Influencing Factors of California Housing Prices in 1990: a Multiple Linear Regression Analysis

Yitan Hao[1, *, †], Luhan Zhuang[2, *, †], Zitao Ying[3, *, ††], Juncheng Zhai[4, *, ††]

* Corresponding author: 118010088@link.cuhk.edu.cn[1], *luhan.zhuang@mail.utoronto.ca[2], *3180103695@zju.edu.cn[3], *l902020040@mail.bnuz.edu.cn[4]

School of Data Science, The Chinese University of Hong Kong, Shenzhen, China[1]
Faculty of Arts and Science, University of Toronto, Toronto, Canada[2]
Electronic Engineering, Zhejiang University, Hangzhou, China[3]
School of Management, Beijing Normal University Zhuhai, Zhuhai, China[4]
†These two authors contributed equally.
††These two authors contributed equally.

Abstract—In 1990, it was a hot topic to study the housing prices of blocks in California. That was mainly because of the recession period of the U.S. economy in 1990, which affected consumption and personal income and impacted the real estate market. The purpose of the study is to find factors that affect the housing price for blocks in California in 1990 with the consideration of properties of the house itself, the income of buyers, and the geographic surroundings. The paper uses the housing price impact theory and multiple linear regression method to study the influencing factors of the California housing price in 1990. This research concludes that median house value is positively correlated with median income, median house age, total bedrooms, households, and distance to San Diego, while negatively correlated with total rooms, population, distance to the coast, and Los Angeles. In addition, there is no statistically significant relationship between the median house value and distance to San Jose. In fact, it gives very specific results about the impact of each specific factor involved in our model. However, the general ideas like the impact of the whole dwelling structures do not involve in this study.

Keywords- house value; income; Dwelling structures; multiple linear regression

1. INTRODUCTION

From 1983 to 1988, the U.S. economy recovered and expanded with the higher consumption caused by baby boomers and the stimulation from the 1986 Tax Reform Act, according to Kamery claimed in 2004 [1]. Higher consumption and demand for money drove up the price and the interest rate and housing price, also surged that all ended in high debts for people. With various other factors combined, such as the increase in oil prices due to the invasion of Kuwait

ICEMME 2022, November 18-20, Nanjing, People's Republic of China

DOI 10.4108/eai.18-11-2022.2327138

by Iraq, the crunch of credits, and so on, as Walsh discussed in 1993, the economy was in the period of recession, which started in July 1990 and ended in 1991 [2]. It resulted in the loss of jobs up to 1.5 million and affected especially white-collar workers in service sectors, as Singleton mentioned in 1993 [3]. Various industries also received negative impacts, thus affect the whole economy.

With the impacts of the recession, consumption behaviors and income of individuals were affected. Since the purchasing of houses was also part of consumption, it affected the real estate market as well. The median house prices of all blocks in California, the U.S., in the year 1990 will be investigated. The research can reflect what variables people may consider when purchasing houses during the time of recession.

The price of houses for a block is the median price of dwellings for a block group in California based on the 1990 census measured in U.S. dollars. Specifically, there are different types of houses, such as condominiums, apartments, townhouses, and the house price refers to the sold price agreed by the seller and the buyer. It can be seen from the line chart (see Figure 1) that the median housing price of California generally showed the trend of increasing during the period from 1975 to 1990 [4]. To be more precise, it started with around $42,000 in 1975 and then rose faster until around the year 1981, which was the recession period. Then the increasing rate decreased from 1981 to 1984 and then surged dramatically and reached its peak around 1989 at approximately $260,000 [4].

