



中国典型城市住房同质价格指数 技术白皮书

北京大学—林肯研究院城市发展与土地政策研究中心

清华大学恒隆房地产研究中心

2014 年 11 月

1 指数编制概要

1.1 指数简介

“中国典型城市住房同质价格指数（Chinese Quality-Controlled Housing Price Index，简称 CQCHPI）”由北京大学林肯研究院城市发展与土地政策研究中心（Peking University-Lincoln Institute, Center for Urban Development and Land Policy）和清华大学恒隆房地产研究中心（Hang Lung Center for Real Estate, Tsinghua University）联合编制。该指数采用国际先进的同质价格指数编制方法进行编制，并针对新建商品住房市场和存量住房市场设计了五类住房市场价格指数产品。作为国内首个由高校学术机构编制并公开发布的住房价格指数，CQCHPI 指数在编制方法上与国际学术前沿接轨，针对我国新建商品住房市场和存量住房市场的数据特点，分别设计了特征价格法（Hedonic 模型）和混合模型法（特征价格方法与重复交易法相结合，Hedonic-Repeat-Sale Hybrid Model，或称“类重复交易模型”）对指数进行编制，能够有效控制住房物理属性以及住房所处的区位属性，避免住房交易样本异质性对价格度量结果的影响。CQCHPI 指数的推出将为市场参与者、研究者和政府决策者更加全面和准确的把握中国住房市场价格走势提供重要的参考依据。

现阶段 CQCHPI 指数覆盖取北京、上海、天津、深圳、成都、大连、武汉和西安八个典型城市，今后还将视需要进行必要的扩充。

1.2 主要参数设定

1.2.1 更新周期

CQCHPI 指数采用季度为更新周期，于每年 2 月、5 月、8 月、11 月的 5 日前发布上一季度的指数。

1.2.2 基期

- 新建商品住房同质价格指数以 2006 年第一季度为基期。

➤ 存量住房同质价格指数根据不同城市的数据起始点不同选取基期如下：

- 北京、上海、天津、深圳：2009 年第一季度；
- 成都、大连：2012 年第二季度；
- 武汉、西安：2012 年第一季度。

1.2.3 价格

CQCHPI 指数所指的价格是指以按住房建筑面积计价的单位面积价格，币种为人民币。价格指数未经 CPI 调整。

1.3 编制范围

1.3.1 新建商品住房同质价格指数

纳入计算的新建商品住房应满足以下条件：

- 该交易单元属于已经取得房地产预售许可证或房地产销售许可证的房地产开发项目；
- 该交易单元由买受人通过“商品房销售合同”（或“商品房预售合同”、“商品房买卖合同”，下同）向物业开发企业或其代理企业直接购买；
- 该交易单元在“商品房销售合同”中规定的用途属于“普通住宅”或“公寓”。

满足上述规则的交易单元，不区分结构形式、户型、面积、所在楼层、总楼层等特征，均属于“新建商品住房同质价格指数”的编制范围。

1.3.2 存量住房同质价格指数

纳入计算的存量住房是指在房地产中介经纪公司挂牌的二手住房，应满足以下条件：

- 有完全产权，可以上市自由交易获得其全部市场价值；

- 该交易单元属于“普通住宅”或者“公寓”。

1.4 细分指数

1.4.1 中心城区指数

中心城区价格指数区分于整个城市层面编制的指数，仅计算位于中心城区的住房样本。总体来说，中心城区的选取按照城市经济学中关于“经济意义上的城市”的概念，即统一的劳动力市场和整合的住房市场所覆盖的空间区域。在实际应用中，参考该城市正在使用的《城市总体规划》中划定的“市区”（不同城市的《城市总体规划》中，使用的类似概念包括“中心城市”、“主城”等）范围。各个城市中心城区选取情况见附表 1。我们编制了 8 个城市中心城区的相应指数，之后利用各城市交易面积总量（表征城市住房市场规模）加权平均获得 8 城市总体指数。

特别指出，随着城市空间结构的动态演进，我们可能会对中心城区的空间范围及边界进行调整。但这种调整不会很频繁。

1.4.2 区位于市场指数

区位于市场同质价格指数（简称“区位于市场指数”）通过选取城市内部以城市主中心或次中心为核心的活跃居住区进行指数编制，以衡量该区域住房价格水平。以北京市为例，选取国贸 CBD、中关村、亚运村、金融街作为区位于市场指数编制的区位对象。这四个区位为人们所熟知和关注，同时在空间上分布在北京市的不同位置，能够反映北京市不同区域的住房价格差异。

