

# 中国各地区人口特征和房价波动的动态关系

许永洪 吴林颖

**内容提要:** 本文分析了人口特征、金融市场和房地产市场三者的相互影响机制,基于 2002–2015 年我国大陆 31 个省、自治区、直辖市的年度数据,建立了面板平滑转换模型,将人口密度作为异质变量构建计量模型,研究房地产市场的非线性影响因素,及各地区人口特征对房价波动的影响机制。实证结果表明:人均 GDP 对房价的影响随人口密度增加呈现非线性提升效应;人口密度小的地区,M2 存量对房地产价格有正向影响;当人口密度较小时,地区中老年人口占比越大,房价下降的可能越大,反映了房地产“年轻人推动房价上涨”的现象,但是极少人口密度比较大城市例外。

**关键词:** 人口特征; 人口—信贷—房价模型; 面板平滑转换模型

**DOI:** 10.19343/j.cnki.11-1302/c.2019.01.003

**中图分类号:** C812      **文献标识码:** A      **文章编号:** 1002-4565(2019)01-0028-11

## A Dynamic Relationship between Chinese Provincial Demographic Characteristics and Housing Price Volatility

Xu Yonghong & Wu Linying

**Abstract:** This paper builds up a panel smoothing transition model based on the study of interactive mechanism among the demographic characteristics, financial market and real estate market and the 2002–2015 annual data from 31 provinces, autonomous regions and municipalities in mainland China. Furthermore, taking the population density as a heterogeneous variable, an econometric model is constructed to study the non-linear factors affecting the real estate market, and the impacts of the Chinese provincial demographic characteristics on the housing price volatility. The results show that as the population density increases, the impact of per capita GDP on housing prices exposes an escalating effect in a non-linear way. In a sparsely populated city, a real estate price increase looks more like money-driven than that in a densely populated city. In a sparsely populated area, the more the aged and middle aged people, the more the housing price is inclined to edge down, reflecting the factual tendency of real estate prices driven by the young generation, but being an exception for a few most populated metropolitan cities.

**Key words:** Demographic Characteristics; Population-Credit-Property Price Model; Panel Smoothing Transition Model

### 一、引言

我国房地产市场化发展的起点是 1998 年 6 月全国城镇住房制度改革与住宅建设工作会议的召开。随后,国务院在同年 7 月 3 日颁发了《关于进一步深化城镇住房制度改革加快住房建设的通知》。正式开启了以“取消福利分房,实现居民住宅货币化、私有化”为核心的住房制度改革。如

今,房地产行业不仅成为我国经济快速发展的重要推动力,而且还是国民财富的重要承载方式。在住房市场化过程中,房地产价格出现持续上涨,特别是一线城市商品房价格在增长量和增速上都超过了大部分城镇居民可支配收入的同比涨幅,房地产“泡沫论”在十余年前就甚嚣尘上。

“泡沫”迟迟不破,让人不禁开始寻找房价上涨背后的支撑因素。从结构看,我国城市住宅价格呈现整体上涨趋势,但是各城市房地产价格涨幅存在明显的差异,呈现出一二线城市房价收入比更高,泡沫大却更能涨,三四线城市房价收入比合理却涨得慢或者属于跟着涨的状况。

城市间教育、医疗和发展机遇等因素对房价的影响越来越明显,但这不容易找到代理变量,人口增量作为这些因素的显性指标其重要性越来越凸显,从居住性的角度看,人口也是住宅需求的核心决定性指标,因此,本文研究人口和人口结构对各地区房价变化的影响。

## 二、文献综述

供需理论认为,供给曲线与需求曲线交点所对应的出清价格反映为市场中可观测到的价格,所有房地产价格的影响因素可以看成是先影响房地产市场的供需关系进而再影响房地产价格。国外的相关研究中,Gabriel等(1999)<sup>[1]</sup>建立房地产定价计量模型后发现大范围的人口迁移是近几十年加利福尼亚房价变化的主要因素,由于气候、教育和多样性等原因,人口的流入使得洛杉矶、旧金山和圣地亚哥等大城市人口逐渐膨胀,房价呈现长期上升趋势。Mankiw和Weil(1989)<sup>[2]</sup>的研究结果表明,人口的年龄结构对住宅需求的影响显著,这也是在房地产市场的相关研究中首次考虑人口特征。闫磊(2011)<sup>[3]</sup>在对房地产价格的各种影响因素的实证分析中,基于多国长达10年的面板数据,发现房地产价格与年末人口数量存在显著的正相关关系。

当然,在考虑名义房价波动时,收入与金融因素也是必不可少的因素。例如Pages和Maza(2003)<sup>[4]</sup>、Mcquinn和Oreilly(2008)<sup>[5]</sup>等分析认为,房价上涨的主要影响因素是人均实际收入和利率。Abeleson等(2005)<sup>[6]</sup>、Posedel和Vizek(2009)<sup>[7]</sup>使用门槛误差修正模型,也得到类似的分析结论:在长期,收入和利率是名义房价的主要决定因素。Meen(2002)<sup>[8]</sup>对美国和英国的长期收入和住房价格的交叉弹性进行估算,认为房价对收入的变化非常敏感,房价收入弹性分别达到2.71和2.51。但Gallin(2003)<sup>[9]</sup>基于协整检验的方法,认为美国城市房价和收入之间的协整关系并不显著。Annett(2005)<sup>[10]</sup>的文章则基于多个发达国家面板数据,并利用结构VAR模型同样得出:在欧盟国家,短期房价波动的主要原因并不是人均实际收入。除了利率和收入因素外,况伟大(2010)<sup>[11]</sup>认为,住户对房价的上涨预期也在相当程度上推动了我国房价上涨。

