

# 基于贝叶斯模型平均(BMA)方法的中国房地产价格影响因素分析

卢二坡,张超

(安徽财经大学 统计与应用数学学院,安徽 蚌埠 233030)

**摘要:**对可能影响中国房价的诸多因素的重要性问题进行识别和检验,基于模型不确定性的视角,使用中国30个省(区)2002~2013年面板数据,采用贝叶斯模型平均(BMA)方法进行模型设定与分析。研究认为,在可能对中国房价产生影响的19个指标中,信贷政策、心理预期、物价水平、房屋竣工面积和产业结构合理化等5个解释变量的后验概率大于90%,它们是影响现阶段中国房地产价格的决定因素;应通过差别化的信贷政策分区域控制房价,通过新闻媒体公开统计和发布房地产数据正确引导人们的心理预期,通过适宜的货币政策有效控制物价,通过保障性住房建设增加房地产供给,通过合理化的产业结构引导房价调控等,促进中国房地产市场的健康发展。

**关键词:**房地产价格;贝叶斯模型平均方法;心理预期;信贷政策;产业结构

**中图分类号:**F224

**文献标志码:**A

**文章编号:**1671-6248(2016)04-0068-09

房地产市场的平稳运行,不只关乎着房地产业的发展问题,还牵涉着国计民生、百姓福利。然而近年来,在房地产市场持续繁荣过程中,房地产价格也一路疯涨,其增长速度已远远高过中国经济、居民收入及消费的增长速度。房地产价格的剧烈波动对中国经济的健康发展和人民群众生活的安定都带来了众多不利影响,纵然政府不断加大对过高和发展过快的房地产价格的调控力度,但仍未平抑房地产价格高涨的态势。因此,为了能够让调控政策更加有针对性,有必要对房地产价格的决定影响因素予以深入研究。通过经验研究找到影响房地产价格的决定因素,对于我们正确看待房地产价格过快上涨、政府如何采取调控措施等方面的问题具有重要的理论意义和现实意义。

## 一、文献综述

国外的房地产市场发展已较为完善,对于房地产价格的经验研究更是早于中国,在研究成果方面亦比中国充分和成熟,大多数的研究成果对他们本国的房地产市场发展起到了积极的引导作用,这些成果值得国内学者去借鉴和学习。Hort使用瑞典1968~1994年的时间序列数据并构建误差修正模型,结果发现,长期来看,居民收入、使用成本以及建筑成本等都对房地产价格产生显著影响<sup>[1]</sup>。Capozza等使用美国1979~1995年62个大都市的面板数据,采用面板数据模型,研究发现,收入水平、人口增长以及建筑成本等是导致房价变动的重要原因<sup>[2]</sup>。Seko选取日本1980~2001年46个县的面板

收稿日期:2016-07-08

基金项目:安徽财经大学研究生科研创新基金项目(CXJJ2014074)

作者简介:卢二坡(1976-),男,河南焦作人,教授,经济学博士。

数据,建立面板 VAR 模型,研究发现,日本各地区的住宅价格受居民收入、人口、物价水平的影响较大<sup>[3]</sup>。Pages 选用西班牙 1976~2002 年的时间序列数据及误差修正模型,研究得出收入和利率是房地产价格的主要影响因素<sup>[4]</sup>。Abelson 等使用澳大利亚 1970~2003 年的时间序列数据并采用普通最小二乘法,研究表明,可支配收入和物价水平对住宅价格产生重要促进效应,而失业率、利率、住宅存量以及股票价格对住宅价格上涨起抑制作用<sup>[5]</sup>。Chen 等采用微型仿真法估算人口老龄化对房价的影响,模拟结果说明,人口老龄化对苏格兰房价的影响不是主要的<sup>[6]</sup>。Ibrahim 等基于 VAR 模型对马来西亚的房价与银行信贷政策的关系进行研究,发现房价与银行信贷之间具有长期均衡关系<sup>[7]</sup>。

国外学者关于房地产价格影响因素的经验研究主要针对的是较为成熟的房地产市场,而对于处于经济转型期的中国来说,房地产市场起步较晚,因此,国外学者在对房地产价格影响因素进行经验研究时所考虑的变量并不一定适合于研究中国房地产价格。针对中国具体国情以及房地产市场的实际发展情况,国内学者也做了大量经验研究。沈悦等使用 1995~2002 年 14 个城市的面板数据并采用混合样本回归分析方法,研究发现,可支配收入、总人口和空置率对房地产价格产生正向作用,而失业率为对房地产价格具有负向作用<sup>[8]</sup>。崔光灿使用 1995~2006 年 31 个省市的面板数据并运用面板数据模型,结果表明利率对房价具有显著负向作用,物价水平具有显著正向作用,房屋竣工面积对房地产价格具有抑制作用,收入对房地产价格产生正向作用但不显著,人口因素对房价影响为负但也不显著<sup>[9]</sup>。邓长荣等使用中国 2002~2007 年的季度时间序列数据并采用误差修正模型,研究表明,收入、银行贷款、房屋供给、土地价格、土地供给是长期影响房价的主要因素<sup>[10]</sup>。罗孝玲等使用中国 2001~2010 年的季度时间序列数据,并采用 VAR 模型,结果发现,房价与 GDP、贷款利率之间均具有双向 Granger 因果关系,与物价水平具有单向因果关系<sup>[11]</sup>。胡晓使用中国 1997~2009 年 29 个省市面板数据,并构建固定效应模型,研究得出,可支配收入和建筑成本对房价具有正向作用,制度性因素对房价上涨产生重要作用,利率对房价存在负向作用但不显著<sup>[12]</sup>。王文雯等使用中国 2003~2010 年 27 个省市面板数据