区位于市场指数给指数产品的受众更加直观的感受，如中关村附近的房价涨了多少或者国贸 CBD 附近的房价涨了多少。由于大城市的住房市场内部存在着显著的空间差异，例如城市中心商务区和城市郊区的房价水平和增速均存在显著差异，因此区位于市场指数的编制弥补了单一全城价格指数的不足，丰富了城市住房价格指数的产品。同时，由于能够针对城市内任意区位编制，我们可以根据不同的需要增加新的指数编制对象。对于每个城市中区位于市场的选取，我们遵循以下原则，同时可根据需要调整：

- 选取的区位对象尽可能覆盖城市的不同区域，不同区位对象之间的距离适中，使得指数产品的受众能够感受到城市不同区域的住房价格差异，并对城市住房价格的空间差异产生直观的感受。
- 选取的区位对象是该城市以主中心或次中心为核心的活跃居住区，即该城市居民和商业机构等认可的有代表性的区位，例如对于上海，“人民广场”是上海市公认的传统意义上的中心，“陆家嘴”是浦东新区的中心等。如此可以使指数产品的受众能够很快通过这些名字联想到相应的区域位置，从而使得区域子市场指数具有较好的现实意义。

目前各城市区位子市场的选取情况见附表 2。

特别指出，随着城市空间结构的动态演进，我们可能会对区位子市场的选择有增补和删减。但这种调整不会很频繁。如果未来有调整，将会补充公布新增区位子市场的历史指数数据。

1.4.3 面积子市场指数

住房市场在不同住房面积区间存在着户型、定位和政府政策的显著差异，这体现了城市中住房市场的档次分异，因此有必要构建按单套面积划分子市场指数（简称“面积子市场指数”）来反映这种差异。

面积子市场指数将住房市场按单套住房面积分为 5 个区间，分别为 70 平方米以下，70-90 平方米，90-120 平方米，120-140 平方米，140 平方米以上。由于新建商品住房的数据分辨率为“预售许可证号-月份”，无单套面积数据，因此只计算存量住房的面积子市场指数。

将住房市场划分成不同面积区间的子市场具有以下两个层面的意义：

- 基于户型、消费者定位差异等。如 70 平方米以下的住房往往采用“两室一（两）厅”的户型设计，90-120 平方米住房有条件采用“三室一（两）厅”的户型设计；此外，不同面积区间的住房针对的消费者人群也显著不同。

- 基于政府政策，尤其是各地纳税征税依据。如国家住建部对开发商拿地建房有相关的要求，70-90 平方米的住房占比不能低于 70%；而 140 平方米以上的住房属于纳税中非普通住宅标准，部分城市的非普通住宅标准的门槛是 144 平方米。

1.5 引用须知

目前中国典型城市住房同质价格指数供公开、免费下载。但在任何刊物或者公开场合引用时，需同时引用以下 2 个参考文献：

- [1] Siqi Zheng, Weizeng Sun, Joyce Yanyun Man and David Geltner. The Resale Home Market Dynamics in Major Chinese Cities: Quality Controlled Index Construction and Analysis. Working Paper. Hang Lung Center for Real Estate at Tsinghua University; PKU-Lincoln Center for Urban Development and Land Policy at Peking University; Center for Real Estate at MIT.
- [2] Xiaoyang Guo, Siqi Zheng, David Geltner and Hongyu Liu. A New Approach for Constructing Home Price Indices: The Pseudo Repeat Sales Model and Its Application in China. *Journal of Housing Economics*, 2014, 25: 20-38.

2 数据采集与数据处理

可靠的数据和严谨的数据处理是 CQCHPI 指数编制工作顺利开展和编制结果优质可靠的重要前提。世联地产作为 CQCHPI 指数编制的战略支持机构，长期提供可靠的数据。CQCHPI 项目组对数据进行了有效的筛选和必要的处理，最终形成指数编制所需的新建商品住房和存量住房数据库。

2.1 数据来源

- 新建商品住房同质价格指数

各城市中心城区新建商品住宅网签全样本价格数据（按月度和预售许可证号为样本分辨率）。

➤ 存量住房同质价格指数

世联地产提供的包括链家地产、麦田、我爱我家等在内的 3 家房地产中介经纪公司的单套挂牌价格数据。

2.2 价格指数

➤ 新建商品住房同质价格指数

新建商品住房数据库中的价格指数为住宅项目月成交均价，即该月该住宅项目的总成交价格/该月该住宅项目的总成交面积。该指标将作为因变量以对数形式引入特征价格模型，作为进一步计算同质性价格指数的基础。

