

# 住宅市场的价格特征:以北京为例\*

蔡 真 汪利娜

〔摘 要〕本文首先对住宅价格三种统计方法进行对比分析,说明特征价格法在住宅价格统计中的优势;然后说明了特征价格法在实践操作中的建模思路;最后对北京的住宅价格在截面和时间序列两个维度进行了计算,并与国家统计局的房价指数进行了对比。具体有以下主要结论:第一,北京住宅价格表现出明显的区位特征,以天安门为中心距离往外每扩展 1 公里,房价下降约 4.5%;第二,将北京划分南北两个区域,北部房价整体高于南部,但北部房价下降的幅度和速度都高于南部;第三,邻里特征并非简单的线性关系,当周边环境设施与小区位置过近时,其对房价的影响为负,只有在适当的距离范围,其对房价的影响最大,例如公共场所对住宅价格影响最大的恰当距离是 800 米范围;第四,建筑特征的表现是,住宅价格随户型增多而下降,但在相同户型内随面积增大而上升,此外整体上毛坯房比装修房便宜。

关键词:住宅市场 特征价格法 建筑特征 邻里特征 区位特征

JEL 分类号:C81 R31 R53

## 一、引 言

住宅价格问题是住宅市场的基本问题,然而由于统计方法缺陷以及主观选择偏差,导致我国目前住宅价格指数并不能客观地反映市场运行情况。就当前房价统计方法而言,采取的主要是非同质方法,一般是平均法。该方法的最大缺陷是忽略了住宅自身的差异因素,在现实中由于新建房样本大都集中在郊区,往往在这一区位因素主导下导致房价统计系统性偏低。住宅特征价格模型是国外房价研究中一种相对成熟的方法,特征价格的英文为 Hedonic Prices, Hedonic 的英文原意为“享乐的”,这里取其“能够给人带来效用”之意。因此特征价格模型的含义可以理解为:商品的价格应等于该商品所包含的一切能够满足人们需求的特征的价格之和。

采用特征价格方法计算住宅价格,具有微观和宏观两方面的实践意义。就微观角度而言,住房购买者可在邻里和区位特征相同的条件下,比较建筑特征和结构特征的差异,从而决定自己购买的住宅是否存在套利空间,比如如果毛坯房相对于精装修的住宅,价格的便宜程度大于装修成本,那么就存在套利空间。住宅开发商可通过一座城市的距离和邻里特征,决定自己的开发选址是否有利。就宏观角度而言,由于我们分解了住宅的价格特征,也就得到了同质条件下的房价,从而也就客观地刻画了住宅市场的变动,这对房地产宏观调控、银行压力测试、土地财政可持续性、房地产税等课题的研究均至关重要。

自上世纪 60 年代以来,国外住宅特征价格研究蓬勃发展,我们将有关文献大致将其分为三个方面进行简要回顾:第一,用于对住宅周边公共产品的分析,这其中研究较多的是特征包括轨道交通、公园、学校等。Grass(1992)研究发现临近地铁、轻轨会提升住房的价格;但 Forrest et al(1996)

\* 蔡真,中国社会科学院金融研究所,副研究员,经济学博士;汪利娜,中国社会科学院经济研究所,研究员。本文基础数据得到禧泰房产数据公司的支持,在此表示感谢。感谢中国社会科学院金融研究所博士后牛新艳、硕士研究生何晓星对基础数据处理工作的贡献。作者感谢程炼、李广子提供的有益评论,当然文责自负。

对曼彻斯特轻轨线路的研究发现邻近轻轨车站会降低住房价格。实际上产生两种对立结果源自于轨道交通对居住环境的影响是两方面的,一方面它为出行提供了交通便利性,但另一方面其产生的噪音影响居住的健康。Lutzenhiser and Netusil(2001)研究了都市公园、天然公园和专业公园对住宅价格的影响,其结论是三者中天然公园的影响最大。Clauretie and Neill(2000)用学生的 SAT 成绩作为学校质量的代理变量,发现其与住房价格之间存在显著的正向关系。第二,对非市场产品的评估,这一类评估主要涉及存在负外部性的产品,在政策上则与环境经济学相关,具体的包括噪音污染、空气污染、水污染等。学界公认的最早将特征价格法应用于住宅领域实证分析的文献即是关于空气质量的,Ridker and Henning(1967)以美国圣路易斯市为样本进行的空气污染(以硫含量为代理变量)对房价影响的分析;Zabel and Kiel(2000)对美国四个城市的空气污染对房价的影响进行了研究,然后根据估计系数和样本得出如下结论,即经过环境治理达到国家环境空气质量标准(NAAQS)后,四个城市从中获益 1.71 亿至 9.53 亿美元。Espey and Lopez(2000)的研究表明,噪音对住宅价格产生负的影响其门槛值至少要达到 65 分贝;Nelson(2004)研究发现,噪音每增加 1 分贝,房价下降 0.51%至 0.67%。关于水质对房价的影响,Michael et al(2000)针对缅因州的湖水透明度产生的影响进行了测度,该文的目的是并不是想说明最终的测度结果,而是通过变化不同的测度方式得出不同的结果,进而说明政策决定必须对其依据进行小心考证。Bin et al(2008),对北卡罗莱州卡特里特郡的海岸线房价受洪水影响的程度进行了分析;该文的创新之处是它不仅分析了洪水这一环境因素的影响,还将其与保险费率联系在一起,其结论是洪水泛滥区的房价明显低于其他地区,相应的保险费率也高于其他地区,这说明住宅市场和保险市场的均衡是一致的。第三,关于住房价格指数编制。在上述第二方面的截面分析中加入时间虚拟变量即可编制住房价格指数。Ferri(1977)将特征价格指数与传统方法编制的房价进行了对比,结果发现 1965 至 1975 年间传统指数上涨 86.2%,而特征价格指数仅为 59.2%,其原因是后者剔除了由于样本住房楼龄的缩短、住房设备的改善等因素对价格的影响。Case et al(1991)对特征价格指数、加权重复销售指数和混合模型方法的结果进行了对比,其结论是三种方法中重复销售指数表现最差,特征价格法优于前者,混合模型效果最好,因为混合模型包含了组内估计和组间估计两方面的信息。Hill et al(2009)对澳大利亚悉尼市 2001 至 2006 年间出现的楼价暴涨和暴跌进行了测度,结果表明三者测度大体一致,差异不大。

住宅特征价格是一项基础性研究,从国外已有文献来看它取得了广泛的应用价值。然而由于房价及其特征数据的可得性和真实性两方面的原因,国内这方面的实证研究未能取得较大进展。本文的主要贡献表现在以下三个方面:第一,基础数据方面,目前二手房价格相对真实的信息主要来自于中介发布的信息(各地建委数据因避税原因明显低于真实成交价),然而中介信息的最大缺陷是重复发布。我们建立了一套重复数据剔除规则,在剔除了 90%以上的数据后,得到北京地区跨时 3 年、800 多个小区,共计 60 余万条包含建筑特征的数据。这一数据库是保证本项研究客观真实的基础。第二,研究思路方面,我们采取了从微观到宏观的实证分析思路,即在相同小区内计算建筑特征差异,在得到每个小区的所有建筑特征组合对应的价格信息后,再进行邻里和区位特征的计算,这么做的原因包括两点:其一,如果采用包括时间虚拟变量的单一方程回归,任何计量软件都难以承担 60 余万条的工作量;其二,在小区内进行数据量的精简可以防止房价系统性高估的可能,因为存量房地产市场数据主要集中在城区,而小区分布则相对均匀。第三,特征指标的全面性。建筑特征除楼龄、朝向、使用面积未包括,几乎包括了所有特征。其中邻里特征包括了商业、教育、医疗、公共场所四个方面,仅仅未考虑环境污染因素。区位特征涉及到市中心的距离和交通便利性,而交通便利性在公交站点方面还区分了上行和下行。

## 二、不同住宅价格统计方法的比较

上文主要阐述了住宅特征价格的基础性作用和广泛的应用价值,并对相关文献进行了简要回顾,在正式开始对北京的住宅特征价格分析前,我们对不同住宅价格统计方法进行对比,说明住宅特征价格在统计上的优势。

### (一)非同质方法

非同质性方法即忽略住房差异性特征,直接以各报告期交易住房单元成交价格的算术平均值、加权平均值或中位数等指标作为价格统计结果。目前国家统计局公布的“商品房平均价格”指标即采用非同质性方法进行计算。该指标将开发企业填报的新建住房销售金额和销售面积分别在城市、省、全国等各层面进行加总后,将销售金额总额除以销售面积,计算得到平均销售价格,实际上等价于以住房单元面积为权重,计算各报告期内成交住房单元单价的加权平均值。

非同质性方法的最大优势是对基础数据源的要求较低,尤其是不需要成交单价、单元面积之外的其他住房属性信息。因为对基础数据的要求低,它的劣势也是非常明显的。由于该方法假设除单元面积以外的属性是匀质的,而现实中可能出现选样的系统性偏差。比如当前的“平均价格指数”,由于统计的新建房大都集中在郊区,因此放在同质可比条件下相当于区位属性的降低,如果考虑时间维度,这意味着系统性的低估房价上涨幅度(即以当前处于城市郊区的住房价格,与以往城市中心地区的住房价格进行比较)。此外,即使报告期内成交的住房单元属性不存在系统性变化,且样本量足够大,非同质性误差还将影响住房价格统计结果的信度(即加大统计结果中包含的随机误差,主要体现为短期市场结构变化引起的统计价格波动)。

综合上述分析,非同质性方法尽管因其低基础数据要求而具有高可行性的优势,但其固有的同质性误差不仅降低了价格统计结果在短期维度上的可靠性,而且可能在长期维度上存在系统性偏误。

### (二)样本匹配法

针对非同质问题的第一种质量调整思路是采取样本匹配(sample matching)方法,即仅选择那些在不同报告期内没有发生任何质量变化的相同样本,据此进行价格统计,从而利用此类样本固有的同质性特点达到质量控制目的。存量住房价格统计中广为应用的是重复交易法,最为典型的是 Case and Shiller(1989)编制的关于美国住宅市场的重复交易指数<sup>①</sup>。这一方法考虑在各报告期内至少交易过两次的样本数据,进而在假设同一物业在不同报告期内保持各项特征不变的前提下,利用相同物业在两次交易间的价格差异编制住房价格指数。

对于新建商品住宅而言,显然无法应用重复交易方法,因为这是首次交易。但是在样本匹配这一思路下,新建住宅可以在统一开发项目的相同单元内进行匹配,这是由于同一住房项目内的住房单元不仅具有相同的区位条件和邻里环境,而且通常在户型、装修、设施等物理属性方面也具有高度相似性,有助于在住房价格统计中进行质量调整。我国目前的“70 个大中城市房地产价格指数”即是采取这种思路进行样本匹配的。该方法的具体做法是:以新建住房项目作为价格统计中的基本单元,首先利用加权平均公式计算每一在售项目在该报告期的平均价格,再将该平均价格与同一项目前一报告期平均价格进行比较,计算其环比变化率。在逐一计算得到各项目本报告期环比变化率的基础上,进一步利用各项目交易量为权重计算环比变化率的加权平均值,作为该报告期全市新建住房环比变化率的最终结果。

<sup>①</sup> 参见 Case, K. and R. Shiller(1989): “The Efficiency of the Market for Single-family Homes”, *American Economic Review*, 79, 125-137.



