

基于Eviews软件的上海住宅商品房平均销售价格影响因素的研究分析

乔飞燕¹, 乔燕铭²

¹上海工程技术大学管理学院, 上海

²河南财政金融学院会计学院, 河南 郑州

收稿日期: 2023年5月5日; 录用日期: 2023年7月6日; 发布日期: 2023年7月13日

摘要

目的/意义: 高房价引发诸多社会问题, 上海市区住宅商品房平均销售价格的研究就显得尤为重要。方法/过程: 本文通过Eviews软件分析所建多元线性模型, 从微观和宏观、需求和供给方面选取变量研究中国上海市区住宅商品房平均销售价格的影响因素。数据为2002~2019年间的的数据, 共18个样本, 摘自国家统计年鉴。对所建模型进行多重共线性检验、异方差和自相关等检验, 解决模型的多重共线性等问题。结果/结论: 研究结果发现在岗职工平均工资对上海市区住宅商品房平均销售价格有显著影响, 据此对国家调控房地产价格提供一些建议(选定 $\alpha = 0.05$)。

关键词

住宅商品房平均销售价格, 线性回归模型, 异方差, 自相关, 多重共线性

Research and Analysis of the Influencing Factors of the Average Sales Price of Residential Commercial Housing in Shanghai Based on Eviews Software

Feiyan Qiao¹, Yanming Qiao²

¹School of Management, Shanghai University of Engineering Science, Shanghai

²School of Accounting, Henan Finance University, Zhengzhou Henan

Received: May 5th, 2023; accepted: Jul. 6th, 2023; published: Jul. 13th, 2023

文章引用: 乔飞燕, 乔燕铭. 基于 Eviews 软件的上海住宅商品房平均销售价格影响因素的研究分析[J]. 建模与仿真, 2023, 12(4): 3522-3531. DOI: 10.12677/mos.2023.124324

Abstract

Purpose/Significance: High housing prices cause many social problems, so the study of the average sales price of residential commercial housing in Shanghai is particularly important. **Method/Process:** This paper analyzes the multivariate linear model through Eviews software, selects variables from the micro and macro, demand and supply aspects to study the influencing factors of the average sales price of residential commercial housing in Shanghai, China. Data are from 2002 to 2019, with a total of 18 samples, extracted from the National Statistical Yearbook. The multicollinearity test, heteroscedasticity and autocorrelation tests are carried out on the established model to solve the problems of multicollinearity of the model. **Results/Conclusion:** The results show that the average salary of on-the-job employees has a significant impact on the average sales price of residential commercial housing in Shanghai, and provide some suggestions for the state to regulate real estate prices (selected $\alpha = 0.05$).

Keywords

Average Sales Price of Residential Commercial Housing, Linear Regression Models, Heteroscedasticity, Autocorrelation, Multicollinearity

Copyright © 2023 by author(s) and Hans Publishers Inc.

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

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



Open Access

1. 引言

2019年,上海商品住宅供应726.85万平方米,较18年同期下降5.39%,为近五年同期中位水平,月均供应量为60.57万方。尽管中央已经出台各项楼市调控的政策,房价仍然未下跌[1]。房价问题一直是上海等发达地区市民关心的问题。房价居高不下的问题,让即将步入社会的大学生们为赖以栖居的住房乃至家庭和婚姻的幸福而担忧。时至今日,“蜗居”仍然是大学生们讨论的热点问题,从另一方面体现了高房价情形之下,大学毕业生的生活压力。甚至很多上海市民,数年来只能蜗居在几十甚至十几平方米的小屋里,生活质量很难有确切保障。社会快速发展的压力也让人们愈加烦躁,贫富差异更加明显,社会在这样的情况下就会慢慢滋生诸多矛盾。因此,上海市区住宅商品房平均销售价格的研究就显得尤为重要。