并建立固定效应和随机效应模型,研究得出,城市化率、收入差距、人均可支配收入、利率政策和房地产开发投资额等都对房地产价格具有显著促进作用,人口老龄化、房屋竣工面积和土地政策等对房地产价格均具有显著抑制作用<sup>[13]</sup>。

然而,由于房地产价格是一个复杂的经济范畴,影响房地产价格的因素极其多元,因而研究者并不能确定应将哪些变量引入所建模型当中,换句话说,研究者们尚不能识别“真正”的模型究竟是什么样子。导致的结果是,在现有的经验研究中,不同的研究者在构建房地产价格回归模型时采用了不同的解释变量,因此在众多研究文献中可以看到各种不同的回归方程:

$$y = \alpha_1 + \beta_1 X + \theta_1 Z_1 + \varepsilon_1 \quad (1)$$

$$y = \alpha_2 + \beta_2 X + \theta_2 Z_2 + \varepsilon_2 \quad (2)$$

其中  $y$  是代表房地产价格的被解释变量,矩阵  $X$  是所有回归模型中共有的解释变量,矩阵  $Z_1$ 、 $Z_2$  代表不同模型中所特有的解释变量。通常来说,矩阵  $Z_1$ 、 $Z_2$  中所含有的解释变量种类以及个数都不一样, $\theta_1$ 、 $\theta_2$  都是系数项, $\varepsilon_1$ 、 $\varepsilon_2$  都是误差项,因此,通过比较这些不同模型的估计结果不难发现:不同模型中共有的解释变量  $X$  的系数不仅在大小、显著性等方面出现差异,而且符号可能存在相反的情况。此时,通过这些不同的单一线性模型分析房地产价格影响因素时往往会造成结果的偏差,会对最终的决策产生不利影响。这一现象的产生主要是由于这些研究者在建立回归模型时忽略了模型的不确定性。

Raftery 等、Fernandez 等证实 BMA 方法可以有效解决模型不确定性问题<sup>[14-15]</sup>。BMA 方法由于考虑了模型自身的不确定性,它在变量选择方面具有线性模型所不具有的优势。BMA 方法对包括不同解释变量的全部可能单一模型预设先验概率,在贝叶斯分析的框架内计算出所有潜在解释变量的后验概率,并通过后验概率的高低对所有解释变量的重要性予以排序;另外,还能够计算出各解释变量对应系数的后验均值、后验标准差以及符号确定率等反映变量重要性和稳健性的指标。BMA 方法由于其在解决模型不确定性问题上的优越表现被大量利用于多个领域,国外主要应用于研究通货膨胀、经济增长、教育回报、能源估算的影响因素等多个领域<sup>[16-20]</sup>,国内应用于研究经济增长、企业信用风险、主权债务偿付能力、通货膨胀、技术创新能力的影响

因素等多个领域<sup>[21-25]</sup>。由于影响房地产价格的因素众多,属于大信息集合,再结合以上有关 BMA 的应用文献不难发现,BMA 方法适合于对房地产价格影响因素进行研究,但国内尚未有人将 BMA 方法运用于房地产价格影响因素研究当中。

基于此,本文摒弃了人为事先设定模型及变量对房地产价格进行解释的传统,考虑了模型的不确定性,采用 BMA 方法并使用中国 30 个省(区)2002 ~ 2013 年面板数据,对所有可能影响中国房地产价格因素按照其后验概率大小进行分类和排序,最终确定影响中国房地产价格的决定因素。

## 二、模型、方法和数据

### (一)模型设定和估计方法

#### 1. 单一模型设定

用中国 30 个省(区)2002 ~ 2013 年面板数据进行 BMA 建模,面板数据可看作是截面数据和时间序列数据的合并,相对于在房地产价格影响因素研究中常用的时间序列分析方法来说,面板数据分析具有多方面的优势:第一,面板数据能够反映个体之间的差异和(或)时间上的差异,但时间序列数据无法反映出这些差异,会导致结果出现偏差;第二,面板数据可以提供更多的样本信息和样本容量,进而提高了数据的自由度并大大减少了解释变量间的共线性,故而提高了模型估计的有效性,而普通的时间序列数据则不能做到这些,当解释变量较多时,时间序列数据将会存在样本数不足和多重共线性问题;第三,面板数据分析能够更好辨识纯时间序列发现不了的影响因素,这在一定程度上降低了因遗漏重要变量导致的偏差。

因为自身拥有诸多优势,面板数据模型不断被众多学者所使用,进而也推动了面板数据模型的更深层次的研究。固定效应和随机效应模型是两种基本的面板数据模型。下文将采用固定效应模型作为 BMA 方法中的单一模型,主要原因为:第一,随机效应模型假定个体效应与模型中的解释变量不相关,而固定效应模型假定个体效应与解释变量相关,显然对于本文研究问题来说后者的假定更为适合;第二,随机效应模型认为个体效应的差异与个体的特征不存在内在联系,即这种差异可能是随机产生的,而固定效应模型反映个体效应的差异与个体特征存

在内在联系,也就是说该个体效应是固定的,结合本文的研究目标发现后者情形更适合于本文研究。

综上所述,为了探究中国房地产价格影响因素,需遵循 Moral-Benito<sup>[26]</sup>和 Cuaresma 等<sup>[27]</sup>的做法,将单一模型设定为个体固定效应模型,即:

$$y_{it} = \beta_0 + \beta_1 X_{it} + \eta_i + v_{it} (t = 1, 2, \dots, T) \\ (i = 1, 2, \dots, N) \quad (3)$$

