，无法为脉冲提供辅助性分析。因此，根据模型因变量和自变量的研究关系进行脉冲分析。得到的脉冲响应曲线如下：

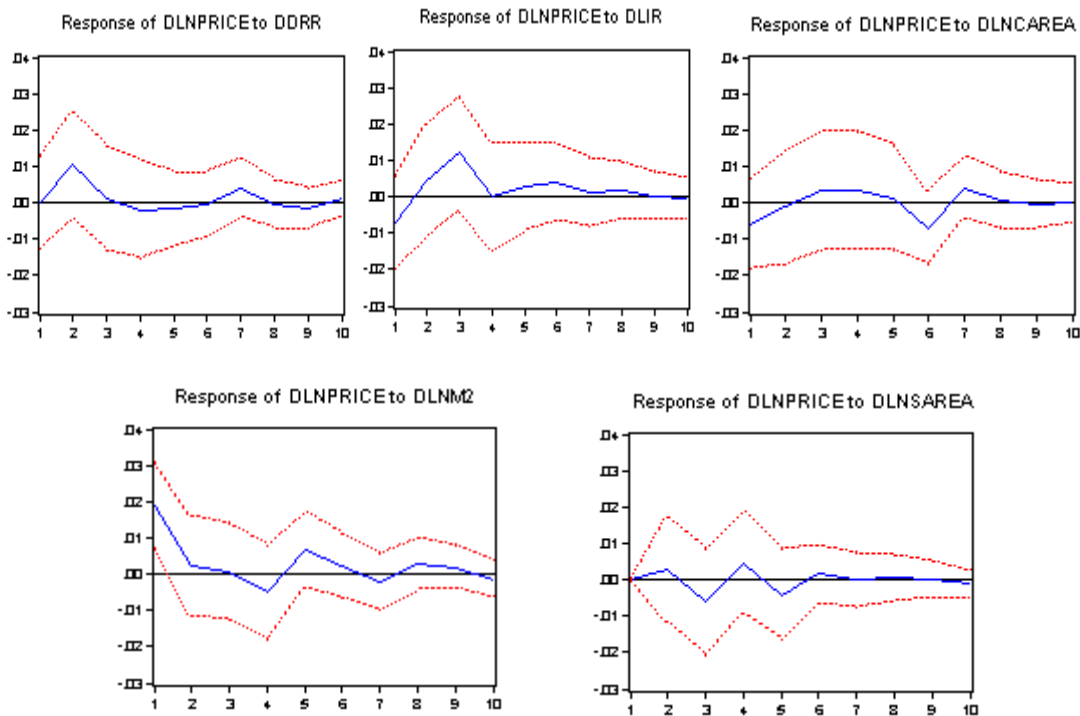


图 24 其它各变量对销售价格的脉冲响应曲线

通过脉冲响应曲线的图形可以看出：

当对本期存款准备金率施加一个正的冲击后，商品房销售价格滞后一期，即在第二期达到最大，随后在第三期恢复，然后趋于平稳，也就是说存款准备金率对房价有滞后一期的正向影响，且持续期较短。实际情况中，存款准备金率提高后，银行的借贷数额减少，房价应该下降。但房价本身有一个上升的趋势，可见第二期的上升不是利率作用的，而是房价上涨的固有趋势。而存款准备金率的提高于第二期抑制了房价的上涨趋势，保证了房价的平稳性。

当对一年期存款利率施加一个正的冲击后，商品房销售价格先下降，然后在第三期上升到最高点后回落，最后趋于稳定。在实际中，利率上涨，无论是房地产的开发商还是购买住房的消费者在投资房地产行业时的成本都增高，即供给和需求都有所下降。特别是消费者的负担过重，购买预期下降，于是房价下降。但房价本身的上升趋势使得房价回调。

当对广义货币量施加一个正的冲击后，商品房的销售价格在第一期即达到最高，但随后即迅速回落，然后趋于小幅度波动。联系实际情况，广义货币供应量增加，表示货币政策放宽，周转资金充足，消费者更愿意投资住房，推动了房价的上涨。

当对销售面积施加一个正的冲击后，商品房的销售价格在 1-5 期震荡波动，随后稳定不变。可见，消费者的需求对房价的影响是由预期决定的。处于“买涨不买跌”以及成本提高等因素综合作用影响房价，影响方向变动复杂。

当对竣工面积施加一个正的冲击后，商品房的销售价格在第一期下降，随后缓慢上升，到第四期后再逐渐下降到最低点，随后在再次上升后趋于稳定。理论上，供应量增加了，价格会有所下降。但是房价主要还是看消费者的预期以及房价的整体发展趋势，为此之后房价还会缓步上升。

(2) 方差分解

进一步分析每个结构冲击对内生变量变化的贡献度，因此采用方差分解方法。用 Eviews 进行操作得到结构图形为：

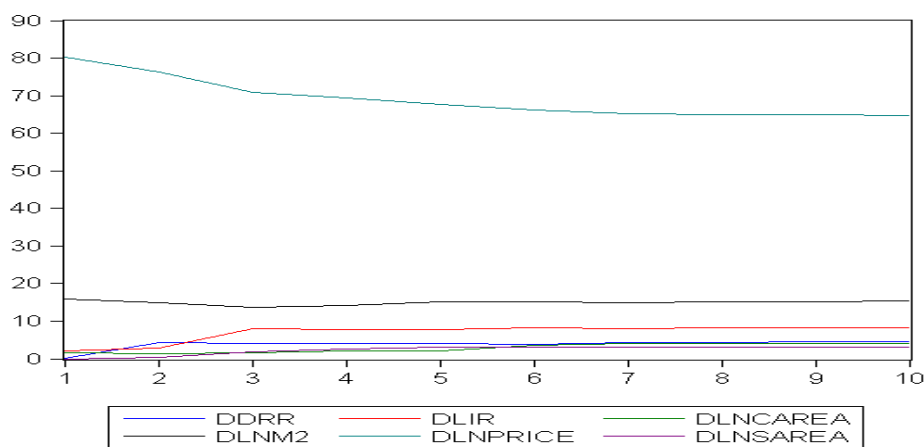


图 25 房价的方差贡献率

由图 25 可以看出供给量、需求量以及政策变量对房价变动的方差贡献程度：房价自身的贡献率始终显著大于其他变量的贡献率，虽然在第二期后其他变量对其方差的贡献有显著的提高，但是仍不是构成房价的影响因素。由此可见，我国的房价增长由投机等非正常因素影响较大，政策对房价有一定的抑制作用，但仅能放缓房价的上涨，却不能阻止其上涨的趋势。

5.2.3 房价预测模型（以西部地区为例）

根据前面建模的分析，利用 $VAR(3)$ 模型建立西部地区房价的预测模型为：

$$\Delta \ln price_t = \begin{pmatrix} 0.043196 \\ 0.009528 \\ -0.012699 \\ 0.847183 \\ -0.202238 \\ 0.010199 \end{pmatrix}' \begin{pmatrix} \Delta drr_{t-1} \\ \Delta lir_{t-1} \\ \Delta \ln carea_{t-1} \\ \Delta \ln m2_{t-1} \\ \Delta \ln price_{t-1} \\ \Delta \ln sarea_{t-1} \end{pmatrix} + \begin{pmatrix} -0.023148 \\ 0.025211 \\ 0.009214 \\ 0.402611 \\ -0.067135 \\ -0.017428 \end{pmatrix}' \begin{pmatrix} \Delta drr_{t-2} \\ \Delta lir_{t-2} \\ \Delta \ln carea_{t-2} \\ \Delta \ln m2_{t-2} \\ \Delta \ln price_{t-2} \\ \Delta \ln sarea_{t-2} \end{pmatrix} + \begin{pmatrix} -0.015232 \\ 0.043133 \\ 0.009369 \\ -1.660323 \\ 0.070347 \\ 0.008132 \end{pmatrix}' \begin{pmatrix} \Delta drr_{t-3} \\ \Delta lir_{t-3} \\ \Delta \ln carea_{t-3} \\ \Delta \ln m2_{t-3} \\ \Delta \ln price_{t-3} \\ \Delta \ln sarea_{t-3} \end{pmatrix} + 0.016160 \quad (10)$$

上式可以充分考察供给和需求以及政策变量的变动对房价的影响。例如：当一年期贷款利率在 0 和 1 期上升的一个单位分别增加第二期的对数价格 0.025 和 0.0095 个单位，可见贷款利率对价格有滞后两期的影响。这与前面的脉冲图像也是吻合的。虽然方程的系数显著性并不强，但考察模型的拟合效果如下图：