对于房地产市场、金融市场和人口特征三者的讨论,Nishimura(2012)<sup>[12]</sup>从一个新的角度来分析社会人口老龄化对金融市场和房地产市场的影响,文章针对人口、信贷和房价三者进行综合分析,但只局限于定性分析,对于影响的具体机制没有系统阐述和实证分析。朱太辉和张晓朴(2015)<sup>[13]</sup>在其分析的基础上,将银行信贷因素嵌套进传统的人口—资产泡沫模型,构建了同时考虑上述因素对金融市场稳定影响的“人口—信贷—房价”模型(Demography-Credit-Property Price Model,DCP模型)。该模型表明,人口特征和信贷因素对房地产市场的影响具有相互推动和放大效应:人口红利时期,青壮年人口数量不断增加,劳动力的增加使得经济增长速率增加,同时引起房地产的需求增加,驱动房地产行业扩展和房地产价格上涨。以此为基础,银行持续的房地产抵押贷款扩张则会进一步增加人们对房地产的投资性需求,引发“炒房”热,导致房地产价格的泡沫不断增加,最后会发展为恶性房地产泡沫,成为房地产市场中的一颗定时炸弹;而当社会逐渐过渡到人口老龄化时期,老年抚养比上升,又会和银行的房地产抵押贷款收缩相互促进,产生与人口红利时期反向的放大效应,最终引发房地产价格下跌。如果房地产价格泡沫破裂,甚至会引起其他资产大

幅贬值,引发金融危机。

### 三、机制分析和计量经济学模型

#### (一) 房价波动影响机制

朱太辉和张晓朴(2015)<sup>[13]</sup>给出了 DCP 模型的方程形式:

$$P_{t+1}/P_t = (1 + V_t) \left( \frac{1 + \lambda_t}{1 + \lambda_{t+1}} \right) \left[ 1 + \frac{g_t(V_t, \lambda_t, d_t) + d_t(l_t/y_t)}{1 + l_t/y_t} \right] \quad (1)$$

式(1)即为“人口—信贷—房价”模型。其中,  $P_t$  表示时期  $t$  的房地产价格,  $V_t$  表示人口增长率,  $\lambda_t$  表示人口结构(老年抚养比),  $g_t$  表示时期  $t$  的经济增速,  $d_t$  表示信贷增长率(银行贷款总量增长率),  $l_t$  表示时期  $t$  的贷款需求,  $y_t$  表示时期  $t$  的收入。该模型可以这样理解:等式左边表示房地产价格上涨速度。右边第一项代表人口增长率,第二项是人口结构变动率,经济增速、信贷增长率和信贷存量三者的综合项组成了等式右边的第三项。可以看出,  $t+1$  期的房地产价格变化速度和人口增长率  $V_t$  存在正相关关系,而与代表人口结构的  $\lambda_t$  存在负相关关系。

对式(1)等号两边取自然对数,两边的分式转化为相关变量的变化率,等号右边的乘法形式转化为加法形式。同时考虑到货币市场供给对房地产价格的影响,可以构建以下计量模型:

$$y_{it} = \alpha + \beta_1 x_{1it} + \beta_2 x_{2it} + \beta_3 x_{3it} + \beta_4 x_{4it} + u_i + \eta_t + \xi_{it} \quad (2)$$

式(2)中,  $y_{it} = \ln(P_{it}/P_{i,t-1})$  代表地区  $i$  在时期  $t$  的房地产价格变化率,  $x_{1it} = \ln(m_{it}/m_{i,t-1})$  为人口总量( $m_{it}$ )的变化率,  $x_{2it} = \ln(\lambda_{it}/\lambda_{i,t-1})$  为人口结构(以老龄人口抚养比  $\lambda_{it}$  表示)的变化率,  $x_{3it} = \ln(l_{it}/l_{i,t-1})$  为银行信贷( $l_{it}$ )的变化率,  $x_{4it} = \ln(GDP_{it}/GDP_{i,t-1})$  为实际人均 GDP 的变化率。由于  $l_{it}/y_{it}$  与银行信贷( $l_{it}$ )、实际人均 GDP 存在较为严重的多重共线性,因此从计量模型中剔除。 $u_i$  为地区  $i$  未观测到的固定效应,  $\eta_t$  为时期  $t$  未观测到的固定效应,  $\xi_{it}$  为误差项。

但是式(1) DCP 模型仍存在三点不足:一是并未考虑货币市场的影响,货币中性假说阐明了货币作为度量单位可能使得两个时间段内相同价值的商品有着不同的价格。二是并未将金融市场考虑全面,因为在分析房地产市场和金融市场的关系时,信贷因素不足以全面地反映金融市场。市场参与者除了可以向金融市场借钱(信贷),也可以像投资房地产市场一样投资金融市场。而在投资金融市场的所有方式中股市最具有代表性,门槛低且直接或间接涉及其他多种金融产品。三是人口总量这个变量应该用人口密度来替代,这样更严谨一些。

综合以上三点不足,本文在参考“人口—信贷—房价”模型的基础上,嵌套货币市场因素和金融市场因素来分析人口结构和金融市场变化对房地产市场的影响。

本文用银行房贷公积金利率(5年以上)和 A 股市场收益率的加权和来替代银行信贷利率,以度量金融市场收益的变化率。在加权平均时用房市和股市的交易规模作为权重(权重每年更新),同时将其他变量设定为当期量(以当期量建模有效,则变化率模型也有效),可得到计量模型:

$$y_{it} = \alpha + \beta_1 x_{1it} + \beta_2 x_{2it} + \beta_3 x_{3it} + \beta_4 x_{4it} + \beta_5 x_{5it} + u_i + \gamma_t + \xi_{it} \quad (3)$$

