

基于潜在因子响应结构驱动的中国城市 二手房价异质性机制研究

石豪灿

华南农业大学数学与信息学院, 广东 广州

收稿日期: 2026年3月9日; 录用日期: 2026年3月29日; 发布日期: 2026年4月14日

摘要

中国城市房价的异质性波动是宏观政策精准调控面临的核心挑战。本文构建了一个“因子提取-结构分析-驱动聚类”的三阶段分析框架,旨在从城市对共性驱动力的差异化响应模式出发,识别由市场内在传导机制定义的城市类型。首先,运用高维时间序列潜在因子模型,从2015~2023年中国70个大中城市二手房价格收益率中提取三个公共因子:全国金融周期因子、区域发展因子和政策情绪因子。其次,基于先验分类的载荷结构分析表明,城市层级、地理区域和人口规模均能有效区分城市对金融周期的敏感度梯度,但对区域发展和政策情绪因子的捕捉存在局限。最后,以因子载荷向量为特征进行无监督聚类,识别出三类由数据内生驱动的城市集群:金融周期敏感型(11城)、政策情绪敏感型(20城)和增长依赖型(39城)。研究发现,传统分类与数据结构分类存在“同形异质”与“异形同质”的双重关系——层级相同的城市可能分属不同集群,而层级不同的城市可能呈现相似的驱动逻辑。本文的方法论贡献在于将因子模型从降维工具拓展为结构发现引擎,实现了从“外部标签”到“内生特征”的范式跃升;实践意义在于为“因城施策”提供了直指传导机制根源的分类依据。

关键词

房价异质性, 高维时间序列因子模型, 因子载荷, 城市分类, 因城施策

Heterogeneous Mechanisms of Second-Hand Housing Prices in Chinese Cities: A Study Driven by Latent Factor Response Structures

Haocan Shi

College of Mathematics and Informatics, South China Agricultural University, Guangzhou Guangdong

Received: March 9, 2026; accepted: March 29, 2026; published: April 14, 2026

Abstract

The heterogeneous fluctuations of urban housing prices in China pose a fundamental challenge to precise macro-policy regulation. This paper develops a three-stage analytical framework—“factor extraction, structural analysis, and driver-based clustering”—to identify city types endogenously defined by market transmission mechanisms based on their differentiated response patterns to common drivers. First, employing a high-dimensional time series latent factor model, we extract three common factors from second-hand housing price return series of 70 major Chinese cities from 2015 to 2023: a national macro-cycle factor, a regional development factor, and a policy sentiment factor. Second, prior-classification-based loading structure analysis reveals that administrative hierarchy, geographic region, and population size effectively differentiate cities' sensitivity gradients to the macro-cycle factor but exhibit limitations in capturing responses to regional development and policy sentiment factors. Finally, unsupervised clustering based on factor loading vectors identifies three data-endogenously driven city clusters: financial-cycle-sensitive (11 cities), policy-sentiment-sensitive (20 cities), and growth-dependent (39 cities). The findings reveal dual relationships between traditional and data-driven classifications: “same form, different nature”—cities with identical administrative tiers may belong to different clusters; and “different form, same nature”—cities across tiers may exhibit similar driving logics. The methodological contribution lies in extending the factor model from a dimension-reduction tool to a structure-discovery engine, achieving a paradigm shift from “external labels” to “endogenous characteristics.” The practical implication provides classification foundations for targeted urban housing policies directly linked to transmission mechanisms.

Keywords

Housing Price Heterogeneity, High-Dimensional Time Series Factor Model, Factor Loading, City Classification, Targeted Urban Housing Policies

Copyright © 2026 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

1.1. 研究背景与问题提出

中国城市房价的波动特征历来是学术界与政策制定者关注的焦点。在探究中国城市房价的决定因素时，一个基本的分析框架已然确立：房价变动是宏观经济基本面及全国性政策等共性力量，与城市自身结构性特征共同作用的结果。然而，这一框架的核心挑战在于，如何精确剥离并度量两者的交互效应——相同的宏观政策或经济冲击，为何在不同城市产生迥异的效果？这正是中国城市房价研究的核心困境所在。

关于房价驱动因素的讨论，学界形成了两条清晰的研究脉络。一条脉络聚焦于经济基本面因素。Yu(2010) [1]系统考察了中国房价的影响因素，发现收入增长、人口结构、建造成本等基础性变量对房价变动具有显著的解释力，这为理解房价的长期趋势提供了理论基础。另一条脉络则强调政策因素的驱动作用。Zhou 等(2020) [2]的研究表明，全国性房地产政策、货币信贷调控等宏观政策冲击，对城市房价波动产生了不可忽视的影响。这两条研究脉络共同揭示了一个基本事实：城市房价是基本面因素与政策因

素交织作用的产物。

然而，一个更为棘手的问题随之浮现：相同的宏观政策或经济冲击，为何在不同城市产生迥异的效果？对此，学界从多个维度展开探索。一部分研究聚焦于人口流动与城市化进程的解释路径。Chen 等(2011) [3]基于 1995~2005 年的数据，系统论证了人口迁移和城市化进程对中国城市房价动态的关键塑造作用。研究隐含地表明，一线城市因其优质的就业机会与公共服务吸引持续的人口流入，而三四线城市则面临人口外流与住房需求疲软的双重压力，这种“人地钱”的空间错配构成了房价分化的微观基础。