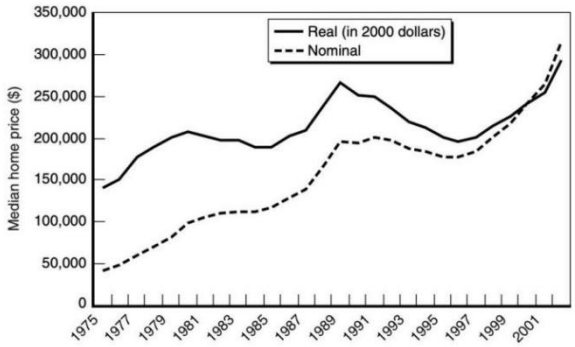

Figure 1. The housing price from 1975 to 2001

Researchers have conducted many investigations for housing prices. Firstly, factors from various aspects that affect the change of housing prices have been discussed. In 1996, Abraham and Hendershott studied from the perspective of macroeconomics which considered how the change in income, construction costs, and interest rate in the real term affect the housing prices in metropolitan [5]. Moreover, Holger and Claus conducted research in 2011 on whether investment products that relate to housing prices can maximize utility for people [6]. From the previous studies, it can be shown that housing prices affect people's welfare, and different factors may attribute to its variation. As a result, the relevant study is beneficial from the perspective of individuals, and it is one of the significant problems to be discussed.

The study conducted by Pace and Barry in 1997 found median housing price in blocks of California is related to several variables such as income, population, and the number of rooms with the sparse spatial autoregressions, with the data from the 1990 census of the U.S. [7]. However, how the geographic factors of the block, for instance, the distances to the coast and major cities, may contribute to the change of price were not considered. While in this study, corresponding variables have been included for analysis, and more specifically, it contains the distances from the centre of blocks to the centers of major cities in California, including LA, San Diego, San Jose, and San Francisco, and also the distance to the nearest coast. Given those new variables, it can further explore whether the housing price can be affected by them. Also, multiple linear regression is used, which differs from previous studies. Based on existing studies and the revision of them, what factors affect the housing price for blocks in California in 1990 with the consideration of properties of the house itself, the income of buyers, and the geographic surroundings will be discussed in this paper.

Based on the existing literature and the gap filled by this study, the housing prices of all blocks in California in 1990 are assumed to be affected by household's income, population, dwelling properties like the house age and the number of rooms, as well as geographic properties, such as distance to coasts and major cities. To verify the above assumption, the influencing factors of California housing prices are analyzed by the theory of housing price influencing factors. The method of multiple linear regression and R software is also applied for the analysis.

2. LITERATURE REVIEW

Over the past several decades, various studies pay particular attention to the factors influencing housing price, such as income, dwelling structures, population, and geographic properties like distance to the coasts and major cities. The detailed literature is as follows.

2.1 House value

The house value can be referred to as the sold price of the dwelling or the monthly payment to the house owner, and in this analysis, the former one is used. For the housing price in California, overall, it increased from 1975 to 1990 with reaching a peak in 1989 from figure1 [4]. Different variables from different perspectives can result in the variation of housing prices according to existing studies.

2.2 Income

Residents' income is an important influencing factor of housing prices. Lamont and Stein, in 1997, found that for cities with higher loan-to-value ratios, housing prices were more sensitive to variations in per capita income, indicating that small changes in income might lead to a big variation in housing prices [8]. Abraham and Hendershott later pointed out that housing price and income are positively related, and that when the gap between the two is too large, housing prices should stagnate or fall to allow income to catch up, and Sani and Rahim in 2015 concluded that they have a positive relationship as well [5, 9]. However, Gallin in 2003 argued that there is little evidence of a significant relationship between housing prices and various fundamentals like income [10]. From the above literature, the positive correlation between income and housing prices is still favored. Therefore, the hypothesis can be made as follows:

H1: The housing price has a positive relationship with income.

2.3 Dwelling structures

The differences in the structures of houses attribute to the variation of price. To be more specific, the structure of houses includes properties such as the age of dwelling, the number of rooms, and its floor level.

1) Total number of rooms: Several studies provide relevant results for the impacts of the number of rooms on housing prices. In 1980, Li and Brown proposed that the number of rooms positively affected housing prices, and the impact declined as the number of rooms increases [11]. Moreover, in 2009, Selim used hedonic pricing and artificial neural network technology with the house prices in Turkey concluded that the number of rooms and house prices are positively correlated [12]. Thus, based on existing studies, it can conclude that housing prices increase as the number of rooms increases. Thus, the hypothesis can be proposed as following：

H2: the housing price and the number of total rooms are positively related.

