

Centralization or Suburbanization? A Multivariable Regression Analysis of Average Residential Housing Prices in Beijing

Aaron Kaiqiang Zhou

York University, Toronto, L3H3R2, Canada

Abstract

In 1978 China initiated and implemented the Opening Up policy, thus accelerating the rural-to-urban migration in the country level. The research uses Beijing as an example to examine the relationship between average residential housing prices and three key influencing factors. Multiple linear regression was selected as the main methodology for the analysis. Results indicate that housing prices is negatively 0.044 associated with newly constructed residential units, positively 0.018 associated with total unemployment, and positively 0.530 associated with disposable income per capita. To compare, New York city was picked to demonstrate the value difference between China and US. It is recommended to Beijing with the strategy of suburbanization, satellite cities, and redistribution of educational and headquarters of business to rebalance.

Keywords

Multiple linear regression, average residential housing price, decentralization, Beijing.

1. Introduction

How housing prices reflect urbanization and industrialization is an intriguing topic in modern economics. In 1867 the concept of urbanization was first described and defined by Spanish engineer Ildefons Cerdà in his seminal work *Teoría General de la Urbanización*^[1]. According to Ildefons Cerdà, urbanization is the process of moving from rural areas to urban areas. Urbanization signals a shift from agrarian societies to industrial or commercial systems. One crucial indicator is the urbanization ratio, which is often applied to measure a nation's level of development.

After Chairman Deng Xiaoping visited Shenzhen in 1978, China ended the 10-year Cultural Revolution and was eager to re-join in the world. Since then, central government designed 5-year plan for the country and aimed to put the entire country into industrialization with aim of economic growth. For many years, urbanization ratio was emphasized as one key performance indicator for local governments and determined whether the official could get promoted or not, as evidenced by documents such as the National New-type Urbanization Plan (2014–2020)^[2]. To benchmark with the rest of the world, average urbanization of 81%^[3] is common in high-income countries whilst low-income countries had an average of 35%^[3] in 2023. Yet this correlation does not imply a causal relationship but was misunderstood by Chinese central government. As a result, a single direction of movement was pushed for the past 40 years until majority of population claimed to be urban residents. Nonetheless, one premier Dr. Li Keqiang argued in his PHD essay *On the Tripartite Structure of My Country's Economy*^[4], China cannot directly shift from a dual-economy to a unitary-economy economic structure. Enterprises should be placed in rural areas to be the third component and agent between Agriculture and Industrial societies.

Before the 20th century, a lot of research on housing prices used multiple linear models. In 1974, Rosen^[5] proposed the hedonic pricing model, which had the advantage of being quite intuitive by breaking down housing prices into multiple feature factors, but the model ignored

the impact of external factors. In 1995, Leamer^[6] had a study that examined the long-term impact of macro variables on housing prices but didn't consider nonlinear models. Then in 1996, Abraham and Hendershott^[7] combined linear regression with time series models, considering financial indicators, but they overlooked the interference of social factors. Sinai and Gyourko^[8] later, in 1998, introduced external variables like taxes and population to improve the multiple linear models, questioning the linearity assumption. Although the model was quite complex, it explained housing prices well. Similarly, in 1999, Logan and Fang^[9] analyzed China's urban housing market, mainly using data from Shanghai and Tianjin, discussing the connection between urbanization and housing prices. The research was limited by China's statistical data, and there might be statistical errors.

Since the 20th century, research on housing prices has shifted more from linear studies to nonlinear modeling. In 2017, Zhang and Yi^[10] built a nonlinear housing price index model using quantile regression, but their 2004-2014 dataset didn't take special events into account, such as the 2008 financial crisis. In 2023, Chen^[11] used 20-30 variables to build a complex XGBoost model, but the model is a black box. In 2024, Tian et al.^[12] created a nonlinear spatial regression model using multiple forms of transportation accessibility.

After going through plenty of literature, a few research gaps were spotted: 1. Lack of long-term data studies on housing prices. Many studies in China use datasets spanning just 3-5 years or 8-10 years. Because of limitations in China's statistical system (methods and accuracy before 1990 weren't great), not many studies look at long-term trends in housing prices. But data over longer periods, such as 20 years, can show more reliable trends and how various factors are related. 2. Not much discussion on culture and social systems. Lots of housing price models are solid, but the link between housing prices and culture or social systems hasn't been explored enough. Based on the research gaps mentioned above, this paper has a few innovative points. 1) Unlike previous studies with shorter time spans, this study uses a 20-year dataset to build trend and prediction models. 2) By comparing between Western and Chinese cultural and values differences, this research fills a gap in the literature and offers a novel interdisciplinary approach to understanding housing prices.

2. Research Design

2.1. Theoretical Framework and Research Hypotheses

Based on the previous literature review, housing prices are influenced by a variety of factors, including supply-related factors, demand-related factors, policy interventions, and market expectations. Empirical studies by Zhang and Yi^[10], Chen^[11], and Tian^[12] suggest that around 20 to 30 indicators might affect housing prices. This study selects some of these variables to explore the determinants of housing prices in Beijing and to build a predictive model.

2.2. Supply-related Factors

According to Chen Xiaoliang's research, key supply-related indicators contain the supply area of residential land, the unit land price, residential construction costs, and the floor area of completed residential units. Among those indicators, some deserve special attention because they are particularly sensitive to government regulation. Developers' land prices and the supply area of residential land may be controlled by the central government, since real estate development must go through the five-step approval process outlined in the Land Administration Law of the PRC^[13]. This study chooses the number of residential units completed by developers as a substitute indicator of supply. Data for other supply-related factors are not easy to obtain, so they haven't been selected for building the model.

According to Law of Supply and Demand^[14], when demand is equal, an increase in supply typically leads to a decrease in price. The number of completed units is an indicator of supply

and increases supply on the market. As a result, average housing price should face a downward pressure.