该方法尽管能够有效控制交易住房单元在项目层面上的质量差异,从而部分满足同质可比要求。但是,该方法在应用于中国新建商品住宅市场统计时,至少存在以下两方面问题。首先,这种方法仍无法实现对单元层面住房属性差异的控制,甚至放大了单元层面住房属性变化对最终统计结果的影响。多数情况下,特定住房项目在取得预售许可证后,其全部单元将同时投放市场。一个合理的假设是,户型、楼层、朝向等方面属性较受欢迎的住房单元将被首先选购,相反滞销单元(“尾盘”)则通常具有某一方面的不利条件。这将导致同一项目在不同报告期内成交的住房单元存在单元层面属性的持续“下降”,从而导致最终统计结果对住房价格涨幅的系统性低估。其次,也是更为重要的,由于该方法仅仅是在单元层面进行特征同质化控制,但在区位和邻里层面比没有进行样本匹配,这实际上也会出现上述非同质化方法中的系统性低估问题。

在我们的研究中,房价中最为主要的特征是区位特征,鉴于当前新建房郊区化的趋势,系统性低估在所难免。事实上,研究存量房市场的价格特征更具实践指导意义。然而,由于存量房市场广泛存在的阴阳合同问题,基础数据的真实性难以保证,而拥有真实交易数据的二手房中介,出于商业保密性的原因,这些数据又不可得。鉴于重复交易方法在实践中的缺陷和数据方面的困境,我们将住宅价格的研究转向特征价格。

### (三)特征价格法

对住宅价格非同质特征进行调整的另一个思路是借助于特征价格模型(hedonic prices model),因此被统称为特征价格法(hedonic method)。特征价格法的核心要义是,特定住房单元价格由其包含的各种特征以及每种特征的特征价格决定,通过大量数据反映的特征信息计算出房价的基础性价格。式(1)给出了其最常见的形式:

$$\log P_i = c + \sum_{n=1}^N \beta_n X_n + \sum_{t=2}^T \alpha_t D_t + \varepsilon \quad (1)$$

其中, $c$ 为常数项, $X_1 \cdots X_n$ 分别为住房的 $n$ 个特征因素, $\beta_1 \cdots \beta_n$ 分别为对应的特征价格, $D_t$ 为时间哑元变量, $\alpha_t$ 为时间哑元变量系数, $\varepsilon$ 为随机误差项。利用特征价格法进行住房价格统计最常见的一种思路(“一次性建模形式特征价格法”),即将各报告期成交住房单元价格和属性数据一次性引入式(1)所示特征价格模型中,再利用估计得到的各时间哑元系数计算住房价格指数。现实中,也可通过截面数据计算出住宅的基础性价格,再通过同比或环比计算出住房价格指数。

理论上,特征价格法能够在很高程度上满足同质可比要求。但特征价格法的准确应用有赖于高质量基础数据源,尤其是除住房交易价格外,还需要包含对各种住房属性数据的完整反映。本项研究获取的数据很大程度可以满足特征价格的计算要求。一方面,关于建筑特征平均一个季度提供的数据就达到5万条;另一方面,通过地图搜索抓取的大量的小区周边的商业、医疗和教育设施数据很好地满足了邻里特征计算的需要。上述公式看似完美,数据也相当充分,似乎计算特征价格没有问题;然而,样本选取的不当仍然可能导致统计结果的系统性偏误,比如每个季度提取的5万多条数据集中在郊区,或集中在大户型住宅上,这都会导致价格的系统性低估。

## 三、住宅价格特征选取、实证研究思路及方法说明

### (一)住宅价格特征选取

特征价格方法有广泛的应用领域,比如国内外文献已有将特征价格用于计算机和汽车这类零部件较多的商品中,然而将特征价格应用到住宅价格决定中是最为复杂和最有难度的。一方面,从商品属性上看,尽管住宅和汽车、计算机这类商品为人们提供的都是综合效用,但住宅与它们的最大区别是商品的不可拆分性。就计算机而言,其特征价格不仅可以通过特征参数由方程解得,也可通过某个部件在不同等级之间的价格差异直接反映,因此方程的估计的特征价格可以进行可靠的

验证;就住宅而言,由于其不可拆分性,在构建特征价格方程时就应该尽可能包括所有特征参数,否则因遗漏导致的价格偏差难以得到印证。另一方面,住宅与汽车、计算机这类商品的极大区别是不可移动性,因此其特征价格很大程度上取决于附着于住宅的周边环境,而对周边环境的估计带有很强的主观性。如果将小区周边范围确定为其中心 20 公里,那么一个行政区都将共享相同的周边设施,这将造成邻里特征参数没有显著性。如果将小区周边范围确定得过小,这又会导致计算邻里特征的样本过少。因此如何使我们的主管判断接近现实中人们对周边设施的认识,这是一个不断试验的过程。

鉴于上述论述,我们在特征参数的选取上应尽可能全面,而在邻里特征计算方面,我们试验不同的周边环境范围。此外,在特征参数的选取方面还应考虑数据的可得性以及已有文献论述的相关特征的重要性。

住宅特征价格大体上可以分为三类:建筑特征、邻里特征和区位特征。这三个特征在空间上遵循着由小到大的顺序,即建筑特征仅限于住宅单元自身,邻里特征涉及小区周边环境,区位特征反映了小区在整个城市中的空间位置。

我们对住宅价格特征文献进行了梳理分析。就建筑特征而言,Sirmans(2005)的总结发现如下特征是学者考虑的重点(按出现频次由高到低排列):房龄、建筑面积、是否有立体车库、是否有壁炉、占地面积、浴室数目、卧室数目、是否有空调、是否有泳池、是否有地下室等。Sirmans 总结的文献主要反映了西方国家的住宅建筑特征,尤其是反映了美国的建筑特征,比如是否有泳池、是否有壁炉、是否有车库。结合中国国情,我们认为建筑特征应该包括:建筑面积,通常认为面积越大,单价越低,这可由规模经济解释;卧室数目,通常认为卧室数目越多,为家庭成员提供的空间和功能越多,效用越大,从而价格越高;是否装修,通常认为装修的住宅比没装修的贵,因为装修的投入可以产生实实在在的效用,但也有观点认为如果装修老旧或不符合购买者需求,装修将成为负担;楼层,在高层住宅并不普遍的时代,人们偏好楼层的中间层,当高层住宅较为流行时,人们往往认为楼层越高越好;建筑年代,老旧的住宅相对便宜;房间朝向,这一点在北京这座城市应该存在显著特征,通常认为南北通透的住宅透气性和采光性都很好,朝南的住宅比朝北的价格更贵;使用面积,在得房率越来越低的背景下,人们更多地考虑使用面积的大小。表 1 列示特征价格计算中考虑的指标以及最终进入计算的指标。

表 1 住房特征价格指标

| 大类   | 指标        | 是否进入最后计算 | 未进入最终计算的原因          |
|------|-----------|----------|---------------------|
| 建筑特征 | 建筑面积      | 是        |                     |
|      | 卧室数目      | 是        |                     |
|      | 是否装修      | 是        |                     |
|      | 楼层        | 否        | 普遍不显著               |
|      | 建筑年代      | 否        | 样本数据不充足             |
|      | 房间朝向      | 否        | 样本不充足,且虚拟变量过多       |
|      | 使用面积      | 否        | 样本数据不充足             |
| 邻里特征 | 商业环境      | 是        |                     |
|      | 教育条件      | 是        |                     |
|      | 医疗条件      | 是        |                     |
|      | 公共场所      | 是        |                     |
|      | 银行服务便利性   | 否        | 样本统计有遗漏,只有中国银行的网点情况 |
| 区位特征 | 宾馆        | 否        | 存在共线性问题             |
|      | 通勤便利性     | 是        |                     |
|      | 达到城市中心的距离 | 是        |                     |

就邻里特征而言,我们考虑如下因素:商业环境,通常认为商业环境对住宅特征价格的影响是正向的,但 Colwell(1985)的研究表明住宅价格和到商业中心的距离并非单调关系,究其原因,靠近商业中心意味着更好的市政设施、娱乐便利性,但同时也面对噪音、污染等负面效应。教育条件在很大程度上会对住宅价格产生正的影响,Clauretie and Neill(2000)用学生的 SAT(Schloastic Assessment Test)成绩作为学校质量的代理变量,发现其与住房价格之间存在显著的正向关系。在中国基础教育的升学是通过片区行政区划定的,因此在计算中如果能够得到某一小区是否属于重点小学或中学的信息,可以直接确定教育条件对房价的影响,但由于信息的缺失,我们使用小区周边的教育设施数量进行替代,即幼儿园得分为 1,小学得分为 2,中学得分为 2.5。这一指标仅仅反映了数量特征,并没有反映出质量特征。关于周边医疗条件对房价的影响,目前文献中还未曾出现,但我们考虑老年人倾向选择在医疗条件好的地区居住,因此将其作为邻里因素考虑。公共场所,公共场所在国外文献中被考虑最多的是公园,并且公园的类型又被分为都市公园、天然公园和专业公园,在三者中天然公园对住宅价格的影响最大(Lutzenhisser and Netusil,2001)。这一研究成果具有典型发达国家特点,西方国家大多地广人稀,在住宅选择上更讲究环境质量。在本项研究中,我们的公共场所不仅包括公园,还包括体育场馆、旅游景点等,这些公共场所的统计采取数量加总的方法,不考虑其在提供效用上的差异。我们综合认为这些设施对人们的生活质量产生正影响,进而推动房价。宾馆,宾馆的选址实际上是商业便利性和旅游便利性两方面的综合,我们选取该指标作为特征因素,但同时也考虑指标之间的共线性问题。

区位特征被认为是影响住宅价格的最重要的因素,房地产领域的至理名言“Location, Location and Location”就是对这一因素的充分阐释。对区位特征的度量可以从两个方面测度,一方面可以通过住宅到城市中心的距离衡量,另一方面可以用住宅周边的交通便利性衡量,因为作为距离的空间是以达到目的地的时间度量的。So(1997)等人通过衡量住宅到最近的地铁、公共汽车、小型公共汽车站点的距离,对香港地区交通可达性的重要性进行研究,揭示了香港地区对公共交通的依赖性很高,人们为了得到更好的工作场所可达性,愿意为住宅支付更高的价格。我们的研究以小区周边交通的便利程度作为代理变量,公交站点得分为 1,地铁站点得分为 5。

## (二)实证研究思路

特征价格的计算可以采取如上文公式(1)所述直接一次计算的方式,这种方式的样本选取要求在各个特征上分布均匀,但从目前提取的样本情况来看,一次性计算可能产生系统性高估房价的情况。图 1 是小区样本的分布情况,根据我们的统计,四环路以内的样本小区数为 426 个,总样本数为 812 个。这意味着四环内外的小区数目是相当的,但是就小区内部的住宅单元来讲,四环以内小区的样本数目明显多于四环以外的小区,因为作为二手房,核心城区的存量和报价量更多,因此如果所有数据直接一次进入方程,那么方程的计算结果将更多地反映核心城区的信息,这会导致北京的房价系统性高估。对此我们采取如下的方法解决:即首先对每个小区内的样本

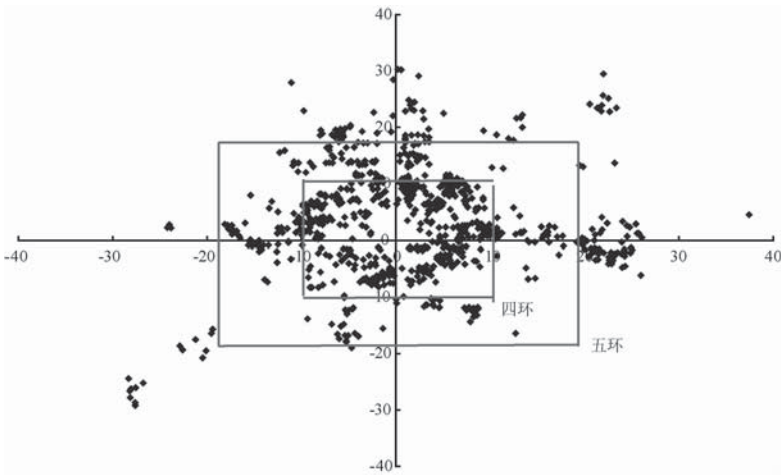


图 1 2010 年小区样本分布情况



进行建筑特征的计算,然后再结合邻里特征和区位特征进行计算,这样可以使核心城区内的样本数有效降低,从而避免房价系统性高估的问题。

为了避免房价系统性高估的问题,我们采取了微观到宏观的计算思路。然而,由于首先要转向微观计算,那么就涉及微观样本的异常处理。如果一次性进入方程计算,数百条或上千条样本的异常对于方程的计算是没有多少影响的,因为我们总体的样本数达到数十万条;但是对于某个小区30至60条样本来讲,若干条样本的异常都会对最终的结果产生影响,因此必须先进行调整。