2) Total number of bedrooms: The relationship between the number of bedrooms and the housing price has been examined in several studies. In 1970, Kain and Quigley constructed research based on housing prices in St. Louis and found that given the condition of the same area, the number of bedrooms positively correlated with housing prices [13]. However, in 2009, Jim and Chen stated that the number of bedrooms and the housing price is negatively related given the average apartment area in Hong Kong was 70m² [14]. Additionally, Liao and Wang in 2012 further verified the conclusion of Kain and Quigley that the number of bedrooms in the unit had a positive impact on the housing price for the city Changsha [15]. Therefore, the hypothesis can be made as:

H3: the housing price and total number of bedrooms are related

3) Housing age： Many studies have been conducted to investigate the relationship between housing age and dwelling price. In 1993, Rubin found that factors such as age and location directly impact households by comparing the rent premium of new and old dwellings [16]. Also, Goodman and Shibodeau, in 1995, concluded that the age of housing would have a significant impact on its depreciation rate that affects the price [17]. Based on their result, in 1997, Clapp and Giaccotto further found that the time series of age coefficients is nonstationary that also shows that they are related [18]. In 2008, Coulson and Mcmillen used real estate sales data in Chicago and concluded the house's value tends to decline during the first few years after the construction; however, it may go up for homes with higher age [19]. These studies can show a relationship between housing age and housing price, and during different periods after the construction, the price will alter. Thus, the corresponding hypothesis can be made as:

H4: The relationship between house price and housing age varies in different periods.

2.4 Population density

Population density is also an important factor that should be considered in the analysis. With more people in the given neighborhood, normally, it has a larger population density. As Sharpe pointed out in 2018, higher population density leads to improvement of transportation and better public services, which leads to an increase in housing prices [20]. Based on previous studies, the

number of people in certain regions will affect the housing price. In addition, the number of households, a group of people residing within a home unit for a block, can be taken as a population indicator. Thus, the corresponding hypothesis can be made as:

H5: The housing price and population density have a positive relationship.

2.5 Geographic factors

Geographic factors may also affect the housing price, like distance to major cities and coasts.

1) Distance to major cities: The impact of distance from the city centre on housing prices has been studied for several decades. Alonso, Mills, Muth proposed the monocentric model theory, which assumed that only a single employment node was located at the central business district (CBD) [21-23]. It showed that the housing prices decreased when houses were located far from the city center. However, subsequent empirical tests produced contradictory results. Kain and Quigley in1970 found no statistically significant relationship between housing price and distance to CBD, while a significant positive correlation was found by Berry in 1976 [11,24]. Richardson, in 1988 argued that the distance to CBD failed to explain the fluctuations in house prices because cities were polycentric, not monocentric [25]. Then, Dubin and Sung, in 1987, considered the impact of sub-centres and eventually demonstrated that the centre produced the expected peak in housing prices. Still, the central effect was limited to a relatively small region [26]. Thus, it is hard to tell the relationship between the housing price and a single center because the housing price is almost affected by more than one center that has limited center effects. Thus, the hypothesis is made to be as follows:

H6: The correlation between the housing price and a single centre is not clearly discernable.

2) Distance to the coast: The distances to the coasts affect housing prices, which existing studies can support. Boarnet and Chalermpong, in 2001, found that in Orange County, California, a one-mile increase in distance from the coast would reduce the housing price by around $42,000 [27]. A study on San Diego housing prices by Conroy and Milosch in 2009 also showed similar results [28]. The study found that houses sold for approximately $8,680 less with each mile from the coast for median-priced houses with an average distance from the coast. All studies show that proximity to the coast has a positive effect on the house value. Therefore, the corresponding hypothesis can be made as:

H7: There is a negative relationship between the housing price and distance to the coast.