Hypothesis 1: The number of completed residential units by real estate developers is negatively correlated with the average residential housing price in Beijing.

2.3. Demand-related Factors

Following the research of Chen Xiaoliang^[11], demand variables include population density, natural population growth rate, proportion of migrants, GDP per capita, average wages, hospital bed availability per capita, public education expenditure per capita, road area per capita, green coverage rate, and unemployment rate. Due to data availability, this study selects disposable income per capita and number of total unemployment in Beijing to represent the effect of demand.

Hypothesis 2: High disposable income increases purchasing power, potentially stimulating housing prices^[15]. Disposable income per capita in Beijing is positively correlated with the average residential housing price.

Hypothesis 3: Previous researcher such as François Geerolf and Thomas Grjebine observed a negative relationship between house prices and unemployment^[16]. The number of total unemployment in Beijing is negatively correlated with the average residential housing price.

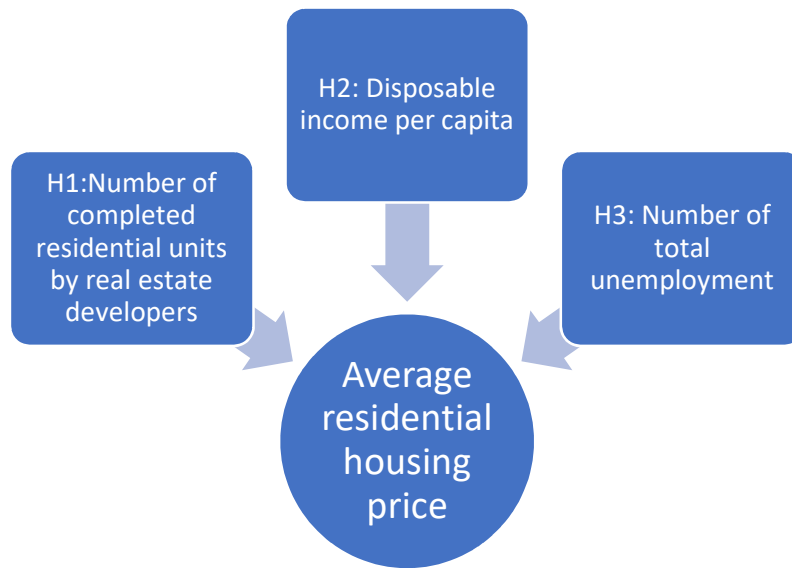


Figure 1. Three Hypotheses

3. Methodology

3.1. Model Construction

To empirically test the hypotheses proposed in this study, a multiple linear regression model is created as shown in Equation (1):

$$Price = \beta_0 + \beta_1 * Construction + \beta_2 * Unemployment + \beta_3 * Income + \varepsilon \quad (1)$$

Where: β_0 is the intercept or constant; β_1 is the coefficient ratio for the number of completed residential units by real estate developers (in ten thousand units); β_2 is the coefficient ratio for the number of total unemployment in Beijing (in ten thousand persons); β_3 is the coefficient ratio for disposable income per capita of all residents in Beijing (in ten thousand yuan); ε is the random error.

3.2. Variable Selection

3.2.1. Dependent Variable

Among several similar indicators provided by the National Bureau of Statistics of China (NBSC)^[17], such as Average Sales Prices of Newly Built Residential Housing, Average Sales Prices of Villas and High-End Apartments, Average Sales Prices of Office Buildings, and Average Sales Prices of Commercial Properties, this study selects the broad indicator: Average Sales Price of Total Residential Housing (P). This metric reflects the average price of newly built and existing residential properties sold by real estate developers during the reporting period, excluding second-hand housing and other property types. It provides a more comprehensive representation of the residential housing market in general.

3.2.2. Core independent Variables

Number Of Completed Residential Units (C):

Among the available indicators, such as Number of Completed High-End Apartments, Number of Completed Residential Units, and Number of Completed Villas, this study selects the Number of Completed Residential Units as it best represents the housing supply stock. This indicator includes all types of residential properties completed and ready for occupancy, aligning with the scope of the dependent variable.

Disposable Income Per Capita of All Residents (I):

Several indicators reflect income levels, including Disposable Income Per Capita for all Residents, Disposable Income Per Capita For Urban Residents, and Disposable Income Per Capita For Rural Residents, as well as various consumption levels and income indexes. Due to data limitations and consistency with the overall population scope, this study selects Disposable Income Per Capita of All Residents as the most comprehensive and representative measure.

Unemployment (U):

Two closely related indicators are available: the Number of Registered Unemployed Individuals and the Registered Unemployment Rate. Given that the unemployment rate is a ratio and lacks data for 2022–2023, this study uses the Number of Total Unemployment as the explanatory variable.

Table 1. Variable Names, Symbols, and Definitions

Variable Name	Symbol	Definition
Average Sales Price of Total Residential Housing (10,000 yuan/m ²)	P	The average price calculated by dividing the total sales contract value of newly built residential housing by the total floor area.
Number of Completed Residential Units (10,000 units)	C	The number of residential units completed by real estate developers that meet occupancy standards and pass inspection during the reporting period.
Disposable Income Per Capita of All Residents (10,000 yuan)	I	The average income available to residents after taxes and social security contributions, reflecting the purchasing power.
Number Of Total Unemployment of All Residents (10,000 persons)	U	The number of unemployed individuals in both city and rural areas who are capable of working and actively seeking employment.

3.3. Data Sources

The data used in this study were obtained from the National Bureau of Statistics of China^[17] (NBSC) for the period 2004 to 2023. The year 2024 was excluded due to incomplete data, which could not be reliably estimated. As it is a 20-year time frame, the data start from the year 2004.