图2是整个住宅特征价格计算的流程。

首先,进行建筑特征价格的计算。由于基础样本数据是混合在一起的,对单个小区的样本进行计算,就需要把这个小区的样本数据分离出来。在分离出单个小区的样本数据后,我们需要对异常数据进行处理,由于异常数据处理涉及的方面较多,下面单独介绍。在完成上述步骤后,我们进行建筑特征的测试。由于建筑特征包含的因素众多,如果一些特征因素在诸多小区都不显著,则没有计算的必要。比如楼层特征普遍不显著,对于装修特征,原始数据标注的有精装修、中等装修、简易装修和毛坯,但检验表明,诸多小区的住宅单元仅仅在装修或非装修之间存在显著差异,这主要是因为装修程度的标识很大程度上是人的主观因素确定的,标准并不统一,而装修或毛坯却是客观事实。在经过100多个小区的测试后,我们最终确定的建筑特征包括:住宅面积、房间数、是否有装修。

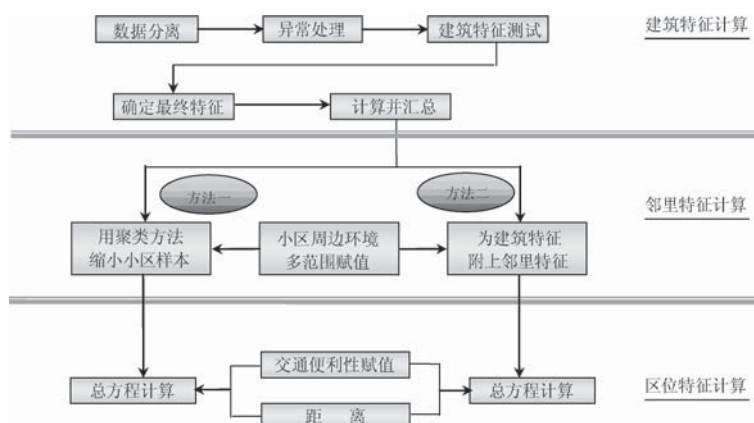


图2 住宅特征价格计算流程图

特征因素在诸多小区都不显著,则没有计算的必要。比如楼层特征普遍不显著,对于装修特征,原始数据标注的有精装修、中等装修、简易装修和毛坯,但检验表明,诸多小区的住宅单元仅仅在装修或非装修之间存在显著差异,这主要是因为装修程度的标识很大程度上是人的主观因素确定的,标准并不统一,而装修或毛坯却是客观事实。在经过100多个小区的测试后,我们最终确定的建筑特征包括:住宅面积、房间数、是否有装修。

其次,我们进行邻里特征的计算。第一步要为小区周边环境进行赋值,我们设定了四种情形:小区周边300米、小区周边500米、小区周边800米、小区周边1000米,通过最后的方程显著性确定各个邻里特征的最终范围。邻里特征的赋值方法具体如下:商业环境,原始数据提供的信息类型包括大型商场、市场、建材超市、古玩市场、珠宝市场,删除一些与生活关系不大的古玩市场、珠宝市场等,只保留大型商场和市场两类,其中大型商场得分为2,市场得分为1;医疗条件,原始数据提供的信息包括医院和诊所两类,医院得分为2,诊所得分为1;教育条件,原始数据提供的信息包括四类,职业教育得分0.5,幼儿园得分1,小学得分2,中学得分2.5;公共场所提供的信息包括公园、景点、场馆、游乐场,这些设施与人们的休闲娱乐相关,不做区分,只统计数量。在为邻里特征赋值完成后,可采用两种方法计算:一是通过聚类方法选取片区中的代表性样本,从而在进入总方程时进一步缩小小区样本,各个小区之间的邻里差异通过回归方程进行计算;二是在每一条建筑特征附上邻里特征,再结合下面的区位特征直接进入总方程计算。方法一的优势是:在进行指数计算时可减少工作量,但前期的基础工作较大;此外从商业运作的角度,它可以为开发商提供区域价值评估。方法二的优势是:它保留了尽可能多的邻里特征信息,从理论探讨的角度这种做法是合适的,本研究采取方法二。

最后,进行区位特征计算并汇总其他特征进行总方程计算。区位特征的计算涉及两个方面:一是到市中心的距离,本项研究中北京的中心确定为天安门;二是交通便利性,交通便利性的赋值方法是,公交站点上行或下行得分0.5,地铁站得分5。地铁站设定数倍公交站得分的原因是,地铁强

大的网络覆盖以及无拥堵带来的准时性。在完成区位特征赋值后,采用方法一或方法二进行总方程的计算。

### (三)有关方法说明

此部分就基础数据的处理方法以及特征计算时的方程形式进行说明。

在基础数据的处理过程中,我们发现存在如下的问题:

第一,价格异常数据。对于价格异常数据的处理,我们认为不应简单采取绝对价格数的标准进行裁定,因为不同小区的价格水平存在很大差异。如果采用较高的绝对数标准,可能导致异常数据不能有效剔除;而采用较低的标准,可能导致一些小区的正常数据被剔除。因此异常数据的剔除应该基于单个小区内的相对价格进行,我们对大量小区的价格水平分析表明,价格分布呈明显的左偏形态,即高价格住宅的数量偏少,因此我们在剔除数据时考虑保留 0.5%分位数和 98%分位数区间内的数据。

第二,户型异常的处理。后文的分析表明,当户型发生跳跃时住宅单价会有显著的下降,因此户型数据是否正确关乎到整体方程的回归效果。在数据处理过程中,我们发现大量的户型错误,即存在大量的户型 A 房间数大于户型 B 房间数,但户型 A 面积却小于户型 B 面积的矛盾现象。我们并不否认上述现象在现实中存在,但我们认为只有在面积非常接近的情况下才会出现上述现象。在处理的程序上,我们采取如下方法修改户型,即将小区住宅按面积进行升序排序,而户型代表的房间数必须呈阶梯状上升。按上述规则修改户型记录,并标记每个小区的修改次数,对于修改次数较多的小区,需要进行关注和复检。即是否存在户型 A 面积很接近户型 B,但户型 A 房间数大于户型 B 房间数的情况,如果出现上述情况,我们视上述情况出现的频次以及是否在不同季度上保持一致,从而确定是否要返回修改。

第三,装修情况的处理。对于没有标记是否装修的记录,我们统一认为是装修的情况。

此外,对于数据是否异常,我们还采取了如下的检测方式。即记录每个季度该小区的均值和标准差,然后进行跨季度的对比。如果两个季度之间的均值和标准差变动很大则要特别关注。首先需要观察两个季度之间是否出现户型突变的情况,在排除了上述情况后即可寻找是否有异常的数据。

关于计算方程的形式说明。

特征价格模型一般包括三种形式:

第一种形式是线性方程式,如下所示:

$$P = \alpha_0 + \beta_1 Z_1 + \beta_2 Z_2 + \cdots + \beta_n Z_n + \varepsilon \quad (2)$$

其中, $Z_1, Z_2, \cdots, Z_n$  分别代表住宅的不同属性, $\beta_1, \beta_2, \cdots, \beta_n$  表示住宅各种属性的影子价格。该方程给出了住宅价格每个属性对住宅价格的贡献值,但是该方程假设每一种属性对价格的贡献量是不变的,而现实中每一种属性的变动存在着边际效应。

第二种形式是对数形式,如下所示:

$$\log P = \log \alpha_0 + \beta_1 \log Z_1 + \beta_2 \log Z_2 + \cdots + \beta_n \log Z_n + \varepsilon \quad (3)$$

其中, $Z_1, Z_2, \cdots, Z_n$  分别代表住宅的不同属性, $\beta_1, \beta_2, \cdots, \beta_n$  不再表示住宅各种属性的影子价格,而是表示每一单位属性增加带来的价格增量,或者说住宅价格对各种属性的弹性。

第三种形式是半对数形式,如下所示:

$$\log P = \log \alpha_0 + \beta_1 Z_1 + \beta_2 Z_2 + \cdots + \beta_n Z_n + \varepsilon \quad (4)$$

其中, $Z_1, Z_2, \cdots, Z_n$  分别代表住宅的不同属性, $\beta_1, \beta_2, \cdots, \beta_n$  表示每一单位属性增加带来的住宅价格对数变化的百分比,将  $\beta_i$  乘以 100 就近似等于住宅价格变化的百分比, $100 \times \beta_i$  又被称为价格  $P$  对  $Z_i$  的半弹性。



对于上述三种形式,具体使用哪一种,可采用 Box-Cox 变换,再通过分析取得最终的函数形式。Box-Cox 变换的原理如下:

$$Z^{(\lambda)} = \begin{cases} (Z^\lambda - 1)/\lambda & \lambda \neq 0 \\ \ln Z & \lambda = 0 \end{cases} \quad (5)$$

当  $\lambda=1$  时,  $Z^{(\lambda)}=Z-1$ ; 当  $\lambda=0$  时,  $Z^{(\lambda)}=\ln Z$ 。

对住宅价格函数的所有变量都进行此种 Box-Cox 变换可得到下列模型:

$$P^{(\lambda_0)} = \beta_0 + \beta_1 S^{(\lambda_1)} + \beta_2 T^{(\lambda_2)} + \beta_3 U^{(\lambda_3)} + \beta_4 V^{(\lambda_4)} + \varepsilon \quad (6)$$

当所有的  $\lambda=0$  时,这个模型就是对数表达式;

当所有的  $\lambda=1$  时,方程转变为:

$$P-1 = \beta_0 + \beta_1(S-1) + \beta_2(T-1) + \beta_3(U-1) + \beta_4(V-1) + \varepsilon, \text{ 该方程与线性表达式是等价的。}$$

当  $\lambda_0=1, \lambda_{1,2,3,4}=0$  时,方程则变成半对数形式。

根据诸多的研究结果,半对数是最为恰当的函数形式。它具有如下优势:第一,半对数模型允许各个特征计量单位的差异性;第二,回归系数很容易理解,系数乘以 100 后表示每一单位特征变化导致的价格变化百分比;第三,半对数模型可以最小化模型的异方差。本文选择半对数模型进行分析。

#### 四、北京市住宅建筑特征分析

上文介绍了住宅特征价格的计算流程,由于我们的计算从微观部分的建筑特征开始,这一节对住宅建筑特征进行分析。

##### (一) 单个小区的建筑特征

根据的大量的试验,小区的建筑特征计算中排除了朝向、建筑年代、使用面积等特征,保留了面积、卧室数目、是否装修三个特征。楼层并不显著,但几乎所有小区计算的楼层系数相对截距项都很小,且在进入总方程计算时,我们都假定是 1 楼,因此不影响结果。

图 3 是农光里小区 2010 年第四季度的住宅价格案例。从图中的走势可以看出,随着住宅面积的增大,房价整体呈下降趋势。我们还可从图中看出房价的下降呈阶梯状,即在某一面积区间范围房价变化幅度不大,这说明房价变动可能存在户型特征。此外,图中的某些突降点可能是装修因素造成的。因此,我们以住宅单价取对数作为被解释变量,面积、楼层、户型虚拟变量、装修虚拟变量作为解释变量,进行回归分析。这里的回归方程一定不存在共线性问题:因为户型和装修都为虚拟变量,而作为连续数据的面积和楼层不可能出现随面积增大楼层相应增大或减小的现象,因此直接进行回归。

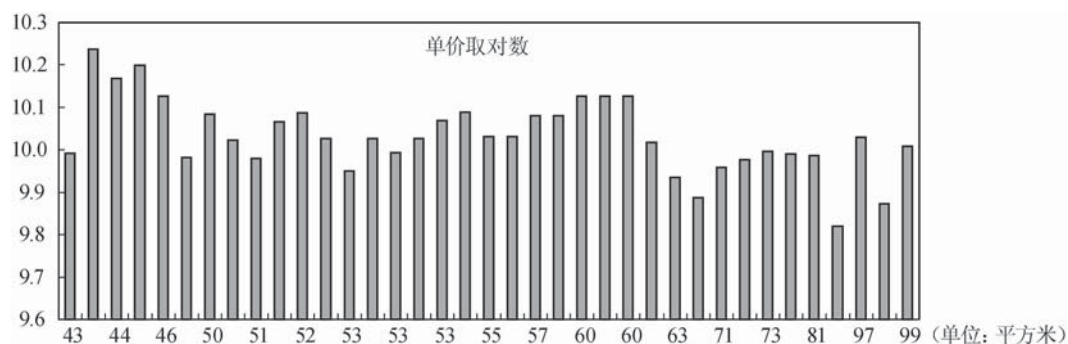


图 3 2010 年第四季度农光里小区单价随面积变化情况

表 2 是 2010 年第四季度农光里小区的建筑特征回归结果,这是有代表性的回归结果。截距项的显著性最为明显,其系数代表了小区单价的整体水平。建筑面积通过 T 检验,其系数为负,表明随着面积的增加单价减少;由于被解释变量取对数,系数意味着变化弹性,表示面积每增加一平方米,单价减少 0.7%,在本案例中即为 229 元,这里体现了住宅价格变化的规模效应。户型虚拟变量系数为正,且两者都通过了 T 检验,表明随着户型增加,住宅为消费者提供的功能增加,效用也增多。建筑面积与户型两者最终对住宅价格的影响取决于综合效应:在这一案例中,住宅由 1 居升为 2 居,住宅单价下降 614 元;住宅由 2 居升为 3 居,住宅单价下降 1018 元。装修特征系数为负,且通过 T 检验,表明毛坯房比装修房便宜,在综合考虑建筑面积、户型因素后,毛坯房比装修房便宜 4848 元。这一价格的便宜程度是很大的,存在着套利空间。因为对于 43 平方米的一居户型而言,毛坯房相对于装修房便宜 21 万,远远可以覆盖装修成本。在所有小区的计算中,这样的案例是很少的,根据我们的统计,毛坯房相对于装修房单价大约便宜 5%,考虑房价和装修成本后,存在装修套利空间的很少。楼层系数并没有通过 T 检验,但其系数很小,比第二小的建筑面积的系数还小 10 倍,因此在其乘以 1 后并不会对结果产生实质性影响。

