

Factors Influencing Real Estate Prices: An Empirical Analysis Based on STATA Using Panel Data from China, Japan, and the U.S.

Lixin Feng

Jurong Country Garden School,
Zhenjiang, Jiangsu, 212426, China
Email: Flxin_0721@126.com

Abstract:

The real estate acts as an essential and indispensable part of the national economy, strongly associated with the economic development. Since the real estate market and the economic development interact and influence each other, the study on the dynamics of real estate prices is very significant for monitor of the economy and formulation of policy making. The aim of this paper is to investigate the key determinants influencing real estate prices in three major economies: China, Japan, and the United States. By identifying and analyzing periods of significant price fluctuations, the study employs Stata software to examine the relationship between selected explanatory variables and housing price trends. The study shows that residential land price, population density and real gross domestic product per capita has the greatest influence on real estate price.

Keywords: Real Estate Price, Real Estate, Stata Analysis, Cross-Country Comparison

1. Introduction

As a crucial component of the national economy, the real estate industry plays a fundamental role in the economic structure of many countries. Various economic theories suggest that changes in a country's economy significantly impact real estate prices. A stable economic environment fosters real estate market expansion, while fluctuations in real estate prices, in turn, affect economic growth. A healthy and stable real estate market not only promotes investment and consumption but also contributes to sustained economic growth. Conversely, market instability

may lead to severe economic fluctuations, disrupting the normal operation of the national economy. As a result, the dynamics of real estate prices have long been a focal point of economic regulation and policy-making worldwide.

Over the past decades, many countries have experienced large-scale fluctuations in their real estate markets. According to existing literature, real estate price trends can be considered leading indicators of the overall economic cycle. Therefore, analyzing the factors influencing housing prices is of great significance. This study selects three critical periods of significant real estate market fluctuations in China,

Japan, and the U.S. Using panel data, it explores the factors influencing housing prices from three dimensions: demand, supply, and monetary policy. A multiple linear regression model is employed to analyze real estate price dynamics, evaluating the varying degrees of impact of different explanatory variables on real estate prices.

2. Variable Selection and Research Hypotheses

2.1 Variable Selection

Researchers worldwide have examined the factors influencing real estate prices from multiple perspectives. Based on existing research, relevant studies mainly focus on demand, supply, regulatory policies, and individual consumer behavior. This paper selects three periods of significant real estate market fluctuations in China, Japan, and the U.S. and utilises national-level panel data from the National Bureau of Statistics of China (2010–2023), the U.S. Bureau of Economic Analysis and the Federal

Reserve Bank of St. Louis (2000–2013), and the Statistics Bureau of Japan (1985–2010). The study constructs explanatory variables from the three dimensions of demand, supply, and monetary policy to explore the mechanisms influencing housing prices.

This article selected real per capita GDP index and population density on the demand side as explanatory variables; the number of residential buildings started two years ago (accounting for construction cycles) and the residential land price index are used as explanatory variables of supply side; On the monetary side, different interest rate indicators are selected for the three countries : China uses the loan prime rate (LPR) for loans exceeding five years, Japan uses the long-term benchmark lending rate, and the U.S. uses the average 30-year and 15-year fixed mortgage rates. The actual real estate price index is chosen as the dependent variable. By constructing a multiple linear regression model, this study aims to verify the significance of these explanatory variables in influencing the actual real estate price index, as shown in Table 1.

Table 1: Variable Names and Descriptions

Variable	Variable Name	Symbol	Unit
Explained Variable	Real Residential Property Price Indexes	Y	(Index Jan 2010=100)
Explanatory Variable			
<i>Demand Side</i>	Real GDP per Capita	X_1	(Index Jan 2010=100)
	Population Density	X_2	(Persons per kilometre)
<i>Supply Side</i>	Constructing Dwellings and Residential Buildings Two Years Ago	X_3	(Buildings)
	Residential Land Price Index	X_4	(Index Jan 2010=100)
<i>Monetary Policy</i>	Interest Rate	X_5	(%/year)

2.2 Research Hypotheses

- i. Real per capita GDP index: This metric represents the inflation-adjusted GDP divided by the total population, reflecting the average economic output per person and serving as a major cost component in housing development.
 - Hypothesis 1: The real per capita GDP index is positively correlated with the actual real estate price index.
- ii. Population density: This measure reflects the number of people per unit area and is commonly used to assess the intensity of residential use and resource pressure.
 - Hypothesis 2: Population density is positively correlated with the actual real estate price index.
- iii. Number of residential buildings started two years prior: Given the construction cycle, this study uses the num-

ber of housing starts from two years prior. For example, if the analysis year is 2023, the data reflects housing starts from 2021.