其中,  $y_{it} = P_{it}$  表示时期  $t$  地区  $i$  的房地产价格,  $x_{1it} = GDP_{it}$  表示地区  $i$  在时期  $t$  的人均实际 GDP,  $x_{2it} = M2_{it}$  表示时期  $t$  的货币市场供给,  $x_{3it} = Den_{it}$  为地区  $i$  在时期  $t$  的人口密度,  $x_{4it} = \lambda_{it}$  表示地区  $i$  在时期  $t$  的社会年龄结构(以老龄人口抚养比  $\lambda_{it}$  表示),  $x_{5it} = \ln(w_{1t}l_{it} + w_{2t}r_{it})$  表示金融市场综合收益率,  $r_{it}$  是时期  $t$  的股市收益率,  $w_{1t}$  表示信贷市场的权重,  $w_{2t}$  表示金融市场的权重,权重以当年房地产市场销售额和 A 股市场成交额为基础计算得到。

#### (二) 模型阐释

模型中关于房地产价格和金融市场之间的关系,在上世纪 80、90 年代出现的“信贷资产观点”

中有类似的阐述。该观点认为, 房地产作为固定资产, 同时具有生产要素和贷款抵押物两种功能, 因此不论是资产价格变化还是金融市场波动, 都会因房地产而相互影响, 使得经济体最初受到的冲击在时间维度上延续下去, 冲击的影响放大并升级。该观点的典型代表有 Bernanke 等(1996)<sup>[14]</sup>的“金融加速器”模型以及 Kiyotaki 和 Moore(1997)<sup>[15]</sup>的“信贷周期”模型。在我国, 房地产作为固定资产, 其存量在国民总财富中占比超过了 60%, 可以认为上述金融市场与房地产市场之间的关系同样适用于个人房产( Iacoviello 和 Pavan, 2013)<sup>[16]</sup>。和“信贷资产观点”的相关模型相比, 本文设定的新模型覆盖了房地产价格与人口因素、货币因素和金融市场因素之间的相互影响, 能够更加全面地理解我国房价的持续上涨以及房地产市场在不同规模城市之间的差异表现。一方面, 社会不断发展使得人口生存周期变长, 其数量和结构的变化至少需要持续几年才能显现, 如果在其他因素稳定的情况下, 同时加上温和的货币政策, 使市场参与者产生稳定的经济增长预期, 这就成为了金融市场进一步扩张的基石, 进而使得房地产抵押贷款规模持续扩大以及房地产价格不断上涨。另一方面, 式(1)中等号右边存在交互项, 说明人口特征的变化还会和社会经济增长、银行抵押贷款扩张产生交互影响相互促进, 进一步放大其影响。因此, 由人口特征和银行贷款规模同时发生同向变化产生的对房地产市场的影响会比两个因素各自独立影响的和还要大, 进而使得市场参与者内心产生更大的房地产价格变动预期, 最终引发对社会经济更大的冲击。

最后, 本文将应用面板平滑转换模型来研究人口特征对房地产市场的影响, 并参考异质性样本的观点进一步完善理论模型。该模型的系数从一种状态向另一种状态的平滑变化, 而不是离散跳跃的, 更符合人口密度变化特征。

## 四、模型的设定、检验和评估

### (一) 模型设定

面板平滑转换模型( Panel Smooth Transition Regression) 是一个含有外生解释变量的固定效应模型, 可以理解为非线性同质面板模型, 也可以理解为一个系数可以随着个体的时间变化的异质面板模型。回归系数的异质性假设系数是转换变量的连续函数, 转换变量随着时间和个体的改变取值不同, 回归系数也将不同。那么, 还可以把面板平滑转换模型看成是面板门限模型的一般扩展。

面板门限模型的广泛适用性得益于门限变量随时间变化, 使得个体可以被归类不同的组, 实现不同机制的转换。但由于模型设定中包含非连续函数, 代表着面板门限模型的机制转换瞬间完成, 这是该模型的一个缺陷。面板平滑转换模型则放松了异质变量的影响系数在阈值点间断的假设, 使得影响系数在不同机制间转换是一个渐变的过程。

一个基本的面板平滑转换模型可写为:

$$y_{it} = u_i + \beta_0 x_{it} + \beta_1 x_{it} g(q_{it}; \gamma, \epsilon) + \xi_{it} \quad (4)$$

$i$  和  $t$  分别代表截面个体和时间维度。转换函数  $g(q_{it}; \gamma, \epsilon)$  是转换变量  $q_{it}$  的连续函数, 取值范围是  $[0, 1]$ , 回归系数在  $\beta_0$  和  $\beta_0 + \beta_1$  之间取值。转换函数一般采用如下的 logistic 函数形式:

$$g(q_{it}; \gamma, \epsilon) = \frac{1}{1 + \exp[-\gamma \prod_{k=1}^m (q_{it} - c_k)]} \quad \gamma > 0, \epsilon_1 \leq c_2 \leq \dots \leq c_m$$

其中  $\gamma$  是转换斜率, 决定了转换的速度。位置参数  $c$  表示转换机制的阈值, 一般只考虑  $m=1$  或  $m=2$  两种情况。

进一步, 模型可由两机制的基本模型扩展到一般化的多机制面板平滑转化模型, 一般形式为:

$$y_{it} = u_i + \beta_0 x_{it} + \sum_{j=1}^r \beta_j x_{it} g_j(q_{it}; \gamma, \epsilon) + \xi_{it} \quad (5)$$

转换函数的一般形式为:

$$g_j(q_{ju}, \gamma, c) = \frac{1}{1 + \exp[-\gamma_j \prod_{k=1}^m (q_{ju} - c_{jk})]} \quad \gamma_j > 0, c_{j1} \leq c_{j2} \leq c_{j3} \leq \dots \leq c_{jm}$$

当  $m = 1, q_{ju} = q_u$  ,且所有的  $\gamma_j$  趋近于无穷大时 ,一般形式的面板平滑转换模型就转变成  $r + 1$  机制的面板门限模型 ,故可以看成是多机制面板门限模型的一般化扩展。