其中, $y_{it}$ 为*i*地区在*t*期的房地产价格,以商品房平均销售价格表示; $\beta_0$ 为常数项; $\beta_1$ 为待估参数向量; $X_{it}$ 表示影响中国房价的因素; $\eta_i$ 代表无法观测的个体固定效应; $T$ 为时期; $N$ 为地区; $v_{it}$ 为独立同分布的经典误差项。设定以上模型能够一定程度上减少遗漏重要解释变量所造成的偏误。

基于对房价影响因素相关文献的研究,罗列影响房价的因素。假定模型空间  $M = \{M_1, M_2, \dots, M_N\}$ ,代表全部潜在解释变量任意结合所生成的模型空间,如果罗列出的作为解释变量的因素共有  $K$  个,由于这些变量的任意一种组合都可以得到一个单一模型,意味着由这  $K$  个解释变量的任意组合将会生成  $2^K$  个模型,即模型空间  $M$  中共包括  $N = 2^K$  个模型,而最优模型是什么模型先前尚不知晓,也就是说存在模型本身的不确定性。根据这  $2^K$  个模型的回归结果,利用 BMA 方法就可计算出所有解释变量的后验概率以及其系数的稳健性指标。

假定  $M_j$  表示模型空间  $M$  中第  $j$  个模型, $p(M_j)$  表示模型  $M_j$  的先验概率,基于贝叶斯理论,便可以得到表示模型  $M_j$  的后验模型概率:

$$p(M_j | y_u, X_u) = \frac{p(y_u | M_j, X_u)p(M_j)}{p(y_u | X_u)} = \frac{p(y_u | M_j, X_u)p(M_j)}{\sum_{s=1}^{2^K} p(y_u | M_s, X_u)p(M_s)} \quad (4)$$

这里, $p(y_u | M_j, X_u) = \int p(y_u | \beta_{ij}, M_j, X_u) p(\beta_{ij} | M_j, X_u) d\beta_{ij}$ 是模型  $M_j$  的边际似然值, $p(y_u | \beta_{ij}, M_j, X_u)$ 是指模型  $M_j$  对应的似然函数, $p(\beta_{ij} | M_j, X_u)$ 是指模型  $M_j$  中参数向量  $\beta_i$  的先验概率分布, $M_s$  指任一单一模型; $\beta_{ij}$ 指模型  $M_j$  的参数变量  $b_i$ ;  $d$  为不定积分中的微分符号。

计算出模型后验概率后,将其作为权重,通过加权平均的形式可以得到参数向量  $\beta_i$  的后验概率、后验均值、后验方差:

$$p(\beta_i | y_u, X_u) = \sum_{j=1}^{2^K} p(\beta_i | M_j, y_u, X_u) p(M_j | y_u, X_u) \quad (5)$$

$$E(\beta_1 | y_u, X_u) = \sum_{j=1}^{2^K} E(\beta_1 | y_u, X_u, M_j) \\ p(M_j | y_u, X_u) \quad (6)$$

$$Var(\beta_1 | y_u, X_u) = \sum_{j=1}^{2^K} Var(\beta_1 | y_u, X_u, M_j) \\ p(M_j | y_u, X_u) \quad (7)$$

这里,  $p(\beta_1 | M_j, y_u, X_u)$  是在模型  $M_j$  条件下参数  $\beta_1$  的后验概率,  $p(M_j | y_u, X_u)$  是模型  $M_j$  后验概率,  $E(\beta_1 | y_u, X_u, M_j)$  与  $Var(\beta_1 | y_u, X_u, M_j)$  为单一模型中参数向量的后验均值和后验方差。

## 2. 参数和模型先验概率设定

由上可知,利用 BMA 方法解决模型不确定性问题之前必须设定先验概率分布。Kass 等研究发现,基于模型的不确定性问题,参数先验概率分布的设定对模型后验分布存在关键影响<sup>[28]</sup>。George 等认为扩散单位信息先验分布(简称 DUIP)作为一种基准先验概率分布具备很高的安全性和稳健性<sup>[29]</sup>,所以为了增加先验概率分布的稳健性和安全性,可将参数先验概率分布设定为 DUIP。参数先验概率分布设定完成,还需要设置模型先验概率分布,该步骤为显现模型不确定性的关键步骤。针对一个包括  $2^K$  个模型的模型空间,通常设定所有模型(以  $M_\gamma$  表示)都服从均匀先验概率分布,为

$$p(M_\gamma) = p_\gamma = 2^{-K}, \gamma = 1, 2, \dots, 2^K, p_\gamma > 0 \\ \text{且} \sum_{\gamma=1}^{2^K} p_\gamma = 1 \quad (8)$$

这意味着,引进或剔除任一模型的先验概率为 0.5,也就是说先前对模型的选择不存在偏好,所以均匀分布为在缺少先验信息的条件下模型先验分布的标准选择。

## 3. 估计方法

在对单一模型、参数和模型的先验概率设定完成后,即可利用式(4)~(7)基于迭代的方法对参数和模型的后验概率进行计算,本文涉及 19 个解释变量,在个体固定效应模型结构的假设下,备选模型总数高达  $2^{19}$  个。但 George 等研究发现,利用迭代的方法计算后验概率的工作量非常之大,对所有模型作加权平均几乎不可能完成<sup>[30]</sup>。因此,处于减轻计算工作量的目的,借鉴 Madigan 等提出的马尔科夫蒙特卡罗模型综合( $MC^3$ )抽样技术对模型空间进行抽签<sup>[31]</sup>,对抽签得到的新样本再根据公式(4)~(7)进行迭代计算参数和模型的近似后验概率。 $MC^3$  算法倾向抽取较高后验概率的模型,当达到一定抽样数时,可以保证抽样结果收敛于迭代计算的结果。