另一部分研究则关注金融因素的异质性传导机制。Meng 等(2015) [4]通过俱乐部收敛分析发现，中国十大重点城市的房价并非同质运行，而是呈现出显著的收敛俱乐部特征——一部分城市在长期中趋向于共同的增长路径，而另一些城市则分化出不同的运行轨迹。这一发现为理解房价异质性提供了动态演进的视角。Huang 等(2021) [5]进一步将研究视野拓展至货币政策与跨境资本流动，证实货币政策冲击与“热钱”流动对房价的影响在不同城市间存在显著的系统性差异。Katagiri (2018) [6]从跨国比较的视角，揭示了金融开放程度对房价联动模式的影响机制，为理解中国城市房价的差异化响应提供了更具一般性的理论参照。

近年来，国内学者在这一领域也取得了重要进展。王一迪等(2024) [7]从微观行为视角切入，发现房价黏性与不确定预期共同塑造了住房需求的波动模式，而这种影响在不同城市间亦呈现系统性差异。蓝天(2022) [8]则运用 TVP-SV-VAR 模型，从经济政策不确定性与金融周期的互动视角，揭示了宏观政策冲击在不同市场环境下的时变传导特征，为理解政策效应的异质性提供了方法论支撑。

上述研究从不同角度深化了我们对房价异质性的认识，但也共同指向一个方法论困境：如何在有效测度全国性共性驱动力的同时，系统揭示不同城市对同一驱动力响应强度的结构性差异？传统研究路径多依赖于先验的城市分类——或基于城市层级，或基于地理区位，或基于人口规模——将这些外部标签作为分析异质性的分组依据。近年来，无监督机器学习方法(如 K-means、层次聚类)被广泛应用于根据多维度经济特征识别城市类型，但这类研究普遍致力于寻找“最优”的单一分类，却忽视了分类本身应服务于特定的理论问题；详见 Wang 等(2022) [9]的研究。更重要的是，多数分类研究与后续的机制分析脱节，所选用的分类变量(如人均 GDP、人口规模)与所要解释的核心经济变量(房价波动)之间缺乏直接的、可解释的传导逻辑，导致“为分类而分类”。然而，Meng 等(2015) [4]的研究已经提示我们：城市房价的内在结构可能存在着由市场传导机制自身定义的、超越传统地理或行政标签的“自然集群”。这意味着，我们需要的不是“一般的”城市分类，而是能够反映房价驱动机制差异的“功能型”分类。

综上所述，现有研究在以下问题上仍存在探索空间：如何突破先验分类的局限，直接从城市对共性驱动力的响应模式出发，识别由市场内在传导机制定义的城市类型？这正是本文试图回答的核心问题。通过构建“因子提取 - 结构分析 - 驱动聚类”的三阶段分析框架，本研究旨在为“因城施策”提供直指机制根源的分类依据。

1.2. 高维时间序列因子模型：方法论契机

正是在这一背景下，高维时间序列因子模型展现出独特的方法论价值。该类模型的核心思想在于：尽管观测到的变量维度(城市数量)可能很高，但其共同变动往往由少数几个不可观测的潜在公共因子所驱动。通过引入因子结构，可以将高维问题转化为低维问题，在有效降维的同时保留数据中的关键信息。这一思想可追溯至经典因子分析，而针对时间序列数据的动态因子模型则进一步拓展了其应用边界。

在众多发展中，Lam 等(2011) [10]提出的高维时间序列潜在因子模型具有里程碑意义。该方法创新性地利用多个非零滞后阶的自协方差矩阵来估计因子载荷与因子序列，其核心优势体现在三个方面：第一，它明确地将具有序列相关性的动态因子与特质性白噪声分离，模型在有限样本下即可识别，无需依赖横

截面维度趋于无穷的渐近假设；第二，理论证明，在存在“强因子”（即每个因子对相当比例的个体有显著影响）时，因子载荷矩阵的估计误差收敛速率独立于横截面维度 N ，即享有“维度的祝福”（ *blessing of dimensionality*）——更多的观测个体不仅不是负担，反而因提供了更多关于共性因子的信息而提高了估计精度；第三，该方法对噪声项的横截面相关性要求较为宽松，具有良好的稳健性。

将这一方法应用于中国城市房价研究，具有天然的理论契合性。70 个大中城市的房价收益率序列，恰好构成了一个 $N=70$ 、 $T \approx 108$ 的高维时间面板。这一维度结构使得 Lam 等(2011)模型“维度的祝福”性质得以充分发挥——我们可以在不依赖传统大样本渐近理论的情况下，从高维数据中有效提取出驱动全国市场同步波动的共性力量。

1.3. 核心挑战：潜在因子的可解释性问题

然而，因子模型的应用并非止步于因子提取。一个根本性挑战随之而来：提取出的潜在因子不可观测，其经济含义往往模糊。在典型应用中，研究者通常仅通过因子与宏观变量的相关性进行事后命名，如将与货币供给增速高度相关的因子称为“货币因子”。但这种命名存在局限：其一，无法回答该因子在不同类型城市间的作用强度差异；其二，无法解释同一因子在不同城市为何产生迥异的传导效果。

这一“可测不可释”的困境，限制了因子模型在政策分析和机制研究中的深度应用。就本文关注的房价问题而言，识别出“存在共同的金融因子”固然重要，但更关键的问题是：这一因子对哪些城市影响最大？传导路径是否存在结构性差异？如何据此设计差异化调控政策？这些问题指向更深层的需求——将抽象因子与城市异质性结构进行关联，实现对因子作用机制的“具象化”解读。

1.4. 本文的研究思路与核心创新

基于上述认识，本研究认为，破解潜在因子可解释性困境的关键，在于将因子载荷矩阵（即各城市对各公共因子的响应强度）本身视为蕴含丰富结构性信息的核心数据对象。因子载荷 $a_{i,j}$ 量化了城市 i 对因子 j 的暴露程度或敏感度，它是由城市自身的经济结构、金融深度、人口特征等内在属性所决定的“响应指纹”。因此，通过系统分析载荷矩阵的结构，我们可以逆向推断出驱动房价异质性传导的深层逻辑。