(二) 模型可识别性检验

采用面板平滑转换模型的前提是面板数据存在变量异质性 ,所以在建立模型之前需要先对面板数据进行线性(同质性)检验。提出原假设  $H_0: r=0$  ,在原假设条件下面板平滑转换模型变为:

$$y_{iu} = u_i + \beta_0 x_{iu} + \xi_{iu} \tag{6}$$

由于面板平滑转换模型中存在其他参数 ,原假设下位置参数  $c$  未定义 ,导致检验无法标准化。为了解决冗余参数问题 ,Luukkonen 等(1998)<sup>[17]</sup>考虑将转换函数  $g$  在  $\gamma = 0$  处进行  $m$  阶泰勒展开 ,从而得到辅助方程:

$$y_{iu} = u_i + \beta_0 x_{iu} + \beta_1 x_{iu} q_{1u} + \dots + \beta_m x_{iu} q_{mu} + \xi_{iu} + R_m \tag{7}$$

其中  $R_m$  为泰勒展开余项。因此检验原假设  $H_0: r=0$  等价于检验辅助方程中  $\beta_1 = \beta_2 = \dots = \beta_m = 0$ 。在原假设下 ,泰勒展开的近似结果不影响  $u$  的渐进分布 ,则可以通过 LM 检验来判断面板数据的异质性特征是否存在。Colletaz 和 Herlin(2006)<sup>[18]</sup>给出了一种易于操作的检验方法 ,具体算法如下:

(1) 定义变量  $\tilde{y}_{iu} = y_{iu} - (\sum_t y_{it}) / T, \tilde{x}_{iu} = x_{iu} - (\sum_t x_{it}) / T$  ,由  $\tilde{y}_{iu}$  对  $\tilde{x}_{iu}$  回归 ,计算得到残差平方和  $SSR_0$ ;

(2)  $\tilde{y}_{iu}$  对  $\tilde{x}_{iu}$  和  $(x_{iu} q_{1u} - (\sum_t x_{it} q_{1t}) / T, \dots, x_{iu} q_{mu} - (\sum_t x_{it} q_{mt}) / T)$  进行回归 ,计算得到残差平方和  $SSR_1$ ;

(3) 计算 LM 统计量和 LMF 统计量:

$$LM = TN(SSR_0 - SSR_1) / SSR_0 \sim \chi^2(mk)$$

$$LMF = \frac{SSR_0 - SSR_1}{mk} \bigg/ \frac{SSR_0}{TN - N - m(k + 1)} \sim F(mk, TN - N - mk)$$

其中  $K$  表示解释变量的个数 ,在原假设下 ,LM 统计量服从渐进卡方分布  $\chi^2(mk)$  ,LMF 统计量渐进服从分布  $F(mk, TN - N - mk)$ 。若在给定的显著性水平下拒绝原假设 ,则说明面板数据存在异质性 ,有显著的非线性影响 ,可以建立面板平滑转换模型。

上述异质性检验还可以用于确定转换函数中位置参数  $m$  的取值。检验以下原假设:  $H_{01}: \beta_1 = 0$  和  $H_{02}: \beta_1 = \beta_2 = 0$  ,如果拒绝了  $H_{02}$  ,则选择  $m=2$  ,否则选  $m=1$ 。

(三) 模型参数估计

以基本的两机制面板平滑转换模型为例 ,模型中需要估计的参数为  $\theta = (\beta_0, \beta_1, \gamma, c)$  ,可以采用固定效应面板模型的估计思想和非线性最小二乘算法的迭代来估计参数。将模型重写为:

$$y_{iu} = u_i + \beta' \tilde{x}_{iu}(\gamma, c) + \xi_{iu}$$

去除个体效应:

$$\tilde{y}_{iu} = u_i + \beta' \tilde{x}_{iu}(\gamma, c) + \tilde{\xi}_{iu}$$

其中  $\tilde{y}_{iu} = y_{iu} - (\sum_t y_{it}) / T, \tilde{x}_{iu}(\gamma, c) = (x_{iu} - (\sum_t x_{it}) / T, x_{iu} g(q_{1u}, \gamma, c) - (\sum_t x_{it} g(q_{1t}, \gamma, c)) / T, \dots, x_{iu} g(q_{mu}, \gamma, c) - (\sum_t x_{it} g(q_{mt}, \gamma, c)) / T)$  ,  $\tilde{\xi}_{iu} = \xi_{iu} - (\sum_t \xi_{it}) / T$ 。则可在  $\gamma$  和  $c$  给定的条件下将模型看成是线性模型 ,因此可以应

用非线性最小二乘法以最小化残差平方和为目标来确定参数  $\gamma$  和  $c$ :

$$(\gamma, c) = \text{Argmin} Q(\gamma, c) = \sum_{i=1}^N \sum_{t=1}^T (\tilde{y}_{it} - \hat{\beta}(\gamma, c) \tilde{x}_{it}(\gamma, c))^2$$

具体算法为: ①给定参数  $\gamma$  和  $c$  的值后, 用线性最小二乘法估计模型得到参数  $\hat{\beta}(\gamma, c)$  的估计值; ②已知  $\hat{\beta}(\gamma, c)$  后, 再通过非线性最小二乘法估计参数  $(\gamma, c)$ ; ③重复以下两步骤, 直至参数向量  $\theta$  收敛。

对  $\gamma$  和  $c$  的初始值设定, 可采用多种算法, 如网格搜索法和模拟退火法。本文采用网格搜索法, 即在一系列  $\gamma$  和  $c$  的可能取值范围中进行划分, 在不同的取值组合下, 分别计算模型拟合得到的残差平方和, 选取使残差平方和达到最小的变量组合作为  $\gamma$  和  $c$  的初始值。网格搜索法可以避免人工赋值的随意性, 更有可能避开局部最优解而达到全局最优解。

#### (四) 模型评估

模型评估的方法以两机制面板平滑转换模型为例, 考虑模型是否已经捕捉到数据的所有异质性成分, 需要对模型进行非保留异质性检验, 即残余异质性检验。若以三机制面板平滑转换模型为检验对象, 则检验模型可写为:

$$y_{it} = u_i + \beta_0 x_{it} + \beta_1 x_{it} g_1(q_{1it}, \gamma_1, c_1) + \beta_2 x_{it} g_2(q_{2it}, \gamma_2, c_2) + \xi_{it} \quad (8)$$