3. METHODOLOGY

3.1 Research design

In this study, the method of multiple linear regression is used to investigate the relationship between the housing price of all blocks of California and relevant factors since this method has been used in many previous studies on influencing factors of housing prices. In 2003, Gallin, Federal Reserve Board used multiple linear regression to investigate the relationship between the housing prices and fundamentals, including income at the national level [10]. In 2009, Hasan used this method to explore the determinants of housing prices in Turkey [12].

3.2 Data collection

The data used in this analysis is secondary data from the Kaggle website, which was partially derived from the paper of Pace and Barry in 1997, and the site user added the distances of California blocks to the nearest coasts and major cities [7]. In this study, latitude and longitude variables are removed and not considered. The following table (see Table 1) shows the variables involved in this study and their detailed descriptions [29].

TABLE I. DEFINITION AND SUMMARY OF VARIABLES

Symbol	Variable	Description
Y	median house value	median house value for households within a block (measured in US Dollars)
X1	median income	the median income for households within a block of houses (measured in tens of thousands of US Dollars)
X2	median age	the median age of a house within a block (year)
X3	total rooms	total number of rooms within a block
X4	total bedrooms	total number of bedrooms within a block
X5	population	total number of people residing within a block
X6	households	total number of households, a group of people residing within a home unit, for a block
X7	distance to coast	distance to the nearest coast point (meter)
X8	distance to Los Angeles	distance to the centre of Los Angeles (meter)
X9	distance to San Diego	distance to the centre of San Diego (meter)
X10	distance to San Jose	distance to the centre of San Jose (meter)
X11	distance to San Francisco	distance to the centre of San Francisco (meter)

3.3 Data collection

First of all, since the range of distinct variable values of raw data varies widely, the objective function will not work efficiently without scaling. The unit length scaling is used to scale the raw data to make each variable vector has length one. Then a multiple linear regression model is established based on scaled data, and the unknown parameters in the regression model are estimated by ordinary least square (OLS). The next step is to test the significance of individual variables to find if there is a linear relationship between each independent variable and the dependent variable in the presence of all other variables based on the p-value. By comparing the p-value and the present significance level α, the significance of each variable is pointed out. Then the final model is determined by removing the statistically insignificant variables from the model, and its fitting accuracy is checked by the coefficient and adjusted coefficient of determinations. The final step is to check the multicollinearity problem, a case of multiple regression in which two or more independent variables are highly linearly correlated, by checking the explanatory variables' variance inflation factors (VIF).

4. RESULTS

Just mentioned above, the data contains 20640 observations and 11 independent variables. The first step is to use unit length scaling to scale independent and dependent variables that produce dimensionless regression coefficients. The scaling process is as follows:

$$X_{ij}^{0} = \frac{X_{ij}-\overline{X_j}}{s_{jj}^{1/2}}, \; i = 1, 2, \ldots, 20640; \; j = 1, \ldots, 11. \tag{1}$$

$$y_i^{0} = \frac{y_i-\bar{y}}{SST^{1/2}}, \; i = 1, 2, \ldots, 20640. \tag{2}$$

where the scaling factors are computed by

$$S_{jj} = \sum_{i=1}^{n}(X_{ij} - \overline{X_j})^2. \tag{3}$$

$$SST = \sum_{i=1}^{n}(y_i - \bar{y})^2. \tag{4}$$

After data processing, a full multiple linear regression model (M0) is built based on 11 independent variables as follows:

$$\text{M0: } \hat{Y} = 0.649X_1 + 0.105X_2 - 0.123X_3 + 0.368X_4 - 0.394X_5 + 0.182X_6 - 0.264X_7 - 0.307X_8 + 0.142X_9 + 0.0475\,X_{10} - 0.191X_{11}. \tag{5}$$

After the model is fitted, the significance of individual variables is tested with the null hypothesis $\beta_j = 0$, alternative hypothesis $\beta_j \neq 0$, and $\alpha = 0.01$. From the following table (see Table 2), the p-value in the significance testing of the variable X_{10} is 0.264, which is larger than 0.01.