首先,通过研究能够发现影响上海市区住宅商品房平均销售价格的主要因素和预测分析,可以了解将来上海市区住宅商品房平均销售价格的发展趋势[2]。广大市民急切希望了解住宅商品房平均销售价格,方便与自身是选择早日买房或是与自身经济实力相比择日另买更加划算。其次,本文是在对前人探究上海市区住宅商品房平均销售价格影响因素的基础上完成的,旨在探索是否有其他更为显著的因素在发挥影响。因此,在剔除了上海居民消费价格指数和上海人均GDP后,选取其他与现实生活更密切相关,广为人知的因素探究上海市区住宅商品房平均销售价格的影响因素。本文从2002年为时间起点,分析了2002到2019年这18年的数据房地产开发住宅投资额,年末常住人口,在岗职工平均工资,房地产开发企业住宅竣工房屋住宅面积,这几个因素对上海市区住宅商品房平均销售价格的影响。通过建立并分析模型,进一步探讨影响上海市区住宅商品房平均销售价格这一社会问题。

2. 理论基础与模型构建

2.1. 数据与变量的选取

本文从《国家统计年鉴》选取了上海住宅商品房平均销售价格 18 年的数据。

一般而言, 微观上房地产开发住宅投资额越大往往会使得住宅商品房平均销售价格越大(正相关), 宏观上年末常住人口越多往往住宅商品房平均销售价格会越大(正相关), 在需求因素方面在岗职工平均工资越多住宅商品房平均销售价格会越大(正相关), 在供给因素方面房屋竣工面积越多住宅商品房平均销售价格会越小(负相关) [1]。因此对房地产开发住宅投资额、年末常住人口、在岗职工平均工资、房屋竣工面积这四个影响因素的研究对制定房价政策具有重要的意义, 理论上我们假设在岗职工平均工资是最显著的影响因素。

首先, 建立了多元线性回归模型[3]如下:

$$Y = B_0 + B_1X_1 + B_2X_2 + B_3X_3 + B_4X_4 + C \tag{1}$$

其中, 被解释变量: 住宅商品房平均销售价格(Y, 元/平方米)。解释变量: 房地产开发住宅投资额(X_1 , 亿元), 年末常住人口(X_2 , 万人), 在岗职工平均工资(X_3 , 元), 房地产开发企业住宅竣工房屋住宅面积(X_4 , 万平方米)和随机误差项(C)。所选取 2002~2019 年各个变量的数据如图 1 所示。

	房地产开发住宅投资额(亿元)	年末常住人口(万人)	在岗职工平均工资(元)	房地产开发企业住宅竣工房屋住宅面积(万平方米)	住宅商品房平均销售价格(元/平方米)
	X_1	X_2	X_3	X_4	Y
2002年	567.76	1334.23	23959	1708.1	4007
2003年	676.28	1341.77	27304	2139.99	4989
2006年	835.63	1368.08	41188	2699.11	7039
2007年	857.53	1378.86	49310	2752.45	8253
2008年	875.46	1391.04	56565	1801.45	8115
2004年	900.67	1352.39	30085	3076.19	5761.21
2009年	918.68	1400.7	63549	1508.81	12364
2005年	920.84	1360.26	34345	2739.91	6698
2010年	1229.83	1412	71874	1396.05	14290
2012年	1451.94	1426.93	80191	1609.13	13869.88
2011年	1465.52	1419.36	77031	1645.47	13565.83
2013年	1615.51	1432.34	91477	1417.41	16192
2014年	1724.65	1438.69	100623	1535.55	16415
2015年	1813.32	1442.97	109279	1588.95	21501
2016年	1965.43	1450	120503	1532.88	25910
2017年	2152.4	1455.13	130765	1862.74	24866
2018年	2225.92	1462.38	142983	1730.27	28981.11
2019年	2318.13	1469.3	151772	1453.28	32926

Figure 1. Data on residential investment in real estate development from 2002 to 2019

图 1. 2002~2019 年房地产开发住宅投资额等数据

2.2. 散点图

通过 Eviews 软件分别做被解释变量(Y)与解释变量(X_1, X_2, X_3, X_4)的散点图(见图 2~图 5)。

从散点图可以看出, 影响因素 X_1, X_2 和 X_3 随着住宅商品房平均销售价格 Y 的上升而增加, 且近似于线性关系。影响因素 X_4 随着住宅商品房平均销售价格 Y 的上升呈现非线性关系。