表 2 2010 年第四季度农光里小区建筑特征回归结果

| 变量             | 系数         | 标准差    | T 统计量   | P 值    |        |
|----------------|------------|--------|---------|--------|--------|
| 截距项            | 10.4039*** | 0.1072 | 97.0205 | 0.0000 |        |
| 建筑面积           | -0.0070*** | 0.0022 | -3.2215 | 0.0030 |        |
| 楼层             | 0.0007     | 0.0020 | 0.3516  | 0.7276 |        |
| 是否 2 居以上       | 0.0721     | 0.0430 | 1.6753  | 0.1039 |        |
| 是否 3 居以上       | 0.1653**   | 0.0811 | 2.0379  | 0.0502 |        |
| 是否装修           | -0.2286*** | 0.0836 | -2.7361 | 0.0102 |        |
| R <sup>2</sup> | 0.4548     | F 统计量  | 5.1727  | F 检验   | 0.0015 |

然而,并不是所有小区的回归结果都很好,表 3 即是一个不好的案例,建筑面积勉强通过检验;楼层并不是关注的变量,却通过检验;户型虚拟变量只有 4 居通过检验,其他都未通过;装修尽管也是勉强通过检验,但系数为正,说明毛坯房比装修房贵。此外,方程的拟合优度也不高。

表 3 2010 年第三季度桃花岛小区建筑特征回归结果

| 变量             | 系数        | 标准差    | T 统计量   | P 值    |        |
|----------------|-----------|--------|---------|--------|--------|
| 截距项            | 9.9157*** | 0.1231 | 80.5255 | 0.0000 |        |
| 建筑面积           | -0.0024   | 0.0015 | -1.6100 | 0.1140 |        |
| 楼层             | -0.0185** | 0.0094 | -1.9598 | 0.0558 |        |
| 是否 2 居以上       | 0.0295    | 0.0691 | 0.4265  | 0.6716 |        |
| 是否 3 居以上       | -0.0023   | 0.0585 | -0.0392 | 0.9689 |        |
| 是否 4 居以上       | 0.2970**  | 0.1582 | 1.8776  | 0.0665 |        |
| 是否装修           | 0.1776    | 0.1119 | 1.5869  | 0.1191 |        |
| R <sup>2</sup> | 0.2511    | F 统计量  | 2.6828  | F 检验   | 0.0251 |

在我们对各个小区的建筑特征进行计算的过程中,通过错误户型的修改以及异常数据的剔除,尽管可以对方程的拟合效果有所提高,但并不是所有小区的建筑特征都拟合的很好。然而,在计算过程中,我们发现一个现象,即所有小区的截距项都是通过检验的。这就意味着每个小区都存

在一个基础价格,而这个价格并不是小区内的特征决定的,恰恰是建筑特征以外的因素决定的。

### (二)户型特征在不同区域的表现

上文的分析表明建筑特征对住宅价格的决定因素仅仅是很小一部分,那么这一节我们通过统计方法对不同建筑特征在不同区域的表现差异进行考察,以便为小区基础价格的决定因素发现一些线索。

上文论述表明,户型提供的功能效用和建筑面积的规模效应交织在一起,但总的表现是随着户型增加单价略微下降。本节考察户型特征在不同区域的表现,我们在相同小区、相同装修特征的前提下,将户型按升序排列,观察大户型相对于小户型是否单价下降,以及下降幅度。

表 4 是按上述方法进行排序后的统计结果。从表中我们可以发现一个有意思的现象:当我们由全市范围向核心城区缩小范围时,随户型增加单价下降的情况相较于随户型增加单价上升的情况以更快的速度递减,这意味着越往主城区靠近,户型增大导致规模效应的可能性越小;价格变动的幅度更能说明问题(第 5、6 两列),越往主城区靠近,随户型增加导致单价下降的幅度越低,相反单价上升的幅度越高,这说明往主城区不仅规模效应的可能性变小,规模效应的程度也变小。这一现象背后反映的是:在城区级差地租逐步抬高的情况下,面积的规模效应逐渐消失。实际上,在建筑特征的多维考察中,室内面积相对于区域来讲是不重要的一个维度。

### (三)装修特征在不同区域的表现

表 4 2010 年第四季度户型在不同区域的表现

| 变量   | 随户型增加单价下降的个数 | 随户型增加单价上升的个数 | 两者之差 | 随户型增加单价下降的幅度(%) | 随户型增加单价上升的幅度(%) | 两者之差(%) |
|------|--------------|--------------|------|-----------------|-----------------|---------|
| 全区域  | 1413         | 808          | 605  | 7.78            | 7.84            | -0.07   |
| 五环以内 | 969          | 637          | 332  | 7.05            | 7.56            | -0.51   |
| 四环以内 | 544          | 413          | 131  | 6.61            | 7.48            | -0.87   |
| 三环以内 | 170          | 160          | 10   | 6.94            | 8.67            | -1.73   |

上文统计了户型在不同区域的表现差异,这里统计装修在不同区域的特征差异。同样,我们固定相同的小区、相同的户型条件下,比较装修房和毛坯房的价格关系。

表 5 是统计结果,整体来看出现毛坯房比装修房贵的情况很少,不到总数的 2%。从区域特点来看,出现毛坯房比装修房贵的情况越往郊区可能性越大。这与郊区别墅较多的情况有关,因为购买别墅的人是价格不敏感的,它们对住房的个性需求是重要的,而这时一座已有装修的房子如果与购房者的偏好不同,对于购房者来说就是个负担。此外,装修房比毛坯房贵的幅度从另一个方面说明了问题,越往核心区靠近,装修房相对毛坯房贵的程度在下降,这也说明装修这种建筑特征相对于级差地租而言是不重要的,反方向的变动也印证了这一事实,越往郊区,毛坯房相对装修房越贵。

表 5 2010 年第三季度装修在不同区域的表现

| 变量   | 装修房比毛坯房贵的个数 | 装修房比毛坯房便宜的个数 | 两者之差 | 装修房比毛坯房贵的幅度(%) | 装修房比毛坯房便宜的幅度(%) | 两者之差(%) |
|------|-------------|--------------|------|----------------|-----------------|---------|
| 全区域  | 1046        | 23           | 1023 | 5.72           | 10.88           | -5.16   |
| 五环以内 | 648         | 14           | 634  | 5.73           | 9.41            | -3.68   |
| 四环以内 | 244         | 4            | 240  | 5.49           | 8.76            | -3.28   |
| 三环以内 | 78          | 0            | 78   | 6.25           | 0.00            | 6.25    |



## 五、北京市住宅建筑特征及区位特征的综合分析

上文的分析已经勾勒出区位重要性的线索,这一节我们先进行建筑特征和区位特征的综合分析,由于我们的特征价格计算采取的是方法二(即混合所有特征计算),且区位特征相对重要。先进行区位特征的计算并不影响最终结果。

### (一)区位特征的整体回归

图 4 是 2010 年第四季度北京市住宅单价取对数后,其随距市中心距离的变化情况。从图中可以看出,住宅单价随距离增加呈逐渐下降趋势,但其下降趋势并不是线性比例的下降,我们给出的拟合线说明住宅单价是加速下降的。这提示我们方程的形式应该包括二次形式。图中的价格下降趋势中呈现锯齿形态,我们设想这是由三个方面的原因造成的:一是户型之间的差异,二是是否装修导致的差异,三是交通便利性导致的。于是我们构建如下的方程形式:

$$\ln P = c + b_1 * D^2 + b_2 * D + b_3 * TC + b_4 * S + \sum b_j * Bedroom_j + b_k * If\_Decor \quad (7)$$

其中,  $D$  表示距离,  $TC$  表示交通便利性,  $S$  表示建筑面积,  $Bedroom_j$  表示户型虚拟变量,  $If\_Decor$  表示是否装修。

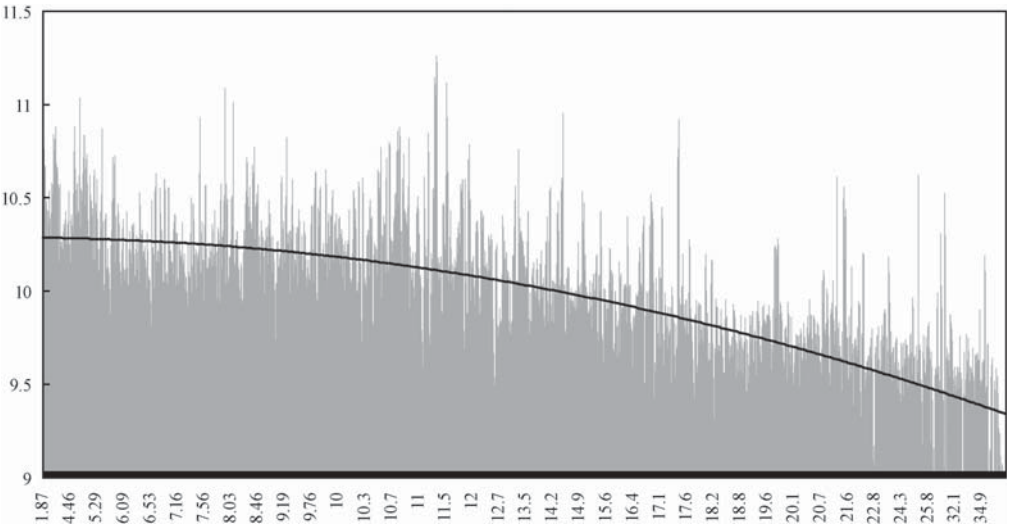


图 4 2010 年第四季度住宅单价在地理空间上的差异

表 6 是包含建筑特征和区位特征的按上述方程回归的结果,由于对于交通便利性我们并不确定多大的赋值范围才是合适的,我们分别计算了四种情形下的模型分析结果,模型一至模型四分别表示小区 300 米范围的交通便利性、小区 500 米范围的交通便利性、小区 800 米范围的交通便利性和小区 1 公里范围的交通便利性。表 6 的检验结果表明,当赋值范围达到小区中心 800 米范围以上时,交通便利性通过 T 检验;而且随着范围扩大,方程的整体拟合优度也在提升。较小的半径范围难以通过检验的原因是,范围过小不能充分的包含交通便利性信息,以小区中心 300 米半径为例,很多大型的小区自身都覆盖了这个面积,因此交通便利性赋值大都为 0,难以起到区分作用。模型三和模型四都是拟合效果很好的方程,考虑到过大的范围也会出现交通便利性共享的问题,我们选用模型三。

模型三的整体拟合效果很好,并且所有系数通过 T 检验。系数所反映的情况是:截距项的系数最大,它体现了北京的住宅单价的整体情况。距离的系数为负,表明随距离的增加,住宅单价是单调递减的;距离的平方系数为正,说明这种递减是加速递减的。交通便利性的系数为正,说明交通

越发达,其对住宅单价有显著的提升作用。建筑面积的系数为正,表示其随单位面积的增大价格上升;这与上文的建筑特征的案例分析结果不同:尽管住宅单价上升,但户型的梯度增加引致单价下降,且户型系数远大于面积,考虑现实户型面积后,两者综合效应是递减的;这说明就北京全市而言,户型越大,房价相对便宜,但相同户型条件下,面积增加引致效用增加,从而住宅价格增加。装修系数为负,说明毛坯房比装修房便宜。