The selection of the 2004–2023 period is based on the availability of long-term provincial data from the NBSC, which is typically segmented into 5-year, 10-year, and 20-year intervals. Research over two decades is relatively rare in China, and this study aims to contribute to that gap by analyzing a 20-year time span.

A total of 80 data was collected. Data from 2005 to 2023 were directly sourced from the NBSC. For the year 2004, where some indicators were missing, supplementary data were obtained from alternative official sources: 1) The Average Sales Price of Total Residential Housing in 2004 was sourced from the *Beijing Municipal Commission of Housing and Urban-Rural Development's 2004 Real Estate Market Report*^[18]. 2) The Number of Completed Residential Units in 2004 was estimated by dividing the total completed floor area (30.67 million m²) by an assumed average unit size of 120 m² per unit. 3) The Number of Registered Unemployed Individuals of All Residents for 2004 was directly obtained from the NBSC. 4) The Disposable Income Per Capita of All Residents for 2004 was not available for all residents; thus, the figure for Disposable Income for Urban Residents^[17] was used as a proxy.

3.4. Research Methodology

This study employs a quantitative empirical analysis approach, primarily utilizing SPSS software to conduct multiple linear regression. The dependent variable is the Average Sales Price of Total Residential Housing (measured in 10,000 yuan per square meter). The three independent variables are: Number of Completed Residential Units (in 10,000 units), Disposable Income Per Capita of All Residents (in 10,000 yuan), Number of Total Unemployment of All Residents (in 10,000 people).

The empirical analysis is divided into three main components. The first component involves descriptive statistical analysis. The second component consists of linear regression analysis, which is carried out in five steps.

4. Empirical Results and Analysis

4.1. Descriptive Statistics

Table 2. Descriptive Statistics of Variables

Variable	Sample Size	Minimum	Maximum	Mean	Standard Deviation	Median
P	20	0.470	4.780	2.422	1.473	1.820
C	20	6.240	25.560	13.863	5.241	13.650
U	20	6.500	37.200	13.635	10.881	8.100
I	20	1.650	8.180	4.509	2.175	4.265

This study conducts a descriptive statistical analysis of the dataset. The column includes key statistical indicators such as sample size, minimum, maximum, mean, standard deviation, and median. The row lists the four variables: one dependent variable and three independent variables.

Among these variables, U or unemployment exhibits the highest standard deviation, indicating the greatest diffusion. In contrast, P or housing price shows the lowest standard deviation, suggesting the least variability.

Data Sources and Reliability

Out of 80 data, 76 originate from the same source, ensuring high reliability. Only the 2004 data come from an alternate source. The data, primarily from the National Bureau of Statistics of China, is considered reliable and free from fabrication. The dataset is complete, with no missing values. Due to high volatility in P and U, the median is recommended for further analysis. I and C are relatively stable and can be described using the mean. Skewness and kurtosis analysis shows that P and I are right-skewed, while U is left-skewed, indicating the presence of extreme values. The coefficient of variation (CV) reveals that U and P exhibit the highest volatility, warranting special attention to regional disparities.

4.2. Correlation Analysis

4.2.1. Exploration of Linear Relationships (Visualized via Scatter Plots)

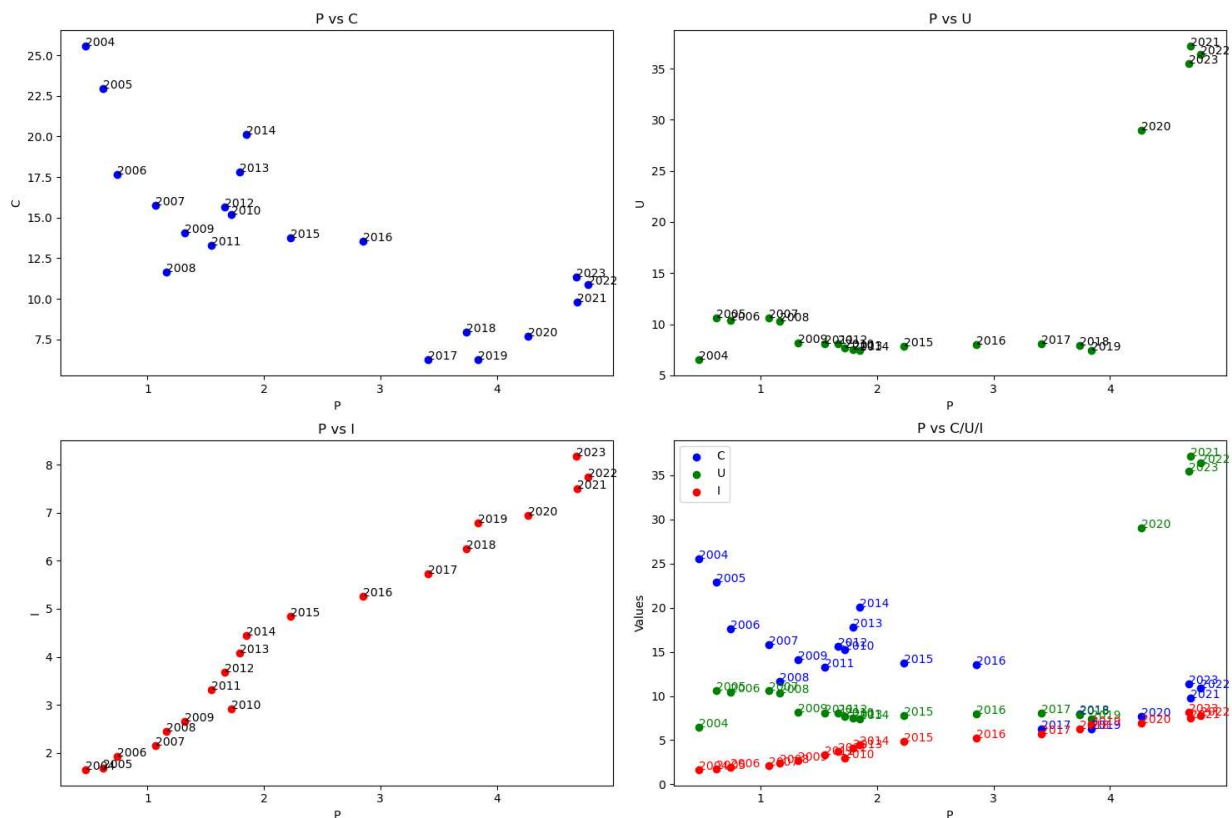


Figure 2. Scatter Points for All Variables

Four scatter plots were constructed to visually examine the linear relationships between housing prices and the three independent variables: P vs. C; P vs. U; P vs. I; P vs. C/U/I.