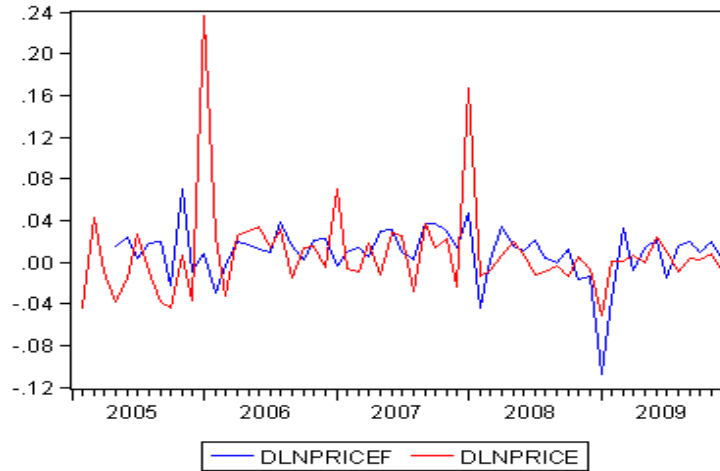


图 26 $VAR(3)$ 拟合效果

可见 $VAR(3)$ 预测的差分项的趋势是基本一致的，且波动在 ± 0.04 个单位内，拟合效果较好。

VAR 模型的静态预测的精确度高，且考虑充分各可能影响变量，引入多个变量的综合作用。联系实际可以发现，各类政策以及供求方面对于房价的影响都存在一定的滞后性，这很大程度的取决于消费者对信息的接纳和反馈的程度与速度。而模型中变量的不同滞后期的不同系数，可以充分体现该变量的滞后影响性。而房价本身的规律在模型中同样通过自相关进行了体现。模型具有一定的分析和实践效用。

5.2.4 东部和中部地区的房价预测模型

根据西部地区的模型构建方法，分别建立东部和中部地区的房价预测模型。根据滞后期的检验，这两个区域均适合采用 VAR(2) 的预测模型。

东部地区的 VAR(2) 房价模型为：

$$\Delta \ln price_t = \begin{pmatrix} 0.002223 \\ 0.016190 \\ -0.017972 \\ 0.461403 \\ -0.212601 \\ 0.007760 \end{pmatrix}' \begin{pmatrix} \Delta drr_{t-1} \\ \Delta lir_{t-1} \\ \Delta \ln carea_{t-1} \\ \Delta \ln m2_{t-1} \\ \Delta \ln price_{t-1} \\ \Delta \ln sarea_{t-1} \end{pmatrix} + \begin{pmatrix} 0.002474 \\ 0.011575 \\ -0.034440 \\ 0.476036 \\ -0.232420 \\ -0.007140 \end{pmatrix}' \begin{pmatrix} \Delta drr_{t-2} \\ \Delta lir_{t-2} \\ \Delta \ln carea_{t-2} \\ \Delta \ln m2_{t-2} \\ \Delta \ln price_{t-2} \\ \Delta \ln sarea_{t-2} \end{pmatrix} + 0.003932 \quad (11)$$

拟合效果如图 27：

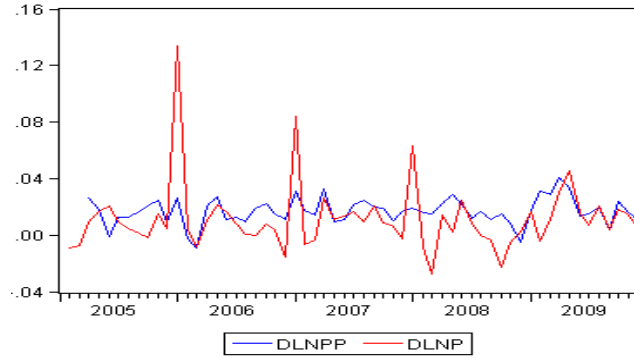


图 27 VAR(2) 拟合效果

中部地区的 VAR(2) 房价模型为：

$$\Delta \ln price_t = \begin{pmatrix} -0.007567 \\ 0.052242 \\ -0.066572 \\ 0.415887 \\ -0.299802 \\ -0.001716 \end{pmatrix}' \begin{pmatrix} \Delta drr_{t-1} \\ \Delta lir_{t-1} \\ \Delta \ln carea_{t-1} \\ \Delta \ln m2_{t-1} \\ \Delta \ln price_{t-1} \\ \Delta \ln sarea_{t-1} \end{pmatrix} + \begin{pmatrix} -0.014941 \\ 0.032924 \\ -0.030285 \\ -0.359286 \\ -0.268346 \\ -0.019652 \end{pmatrix}' \begin{pmatrix} \Delta drr_{t-2} \\ \Delta lir_{t-2} \\ \Delta \ln carea_{t-2} \\ \Delta \ln m2_{t-2} \\ \Delta \ln price_{t-2} \\ \Delta \ln sarea_{t-2} \end{pmatrix} + 0.020039 \quad (12)$$

拟合效果如图 28：

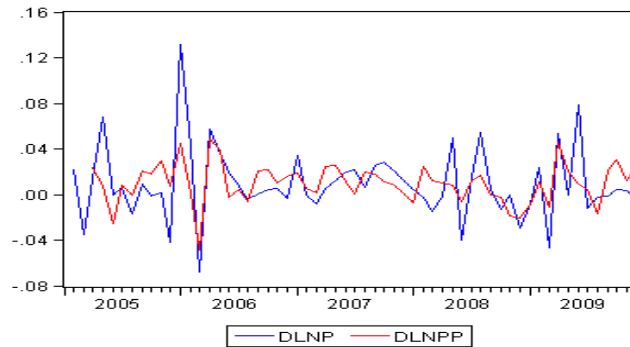


图 28 VAR(2) 拟合效果

通过对不同区域的预测模型和拟合图形的比较发现：三个区域的变量滞后期的期数选择不同。西部地区的存在影响的滞后期数较长，这主要反应了西部地区房价变动敏感度较低，与西部的经济发展水平相一致，而经济较发达的东部和中部地区的指标影响的滞后期较短。同时，三个区域对同一政策的反应情况不同，如存款准备金率前面的系数的差异。这都是地区发展特征的结果。

通过对房价模型的分析进一步的验证了区域划分的必要性。同时也表明：国家现在提出的不同经济发展状况的城市采用不同的房价调控政策是明智的，同一政策的调控是不可取的。

5.3 房地产行业关联度与发展态势投入产出模型与求解

在本部分，我们主要基于全国 2007 年投入产出表建立投入产出模型对房地产行业的关联度与发展态势进行分析，借鉴中国统计年鉴上的 19 行业划分标准，将 42×42 行业投入产出表中的相关行业进行合并处理，得到 19×19 行业投入产出表。

5.3.1 投入产出模型理论分析

1. 直接消耗系数矩阵

直接消耗系数是行业 j 每生产一单位的总产品对行业 i 产品的直接消耗数量，记作 a_{ij} ，计算公式如下：

$$a_{ij} = \frac{x_{ij}}{q_j}, \quad i, j = 1, 2, \dots, n \quad (13)$$

其中， x_{ij} 为行业 j 为生产一定数量的总产品所消耗的行业 i 产品的数量； q_j 为行业 j 的总产出。直接消耗系数矩阵为：

$$\mathbf{A} = (a_{ij})_{n \times n} \quad (14)$$

2. 自产率

在直接消耗系数 $a_{ij} = \frac{x_{ij}}{q_j}$ 中， x_{ij} 既包括本地区产出，也包括进口和其他地区

调入的部分。本文考察的是本地区各行业吸纳劳动力的能力，故需引入自产率来进行调整。根据投入产出表可知：总产出+行业进口额+行业调入额=行业中间使用+行业最终使用，故自产率=总产出/(总产出+行业进口额+行业调入额)=总产出/(行业中间使用+行业最终使用)。

设行业 i 自产率为 m_i ，则调整后的直接消耗系数为 a_{ij}^* ，调整后直接消耗矩阵

$$\mathbf{A}^* = \mathbf{M} \times \mathbf{A} \quad (15)$$

其中 $\mathbf{M} = \text{diag}(m_1, m_2, \dots, m_n)$ 。

3. 里昂惕夫逆矩阵

里昂惕夫逆矩阵计算公式为

$$\mathbf{C} = (\mathbf{I} - \mathbf{A}^*)^{-1} = (c_{ij})_{n \times n} \quad (16)$$