其中  $q_{1it}$  和  $q_{2it}$  不一定相同。原假设  $H_0: \gamma_1 = 0$  将  $g_2$  在  $\gamma_2 = 0$  的  $m$  阶泰勒展开式替代  $g_2$ , 可得检验的辅助方程:

$$y_{it} = u_i + \beta_0 x_{it} + \beta_1 x_{it} g_1(q_{1it}, \gamma_1, c_1) + \beta_{21} x_{it} q_{21it} + \dots + \beta_{2m} x_{it} q_{2mit} + \xi_{it} \quad (9)$$

其中, 通过两机制面板平滑转换模型的参数估计得到  $\gamma_1$  和  $c_1$ 。对于原假设  $H_0: \gamma_2 = 0$  等价于检验  $H_0: \beta_{21} = \dots = \beta_{2m} = 0$ , 检验过程类似于异质性检验, 检验过程如下:

- (1) 对两机制面板平滑转换模型作参数估计, 得到模型残差平方和  $SSR_0$ ;
- (2) 对以上辅助方程进行拟合, 计算得到残差平方和  $SSR_1$ ;
- (3) 计算 LM 统计量和 LMF 统计量:

$$LM = TN(SSR_0 - SSR_1) / SSR_0 \sim \chi^2(mk)$$

$$LMF = \frac{SSR_0 - SSR_1}{mk} \bigg/ \frac{SSR_0}{(TN - N - m(k + 1))} \sim F(mk, TN - N - mk)$$

其中  $k$  为解释变量的个数, 在原假设下, LM 统计量服从渐进卡方分布  $\chi^2(mk)$ , LMF 统计量渐进服从分布  $F(mk, TN - N - m(k + 1))$ 。如果在给定的显著性水平下无法拒绝原假设, 则说明两机制面板平滑转换模型的设定是合理的。

模型的非保留异质性检验, 也可以理解为对转换函数个数的检验。如果估计的面板平滑转换模型中含有  $r^*$  个转换函数, 则先检验原假设  $r = r^*$ , 若无法拒绝则认为采用  $r^*$  个转换函数的设定是合理的, 否则继续检验  $H_0: r = r^* + 1 \dots$  直到无法拒绝原假设为止。

## 五、实证分析

### (一) 变量选择与数据来源

住宅价格是我国房地产价格的核心问题, 因此本文选取《中国统计年鉴》中的住宅价格作为研究对象; 人均实际 GDP 以当年人均 GDP 剔除通货膨胀因素计算得到; 货币因素为货币市场的 M2 存量; 人口密度以人每平方公里为单位度量; 人口结构使用老年抚养比来表示; 金融市场收益率是房贷公积金利率(5 年以上) 和股市(文中采用沪市大盘数据) 收益率的加权平均, 以房市和股市的交易规模作为权重, 并且每年更新权重。

房地产价格、人均实际 GDP、M2、人口密度、人口结构和金融市场收益率的基础数据来源于国家统计局官网和《中国统计年鉴》。因老年抚养比的统计指标从 2002 年才开始有,且只有年度数据,故本文数据覆盖全国 31 个省、自治区、直辖市(不含港澳台),时间从 2002 年到 2015 年总计 14 年,共 434 条数据。

## (二) 模型检验与识别

建立面板回归模型必须要通过单位根检验和协整检验才能避免伪回归结果的出现。本文采用 LLC 和 IPS 检验进行面板单位根检验,若两个检验都通过则可以认为面板数据不存在单位根过程。检验时设置最大滞后期为 12,检验结果见表 1。

表 1 面板单位根检验结果

检验形式	原假设	统计量	P 值	是否拒绝原假设
LLC,有截距和趋势	存在单位根过程	-11.985	0.000	是
IPS,有截距和趋势	存在单位根过程	-28.474	0.000	是

从表 1 可以看出,无论是 LLC 检验还是 IPS 检验的 P 值都很小,故拒绝存在单位根的原假设,认为面板数据不存在单位根过程。对于协整检验本文采用 PO 检验,检验结果见表 2。

表 2 面板协整关系检验结果

原假设	残差滞后期	PO 统计量	P 值	是否拒绝原假设
不存在协整关系	4	-64.36	0.01	是

从表 2 可以看出,在 5% 的显著性水平下 PO 检验是拒绝原假设的,即认为面板数据存在协整关系,可以进行回归建模。

面板平滑转化模型的本质是固定效应面板模型,所以还需要进行固定效应和随机效应的检验,一般用 Hausman 检验来判断面板数据更适合随机效应模型还是固定效应模型。以房地产价格为被解释变量,人均实际 GDP、M2 存量、人口密度、人口结构和金融市场收益率作为解释变量进行 Hausman 检验,检验方程如式(3)。进行 Hausman 检验之前先对面板混合估计模型作 LM 检验来决定面板模型的效应形式,结果见表 3。

表 3 混合模型 LM 检验结果

效应形式	LM 统计量	P 值	效应是否显著
时点	-0.255	0.799	否
个体	36.719	0.000	是
时点和个体	1348.3	0.000	是

从表 3 可以看出,时点效应并不显著,故考虑建立个体固定效应模型和个体随机效应模型来进行 Hausman 检验,结果见表 4。

表 4 Hausman 检验结果

原假设	卡方统计量	P 值	是否拒绝原假设
随机效应更有效	45.56	0.000	是

由表 4 可知,个体固定效应模型比个体随机效应模型更有效,故可以建立面板平滑转换模型。最后,由于本文只考虑人口密度作为模型的异质变量,即认为  $r=1$ ,故模型识别只检验  $m=1$  和  $m=2$  两种情况。检验结果如表 5 所示。