TABLE II. SIGNIFICANCE OF INDIVIDUAL VARIABLES IN MODEL M0

Variables	Estimate	t-value	p-value	Significance
(Intercept)	-4.66E-18	0	1	
X1	6.49E-01	115.164	< 2E-16	***
X2	1.05E-01	21.054	< 2E-16	***
X3	-1.23E-01	-8.168	3.32E-16	***
X4	3.68E-01	14.568	< 2E-16	***
X5	-3.94E-01	-36.822	< 2E-16	***
X6	1.82E-01	7.291	3.18E-13	***
X7	-2.64E-01	-48.791	< 2E-16	***
X8	-3.07E-01	-19.607	< 2E-16	***
X9	1.42E-01	6.234	4.64E-10	***
X10	4.75E-02	1.117	0.264	
X11	-1.91E-01	-3.573	3.54E-04	***

*** significant at the 0.001 level (two-tailed)

Therefore, the null hypothesis is failed to be rejected, and it can be concluded that there is no linear relationship between the dependent variable Y and X_{10}. Thus, X_{10} can be removed from this model. By fitting the model again, the reduced model (M1) is as follows

$$\text{M1: } \hat{Y} = 0.648X_1 + 0.106X_2 - 0.123X_3 + 0.368X_4 - 0.395X_5 + 0.182X_6 - 0.266X_7 - 0.302X_8 + 0.150X_9 - 0.133X_{11}. \quad (6)$$

Then the significance of individual variables is tested again (see Table 3) and the result shows that there is a significant relationship between the dependent and the remaining independent variables. In addition, all remaining independent variables significantly improve the model and cannot be removed. Then, the coefficient and adjusted coefficient of determinations for this model are calculated as 0.6355 and 0.6354. A less high fitting accuracy could be observed from these two coefficients.

TABLE III. SIGNIFICANCE OF INDIVIDUAL VARIABLES IN MODEL M1

Variables	Estimate	t-value	p-value	Significance
(Intercept)	-4.57E-18	0	1	
X1	6.48E-01	115.175	< 2E-16	***
X2	1.06E-01	21.787	< 2E-16	***
X3	-1.23E-01	-8.139	4.20E-16	***
X4	3.68E-01	14.587	< 2E-16	***
X5	-3.95E-01	-36.983	< 2E-16	***
X6	1.82E-01	7.303	2.92E-13	***
X7	-2.66E-01	-52.255	< 2E-16	***
X8	-3.02E-01	-20.101	< 2E-16	***
X9	1.50E-01	6.84	8.13E-12	***
X11	-1.33E-01	-10.293	< 2E-16	***

*** significant at the 0.001 level (two-tailed)

The next step is to test multicollinearity, which leads to imprecise estimations of the OLS estimators, thereby affecting the understanding of how the X variable affects Y, but this is a matter of degree, not a matter of presence or absence [30]. Conclusions about the positive and negative effects of explanatory variables on Y could also be drawn even with the existence of multicollinearity. One of the useful ways to detect multicollinearity is to examine the VIF of explanatory variables, and multicollinearity exists if VIF is larger than 5. The VIF of each explanatory variable of two models (see Table 4) are shown as follows:

TABLE IV. VIF OF EACH INDEPENDENT VARIABLE

Variable	VIF in Model M0	VIF in Model M1
X1	1.795	1.794
X2	1.399	1.336
X3	12.913	12.903
X4	36.065	36.057
X5	6.495	6.465
X6	35.129	35.125
X7	1.652	1.646
X8	13.859	12.768
X9	29.504	27.084
X10	102.238	
X11	161.656	9.545

As can be seen from the table, VIF is reduced by removing the variable. However, there still exists a multicollinearity problem in model M0, which could be solved by more advanced techniques in the future study. However, as mentioned earlier, the positive or negative effects of explanatory variables on Y could be explained.