表 6 包含建筑特征和区位特征的回归结果

|                     | 模型一                   | 模型二                   | 模型三                   | 模型四                   |
|---------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| 截距项                 | 10.4918*** (650.8188) | 10.4894*** (628.4954) | 10.4687*** (601.6705) | 10.4621*** (581.9495) |
| 距离平方                | 0.0003*** (7.6499)    | 0.0003*** (7.6562)    | 0.0003*** (7.6791)    | 0.0003*** (7.5647)    |
| 距离                  | -0.0459*** (-29.6215) | -0.0459*** (-29.6118) | -0.0457*** (-29.4732) | -0.0455*** (-29.2337) |
| 交通便利性               | -0.0003 (-0.6866)     | 0.0000 (-0.0742)      | 0.0003*** (2.5782)    | 0.0003*** (2.9679)    |
| 建筑面积                | 0.0019*** (21.2352)   | 0.0019*** (21.2068)   | 0.0020*** (21.4310)   | 0.0020*** (21.4693)   |
| 是否 2 居以上            | -0.0660*** (-6.3215)  | -0.0660*** (-6.3206)  | -0.0664*** (-6.3624)  | -0.0665*** (-6.3698)  |
| 是否 3 居以上            | -0.1040*** (-9.8463)  | -0.1041*** (-9.8503)  | -0.1048*** (-9.9219)  | -0.1049*** (-9.9336)  |
| 是否 4 居以上            | -0.1437*** (-10.3661) | -0.1436*** (-10.3509) | -0.1451*** (-10.4649) | -0.1451*** (-10.4682) |
| 是否装修                | -0.0648*** (-7.7058)  | -0.0645*** (-7.6714)  | -0.0630*** (-7.4963)  | -0.0627*** (-7.4601)  |
| R <sup>2</sup>      | 0.6171                | 0.6171                | 0.6177                | 0.6179                |
| 调整后的 R <sup>2</sup> | 0.6164                | 0.6163                | 0.6170                | 0.6172                |
| F 统计量               | 830.0789              | 829.9268              | 832.0949              | 832.8003              |
| F 值的 p 检验值          | 0.0000                | 0.0000                | 0.0000                | 0.0000                |

(二) 区位特征的南北差异

对于北京的住宅价格,民众普遍的说法是“上风上水上海淀”,其表明的含义是北京居民偏好于在北部购房,那么相应地北部区域的房价相对更高,下降的速度也应该更慢。图 5 绘制了 2010 年第四季度北京住宅单价在南北分布之间的差异,以黑色的竖线为中心,右手表示北部区域,左手表示南部区域。图中也简单地拟合了住宅单价随距离变动的二次曲线,从拟合的结果来看,北部的房价随距离下跌的速度更快,似乎并不能得出北部房价更高的结论。

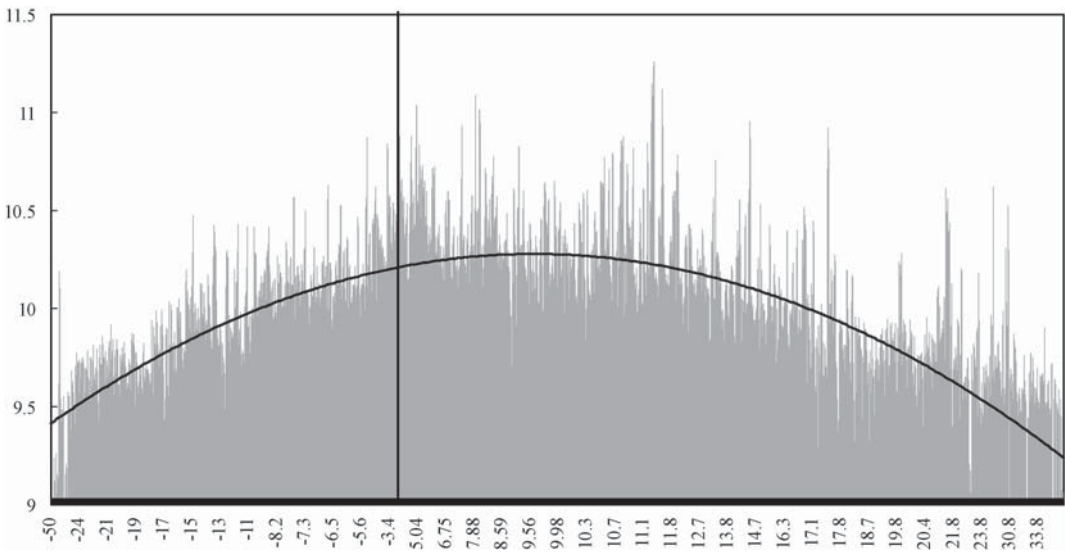


图 5 2010 年第四季度北京住宅单价在南北区域分布的差异

为了更深入地说明问题,我们需要通过回归分析进一步地观察。表 7 是应用模型三分别对南北两个区域进行估计的结果:从截距项来看,北部高于南部,说明北部的房价整体高于南部区域;距离系数表现出有意思的现象,距离本身的系数北部的绝对值大于南部,说明北部房价下降的幅度大,距离平方也同样说明北部的房价越往城外走下降的速度越快。结合截距项的结果可以说明如下问题:所谓的北部区位优势好,实际上仅仅限于核心城区部分,而到郊区,北部的房价下降的更快;对于南部而言,城区的房价相对较低,而郊区的下降幅度也不大,这与当前北京南部的开发力度加大有关。就交通便利性特征而言,系数反映的结果是南部地区的交通便利性对房价有正向作用,北部地区的交通便利性对房价却是负向作用,这可能是北部的郊区交通已经很发达,但房价还未达到相应的估计水平造成的。综合建筑面积和户型虚拟变量来看,在北部区域买更大的住宅相对于南部会更便宜。装修特征的南北部差异是,北部区域的毛坯房相对于装修房更便宜。

表 7 区位特征在南北之间的差异

|                     | 南部                    | 北部                    |
|---------------------|-----------------------|-----------------------|
| 截距项                 | 10.3367*** (506.4673) | 10.6975*** (431.3466) |
| 距离平方                | 0.0002*** (3.9947)    | 0.0006*** (11.0896)   |
| 距离                  | -0.0405*** (-23.5473) | -0.0605*** (-27.0540) |
| 交通便利性               | 0.0001 (0.7456)       | -0.0001 (-0.5853)     |
| 建筑面积                | 0.0018*** (11.3952)   | 0.0018*** (16.6594)   |
| 是否 2 居以上            | -0.0600*** (-4.3912)  | -0.0631*** (-4.8040)  |
| 是否 3 居以上            | -0.0969*** (-6.7953)  | -0.0968*** (-7.3737)  |
| 是否 4 居以上            | -0.1314*** (-6.6314)  | -0.1337*** (-7.9694)  |
| 是否装修                | -0.0480*** (-4.4290)  | -0.0558*** (-5.2497)  |
| R <sup>2</sup>      | 0.7149                | 0.6387                |
| 调整后的 R <sup>2</sup> | 0.7133                | 0.6376                |
| F 统计量               | 441.6975              | 597.0136              |
| F 值的 p 检验值          | 0.0000                | 0.0000                |

### (三)不同户型组在区位特征上的差异

第四节中我们分析了户型特征在不同区域的表现,结果表明越往中心城区,户型增加导致单价下降的可能性越小,甚至可能还是上升的,我们将其解释为中心城区的地租是房价决定中的更重要因素。在这一部分我们进行同样目的的检验,但方法略有不同:我们固定住某一户型,看其在距离增大的过程中价格下降的程度,对于较大的户型而言,其下降的幅度和速度都应该更大,因为越往中心,扩大面积的地租代价越大,或者换句话讲,越往郊区,大户型才有发展的空间。

图 6 是不同户型组在距离特征上的差异,根据前文的结论以及这一部分的假设,4 居以上相对于 3 居以及 2 居和 1 居应该有更大的负斜率,并且有更大的凸度。上述四个子图基本符合这一趋势。

表 8 是四个方程的拟合结果:距离平方即为凸度表示,确实是随着户型增加逐渐增加;距离的系数是负数,但其绝对值是随着户型增加逐渐增加的,这说明下降的幅度是增加的。交通便利性的系数随户型增大逐渐减少的,甚至当达到一定户型后其系数为负,这背后反映了一个有意思的现象:如果各种户型对应于不同的购买者,单身族或上班族更看重通勤便利性,因此交通便利性对房价产生正向影响;对于两口或三口之家而言,在该方程中他们对交通便利性并不敏感;对于那些有能力购买大户型住宅的消费者而言,交通便利性还会产生负面影响,因为越便利的交通意味着更多的污染、噪音等,这类奢侈住宅大多建在较为幽静的地方。建筑面积的系数在各种户型下都为正

数,这意味着住宅提供的功能整体不变的条件下,增加面积是增加效用的,因此需要支付更高的价格;只有当户型发生改变时,单价才随面积增加而下降。装修系数都为负,且随着户型增加其相对于装修房的便宜程度逐渐降低。

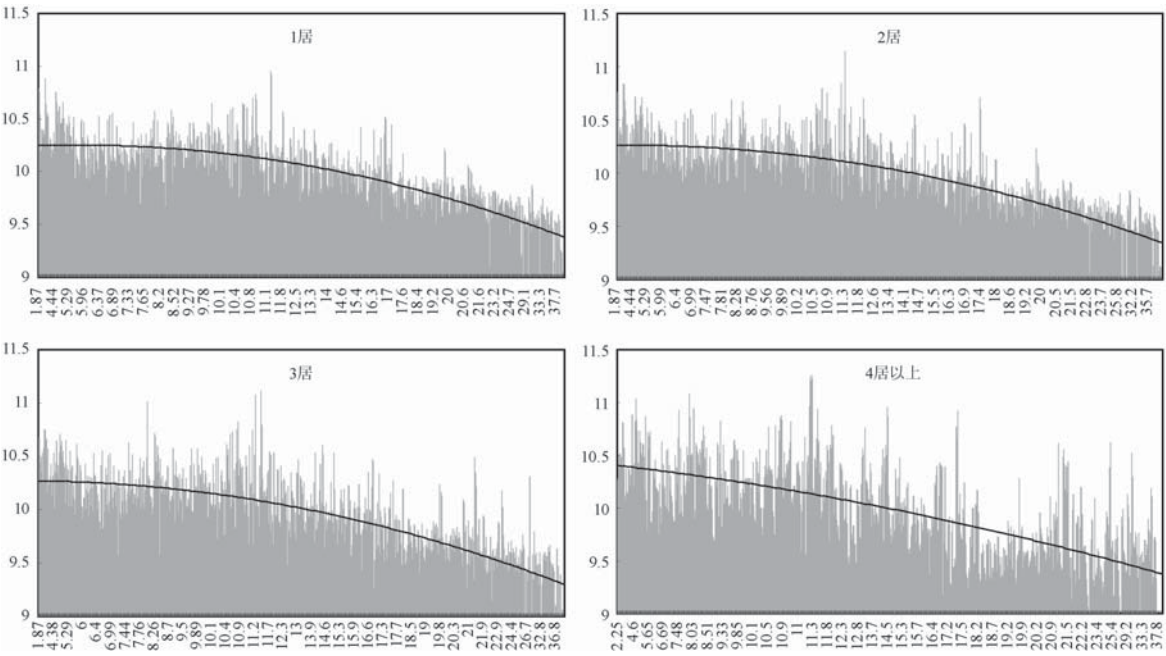


图 6 不同户型组在区位特征上的差异

表 8 包含建筑特征和区位特征的回归结果

|                     | 1 居                   | 2 居                   | 3 居                   | 4 居以上                 |
|---------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| 截距项                 | 10.3984*** (316.4124) | 10.3233*** (290.2520) | 10.2742*** (244.8237) | 10.3371*** (194.6951) |
| 距离平方                | 0.0001** (1.9139)     | 0.0002*** (3.8998)    | 0.0003*** (4.6599)    | 0.0005*** (4.7769)    |
| 距离                  | -0.0372*** (-13.9784) | -0.0435*** (-16.9619) | -0.0484*** (-16.4922) | -0.0558*** (-13.3692) |
| 交通便利性               | 0.0007* (1.6679)      | 0.0001 (0.3871)       | 0.0002 (0.4145)       | -0.0011* (-1.6037)    |
| 建筑面积                | 0.0019*** (3.8847)    | 0.0029*** (8.6007)    | 0.0025*** (9.8620)    | 0.0018*** (13.0370)   |
| 是否装修                | -0.0673*** (-4.6058)  | -0.0665*** (-4.7692)  | -0.0630*** (-3.9714)  | -0.0619*** (-2.8259)  |
| R <sup>2</sup>      | 0.6670                | 0.6737                | 0.6429                | 0.5383                |
| 调整后的 R <sup>2</sup> | 0.6653                | 0.6722                | 0.6413                | 0.5359                |
| F 统计量               | 384.9657              | 457.0962              | 394.6791              | 219.4369              |
| F 值的 p 检验值          | 0.0000                | 0.0000                | 0.0000                | 0.0000                |