可以看出无论对于  $m=1$  还是  $m=2$ ,检验结果都显著拒绝“线性模型”的原假设,认为模型中的确存在以人口密度为异质变量的非线性关系。且  $m=1$  时拒绝原假设的力度更大(二者 LM 统计量的 P 值相近,但是 LMF 统计量的 P 值相差一个数量级),故选定  $r=1$  及  $m=1$  进行建模。

表 5 模型机制检验

模型机制	统计量类型	统计量	P 值
m=1	LM	105.5484	$3.57 \times 10^{-21}$
m=1	LMF	19.31	$3.212 \times 10^{-17}$
m=2	LM	121.9931	$1.987 \times 10^{-21}$
m=2	LMF	10.99	$1.44 \times 10^{-16}$

(三) 模型参数估计与结果分析

采用网格搜索法确定位置参数初值后,使用 R 语言进行建模,估计模型形式为:

$$y_{it} = u_i + \beta_0 x_{it} + \beta_1 x_{it} g(q_{it}, \gamma, \rho) + \xi_{it} \quad (10)$$

其中,  $\beta_0 = (\beta_{01} \beta_{02} \beta_{03} \beta_{04} \beta_{05})$ ,  $\beta_1 = (\beta_{11} \beta_{12} \beta_{13} \beta_{14} \beta_{15})$  为解释变量的回归系数,  $x_{it} = (x_{1it}, x_{2it}, x_{3it}, x_{4it}, x_{5it})$  为解释变量向量。参数估计结果见表 6。

表 6 面板平滑转换模型参数估计结果

变量	$\beta_0$	$\beta_1$
人均 GDP	0.016*** (0.007)	1.084*** (0.092)
M2	0.003*** (0.0002)	-0.008*** (0.002)
人口密度	19.779*** (1.169)	-16.404*** (1.265)
人口结构	-153.378*** (27.990)	710.268*** (59.141)
金融收益率	1.016 (0.747)	-11.162*** (4.022)
转化速率 r	0.237	
位置参数 c	1201.029	
回归方程调整 R <sup>2</sup>	0.891	
F 统计量	352.497***	

注: 括号内为 t 检验值, \*, \*\*, \*\*\* 分别表示在 10%、5% 和 1% 的水平上显著。

在此模型中,回归方程的调整 R<sup>2</sup> 达到了 0.891, F 统计量显著,且除金融收益率的  $\beta_1$  不显著外,其他所有参数的估计值都是显著异于 0 的,说明方程拟合程度较好。转换函数 g 的位置参数为 1201.029, 转换速率为 0.237, 可能与度量单位有关。即当人口密度大于约 1200 人每平方公里时,可以认为该地区处于模型的“高处”,要重点考虑  $\beta_1$  在模型中的作用。

图 1~5 是式(10)估计结果,其中  $\beta_0 + \beta_1 g(q_{it}, \gamma, \rho)$  即为  $x_{it}$  的影响系数随  $q_{it}$  (人口密度) 变化曲线的函数解析形式,转换函数 g( ) 在模型设定部分以 logistic 函数形式给出。针对各个变量的回归系数作进一步分析,可以得到以下几点结论。

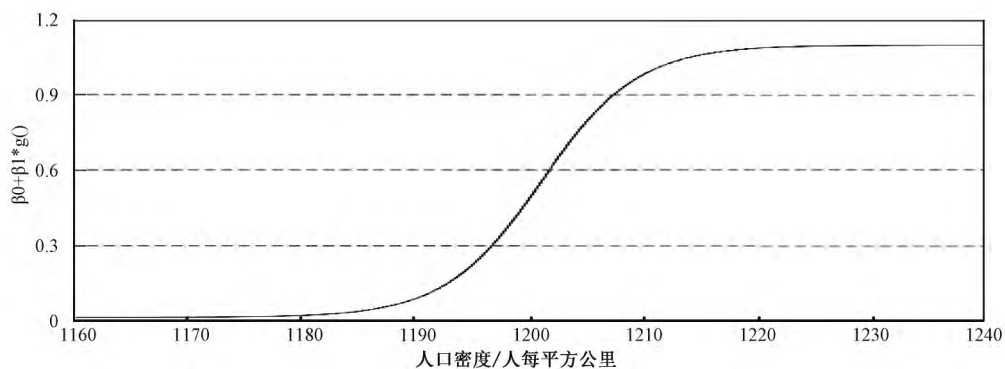


图 1 人均实际 GDP 影响系数变化

第一,在人口密度较小时人均 GDP 对房地产价格的影响不大,但随着人口密度提高影响逐渐变大,当人口密度接近于 1191.35 人每平方公里时,人均实际 GDP 每增加 1 元,房价上涨 1 元。而在人口密度小的地区,人均收入对房价的影响并不大,可能是因为财富并不集中于当地人手中,也