Among these, Plot 1 and Plot 3 exhibit relatively strong linear relationships: Plot 1 suggests a negative correlation between average residential housing prices and the number of completed residential units. The data points are symmetrically distributed around a downward-sloping trend line, though with considerable dispersion. The graphical analysis suggests a generally negative correlation between variables P and C across most years. However, the precise quantitative relationship remains to be examined and will be further elucidated through the derivation of relevant formulas in the subsequent sections. Plot 3 indicates a positive correlation between average residential housing prices and disposable

income per capita. The data points are more tightly clustered around the trend line, reflecting lower variability.

In contrast, Plot 2 does not reveal a clear linear relationship. Most data points from 2004 to 2019 are concentrated along a horizontal line, indicating low unemployment levels during this period while housing prices steadily increased. After 2020, unemployment numbers rose sharply, yet housing prices remained relatively stable, further weakening the apparent correlation.

4.2.2. Pearson Correlation Coefficient Analysis

Table 3. Pearson Correlation Coefficients

Variable	Mean	SD	C	U	I
C	13.247	4.582	1	—	—
U	14.011	11.045	-0.330	1	—
I	4.659	2.125	-0.696**	0.682**	1

$p < 0.05$, $p < 0.01$

This analysis evaluates the degree of correlation among the three independent variables. The disposable income shows correlation coefficients greater than 0.6 with both the new construction and unemployment, indicating a moderate level of association. However, since all absolute values are below 0.85, the correlations are considered acceptable and do not indicate severe multicollinearity.

Specifically: The correlation between Number of Completed Residential Units (C) and Number of Total Unemployment (U) is -0.330, indicating a weak negative relationship. The correlation between Number of Completed Residential Units (C) and Disposable Income Per Capita (I) is -0.696, significant at the 1% level, suggesting a moderately strong negative relationship. The correlation between Number of Total Unemployment (U) and Disposable Income Per Capita (I) is 0.682, also significant at the 1% level, indicating a moderately strong positive relationship. In terms of variance, a larger variance implies greater deviation between observed data and theoretical correlation. The U–I relationship shows more deviation from theoretical expectations, while the C–I relationship aligns more closely with theoretical predictions.

4.2.3. Multicollinearity Diagnostics

All three independent variables meet the accepted thresholds for multicollinearity assessment: A Variance Inflation Factor (VIF) greater than 10 (or more conservatively, greater than 5) indicates potential multicollinearity^[19].

The tolerance value, calculated as $1/\text{VIF}$, should be greater than 0.1 (or more strictly, greater than 0.2) to rule out multicollinearity^[19].

In this study, all three independent variables—new construction (C), unemployment (U), and disposable income (I)—have VIF values below 5 and tolerance values above 0.1. This confirms that no multicollinearity issues are present, and all variables are retained in the model.

4.2.4. Multiple Linear Regression Analysis

Table 4. Multivariate Regression Results

Variable	Unstandardized Coefficient (B)	Std. Error	Standardized Coefficient (Beta)	t	p	VIF	Tolerance
Constant	0.390	0.277	—	1.407	0.179	—	—
C	-0.044	0.012	-0.155	-3.668	0.002**	2.295	0.436
U	0.018	0.005	0.133	3.320	0.004**	2.067	0.484
I	0.530	0.037	0.782	14.407	0.000**	3.804	0.263

$R^2 = 0.988$, Adjusted $R^2 = 0.985$

$F(3,16) = 424.474$, $p < 0.001$

Durbin-Watson = 1.639

Dependent Variable: Average apartment price (10,000 yuan/m²)

$p < 0.05$, $p < 0.01$

The standard errors across the three predictors are relatively small, indicating stable estimates. Among the independent variables, Number of Total Unemployment shows the least dispersion, followed by Number of Completed Residential Units, while Disposable Income Per Capita exhibits the highest variability.

Both unstandardized and standardized coefficients are reported. The standardized coefficients (Beta), which are normalized using z-scores, allow for direct comparison of the relative influence of each predictor. Among the three, Disposable Income Per Capita has the strongest positive effect on housing prices, while Number of Total Unemployment has the weakest.

4.2.5. Model Robustness Test

The F-test checks the overall significance of the model. The F value of this model is 424.474 with $p < 0.001$, indicating that the model is statistically significant. A higher F value suggests that the variance between groups is larger compared to the variance within groups, so the null hypothesis that all regression coefficients are zero is rejected. The t-test checks the significance of individual coefficients. All three independent variables are significant at the 1% level, meaning they all have a significant impact on the dependent variable. Model fit is assessed using R^2 , which measures the proportion of variation in the dependent variable explained by the predictors^[20]. The R^2 value is 0.988, and the adjusted R^2 is 0.985, indicating a high in-sample fit.

4.2.6. Residual Independence Test

Residual independence is assessed using the Durbin-Watson (D-W) value. A D-W value close to 2, usually between 1.5 and 2.5, is used as a practical indication of limited first-order residual autocorrelation^[21]. In this model, the D-W value is 1.639, within this practical range, although below 2. One possible explanation is that the 2004 data point came from a different source than the other years, which might have introduced estimation errors. When the 2004 data is excluded, the D-W value for the remaining 19 years is 2.054, well within the acceptable range and very close to 2, showing that the residuals are quite independent.