基于这一思想，本文构建了一个递进式的“因子提取 - 结构分析 - 驱动聚类”三阶段分析框架：

第一阶段：因子提取。运用高维时间序列潜在因子模型，从 70 个城市的房价收益率面板数据中提取出关键的公共因子及其核心产出——因子载荷矩阵。这一步实现了对共性驱动力的有效测度与降维。

第二阶段：结构分析。为深入理解载荷矩阵的结构信息，我们首先将其嵌入城市层级、地理区域、人口规模等多种先验分类方案中进行比较分析。通过方差分析(ANOVA)和组间比较，检验传统分类在捕捉因子异质性方面的有效性，并初步揭示因子影响的系统性差异。这一阶段的作用是“诊断”——评估既有认知框架的局限，并为后续的数据驱动分类提供参照基准。

第三阶段：驱动聚类。这是本研究的核心创新环节。我们摒弃依赖外部标签进行分类的传统路径，转而以各城市对核心公共因子的载荷向量作为其特征画像，进行无监督聚类分析。由此产生的城市集群，完全由数据内在的“因子 - 城市”响应结构所驱动，其划分依据是城市对共性驱动力的真实响应模式，而非任何先验假设。

1.5. 理论贡献与实践意义

本研究的理论贡献体现在三个层面：

第一，范式创新。确立了“先测度共性响应，再依据响应模式定义结构”的分析新路径，将因子模型从降维工具拓展为结构发现引擎，实现了从机制测量到结构定义的逻辑闭环，真正实现了从“提取共性”到“解释异质”的方法论跃升。

第二，方法创新。首次系统地将 Lam 等(2011)高维时间序列潜在因子模型与基于因子载荷的聚类分析相结合，开创了以“对共性驱动力响应模式”为基准的数据结构驱动型城市分类方法。

第三，洞察创新。通过该方法，本研究能够识别出如“金融周期敏感型”“政策情绪敏感型”等由市场传导机制内生定义的城市集群，深化了对货币政策区域异质性效应的微观基础理解，为“因城施策”提供了超越传统标签、直指风险根源的决策依据。

1.6. 文章结构安排

本文后续章节安排如下：第二章详细阐述“因子提取 - 结构分析 - 驱动聚类”的三阶段方法论框架；第三章报告实证分析结果，包括因子提取、载荷结构分析与数据结构驱动聚类发现；第四章深入讨论异质性传导机制及其政策含义；第五章总结全文并提出研究展望。

2. 研究设计与方法论框架

2.1. 理论框架与核心命题

本研究的核心思想是：城市房价对共性驱动力的差异化响应模式(即因子载荷)，是刻画其市场本质和风险属性的最有效特征。基于此，提出以下命题：

P1：存在少数几个潜在的、动态的公共因子(如全国金融周期因子、区域增长动能因子)，共同驱动着中国主要城市的房价波动。

P2：不同城市对同一公共因子的响应强度(因子载荷)存在显著且系统的差异，这种差异反映了城市在经济结构、金融深度、政策环境等方面的异质性，构成了房价异质性传导的微观基础。

P3：以城市对核心公共因子的载荷向量为特征进行聚类，所得到的城市集群，比任何基于先验规则或一般经济指标的分类，都更能揭示房价波动的内在动力结构和风险同质性，因而具有更强的经济学解释力和政策相关性。

2.2. 数据体系与预处理

样本与数据来源：本研究选取 2015 年 1 月至 2023 年 12 月中国 70 个大中城市的月度面板数据。房价数据(二手房价格指数)来源于国家统计局，宏观经济与城市特征数据来源于 Wind 数据库、CEIC 数据库及各城市统计年鉴。

核心变量构造：

被解释变量：各城市月度二手房价格对数收益率，构成 70 维时间序列。

先验分类变量：作为第二阶段结构分析的对比基准，包括行政与房地产市场层级(核心/中心/三线/四线)、地理区域(东部沿海/中部/西部/东北)和人口规模(超大城市/特大城市/I 型大城市/II 型大城市)。

宏观标识变量：用于辅助解释提取的潜在因子，包括 M2 增速、社会融资规模、国房景气指数、制造业 PMI 等。

2.3. 共性潜在因子提取

基于 Lam 等(2011)的方法，观测序列 y_t 估计如下模型：

$$y_t = Ax_t + \varepsilon_t,$$

A 是因子载荷矩阵， x_t 是潜在因子序列， ε_t 为特质噪声。

其中，A 为因子载荷矩阵， x_t 为不可观测的潜在因子序列， ε_t 为特质噪声。模型旨在通过少数公共因子驱动高维数据的共同波动，实现降维。

估计采用 Lam 等(2011)基于自协方差矩阵的特征分析方法, 利用因子序列相关性与噪声无自相关的性质。首先计算滞后 k 的样本自协方差矩阵 $\Sigma_y(k)$, 进而构建信息矩阵 $M = \sum_{k=1}^{k_0} \Sigma_y(k) \Sigma_y(k)^T$ 。对 M 进行特征值分解, 采用特征值比率法确定因子个数 r , 详见 Lam 等(2012) [11]。并取前 r 个特征向量构成载荷矩阵估计 \hat{A} , 最后得到潜在因子序列估计 $x_t = \hat{A}x_t$ 。

2.4. 基于先验分类的载荷结构分析

在完成共性潜在因子提取并获得因子载荷矩阵 \hat{A} 后, 本阶段旨在系统探究城市异质性特征如何塑造其对公共因子的差异化响应模式。因子载荷 $\hat{a}_{i,j}$ (对应 \hat{A} 的 (i, j) 元素) 具有明确的经济含义——它量化了城市 i 对第 j 个公共因子的敏感度或暴露程度, 是城市自身结构性特征在因子空间中的投射。因此, 通过对载荷矩阵进行结构性分解, 可以揭示不同类型城市在房价驱动机制上的系统差异。