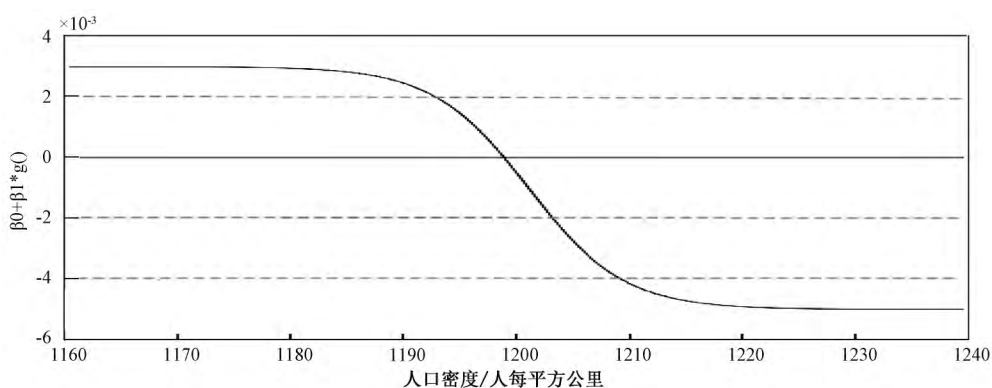


图 2 M2 影响系数变化

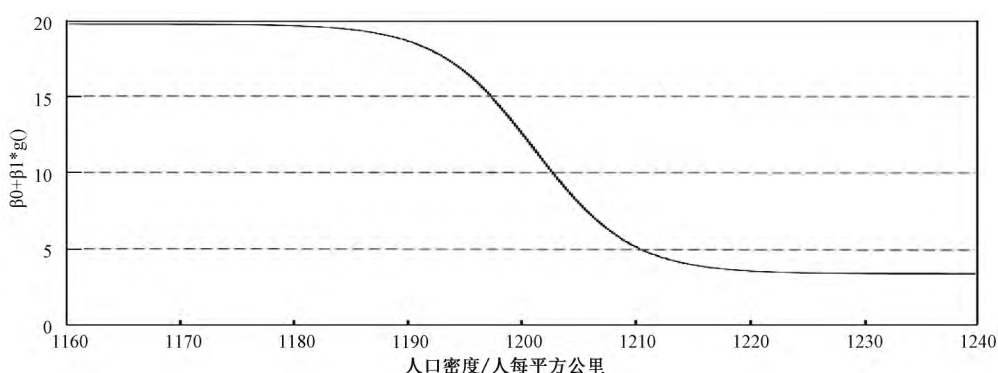


图 3 人口密度影响系数变化

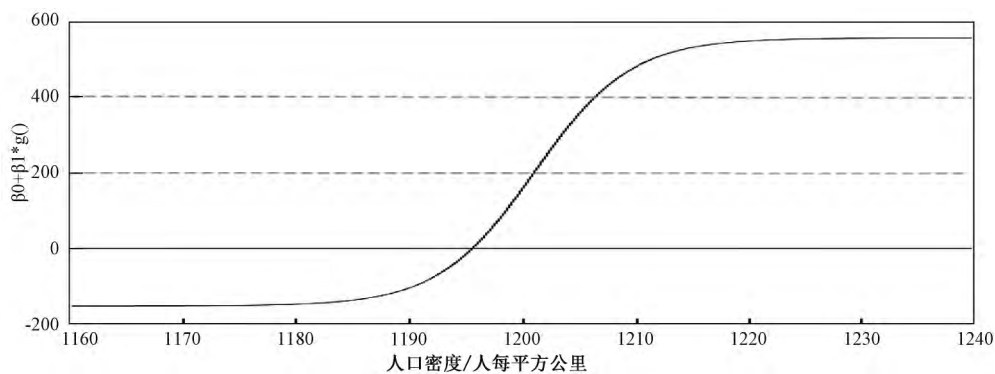


图 4 人口结构影响系数变化

有可能是因为人口稀少使得房地产市场本身体量就很小房子的使用属性大于其资产属性。

第二, M2 存量作为货币市场的代表, 对房地产价格的影响随着人口密度的变化而变化。在人口密度小的地区, M2 存量对房地产价格有正向的影响, 每 330 亿的货币存量增加将使得房价上涨 1 元。当人口密度大于 1198.87 人每平方公里时, 正向影响转变为负向影响, 如北京、天津、上海等直辖市。这意味着随人口密度的增加, 宽松的货币政策对房价的正向影响将越来越小, 甚至变为负向影响。对此笔者注意到模型的设定中考虑了人均 GDP 对房价的影响, 也就是说剔除人均 GDP 增加对应的货币后, 多发放的货币不是流向消费者主体, 而是流向众多行业的生产者(包括房地产行业)。在人口密度大的城市很大一部分货币流向非房地产行业去追求相比房地产行业更高质量的回报, 这样一来房地产行业就会受到负向影响。

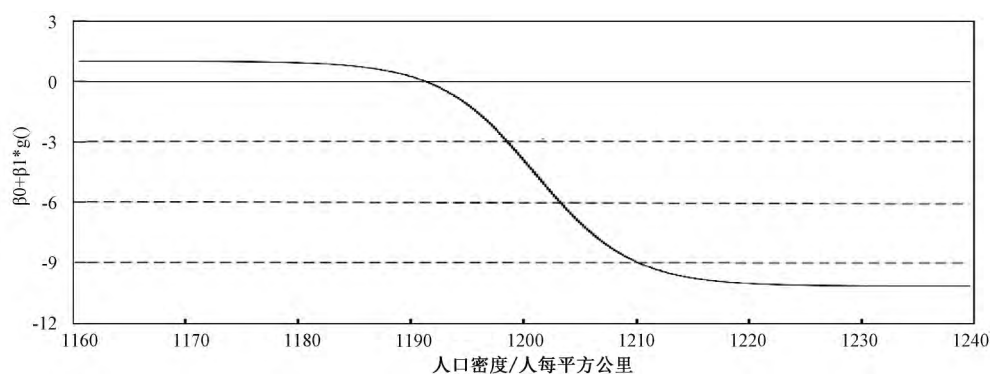


图 5 金融收益率影响系数变化

第三,人口密度本身对房地产价格具有正向影响,但随着人口密度的增加,正向影响逐渐减弱,但始终为正。这表明人口密度对房地产的正向影响具有衰减效应:随着人口密度的增加,一方面人口压力导致社会对房地产的需求增加,这种由人口密度的改变引发的需求增加使得房价上涨;另一方面,人口压力使得居住环境变差,削弱了房价上涨的力度,但仍无法改变房价上涨的趋势。

第四,随着人口密度的增加,人口结构对房地产市场的影响从负向转向正向。当人口密度较小时,老年人口(65岁以上)占比越大,房价越有下降可能。因为在人口密度小的地区,老年人口占比增加意味着该地区年轻人外出寻求就业机会,失去了对房产的刚需,房价自然会下降。当人口密度增加到 1195.59 人每平方公里时,影响反转为正向,老年人口占比越大,房价越有可能上涨。这是因为在人口密度大的地区,老年人口比重的增加意味着有更多的人选择在这个地区安定下来,而老年人口掌握着社会的主要资源,已经成为社会中的中产阶级,他们中相当一部分人甚至都为子女置备了房产,所以这部分人的占比对房地产价格的影响是正向的。