➤ 存量住房同质价格指数

存量住房数据库中的价格指数为单套住房的“挂牌”价格（总价），即房地产中介经纪公司提供的报价。报价和真实交易价格会有一定的差异，这种差异来源于两个方面：第一，为了吸引购房者，房地产中介经纪公司会降低报价，而在实际住房交易过程中以其他形式提高价格，使得最终成交价格高于报价；第二，房地产中介经纪公司有时也会提高报价，以形成一定的议价空间，满足购房者的“还价心理”，此时，最终成交价格低于报价。所以，报价可能高于报价，也可能低于真实成交价格。尽管如此，报价仍然是目前市场上能够获得的较为准确的存量住房价格指数，原因主要有以下两个方面：

第一，真实的存量住房成交易价格是很难的，甚至不可能大量获得。由于避税行为等，房地产交易管理部门登记的成交价格与真实成交价格在不同程度上偏低。

第二，房地产中介经纪公司的报价是能够批量获得的反映市场情况的价格指标。因为房地产中介经纪公司所提供的报价是其与房主沟通协商的结果，既反映了卖家的预期，也考虑了房地产中介经纪公司根据市场情况、实际住房属性、已经成交的类似住房交易情况等对卖家预期的“修正”。

一些学术研究表明，报价与真实交易价存在较稳定的关系，因此基于报价编制房价指数能较好反映市场走势。该指标将作为因变量以对数形式引入此后构建

的混合模型中，作为进一步计算同质性价格指数的基础。

2.3 住房属性

除成交价格外，新建商品住房数据库和存量住房数据库中的每一个样本还包括以下属性信息。这些属性将在同质性价格指数编制过程中作为控制变量引入。

➤ 新建商品住房同质价格指数

- 物理属性：
 - 小区规模：以该小区住宅总套数来表征，取对数。
 - 住房面积：住宅项目平均每套住房的建筑面积（住房面积=当期成交总面积/当期成交住房套数），单位为平方米，取对数。
 - 容积率：住宅小区的容积率，即总住房面积与用地面积的比率。
 - 绿化率：住宅小区的绿化率。
 - 装修状况：住宅小区的装修类型，含 2 个虚拟变量。
- 区位属性利用空间栅格予以控制，具体参见 2.4.3。如该城市为典型的单中心城市，增加“到城市中心距离”变量。
- 成交季度：成交价格所发生的季度，为一系列季度虚拟变量，该变量前面的系数用于生成指数序列。

➤ 存量住房同质价格指数

- 物理属性：
 - 住房面积：单套住房的建筑面积，单位为平方米。
 - 所在楼层：单套住房在楼栋中所处的楼层。
- 区位属性利用项目虚拟变量予以控制，具体参见 2.4.3。
- 成交季度：单套住房成交所在季度，为一系列季度虚拟变量，该变量前面的系数用于生成指数序列。

2.4 数据筛选和处理

2.4.1 样本筛选

由于原始数据存在变量缺失和数据错漏情况，在对样本进行进一步的处理之前，首先需要进行样本筛选。样本筛选按照以下步骤进行：

- 删除缺失关键住房属性或住房价格的样本数据，如某新建商品住房样本中装修属性缺失，则删除整条样本。
- 删除住房属性异常的样本数据，如删除存量住房样本中单套住房面积小于 20 平方米的样本。
- 删除价格严重偏离普通商品住房市场价格的样本，以北京市为例，删除总价超过 5000 万元和低于 20 万元的住房样本。

具体样本筛选标准视样本情况而定。

2.4.2 样本空间化和区位属性计算

样本的空间化的目的是为计算区位属性奠定基础，利用经纬度查询工具，批量查询每个项目所对应的空间坐标位置，由此构建可以 GIS 化的新建商品住房和存量住房的空间数据库。

新建商品住房是项目层面数据，采用特征价格法。为保证区位属性得到有效的控制，我们将各个城市的整个市域空间划分为 $2\text{km} \times 2\text{km}$ 的栅格 (LATTICE) 并编号，每一个新建商品住房样本都将位于市域空间中的某一个栅格内，对应相应栅格编号。如此，每一个样本都拥有一个栅格变量，用于控制空间区位属性。

存量住房是微观数据，采用混合模型法，住房样本在项目层面进行匹配。

2.4.3 生成虚拟变量

虚拟变量 (Dummy Variables) 又称虚设变量、名义变量或哑变量，用以反映质的属性的一个人工变量，是量化了的自变量，通常取值为 0 或 1。引入哑变量可使线形回归模型变得更复杂，但对问题描述更简明，且接近现实。需要生成虚