5. DISCUSSION

Based on the above regression model M1, the correlation between house value and influencing factors is discussed. As for the income for households, the model shows the house value and households income are positively correlated, verifying the hypothesis H1 proposed above. Turning to the effects of total rooms and bedrooms, the model reveals that the number of total rooms has a negative impact on house value. In contrast, the number of total bedrooms has a positive impact, which seems to conflict with hypothesis H2. However, previous studies on the total number of bedrooms were conducted under the same floor area. In contrast, in this study and previous studies on the total number of rooms, the floor area is not fixed to be approximately the same, leading to less accurate results and lack of credibility. Regarding to the housing age, the house value increases with housing age, which is not contradictory with hypothesis H4, which proposes that the correlation between house price and housing age changes in different time periods and does not state the overall trend. In terms of population and number of households, there is a negative relationship between the population and housing price. In contrast, a positive impact on the number of households can be found. This finding is inconsistent with hypothesis H5 because previous studies have discussed population density rather than population and the total number of households. As to the distance to big cities, the housing price is negatively correlated with the distance to Los Angeles and San Francisco, while positively correlated with the distance to San Diego. In addition, there is no statistically significant correlation between house value and distance to San Jose. The above results clearly verify hypothesis H6 because it is hard to tell the correlation between the housing price and a single centre under polycentric model theory. When it comes to distance to the coasts, the model

shows the distance to coasts has a significant but declining effect on house value, namely the houses in close proximity to the coast are more expensive, which proves the hypothesis H7. Such universal conclusions consistent with previous studies can be applied in future studies.

However, the fitting accuracy of the model still needs to be improved. Ridge regression is one of the mainstream methods to improve the accuracy of fitting caused by multicollinearity. In addition, spatial autoregression can be fitted to derive more accurate conclusions. In 1994, Olmo proposed that in the presence of spatial autocorrelation, the OLS estimation of the housing price prediction model parameter is inefficient [31].

6. CONCLUSION

Based on the background of the US economic recession in 1990, this paper explores the influencing factors of California housing prices in 1990. Using the housing price impact theory and multiple regression method, it is concluded that the California housing prices are indeed affected by buyers' income, the structure of the house itself, population density and geographic surroundings, such as the distance to major cities the coast. Specifically, the California housing price is positively correlated with resident income, housing age, the number of total bedrooms, households, and distance to San Diego. At the same time, it is negatively related with the number of total rooms, population, distance to the coast and Los Angeles. In addition, there is no statistically significant relationship between the house value and distance to San Jose. Since this paper focuses on California housing prices during the recession, it can be referred to when California experience a similar economic situation. At the same time, the common conclusions consistent with previous studies can also be used as a reference for future research on housing prices in a certain area. It can also provide real estate developers with decisions about the location, structure, and price for the houses to be built and sold in the future. It should be noted that this paper avoids other factors that affect housing prices because the data used in this paper is not comprehensive enough. For instance, neighborhood quality is a frequently explored topic in housing prices, including public school achievement, major crime rates, and the educational level of neighbors. More influencing factors can be discussed in future studies.

REFERENCES

[1] Kamery, Robert H. "A brief review of the Recession of 1990–1991." Proceedings of the Academy of Legal, Ethical and Regulatory Issues 8.2 (2004): 61.

[2] Walsh, Carl E. "What caused the 1990-1991 recession?." Economic Review-Federal Reserve Bank of San Francisco 2 (1993): 33.

[3] Singleton, Christopher J. "Industry employment and the 1990-91 recession." Monthly labor review 116.7 (1993): 15-25.

[4] Johnson, Hans P., Rosa Maria Moller, and Michael Dardia. In Short Supply?: Cycles and Trends in California Housing. San Francisco: Public Policy Institute of California, 2004.

[5] Abraham, Jesse M., and Patric H. Hendershott. "Bubbles in metropolitan housing markets." Journal of Housing Research 7.2 (1996): 191-207.

[6] Kraft, Holger, and Claus Munk. "Optimal housing, consumption, and investment decisions over the life cycle." Management Science 57.6 (2011): 1025-1041.

[7] Pace, R. Kelley, and Ronald Barry. "Sparse spatial autoregressions." Statistics & Probability Letters 33.3 (1997): 291-297.