4. 完全消耗系数矩阵

完全消耗系数反映生产单位第 j 种最终产品对第 i 种产品的直接与全部间接消耗量之和,完全消耗系数矩阵的计算公式为

$$\mathbf{B} = (\mathbf{I} - \mathbf{A}^*)^{-1} - \mathbf{I} = (b_{ij})_{n \times n} \quad (17)$$

5. 产出乘数

行业 j 的总产出乘数为

$$F_j = \sum_{i=1}^n c_{ij} \quad (18)$$

含义为配合行业 j 最终使用增加 1 个单位, 其他行业最终使用均不增加, 各个行业总产出的增加额之总和(也即社会总产出的增加额)。它反映了行业 j 带动国民经济规模的能力, 也能间接反映行业 j 吸纳就业的能力。

6. 影响力系数

行业 j 的影响力系数为

$$\lambda_j = \frac{\sum_{i=1}^n c_{ij}}{\frac{1}{n} \sum_{j=1}^n \sum_{i=1}^n c_{ij}} \quad (19)$$

表明某个行业对国民经济各部门的影响力的相对水平。当 $\lambda_j > 1$, 表明该行业的影响力超过各行业的平均水平; 当 $\lambda_j < 1$, 表明该行业的影响力低于各行业的平均水平。

7. 直接分配系数矩阵

直接分配系数矩阵反映第 i 种产品分配给第 j 种产品生产消耗数量与第 i 种产品国内总产出的比重, 记作 r_{ij} , 计算公式为

$$r_{ij} = \frac{x_{ij}}{x_i} \quad (20)$$

其中, x_{ij} 为行业 i 分配给行业 j 使用的产出量; x_i 为行业 i 的总产出。直接分配系数矩阵为

$$\mathbf{R} = (r_{ij})_{n \times n} \quad (21)$$

8. 完全分配系数矩阵

完全分配系数反映第 i 行业直接和全部间接分配给行业 j 生产消耗数量与第 i 行业国内总产出的比重，完全分配系数矩阵为

$$\mathbf{D} = (\mathbf{I} - \mathbf{R})^{-1} - \mathbf{I} = (d_{ij})_{m \times n} \quad (22)$$

9. 行业贡献率

行业 j 的贡献率为

$$\delta_j = \frac{\Delta y_j}{\Delta y} \quad (23)$$

其中， Δy_j 为行业 j 的增加值增量， Δy 为国内生产总值增量。贡献率越大表明该行业对经济增长的贡献和作用越大，反之则小。

5.3.2 投入产出模型求解

根据前文设计的研究方法，计算出各个指标的数值，分析结果如下：

1. 房地产行业与其他行业关联度

表 14 房地产行业对各行业消耗系数表

行业	直接消耗系数	排名	完全消耗系数	排名
农林牧渔业	0.0000	19	0.0137	7
采矿业	0.0001	18	0.0123	9
制造业	0.0538	1	0.1849	1
电力、燃气及水的生产和供应业	0.0043	8	0.0198	4
建筑业	0.0121	4	0.0128	8
交通运输、仓储和邮政业	0.0040	9	0.0141	6
信息传输、计算机服务和软件业	0.0027	12	0.0053	13
批发和零售业	0.0029	11	0.0092	11
住宿业和餐饮业	0.0106	5	0.0156	5
金融	0.0247	2	0.0331	2
房地产业	0.0089	6	0.0116	10
租赁和商务服务业	0.0182	3	0.0229	3
科学研究、技术服务和地质勘查业	0.0012	13	0.0027	15
水利、环境和公共设施管理业	0.0002	17	0.0006	18
居民服务和其他服务业	0.0061	7	0.0085	12
教育	0.0004	14	0.0009	16
卫生、社会保障和社会福利事业	0.0003	16	0.0007	17
文化、体育和娱乐业	0.0035	10	0.0045	14

公共管理和社会组织	0.0004	15	0.0005	19
-----------	--------	----	--------	----

注：房地产行业对农林牧渔业的直接消耗系数为 0.0000，表明其对农林牧渔业的直接消耗可以忽略。

观察表 14 中的直接消耗系数排名，制造业、金融业、租赁和商务服务业、建筑业、住宿业和餐饮业对房地产行业的发展投入相对较多，也即房地产行业的发展对这些行业直接以依赖性强，而观察完全消耗系数排名发现，制造业、金融业、租赁和商务服务业、电力、燃气和水的生产和供应业、住宿业和餐饮业、交通运输、仓储和邮政业、建筑业对旅游业的全部间接投入相对较多，其中，制造业、电力、燃气和水的生产和供应业、建筑业都属于第二产业，可见，房地产行业对第二产业的依赖性很强；同时，无论是直接消耗系数排名还是间接消耗系数排名，金融业均位居第二，说明金融业对于我国房地产业的发展具有不可忽视的作用。

表 15 房地产行业对各行业分配系数表

行业	直接分配系数	排名	完全分配系数	排名
农林牧渔业	0.0000	19	0.0059	13
采矿业	0.0001	18	0.0117	7
制造业	0.0020	10	0.0084	11
电力、燃气及水的生产和供应业	0.0019	12	0.0109	8
建筑业	0.0029	9	0.0031	16
交通运输、仓储和邮政业	0.0019	12	0.0081	12
信息传输、计算机服务和软件业	0.0042	7	0.0089	10
批发和零售业	0.0015	14	0.0058	14
住宿业和餐饮业	0.0109	4	0.0172	4
金融	0.0188	2	0.0268	2
房地产业	0.0089	6	0.0120	6
租赁和商务服务业	0.0274	1	0.0361	1
科学研究、技术服务和地质勘查业	0.0033	8	0.0094	9
水利、环境和公共设施管理业	0.0016	13	0.0052	15
居民服务和其他服务业	0.0105	5	0.0156	5
教育	0.0005	15	0.0012	17
卫生、社会保障和社会福利事业	0.0003	17	0.0011	18
文化、体育和娱乐业	0.0161	3	0.0214	3
公共管理和社会组织	0.0004	16	0.0005	19

注：房地产行业对农林牧渔业的直接分配系数为 0.0000，表明其对农林牧渔业的直接分配可以忽略。

观察表 15 中分配系数排名，易知房地产行业分配系数较高的行业是租赁和商务服务业、金融业、文化、体育和娱乐业、住宿业和餐饮业、居民服务和其他服务业等，说明房地产行业的发展有力地带动了这些第三产业的发展。

结合消耗系数表和分配系数表，我们知道，我国房地产行业与金融业、租赁和商务服务业、住宿业和餐饮业之间存在着相互制约、相互影响的关系，无论从消耗角度还是分配角度，它们之间的关联度都很高。

根据投入产出表，我们还可以计算出房地产行业的产出乘数和影响力系数，分别是 1.374 和 0.618，产出乘数 1.374 表示房地产行业带动各个行业生产的能力，该行业最终使用增加 1 个货币单位，由此引发的各个行业产出增加的总额为 1.374，影响力系数 0.618，它小于 1 说明房地产行业对其它行业的影响小于社会平均水平。

2. 房地产行业发展态势分析

基于 2004 至 2009 年 19 行业的增加值数据^①，我们可以对房地产行业的发展态势进行研究，表显示了房地产行业从 2004 年至 2009 年的行业增加值和行业贡献率数据，为了更加直观的分析，图绘制了房地产行业增加值变化趋势图和房地产行业贡献率变化趋势图。观察行业增加值图表可知，房地产行业的增加值呈现稳步升高的态势，只是在 2008 年，增加值增量甚微，原因在于，2008 年国家宏观调控政策以及美国次贷危机作用显现，市场供给方资金链条绷紧，需求方处于明显观望状态，商品房交易量大幅下降，房价上涨幅度放缓，房地产市场景气回落，导致房地产行业 2008 年增加值没有显著提升。同样地，观察房地产行业贡献率变化趋势，从 2004 年至今，房地产行业对于国民经济的贡献一直处于上升的态势，只是在 2008 年由于美国次贷危机的影响出现陡然降低，而 2009 年则快速回升并超过 2007 年度的行业贡献率，行业贡献率达到 13.83%，从长远来看，我国房地产行业在国民经济中的重要性会日益凸显，因此，房地产行业的长期可持续发展研究就显得很有必要。