拟变量的属性如下：

- **装修状况变量：**以北京市为例，原始数据包含每个住房样本的装修状况：毛坯房、精装修、简装修。分别对应装修状况属性（DECO）的值的 1、2、3，我们可以通过 STATA 软件生成装修状况的虚拟变量，原始数据中的 DECO 属性被分解成 DECO1、DECO2、DECO3。如果住房样本的装修状况为毛坯，则 DECO1 属性值为 1，其他 DECO 属性值为 0；如果住房样本装修状况为精装修，则 DECO2 属性值为 1，其他属性 DECO 属性值为 0，以此类推所有样本点都生成相应 DECO1-DECO3 的属性值。
- **成交季度虚拟变量：**以北京新建商品住房同质价格指数为例，其原始数据中成交季度属性（TQUARTER）为 1-35，分别对应从基期 2006 年第一季度开始到 2014 年第三季度的 35 个季度。如某个住房样本在 2012 年第一季度交易并记录，则该样本所对应的成交季度属性值为 25。通过 STATA 软件生成成交季度的虚拟变量，原数据中的 TQUARTER 属性被分解成 TQUARTER1、TQUARTER2 至 TQUARTER35，如果住房样本在 2012 年第一季度交易并记录，TQUARTER25 的属性值为 1，其他 TQUARTER 属性值为 0，以此来推所有样本点都生成相应的 TQUARTER1-TQUARTER35 的属性值。
- **栅格虚拟变量：**以北京新建商品住房同质价格指数为例，其原始数据中栅格变量（LATTICE）为 1-395，栅格变量的每个数值（即栅格的编号）对应一个 2km×2km 的栅格，如果某个住房样本处在第 30 个栅格里，则该样本所对应的栅格变量值为 30。通过 STATA 软件生成栅格的虚拟变量，原数据中的 LATTICE 属性被分解成 LATTICE1、LATTICE2 至 LATTICE395。如果住房样本在第 30 个栅格里，则 LATTICE30 的属性值为 1，其他 LATTICE 属性值为 0，以此来推所有样本点都生成相应的 LATTICE1-LATTICE395 的属性值。
- **项目虚拟变量：**以北京存量住房同质价格指数为例，其原始数据中项目变量（PROJECT）为 1-561（561 为 2014 年第三季度更新时样本中项目

数), 每个项目变量对应 $2\text{km} \times 2\text{km}$ 的项目, 如某个住房样本处在第 30 个项目里, 则该样本所对应的项目属性值为 30。通过 STATA 软件生成项目的虚拟变量, 原数据中的 PROJECT 属性被分解成 PROJECT1、PROJECT2 至 PROJECT561。如果住房样本在第 30 个项目里, 则 PROJECT30 的属性值为 1, 其他 PROJECT 属性值为 0, 以此来推所有样本点都生成相应的 PROJECT1-PROJECT561 的属性值。

2.4.4 属性的对数化

无论是特征价格法还是特征价格法和重复交易法相结合的混合模型 (或称“类重复交易模型”), 都涉及到统计上的回归分析, 为了使数据的呈现方式接近我们所希望的前提假设, 从而更好的进行统计推断, 我们需要对某些属性进行对数化再引入模型。

对于新建商品住房, 我们对住房价格、住房面积、小区规模取对数。其他百分比形式或者需要转换为虚拟变量的属性不进行对数化。

对于存量住房, 我们对住房价格、住房面积、所在楼层取对数。其他百分比形式或者需要转换为虚拟变量的属性不进行对数化。

3 指数编制方案

CQCHPI 指数采用特征价格法 (针对新建商品住房) 和特征价格法与重复交易法相结合的混合模型法 (或称“类重复交易模型”, 针对存量住房) 编制, 其核心思想都在于构造同质可比的价格指数。由于住房是一种典型的异质性商品, 严格意义上没有任何两宗住房单元是完全一致的, 而且很多差异都会对住房价格产生巨大的影响。因此, 不同时期观察到的住房市场价格水平的波动不仅源于市场供求情况的改变, 也会源于市场中物业品质、档次的整体提升或物业聚集地的整体变化, 还可能源于各期交易量中不同特征住房组成结构的变化。所以, 这就违背了价格指数编制中应当遵循的“同质可比”的前提条件, 因而需要尽可能在相似的样本空间内比较住房价格的变化, 并对住房样本进行“质量调整”, 即在

指数编制过程中剔除不同样本（住房单元）之间的质量差异。

3.1 编制方法

3.1.1 指数编制模型总览

CQCHPI 指数中新建商品住房数据是项目层面数据，每一条交易样本表示某月某项目的成交情况和项目属性；而存量住房数据是微观交易数据，每一条交易样本表示某月某项目某套住房的成交情况和住房属性。由于数据形式的不同，在指数编制时针对新建商品住房和存量住房采用不同的指数编制模型。