2.4.1. 先验分类体系的构建

为评估传统城市分类框架在捕捉因子异质性方面的表现, 本文构建了多维度的先验分类体系作为分析基准:

(1) 依据城市行政地位、经济辐射力及房地产市场成熟度, 将 70 个城市划分为四个层级: 核心城市(北京、上海、广州、深圳、厦门, 共 5 城)、中心城市(天津、南京、杭州、武汉、郑州、福州等, 共 16 城)、三线城市(共 32 个, 以省会城市及区域中心城市为主体)、四线城市(共 13 个, 多为经济规模较小或区位边缘的城市)。需要说明的是, 这一划分侧重于房地产市场成熟度与平均房价等指标, 与通常基于 GDP 等经济指标的分类有所区别。该分类与政府调控和房地产市场认知中常用的“一线/二线/三线”话语体系基本对应。

(2) 地理区域分类: 依据传统地理分区, 将城市划分为东部沿海、中部、西部和东北四大板块, 以捕捉区域经济发展水平、产业结构及政策倾斜等因素的潜在影响。

(3) 人口规模分类: 依据国务院发布的《关于调整城市规模划分标准的通知》, 以城区常住人口为基准, 将城市划分为超大城市(1000 万以上)、特大城市(500 万~1000 万)、I型大城市(300 万~500 万)和II型大城市(100 万~300 万), 以考察人口集聚程度对房价驱动机制的影响。

2.4.2. 组间载荷模式的比较分析

为直观呈现不同城市类别在因子响应强度上的差异, 本文计算各类别在三个公共因子上的平均载荷, 通过比较均值大小来揭示因子影响的层级梯度。具体而言, 对于每个先验分类维度, 分别计算:

$$\bar{a}_j^{(k)} = \frac{1}{n_k} \sum_{i \in \text{类别}k} \hat{a}_{ij}$$

其中 n_k 为第 k 类城市的数量, $\bar{a}_j^{(k)}$ 表示第 k 类城市在第 j 个因子上的平均敏感度。通过比较不同类别在同一因子上的均值大小, 可以识别因子影响的系统性差异; 通过观察同一类别在不同因子上的均值结构, 可以刻画该类城市的驱动模式特征。

2.5. 基于因子载荷的数据结构驱动聚类

本节旨在突破先验分类的外部标签局限, 直接从数据内在结构出发, 识别由城市对共性驱动力响应模式所定义的城市集群。

聚类分析: 以第一阶段估计得到的因子载荷矩阵 \hat{A} 作为输入, 该矩阵的每一行 $\hat{a}_i = (\hat{a}_{i1}, \hat{a}_{i2}, \dots, \hat{a}_{ir})$ 即为城市 i 在 r 个公共因子上的载荷向量, 量化了该城市对所有共性驱动力的“响应指纹”。基于此, 本文

以 $70 \times r$ 的特征向量矩阵为对象, 采用 K-means++ 算法进行无监督聚类。最优聚类数 k 通过轮廓系数 (Silhouette Coefficient) 与 Gap Statistic 综合确定, 以确保集群划分的稳健性与内在合理性。

集群的经济学解读与命名: 聚类完成后, 通过两步赋予其经济含义:

(1) 剖面分析: 计算各集群内城市在因子载荷上的均值向量, 得到“平均响应剖面”, 刻画集群整体的因子敏感结构。

(2) 集群命名: 依据最突出的因子载荷特征进行命名, 使集群名称直接反映房价波动的核心驱动逻辑。

3. 实证结果分析

3.1. 共性潜在因子提取结果

本文基于高维时间序列潜在因子模型, 从 70 个大中城市二手房价格对数收益率序列中提取公共因子。模型估计结果表明, 三个公共因子足以捕捉城市房价波动的核心共性信息。

因子贡献度分析: 三个因子的贡献度呈现明显的梯度分布(见表 1)。因子 F1 贡献度达 79.83%, 在解释城市房价共同波动中起绝对主导作用; 因子 F2 和 F3 分别为 9.13% 和 6.41%, 起辅助解释作用。三者累计方差贡献率 95.37%, 表明提取的因子能够有效刻画 70 个城市房价波动的共性成分, 模型拟合良好。

Table 1. Factor contribution

表 1. 因子贡献度

因子	F1	F2	F3	累计
贡献度(%)	79.83	9.13	6.41	95.37

3.2. 基于先验分类的因子载荷结构分析

为探究城市异质性特征如何塑造其对公共因子的差异化响应, 本文分别基于城市层级、地理区域和人口规模三种先验分类方案, 计算各类别城市在三个因子上的平均载荷。

3.2.1. 城市层级分类结果

依据城市行政地位与市场辐射力, 将 70 个城市划分为核心城市、中心城市、三线城市和四线城市四类, 各类别平均载荷计算结果见表 2。

Table 2. Average factor loadings by city hierarchy

表 2. 城市层级分类的因子平均载荷

类别	城市层级	因子 F1	因子 F2	因子 F3
1	核心城市	-0.1690	-0.1096	-0.2209
2	中心城市	-0.1577	-0.1050	0.0448
3	三线城市	-0.1032	0.0684	0.0349
4	四线城市	-0.0566	0.1277	0.0174

因子 F1 呈现严格梯度递减格局: 城市层级越高, 对全国金融周期越敏感, 表现为核心城市 > 中心城市 > 三线城市 > 四线城市。核心城市作为资金、信息与政策的集散地, 对宏观政策变动反应最为迅速剧烈; 四线城市市场相对封闭, 受全国性波动直接影响较小。

