

Analysis of the Influence of Multiple Linear Regression on Construction Price

Yalu Ping

School of Statistics and Mathematics, Yunnan University of Finance and Economics, Kunming Yunnan
Email: p_yalu@163.com

Received: Dec. 19th, 2019; accepted: Jan. 1st, 2020; published: Jan. 8th, 2020

Abstract

After the reform and development, China's construction industry has made great progress that made an important contribution to the rapid growth of the national economy and has become a pillar industry of the national economy, which plays an important role in the national economy and social development. So, it is of great significance to study the influencing factors of construction price for the development of construction industry. Through building multiple linear model, we obtain the result that the main factors are Wholesale price index (WPI) of building materials for the base year, Land price index for the base year, Total floor areas of building permits issued by the city/municipality, the interest rate for loan in a time resolution and Stock market index.

Keywords

Multiple Linear Model, Least Squares Estimation, Construction Price

基于多元线性回归对建筑价格的影响分析

平雅露

云南财经大学统计与数学学院, 云南 昆明
Email: p_yalu@163.com

收稿日期: 2019年12月19日; 录用日期: 2020年1月1日; 发布日期: 2020年1月8日

摘要

改革开放以后, 我国建筑业取得了长足的进步, 为国民经济高速增长做出了重要贡献, 其产业影响力日益突出, 已成为国民经济支柱产业, 对于国民经济和社会发展具有重要作用。因此探讨建筑价格的影响因素对建筑业以及经济的发展具有重要的意义。本文通过建立多元线性模型, 得出影响建筑价格的主要

因素是基准年建筑材料批发价格指数(WPI)、基准年土地价格指数、城市/市政当局颁发建筑许可证的总建筑面积、一段时间内贷款的利率和股票市场指数。

关键词

多元回归模型, 最小二乘估计, 建筑价格

Copyright © 2020 by author(s) and Hans Publishers Inc.

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

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



1. 引言

改革开放以后, 随着市场经济制度的确立, 我国建筑业取得了长足的进步, 为国民经济高速增长做出了重要贡献, 其产业影响力日益突出, 已成为国民经济支柱产业, 对于国民经济和社会发展具有重要作用。我国加入 WTO 后, 各行各业都面临着机遇和挑战, 作为国民经济产业支柱之一的建筑业也不例外。2015~2020 年, 是我国实现全面建设小康社会和现代化建设第三步战略目标的关键时期, 也是经济社会发展的一个重要战略机遇期。这一时期, 我国经济社会发展的大环境将有所改变。在这样的背景下, 研究建筑价格的影响因素对建筑业的发展具有重要的意义。

建设项目是劳动生产产品, 在市场经济条件下同时又是市场交易商品。产品的成本价和交易价是建设项目价格确定与控制的双步骤、双要素。建设项目成本价是计划经济条件下工程计价的主导定价方式, 项目设计、施工属于技术、静态和单向抉择过程, 归于科学范畴; 建设项目交易价, 即供需双方的交易、竞争定价, 属于商品市场博弈与交易的主观意愿评判、交换及双向抉择的动态过程, 归于社会学范畴[1]。