➤ 新建商品住房同质价格指数

新建商品住房同质价格指数采用含栅格固定效应的特征价格法，其核心思想是：把异质性的住宅分解为众多同质性的特征，不同特征在各自的隐含市场上各有其特征价格，这些特征价格的集合构成了异质性商品的价格。

特征价格法具有多种形式，CQCHPI 指数的编制主要采用时间哑元法：在多元线性回归模型中引入时间虚拟变量，在控制住宅的区位属性和物理属性等异质性特征后，时间虚拟变量的系数估计值反映供求关系决定的市场价格的整体走势。本白皮书中所提到的特征价格法均指的是时间哑元法。

特征价格法计算价格指数的方程形式如下：

$$\log(HP_i) = X_i\alpha + \sum_t \gamma_t T_{it} + \varepsilon_i$$

其中， HP 代表住房价格；向量 X 表示特征价格属性变量，用于控制住房的区位属性和物理属性（这里假设住房属性的系数不随时间发生变化）；向量 T 表示时间虚拟变量，每个住房交易样本在该期成交取 1，否则取 0， β_t 为第 t 期相应的时间虚拟变量系数；标准误差项用 ε_i 表示。

特征价格法往往存在缺失变量问题，即实际数据中的某些特征缺失，使得指数编制结果出现偏误。尽管如此，在控制了足够多的区位属性和项目属性之后，特征价格法仍然远比简单的平均价格指数或加权平均价格指数等具有更好的同

质性。新建商品住房数据库中有大量控制项目属性的变量，包括住房面积、小区规模、容积率、绿化率、装修状况，可以认为，不同新建商品住宅项目间的项目属性差异得到了很好的控制。

为更好的控制区位属性，我们在特征价格模型中引入了栅格虚拟变量。 $2\text{km} \times 2\text{km}$ 的栅格虚拟变量的引入，剔除了由于空间位置差异（样本位于不同栅格）而带来的住房价格差异，从而大大提高了特征价格法计算的住房价格指数的准确性。修正后的特征价格模型如下：

$$\log(HP_i) = \sum_j \beta_j D_{ij} + X_i \alpha + \sum_t \gamma_t T_{it} + \varepsilon_i$$

其中， D 代表栅格虚拟变量，用于控制区位属性。

此外，考虑到不同住宅项目在各期成交的套数不同，在回归中使用加权最小二乘回归的方式对参数进行估计，以项目当期成交套数作为权重。

➤ 存量住房同质价格指数

存量住房同质价格指数采用混合模型法。由于存量住房数据是微观交易数据，所以存量住房同质价格指数有条件采用混合模型法，更有效的剔除不同样本（住房单元）之间的质量差异。

与特征价格法相比，重复交易法可以很好的解决“缺失变量”问题，其核心思想是：根据同一住宅在不同时期交易价格的差异来计算价格指数，在不考虑折旧的影响时，根据重复交易法编制的指数可以完全满足住宅价格“同质性”需求。但是由于中国住房市场中短期内多次重复交易样本较少，抽样数据具有严重的样本选择偏误，特别是对于新建商品住宅不存在重复交易问题，因此难以使用重复交易法。

混合模型法（类重复交易模型）结合了重复交易法和特征价格法的优势，选取项目（即住宅小区）作为住房交易样本的空间匹配单元，通过差分将项目以外的不可观测或者较难观测的特征属性消除，对余下少量方便观测的特征属性加以控制。“混合模型法”既避免了特征价格模型的缺失变量问题，又解决了重复交易法样本量过少的问题。

一种精确但是相对复杂的类重复交易模型请参见论文：[Xiaoyang Guo, Siqi Zheng, David Geltner and Hongyu Liu. A New Approach for Constructing Home Price Indices: The Pseudo Repeat Sales Model and Its Application in China. *Journal of Housing Economics*, 2014, 25: 20-38.](#)

为了保证大样本情况下指数计算的效率，在存量住房同质价格指数计算时，我们采用了一种简化的类重复交易模型：利用特征价格方法的基本模型框架，引入项目虚拟变量，控制项目层面的特征属性，实际等效于选取项目作为住房交易样本的空间匹配单元。修正后的特征价格模型如下：

$$\log(HP_i) = \sum_j \beta_j D_{ij} + X_i \alpha + \sum_t \gamma_t T_{it} + \varepsilon_i$$

其中， D 代表项目虚拟变量， HP 代表住房价格；向量 X 表示特征价格属性变量，用于现有数据可以控制项目内住房属性差异（这里假设住房属性的系数不随时间发生变化）；向量 T 表示时间虚拟变量，每个住房交易样本在该期成交取 1，否则取 0， β_t 为第 t 期相应的时间虚拟变量的系数；标准误差项用 ε_i 表示。