- Hypothesis 3: The number of residential buildings started two years prior is negatively correlated with the actual real estate price index.
- iv. Residential land price index: This index measures changes in land prices for residential use. The data is indexed with January 2010 as the base period.
 - Hypothesis 4: The residential land price index is positively correlated with the actual real estate price index.
- v. Interest rate: This reflects loan interest rates in each country.
 - Hypothesis 5: Interest rates are negatively correlated

with the actual real estate price index.

3. Data Sources and Model Construction

3.1 Data Sources

This study uses data from China (2010–2023), Japan

(1985–2010), and the U.S. (2000–2013) to conduct an empirical analysis of housing price determinants. The data sources include the National Bureau of Statistics of China, the U.S. Bureau of Economic Analysis, the Federal Reserve Bank of St. Louis, and the Statistics Bureau of Japan, as shown in Tables 2, 3, and 4.

Table 2 : Data of influencing factors of real estate in China from 2010 to 2023

year	Real GDP per capita	Population Density (Persons per kilometre)	Constructing Buildings Two Years Ago	Residential land price index	Loan Prime Rate(,5 years)	Real Residential Property Price Indexes
2010	100	141.94	283266	100	6.27	100
2011	109	142.71	320368	107	6.82	98.65
2012	116.7	143.69	405356	111	6.68	95.49
2013	125	144.65	5057959	123	6.15	98.62
2014	133.4	145.56	573418	130	5.4	99.13
2015	142	146.41	665572	136	4.9	93.92
2016	150.8	147.25	726483	150	4.9	98.03
2017	160.3	148.14	735693	166	4.9	103.3
2018	170.3	148.84	758975	178	5.46	106.72
2019	179.8	149.37	781484	187	4.85	109.78
2020	183.4	149.72	822300	193	4.75	109.65
2021	198.7	149.86	893821	202	4.65	111.82
2022	204.6	150.44	926759	187	4.6	107.16
2023	215.6	149.18	975387	175	4.2	103.45

Table 3: Data of influencing factors of real estate in Japan from 1985 to 2010

year	Real GDP per capita	Population Density (Persons per kilometre)	Constructing Buildings Two Years Ago	Residential land price index	Long-term Prime Lending Rates	Real Residential Property Price Indexes
1985	68.9	331.42	94733	159	7.36	122.2
1986	70.4	333.1	98940	164	6.56	125.41
1987	73	334.8	103006	173	5.4	134.75
1988	77.8	336.2	113717	190	5.6	142.58
1989	81.7	337.55	139525	205	5.93	153.66
1990	86	338.67	140387	233	7.97	174.55
1991	88.5	340	138551	257	7.14	181.98
1992	88.9	341.26	142259	253	6.02	174.95
1993	88.8	342.37	114177	228	4.71	167.36
1994	89.3	343.33	116883	224	4.45	163.44
1995	90.8	344.14	123807	220	3.31	160.94
1996	92.9	345.01	130854	210	3.01	158.02
1997	94.2	345.84	122527	201	2.7	155.95
1998	92.1	346.78	136939	194	2.41	153.33
1999	91.7	347.41	115585	185	2.39	148.56
2000	93.6	347.99	99858	174	2.23	142.91
2001	93.7	348.83	101217	163	2.09	136.64
2002	93.8	349.64	102487	151	1.93	129.41
2003	95.2	350.39	97822	139	1.56	121.36
2004	97.4	350.51	95918	127	1.71	113.93
2005	98.7	350.54	96674	118	1.65	108.42
2006	100.3	350.77	99087	112	2.31	105.18
2007	102.5	351.17	103015	110	2.34	104.09
2008	101.5	351.34	107533	109	2.29	102.45
2009	96	351.29	88395	105	1.94	98.58
2010	100	351.36	91127	100	1.49	100