表 16 房地产增加值及行业贡献率表

	2004 年	2005 年	2006 年	2007 年	2008 年	2009 年
房地产行业增加值（亿元）	8175.83	9526.43	11428.90	14522.15	14635.53	18654.90
房地产行业贡献率（%）	---	5.48	6.04	7.52	0.35	13.83

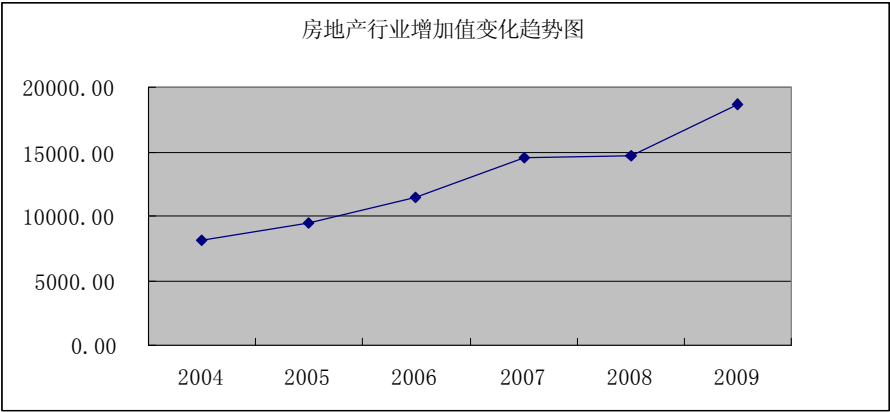


图 29 房地产行业增加值变化趋势图

^①对于 2004 至 2009 年分行业增加值数据，利用 2009 物价指数进行调整，以消除价格变动影响。

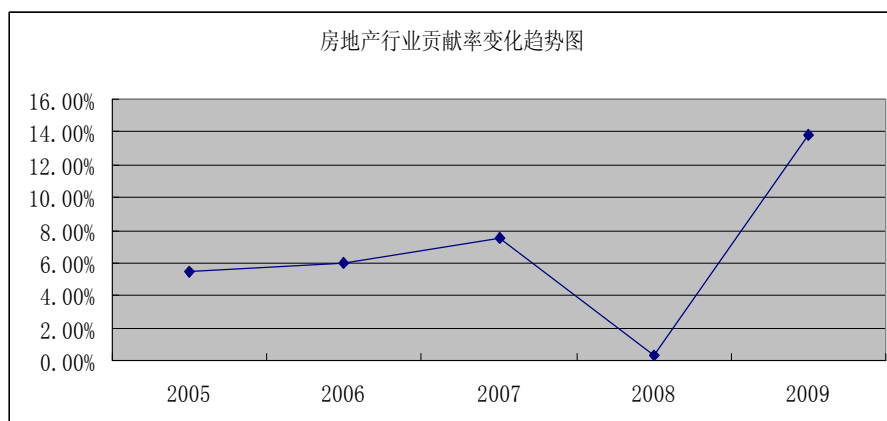


图 30 房地产行业贡献率变化趋势图

5. 4 房地产行业关联度与发展态势空间面板数据模型与求解

经过对国内外文献的研究,我们发现相关学者对房地产行业与其他行业间关联度的专门、定量研究还不是很多,大部分都是一些定性描述和分析,对于房地产行业关系密切的行业和关联程度并没有一致的结论。本文把面板数据模型应用于房地产行业与其他行业的关联度分析上,旨在扩大样本信息和样本容量,提高估计量的有效性和准确性,模型拟合出固定效应的变系数模型后得到了房地产行业于其他行业相关程度的排名,并对其相关性进行了一定的定性分析。

5. 4. 1 面板数据模型的理论分析

1. 面板数据模型的概述

面板数据也称作时间序列与界面混合数据(pooled time series and cross section data),它是截面上个体在不同时间的重复观测数据。面板数据利用双下标变量 $y_{it}(i=1,2,\dots,N;t=1,2,\dots,T)$ 表示,其中 i 对应面板数据中的个体, t 对应面板数据中的不同时间。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, $y_t(t=1,2,\dots,T)$ 是纵剖面上的一个时间序列。

利用面板数据建立模型的好处在于,首先由于观测值得增多,可以增加估计量的抽样精度。其次,对于固定效应模型能得到参数的一直估计量,甚至是有效估计量。最后,面板数据建模比单截面数据建模可以获得更多的动态信息。

面板的一般表现形式为

$$y_{it} = \alpha_i + X'_{it}\beta_i + u_{it}, i=1,2,\dots,N; t=1,2,\dots,T \quad (24)$$

式中, y_{it} 为被解释变量, α 表示截距项, X'_{it} 为 $k \times 1$ 阶解释变量列向量(包括 k 个解释变量), β 为 $k \times 1$ 阶回归系数列向量(包括 k 个回归系数), u_{it} 为随机误差项。

根据截距项和回归系数的不同，面板数据可以表现为以下三种形式：

$$\text{模型一: } y_{it} = \alpha_i + X'_{it}\beta_i + u_{it} \quad (25)$$

$$\text{模型二: } y_{it} = \alpha_i + X'_{it}\beta + u_{it} \quad (26)$$

$$\text{模型三: } y_{it} = \alpha + X'_{it}\beta + u_{it} \quad (27)$$

其中，模型一称为变系数模型，其截距项和回归系数均不同，除了存在个体影响外，截面上也存在变化的结构，因而结构参数在不同横截面上是不同的；模型二称为变截距模型，其回归系数相同但截距项不同，即在截面上个体影响不同，变截距模型又可分为固定影响变截距模型和随机影响变截距模型。以上模型一和模型二又统称为变参数模型。模型三在截面上无个体影响，也无结构变化，称为不变参数模型。

2. 面板数据模型的设定

构造和检验面板数据模型比单独使用截面数据或时间序列数据更加丰富了计量经济学的内容，但是由于面板数据包括二维数据，假如模型的设定不正确会贵结果造成很大的偏差，估计结果与实际可能相差甚远。因此，在建立面板数据模型时一定要正确选择模型的形式，以避免模型的设定偏差，同时改进参数估计的准确性与有效性。

(1) 确定模型的影响形式

面板数据模型分为混合模型、固定效应模型和随机效应模型。这里我们利用 F 统计量来确定模型的影响形式是固定的还是随机的。

F 统计量用来检验对于一组面板数据应该建立混合模型还是个体固定效应模型。混合模型属于约束模型，个体固定效应模型属于非约束模型（ α_i 可以随个体不同）。建立假设

$$H_0: \alpha_i = \alpha \text{ 模型中不同个体的截距相同（即混合模型）}$$

$$H_1: \text{模型中不同个体的截距项 } \alpha_i \text{ 不同（即个体固定效应模型）}$$

F 统计量定义为：

$$F = \frac{(RRS_r - RRS_u) / [(NT - k) - (NT - N - k)]}{RRS_u / (NT - N - k)} \sim F(N, NT - N - k) \quad (28)$$

其中 RRS_r 表示约束模型，即混合模型的残差平方和， RRS_u 表示非约束模型，即个体固定效应模型的残差平方和， k 表示混合模型中回归参数个数。 F 统计量在 H_0 成立条件下服从自由度为 $(N, NT - N - k)$ 的 F 分布。

若用样本计算的 $F \leq F_\alpha(N, NT - N - k)$ ，则接受原假设，建立混合模型；

若用样本计算的 $F > F_\alpha(N, NT - N - k)$ ，则拒绝原假设，建立个体固定效应模型。

(2) 确定模型的形式

按照模型的参数是否可变，面板数据模型又分为变系数模型、变截距模型和不变参数模型。 F 统计量同样可以用于检验回归参数的约束条件是否成立，因此我们再次利用 F 统计量来确定模型的形式。建立假设

$$H_0: \alpha_i, \beta_i \text{ 均不相等}$$

$$H_1: \beta_1 = \beta_2 = \cdots = \beta_N$$

$$H_2: \alpha_1 = \alpha_2 = \cdots = \alpha_N, \beta_1 = \beta_2 = \cdots = \beta_N$$

F 统计量定义为：

$$F_2 = \frac{(RRS_3 - RRS_1) / [(N-1)(k+1)]}{RRS_1 / [NT - N(k+1)]} \sim F[(N-1)(k+1), N(T-k-1)] \quad (29)$$

$$F_1 = \frac{(RRS_2 - RRS_1) / [(N-1)k]}{RRS_1 / [NT - N(k+1)]} \sim F[(N-1)k, N(T-k-1)] \quad (30)$$