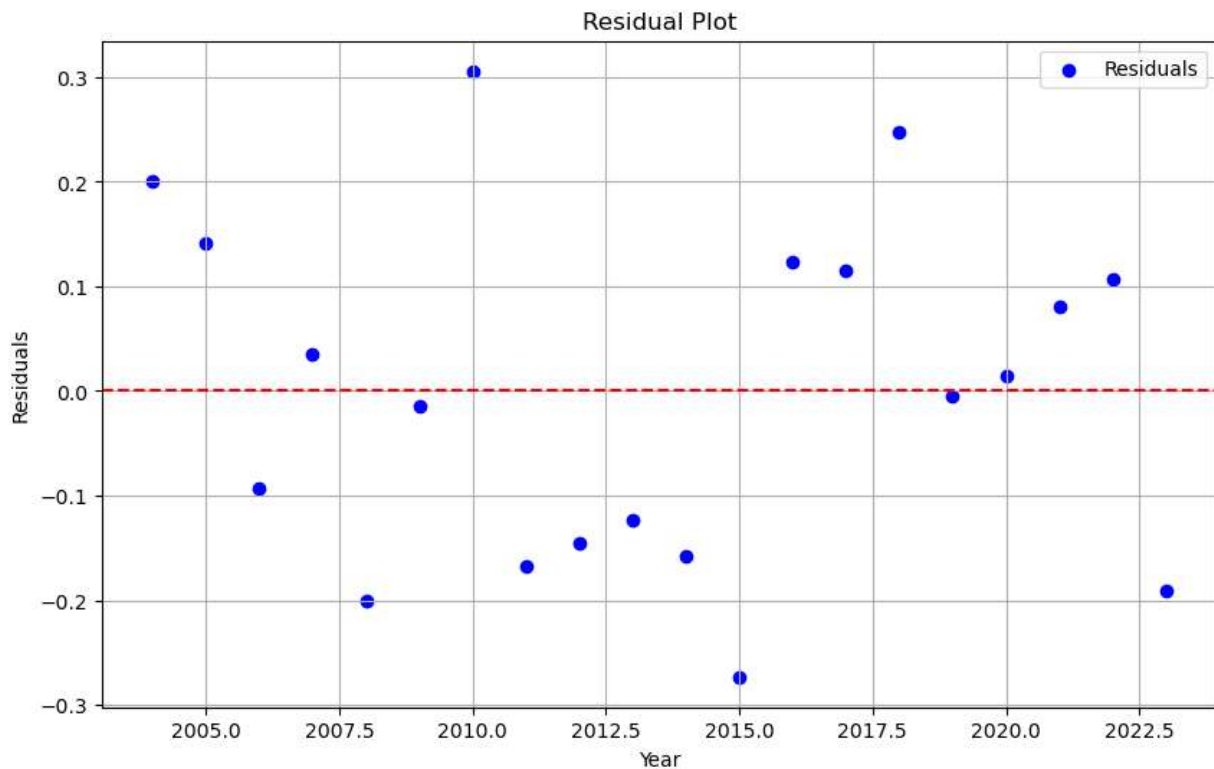


Figure 3. Residual Plot

The residual plot shows the differences between the model's predicted values and the actual observations over time. The red dashed line represents the baseline where the residual is zero. The closer the data points are to this line, the more accurate the model's predictions. Figure 3 demonstrates that the residuals are evenly distributed on both sides of the red line, with no obvious trends, which means the model fits relatively well.

4.2.7. Regression Equation

The estimated regression model is shown as follows:

$$P=0.3902-0.0435C+0.0180U+0.5300I$$

Where P means Average Residential Housing Price (1,000 yuan/m²), C is Number of Completed Residential Units (10,000 units), U means Number of Total Unemployment (10,000 people), and Disposable Income Per capita (10,000 yuan) is denoted as I. For every 10,000 yuan increase in per capita disposable income, the average residential housing price rises by about 5,300 yuan/m², showing a strong positive correlation. For every 10,000 increases in total unemployment, the average residential housing price rises by about 180 yuan/m², indicating a slight positive correlation. For every 10,000 new housing units added, the average residential housing price drops by about 430 yuan/m², indicating a negative correlation. Among all three variables, disposable income per capita has the most influence and is a dominant factor.

4.2.8. Further Analysis in the Socioeconomic Context

Among all the variables in the model, disposable income is the most leading factor, accounting for about 53% of the changes in housing prices. This aligns with the Price-to-Income Ratio, which reflects housing affordability. According to data in Numbeo^[22], Beijing ranks third in housing unaffordability among 257 cities worldwide in 2025, followed by Shanghai (4th), Hong Kong (10th), and Shenzhen (12th). In contrast, cities like London (28th), Tokyo (63rd), New York (64th), and Los Angeles (179th) are much more affordable.

Even though housing prices are so high, the prices in China's first-tier cities haven't crashed. One possible explanation is intergenerational wealth transfer. A 2023 survey^[23] found that almost 70% of first-time homebuyers relied on their families to help with the down payment. However, as Yuan Shi^[23] pointed out in his research, the proportion of young people aged 25 to 34 who own homes dropped from 70% in 2010 to 50% in 2020. With fewer people buying homes, it means more young people can't afford to buy and choose to rent instead. This social phenomenon is further explained under China's centralized urban development strategy. As Cao Zhenhan^[24] pointed out, the Chinese government has historically leaned toward centralization to maintain national stability. Nowadays, this centralization shows up in the four-tier city classification system, with resources and government investments heavily concentrated in first- and second-tier cities. As a result, migration from rural areas to cities has increased significantly. According to data from the National Bureau of Statistics, China's urbanization rate^[17] rose from 26.41% in 1989 to 66.16% in 2024. While this rapid urbanization reflects the modernization process, it also worsens overpopulation and causes socio-economic pressure in big cities.