2.3. 回归方程

运用 Eviews 的最小二乘法程序, 输出的结果如图 6 所示。

由图 6, 得到的回归方程为

$$Y = 51462.93 + 2.484022X_1 - 41.60827X_2 + 0.223468X_3 + 0.562526X_4 \tag{2}$$

$$(1.386233) (0.570214) (-1.482637) (4.184236) (0.599594)$$

残差平方和为: Sum squared resid = 30067939

回归标准差为: S.E. of regression = 1520.828

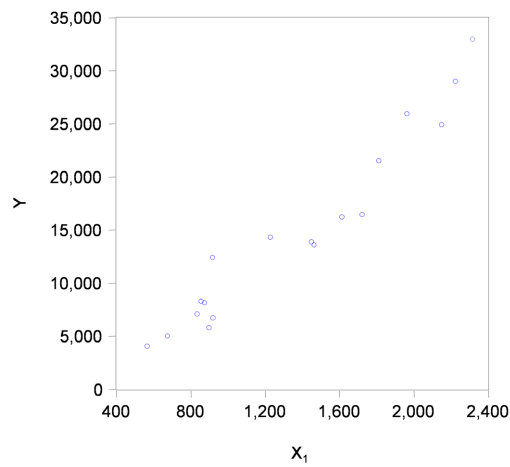


Figure 2. Scatterplot of X_1 and Y
图 2. X_1 与 Y 的散点图

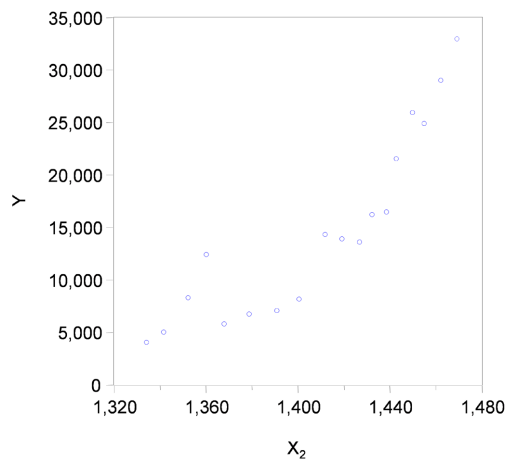


Figure 3. Scatterplot of X_2 and Y
图 3. X_2 与 Y 的散点图

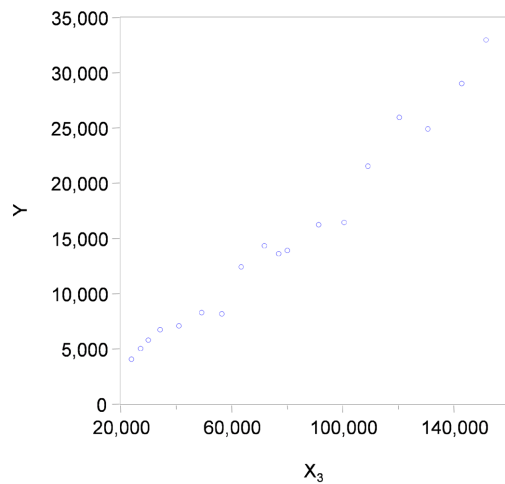


Figure 4. Scatterplot of X_3 and Y
图 4. X_3 与 Y 的散点图

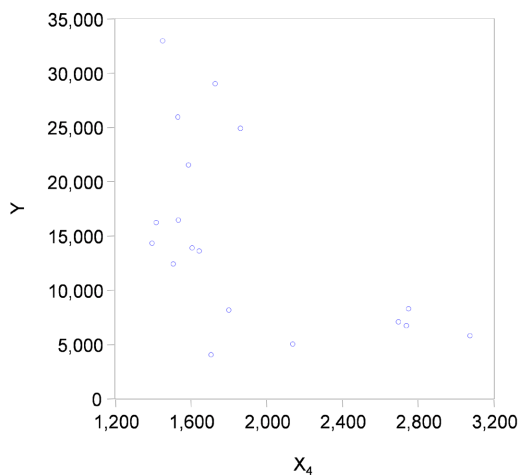


Figure 5. Scatterplot of X_4 and Y
图 5. X_4 与 Y 的散点图

Dependent Variable: Y

Method: Least Squares

Date: 12/21/20 Time: 10:57

Sample: 2002 2019

Included observations: 18

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	51462.93	37124.29	1.386233	0.1890
X_1	2.484022	4.356300	0.570214	0.5783
X_2	-41.60827	28.06370	-1.482637	0.1620
X_3	0.223468	0.053407	4.184236	0.0011
X_4	0.562526	0.938178	0.599594	0.5591
R-squared	0.977226	Mean dependent var		14763.50