六、北京市住宅建筑特征、邻里特征及区位特征的综合分析

上文对北京市住宅的建筑特征和区位特征进行了分析,这一节我们将考虑对住宅价格影响仅次于区位因素的小区周边环境,即邻里特征。这些邻里特征包括:商业环境、医疗条件、教育条件和公共场馆情况。这些因素与上文区位特征中的交通便利性很可能反映出共同的信息,因为便利的交通除了是为了快速的达到工作场所,也同样会指向商场、医院、学校、公园、体育场馆等,因此这里的分析尤其要注意共线性问题。



(一)加入商业环境的住宅特征价格

在上述第五节的方程中加入商业环境因素,从而分析这种邻里因素对住宅价格的影响。由于我们并不确定小区多大范围内的商业环境会对住宅价格产生影响以及其影响方向,我们设计了300米范围、500米范围、800米范围和1公里范围四个方程进行分析。本节开头指出潜在的多重共线性问题,我们首先进行共线性检验。表9是四个方程的方差膨胀因子计算结果,通常认为方差膨胀因子整体大于10存在显著的共线性,表9的结果表明各个方程的方差膨胀因子平均在4左右,不存在共线性问题。

表9 包含商业环境方程的方差膨胀因子

| 方差膨胀因子   | 小区 300m 范围 | 小区 500m 范围 | 小区 800m 范围 | 小区 1km 范围 |
|----------|------------|------------|------------|-----------|
| 距离平方     | 11.8303    | 11.9223    | 11.9867    | 12.1742   |
| 距离       | 11.9007    | 12.0306    | 12.1831    | 12.3963   |
| 交通便利性    | 1.0516     | 1.0610     | 1.0836     | 1.0773    |
| 商业环境     | 1.0159     | 1.0493     | 1.0941     | 1.1015    |
| 建筑面积     | 3.6437     | 3.6607     | 3.6452     | 3.6402    |
| 是否 2 居以上 | 1.5176     | 1.5179     | 1.5176     | 1.5174    |
| 是否 3 居以上 | 2.1642     | 2.1650     | 2.1640     | 2.1639    |
| 是否 4 居以上 | 2.6418     | 2.6416     | 2.6392     | 2.6382    |
| 是否装修     | 1.0220     | 1.0219     | 1.0218     | 1.0218    |
| 均值       | 4.0875     | 4.1189     | 4.1484     | 4.1923    |

表10给出了包括商业环境的住宅特征价格回归结果,模型一至模型四分别表示小区300m范围至1公里范围的商业环境得分情况。从回归结果看,在加入商业环境后方程的整体回归结果与第五节的回归结果相比没有大的变化,截距项有略微波动,距离系数有略微差异,距离平方系数没有变化,其他系数变化也不大。这说明商业环境的加入对北京住宅特征价格的整体影响不大。但一个有意思的现象是,如果商业环境距离小区越近,其对小区住宅价格的影响是负向的,这与我们通常想象的周边商业越发达住宅价格越高的观念不相一致。随着周边范围的扩大,商业环境对住宅价格的系数转为正,但这些方程都没有通过T检验,这说明房价因素决定中,商业环境并不是一个主要因素,但距离周边商业过近的小区,房价会有负面影响。

(二)加入教育环境的住宅特征价格

教育环境通常是居民购房选择过程中重点考虑的因素,当前学区房价格普遍高于周边住房价格即是明证。遵照上一部分的方法,我们单独加入教育环境进行分析。首先依然是进行共线性检验,从表11给出的计算结果来看,加入教育环境后的方程不存在多重共线性问题。

表12给出了不同小区半径范围的加入教育环境的住宅特征价格回归结果,同样地,与第五节回归结果相比,方程整体的改变不大。回归结果显示,在小区周边800米范围内,教育对住宅价格的影响是不显著的,但是在小区周边1公里的范围内,教育环境对住宅价格有正向的影响。从系数来看,这种影响效果并不强,甚至小于建筑面积递增产生的正的影响。这一结果与现实感受存在一定的差距,很大原因是由于,我们只标志了教育设置的物理距离,而现实中哪些小区能够与重点学校相连很大程度上是行政直接划分的结果。如果我们能够明确得到小区是否是学区房的明确信息,相信这一回归结果的效果会更明显。

(三)加入医疗条件的住宅特征价格

医疗条件是购房过程中考虑的另一重要邻里因素,先验的想法是:老年人很愿意在靠近医院

表 10 包含商业环境的住宅特征价格回归结果

|                     | 模型一                   | 模型二                   | 模型三                   | 模型四                   |
|---------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| 截距项                 | 10.4737*** (593.7655) | 10.4668*** (579.9190) | 10.4725*** (565.9543) | 10.4732*** (554.0179) |
| 距离平方                | 0.0003*** (7.8022)    | 0.0003*** (7.5675)    | 0.0003*** (7.6847)    | 0.0003*** (7.6535)    |
| 距离                  | -0.0459*** (-29.5077) | -0.0456*** (-29.1176) | -0.0459*** (-29.1003) | -0.0459*** (-28.8755) |
| 交通便利性               | 0.0003*** (2.6684)    | 0.0003** (2.5184)     | 0.0003*** (2.6440)    | 0.0003*** (2.6419)    |
| 商业环境                | -0.0100*** (-1.7070)  | 0.0010 (0.4067)       | 0.0000 (-0.6005)      | 0.0000 (-0.6028)      |
| 建筑面积                | 0.0020*** (21.3344)   | 0.0020*** (21.3895)   | 0.0020*** (21.3716)   | 0.0020*** (21.3959)   |
| 是否 2 居以上            | -0.0662*** (-6.3415)  | -0.0665*** (-6.3685)  | -0.0663*** (-6.3524)  | -0.0664*** (-6.3567)  |
| 是否 3 居以上            | -0.1045*** (-9.8984)  | -0.1049*** (-9.9278)  | -0.1048*** (-9.9133)  | -0.1047*** (-9.9157)  |
| 是否 4 居以上            | -0.1442*** (-10.3978) | -0.1453*** (-10.4717) | -0.1447*** (-10.4502) | -0.1451*** (-10.4626) |
| 是否装修                | -0.0628*** (-7.4704)  | -0.0630*** (-7.4908)  | -0.0630*** (-7.4903)  | -0.0630*** (-7.4932)  |
| R <sup>2</sup>      | 0.6180                | 0.6177                | 0.6177                | 0.6177                |
| 调整后的 R <sup>2</sup> | 0.6171                | 0.6169                | 0.6169                | 0.6169                |
| F 统计量               | 740.3073              | 739.5084              | 739.5652              | 739.5660              |
| F 值的 p 检验值          | 0.0000                | 0.0000                | 0.0000                | 0.0000                |

表 11 包含教育环境方程的方差膨胀因子

| 方差膨胀因子   | 小区 300m 范围 | 小区 500m 范围 | 小区 800m 范围 | 小区 1km 范围 |
|----------|------------|------------|------------|-----------|
| 距离平方     | 12.1644    | 12.5936    | 13.0344    | 13.3614   |
| 距离       | 12.4615    | 13.1471    | 13.9478    | 14.6225   |
| 交通便利性    | 1.0685     | 1.0768     | 1.0967     | 1.0916    |
| 教育环境     | 1.1362     | 1.2235     | 1.3569     | 1.4479    |
| 建筑面积     | 3.6980     | 3.6797     | 3.6717     | 3.6514    |
| 是否 2 居以上 | 1.5182     | 1.5182     | 1.5179     | 1.5175    |
| 是否 3 居以上 | 2.1677     | 2.1673     | 2.1667     | 2.1653    |
| 是否 4 居以上 | 2.6424     | 2.6410     | 2.6409     | 2.6384    |
| 是否装修     | 1.0218     | 1.0222     | 1.0220     | 1.0218    |
| 均值       | 4.2087     | 4.3410     | 4.4950     | 4.6131    |

表 12 包含教育环境的住宅特征价格回归结果

|                     | 模型一                   | 模型二                   | 模型三                   | 模型四                   |
|---------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| 截距项                 | 10.4743*** (567.8639) | 10.4744*** (541.4644) | 10.4709*** (514.8214) | 10.4490*** (496.7333) |
| 距离平方                | 0.0003*** (7.7121)    | 0.0003*** (7.5874)    | 0.0003*** (7.3519)    | 0.0003*** (6.6167)    |
| 距离                  | -0.0460*** (-28.8781) | -0.0460*** (-28.1248) | -0.0458*** (-27.1777) | -0.0444*** (-25.7374) |
| 交通便利性               | 0.0003*** (2.6777)    | 0.0003** (2.6525)     | 0.0003*** (2.5645)    | 0.0003** (2.1957)     |
| 教育环境                | -0.0020 (-0.9036)     | 0.0000 (-0.6697)      | 0.0000 (-0.2092)      | 0.0010* (1.6688)      |
| 建筑面积                | 0.0019*** (21.1329)   | 0.0020*** (21.2285)   | 0.0020*** (21.3031)   | 0.0020*** (21.4988)   |
| 是否 2 居以上            | -0.0662*** (-6.3384)  | -0.0662*** (-6.3442)  | -0.0664*** (-6.3565)  | -0.0666*** (-6.3815)  |
| 是否 3 居以上            | -0.1044*** (-9.8741)  | -0.1045*** (-9.8857)  | -0.1047*** (-9.9062)  | -0.1052*** (-9.9653)  |
| 是否 4 居以上            | -0.1446*** (-10.4203) | -0.1448*** (-10.4370) | -0.1450*** (-10.4518) | -0.1453*** (-10.4826) |
| 是否装修                | -0.0631*** (-7.5015)  | -0.0632*** (-7.5089)  | -0.0631*** (-7.4978)  | -0.0629*** (-7.4842)  |
| R <sup>2</sup>      | 0.6178                | 0.6177                | 0.6177                | 0.6180                |
| 调整后的 R <sup>2</sup> | 0.6169                | 0.6169                | 0.6169                | 0.6171                |
| F 统计量               | 739.6977              | 739.5907              | 739.4731              | 740.2698              |
| F 值的 p 检验值          | 0.0000                | 0.0000                | 0.0000                | 0.0000                |

的周边居住,因为他们有更多的健康需求,并且他们相对于年轻人在更早的时机有更强的购买力。因此我们推断医疗条件对住宅价格有正向影响。下面通过方程进行检验,首先依然是考虑共线性问题,表 13 给出的方差膨胀因子表明方程不存在共线性关系。

表 13 包含医疗条件方程的方差膨胀因子

| 方差膨胀因子   | 小区 300m 范围 | 小区 500m 范围 | 小区 800m 范围 | 小区 1km 范围 |
|----------|------------|------------|------------|-----------|
| 距离平方     | 11.7626    | 11.7463    | 11.7979    | 11.9395   |
| 距离       | 11.8197    | 11.8215    | 11.9327    | 12.1346   |
| 交通便利性    | 1.0526     | 1.0486     | 1.0552     | 1.0531    |
| 医疗条件     | 1.0191     | 1.0132     | 1.0369     | 1.0509    |
| 建筑面积     | 3.6680     | 3.6512     | 3.6394     | 3.6364    |
| 是否 2 居以上 | 1.5176     | 1.5179     | 1.5173     | 1.5174    |
| 是否 3 居以上 | 2.1658     | 2.1647     | 2.1639     | 2.1638    |
| 是否 4 居以上 | 2.6425     | 2.6400     | 2.6394     | 2.6383    |
| 是否装修     | 1.0223     | 1.0219     | 1.0218     | 1.0217    |
| 均值       | 4.0745     | 4.0695     | 4.0894     | 4.1284    |