This model takes unemployment rates into consideration as well. The test results, contrary to classic expectations (and contrary to Hypothesis 3), show a positive correlation with housing prices. Traditionally, high unemployment is commonly linked to poor economic performance and lower housing demand. Nonetheless, the data of past 20 years in Beijing shows a surprisingly structural shift. From 2005 to 2019, Beijing's unemployment rate stayed low and steady. Then in 2020, unemployment soared to four times the previous year's level, mainly because of sudden pandemic lockdowns. This surge might have intervened in the expected negative correlation (hypothesis 3). When unemployment is stable, it barely affects housing prices; however, a sharp rise in unemployment could boost the overall correlation and explain the positive link seen in the model.

Overall, this linear regression model captures the structural and cyclical factors shaping China's urban housing market. It highlights the dominant role of income, the complex effects of unemployment, and the ongoing impact of supply dynamics within a highly concentrated urban development framework.

4.3. Comparative Interpretation of Housing Affordability

To better understand the broad implications of the empirical findings, New York, a representative international city characterized by a more decentralized urban structure, is selected as benchmark.

4.3.1. Affordability Metrics

Table 5. Housing Ratios in New York & Beijing

Indicator	New York, NY	Beijing
Price-to-Income Ratio	15.04	38.46
Mortgage as % of Income	134.39%	271.16%
Loan Affordability Index	0.74	0.37
Price-to-Rent Ratio (City Centre)	18.65	65.70
Price-to-Rent Ratio (Outside Centre)	21.01	55.35
Gross Rental Yield (City Centre)	5.36%	1.52%
Gross Rental Yield (Outside Centre)	4.76%	1.81%

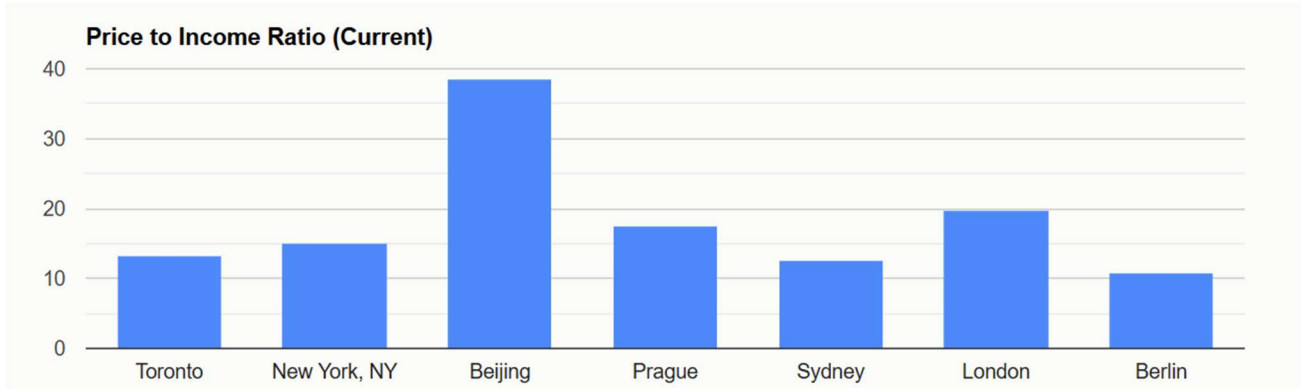


Figure 4. Price to Income Ratio for Various International Mega-cities

According to 2025 data from Numbeo^[25], the price-to-income ratio of Beijing is 38.46 and is more than twice that of that ratio in New York (15.04), making Beijing one of the least affordable cities in the world, surpassing London, Prague, Toronto, Sydney, and Berlin. The mortgage pressure in Beijing and the low rental yields further highlight the financial strain on homeowners and investors.

4.3.2. Value and Cultural Differences

The gap between New York and Beijing in terms of affordability and urban development stems from fundamental differences in government urban planning and cultural values. In East Asian countries like China, Japan, and South Korea, centralization is part of traditional routine. This is evident in the population concentration in megacities such as Tokyo, Seoul, and Beijing, whereas the urban structure in the U.S. is much more dispersed. In terms of medical resources, Beijing has 11 top-tier (3A+++) hospitals, and 23 of hospitals are ranked in China's top 100^[26]; in comparison, only 4 of the top 20 hospitals in the U.S. are in New York^[27]. When it comes to education, Beijing has 26 Project 211 universities and 8 Project 985 universities^[28]. For business headquarters, Beijing hosts 53 Fortune Global 500 headquarters, while New York will only have 16 in 2024^[29]. In contrast, medical, educational, and job resources in the U.S. are more evenly distributed across states. For example, Berkshire Hathaway's headquarters is in Omaha, Nebraska, which shows the dispersed nature of the U.S. economic geography. Yan Ligang, Han Zhenglong, and Liu Di^[30] studied 35 large and medium-sized Chinese cities and found that housing prices are positively correlated (0.16%) with the concentration of public services and institutions, further highlighting the central agglomeration effect.

The urban-structure comparison can be situated in a broader literature. Xiao et al.^[31] compare urban land expansion in Tokyo, Seoul, Beijing, and Shanghai, while Shen et al.^[32] examine the socioeconomic determinants of Chinese city size using an equilibrium framework. Wang and Fu^[33] discuss industrial relocation and satellite-town development as ways to decentralize Beijing's population and industries. These studies provide context for considering the spatial distribution of resources alongside housing affordability. Gollin, Jedwab, and Vollrath^[34] distinguish urbanization with industrialization from urbanization without it, suggesting that the urbanization rate alone is insufficient to characterize industrial development.