Adjusted R-squared	0.970219	S.D. dependent var		8812.732
S.E. of regression	1520.828	Akaike info criterion		17.72203
Sum squared resid	30067939	Schwarz criterion		17.96936
Log likelihood	-154.4983	Hannan-Quinn criter.		17.75613
F-statistic	139.4585	Durbin-Watson stat		1.735157
Prob(F-statistic)	0.000000			

Figure 6. Results of a multiple linear regression
图 6. 一次多元线性回归结果

$R^2 = 0.977226$, 修正后的 $R^2 = 0.970219$, $n = 18$, $df = 13$, $F = 139.4585$ 。

2.4. 经济意义检验

由经济分析可知, 住宅商品房平均销售价格(Y)与房地产开发住宅投资额(X_1), 年末常住人口(X_2), 在岗职工平均工资(X_3)应呈现正相关关系, 与房地产开发企业住宅竣工房屋住宅面积(X_4)应呈现负相关关系。但是, 多元回归模型中年末常住人口(X_2)前的系数符号为负, 证明成负相关关系, 住宅竣工房屋住宅面积(X_4)前的系数为正, 证明成正相关关系, 这与经济学意义不符合, 故认为原模型存在错误, 需要进行调整。

3. 检验与修正模型

3.1. 统计检验

由上述回归结果可知, 可决系数 R^2 与修正的可决系数的值都接近于 1, 表明模型的拟合度很高。在 $\alpha = 0.05$ 的显著性水平下, X_3 的 t 统计量值通过了显著性检验, X_1, X_4 的 P 值分别为 0.5783, 0.5591, 明显大于 0.05, 所以结果不显著。以上表明所建立的模型存在误差, 需要进行修正。

3.2. 多重共线性检验

X_1, X_2, X_3, X_4 的相关系数如图 7 所示。

	X_1	X_2	X_3	X_4
X_1	1.000000	0.950962	0.980724	-0.526643
X_2	0.950962	1.000000	0.928564	-0.534784
X_3	0.980724	0.928564	1.000000	-0.597489
X_4	-0.526643	-0.534784	-0.597489	1.000000