其中 RRS_1 表示变系数模型的残差平方和， RRS_2 表示变截距模型的残差平方和， RRS_3 表示不变参数模型的残差平方和。

若 $F_2 \leq F_{\alpha}[(N-1)(k+1), N(T-k-1)]$ ，则接受假设 H_2 ，建立不变参数模型；

若 $F_2 > F_{\alpha}[(N-1)(k+1), N(T-k-1)]$ ，则拒绝假设 H_2 ，此时检验假设 H_1 。

若 $F_1 \leq F[(N-1)k, N(T-k-1)]$ ，则接受假设 H_1 ，建立变截距模型；

若 $F_1 > F[(N-1)k, N(T-k-1)]$ ，则拒绝假设 H_1 ，建立变系数模型。

通过以上模型设定的检验，我们可以确定出更为合适的面板数据模型形式，从而减小模型设定偏差，取得更好的模型拟合效果。

5.4.2 模型变量选择与数据处理

本文研究所涉及的行业是根据《国民经济行业分类》国家标准来分类的，该标准（GB/T 4754-2002）由国家统计局起草，国家质量监督检验检疫总局、国家标准化委员会批准发布，并于 2002 年 10 月 1 日实施。该标准将国民经济行业分为 19 类，除房地产行业外，我们分别用代码 A 、 B 、 \cdots 、 S 表示其他的 18 个行业，如表 17 所示：

表 17 行业代码及名称说明

代码	行业名称
A	农林牧渔业
B	采矿业
C	制造业

D	电力、燃气及水的生产和供应业
E	建筑业
F	交通运输、仓储和邮政业
G	信息传输、计算机服务和软件业
H	批发和零售业
I	住宿和餐饮业
J	金融业
L	租赁和商务服务业
M	科学研究、技术服务和地质勘查业
N	水利、环境和公共设施管理业
O	居民服务和其他服务业
P	教育
Q	卫生、社会保障和社会福利业
R	文化、体育和娱乐业
S	公共管理和社会组织

我们选取 2004-2009 年我国各行业生产总值的增加值为研究变量。增加值是各个行业在生产过程中创造的新增价值，各行业的增加值是一个国家国内生产总值的组成部分。增加值可以按生产法计算，也可以按收入法计算，根据国家统计局有关规定按收入法计算，即增加值为劳动者报酬、生产税净额、固定资产折旧和营业盈余四部分之和，所有数据均来自《中国统计年鉴》。

5.4.3 基于面板数据模型的房地产行业与其他行业相关性效应实证分析

1. 模型的设定

(1) 混合模型和个体固定效应模型的验证

我们分别通过混合模型与固定效应模型的建立，得到

$$SSE_r = 3.61 \times 10^{10} \quad SSE_u = 2.45 \times 10^9$$

计算出 F 统计量的值为：

$$F = \frac{(3.61 \times 10^{10} - 2.45 \times 10^9) / 18}{2.45 \times 10^9 / (108 - 18 - 1)} = 67.71 > F_{0.05}(18, 89) = 1.72$$

因此，不能接受原假设 H_0 ，即建立个体固定效应模型更加合适。同时，由于本文主要是研究我国房地产行业和其他行业间的关系，而研究的行业间相关性紧集中在现有的 19 个行业，所选样本并不具有随机性，我们仅对所选样本本身进行分析，因此选择建立固定效应模型更合理。

(2) 确定模型的形式

由变系数模型、固定影响模型和不变参数模型的建立，可得

$$SSE_1 = 1.65 \times 10^8 \quad SSE_2 = 2.45 \times 10^9 \quad SSE_3 = 3.61 \times 10^{10}$$

计算得 F 统计量的值为：

$$F_2 = \frac{(3.61 \times 10^{10} - 1.65 \times 10^8) / [(18-1)(1+1)]}{1.65 \times 10^8 / [108 - 18 \times (1+1)]} = 461.1979 > F_{0.05}(34, 72) = 1.59$$

$$F_1 = \frac{(2.45 \times 10^9 - 1.65 \times 10^8) / (18-1)}{1.65 \times 10^8 / [108 - 18 \times (1+1)]} = 58.6524 > F_{0.05}(17, 72) = 1.77$$

由于同时拒绝假设 H_2 和假设 H_1 ，所以建立变系数模型更加合适。

综上所述，我们认为针对房地产行业与国民经济其他行业的关系建立固定效应的变系数模型更为合适。

2. 面板数据模型的实证结果

通过上面对模型形式的分析，我们将针对房地产行业与国民经济其他行业的关系建立固定效应的变系数模型。

首先，我们利用原始数据建立模型得到结果如图 31 所示：

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	2551.594	474.7404	5.374714	0.0000
A--FAVA	1.325002	0.157001	8.439455	0.0000
B--FAVB	0.863048	0.157001	5.497087	0.0000
C--FAVC	5.328756	0.157001	33.94092	0.0000
D--FAVD	0.222773	0.157001	1.418927	0.1602
E--FAVE	1.197919	0.157001	7.630013	0.0000
F--FAVF	0.687118	0.157001	4.376521	0.0000
G--FAVG	0.358406	0.157001	2.282826	0.0254
H--FAVH	1.525772	0.157001	9.718236	0.0000
I--FAVI	0.308900	0.157001	1.967503	0.0530
J--FAVJ	1.154709	0.157001	7.354788	0.0000
L--FAVL	0.321001	0.157001	2.044579	0.0446
M--FAVM	0.260867	0.157001	1.661565	0.1009
N--FAVN	0.062053	0.157001	0.395242	0.6938
O--FAVO	0.231369	0.157001	1.473679	0.1449
P--FAVP	0.480259	0.157001	3.058958	0.0031
Q--FAVQ	0.220895	0.157001	1.406965	0.1637
R--FAVR	0.103721	0.157001	0.660639	0.5110
S--FAVS	0.812689	0.157001	5.176329	0.0000
Fixed Effects (Cross)				
A--C	8840.452			
B--C	217.5489			
C--C	12914.15			
D--C	2518.008			
E--C	-2527.611			
F--C	2365.560			
G--C	-669.1750			
H--C	-1339.824			
I--C	-1000.899			
J--C	-5893.562			
L--C	-2130.937			
M--C	-2609.535			
N--C	-2239.140			
O--C	-1535.575			
P--C	-1062.287			
Q--C	-1472.318			
R--C	-2252.247			
S--C	-2122.606			

Cross-section fixed (dummy variables)			
R-squared	0.995588	Mean dependent var	13042.82
Adjusted R-squared	0.993444	S.D. dependent var	18686.54
S.E. of regression	1513.057	Akaike info criterion	17.74285
Sum squared resid	1.65E+08	Schwarz criterion	18.63690
Log likelihood	-922.1141	Hannan-Quinn criter.	18.10536
F-statistic	464.2398	Durbin-Watson stat	3.278080
Prob(F-statistic)	0.000000		

图 31 固定效应的变系数模型拟合结果

上图的结果说明，模型的拟合效果很好， F 值的伴随概率趋近于零，说明模型是显著的。但是模型中许多变量的系数没有通过 t 检验，说明这些系数并不显著。因此我们考虑取对数再次考察模型的效果，于是对模型的自变量和因变量取对数后进行固定效应的变系数模型的建立，得到模型结果如图 32 所示：

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	1.307204	0.178857	7.308652	0.0000
A--LOG(FAVA)	0.569500	0.081050	7.026547	0.0000
B--LOG(FAVB)	0.858300	0.081050	10.58978	0.0000
C--LOG(FAVC)	0.823255	0.081050	10.15739	0.0000
D--LOG(FAVD)	0.397214	0.081050	4.900866	0.0000
E--LOG(FAVE)	0.986760	0.081050	12.17474	0.0000
F--LOG(FAVF)	0.650287	0.081050	8.023301	0.0000
G--LOG(FAVG)	0.714570	0.081050	8.816433	0.0000
H--LOG(FAVH)	0.931952	0.081050	11.49851	0.0000
I--LOG(FAVI)	0.707668	0.081050	8.731276	0.0000
J--LOG(FAVJ)	1.345915	0.081050	16.60602	0.0000
L--LOG(FAVL)	0.920779	0.081050	11.36066	0.0000
M--LOG(FAVM)	1.038490	0.081050	12.81298	0.0000
N--LOG(FAVN)	0.681409	0.081050	8.407286	0.0000
O--LOG(FAVO)	0.741970	0.081050	9.154496	0.0000
P--LOG(FAVP)	0.778600	0.081050	9.606433	0.0000
Q--LOG(FAVQ)	0.708699	0.081050	8.743993	0.0000
R--LOG(FAVR)	0.790639	0.081050	9.754977	0.0000
S--LOG(FAVS)	0.967330	0.081050	11.93500	0.0000
Fixed Effects (Cross)				
A--C	3.569824			
B--C	0.111622			
C--C	2.249545			
D--C	3.924326			
E--C	-1.001029			
F--C	2.081159			
G--C	0.720515			
H--C	-0.179858			
I--C	0.623397			
J--C	-4.714065			
L--C	-1.592660			
M--C	-3.033370			
N--C	-0.733264			
O--C	-0.026188			
P--C	0.277231			
Q--C	0.270748			
R--C	-1.383310			
S--C	-1.164622			