At the housing-market level, Chen et al.^[35] use XGBoost and GeoShapley to examine nonlinear and spatially varying relationships in Beijing housing prices. Their analysis highlights forms of heterogeneity that the present aggregate linear model does not capture. This provides a methodological qualification to the interpretation of the regression results, rather than direct evidence explaining the positive unemployment coefficient.

5. Conclusion and Discussion

5.1. Conclusion

This study analyzed the factors affecting residential housing prices in Beijing using a multiple linear regression model with data from 2004 to 2023. The results show that per capita disposable income has the strongest positive association with housing prices, while the number of completed residential units has a negative association. The socioeconomic discussion and comparison between Beijing and New York suggest that the concentration of population and resources is relevant to understanding housing affordability, although the regression does not directly test this relationship. A more dispersed urban structure may help ease housing burdens and support more balanced regional development. Beijing's high price-to-income ratio also indicates that a high urbanization rate can coexist with low housing affordability. Assessments of urban development should therefore consider multiple economic and social indicators.

Contrary to Hypothesis 3, the unemployment measure is positively associated with housing prices in this sample. This result may reflect unusual labor-market fluctuations during the 2020 pandemic rather than a stable long-term relationship. Further analysis would be needed to establish the underlying mechanism.

This research combines a 20-year observation period, a city-focused regression analysis, and a comparative discussion of urban structure in different countries. Its contribution is to consider housing affordability alongside broader urban development outcomes and planning strategies.

5.2. Policy Recommendations

This study proposes a few policy suggestions based on the research findings to improve the effectiveness and efficiency of China's urban structure and industrial development. Firstly, it recommends that policymakers should not rely on a single indicator, i.e. the urbanization rate, as a substitute for assessment of industrialization. While the urbanization rate is a useful reference, relying on just one indicator might overlook the multiple angles of industrial transformation, including sustainable development and residents' well-being.

Secondly, in terms of urban development, a suburbanization strategy is recommended to be adopted. This includes promoting the development of suburbs and satellite cities to ease the pressure on core urban areas, redistributing educational and public resources to lower-tier cities to improve regional equity, and relocating non-core administrative and corporate headquarters out of megacities, especially the capital, to reduce population density and infrastructure load.

A key part of this strategy is strengthening rail connections to enable efficient commuting between suburban residences and urban employment centers.

Indeed, the Chinese government has already started implementing part of this strategy. Since 2017, the establishment of Xiong'an New Area has served as a pilot project to relocate some of Beijing's non-capital functions. Xiong'an is strategically located in the Beijing-Tianjin-Baoding corridor, about 105 km from both Beijing and Tianjin, 155 km from Shijiazhuang, and 55 km from Beijing Daxing International Airport. By 2025, the area had successfully attracted 4 state-owned enterprises and 3 Project 211 universities to relocate, and there are plans to bring in medical institutions, software companies, and financial headquarters in the coming years.

The development of Xiong'an New Area provides empirical support for the recommendations proposed in this study, showing the feasibility and effectiveness of addressing the over-concentration in mega cities through spatial redistribution and multi-node urban planning. Besides that, Tongzhou will also play a role in easing the political centralization of Beijing. Various municipal government-related institutions, other than the central government, will gradually move to Tongzhou. Tongzhou will build a fast urban rail that allows quick access to Beijing's CBD. Residents or migrant workers can choose to live in Tongzhou and commute into

the city for work. Tianjin, as a satellite city of Beijing, should also, under policy allowance, help divert some of the people who have moved to Beijing for work.

5.3. Limitations

This study has three limitations. The first one is timeline of data. The study would significantly improve if the research had a wider time range, ideally covering 30–50 years of economic trends since the 1978 opening-up policy. Secondly, data consistency is preferred. The analysis would be more robust if there were a unified dataset for at least 20 years (data from 2004 comes from a second source). Lastly, variables could be further explored. Adding more variables and building a more comprehensive set of indicators could provide deeper insights.

References

- [1] Ildefons, C. (2021). General theory of urbanization 1867. Actar D, Inc.
- [2] Central Committee of the Communist Party of China, & State Council. (n.d.). National new-type urbanization plan (2014–2020) [State Council Gazette Issue No. 9, 2014]. The Central People's Government of the People's Republic of China. https://www.gov.cn/gongbao/content/2014/content_2644805.htm
- [3] World Bank Open Data. (2018). World bank open data. https://data.worldbank.org/indicator/SP.URB.TOTL.IN.ZS?locations=XD&name_desc=false
- [4] Lin, Q. (2016, April 7). Li Keqiang: On the ternary structure of China's economy. The State Council of the People's Republic of China. http://big5.www.gov.cn/gate/big5/www.gov.cn/premier/2016-04/07/content_5062083.htm
- [5] Rosen, S. (1974). Hedonic prices and implicit markets: Product differentiation in pure competition. *Journal of Political Economy*, 82(1), 34–55. <https://doi.org/10.1086/260169>
- [6] Leamer, E. E. (1995). Housing prices and the macroeconomy [Unpublished manuscript]. Anderson School of Management, University of California, Los Angeles.
- [7] Abraham, J. M., & Hendershott, P. H. (1996). Bubbles in metropolitan housing markets. *Journal of Housing Research*, 7(2), 191–207.
- [8] Sinai, T., & Gyourko, J. (1998). House price dynamics: The role of tax policy and demographics. *Brookings Papers on Economic Activity*, 1998(1), 1–50.
- [9] Logan, J. R., Bian, Y., & Bian, F. (1999). Housing inequality in urban China in the 1990s. *International Journal of Urban and Regional Research*, 23(1), 7–25. <https://doi.org/10.1111/1468-2427.00176>
- [10] Zhang, L., & Yi, Y. (2017). Quantile house price indices in Beijing. *Regional Science and Urban Economics*, 63, 85–96. <https://doi.org/10.1016/j.regsciurbeco.2017.01.002>
- [11] Chen, X. (2023). A study on the influencing factors of housing prices in first-tier cities based on machine learning methods. *Nankai Journal (Philosophy and Social Sciences Edition)*. <http://222.198.130.40:81/>
- [12] Tian, D., Wang, J., Xia, C., Zhang, J., Zhou, J., Tian, Z., Zhao, J., Li, B., & Zhou, C. (2024). The relationship between green space accessibility by multiple travel modes and housing prices: A case study of Beijing. *Cities*, 145, 104694. <https://doi.org/10.1016/j.cities.2023.104694>
- [13] Cao, L. (2019). Land administration law of the People's Republic of China. Ministry of Ecology and Environment of the People's Republic of China. https://www.mee.gov.cn/ywgz/fgbz/fl/201904/t20190428_701293.shtml
- [14] Turvey, R. (2022). Demand and supply. Routledge.
- [15] Zhang, C., Jia, S., & Yang, R. (2016). Housing affordability and housing vacancy in China: The role of income inequality. *Journal of Housing Economics*, 33, 4–14. <https://doi.org/10.1016/j.jhe.2016.05.005>
- [16] Geerolf, F., & Grjebine, T. (n.d.). Assessing house price effects on unemployment dynamics 1 [Working paper].