因子 F2 呈现明显等级分化: 高层级城市为负, 低层级城市为正。低层级城市房价对本地发展因素具

有更强的正向依赖，而高层级城市则受政策调控、供给约束等因素制约，削弱了区域发展因子的正向作用。

因子 F3 呈现核心城市特异性：仅核心城市显著为负，其他等级城市载荷接近零。这反映了核心城市作为房地产调控重点对象，面临更严格的限购限贷政策和紧张的土地供给。

3.2.2. 地理区域分类结果

根据传统地理分区，将 70 个城市划分为东部沿海、中部、西部和东北四大板块，各类别平均载荷计算结果见表 3。

Table 3. Average factor loadings by geographic region

表 3. 地理区域分类的因子平均载荷

类别	区域	因子 F1	因子 F2	因子 F3
1	东部沿海	-0.1306	-0.0363	0.0206
2	中部地区	-0.1144	0.0492	-0.0452
3	西部地区	-0.0810	0.1081	0.0392
4	东北地区	-0.0842	0.1113	0.0405

因子 F1 敏感度排序：东部(0.1306) > 中部(0.1144) > 东北(0.0842) \approx 西部(0.0810)。东部沿海地区对全国金融周期最为敏感，中部次之，西部和东北地区敏感度较低，呈现清晰的“东高西低”梯度格局。这一结果符合经济直觉：东部地区经济开放度高、金融深化程度强、与全国市场联动性更为紧密，因此对全国性宏观冲击的反应更为灵敏。

因子 F2 呈现明显的东西分化：东部为负(-0.0363)，中部、西部、东北均为正，且西部和东北地区载荷最高(0.11 左右)。这表明西部和东北地区房价对区域性发展因素(可能涉及产业转移、政策倾斜、基础设施建设等)具有更强的正向依赖，而东部地区受此类因素的影响相对较弱甚至为负向。

因子 F3：各区域差异不大，载荷值均在零附近，表明该因子未呈现显著的地理分区特征。

3.2.3. 人口规模分类结果

依据城区常住人口规模，将 70 个城市划分为超大城市、特大城市、I型大城市和II型大城市四类，各类别平均载荷计算结果见表 4。

Table 4. Average factor loadings by population size

表 4. 人口规模分类的因子平均载荷

类别	规模等级	因子 F1	因子 F2	因子 F3
1	超大城市	-0.1501	-0.0851	-0.1243
2	特大城市	-0.1365	0.0019	0.0058
3	I型大城市	-0.1346	0.0056	-0.0385
4	II型大城市	-0.0900	0.0672	0.0513

因子 F1 呈现明显规模梯度：超大城市(0.1501) > 特大城市(0.1365) > I型大城市(0.1346) > II型大城市(0.0900)。前三类城市敏感度差异不大，但II型大城市出现明显下降。这表明人口规模达到一定门槛后，城市对宏观政策的敏感度趋于稳定；规模较小的城市则因市场相对封闭，受金融周期影响较弱。

因子 F2 同样呈现规模梯度：超大城市为负(-0.0851)，其余规模城市为正，II型大城市载荷最高(0.0672)。中小城市房价更依赖本地发展因素，而超大城市受调控政策制约，削弱了区域发展因子的正向作用。

因子 F3 具有明显的规模特异性：仅超大城市显著为负(-0.1243)，其他规模城市载荷接近零。这捕捉到超大城市特有的政策调控与供给约束因素，独立于金融周期和区域发展，构成影响其房价的特异性力量。

3.3. 因子的异质性特征分析

综合上述分析，三种先验分类方案在揭示因子响应异质性方面呈现以下特征：

第一，因子 F1 敏感度梯度高度一致。城市层级、地理区域和人口规模三个维度均呈现“中心 - 边缘”递减格局：核心城市/东部/超大城市最敏感，四线城市/西部/小规模城市最不敏感。该因子与 M2 增速、社会融资规模高度相关，可命名为全国宏观周期因子。

第二，因子 F2 结构性差异更为丰富。三个分类均呈现高层级/高发展水平城市为负、低层级/低发展水平城市为正的的模式。该因子与工业增加值、居民收入增速高度相关，可命名为区域增长因子。

第三，因子 F3 主要捕捉核心城市/超大城市的特有因素。该因子仅在城市层级的核心城市和人口规模的超大城市上显著为负，在地理区域分类中无显著差异，可命名为政策情绪因子。

上述发现验证了传统分类框架的有效性，也揭示了其局限：不同分类维度在不同因子上呈现互补的区分能力，单一分类难以完整刻画城市房价驱动机制的复杂结构。这为后续数据驱动聚类分析提供了参照基准

3.4. 数据结构驱动聚类结果

前文基于先验分类的分析揭示了因子响应的系统性差异，但城市层级、地理区域等外部标签能否真实反映城市房价波动的内在驱动结构，仍有待检验。为此，本文进一步以各城市在三个因子上的载荷向量为特征，采用 K-means 聚类方法进行无监督学习，让数据自身“说话”，识别由因子响应模式内生定义的城市集群。

3.4.1. 聚类结果与集群特征

各集群的因子载荷均值及特征描述见表 5。

Table 5. Cluster characteristics based on data-driven clustering

表 5. 数据结构驱动聚类的集群特征

集群	集群特征	城市数量	因子 F1	因子 F2	因子 F3
集群 1	金融周期敏感型	11	-0.1750	-0.1127	-0.1615
集群 2	政策情绪敏感型	20	-0.1199	-0.0517	0.1228
集群 3	增长依赖型	39	-0.0884	0.1141	0.0038