第五,金融收益率的影响在人口密度低的情况下并不显著,这是因为在人口密度低的地区,房地产的使用属性远大于资产属性,这与人均 GDP 的影响类似。同时,正向的影响意味着金融收益率的增加将带动房地产市场,放大了金融市场的红利。随着人口密度的逐渐增大,金融收益率的影响经历了从正向到负向的变化,人口密度约为 1191.3 人每平方公里时为转折点。在人口密度大的地区,房地产的资产属性更受重视,如果金融市场的收益率高,人们更愿意投资于金融市场,相对冷落房地产市场,房价有下降的可能。

## 六、结论

本文考虑了人口结构、货币政策和金融收益率等因素对房地产市场的非线性影响,并把人口密度作为重点考察对象,分析了地区间异质性特征。采用了 2002-2015 年的宏观数据,以面板平滑转化模型为计量工具,考察了我国 31 个省(区、市)影响房地产市场的潜在因素,其结论如下:

第一,人均 GDP 对房价的影响在不同人口密度地区的效应不同,在低人口密度地区,人均 GDP 对房价的影响较小,但随着人口密度增加而增大。相同的货币政策在人口密度不同的地区有不同的影响,说明房地产兼具居住和资产属性,且在人口密度大的地区更易受到产出和收入影响。

第二,M2 存量作为货币市场的代表,对房地产价格的影响随着人口密度的变化而变化。在人口密度小的地区,M2 存量对房地产价格有正向的影响,但是在人口密度大的地区,房价的上涨更加不受制于 M2 存量,说明这些城市房价上涨更倾向于是由人口的需求和人均 GDP 等基本面推动。

第三,人口结构对房地产市场的影响随着人口密度的增加有所不同。当人口密度较小时,中老

年人口(65岁以上)占比越大的地区,房价越有下降的可能,反映了房地产“年轻人推动房价上涨”的现象。但是少数超过人口密度阈值的城市,老年人占比对房价呈现正向提升作用。

#### 参考文献

- [1] Gabriel S. A., Matthey J. P. and Wascher W. L.. House Price Differentials and Dynamics: Evidence from the Los Angeles and San Francisco Metropolitan Areas[J]. *Economic Review*, 1999 (1): 3-22.
- [2] Mankiw N. G. and Weil D. N.. The baby boom, the baby bust, and the housing market: A reply to our critics. NBER Working Papers, 1989, 19(2): 235-258.
- [3] 闫磊. 房地产价格影响因素分析: 基于1997-2008年面板数据的实证研究[J]. *商业文化*, 2011 (1): 138-138.
- [4] Pages J. and Maza L.. Analysis of House Prices in Spain. Banco de Espana Working Papers, 2003: No.0307.
- [5] McQuinn K. and O'Reilly G.. Assessing the Role of Income and Interest Rates in Determining House Prices[J]. *Economic modelling*, 2008 (25): 377-390.
- [6] Abelson P., Joyeux R., Milunovich G., and Chung D.. Explaining House Prices in Australia: 1970-2003[J]. *Economic Record*, 2005, (81): 96-103.
- [7] Posedel P. and Vizek M.. House Price Determinants in Transition and EU-15 Countries[J]. *Post-Communist Economies*, 2009 (21): 327-343.
- [8] Meen G.. The Time-series Behaviour of House Prices: A Transatlantic Divide[J]. *Journal of Housing Economics*, 2002 (11): 1-23.
- [9] Gallin J.. The Long-run Relationship between House Prices and Incomes: Evidence from Local-house Markets[J]. *Real Estate Economics*, 2003 (34): 417-439.
- [10] Annett A.. House Prices and Monetary Policy in the Euro Area. IMF country report, 2005: No.05/266.
- [11] 况伟大. 利率对房价的影响[J]. *世界经济*, 2010 (4): 134-145.
- [12] Nishimura K.. How to Detect and Respond to Property Bubbles[Z]. Presentation at the Reserve Bank of Australia-BIS Research Conference "Property Markets and Financial Stability", Sydney, 21 August 2012.
- [13] 朱太辉, 张晓朴. 金融危机的人口-信贷-房价模型[J]. *金融评论*, 2015 (5): 31-44.
- [14] Bernanke B., Gertler M. and Gilchrist S.. The Financial Accelerator and the Flight to Quality[J]. *Review of Economics and Statistics*, 1996 (78): 1-15.
- [15] Kiyotaki N. and Moore J.. Credit Cycles[J]. *Journal of Political Economy*, 1997 (105): 211-248.
- [16] Iacoviello M. and Pavan M.. Housing and Debt over the Life Cycle and over the Business Cycle[J]. *Journal of Monetary Economics*, 2013 (60): 221-238.
- [17] Luukkonen R., Saikkonen P. and Teräsvirta T.. Testing linearity against smooth transition autoregression models[J]. *Biometrika*, 1988, 75(3): 491-499.
- [18] Colletaz G. and Hurlin C.. Threshold effects of the public capital productivity: an international panel smooth transition approach[M]. Document de recherche du LEO, 2006, 01.

#### 作者简介

许永洪,男,厦门大学经济学院统计系副教授、博士生导师,福建省高等学校人文社会科学研究基地“厦门大学数据挖掘研究中心”副主任,计量经济学教育部重点实验室(厦门大学)和福建省统计科学重点实验室(厦门大学)研究员。研究方向为房地产价格理论、宏观经济统计分析、大数据应用和微观金融市场交易机制设计。

吴林颖,男,上海财经大学数量经济学在读硕士研究生。研究方向为宏观经济趋势分析、大数据算法应用和房地产价格理论。

(责任编辑:黄煌)