$MC^3$  抽样技术以特定的概率分布对模型空间  $M = \{M_1, M_2, \dots, M_N\}$  进行抽签抽取备选模型  $M^*$ ,

再以特定的概率接受此备选模型。在第  $s$  次抽签中,如若备选模型被接受,则  $M^* = M^{(s)}$ ,如若没有被接受,则接受上一次抽签模型,即  $M^* = M^{(s-1)}$ 。第  $s$  次抽签备选模型有如下 3 种模型选择:(1)前一次抽样模型  $M^{(s-1)}$ ;(2)从  $M^{(s-1)}$  中减少一个解释变量的模型;(3)从  $M^{(s-1)}$  中增加一个解释变量的模型。当备选模型成功抽取后,将会以如下接受概率接受此备选模型:

$$\alpha(M^{(s-1)}, M^*) = \min \left[ \frac{p(y | M^*)p(M^*)}{p(y | M^{(s-1)})p(M^{(s-1)})}, 1 \right] \quad (9)$$

其中  $p(y | M^*)$  和  $p(y | M^{(s-1)})$  为模型边际似然值。由于假定所有模型均服从均匀先验分布,即所有模型的先验概率相同,所以式(9)简化为计算模型  $M^*$  对模型  $M^{(s-1)}$  的贝叶斯因子:  $\frac{p(y | M^*)}{p(y | M^{(s-1)})}$ 。

$MC^3$  的抽样原理将会使抽签过程中拥有更多有效信息的变量被接受,进而会促使具有更高后验概率的模型被抽中。

## (二) 样本变量和数据说明

样本资料是中国 2002~2013 年 30 个省(区)的面板数据,由于西藏自治区资料缺失较多,故没有将其列入。所有的原始资料都来源于《中国统计年鉴》、各地区相关年份的《统计年鉴》以及中宏网统计数据库。

### 1. 被解释变量设定

能作为房地产价格最合理的指标是房屋重复销售价格或价格指数,由于中国房地产市场发展时间较短,缺乏相应的指标,故无法获得全面的数据资料,但由于中国房地产市场上的交易以新建商品房为主,商品房平均销售价格基本上可以反映一个地区的房地产价格情况。因此,本文被解释变量选取了中国 30 个省(区)的商品房平均销售价格作为代理变量,记为  $p_u$ ,与式(3)中的  $y_u$  意义一致。

### 2. 解释变量设定

对房地产价格产生影响的变量诸多,基于变量数据的可靠性与可得性,以及结合经济学理论框架,从需求层面、供给层面、经济发展层面、政策层面以及其他层面等 5 个层面选取了 19 个可解释变量。其中包括需求层面的消费者收入水平  $A_u$ 、城乡收入差距  $B_u$ 、城镇化率  $C_u$ 、人口老龄化  $P_u$ 、失业率  $D_u$  以及人口密度  $E_u$ ;供给层面的土地价格  $F_u$ 、建筑成本  $G_u$ 、房地产投资额  $H_u$  以及房屋竣工面积  $I_u$ ;经济发展层面的经济增长  $J_u$ 、产业结构合

理化  $L_{it}$ 、产业结构升级化  $N_{it}$ 以及物价水平  $O_{it}$ ;政策层面的利率政策  $Q_{it}$ 、信贷政策  $R_{it}$ 、土地政策  $S_{it}$ 以及税收政策  $U_{it}$ ;其他层面的心理预期  $W_{it}$ 。表 1 是指标来源说明以及处理方法。

表 1 指标来源说明以及处理方法

变量类型	指标名称	指标来源说明	处理方式
被解释变量	商品房平均销售价格( $p$ )	商品房平均销售价格	1
	消费者收入水平( $A$ )	城镇居民人均可支配收入	1
	城乡收入差距( $B$ )	泰尔指数计算得来	0
需求层面解释变量	城镇化率( $C$ )	城镇总人口除以总人口	0
	人口老龄化( $P$ )	城镇老年抚养比	0
	失业率( $D$ )	城镇登记失业率	0
	人口密度( $E$ )	年末总人口除以总面积	0
	土地价格( $F$ )	房地产土地购置费用除以购置费用	1
供给层面解释变量	建筑成本( $G$ )	建筑安装工程投资价格指数	0
	房地产投资额( $H$ )	房地产开发投资总额	1
	房屋竣工面积( $I$ )	商品房住宅竣工面积	0
	经济增长( $J$ )	人均国内生产总值	1
经济层面解释变量	产业结构合理化( $L$ )	泰尔指数计算得来 <sup>[32]</sup>	0
	产业结构升级化( $N$ )	非农产业产值占国民经济总产值比重	0
	物价水平( $O$ )	消费者价格指数	0
	利率政策( $Q$ )	实际利率水平	0
政策层面解释变量	信贷政策( $R$ )	房地产开发非自筹资金与商品房销售额的 80%之和 <sup>[33]</sup>	1
	土地政策( $S$ )	商品房开发企业购置土地面积	0
	税收政策( $U$ )	房产税	1
其他层面解释变量	心理预期( $W$ )	滞后一期的房地产价格	1