3.1.2 细分指数编制方法

CQCHPI 指数共编制三种细分指数产品，包括中心城区价格指数、区位于市场指数和面积子市场指数，其中新建商品住房同质价格指数针对两个细分指数编制：中心城区价格指数和区位于市场指数，而存量住房同质价格指数针对全部三个细分指数编制：中心城区价格指数、区位于市场指数和面积子市场指数。

细分指数产品的编制根据新建商品住房和存量住房的区别采用特征价格法和混合模型法，在此基础上，由于编制对象和目标的差异，三种细分指数产品的编制方法略有区别。

➤ 中心城区指数

中心城区价格指数的编制对象是位于中心城区的住房市场，在应用统计模型计算指数时，需筛选出位于中心城区的样本进入模型，具体的筛选标准见 1.4.1。

➤ 区位于市场指数

区位子市场指数的思想来源于局部加权回归，“局部”表明并非全样本，而是将局部的样本点引入到模型中；“加权”表明样本并非等权重进入模型，而存在着权重的变化。对于区位子市场指数，采用距离加权的形成，即位于中心位置的样本以 100%的权重进入回归，随着样本到子市场中心的距离增加，样本在回归中的权重逐步衰减至 0，具体的衰减速率由权重函数所决定。

在加权回归的参数设定中，为了保证指数模型中有充分的样本点，在计算时选取 30%样本比例作为回归“带宽”（bandwidth）。增大“带宽”使得回归中放入了更多的样本点，有利于提高指数的平滑性；但另一方面，减小带宽，即放入更少（局部）的指数样本进入回归，意味着指数越能反映子市场中心点的住房价格水平。在现有样本条件下，不同带宽分区位同质价格指数的测算表明，在带宽达到 30%样本时能够保证各区位子市场指数的平稳性。

参考国外学者的研究（McMillen, 2003），考虑到权重的平滑衰减，选取“三三”权重衰减函数，如下：

$$w_i = (1 - (d_i / d_b)^3)^3$$

其中， w_i 表示每个样本点在回归中的权重， d_i 为样本点到回归中心的距离， d_b 为样本点的最远距离。由于新建商品住房和存量住房的所有样本都是空间化的，相应的距离可以通过计算得到。

可以看到，通过局部加权回归的方法，我们可以针对空间任意“热点”构造价格指数。指数使用者如有需求可联系项目组进行测算。

➤ 面积子市场指数

面积子市场指数的编制对象是城市中住房面积不同的子市场。由于新建商品住房数据是项目层面数据，无单套面积数据，因此仅针对存量住房进行编制。

面积子市场指数将住房市场按单套住房面积分为 5 个区间，分别为 70 平方米以下，70-90 平方米，90-120 平方米，120-140 平方米，140 平方米以上，具体的划分依据见 1.4.3。在编制面积子市场指数时，需要将城市样本按面积区间筛选，针对每个面积区间的样本应用混合模型法进行计算。

3.2 变量设定

3.2.1 新建商品住房指数变量设定

新建商品住房同质价格指数包括中心城区价格指数和区位于市场指数，模型中包含的变量、权重及变量形式见下表。

新建商品住房同质价格指数变量表

变量类型	变量名	变量形式	中心城区价格指数	区位于市场指数
因变量	项目成交均价	对数	√	√
待估参数	成交季度虚拟变量	虚拟变量	√	√
控制变量	小区规模	对数	√	√
	该月度成交面积均值	对数	√	√
	容积率	—	√	√
	绿化率	—	√	√
	装修状况	虚拟变量	√	√
	栅格虚拟变量	虚拟变量	√	√
回归权重	项目当期成交套数	—	√	√
	权重衰减函数	函数	×	√

注：变量形式中“—”表示未对变量进行任何处理；“√”表示该指数模型中包含该变量或权重，“×”表示不包含。

3.2.2 存量住房指数变量设定

存量住房同质价格指数包括中心城区价格指数、区位于市场指数和面积子市场指数，模型中包含的变量、权重及变量形式见下表。

存量住房同质价格指数变量表

变量类型	变量名	变量形式	中心城区	区位于市场	面积子市场
因变量	卖方报价	对数	√	√	√
待估参数	成交季度虚拟变量	虚拟变量	√	√	√
控制变量	住房面积	对数	√	√	√
	所在楼层	对数	√	√	√
	项目虚拟变量	虚拟变量	√	√	√
权重	权重衰减函数	函数	×	√	×

注：变量形式中“—”表示未对变量进行任何处理；“√”表示该指数模型中包含该变量或权重，“×”表示不包含。

3.3 指数计算

在计算指数时，每个城市单独建立特征价格模型或混合模型，并采用加权最小二乘法或普通最小二乘法进行估计。模型形式如下式所示：

$$\log HP_i = \sum_j \beta_j D_{ij} + X_i \alpha + \sum_t \gamma_t T_{it} + \varepsilon_i$$