集群 1 (金融周期敏感型)：包含北京、上海、广州、深圳、天津、南京等 11 个城市。该集群在因子 F1 上载荷绝对值最大(0.1750)，在 F2 和 F3 上也呈现较高负向载荷，表明其对全国金融周期最为敏感，同时对区域发展和政策情绪因素也有较强负向响应。这类城市主要为直辖市、一线城市及经济强省会，房地产市场与全国金融体系高度关联，是宏观政策传导的“前沿阵地”。

集群 2 (政策情绪敏感型)：包含杭州、宁波、成都、重庆、西安、长沙等 20 个城市。该集群最显著的特征是在因子 F3 上呈现较高正向载荷(0.1228)，因子 F1 载荷居中(0.1199)，因子 F2 为微弱负向，表明

房价对市场情绪、政策预期等非基本面因素反应更为灵敏，同时受金融周期影响较强。这些城市多为近年来房地产市场热度较高的强二线城市或区域中心城市。

集群 3 (增长依赖型): 包含 39 个城市，是规模最大的集群。该集群在因子 F2 上呈现显著正向载荷 (0.1141)，因子 F1 和 F3 载荷均较低，表明房价主要受本地经济增长、居民收入提升等基本面因素驱动，对全国金融周期和市场情绪敏感度较低。城市构成上，该集群涵盖中西部、东北地区的三线及四线城市，以及部分东部普通地级市，房地产市场相对独立，更依赖本地内生增长动力。

3.4.2. 集群与传统分类的对比分析

将聚类结果与传统分类进行交叉比对，有助于理解数据结构驱动分类的经济内涵(见表 6)。

Table 6. Cross-distribution of clustering results and traditional classifications

表 6. 聚类结果与传统分类的交叉分布

集群	特征	城市层级分布	地理区域分布	人口规模分布
集群 1	金融周期敏感型	核心城市(5)、 中心城市(5)、三线(1)	东部(8)、中部(3)	超大(5)、特大(3)、I型(3)
集群 2	政策情绪敏感型	中心城市(7)、 三线(11)、四线(2)	东部(13)、中部(3)、西部(4)	超大(2)、特大(4)、 I型(2)、II型(12)
集群 3	增长依赖型	中心城市(2)、 三线(23)、四线(14)	东部(10)、中部(10)、西部 (11)、东北(8)	特大(5)、I型(6)、II型(28)

集群 1 与先验分类的对应: 该集群囊括 5 个核心城市，以及天津、南京、合肥、郑州、武汉等 5 个中心城市，仅石家庄一个三线城市入围。地理上高度集中于东部沿海，人口规模以超大城市和特大城市为主。金融周期敏感型城市与传统认知中的“高层级城市”高度重合，房地产市场与宏观金融环境深度绑定。

集群 2 的独特定位: 该集群 20 个城市呈现跨层级特征，既有杭州、宁波等中心城市，也有青岛、长沙、重庆、成都等三线城市，甚至包括兰州、三亚等四线城市。地理分布以东部为主，向中西部扩散；人口规模多元，II型大城市占主导。这一分布揭示，在传统等级分类中处于不同层次的城市，在市场情绪层面可能表现出相似行为模式，形成了传统分类难以捕捉的“同质异级”特征。

集群 3 的广泛覆盖: 该集群规模最大，涵盖 39 个城市，占样本一半以上。构成以三线、四线城市为主体，仅济南、韶关两个中心城市入围。地理分布均衡，人口规模上II型大城市占绝大多数。这类城市远离金融周期中心，对宏观政策情绪传导相对钝化，房价波动更多由本地经济增长、人口流动等基本面因素决定。

3.4.3. 数据结构驱动分类的理论内涵

上述聚类结果具有重要的理论含义:

第一，因子响应模式是城市房价驱动机制的本质刻画。以因子载荷为特征的聚类，直接基于城市对金融周期、区域发展和政策情绪三类力量的敏感度进行划分，所得集群具有清晰的经济解释——金融周期敏感型、政策情绪敏感型、增长依赖型分别对应了不同的房价驱动逻辑。

第二，传统分类与数据结构分类存在“同形异质”与“异形同质”的双重关系。“同形异质”指传统同一类别的城市(如中心城市)可能分属不同集群，表明外部标签相同的城市内在驱动机制可能存在差异；“异形同质”指传统不同类别的城市(如部分三线城市与强二线城市)可能归于同一集群，表明其房价波动

具有相似的驱动逻辑。这解释了为何单一先验分类难以完整刻画房价异质性。

第三，识别出了传统分类无法揭示的新型城市群。尤其是“政策情绪敏感型”集群的发现，揭示了部分城市房价对市场预期和非理性因素的特殊敏感性，这类城市在宏观审慎管理中应予以特别关注。

综上，数据结构驱动聚类从因子载荷中提取了城市房价波动的“响应指纹”，实现了从“外部标签”到“内生特征”的范式跃升，为“因城施策”提供了更具针对性的分类依据。

4. 机制讨论、政策含义与研究展望

4.1. 异质性传导机制的综合阐释

本研究揭示了中国城市房价异质性传导的两层核心机制：

驱动力层面的二分：全国性金融冲击与区域性增长冲击，构成了房价波动的两个最主要、相对独立的共性来源。

城市响应层的三分：根据对上述驱动力响应模式的差异，城市分化为三类传导“受体”：金融传导主导型、增长传导主导型和本地化特质型。这种“受体”类型由城市的经济结构、金融深化程度和发展阶段内生决定，是理解“同策异效”的钥匙。

4.2. 对“因城施策”的精准启示

本研究揭示的城市房价驱动机制异质性，为“因城施策”提供了超越传统分类的决策依据。不同城市群城市对宏观政策的响应模式存在系统性差异，调控需从“一刀切”转向“因类施策”。