Table 4: Data of influencing factors of real estate in United States from 2000 to 2013

year	Real GDP per capita	Population Density (Persons per kilometre)	Constructing Buildings Two Years Ago	Residential land price index	30-Year Fixed Rate Mortgage	15-Year Fixed Rate Mortgage	Real Residential Property Price Indexes
2000	92.1	30.8	1621000	66	8.05	7.72	97.35
2001	92.1	31.1	1647000	72	6.97	6.5	102.99
2002	92.8	31.39	1573000	85	6.54	5.98	109.95
2003	94.5	31.66	1601000	95	5.83	5.17	117.95
2004	97.2	31.96	1710000	108	5.84	5.21	131.16
2005	99.7	32.25	1854000	124	5.87	5.42	147.01
2006	101.5	32.57	1950000	125	6.41	6.07	150.67
2007	102.5	32.88	2073000	120	6.34	6.03	138.06
2008	101.7	33.24	1812000	109	6.03	5.62	113.77
2009	98.2	33.54	1342000	101	5.04	4.57	103.08
2010	100	33.82	900000	100	4.69	4.1	100
2011	100.8	34.06	554000	97	4.45	3.68	94.27
2012	102.4	34.31	586000	96	3.66	2.93	96.73
2013	103.7	34.55	612000	102	3.98	3.1	104.75

3.2 Model Construction

To explore the relationships between the dependent variable (actual real estate price index) and the explanatory variables (real per capita GDP index, population density, number of residential buildings started two years prior, residential land price index, and interest rates), this study uses STATA software to analyze real estate data from Chi-

na, Japan, and the U.S. A multiple linear regression model is established as the follow equation:

$$Y = \beta_0 + \beta_1 * X_1 + \beta_2 * X_2 + \beta_3 * X_3 + \beta_4 * X_4 + \beta_5 * X_5 + e$$

4. Empirical Analysis of Factors Affecting the Actual Real Estate Price Index

Table 5 : Regression Results Of China

Y	Coef.	Std. Err.	t	P > t	[95% Conf. Interval]	
X ₁	0.0324028	0.0494545	0.66	0.531	[-0.0816, 0.1464]	
X ₂	-3.910428	1.255865	-3.11	0.014	[-6.8065, -1.0144]	
X ₃	0.000000628	0.000000454	0.14	0.894	[0.00000111, 1.11E-06]	
X ₄	0.4464531	0.0831197	5.37	0.001	[0.2548, 0.6381]	
X ₅	1.378517	1.579984	0.87	0.408	[-2.2649, 5.0219]	
_cons	596.4454	176.378	3.38	0.01	[189.7171, 1003.174]	
R-squared			0.9319	Adj R-squared		0.8893

Table 6 : Regression Results Of Japan

Y	Coef.	Std. Err.	t	P > t	[95% Conf. Interval]	
X ₁	-0.3963098	0.3619084	-1.1	0.287	[-1.1512, 0.3586]	
X ₂	1.086888	0.7286918	1.49	0.151	[-0.4331, 2.6069]	
X ₃	0.0000707	0.0000711	0.99	0.332	[-0.0000776, 0.0002191]	
X ₄	0.559382	0.0246391	22.7	0.000	[0.5080, 0.6108]	
X ₅	0.0395681	0.8903325	0.04	0.965	[-1.8176, 1.8968]	
_cons	-306.107	225.5493	-1.36	0.190	[-776.5946, 164.3807]	
R-squared			0.9912	Adj R-squared		0.9890

Table 7: Regression Results Of United States

Y	Coef.	Std. Err.	t	$P > t $	[95% Conf. Interval]	
X_1	2.597015	1.210561	2.15	0.069	[-0.2655, 5.4595]	
X_2	-20.25804	5.572227	-3.64	0.008	[-33.4343, -7.0818]	
X_3	0.00000839	-0.66	0.532	-0.0000253	[-0.0000253, 0.0000143]	
X_4	0.9768664	0.2057637	4.75	0.002	[0.4903, 1.4634]	
X_5	-6.012484	20.21503	-0.3	0.775	[-53.8134, 41.7885]	
X_6	3.791168	16.9029	0.22	0.829	[-36.1778, 43.7602]	
_cons	446.7338	139.3782	3.21	0.015	[117.1567, 776.311]	
R-squared			0.9788	Adj R-squared		0.9607

4.1 Model Fit Analysis

The R-squared and adjusted R-squared values indicate the model's explanatory power:

- China (2010–2023): R-squared = 0.9319, Adjusted R-squared = 0.8893, indicating strong explanatory power (93.19% of housing price variations explained).
- Japan (1985–2010): R-squared = 0.9912, Adjusted R-squared = 0.9890, suggesting an extremely high fit.
- U.S. (2000–2013): R-squared = 0.9788, Adjusted R-squared = 0.9607, demonstrating strong explanatory power.