Cross-section fixed (dummy variables)			
R-squared	0.997341	Mean dependent var	8.903425
Adjusted R-squared	0.996048	S.D. dependent var	1.041357
S.E. of regression	0.065465	Akaike info criterion	-2.353415
Sum squared resid	0.308564	Schwarz criterion	-1.459371
Log likelihood	163.0844	Hannan-Quinn criter.	-1.990913
F-statistic	771.5182	Durbin-Watson stat	3.011867
Prob(F-statistic)	0.000000		

图 32 取对数后的模型拟合结果

由以上模型建立结果可以看出，对变量取对数后的模型的拟合效果比之前的模型更优， F 值的伴随概率依然趋于零，而此时各变量系数均通过了 t 检验，说明系数都显著。

由此我们可以得到房地产行业与国民经济其他行业的关联效应方程如下：

$$\begin{aligned}
\ln(oava) &= 3.5698 + 0.5695 \times \ln(fava) \\
\ln(oavb) &= 0.1116 + 0.8583 \times \ln(favb) \\
\ln(oavc) &= 2.2495 + 0.8233 \times \ln(favc) \\
\ln(oavd) &= 3.9243 + 0.3972 \times \ln(favd) \\
\ln(oave) &= -1.0010 + 0.9868 \times \ln(fave) \\
\ln(oavf) &= 2.0811 + 0.6503 \times \ln(favf) \\
\ln(oavg) &= 0.7205 + 0.7146 \times \ln(favg) \\
\ln(oavh) &= -0.1799 + 0.9320 \times \ln(favh) \\
\ln(oavi) &= 0.6234 + 0.7077 \times \ln(favi) \\
\ln(oavj) &= -4.7147 + 1.3459 \times \ln(favj) \\
\ln(oavl) &= -1.5927 + 0.9208 \times \ln(favl) \\
\ln(oavm) &= -3.0333 + 1.0385 \times \ln(favm) \\
\ln(oavn) &= -0.7333 + 0.6814 \times \ln(favn) \\
\ln(oavo) &= -0.0262 + 0.7420 \times \ln(favo) \\
\ln(oavp) &= 0.2772 + 0.7786 \times \ln(favp) \\
\ln(oavq) &= 0.2707 + 0.7087 \times \ln(favq) \\
\ln(oavr) &= -1.3833 + 0.7906 \times \ln(favr) \\
\ln(oavs) &= -1.1646 + 0.9673 \times \ln(favs)
\end{aligned} \tag{31}$$

以上关联效应方程（31）表明，房地产行业每增加 1% 的增加值，农林牧渔业增加 0.5695% 的增加值，采矿业增加 0.8583% 的增加值，制造业增加 0.8233% 的增加值，电力、燃气及水的生产和供应业增加 0.3972% 的增加值，建筑业增加 0.9868% 的增加值，交通运输、仓储和邮政业增加 0.6503% 的增加值，信息传输、计算机服务和软件业增加 0.7146% 的增加值，批发和零售业增加 0.9320 的增加值，其他行业同理可知。

从以上结果我们可以看出，我国房地产行业与国民经济中其他 18 个行业都具有一定的相关性，其中金融业、科学研究技术服务和地质勘查业、建筑业、公共管理和社会组织及批发和零售业与房地产业的相关性最大，而现实中这些行业与房地产业的密切关系也是显而易见的。根据模型结果，我们可以将房地产业与

其他行业相关性大小用斜率 β 系数从大到小排列，如表 18：

表 18 房地产业与其他行业相关性大小

排名	行业名称	行业代码	β 系数
1	金融业	J	1.3459
2	科学研究、技术服务和地质勘查业	M	1.0385
3	建筑业	E	0.9868
4	公共管理和社会组织	S	0.9673
5	批发和零售业	H	0.923
6	租赁和商务服务业	L	0.9208
7	采矿业	B	0.8583
8	制造业	C	0.8233
9	文化、体育和娱乐业	R	0.7906
10	教育	P	0.7786
11	居民服务和其他服务业	O	0.742
12	信息传输、计算机服务和软件业	G	0.7146
13	卫生、社会保障和社会福利业	Q	0.7087
14	住宿和餐饮业	I	0.7077
15	水利、环境和公共设施管理业	N	0.6814
16	交通运输、仓储和邮政业	F	0.6503
17	农林牧渔业	A	0.5695
18	电力、燃气及水的生产和供应业	D	0.3972

房地产业和金融业是相互支持、相互促进的关系。一方面房地产开发和住房消费需要金融业的大力支持，另一方面房地产行业又成为金融业新的增长点。金融业在支持房地产行业的发展过程中也为自身的发展开辟了广阔的前景。由房地产业和金融业延伸而来的即房地产金融市场，是指房地产资金供求双方运用金融工具进行各类房地产资金交易的总和。随着信用工具的日益发达和不断创新，房地产金融市场的业务范围日益扩大，包括了各类住房储蓄存款、住房贷款、房地产抵押贷款，房地产信托、房地产证券，房地产保险、房地产典当等。如今房地产金融已承担着宏观调控的职能，利用金融政策手段调节房地产市场运行，并促进金融业与房地产业同时实现快速、稳健的增长。

科学研究技术服务和地质勘查业中包含了工程勘察设计和规划管理，指建筑施工前的工程地质勘查和工程设计等活动以及对区域和城市、集镇、村庄的规划及其他规划活动，这些方面的工作也是与房地产行业密切相关的。

建筑业与房地产行业的相关性更是显而易见。建筑业是建筑工程从破土动工到工程主体竣工（或封顶）的活动过程。由于房地产业中的房产建设是由建筑业来完成的，所以，建筑业和房地产业有着一种很明显的相互依存关系。它们是相互邻近的行业。

公共管理和社会组织包括国家机构、群众团体、社会团体和基层群众自治组织等，房地产行业的运作过程中必然不能缺少这些公共、社会组织的参与。

批发和零售业指批发商向批发、零售单位及其他企业、事业、机关批量销售商品以及百货商店、零售商店面向最终消费者销售的活动。房地产运作过程中必

然会向批发商或零售商购买大量基础设施建设及房屋建设过程中所需的商品,这就使得两个行业间也具有密不可分的关系。

总而言之,在我国国民经济的运作过程中,不管是与房地产行业相关关系非常密切的金融业、建筑业,还是与房地产行业有较弱相关关系的农林牧渔业,各行业的发展都或多或少地受到房地产行业的影响。

5.5 基于熵值法的房地产可持续发展模型

所谓可持续发展是指在满足现代人的需求的情况下,不损害后代人满足需求的能力。即指经济、社会、资源和环境保护协调发展,它们是一个不可分割的系统,现代社会的发展既要达到发展经济的目的、又要保护好资源和环境。扩展到房地产行业的可持续发展性可以包含土地资源的永久利用、产业自身的稳定协调发展、房地产市场完善与人居环境改善等多方面内容,即从房地产的自然属性和社会属性两方面出发实现房地产的可持续性发展。

在房地产与环境的可持续发展中,可以基于生态足迹模型进行研究。但房地产行业消耗资源的数据有待完善,现阶段无法获得充分的数据。于是本部分主要立足于产业自身与社会和经济的稳定协调发展方面,运用熵值法进行评分和计算近些年房地产行业的协调和可持续的分数情况。

5.5.1 基于熵值法的房地产行业可持续发展模型理论分析

在综合评价体系中,应用信息熵可以评价所获系统信息的有序程度和信息的信息熵。信息熵是信息无序度的度量,信息熵越大,信息的无序度越高,其信息的效用越小;反之,信息的熵越小,信息的无序度越低,其信息的效用值越大。这里使用了熵值法确定权重系数,用评价指标构成的判断矩阵来确定指标的权重,它可以尽可能的消除各指标权重计算的认为干扰,使房地产的可持续性分析更加符合实际。