注:处理方式:“0”代表原始数据,“1”代表消除价格因素。

对于以上所有价值型变量,利用各省以 2002 年

为基期的 CPI 剔除相关价格和收入指标中的价格因素的影响。为了剔除异方差和量纲的干扰,本文对全部指标均作对数化处理。

三、实证结果及分析

19 个解释变量在个体固定效应模型结构的假设下,备选模型总计  $2^{19}$  个。采用  $MC^3$  抽样技术进行模型抽样抽取其中 501 000 个模型。初始的 1 000 次抽样模型作为预热实验,其后的 500 000 次抽签作为模拟样本。

(一) 实证结果描述

本文使用 R-3. 1. 1 软件,调用综合了 BMA 方法和  $MC^3$  抽样技术的 BMS 软件包,对房地产价格影响因素模型进行估计,估计结果描述如下。

图 1 为中国房地产价格后验模型规模分布均值图,容易发现模型先验概率分布图确实是呈均匀分布的,这与之前的先验假设相符;模型后验概率分布图在横轴为 10. 093 4 时达到最高点,因此后验模型规模分布均值为 10. 093 4,即在抽取的 500 000 个模型中,每个单一模型平均大概含有 10 个解释变量。解释变量的后验概率就是所有包含该变量的模型后验概率之和。

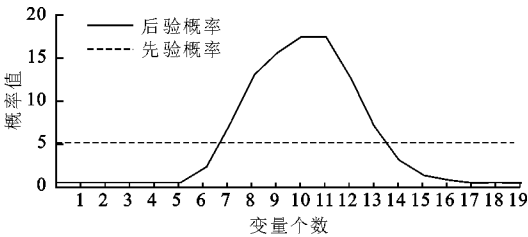


图 1 中国房地产价格后验模型规模分布均值

表 2 给出了全国房地产价格对各影响因素回归的 BMA 估计结果,该结果是在所抽到的全部单一模型回归结果的基础上采用 BMA 方法估计得到的。

表 2 中,第二列的后验概率是根据公式(5)计算而来,是所有含有该变量的模型后验概率之和,它表示解释变量对被解释变量的解释力度(也可理解为对被解释变量的重要程度),且表 2 中 19 个解释变量是按后验概率大小进行排序的。第三列表示解释变量的后验均值,根据公式(6)计算得来;第四列为解释变量的后验标准差,是由公式(7)所得结果开方得到;第五列为解释变量符号确定率,它是指所有含有该解释变量的估计结果中,该变量的系数符号与后验均值符号一致的几率,以此反映变量的稳

健性。

表 2 全国房地产价格对各影响因素回归的 BMA 估计结果

序号	变量名称	后验概率	后验均值	后验标准差	符号确定率
1	信贷政策( $R$ )	1.000 0	0.267 9	0.025 2	1.000 0
2	心理预期( $W$ )	1.000 0	0.544 1	0.035 1	1.000 0
3	物价水平( $O$ )	0.999 9	-0.266 1	0.062 4	1.000 0
4	房屋竣工面积( $I$ )	0.998 3	-0.120 5	0.031 7	1.000 0
5	产业结构合理化( $L$ )	0.929 0	-0.053 6	0.022 5	1.000 0
6	人口密度( $E$ )	0.728 5	0.248 9	0.187 7	1.000 0
7	房地产投资额( $H$ )	0.712 1	-0.078 2	0.062 7	1.000 0
8	土地价格( $F$ )	0.663 4	0.028 2	0.026 1	1.000 0
9	失业率( $D$ )	0.580 3	0.061 3	0.061 5	1.000 0
10	消费者收入水平( $A$ )	0.532 2	0.103 9	0.113 7	0.999 9
11	税收政策( $U$ )	0.399 0	0.026 0	0.037 2	1.000 0
12	土地政策( $S$ )	0.394 4	0.013 7	0.020 5	0.947 5
13	城乡收入差距( $B$ )	0.390 5	-0.023 1	0.033 7	1.000 0
14	建筑成本( $G$ )	0.220 5	0.037 6	0.086 2	1.000 0
15	经济增长( $J$ )	0.147 9	0.008 7	0.030 7	0.916 4
16	人口老龄化( $P$ )	0.127 3	-0.007 0	0.025 2	0.986 8
17	产业结构升级化( $N$ )	0.106 0	0.029 5	0.128 1	0.976 1
18	城镇化率( $C$ )	0.089 5	0.002 5	0.014 1	0.963 5
19	利率政策( $Q$ )	0.074 3	-0.000 8	0.011 8	0.771 2

由表 2 中的后验概率可将 19 个解释变量大致划为 4 组:

第一组的解释变量包括信贷政策( $R$ )、心理预期( $W$ )、物价水平( $O$ )、房屋竣工面积( $I$ )以及产业结构合理化( $L$ )等 5 个变量。不难发现,这 5 个解释变量的后验概率都高于 0.9,意味着它们对中国房地产价格的解释能力较强;同时,这些变量的符号确定率也都达到 100%,表明它们的稳健性非常好。此外,这 5 个变量的后验均值与后验标准差的比值(取绝对值,下同)均大于 2。综上,可以推出信贷政策( $R$ )、心理预期( $W$ )、物价水平( $O$ )、房屋竣工面积( $I$ )以及产业结构合理化( $L$ )这 5 个变量是中国房地产价格影响因素模型中极为重要的解释变量。

第二组的解释变量有人口密度( $E$ )、房地产投资额( $H$ )、土地价格( $F$ )、失业率( $D$ )以及消费者收入水平( $A$ )等 5 个变量。这 5 个解释变量的后验概率均大于 0.5 且小于 0.9,表示它们对中国房地产价格同样具有较好的解释能力;观察符号确定率,它们都表现出较好的稳健性。同时,这些变量的后验均值与后验标准差的比值大部分都大于 1。因此,以上 5 个变量也是影响中国房价的重要因素。