其中： HP 代表住房价格；对于新建商品住房同质价格指数，向量 D 表示价格虚拟变量，对于存量住房同质价格指数，向量 D 表示项目虚拟变量；向量 X 表示用于控制住房交易样本特征属性的变量（这里假设住房属性的系数不随时间发生变化）；向量 T 表示月份虚拟变量，每个住房交易样本在该月成交取 1，否则取 0， β_t 为第 t 月相应的时间虚拟变量系数；标准误差项用 ε_i 表示。

最终第 t 月的对应的定基比指数 $index_t$ 由下式中时间虚拟变量 β_t 的系数计算得到：

$$index_t = 100 \cdot \exp(\beta_t)$$

3.4 指数更新

对于指数更新，主要有两点需要解释和声明，以便于指数受众更好地理解和使用。

- 指数每一期更新都会利用特征价格法和混合模型法进行参数估计。采用五年期作为指数计算模型中的样本时序长度。例如，在更新 2014 年第 3 季度的指数时，用于构建模型的样本时间范围为 2009 年第 3 季度至 2014 年第 3 季度。
- 上述模型可给出更新使用的定基比价格指数序列。根据该更新使用的定基比指数序列计算当期相对于上一期的环比增长率，并利用该环比增长率在历史指数基础上计算当期定基比指数。
- 为了避免误差积累导致指数准确性降低，我们会每 5 年对指数进行一次修正。

4 参考文献

- [1] 程亚鹏. 基于住房子市场的特征价格住房价格指数[J]. 经济问题, 2010 (10): 66-68.
- [2] 李文斌, 杨春志, 史红, 等. 基于子市场的住房价格模型研究[J]. 城市问题, 2010 (10): 61-66.
- [3] 刘洪玉, 吴璟. 中国城市商品住宅价格的统计方法与变化规律[J]. 中国市场, 2012 (37): 75-82.
- [4] 吴璟. 中国城市住房价格短期波动规律研究[D]. , 2009.
- [5] 吴璟, 刘洪玉, 马亚男. 住房价格指数的主要编制方法及其选择[J]. 建筑经济, 2007 (7): 27-30.
- [6] 吴璟, 郑思齐, 刘洪玉. 编制住房价格指数的特征价格法细解[J]. 统计与决策, 2007, 24: 16-18.
- [7] 郑思齐, 孔鹏, 郭晓旸. 类重复交易房价指数编制方法与应用[J]. 统计研究, 2013, 30(12): 41-47.
- [8] Bourassa S C, Hamelink F, Hoesli M, et al. Defining housing submarkets[J]. Journal of Housing Economics, 1999, 8(2): 160-183.
- [9] Case B, Pollakowski H O, Wachter S M. On choosing among house price index methodologies[J]. Real estate economics, 1991, 19(3): 286-307.
- [10] Coulson N E, McMillen D P. The dynamics of intra urban quantile house price indexes[J]. Urban Studies, 2007, 44(8): 1517-1537.
- [11] Guo X, Zheng S, Geltner D, et al. A New Approach for Constructing Home Price Indices in China: The Pseudo Repeat Sales Model[J]. Available at SSRN 2127336, 2012.
- [12] Knerr W, Gardner F, Cluver L. Improving positive parenting skills and reducing harsh and abusive parenting in low-and middle-income countries: A systematic review[J]. Prevention science, 2013, 14(4): 352-363.
- [13] Liao W C, Wang X. Hedonic house prices and spatial quantile regression[J]. Journal of Housing Economics, 2012, 21(1): 16-27.
- [14] Maclellan D, Tu Y. Economic perspectives on the structure of local housing systems[J]. Housing studies, 1996, 11(3): 387-406.
- [15] McMillen D P. Neighborhood house price indexes in Chicago: a Fourier repeat sales approach[J]. Journal of Economic Geography, 2003, 3(1): 57-73.
- [16] McMillen D P. Changes in the distribution of house prices over time: Structural characteristics, neighborhood, or coefficients?[J]. Journal of Urban Economics, 2008, 64(3): 573-589.
- [17] Michaels R G, Smith V K. Market segmentation valuing amenities with hedonic models: the case of hazardous waste sites[J]. Journal of Urban Economics 1990, 28(2): 223-242.
- [18] Palm R. Spatial segmentation of the urban housing market[J]. Economic Geography, 1978: 210-221.
- [19] Quigley, John M. A simple hybrid model for estimating real estate price indexes[J]. Journal of Housing Economics, 1995, 4(1): 1-12.
- [20] Rothenberg, Jerome. The maze of urban housing markets: Theory, evidence, and policy[M]. University of Chicago Press, 1991.
- [21] Straszheim M. Hedonic estimation of housing market prices: A further comment. The Review of Economics and Statistics. 1970, 56(3): 404-406.
- [22] Tu Y. Segmentation, adjustment and disequilibrium[J]. Housing economics and public policy, 2003: 38-55.
- [23] Zheng S, Sun W, Wang R. Land supply and capitalization of public goods in housing prices: evidence from beijing[J]. Journal of Regional Science, 2013.