<https://www.anderson.ucla.edu/sites/default/files/documents/areas/ctr/ziman/2015-15WP.pdf>

- [17] National Bureau of Statistics of China. (2023). National data. <https://data.stats.gov.cn/easyquery.htm?cn=E0103>
- [18] Beijing Municipal Commission of Housing and Urban-Rural Development. (2005, February 16). Beijing real estate market information, January-December 2004. <http://zjw.beijing.gov.cn/bjjs/xxgk/sjtj/319219/index.shtml>
- [19] Pennsylvania State University. (2018). 10.7-Detecting multicollinearity using variance inflation factors | STAT 462. <https://online.stat.psu.edu/stat462/node/180/>
- [20] Fernando, J. (2024, November 13). R-squared: Definition, calculation, and interpretation. Investopedia. <https://www.investopedia.com/terms/r/r-squared.asp>
- [21] SAP. (n.d.). Durbin-Watson. https://help.sap.com/doc/saphelp_scm700_ehp02/7.0.2/en-US/e8/3cc95360267614e10000000a174cb4/content.htm?no_cache=true
- [22] Numbeo. (2025). Current property price to income ratio by city. https://www.numbeo.com/quality-of-life/rankings_current.jsp?displayColumn=5
- [23] Shi, Y. (2024, December 21). China's real estate crisis: Why the younger generation is not buying houses anymore. The Diplomat. <https://thediplomat.com/2024/12/chinas-real-estate-crisis-why-the-younger-generation-is-not-buying-houses-anymore/>
- [24] Cao, Z. (2018). Centralization and decentralization of power structure: A theory of ruling risks and empirical evidence from Chinese history. *Chinese Journal of Sociology*, 4(4), 506–564. <https://doi.org/10.1177/2057150x18789048>
- [25] Numbeo. (2025). Property prices comparison between Beijing, China and New York, NY, United States. https://www.numbeo.com/property-investment/compare_cities.jsp?country1=China&city1=Beijing&country2=United+States&city2=New+York%2C+NY&displayCurrency=USD
- [26] Zhang, Z. (2021, November 20). 2020 China hospital rankings released: 23 top-tier hospitals in Beijing make the list. *Bjnews.com.cn*. <https://m.bjnews.com.cn/detail/163740603514790.html>
- [27] U.S. News & World Report. (2024). Find the best hospitals in your state. <https://health.usnews.com/best-hospitals/area>
- [28] Ministry of Education of the People's Republic of China. (n.d.). List of “Project 211” universities. https://www.moe.gov.cn/srcsite/A22/s7065/200512/t20051223_82762.html
- [29] Fortune. (2024). Fortune global 500. <https://fortune.com/ranking/global500/>
- [30] Yan, L., Han, Z., Liu, D., & Yang, L. (2020). Spatiotemporal differences in the impact of educational resource allocation on housing prices: A case study of 35 large and medium-sized cities. *Human Geography*, 35(2), 22–31. <https://doi.org/10.13959/j.issn.1003-2398.2020.02.003>
- [31] Xiao, R., Murayama, Y., Qin, K., Su, J., Gao, Z., Liu, L., Xu, G., & Jiao, L. (2024). Urban expansion in highly populous East Asian megacities during 1990–2020: Tokyo, Seoul, Beijing, and Shanghai. *Ecological Informatics*, 83, 102843. <https://doi.org/10.1016/j.ecoinf.2024.102843>
- [32] Shen, D., Thill, J.-C., & Sun, J. (2019). Are Chinese cities oversized? *International Regional Science Review*, 43(6), 632–654. <https://doi.org/10.1177/0160017619884108>
- [33] Wang, X., & Fu, M.-C. (2017). Analysis on the industrial structure optimizing and the population decentralizing strategy of Beijing. In *Proceedings of the 2017 International Conference on Innovations in Economic Management and Social Science (IEMSS 2017)* (pp. 546-549). Atlantis Press. <https://doi.org/10.2991/iemss-17.2017.109>
- [34] Gollin, D., Jedwab, R., & Vollrath, D. (2015). Urbanization with and without industrialization. *Journal of Economic Growth*, 21(1), 35–70. <https://doi.org/10.1007/s10887-015-9121-4>

- [35]Chen, Y., Ye, Y., Liu, X., Yin, C., & Jones, C. A. (2025). Examining the nonlinear and spatial heterogeneity of housing prices in urban Beijing: An application of GeoShapley. *Habitat International*, 162, 103439. <https://doi.org/10.1016/j.habitatint.2025.103439>