第三组包括税收政策( $U$ )、土地政策( $S$ )、城乡

收入差距( $B$ )、建筑成本( $G$ )等 4 个变量。这 4 个变量的后验概率都是介于 0.2 与 0.5 之间,说明它们可以对房价进行一定的解释,观察符号确定率发现,这 4 个变量均显示良好的稳健性。另外,这些变量的后验均值与后验标准差的比值都在 0.2 与 0.7 之间。因而这些变量对中国房价来说仍具有一定的重要性。

第四组包括经济增长( $J$ )、人口老龄化( $P$ )、产业结构升级化( $N$ )、城镇化率( $C$ )和利率政策( $Q$ )等 5 个解释变量。由于这些变量的后验概率均不足 0.2 且符号确定率基本上都小于之前三组,意味着本组变量的解释力度最弱,且稳健性也不高,除非我们对这 5 个变量抱有较强的理论根据或较强烈的主观信念,否则将这些变量引入回归方程的几率不是很大。

(二) 结果分析

Raftery 等、Fernadez 等国外学者定义<sup>[14-15]</sup>:后验概率大于 90% 的解释变量称作最有效解释变量,大于 20% 的解释变量称作有效解释变量。从表 2 的实证结果可知,在可能对中国房价产生影响的 5 个层面的 19 个指标中,信贷政策( $R$ )、心理预期( $W$ )、物价水平( $O$ )、房屋竣工面积( $I$ )以及产业结构合理化( $L$ )等 5 个解释变量的后验概率超过 90%,这意味着这 5 个解释变量对中国房地产价格具有长期稳健影响,它们是最有效解释变量,即此 5 个解释变量是中国房地产价格的决定影响因素,具体分析如下。

信贷政策( $R$ )的后验概率为 1.000 0,且该指标对房地产价格的影响为正,说明信贷政策对房地产价格调控作用明显。在中国房地产市场发展的过程中,经济调控政策是影响房价的主要政策,而信贷政策是最有效的调控政策,信贷规模已成为中国房地产市场的向导,因此采取紧缩的信贷政策可以有效抑制房价上涨。

心理预期( $W$ )的后验概率为 1.000 0,且该指标对房地产价格的影响为正,揭示出人们看涨的心理预期对中国房价上涨起到主导作用。“未来房价肯定上涨”这是现在只要对房地产有一定了解的大多数中国人都能够得出的“结论”,但“未来房价肯定上涨”这一“简单”预测并不只是房价变动的“结果”,更是造成房价上涨的重要原因之一。回顾中国房地产市场实际发展过程,2007 年房地产价格过快上涨后,人们普遍预期房地产价格还会大涨,对于消费者来说,对房地产市场需求量增长迅速,对房地

产开发企业来说,它们会大量增加房地产供给量,带来的结果就是房价继续快速增长。

物价水平( $O$ )的后验概率为 0.999 9,且该指标对房地产价格的影响为负,说明物价水平的上升对房地产价格产生了重要抑制作用。可能的原因是,从需求角度,近年来中国政府不断实施严厉的房价政策对投机性需求进行镇压,世联地产统计显示 2013 年投资投机性需求占比从 2010 年的 46.7% 降至 14%,由此可见投机性需求逐渐被挤出市场,导致的结果是即使物价上涨,投机性需求也不会上升很多,而与此同时,中低收入人群的自主性和改善性需求受物价上升的影响较大,因此,总的来说物价水平将会导致房地产总需求下降;从供给角度,物价的上涨抬高了房地产成本,开发商受资金限制便会缩减房地产供应。综上,物价从供需两个方面对房价上涨起抑制作用。此外,物价上涨将会驱使央行采取紧缩的货币政策,进而也会造成房价下跌。但中国政府并不能依靠提高物价水平来遏制房价上涨,因为这将“误伤”中低等人群的自主性和改善性需求,应当将物价控制在合理区间并结合相关调控政策对房价进行理性调控。

房屋竣工面积( $I$ )的后验概率为 0.998 3,且该指标对房地产价格的影响为负,现阶段中国房地产市场总体上供过于求,供需失衡的矛盾集聚凸显,虽然目前各个省市的房地产价格仍呈上涨趋势,但上涨速度有所减缓,这是由于明显偏多的房屋供给量遏制了房地产价格的快速上涨,若房地产供给量过度增多,甚至可能会出现房地产价格不断下跌现象,让房价暴涨固然不好,但让房价暴跌更不好,而是让它适当的上涨,因此政府应该通过适当增加房地产供给量对高房价进行遏制。

产业结构合理化( $L$ )的后验概率为 0.929 0,且该指标对房地产价格的影响为负,说明产业结构合理化的加快是推动中国房价上涨的关键“引擎”。由此可见,地区房地产市场的发展需要一定的产业支撑,产业之间的合理配置是维系房地产市场价格秩序的关键,若仅仅依靠某一项或几项支柱产业的支撑发展房地产必将受挫。内蒙古鄂尔多斯市就是一个典型个案,其在煤炭被当做“黑金”的时候曾风光一时,该城市由此得到快速发展,并一时成为资本的聚集地,“炒房团”蜂拥而至,房地产业不断扩大,但当煤炭风光不在,城市经济大幅下滑,房地产市场几近“崩溃”,大量闲置的楼盘更是被冠以“鬼城”恶名。因此,合理的产业结构是地区经济发展和房地