5 附件

附表 1：中心城区范围划定

城市名称	中心城区范围
北京	内城 6 区（西城区、东城区、海淀区、朝阳区、丰台区、石景山区）
上海	包含外环以内 9 区（黄浦区、卢湾区、徐汇区、长宁区、静安区、普陀区、闸北区、虹口区、杨浦区）以及浦东新区在外环以内的部分
天津	包含内城 6 区（和平区、河西区、南开区、河东区、河北区、红桥区）以及环城 4 区（北辰区、东丽区、津南区、西青区）在外环以内的部分
深圳	包含关内 4 区（福田区、罗湖区、南山区、盐田区）以及关外 2 区（宝安区、龙岗区）的核心区域
成都	包含内城 5 区（成华区、金牛区、青羊区、武侯区、锦江区）和周边 3 区 1 县（龙泉驿区、高新区、高新西区、郫县）
大连	包含旧城 4 区（西岗区、中山区、沙河口区、甘井子区）和旅顺口区
武汉	包含中环以内 7 个中心区（江岸区、江汉区、硚口区、汉阳区、武昌区、青山区、洪山区）以及 3 个新城（东西湖区、江夏区、蔡甸区）位于中环以内的部分
西安	包含内城 5 区（城内、城东区、城南区、城西区、城北区）以及高新区和长安区

附表 2：区位于市场中心的选取

城市名称	区位于市场中心
北京	中关村、亚运村、国贸、金融街
上海	人民广场、徐家汇、陆家嘴、新客站、五角场、中山公园
天津	友谊路、中山路、滨江道、滨海新区
深圳	景田、后海、宝安中心区、东门-翠竹
成都	高新南区、春熙路、建设路口
大连	青泥洼、华南、高新园区-西安路
武汉	白沙洲、光谷、街道口、江汉路
西安	小寨、钟楼、曲江新区

附表 3:

指数顾问团队（Advisory Group，按姓氏拼音排序）

Karl Case, 教授, 美国 Case-Shiller 房地产价格指数创始人

邓永恒, 讲座教授, 新加坡国立大学房地产研究院院长

David Geltner, 教授, 美国麻省理工学院 (MIT) 房地产研究中心主任, 穆迪 RCA CPPI 商业地产指数设计者

Gregory K. Ingram, 博士, 美国林肯土地政策研究院前总裁

Matthew Kahn, 教授, 美国加州大学洛杉矶分校 (UCLA) Ziman 房地产研究中心研究主任, NBER 研究员

李娜, 深圳市世联土地房地产评估有限公司, 董事总经理

刘洪玉, 教授, 清华大学房地产研究所所长, 清华大学恒隆房地产研究中心理事会常务副主席

刘建伟, 高级统计师, 国家统计局城市司副司长

刘志, 博士, 北京大学-林肯研究院城市发展与土地政策研究中心主任

满燕云, 教授, 林肯土地政策研究院资深研究员, 北京大学-林肯研究院城市发展与土地政策研究中心前主任, 美国印第安纳大学公共及环境事务学院教授

Daniel McMillen, 教授, 伊利诺伊大学厄巴纳-香槟分校; Regional Science and Urban Economics 杂志主编

指数工作团队（Working Group）

郑思齐, 教授, 清华大学恒隆房地产研究中心, 主任

孙伟增, 清华大学恒隆房地产研究中心, 博士生

满燕云, 教授, 林肯土地政策研究院资深研究员, 北京大学-林肯研究院城市发展与土地政策研究中心前主任, 美国印第安纳大学公共及环境事务学院教授

吴璟, 副教授, 清华大学恒隆房地产研究中心

孔鹏, 高级统计师, 国家统计局城市司房地产价格处, 处长

张博, 清华大学恒隆房地产研究中心, 主任助理

胡晓珂, 清华大学恒隆房地产研究中心, 博士生

张索迪, 清华大学恒隆房地产研究中心, 硕士生

孙聪, 清华大学恒隆房地产研究中心, 博士生

夏仕文, 深圳市世联土地房地产评估有限公司

朱焱东, 深圳市世联土地房地产评估有限公司