价格问题始终是影响建筑业成长发展和竞争力的至关重要的问题。建筑商品价格的决定因素不仅在于其自身的价值, 而且在于其对于消费者所产生的效用。

2. 文献综述

对于影响建筑物价格的因素, 国内外学者做了大量的研究。国外学者根据研究对象所处的经济环境、地理位置、所关注的角度不同, 得到不同的结论。Kilian, W. F. and Snyman [2]和 Taylor and Bowen [3]在研究 BER (bureau for Economic Research, University of Stellenbosch)的建造成本指数 BCI 都得出建筑需求对建筑价格有决定性的作用的结论。Fellows 用相关分析和回归技术鉴别出一组影响建筑价格的先导指标, 即利息率、建筑业投资、授予合同数量、建筑业产出及建造成本式建筑价格的主要影响因素[4]。Skitmore, R. M.认为建筑业的需求和供给共同决定了建筑价格, 且建筑需求与建筑价格正相关[5]。Berk, J. M. and Bikker 认为影响建筑价格的经济变量分为三大类: 能引起经济活动水平的变动的经济变量, 如利息率, 通货膨胀率等; 研究对象的先导变量, 如建筑业产出等; 对未来有一定预示作用的经济变量, 如股票价格指数等[6]。Akintola Akintoye, Paul Bowen and Cliff Hardcastle 认为失业率、建筑业产出量、制造业价格成本比率构成建筑业价格的连续性、先导性经济变量; 建造成本指数、国民生产总值构成建筑价格的同期经济变量, 而常用的宏观经济变量如名义利率、通货膨胀率等与建筑价格之间没有明显的相关关系[7]。

国内的研究也是丰富多彩, 主要从定性和定量两方面研究了影响建筑价格的因素。定性方面, 中国建筑产品的价格水平普遍受到建筑业投资、人均国民收入、建筑业产值、工业品价格与成本比率等宏观

经济变量的影响。吴韵侠, 吴汉雄等认为政策法规性因素, 地区性与市场性因素, 设计因素, 施工因素, 人员素质因素是影响建筑价格的主要因素[8]。陈光林认为, 影响公路建设项目造价的三因素为设计因素, 管理因素和施工因素[9]。郭计云等把材料价格, 工程量因素, 人为因素, 概算指标, 定额, 国家及主管部门的政策性因素等作为影响建筑价格的主要因素。陈光认为影响建筑价格的主要因素有经济因素, 如经济发展状况, 居民收入状况, 通货膨胀及物价状况等; 供需关系; 政策因素, 如土地制度政策, 住房制度政策, 税收政策等[10]。定量方面, 尚梅以经济分析为基础, 对定性分析的 29 个影响因素作 Granger 因果检验, 鉴别出建筑业产值、广义货币供给量、建筑业劳动生产率、等七个变量构成建筑价格的同步预测变量, 建筑业劳动生产率、广义货币供给量建筑产品出厂价格等构成建筑价格的滞后先导变量[11]。尚梅, 张金锁基于修正的 Granger 因果关系, 结合中国建筑产品价格时间序列的特征, 引入了表征国内建筑市场成熟度的虚拟变量, 发现广义货币供给量、建筑材料出厂价格指数、建筑业劳动生产率、已完大中型建设项目个数及城镇居民消费价格指数构成中国建筑产品价格连续的先导变量[12]。王振鑫运用灰色系统理论建立了灰色 GM(1,1)模型预测价格, 预测精度达到了 95%左右[13]。钱永峰运用灰色系统理论, 构建了城镇建筑造价指数的灰色系统预测模型, 对莫以城镇建筑造价进行了中、长期预测[14]。

3. 模型建立

3.1. 模型建立与参数估计

线性回归是研究变量间相互关系的一种有力工具, 是利用数理统计中回归分析, 来确定两种或两种以上变量间相互依赖的定量关系的一种统计分析方法。

为了更好找到变量之间明确的关系以便预测未来, 本文引用线性回归方程。

$$Y = X\beta + e$$

其中, Y 为因变量或响应变量, X 称为自变量, e 为随机误差, 满足 $E(e) = 0$, $Cov(e) = \sigma^2 I$ 。

本文采用最小二乘法[15]来估计模型中的未知参数 β 。

3.2. 变量选择

模型选择是近年来统计学和计量经济学领域的核心问题之一。在众多的备选模型中如何选出最好的模型一直是统计学家和计量经济学家关注的问题。对于模型变量的选择, 统计学家从不同角度提出了多种回归自变量的选择, 常见的有平均残差平方和准则(RMS_q), C_p 准则, AIC 准则等等。本文将使用 AIC 准则选取变量。