产市场发展的重要基石,但将产业结构合理度控制在一定区间内是房地产市场健康发展和房价回归合理的必要前提。

此外,人口密度( $E$ )、房地产投资额( $H$ )、土地价格( $F$ )、失业率( $D$ )、消费者收入水平( $A$ )、税收政策( $U$ )、土地政策( $S$ )、城乡收入差距( $B$ )以及建筑成本( $G$ )9 个解释变量的后验概率低于 90%,但都高于 20%,说明这 9 个解释变量对中国房地产价格亦具有一定的解释能力,但它们的解释能力远不及 5 个长期决定因素的解释能力,它们属于有效解释变量。剩余的 5 个解释变量的后验概率都低于 20%,意味着这些变量的解释力度最弱,尚不能成为影响中国房地产价格的有效解释变量。

## 四、对策和建议

从影响中国房地产价格的 5 个决定因素着手,针对如何科学有效控制中国房地产价格以及如何促进中国房地产市场的健康发展提出以下政策建议。

第一,执行差别化信贷政策。中国目前的信贷政策对中国房价产生重要促进作用,因此为了控制房价在合理区间,中国政府应当严格执行差别化信贷政策,对于房地产市场发展过快,房价过高的区域,应该限制该区域的信贷规模,防止房地产投机带来金融风险。对于房地产市场发展落后于当地经济发展水平的区域,应适时松绑信贷政策,推动该地区房地产市场健康稳定发展。

第二,正确引导人们心理预期。心理预期因素对中国房价产生重要的促进作用。因此为了控制房价在合理区间,中国政府部门应当将土地供应计划和城市规划的制定原则、具体内容和实现步骤公布于众,避免人们对市场供应能力的担忧,进而理性的安排购买行为。中国政府机构应当改善房地产数据的统计和发布,使公众明白房价上涨的缘由和基本趋势,帮助人们合理心理预期的形成。新闻媒体要如实报道,不应帮助房地产开发商煽风点火,推涛作浪,而应该正确引导投资者进行理性投资。

第三,对物价进行有效控制。物价水平对中国房价产生重要抑制作用,为了减轻物价水平对房地产市场的影响,政府须制定适宜的货币政策,避免物价水平产生虚高现象,构建合适的物价监督机制,有关物价监督部门必须严格依据法律法规对物价进行有效控制,使物价水平处于合理区间,以促进房地产市场的稳定。

第四,落实保障性住房政策,增加房地产有效供给。房屋竣工面积(房地产供给量)对中国房价产生重要抑制作用,因此为了控制房价在合理区间,中国政府应加大普通商品房、经济适用房以及廉住房的供给比重,激活二手房存量住房市场,增加房地产有效供给,将保障性住房建设及房地产供给结构调整政策的落实情况列入地方政府官员政绩考评体系中。

第五,将产业结构合理化引入房价调控决策当中。产业结构合理化对中国房价产生重要促进作用,因此中国政府在对房价进行调控时,应因地制宜考虑产业结构合理化的区域差异,把产业结构合理化引入房价调控的决策集中,除去产业结构合理化对房价上涨的合理驱动,以考察各区域房价是否理性上涨。

## 五、结语

本文以中国 30 个省(区)2002~2013 年房地产价格及相关影响因素作为样本,引入基于模型不确定性的贝叶斯模型平均方法(BMA)对中国房价的影响因素进行研究。通过分析研究,本文得出了如下结论:第一,信贷政策、心理预期、物价水平、房屋竣工面积以及产业结构合理化等 5 个解释变量对中国房地产价格具有长期稳健的影响。第二,人口密度、房地产投资额、土地价格、失业率、消费者收入水平、税收政策、土地政策、收入差距以及建筑成本等 9 个解释变量显示出良好的稳健性,这些变量对中国房价来说仍具有一定的重要性。第三,经济增长、人口老龄化、产业结构升级、城镇化率和利率政策等 5 个解释变量对中国房地产价格的解释力度最弱,且稳健性也不高。基于以上结论,本文针对如何科学有效控制中国房地产价格以及如何促进中国房地产市场的健康发展提出若干政策建议。

### 参考文献:

[1] Hort K. The determinants of urban house price fluctuations in Sweden 1968-1994[J]. Journal of Housing Economics, 1998, 7(2): 93-120.

[2] Capozza D R, Hendershott P H, Mack C, et al. Determinants of real house price dynamics[EB/OL]. (2002-10-01) [2016-06-03]. <http://www.nber.org/papers/w9262.pdf>

[3] Seko M. Effects of land taxes on housing markets in Ja-

pan[J]. Toshi Mondai the Journal of the Tokyo Institute for Municipal Research, 2001(92):55-67.

[4] Pagés J M, Maza L A. Analysis of house prices in Spain [EB/OL]. (2003-01-01) [2016-06-03]. <http://www.bde.es/f/webbde/SES/Secciones/Publicaciones/PublicacionesSeriadas/DocumentosTrabajo/03/Fic/dt0307e.pdf>

[5] Abelson P, Joyeux R, Milunovich G, et al. Explaining house prices in Australia: 1970-2003[J]. Economic Record, 2005, 81(s1): S96-S103.

[6] Chen Y, Gibb K, Leishman C, et al. The impact of population ageing on house prices: a micro-simulation approach[J]. Scottish Journal of Political Economy, 2012, 59(5): 523-542.

[7] Ibrahim M H, Law S H. House prices and bank credits in Malaysia: an aggregate and disaggregate analysis[J]. Habitat International, 2014(42): 111-120.