Figure 7. Correlation coefficients of X_1, X_2, X_3, X_4

图 7. X_1, X_2, X_3, X_4 的相关系数

由图 7 可知, 各相关变量相互之间的相关系数较高, 所以存在严重的多重共线性问题, 为了检验和处理多重共线性, 采用修正 Frisch 法。

1) 对 Y 分别做关于 X_1, X_2, X_3, X_4 做最小二乘回归, 得

$$(1) Y_1 = -5531.093 + 14.90089X_1$$

$$(-3.821173) \quad (15.14270)$$

$R^2 = 0.934774$, 修正后的 $R^2 = 0.930698$, $DW = 1.168786$, $F = 229.3015$

$$(2) Y_2 = -241207.0 + 181.8516X_2$$

$$(32005.94) \quad (22.72811)$$

$R^2 = 0.800047$, 修正后的 $R^2 = 0.787550$, $DW = 1.597849$, $F = 64.01882$

$$(3) Y_3 = -1771.149 + 0.212164X_3$$

$$(-2.247686) \quad (23.56104)$$

$R^2 = 0.971985$, 修正后的 $R^2 = 0.970234$, $DW = 1.356840$, $F = 555.1227$

$$(4) Y_4 = 32199.58 + -9.177489X_4$$

$$(4.831704) \quad (-2.714044)$$

$$R^2 = 0.315245, \text{ 修正后的 } R^2 = 0.272448, DW = 0.294834, F = 7.366035$$

其中括号内的数字是 t 值, 根据经济理论分析和回归结果易知 X_3 是最重要的解释变量, 所以选取第 3 个回归方程为基本回归方程。

2) 加入房地产开发住宅投资额 X_1 , 对 Y 关于 X_1, X_3 作最小二乘回归, 得

$$Y = -1765.015 - 0.020812X_1 + 0.212449X_3$$

$$(-1.365081) \quad (-0.006106) \quad (4.463610)$$

$$R^2 = 0.971985, \text{ 修正后的 } R^2 = 0.968250, DW = 1.356410, F = 260.2144$$

可以看出, 在加入 X_1 后, 拟合优度 R^2 增加不显著, 修正后的 R^2 有所减小, 并且 X_1 和 X_3 系数均不显著说明存在严重的多重共线性, 比较 X_1 和 X_3 , 在岗职工平均工资比房地产开发住宅投资额对住宅商品房平均销售价格的影响大, 所以在模型中保留解释变量 X_3 略去解释变量 X_1 。

3) 加入年末常住人口 X_2 , 对 Y 关于 X_2, X_3 作最小二乘回归, 得

$$Y = 39497.99 - 31.00655X_2 + 0.242639X_3$$

$$(1.331686) \quad (-1.391865) \quad (10.29016)$$

$$R^2 = 0.975189, \text{ 修正后的 } R^2 = 0.971881, DW = 1.423973, F = 294.7897$$

可以看出, 在加入 X_2 后, 虽然提高了回归方程的拟合优度, 使 R^2 , 修正后的 R^2 均有所增加, 但是 X_2 参数估计值的符号为负, 与实际经济情况不符。比较 X_2 和 X_3 , 在岗职工平均工资比年末常住人口对住宅商品房平均销售价格的影响大, 所以在模型中保留解释变量 X_3 略去解释变量 X_2 。

4) 加入房地产开发企业住宅竣工房屋住宅面积 X_4 , 对 Y 关于 X_3, X_4 作最小二乘回归, 得

$$Y = -3533.812 + 0.217681X_3 + 0.701439X_4$$

$$(-1.530993) \quad (19.17856) \quad (0.813628)$$

$$R^2 = 0.973169, \text{ 修正后的 } R^2 = 0.969592, DW = 1.496035, F = 272.0287$$

可以看出, 在加入 X_4 后, 拟合优度 R^2 增加不显著, 修正后的 R^2 有所减小, 并且 X_4 参数估计值的符号为负, 与实际经济情况不符。比较 X_3 和 X_4 , 在岗职工平均工资比房地产开发企业住宅竣工房屋住宅面积对住宅商品房平均销售价格的影响大, 所以在模型中保留解释变量 X_3 略去解释变量 X_4 。

综上所述, 得到 Y 关于 X_3 的回归方程, 运用 Eviews 的最小二乘法程序, 输出的结果如图 8 所示。

$$Y = -1771.149 + 0.212164X_3$$

$$(-2.247686) \quad (23.56104)$$

$$R^2 = 0.971985, \text{ 修正后的 } R^2 = 0.970234, DW = 1.356840, F = 555.1227$$

该模型中系数均显著, 且符号正确, 仅余一个解释变量, 因此不再存在高度线性关系。也说明房地产开发住宅投资额, 年末常住人口, 与房地产开发企业住宅竣工房屋住宅面积不适合用来作为住宅商品房平均销售价格的解释变量。