AIC 准则, 又称为 Akaike 信息量准则, 是利用极大似然原理选择回归模型自变量。即:

$$AIC = n \ln(SS_{eq}) + 2q$$

其中, SS_{eq} 为选模型下的残差平方和, $SS_{eq} = Y'(I - P_{X_q})Y = \|Y - X_q \tilde{\beta}_q\|^2$ 。

4. 实证分析

4.1. 数据来源

数据来源于 UCI 数据库(<https://archive.ics.uci.edu/ml/datasets.php>)中居住建筑数据集(Residential Building Data Set)。本文共选取了数据集中项目地区(Project locality defined in terms of zip codes)编码为 01 的数据, 数据包含 29 个变量, 其中, 实际销售价格(Actual sales prices)和实际建筑成本(Actual construction costs)为响应变量。由于本文只对建筑物实际销售价格进行分析和讨论, 因此选取了以下变量:

X_1 : 建筑物总建筑面积。

X_2 : 根据项目开始时的价格估算的建筑成本。

X_3 : 施工期。施工期的长短直接影响着建筑物的成本。

X_4 : 实际建设成本是直接影响销售价格的一个重要因素, 成本越高, 价格越高。

X_5 : 基准年建筑材料批发价格指数(WPI)。建筑材料批发价格直接影响着建筑物的初期投放资本, 也影响着建筑物的成本, 成本越高, 销售价格就会随着增加。

X_6 : 基准年土地价格指数。土地作为建筑物的基石, 其价格的高低直接影响着建筑物的价格。而土地交易价格指数, 指一定时期内土地交易价格变动趋势和变动程度的相对数, 反映了土地在一定时期内价格的涨幅, 对建筑物的销售价格有很大的影响。

X_7 : 城市/市政当局颁发建筑许可证的总建筑面积, 会影响建筑物的供需关系, 如果建筑面积供不应求, 将会导致建筑物的价格上涨, 而供过于求则会导致建筑物价格的下降。

X_8 : 贷款利率也会影响建筑物的价格。在宽松的货币政策下, 银行存贷利率都较低, 建筑物等投资市场领域就会吸引流动性资金进入, 继而拉高建筑物价格。

X_9 : 基准年消费价格指数(CPI)。消费价格指数反映居民家庭购买消费商品及服务的价格水平的变动情况, 消费者物价指数的提高意味着实际工资的减少, 说明市场上投资者购买建筑物的能力下降, 因此也是影响建筑物价格的一个重要因素。

X_{10} : 股票市场指数。股票市场仍是一个重要影响因素, 这是由于股票市场对建筑物市场存在财富效应, 当股票价格上升时, 投资者的财富总量增加, 会增加对建筑物的消费支出, 从而引起建筑物价格上涨。

X_{11} : 城市人口。人口数量直接影响着建筑物的供需, 从而影响建筑物价格。

Y : 实际销售价格。

4.2. 变量选择和模型建立

由于选取的变量中仍存在较强的相关性, 且一些变量对建筑物价格的影响很小, 所以需要对数据进行变量选择, 剔除模型中这些自变量。根据 AIC 准则, 当 $X = (X_1, X_4, X_5, X_6, X_7, X_8, X_{10})$, AIC 最小, 为 495.9955。对变量进行最小二乘回归, 得出以下结果:

Table 1. The regression results of selected model

表 1. 选模型回归结果

变量	系数	T 值	p 值	方程显著性
$\hat{\beta}_0$	5082.4778	2.3054	0.0305	R^2
$\hat{\beta}_1$	-0.1020	-1.883	0.0724	0.9106
$\hat{\beta}_4$	1.8134	1.607	0.1217	F 统计量
$\hat{\beta}_5$	64.5518	3.5694	0.0016	32.0250
$\hat{\beta}_6$	213.3041	2.2204	0.0365	p 值
$\hat{\beta}_7$	-40.9152	-3.1585	0.0044	0.0000
$\hat{\beta}_8$	-400.9235	-2.6888	0.0131	
$\hat{\beta}_{10}$	0.1420	2.2448	0.0347	