1. 信息熵及权重函数的确定

假设评价房地产行业 m 年的发展状况,评价的指标体系包括 n 个指标,即 m 个样本、 n 个指标的综合评价问题。设评价系统的初始数据矩阵为 $X = \{x_{ij}\}_{m \times n}$,

其中 x_{ij} 表示第 i 个样本第 j 项评价指标的数值。由于各指标的量纲、数量级和指标的符号存在差异,首先要对初始数据做标准化处理。

(1) 数据标准化处理

假定评价指标 j 的理想值为 x_j^* ,其大小因评价指标性质不同而异。当指标为正时, x_j^* 越大越好;反之, x_j^* 越小越好。定义 x'_{ij} 为 x_{ij} 对于 x_j^* 的接近度。于是定义标准化值: $y_{ij} = x'_{ij} / \sum_{i=1}^m x'_{ij}, (0 \leq y_{ij} \leq 1)$, 得标准化矩阵 $Y = \{y_{ij}\}_{m \times n}$ 。

(2) 信息效用值 d 和评价指标的权重

第 j 项指标的信息熵值为:

$$e_j = -K \sum_{i=1}^m y_{ij} \ln y_{ij} \quad (32)$$

其中，常数 K 与系统的样本数 m 有关。对一个信息完全无序的系统，有序度为零，其熵值最大， $e=1$ ， m 个样本处于完全无序分布状态时， $y_{ij} = \frac{1}{m}$ ，于是有：

$$K = 1 / \ln m \quad (33)$$

指标的信用效用价值取决于信息熵与 1 之间的关系：

$$d_j = 1 - e_j \quad (34)$$

此时，根据信用价值系数确定权重大小，价值系数越高对评价的贡献越大。因此，第 j 项指标的权重为：

$$w_j = d_j / \sum_{j=1}^m d_j \quad (35)$$

(3) 多层评价系统

用第 j 项指标权重与标准化矩阵中第 i 个样本第 j 项评价指标接近度 x'_{ij} 的乘

积作为 x_{ij} 的评价值 f_{ij} ，而第 i 个样本的评价值就为 $f_i = \sum_{j=1}^n f_{ij}$ 。显然， f_i 越大，

样本效果越好。

进一步，根据熵的可加性，可以利用下层结构的信息效用值按比例计算对应于上层结构的权重 W_j 。于是有 $D_k = \sum_j^t d_j$ ，其中 t 为第 k 层的指标个数，那么全

部指标的总和为 $D = \sum_{k=1}^n D_k$ ，此处的 n 为层数。

相应类的权重为： $W_k = D_k / D$

指标对于上层结构的评价值： $W_k = d_j / D$

该指标对应于上层结构的评级值： $f_{ij} = \sum_{j=1}^n x_{ij} W_j$

最后总的评价数可由上述过程层层评级得到。

(4) 基于房地产可持续发展的综合评价方法

参考各类文献中关于经济和城市发展的可持续性分析，这里提出的房地产可持续发展的综合评价的指标主要包括：发展指数、发展可持续系数和发展可协调系数。

发展指数：

$$F_{ki} = \sum_{j=1}^n x'_{ij} W_j \quad (36)$$

其中 F_{ki} 表示第 i 个样本相应的指数， k 表示第 i 个样本的各类发展指数； W_j 表示第 j 个指标的权重； x'_{ij} 表示第 i 个样本中第 j 个指标的接近度； n 为各指数包含的指标数目。

在研究房地产行业的可持续发展时，主要考虑社会进步指数、行业自身发展指数和经济发展指数三类指标，用来反映房地产行业与社会及经济的联系以及自身的发展水平、速度和协调状况。

发展可持续系数：

$$S_i = \sum_k W_k F_{ki} \quad (37)$$

其中， k 为房地产可持续发展系数所包含指数的数目。

房地产行业的发展可持续系数可以衡量该行业发展系统结构和功能的状态的综合指标。该系数综合反映了评价房地产行业在评价期发展的总体水平和能力，用来描述行业的发展水平状况和发展速度。连续观察评价房地产行业在一段时间内的发展可持续系数的变化，可以度量出该行业发展速度的快慢和强弱程度，能够发现行业发展持续性总体上的趋势和方向。

发展可协调系数：

$$C_i = 1 - S_i / \bar{F}_i \quad (38)$$

其中， S_i 表示样本 i 的各指数的标准差； \bar{F}_i 表示样本 i 各指数的平均值。

协调指数用来研究行业与社会和经济等的协调性。计算出的系数越接近于 1，表明各系统之间发展越协调。

联系可持续发展系数和协调系数绘制二维评价空间，可以得出对房地产行业可持续发展的评价结论，进而分析行业的发展策略，提出政策性的指导意见。

2. 模型指标选择和数据处理

针对房地产行业可持续发展的内涵和我国目前房地产行业的现状，依据设计评价指标体系的简明性、科学性、定性与定量相结合以及政策相关性的原则，本文提出了一个三层次的房地产行业可持续发展综合评价指标体系。（见表 19）

本部分主要是基于天津市房地产行业 2005 至 2009 年相关数据建立可持续发展模型，数据来源于天津市统计年鉴、国家统计局数据库、巨星金融平台数据库，同时考虑到包含价格因素在内的指标，如人均住房年消费支出、房地产行业年平均工资、商品住房年销售额、房地产年开发投资额、房地产行业增加值等，我们均利用物价指数对其进行处理，即以 2009 年价格作为基期进行调整，以消除物价变动影响。

表 19 房地产行业可持续发展指标体系

一级指标	二级指标	三级指标
可持续发展系数 &发展协调系数	社会进步指数	人均住房年消费支出（元）

		房地产行业就业人口数（万人）
		房地产行业年平均工资（元）
	房地产业发展指数	商品住房年销售面积（万平方米）
		商品住房年销售额（亿元）
		房地产年开发投资额（亿元）
	经济发展指数	房地产行业增加值（亿元）
		房地产业发展速度（%）
		房地产行业占 GDP 比值（%）

5.5.2 基于熵值法的房地产行业可持续发展模型求解

本文以天津市 2005-2009 年房地产相关数据作为基础，采用熵值法对此期间的天津市房地产行业可持续发展状况进行综合评价，我们分别得到了 2005 年至 2009 年各年一级指标、二级指标的取值结果，同时也得到了的房地产行业综合评分，结果见表 20 和图 33。

表 20 基于熵值法的房地产行业可持续发展综合评价表

年份	二级指标			一级指标		熵值法综合评分
	社会进步指数	房地产业发展指数	经济发展指数	可持续发展系数	发展协调系数	
2005	0.763	0.623	0.575	0.627	0.851	0.116
2006	0.752	0.732	0.726	0.733	0.981	0.138
2007	0.844	0.831	0.702	0.779	0.901	0.142
2008	0.979	0.781	0.759	0.806	0.855	0.137
2009	1.000	1.000	1.000	1.000	1.000	0.172

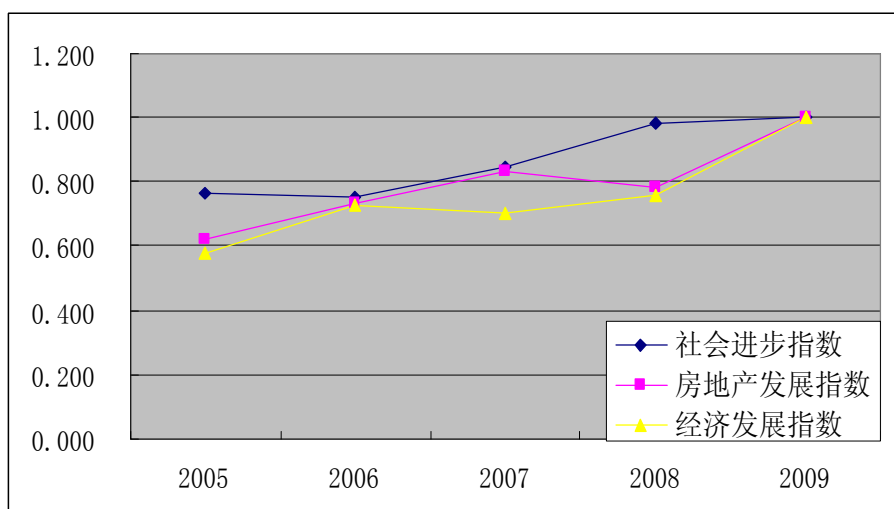


图 33 天津市 2005—2009 年房地产行业可持续发展评价指数图

首先，根据表中二级指标数据，并观察二级指标发展趋势图，易得如下分析：
社会进步指数——在 2006 年略微有下降，其余年份均稳步提高，可见，房地产行业的发展对于改善人民生活，促进社会和谐发展存在着积极的作用。