[8] 沈悦, 刘洪玉. 住宅价格与经济基本面: 1995~2002 年中国 14 城市的实证研究[J]. 经济研究, 2004(78): 78-86.

[9] 崔光灿. 房地产价格与宏观经济互动关系实证研究[J]. 经济理论与经济管理, 2009(1): 57-62.

[10] 邓长荣, 马永开. 我国住宅价格多层次因素模型及其实证研究[J]. 系统工程理论与实践, 2010(1): 48-55.

[11] 罗孝玲, 洪波, 马世昌. 基于 VAR 模型的房地产价格影响因素研究[J]. 中南大学学报: 社会科学版, 2012(4): 01-07.

[12] 胡晓. 我国房地产价格上涨背后的制度性因素——兼论房地产价格泡沫[J]. 中央财经大学学报, 2014(7): 91-99.

[13] 王文雯, 陶永亮. 中国房地产价格因素分析[J]. 统计研究, 2014(7): 111-112.

[14] Raftery A E, Madigan D, Hoeting J A. Bayesian model averaging for linear regression models[J]. Journal of the American Statistical Association, 1997, 92(437): 179-191.

[15] Fernandez C, Ley E, Steel M F J. Benchmark priors for Bayesian model averaging[J]. Journal of Econometrics, 2001, 100(2): 381-427.

[16] Cogley T, Sargent T J. The conquest of US inflation: learning and robustness to model uncertainty[J]. Review of Economic Dynamics, 2005, 8(2): 528-563.

[17] Koop G, Korobilis D. Forecasting inflation using dynamic model averaging[J]. International Economic Review, 2012, 53(3): 867-886.

[18] Ley E, Steel M F J. On the effect of prior assumptions in Bayesian model averaging with applications to growth regression[J]. Journal of applied econometrics, 2009, 24(4): 651-674.



- [19] Ley E, Steel M F J. Mixtures of g-priors for Bayesian model averaging with economic applications[J]. Journal of Econometrics, 2012, 171(2): 251-266.
- [20] Culka M. Applying Bayesian model averaging for uncertainty estimation of input data in energy modelling[J]. Energy, Sustainability and Society, 2014,4(1): 1-17.
- [21] 王亮,刘金全. 中国经济增长的决定因素分析——基于贝叶斯模型平均(BMA)方法的实证研究[J]. 统计与信息论坛, 2010, 25(9): 3-7.
- [22] 高丽君. 基于贝叶斯模型平均生存模型的中小企业信用风险估计[J]. 中国管理科学, 2012, 20(S1): 327-331.
- [23] 司明,孙大超. 发达国家主权债务危机成因分析及启示——基于贝叶斯模型平均方法的实证研究[J]. 中南财经政法大学学报, 2013, 4: 86-92.
- [24] 陈伟,牛霖琳. 基于贝叶斯模型平均(BMA)方法的中国通货膨胀的建模及预测[J]. 金融研究, 2013(11): 15-27.
- [25] 许培源,张华. 福建省技术创新能力的影响因素——基于贝叶斯模型平均方法的研究[J]. 东南学术, 2014(1): 124-133.
- [26] Moral-Benito E. Determinants of economic growth: a Bayesian panel data approach[J]. Review of Economics and Statistics, 2012, 94(2): 566-579.
- [27] Cuaresma J C, Doppelhofer G, Feldkircher M. The determinants of economic growth in European regions[J]. Regional Studies, 2014, 48(1): 44-67.
- [28] Kass R E, Raftery A E. Bayes factors[J]. Journal of the American Statistical Association, 1995(430): 773-795.
- [29] George E I, McCulloch R E. Approaches for Bayesian variable selection[J]. Statistica Sinica, 1997(2): 339-373.
- [30] George E I, McCulloch R E. Approaches for Bayesian variable selection[J]. Statistica Sinica, 1997(2): 339-373.
- [31] Madigan D, York J, Allard D. Bayesian graphical models for discrete data[J]. International Statistical Review, 1995(63): 215-232.
- [32] 干春晖,郑若谷,余典范. 中国产业结构变迁对经济增长和波动的影响[J]. 经济研究, 2011(5): 4-16.
- [33] 周京奎. 房地产泡沫生成与演化——基于金融支持过度假说的一种解释[J]. 财贸经济, 2006(5): 3-10.

## Analysis of determinants of real estate prices in China based on Bayesian Model Averaging(BMA)

LU Er-po, ZHANG Chao

(Institute of Statistics and Applied Mathematics, Anhui University of Finance and Economics, Bengbu 233030, Anhui, China)

**Abstract:** This paper identified and checked the determinants of real estate prices in China, and adopted thirty provinces (autonomous regions) panel data from 2002 to 2013 to do model specification and analysis by use of bayesian model averaging method from the perspective of model uncertainty. The results show that in the nineteen factors of five levels that may have an impact on China's real estate prices, the posterior probability of five variables is more than 90%, which are credit policy, psychological expectations, price levels, completion of housing area, and the rationalization of the industrial structure. They are the determinants of China's real estate prices at the present stage. Therefore, in order to promote the good development of China's real estate market, differential credit policy should be adopted to control real estate price according to different zones, the real estate data should be collected publicly and published via the media to lead masses' psychological expectation, appropriate monetary policy should be made to control prices levels effectively, indemnificatory housing should be increased to meet the real estate supply, and industrial structure rationalization should be improved to adjust housing price.

**Key words:** real estate prices; bayesian model averaging method; determinants; psychological expectation; credit policy; industrial structure