All three models show high goodness-of-fit, confirming that the selected explanatory variables effectively explain real estate price fluctuations.

4.2 Significance of Variables

4.2.1 Regression results for China

Based on the above regression results, in the regression model for China, the p-values of variables X_2 and X_4 are 0.014 and 0.001 respectively, both of which are below the 0.05 significance threshold, indicating statistical significance within the model. The corresponding regression coefficients are -3.9104 for X_2 and 0.4465 for X_4 , suggesting that population density (X_2) exerts a significant negative impact on the housing price index, while the land price index on the supply side (X_4) has a significant positive effect.

In contrast, the p-values for variables X_1 , X_3 , and X_5 are 0.531, 0.894, and 0.408 respectively, all exceeding 0.1, implying that these variables are not statistically significant within the model.

4.2.2 Regression results for Japan

In the case of Japan, variable X_4 similarly demonstrates a high level of statistical significance, with a p-value of 0.000 and a regression coefficient of 0.5594, indicating a significant positive effect on the housing price index. This result is consistent with the findings from the Chinese model. However, variable X_2 has a regression coefficient of 1.0869 and a p-value of 0.151, which does not meet the 0.05 significance level. Moreover, the sign of the coefficient differs from that observed in the Chinese model.

The remaining variables X_1 , X_3 , and X_5 also exhibit p-values above 0.05 and are thus statistically insignificant.

4.2.3 Regression results for the United States

Regarding the United States, variable X_2 exhibits a regression coefficient of -20.2580 with a p-value of 0.008, indicating a statistically significant negative impact on the housing price index. This aligns with the findings from the Chinese model, in which X_2 also shows a significant negative effect. In addition, the regression coefficient of variable X_4 is 0.9769 with a p-value of 0.002, again revealing a significant positive effect on housing prices. Variable X_1 has a p-value of 0.069, which falls below the 0.1 threshold, indicating weak statistical significance.

The remaining variables— X_3 , X_5 (average 30-year fixed mortgage rate), and X_6 (average 15-year fixed mortgage rate)—have p-values of -0.000, 0.775, and 0.829 respectively, none of which meet conventional significance levels.

4.2.4 Comparative analysis

Based on the regression results of the three models, variables X_2 and X_4 demonstrate significant effects on housing

prices. All three models confirm a positive relationship between X_4 (residential land price index) and the housing price index, which is consistent with the research hypothesis. However, regarding variable X_2 (population density), the sign of the regression coefficient varies across models. In the Japan-based model, X_2 shows a significant positive impact on housing prices, in alignment with the hypothesis. Conversely, the models for China and the United

States indicate a negative relationship between population density and housing prices, contradicting the initial hypothesis.

Given the statistical insignificance of the remaining explanatory variables, separate regressions were conducted for variables X_1 , X_3 , and X_5 to eliminate the possibility of multicollinearity, as shown in Tables 8, 9, and 10.

Table 8 : Regression Results Of X_1 X_3 X_5 Of China

Y	Coef.	Std. Err.	t	P > t	[95% Conf. Interval]
X_1	0.1147829	0.0295334	3.89	0.002	[0.0504351 , 0.1791307]
X_3	-0.000000385	0.00000138	-0.28	0.785	[-0.00000339 , 0.00000262]
X_5	-3.727852	1.656996	-2.25	0.044	[-7.338136 , -0.1175676]
R-squared (X_1)			0.5573	Adj R-squared (X_1)	0.5204
R-squared (X_3)			0.0065	Adj R-squared (X_3)	-0.0763
R-squared (X_5)			0.2967	Adj R-squared (X_5)	0.2380

Table 9 : Regression Results Of X_1 X_3 X_5 Of Japan

Y	Coef.	Std. Err.	t	P > t	[95% Conf. Interval]
X_1	-0.9353652	0.5411724	-1.73	0.097	[-2.05229 , .1815596]
X_3	0.001285	0.000163	7.88	0.000	[0.0009486 , 0.0016215]
X_5	6.95563	2.087399	3.33	0.003	[2.647451 , 11.26381]
R-squared (X_1)			0.1107	Adj R-squared (X_1)	0.0736
R-squared (X_3)			0.7214	Adj R-squared (X_3)	0.7098
R-squared (X_5)			0.3163	Adj R-squared (X_5)	0.2878