3.3. 异方差分析

观察图 8 的残差 e_i 取值, 好像随 X_3 的变化而变化, 怀疑模型存在异方差性, 接下来用怀特检验随机误差项的异方差性。

运用 Eviews 软件, 输出的结果如图 9 所示。

因为式中只含有一个解释变量, 所以 White 检验辅助回归式中应该包括两个解释变量。辅助回归式估计结果如下:

$$u_3^2 = -411171.9 - 0.000111X_3 + 42.59800X_3^2$$

$$(-0.149376) \quad (-0.257417) \quad (0.568533)$$

Dependent Variable: Y
 Method: Least Squares
 Date: 12/21/20 Time: 16:18
 Sample: 2002 2019
 Included observations: 18

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-1771.149	787.9877	-2.247686	0.0390
X ₃	0.212164	0.009005	23.56104	0.0000
R-squared	0.971985	Mean dependent var		14763.50
Adjusted R-squared	0.970234	S.D. dependent var		8812.732
S.E. of regression	1520.444	Akaike info criterion		17.59583
Sum squared resid	36987980	Schwarz criterion		17.69476
Log likelihood	-156.3625	Hannan-Quinn criter.		17.60947
F-statistic	555.1227	Durbin-Watson stat		1.356840
Prob(F-statistic)	0.000000			

Figure 8. Multiple linear regression results after model correction
图 8. 模型修正后的一次多元线性回归结果

Heteroskedasticity Test: White

F-statistic	1.209817	Prob. F(2,15)	0.3258
Obs*R-squared	2.500249	Prob. Chi-Square(2)	0.2865
Scaled explained SS	1.543794	Prob. Chi-Square(2)	0.4621

Test Equation:
 Dependent Variable: RESID^2
 Method: Least Squares
 Date: 12/21/20 Time: 16:09
 Sample: 2002 2019
 Included observations: 18

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-411171.9	2752592.	-0.149376	0.8832
X ₃ ²	-0.000111	0.000433	-0.257417	0.8004
X ₃	42.59800	74.92613	0.568533	0.5781
R-squared	0.138903	Mean dependent var		2054888.
Adjusted R-squared	0.024090	S.D. dependent var		2643446.
S.E. of regression	2611412.	Akaike info criterion		32.53969
Sum squared resid	1.02E+14	Schwarz criterion		32.68809
Log likelihood	-289.8572	Hannan-Quinn criter.		32.56015
F-statistic	1.209817	Durbin-Watson stat		2.618020
Prob(F-statistic)	0.325755			

Figure 9. White test output
图 9. 怀特检验输出结果

$$R^2 = 0.138903, T = 18$$

因为 $TR^2 = 18 \times 0.138903 = 2.500254$, 且概率大于 0.05, 则拒绝原假设, 表明模型中随机误差不存在异方差, 不需要对模型异方差性修正。

3.4. 自相关检验

3.4.1. DW 检验

对模型进行 DW 检验, 因为 $T = 18$, 解释变量的个数为 1, 根据 DW 图表得到下限临界值 $d_L = 1.16$, $d_U = 1.39$, $DW = 1.356840$, 取值在 (d_L, d_U) 之间, 这种检验没有结论, 即不能判别 u_3 是否存在一阶自相关。

3.4.2. LM 检验

运用 Eviews 软件, 输出的结果如图 10 所示。

可知 $nR^2 = 2.10444 < 5.9915$, 并且概率值 P 远大于 0.05, 说明不存在自相关, 不需要对模型进行修正。

Breusch-Godfrey Serial Correlation LM Test:

F-statistic	0.926742	Prob. F(2,14)	0.4188
Obs*R-squared	2.104441	Prob. Chi-Square(2)	0.3492

Test Equation:

Dependent Variable: RESID

Method: Least Squares

Date: 12/21/20 Time: 16:50

Sample: 2002 2019

Included observations: 18

Presample missing value lagged residuals set to zero.