表 14 是包含医疗条件的住宅特征价格回归结果。与不加入医疗条件的回归结果相比,方程的系数和各变量的显著性检验差异不大,同样说明了邻里特征并不是价格决定的主要因素。但是对不同小区半径的回归结果却与我们的先验观念不一致:医院或诊所距离小区越近,其对小区住宅价格的影响是负向的,且程度越大;当医院或诊所距离小区 1 公里范围的情况下,其对住宅价格的影响是正向的,但却不显著。对此我们有两点解释:第一,距离医院越近,存在医疗污染的风险越大,这是负面的效应。第二,对于有大量医疗服务需求的老年人来说,他们的另一个属性是有大量的时间,空间上的距离是通过时间上的可达性丈量的,在时间相对充裕的情况下,距离的远近也就并非重要;此外即使存在急救的需求,这一方面是一个概率事件,花费很大的住宅价格成本去换取急救时间可能是不值得的,另一方面在发生急救后,治疗就不存在距离上的远近问题,因为这时已经转为住院服务了。

表 14 包含医疗条件的住宅特征价格回归结果

|                     | 模型一                   | 模型二                   | 模型三                   | 模型四                   |
|---------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| 截距项                 | 10.4764*** (602.8980) | 10.4853*** (589.3814) | 10.4780*** (570.3098) | 10.4665*** (548.7223) |
| 距离平方                | 0.0003*** (7.4245)    | 0.0003*** (7.8147)    | 0.0003*** (7.7733)    | 0.0003*** (7.5752)    |
| 距离                  | -0.0452*** (-29.2660) | -0.0460*** (-29.7149) | -0.0460*** (-29.4735) | -0.0456*** (-29.0010) |
| 交通便利性               | 0.0004*** (2.9558)    | 0.0003*** (2.6351)    | 0.0003*** (2.6955)    | 0.0003*** (2.5532)    |
| 医疗条件                | -0.0240*** (-5.9121)  | -0.0100*** (-4.2810)  | -0.0020 (-1.5669)     | 0.0000 (0.2880)       |
| 建筑面积                | 0.0019*** (20.8724)   | 0.0019*** (21.1537)   | 0.0020*** (21.3765)   | 0.0020*** (21.4305)   |
| 是否 2 居以上            | -0.0655*** (-6.3006)  | -0.0655*** (-6.2914)  | -0.0664*** (-6.3611)  | -0.0664*** (-6.3600)  |
| 是否 3 居以上            | -0.1028*** (-9.7738)  | -0.1038*** (-9.8524)  | -0.1046*** (-9.9098)  | -0.1048*** (-9.9214)  |
| 是否 4 居以上            | -0.1418*** (-10.2619) | -0.1436*** (-10.3728) | -0.1446*** (-10.4315) | -0.1451*** (-10.4650) |
| 是否装修                | -0.0642*** (-7.6602)  | -0.0626*** (-7.4587)  | -0.0629*** (-7.4858)  | -0.0630*** (-7.4956)  |
| R <sup>2</sup>      | 0.6209                | 0.6194                | 0.6179                | 0.6177                |
| 调整后的 R <sup>2</sup> | 0.6201                | 0.6186                | 0.6171                | 0.6169                |
| F 统计量               | 749.6189              | 744.7869              | 740.1739              | 739.4845              |
| F 值的 p 检验值          | 0.0000                | 0.0000                | 0.0000                | 0.0000                |

## (四)加入公共场所的住宅特征价格

公共场所对于居民生活质量的提升是显然的,因此其对周边小区的住宅价格应该有显著正向影响。下面对加入公共场所的住宅价格特征进行分析,表 15 是共线性检验结果,方差膨胀因子显示并不存在共线性问题。

表 15 包含公共场所方程的方差膨胀因子

| 方差膨胀因子   | 小区 300m 范围 | 小区 500m 范围 | 小区 800m 范围 | 小区 1km 范围 |
|----------|------------|------------|------------|-----------|
| 距离平方     | 11.8200    | 11.8418    | 11.9408    | 11.9209   |
| 距离       | 11.8685    | 11.9094    | 12.0863    | 12.1028   |
| 交通便利性    | 1.0490     | 1.0487     | 1.0552     | 1.0546    |
| 公共场所     | 1.0092     | 1.0135     | 1.0350     | 1.0417    |
| 建筑面积     | 3.6366     | 3.6386     | 3.6364     | 3.6360    |
| 是否 2 居以上 | 1.5173     | 1.5173     | 1.5175     | 1.5174    |
| 是否 3 居以上 | 2.1638     | 2.1639     | 2.1638     | 2.1638    |
| 是否 4 居以上 | 2.6384     | 2.6384     | 2.6397     | 2.6392    |
| 是否装修     | 1.0217     | 1.0219     | 1.0222     | 1.0221    |
| 均值       | 4.0805     | 4.0882     | 4.1219     | 4.1221    |

回归结果如表 16 所示,与不加入公共场所的回归方程相比,系数和方程稳定性变化都不大。但一个有意思的现象是:公共场所距离小区越近,其对小区住宅价格的影响是负向的,这可能是因人流、车流形成的各种污染造成的;随着距离的扩大,这种负面影响效应在下降,对生活品质的正向影响效应在上升;但是达到一定距离后,正向影响效应也成下降趋势。表 16 反映的各种距离条件下的回归结果表明,在公共场所 800 米范围内,其对住宅价格的影响效应最大。

表 16 包含公共场所的住宅特征价格回归结果

|                     | 模型一                   | 模型二                   | 模型三                   | 模型四                   |
|---------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| 截距项                 | 10.4738*** (595.5697) | 10.4710*** (587.8696) | 10.4566*** (566.8424) | 10.4566*** (559.0061) |
| 距离平方                | 0.0003*** (7.8187)    | 0.0003*** (7.7006)    | 0.0003*** (7.2060)    | 0.0003*** (7.4013)    |
| 距离                  | -0.0459*** (-29.5449) | -0.0458*** (-29.3834) | -0.0452*** (-28.6460) | -0.0452*** (-28.8140) |
| 交通便利性               | 0.0003** (2.5354)     | 0.0003*** (2.5681)    | 0.0003*** (2.8258)    | 0.0003*** (2.7065)    |
| 公共场所                | -0.0150* (-1.9559)    | -0.0020 (-0.5972)     | 0.0060*** (3.1784)    | 0.0030* (1.7711)      |
| 建筑面积                | 0.0020*** (21.4099)   | 0.0020*** (21.4052)   | 0.0020*** (21.4158)   | 0.0020*** (21.4423)   |
| 是否 2 居以上            | -0.0665*** (-6.3692)  | -0.0664*** (-6.3600)  | -0.0663*** (-6.3348)  | -0.0663*** (-6.3510)  |
| 是否 3 居以上            | -0.1047*** (-9.9203)  | -0.1047*** (-9.9149)  | -0.1048*** (-9.9270)  | -0.1048*** (-9.9282)  |
| 是否 4 居以上            | -0.1453*** (-10.4848) | -0.1452*** (-10.4687) | -0.1446*** (-10.3969) | -0.1446*** (-10.4311) |
| 是否装修                | -0.0630*** (-7.4955)  | -0.0630*** (-7.4882)  | -0.0633*** (-7.5728)  | -0.0633*** (-7.5291)  |
| R <sup>2</sup>      | 0.6181                | 0.6177                | 0.6186                | 0.6180                |
| 调整后的 R <sup>2</sup> | 0.6172                | 0.6169                | 0.6178                | 0.6172                |
| F 统计量               | 740.5723              | 739.5640              | 742.3964              | 740.3720              |
| F 值的 p 检验值          | 0.0000                | 0.0000                | 0.0000                | 0.0000                |



七、住宅特征价格在时间序列上的表现

以上第四节至第六节都是基于截面数据的住宅特征价格分析,或者说北京的住宅价格在空间特征上的表现,本节加入时间序列因素,在确定空间特征因素后重点观察其在时间上的动态表现。

(一)考虑时间因素的总方程回归结果

为不同季度的住宅特征数据标记上时间虚拟变量,运用第二节中的一次性建模方法计算住宅特征价格。第六节对住宅的邻里特征进行了分析,在这章的方程建模中,我们选择对住宅价格正向影响程度最大的邻里特征进入方程,即小区 1 公里半径这节范围的教育环境和小区 800 米半径范围的公共场所情况,其他建筑特征和区位特征保持不变。

在进行总方程回归结果考察前,首先还是观察共线性的问题,表 17 给出了方程中共线性检验的结果,结果表明方程各因素之间并不存在共线性问题。

表 17 总方程回归中的方差膨胀因子

|                  |                  |                  |                  |                  |                  |                  |
|------------------|------------------|------------------|------------------|------------------|------------------|------------------|
| 距离平方             | 距离               | 交通便利性            | 教育环境             | 公共场所             | 建筑面积             | 是否 2 居以上         |
| 9.7390           | 10.9939          | 1.1199           | 1.5593           | 1.0775           | 3.4182           | 1.5108           |
| 是否 3 居以上         | 是否 4 居以上         | 是否装修             | 是否 2009<br>Q2 以后 | 是否 2009<br>Q3 以后 | 是否 2009<br>Q4 以后 | 是否 2010<br>Q1 以后 |
| 2.1557           | 2.6155           | 1.0484           | 1.7293           | 3.0921           | 4.1938           | 5.1436           |
| 是否 2010<br>Q2 以后 | 是否 2010<br>Q3 以后 | 是否 2010<br>Q4 以后 | 是否 2011<br>Q1 以后 | 是否 2011<br>Q2 以后 | 是否 2011<br>Q3 以后 | 是否 2011<br>Q4 以后 |
| 5.8170           | 6.0827           | 5.8687           | 5.4283           | 4.5479           | 3.2100           | 1.7702           |

表 18 给出加入时间因素的住宅特征价格回归结果。从系数的估计结果来看,所有的表示空间特征的因素系数都没有发生大的变化;方程的截距项显著下降,这是因为之前所有的空间特征分析都是基于 2010 年第四季度数据,而加入时间因素的初始起点在 2009 年第一季度,截距项的减少都在后续的时间哑变量中得到了体现。由于方程的被解释变量为单价取对数,因此方程的系数可解释为弹性或变化百分比。比如距离系数为-0.0491,表示距离每增加一公里价格下降 4.91%,当然这里还得考虑二次项的变化系数;交通便利性系数为 0.0003,表示小区 800 米半径范围上每增加一个公交站点(或 1/5 座地铁站),价格增加 0.03%。对于时间虚拟变量系数,实际上代表了某一季度相对于前一季度的环比变化情况。比如“是否 2009Q2 以后”系数为 0.0566,其含义是 2009 年第二季度相对于 2009 年第一季度增长 5.66%;“是否 2011Q2 以后”系数为 0.0037,表示相对于上一季度上涨 0.37%,该数值非常小并且并没有通过 T 检验。这一数据反映的一个有趣的现象是,2011 年 1 月 26 日是国务院以“限购”为核心的“国八条”出台时间,2011 年第二季度相对于第一季度的数据表现是上涨几乎停滞,P 值未能拒绝原假设其实更好地描述了当时的市场分歧状态。

(二)基于特征价格的二手房价格趋势及与统计局结果的比较

我们将上文时间虚拟变量的系数绘制在图中,就可以得到基于特征价格的二手房价格环比变动图。同时我们将统计局的二手房价格环比数据进行月度相乘再减去 1,即可得到季度的环比价格图,因为季度环比变化是月度环比的累积。

图 7 是应用特征价格法和统计局的样本匹配法得到的二手房价格环比情况,对比两个时间序列可以发现如下差异:第一,在 2009 年至 2010 年上半年这段时期,统计局的样本匹配法的上涨幅