Table 10 : Regression Results Of X_3 X_5 X_6 Of United States

Y	Coef.	Std. Err.	t	P > t	[95% Conf. Interval]
X_3	0.0000262	0.00000726	3.61	0.004	[0.0000104 , 0.000042]
X_5	4.752211	4.378379	1.09	0.299	[-4.787457 , 14.29188]
X_6	4.676739	3.881612	1.20	0.251	[-3.780567 , 13.13404]
R-squared (X_3)			0.5205	Adj R-squared (X_3)	0.4905
R-squared (X_5)			0.0894	Adj R-squared (X_5)	0.0135
R-squared (X_6)			0.1079	Adj R-squared (X_6)	0.0336

By conducting separate regressions for the previously insignificant explanatory variables, the results presented above were obtained. Examination of the R-squared and adjusted R-squared values indicates that the explanatory power of the models constructed with these variables remains relatively weak.

It is noteworthy that variable X_1 in both the Chinese and

Japanese models yielded p-values of 0.002 and 0.097, respectively, both of which are below the 0.1 significance threshold. This aligns with the results from the U.S. model and suggests that real per capita GDP exhibits a weak correlation with housing prices. Additionally, the regression coefficients for this variable are positive in the models for China and the United States, which is consistent with the

research hypothesis. Regarding variable X_5 (and X_6), the interest rate variables in the Chinese and Japanese models produced p-values of 0.044 and 0.003, respectively, both of which are statistically significant at the 0.05 level. However, in the U.S. model, neither the average 30-year fixed mortgage rate (X_5) nor the 15-year counterpart (X_6) passed the significance test. Moreover, the regression coefficients for variable X_3 are all close to zero across the models, suggesting that the number of residential buildings started two years prior has only a negligible impact on housing prices.

4.3 Interpretation of Regression Model Results

In summary, the empirical results support theoretical hypotheses 1 and 4 proposed in Section 2.2, namely, that there is a positive relationship between real per capita GDP, the residential land price index, and the real housing price index. However, hypothesis 2 is not supported: the regression models reveal a negative correlation between population density and real housing prices, contrary to the hypothesized positive relationship.

Among the explanatory variables X_1 , X_2 , and X_4 , variable X_4 demonstrates the strongest explanatory power, followed by X_2 and then X_1 . This implies that the residential land price index has the greatest influence on housing prices, followed by population density, and lastly, real per capita GDP.

5. Conclusion

In conclusion, the above findings indicate that both sup-

ply-side and demand-side factors exert substantial influence on housing prices. However, the lack of empirical support for hypothesis 2 may be attributable to lagging effects stemming from country-specific economic structures and stages of urbanization. These factors may have contributed to the observed negative correlation between population density and real housing prices, as opposed to the anticipated positive relationship.

Similarly, the lack of support for hypothesis 5 may also be due to varying degrees of lagged responses in the housing markets of different countries to interest rate adjustments.

References

- [1] Meng Suo. (2022). Analysis of Influencing Factors of Housing Price—An empirical study of China's panel data based on STATA. *Advances in Applied Mathematics*, 11(04), 1503–1511. <https://doi.org/10.12677/aam.2022.114164>
- [2] Alkali, M. A., Sipan, I., & Razali, M. N. (2018). An overview of macro-economic determinants of real estate price in Nigeria. *International Journal of Engineering & Technology*, 7(3.30), 484-488.
- [3] Mavrodiy, A. (2005). Factor analysis of real estate prices. Kyiv: National University.
- [4] Dobrovolska, O., & Fenenko, N. (2024). Forecasting trends in the real estate market: analysis of relevant determinants. *Financial Markets, Institutions and Risks*, 8(3), 227-253.
- [5] Guerrieri, V., Hartley, D., & Hurst, E. (2013). Endogenous gentrification and housing price dynamics. *Journal of Public Economics*, 100, 45–60. <https://doi.org/10.1016/j.jpubeco.2013.02.001>
- [6] Adams, Z., & Füss, R. (2009). Macroeconomic determinants of international housing markets. *Journal of Housing Economics*, 19(1), 38–50. <https://doi.org/10.1016/j.jhe.2009.10.005>