金融周期敏感型集群(11个，强经济城市)对金融周期因子响应最强，调控核心在于熨平金融冲击：宽松周期中率先启动逆周期调节工具，紧缩周期中建立动态预警机制。

政策情绪敏感型集群(20个，强二线城市)对政策情绪因子响应突出，房价易受预期放大：需保持政策连续性，加强信息发布与预期引导，抑制投机情绪蔓延。

增长依赖型集群(39个，三四线城市)对本地增长因子响应主导，对宏观政策相对钝化：着力点应从短期调控转向长期制度建设，推动产业升级夯实房价基础，培育内生增长动力。

政策协同上，需警惕“同形异质”(同级城市分属不同集群)与善用“异形同质”(不同级城市属同一集群)，实现区域政策协同。动态调控框架可建立“诊断-识别-配置”三阶机制，将城市分类从“静态标签”升级为“动态画像”，实现从“因城施策”到“因类施策”再到“因时施策”的跃升。

4.3. 局限与展望

局限：首先，研究基于历史数据，因子载荷被视为静态，未能捕捉其随时间的变化。其次，聚类分析本身对初始值和 k 值选择有一定敏感性。

展望：未来研究可沿以下方向拓展：

- (1) 采用滚动窗口或时变参数因子模型，动态追踪城市因子载荷与集群归属的演变，研究城市类型的“跃迁”。
- (2) 引入复杂网络分析，基于因子载荷的相似性构建城市关联网络，研究风险传染路径。
- (3) 将框架扩展至商业地产、租赁市场等不同业态，进行对比研究。

5. 结论

本研究通过构建“因子提取-结构分析-驱动聚类”的递进框架，成功地将抽象的房价共性驱动力与具体的城市异质性结构联系起来。研究发现，中国城市可以根据其对“金融周期”和“经济增长”两大核心驱动力的响应模式，清晰地划分为三个内在动力机制迥异的集群。这一由数据结构驱动产生的分类，

不仅比传统先验分类更具经济学解释力，更重要的是，它直接映射出差异化的风险传导渠道和政策敏感点，实现了学术发现与政策需求的精准对接。

研究最终表明，理解中国房地产市场分化的关键，不在于寻找更多的解释变量，而在于重新审视“城市”本身应如何被科学定义。从依据外在标签分类，转向依据其对内生驱动力的响应模式来分类，是迈向精准认知和精准治理的必由之路。本研究为此提供了一套完整、可操作的方法论与实证范例。

参考文献

- [1] Yu, H. (2010) China's House Price: Affected by Economic Fundamentals or Real Estate Policy? *Frontiers of Economics in China*, **5**, 25-51. <https://doi.org/10.1007/s11459-010-0002-7>
- [2] Zhou, Q., Shao, Q., Zhang, X. and Chen, J. (2020) Do Housing Prices Promote Total Factor Productivity? Evidence from Spatial Panel Data Models in Explaining the Mediating Role of Population Density. *Land Use Policy*, **91**, Article ID: 104410. <https://doi.org/10.1016/j.landusepol.2019.104410>
- [3] Chen, J., Guo, F. and Wu, Y. (2011) One Decade of Urban Housing Reform in China: Urban Housing Price Dynamics and the Role of Migration and Urbanization, 1995-2005. *Habitat International*, **35**, 1-8. <https://doi.org/10.1016/j.habitatint.2010.02.003>
- [4] Meng, H., Xie, W. and Zhou, W. (2015) Club Convergence of House Prices: Evidence from China's Ten Key Cities. *International Journal of Modern Physics B*, **29**, Article ID: 1550181. <https://doi.org/10.1142/s0217979215501817>
- [5] Huang, X., Jin, T. and Zhang, J. (2021) Monetary Policy, Hot Money and Housing Price Growth across Chinese Cities. *Applied Economics*, **53**, 6855-6877. <https://doi.org/10.1080/00036846.2021.1949433>
- [6] Katagiri, M. (2018) House Price Synchronization and Financial Openness: A Dynamic Factor Model Approach. *IMF Working Papers*, **18**, 1-28. <https://doi.org/10.5089/9781484378243.001>
- [7] 王一迪, 杨赞, 樊颖. 房价黏性、不确定性与住房需求波动[J]. *统计与决策*, 2024, 40(2): 155-159.
- [8] 蓝天. 经济政策不确定性、金融周期与房地产价格——基于 TVP-SV-VAR 模型的分析[J]. *区域金融研究*, 2022(2): 19-27.
- [9] Wang, J. and Biljecki, F. (2022) Unsupervised Machine Learning in Urban Studies: A Systematic Review of Applications. *Cities*, **129**, Article ID: 103925. <https://doi.org/10.1016/j.cities.2022.103925>
- [10] Lam, C., Yao, Q. and Bathia, N. (2011) Estimation of Latent Factors for High-Dimensional Time Series. *Biometrika*, **98**, 901-918. <https://doi.org/10.1093/biomet/asr048>
- [11] Lam, C. and Yao, Q. (2012) Factor Modeling for High-Dimensional Time Series: Inference for the Number of Factors. *The Annals of Statistics*, **40**, 694-726. <https://doi.org/10.1214/12-aos970>