表 18 加入时间因素的住宅特征价格回归结果

| 行业           | 系数      | 标准差    | T 统计量     | P 值    |
|--------------|---------|--------|-----------|--------|
| 截距项          | 9.7754  | 0.0071 | 1385.1762 | 0.0000 |
| 距离平方         | 0.0004  | 0.0000 | 33.6181   | 0.0000 |
| 距离           | -0.0491 | 0.0005 | -105.7045 | 0.0000 |
| 交通便利性        | 0.0003  | 0.0000 | 6.9274    | 0.0000 |
| 教育环境         | 0.0003  | 0.0001 | 2.5021    | 0.0123 |
| 公共场所         | 0.0053  | 0.0006 | 8.8184    | 0.0000 |
| 建筑面积         | 0.0019  | 0.0000 | 75.2280   | 0.0000 |
| 是否 2 居以上     | -0.0538 | 0.0032 | -16.9452  | 0.0000 |
| 是否 3 居以上     | -0.0950 | 0.0032 | -29.8272  | 0.0000 |
| 是否 4 居以上     | -0.1350 | 0.0041 | -33.2131  | 0.0000 |
| 是否装修         | -0.0545 | 0.0024 | -22.6401  | 0.0000 |
| 是否 2009Q2 以后 | 0.0566  | 0.0050 | 11.2455   | 0.0000 |
| 是否 2009Q3 以后 | 0.1105  | 0.0049 | 22.5722   | 0.0000 |
| 是否 2009Q4 以后 | 0.1204  | 0.0049 | 24.4740   | 0.0000 |
| 是否 2010Q1 以后 | 0.1738  | 0.0051 | 34.3209   | 0.0000 |
| 是否 2010Q2 以后 | 0.2184  | 0.0052 | 41.6886   | 0.0000 |
| 是否 2010Q3 以后 | -0.0156 | 0.0054 | -2.8943   | 0.0038 |
| 是否 2010Q4 以后 | 0.0373  | 0.0054 | 6.8489    | 0.0000 |
| 是否 2011Q1 以后 | 0.0702  | 0.0056 | 12.6113   | 0.0000 |
| 是否 2011Q2 以后 | 0.0037  | 0.0056 | 0.6652    | 0.5059 |
| 是否 2011Q3 以后 | -0.0119 | 0.0056 | -2.1327   | 0.0330 |
| 是否 2011Q4 以后 | -0.0289 | 0.0058 | -4.9734   | 0.0000 |

度明显小于特征价格法,大家对存量房价格变动的感受可能更接近特征价格法。第二,2010 年第二季度以“限贷”为核心的“国十条”政策出台后,两种统计方法表现出差异。特征价格法的表现是,2010 年第三季度相对于第二季度略微下降后又保持了几乎 5%的增长率,反映了“限贷”政策没有起到抑制房价上涨作用的事实。蔡真(2011)<sup>①</sup>论述“限贷”政策不起作用的大经济背景是流动性过剩,具体的佐证事实是链家地产和光大银行 2010 年 8 月发布的《北京市二手房房贷市场分析报告》:其一,在“限贷”政策实施后一次性付款比例由新政前 38.9%上升至新政后 59.9%;其二,市场更加倾向小户型,表明那些受到一定资金约束的人依然有投资热情。统计局的样本匹配法反映的情况是限贷政策起到应有的作用,即房价停止上涨了。第三,2011 年第一季度以“限购”为核心的“国八条”政策出台后,两种统计方法也表现出差异。特征价格法反映的价格下降幅度明显大于统计局的样本匹配法。我们认为“限购”在执行之初是起到打击房价作用的,因为这一政策明显抑制了住房需求。总体而言,我们倾向于认为特征价格法较好地反映了住宅这种资产价格的波动特征。

此外,图 7 还展示了统计局的二手房指数和新建房指数的对比情况,整体而言二手房指数的上升幅度和下降幅度都小于新建房的变动情况。这似乎与人们的一般认识存在偏差:这一轮房价变动周期中,大家普遍感觉存量房为主的城区的房价变动要大于新建房为主的郊区。

## 八、总结及进一步研究展望

本文对北京存量住宅市场的价格特征进行了详细深入的分析。在时间序列上,时间虚拟变量

<sup>①</sup> 参见蔡真(2011):《房价调控失效的宏观原因及房价上涨的微观机制》,《中国房地产业》,第 11 期。

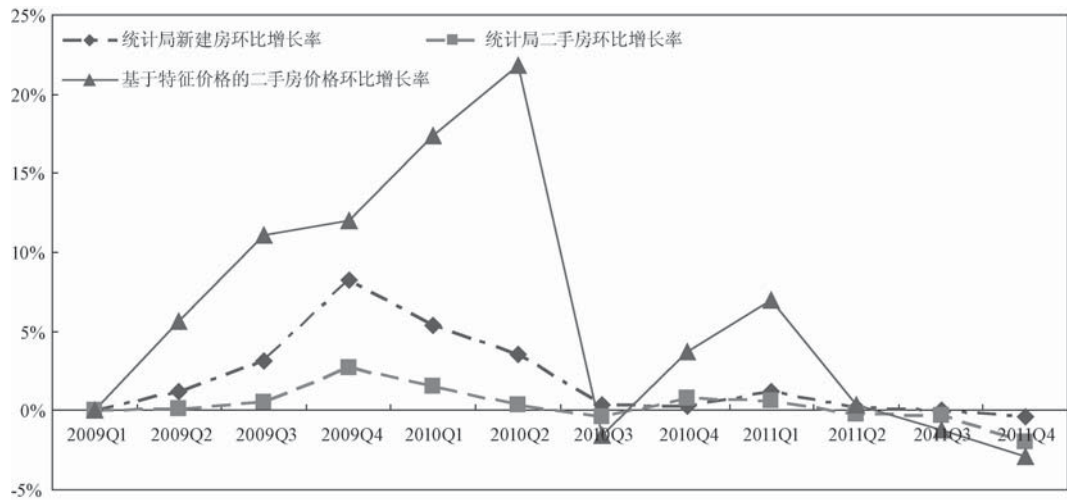


图7 两种统计方法的二手房价格变化对比图

反映的价格动态变化,一方面很好地刻画了2009年至2010年上半年房价大幅上涨的特征,与人们对市场的感受认知相符;另一方面,它也很好反映了“限贷”和“限购”两种政策的不同效果。我们倾向于认为特征价格法描述的价格变动更接近客观事实。

在时点层面上,截面数据表现出明显的空间和区位特征:以天安门为中心距离每增加一公里,住宅单价大约下降5%,且这种下降呈加速趋势;交通便利性对房价也存在显著正影响,即在小区800m半径范围内,公交站点每增加一个,房价上涨0.03%;北京的住宅价格在南北两部分也表现出差异,所谓的“上风上水上海淀”仅仅描述出北部整体的房价高于南部,事实上北部房价下降的速度和加速度都高于南部(这一点可以用股票的阿尔法和贝塔理解)。

邻里特征对房价的作用并不是简单的线性关系,事实上周边环境对房价起正向作用发生在一定范围内,这反映了空间经济理论中的推拉两种效应共同作用的结果,也与Lutzenhiser(2001)关于公园与住宅价格的关系一致。比如本文的实证表明,如果公共场所距小区800米范围,其对房价的影响是最大的,过大或过小都会产生负面影响;关于教育,太近的距离没有显著性,对1公里范围内的小区有轻微的正影响;关于医疗条件,则是完全的负面影响;商业环境对300米范围内的小区有负面影响,更大的半径条件下它不是房价的决定因素。

关于建筑特征的实证分析结论如下:装修的住宅比未装修的贵;面积大的住宅比面积小的住宅便宜,这种便宜仅仅在住宅功能发生改变时才成立,即在户型扩大的情况下住宅单价才有下降趋势;而在相同户型条件下,增加户型面积是需要多支付价格的。此外,这些户型和装修特征还在空间和区位上表现出差异:大户型由中心向郊区发展其价格相对于小户型有更快的下降幅度和速度;装修特征对于城区和郊区来说有不同的表现形式,城区中装修住宅比未装修住宅贵的可能性和幅度都大;而在郊区,毛坯房更贵的可能性和幅度都大。

以上实证分析呈现出如此完美的结果,实际上归因于住宅特征的相互独立性,或者说这些特征几乎处于正交空间中,因此不存在共线性问题。正是基于这一点认识,下一步的研究中无须担心加入新的特征后会出现实证分析效果下降的情形。对于当前的研究结果,我们认为还有如下三点值得进一步改进:第一,需要加入建筑年代特征,距离接近的两个小区,在其他建筑特征相同的条件下,旧住宅与新住宅存在显著差异,根据切身感受,我们认为这一影响因素的系数至少大于建筑面积的系数;第二,当前教育特征的标注是将价格与距教育设施的距离联系一起的,现实中完全存

在一线之隔的两栋住宅因学区划片不同导致价格显著差异的情况,因此关于教育特征的标注需要用学区划片标识,而非距离标识;第三,当前的邻里特征反映的是最后时点的情况,并没有动态变化,而这就可能忽略了城市生长对住宅价格的推动作用。

#### 参考文献

- 奥沙利文(2008):《城市经济学(第6版)》,北京大学出版社。
- 蔡真、汪利娜(2011):《房价与信贷关系研究——兼论当前房价调控政策的有效性》,《金融评论》,第1期。
- 蔡真(2011):《房价调控失效的宏观原因及房价上涨的微观机制》,《中国房地产业》,第11期。
- 吴璟、刘洪玉、马亚男(2007):《住房价格指数的主要编制方法及其选择》,《建筑经济》,第7期。
- 张宏斌、贾生华(2000):《编制城市房地产价格指数的理论模型和实用方法》,《中国软科学》,第4期。
- Adair, S., J. Berry and W. McGreal (1996): “Hedonic Modeling, Housing Submarkets and Residential Valuation”, *Journal of Property Research*, 13, 67–83.
- Bin, O., J. Kruse and C. Landry (2008): “Flood Hazards, Insurance Rates, and Amenities: Evidence from the Coastal Housing Market”, *The Journal of Risk and Insurance*, 75, 63–82.
- Case, B., H. Pollakowski and S. Wachter (1991): “On Choosing among House Price Index Methodologies”, *Real Estate Economics*, 19, 286–307.
- Case, E. and R. Shiller (1989): “The Efficiency of the Market for Single-Family Homes”, *American Economic Review*, 79, 125–137.
- Clauretie, M. and H. Neill. (2000): “Year-round School Schedules and Residential Property Values”, *Journal of Real Estate Finance and Economics*, 20, 311–322.
- Colwell, F., S. Gujral and C. Coley (1985): “The Impact of a Shopping Center on the Value of Surrounding Properties”, *Real Estate Issues*, 10, 35–39.
- Espey, M. and H. Lopez (2000): “The Impact of Airport Noise and Proximity on Residential Property Values”, *Growth and Change*, 31, 408–419.
- Ferri, M (1977): “An Application of Hedonic Indexing Methods to Monthly Changes in Housing Prices: 1965–1975”, *Real Estate Economics*, 5, 455–462.
- Forrest, D., J. Glen and R. Ward (1996): “The Impact of a Light Rail System on the Structure of House Prices: A Longitudinal Study”, *Journal of Transport Economics and Policy*, 30, 15–29.
- Grass, G (1992): “The Estimation of Residential Property Values around Transit Station Sites in Washington, D.C.”, *Journal of Economics and Finance*, 16, 139–146.
- Hill, J., D. Melser and I. Syed (2009): “Measuring a Boom and Bust: The Sydney Housing Market 2001–2006”, *Journal of Housing Economics*, 18, 193–205.
- Lutzenhiser, M. and N. Netusil (2001): “Effect of Open Space on a Home’s Sale Price”, *Contemporary Economic Policy*, 19, 291–298.
- Michael, J., K. Boyle and R. Bouchard (2000): “Does the Measurement of Environmental Quality Affect Implicit Prices Estimated from Hedonic Models”, *Land Economics*, 76, 283–298.
- Nelson, J. (2004): “Meta-Analysis of Airport Noise and Hedonic Property Values: Problems and Prospects”, *Journal of Transport Economics and Policy*, 38, 1–28.
- Ridker, G. and J. Henning (1967): “The Determinants of Residential Property Values with Special Reference to Air Pollution”, *Review of Economics and Statistics*, 49, 246–257.
- Sirmans, S. and D. Macpherson (2005): “The Composition of Hedonic Pricing Models”, *Journal of Real Estate Literature*, 13, 3–41.
- So, M., R. Tse and S. Ganesan (1997): “Estimating Influence of Transport on House Prices: Evidence from Hong Kong”, *Journal of Property Valuation and Investment*, 15, 40–47.
- Zabel, E. and K. Kiel (2000): “Estimating the Demand for Air Quality in Four U.S.”, *Cities Land Economics*, 76, 174–194.

(责任编辑:罗 滢)