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	11.27271	820.5523	0.013738	0.9892
X ₃	-7.97E-05	0.009701	-0.008211	0.9936
RESID(-1)	0.341021	0.289432	1.178242	0.2583
RESID(-2)	-0.260972	0.298317	-0.874814	0.3964
R-squared	0.116913	Mean dependent var		2.93E-12
Adjusted R-squared	-0.072319	S.D. dependent var		1475.047
S.E. of regression	1527.453	Akaike info criterion		17.69372
Sum squared resid	32663590	Schwarz criterion		17.89158
Log likelihood	-155.2435	Hannan-Quinn criter.		17.72100
F-statistic	0.617828	Durbin-Watson stat		1.703704
Prob(F-statistic)	0.614851			

Figure 10. LM test output

图 10. LM 检验输出结果

4. 结论

上文首先构建被解释变量: 住宅商品房平均销售价格(Y , 元/平方米), 解释变量: 房地产开发住宅投资额(X_1 , 亿元), 年末常住人口(X_2 , 万人), 在岗职工平均工资(X_3 , 元), 房地产开发企业住宅竣工房屋住宅面积(X_4 , 万平方米)和随机误差项(C)的多元线性回归模型, 然后分别进行了经济意义检验和统计检验, 发现该模型存在错误, 需要调整。用 Eviews 软件对模型进行多重共线性检验, 剔除不显著的影响因素后, 对模型进行异方差分析和自相关检验, 得到修正后的模型不需要进行异方差性修正和自相关性修正。

综上所述, 可知最后建立的模型为

$$Y = -1771.149 + 0.212164X_3 \\ (-2.247686) (23.56104)$$

$$R^2 = 0.971985, \text{修正后的 } R^2 = 0.970234, DW = 1.356840, F = 555.1227$$

该模型表示, 当在岗职工平均工资每增加一元时, 住宅商品房平均销售价格将上升 0.212164 元/平方米。

不足之处, 本文所选取的住宅商品房平均销售价格影响因素较少, 因试图将住宅商品房人均居住面积影响因素与住宅商品房平均销售价格影响因素比较, 观察前者影响因素是否适用于后者, 导致所选影响因素较为片面。诸多因素并未加以考虑, 例如: 狭义货币供应量, 上海市区人均实际 GDP, 货币利率等等。

本文最终确定了影响住宅商品房平均销售价格的一个因素, 即在岗职工平均工资, 接下来根据所建立的最终模型给出相关的政策建议。

在岗职工平均工资是根据现行统计制度, 城镇单位在岗职工工资统计范围包括国有单位、城镇集体单位、以及联营经济、股份制经济、外商投资经济、港澳台投资经济单位, 尚未包括城镇的私营企业和个体工商户。而工资总额则是指这些单位在一定时期内直接支付给本单位全部职工的劳动报酬总额, 包括计时工资、计件工资、奖金、津贴和补贴、加班加点工资、特殊情况下支付的工资, 不论是否计入成本, 也不论是以货币形式还是以实物形式支付, 均包括在内。

国家调控房地产价格的手段不能只有提高在岗职工平均工资, 因为当在岗职工平均工资提高时, 住宅商品房平均销售价格也在提高, 不能很好地解决住宅商品房平均销售价格过高的问题, 应当从其他方面入手, 如改变货币供应量等等。

致 谢

这次基于 Eviews 软件的上海住宅商品房平均销售价格影响因素的研究分析, 使我对 Eviews 软件如何制作散点图, 运用最小二乘法做回归方程, 并且对异方差、自相关性、多重共线性的检验如何操作, 结果如何处理, 有了更深一步的了解。同时, 对上海住宅商品房平均销售价格影响因素有了初步的认识。感谢我的指导老师吴常群和给予我能够利用图书馆知网免费查询参考论文的学校。感谢给我提供诸多帮助的同学。

参考文献

- [1] 王晴, 朱家明. 基于 EViews 的我国房地产价格影响因素计量分析[J]. 洛阳师范学院学报, 2017, 36(5): 64-68.
- [2] 李翔. 基于 Eviews 软件分析的上海市区人均住房居住面积影响因素研究[J]. 安阳工学院学报, 2013, 12(1): 59-62.
- [3] 张晓峒. 计量经济学基础[M]. 第 4 版. 天津: 南开大学出版社, 2014.