[8] Lamont, Owen, and Jeremy C. Stein. "Leverage and house-price dynamics in US cities." (1997).

[9] [Sani, M. M., and A. Rahim. "Price to income ratio approach in housing affordability." Journal of Economics, Business and Management 3.12 (2015): 1190-1193.
[10] Gallin, Joshua. "The Long-Run Relation between House Prices and Income: Evidence from Local Housing Markets." Washington, DC: Board of Governors of the Federal Reserve System, Finance and Economics Discussion Paper Series 2003.17 (2003).
[11] Li, Mingche M., and H. James Brown. "Micro-neighborhood externalities and hedonic housing prices." Land economics 56.2 (1980): 125-141.
[12] Selim, Hasan. "Determinants of house prices in Turkey: Hedonic regression versus artificial neural network." Expert systems with Applications 36.2 (2009): 2843-2852.
[13] Kain, John F., and John M. Quigley. "Measuring the value of housing quality." Journal of the American statistical association 65.330 (1970): 532-548.
[14] C.Y. Jim, ,and Wendy Y. Chen."Value of scenic views: Hedonic assessment of private housing in Hong Kong." Landscape and Urban Planning 91.4(2009): doi:10.1016/j.landurbplan.2009.01.009.
[15] Liao, Wen-Chi, and Xizhu Wang. "Hedonic house prices and spatial quantile regression." Journal of Housing Economics 21.1 (2012): 16-27.
[16] Rubin, Geoffrey M. "Is housing age a commodity? Hedonic price estimates of unit age." Journal of Housing Research (1993): 165-184.
[17] Goodman, Allen C., and Thomas G. Thibodeau. "Age-related heteroskedasticity in hedonic house price equations." Journal of Housing Research (1995): 25-42.
[18] Clapp, John M., and Carmelo Giaccotto. "Residential hedonic models: A rational expectations approach to age effects." Journal of Urban Economics 44.3 (1998): 415-437.
[19] Coulson, N. Edward, and Daniel P. McMillen. "Estimating time, age and vintage effects in housing prices." Journal of Housing Economics 17.2 (2008): 138-151.
[20] Sharpe, Ryan G. The effects of population density on home prices. MS thesis. 2019., unpublished.
[21] Alonso, William. Location and land use. Harvard University Press, 1964.
[22] Mills, Edwin S. "Transportation and patterns of urban development." American Economic Review 57.2 (1967): 197-210.
[23] Muth, Richard F. "Cities and housing university of chicago press." Chicago, IL (1969).
[24] Berry, Brian JL. "Ghetto expansion and single-family housing prices: Chicago, 1968–1972." Journal of urban Economics 3.4 (1976): 397-423.
[25] Richardson, Harry W. "Monocentric vs. policentric models: The future of urban economics in regional science." The Annals of Regional Science 22.2 (1988): 1-12.
[26] Dubin, Robin A., and Chein-Hsing Sung. "Spatial variation in the price of housing: rent gradients in non-monocentric cities." Urban Studies 24.3 (1987): 193-204.
[27] Boarnet, Marlon G., and Saksith Chalermpong. "New highways, house prices, and urban development: A case study of toll roads in Orange County, CA." Housing policy debate 12.3 (2001): 575-605.
[28] Conroy, Stephen J., and Jennifer L. Milosch. "An estimation of the coastal premium for residential housing prices in San Diego County." The Journal of Real Estate Finance and Economics 42.2 (2011): 211-228.
[29] Fedesoriano. “California Housing Prices Data - Extra Features.” Kaggle, 3 July 2021, www.kaggle.com/fedesoriano/california-housing-prices-data-extra-features.
[30] Paul, Ranjit Kumar. "Multicollinearity: Causes, effects and remedies." IASRI, New Delhi 1.1 (2006): 58-65.
[31] Chica Olmo, Jorge. "Spatial estimation of housing prices and locational rents." Urban studies 32.8 (1995): 1331-1344.