房地产行业发展指数——房地产行业在 2008 年由于政府宏观调控政策和美国的次贷危机干扰，故房地产行业发展指数在 2008 年出现了下跌，其余年份该行业都稳步健康发展，从中我们也可以看出，天津市房地产行业的发展体现了可持续发展的态势。

经济发展指数——经济发展指数体现了房地产行业对于经济的发展促进作用，经济发展指数在 2007 年的下滑同样也是由于美国次贷危机引发的金融危机影响。

综上可知，要加强天津市房地产行业的可持续发展性，除了自身平稳健康发展以外，还要建立健全完善的对抗外部突发因素冲击和干扰的机制。

然后，我们集中研究可持续发展系数和发展协调系数这两个一级指标，以 2009 年数据作为评价的目标值来看待，并制作了天津市 2005—2009 年度房地产行业可持续发展状况图。

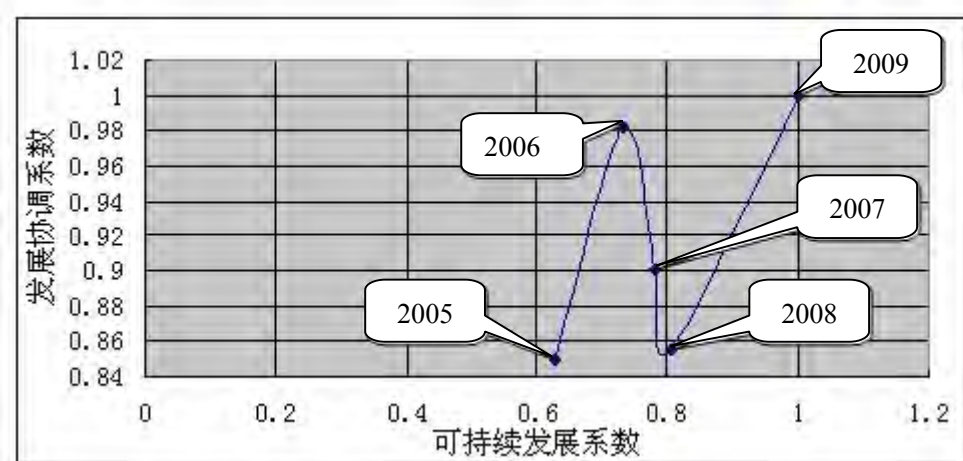


图 34 天津市 2005—2009 年房地产行业可持续发展状况图

观察图 34，我们可以发现 2005 年以来天津市房地产行业可持续发展的过程中出现了两次拐点，分别是 2006 年和 2008 年，2006 年开始，天津市房地产行业发展协调系数开始走低，其主要原因仍然在于美国次贷危机的消极影响，使得房地产行业一片低迷，在此期间我国政府相继出台了一系列的政策来鼓励和刺激房地产行业的发展，2008 年底，这些政策取得积极成效，新建商品住房成交面积大幅度增加，保障性安居工程建设进度进一步加快，因此，在 2008 年房地产市场再次出现拐点，房地产行业开始向上运行，至 2009 年，房地产行业迅速突破历史记录达到新的水平，从此，天津市房地产行业开始了一个新的阶段，同时，关注我国政府 2010 年出台的一系列促进房地产市场健康稳定发展的政策，可以预测我国房地产行业将进入全面调整阶段，将朝着可持续发展进一步迈进。

六、模型的评价与推广

本文紧密围绕我国房地产行业的相关问题进行全面研究,从当前经济社会发展现实出发,运用现实经济数据,分别建立了五个模型,分别是需求与供给模型、定价模型、房地产行业关联度及发展态势投入产出模型、房地产行业关联度及发展态势空间面板数据模型、基于熵值法的房地产行业可持续发展模型。

在模型一中,我们首先对影响房地产行业需求和供给的变量进行空间异质性和空间自相关性检验,在验证了空间自相关性存在的基础上,利用地理加权回归(*GWR*)方法建立了需求模型和供给模型,此方法分省份研究房地产市场的需求和供给,充分考虑了房地产市场地区差异性,模型结果更加符合现实,其次,在需求和供给模型求解的基础上,我们利用分层聚类分析的方法对我国房地产市场进行区域划分,为下文建立分区域的房地产定价模型做了有力的铺垫。

在对模型二的研究中,我们基于模型一对房地产市场的划分结果,主要采用了计量经济学中的向量自回归模型分区域进行建模和预测,其中引入了脉冲响应分析和方差分解,不但整体地考虑了各变量对房价的综合影响,更通过脉冲图形清晰地反映了单个变量变动对房价的作用,这相对单独使用时间序列的方法进行回归,更具有指导意义。在今后的问题的研究中,为了进一步增加实际作用,还应提高精确度,可以尝试不同变量组合的*VAR*模型的构建,再进行比较分析。

模型三则从国民经济核算角度入手建立了投入产出模型来研究房地产业发展态势和行业关联度,首先基于消耗系数矩阵和分配系数矩阵来考察房地产行业与国民经济其他行业的关联性强弱以及房地产行业对国民经济的拉动作用,其次,利用行业贡献率指标的变化趋势来分析房地产业的发展态势,模型求解的结果与现实经济十分吻合。

同样为了探讨房地产业发展态势和行业关联度,在模型四的研究中,我们从计量经济学的角度出发建立了面板数据模型,为了避免模型设定偏差,首先对模型的影响因素及模型的形式进行了检验,之后建立了固定效应的变系数面板数据模型,对房地产行业与其他行业的关联度进行了拟合。面板数据作为截面数据和时间序列数据的结合,有效克服了两者的不足,该模型获得了比单截面数据模型更多的动态信息,其运用也越来越广泛。

在模型五的房地产行业可持续发展模型中运用综合评价的方法,富有创新性地引入熵值法赋权重,避免了综合评中主观性赋权重的随机性,同时,通过多层评价系统中信息熵的构建,可以计算出具有实际意义的可持续性发展系数和协调系数,综合评价结果具有说服力。在此研究方法中,随着层次的增多,层次中代表变量选择的增加和细化,评分更细,效果更好,为此可在今后的研究中增加层次和细分变量,如房地产行业对相关自然环境的影响的量化指标,由于时间仓促,我们并未在评价体系中未引入环境指数,这也为今后深入研究提供了一个方向。

七、参考文献

- [1] 张晓峒, 应用数量经济学[M], 北京: 机械工业出版社, 2009
- [2] 刘玉红, 张红, 房地产业与社会经济[M], 北京: 清华大学出版社, 2006
- [3] 魏巍贤, 李阳, 我国房地产需求的地区差异分析 [J], 统计研究, 2005 (9): 56-59, 2005
- [4] “房地产价格发展趋势研究” 课题组, 我国房地产价格发展趋势研究 [J], 统计研究, 2008 (25): 19-25, 2008
- [5] 曲闻, 影响我国房地产价格的宏观经济因素实证分析 [J], 经济分析, 2006 (9): 15-17, 2006
- [6] 张红, 翁少群, 基于均衡价格形成机制的住宅价格变化特征研究[J], 土木工程学报, 2007 (8): 100-105, 2007
- [7] 王勇, 龙奋杰, 中国城市住宅价格供需关系模型极其应用 [J], 土木工程学报, 2002 (2): 102-108, 2002
- [8] 洪涛, 西宝, 高波, 房地产价格区域间联动与泡沫的空间扩散[J], 统计研究, 2007 (24): 64-67, 2007
- [9] 张卫民, 基于熵值法的城市可持续发展评价模型, 厦门大学学报, 2004(2): 109-115, 2004
- [10] Abraham J M, Hendershott P H, Bubbles in metropolitan housing markets [J], Journal of Housing Research, 1996, 7 (2): 191-207, 1996
- [11] Geoff Kenny, Modeling the demand and supply sides of the housing market : evidence from Ireland , Economic Modeling, 1999(16), 389-409, 1999