附录

Table A1. Full list of 70 sample cities

表 A1. 70 个研究城市的完整列表

编号	城市	编号	城市	编号	城市	编号	城市
1	北京	19	青岛	37	秦皇岛	55	洛阳
2	天津	20	郑州	38	包头	56	平顶山
3	石家庄	21	武汉	39	丹东	57	宜昌
4	太原	22	长沙	40	锦州	58	襄阳
5	呼和浩特	23	广州	41	吉林	59	岳阳
6	沈阳	24	深圳	42	牡丹江	60	常德
7	大连	25	南宁	43	无锡	61	韶关
8	长春	26	海口	44	徐州	62	湛江
9	哈尔滨	27	重庆	45	扬州	63	惠州
10	上海	28	成都	46	温州	64	桂林
11	南京	29	贵阳	47	金华	65	北海
12	杭州	30	昆明	48	蚌埠	66	三亚
13	宁波	31	西安	49	安庆	67	泸州
14	合肥	32	兰州	50	泉州	68	南充
15	福州	33	西宁	51	九江	69	遵义
16	厦门	34	银川	52	赣州	70	大理
17	南昌	35	乌鲁木齐	53	烟台		
18	济南	36	唐山	54	济宁		

Table A2. Summary of data-driven clustering results

表 A2. 数据驱动聚类结果汇总表

集群类别	城市
政策情绪敏感型	杭州、宁波、福州、青岛、长沙、海口、重庆、成都、西安、兰州、无锡、徐州、扬州、温州、金华、泉州、九江、赣州、烟台、三亚
金融周期敏感型	北京、天津、石家庄、上海、南京、合肥、厦门、郑州、武汉、广州、深圳
增长依赖型	太原、呼和浩特、沈阳、大连、长春、哈尔滨、南昌、济南、南宁、贵阳、昆明、西宁、银川、乌鲁木齐、唐山、秦皇岛、包头、丹东、锦州、吉林、牡丹江、蚌埠、安庆、济宁、洛阳、平顶山、宜昌、襄阳、岳阳、常德、韶关、湛江、惠州、桂林、北海、泸州、南充、遵义、大理

Table A3. Prior classification of sample cities

表 A3. 城市先验分类归属表

编号	城市	行政市场层级	地理区域	人口规模
1	北京	核心城市	东部沿海	超大城市

续表

2	天津	中心城市	东部沿海	超大城市
3	石家庄	三线城市	东部沿海	I型大城市
4	太原	三线城市	中部地区	I型大城市
5	呼和浩特	三线城市	西部地区	II型大城市
6	沈阳	三线城市	东北地区	特大城市
7	大连	三线城市	东北地区	特大城市
8	长春	三线城市	东北地区	I型大城市
9	哈尔滨	三线城市	东北地区	特大城市
10	上海	核心城市	东部沿海	超大城市
11	南京	中心城市	东部沿海	特大城市
12	杭州	中心城市	东部沿海	特大城市
13	宁波	中心城市	东部沿海	I型大城市
14	合肥	中心城市	中部地区	I型大城市
15	福州	中心城市	东部沿海	I型大城市
16	厦门	核心城市	东部沿海	I型大城市
17	南昌	三线城市	中部地区	I型大城市
18	济南	中心城市	东部沿海	特大城市
19	青岛	三线城市	东部沿海	特大城市
20	郑州	中心城市	中部地区	特大城市
21	武汉	中心城市	中部地区	特大城市
22	长沙	三线城市	中部地区	特大城市
23	广州	核心城市	东部沿海	超大城市
24	深圳	核心城市	东部沿海	超大城市
25	南宁	三线城市	东部沿海	I型大城市
26	海口	三线城市	东部沿海	II型大城市
27	重庆	三线城市	西部地区	超大城市
28	成都	三线城市	西部地区	超大城市
29	贵阳	三线城市	西部地区	I型大城市
30	昆明	三线城市	西部地区	特大城市
31	西安	三线城市	西部地区	特大城市
32	兰州	四线城市	西部地区	II型大城市
33	西宁	四线城市	西部地区	II型大城市
34	银川	四线城市	西部地区	II型大城市
35	乌鲁木齐	四线城市	西部地区	I型大城市
36	唐山	三线城市	东部沿海	II型大城市
37	秦皇岛	三线城市	东部沿海	II型大城市

续表

38	包头	四线城市	西部地区	II型大城市
39	丹东	四线城市	东北地区	II型大城市
40	锦州	四线城市	东北地区	II型大城市
41	吉林	四线城市	东北地区	II型大城市
42	牡丹江	四线城市	东北地区	II型大城市
43	无锡	中心城市	东部沿海	II型大城市
44	徐州	三线城市	东部沿海	II型大城市
45	扬州	三线城市	东部沿海	II型大城市
46	温州	中心城市	东部沿海	II型大城市
47	金华	中心城市	东部沿海	II型大城市
48	蚌埠	三线城市	中部地区	II型大城市
49	安庆	三线城市	中部地区	II型大城市
50	泉州	中心城市	东部沿海	II型大城市
51	九江	三线城市	中部地区	II型大城市
52	赣州	三线城市	中部地区	II型大城市
53	烟台	三线城市	东部沿海	II型大城市
54	济宁	三线城市	东部沿海	II型大城市
55	洛阳	三线城市	中部地区	II型大城市
56	平顶山	三线城市	中部地区	II型大城市
57	宜昌	三线城市	中部地区	II型大城市
58	襄阳	三线城市	中部地区	II型大城市
59	岳阳	三线城市	中部地区	II型大城市
60	常德	三线城市	中部地区	II型大城市
61	韶关	中心城市	东部沿海	II型大城市
62	湛江	三线城市	东部沿海	II型大城市
63	惠州	三线城市	东部沿海	II型大城市
64	桂林	四线城市	东部沿海	II型大城市
65	北海	四线城市	东部沿海	II型大城市
66	三亚	四线城市	东部沿海	II型大城市
67	泸州	四线城市	西部地区	II型大城市
68	南充	四线城市	西部地区	II型大城市
69	遵义	四线城市	西部地区	II型大城市
70	大理	四线城市	西部地区	II型大城市