得到关于建筑价格的多元线性模型为:

$$\hat{Y} = 4372.1888 + 51.9924X_5 + 216.1359X_6 - 28.7767X_7 - 357.6507X_8 + 0.1848X_{10} \quad (1)$$

从表 1 中可以看出实际建设成本, 基准年建筑材料批发价格指数(WPI), 基准年土地价格指数和股票市场指数与建筑物价格呈正相关性关系, 且基准年土地价格指数的影响较大, 说明土地价格指数是影响建筑物价格的重要因素。而建筑物总建筑面积, 城市/市政当局颁发建筑许可证的总建筑面积和贷款利率也会影响建筑物的价格与建筑物价格成负相关关系, 且贷款利率的影响较大。

从上述拟合结果可以看出, $R^2 = 0.9106$, 说明方程的拟合效果比较好, $F = 32.0250$, 其 p 值为 0.0000, 说明方程比较显著, 在置信水平为 0.05 的情况下, 只有 $\hat{\beta}_1$, $\hat{\beta}_4$ 的 p 值大于 0.05, 认为其不显著。这可能是由于不满足 Gauss-Markov 假设或者复共线性等因素造成的。因此, 我们需要对回归模型进行修正。

4.3. 模型修正

从残差图 1 中可以看出, 除了第 18 和 29 组数据, 其余数据的学生化残差 r_i 落在 $[-2, 2]$ 中, 且没有明显的趋势, 可认为 \hat{Y} 与残差项 \hat{e} 不相关, 即满足 Gauss-Markov 假设。并通过对均值漂移线性回归模型进行检验来检验异常点, 第 18 和 29 组数据的 t 值分别为 2.1278, 2.6445, 大于 $t_{0.025}(22) = 2.0739$, 为异常点。本文将删除异常点后再进行回归分析。

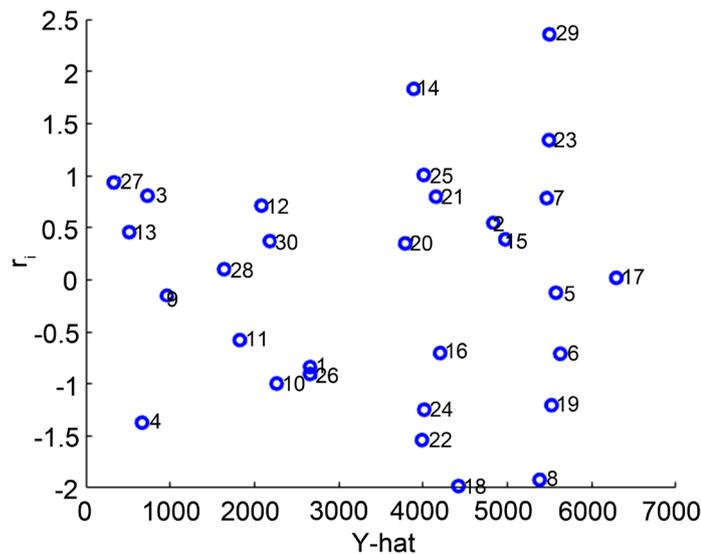


Figure 1. Residual plots

图 1. 残差图

从表 1 中可以看出, 线性回归方程是显著的, 而变量的系数并不全部是显著的, 这很有可能是复共线性的原因造成的。一般用 $X'X$ (X' 为自变量矩阵的转置矩阵) 的最大特征值与最小特征值之比来判断复共线性是否存在以及复共线性严重程度, 经计算, 得到如下结果:

$$1000 > k = \frac{\lambda_1}{\lambda_{p-1}} = \frac{4.3251}{0.0070} = 617.8714 > 100$$

因此可认为自变量 X 之间存在中等程度或较强的复共线性, 而且 λ_6 , λ_7 的值很小, 本文将通过逐步回归法对变量进行选择, 使用最小二乘法进行回归, 得到如下结果:

Table 2. The regression results of ultimate model
表 2. 模型回归结果

变量	系数	T 值	p 值	方程显著性
$\hat{\beta}_0$	4372.1888	1.9343	0.0645	R^2
$\hat{\beta}_3$	51.9924	2.9887	0.0062	0.9106
$\hat{\beta}_6$	216.1359	2.2005	0.0372	F 统计量
$\hat{\beta}_7$	-28.7767	-2.4100	0.0236	32.0250
$\hat{\beta}_8$	-357.6507	-2.3321	0.0280	p 值
$\hat{\beta}_{10}$	0.1848	3.0429	0.0054	0.0000

从表 2 中可以看出, 在置信水平为 0.05 的情况下, 模型显著和各变量都是显著的。得到修正后的多元线性模型为:

$$\hat{Y} = 4372.1888 + 51.9924X_3 + 216.1359X_6 - 28.7767X_7 - 357.6507X_8 + 0.1848X_{10} \quad (2)$$

从(2)式中可以看出基准年建筑材料批发价格指数(WPI), 基准年土地价格指数和股票市场指数与建筑物价格呈正相关性关系, 且基准年土地价格指数的系数较大, 说明土地价格指数是影响建筑物价格的重要因素。而城市/市政当局颁发建筑许可证的总建筑面积和贷款利率也会影响建筑物的价格与建筑物价格成负相关关系, 且贷款利率的系数绝对值较大, 说明贷款利率也是影响建筑物价格的重要因素。

5. 结论

通过建立多元线性模型对建筑价格的影响因素进行分析, 得出以下结论:

1) 建筑材料批发价格指数与建筑物销售价格有正相关关系。建筑材料批发价格直接影响着建筑物的初期投放资本, 也影响着建筑物的成本, 成本越高, 销售价格就会随着增加, 因此, 建筑材料批发价格越高, 建筑物销售价格越高。

2) 城市/市政当局颁发建筑许可证的总建筑面积与建筑物销售价格有负相关关系。总建筑面积增大, 使得市场上建筑物供给增加, 导致供大于求, 从而使得建筑物的价格降低。

3) 土地价格指数与建筑物销售价格有正相关关系。首先, 地方政府通过转让土地使用权获得一笔不菲的土地出让金; 其次, 建筑地产开发商不会自己承担购买土地使用权的费用支出, 所以把这部分费用当作建筑成本转嫁到建筑物价格上, 从而形成建筑物价格比预期高、建筑行业高利润率的局面, 即高的土地价格必然会导致高的建筑销售价格。

4) 贷款的利率与建筑物售价成反向相关关系。这是因为, 在宽松的货币政策下, 银行存贷利率都较低, 那么银行对社会流动性资金吸引力就会降低, 银行储蓄存款总额甚至会出现负增长。而房地产等投资市场领域就会吸引流动性资金进入, 继而拉高房价。而在紧缩的货币政策下, 银行存贷利率都较高, 对社会流动性资金吸引力较高, 在房地产的投资市场的流动性资金就会进入到银行体系中, 且上调利率对于购房者来说意味着购房支出的增加, 这实际上是央行在削减对贷款购房者的支持力度, 降低其购买力, 而购买力的降低必然会导致对房地产需求, 这种对需求的抑制作用将最终决定供给, 从而拉低建筑物价格。

5) 股票市场指数与建筑物销售价格成正相关关系。由于股票市场对房地产市场存在财富效应, 当股票价格上升时, 投资者的财富总量增加, 会增加对房地产的消费支出, 从而引起房价上涨。

参考文献

- [1] 李冬青, 张红标. 市场、价格及建筑商品价格确定研究[J]. 工程经济, 2016, 26(6): 5-9.
- [2] Kilian, W.R. and Snyman, G.J.J. (1984) Revised Forecast of the CPAP (Haylett Formula) and the BER Building Cost Index.
- [3] Taylor, R.G. and Bowen, P.A. (1987) Building Price Level Forecasting: An Examination of Techniques and Applications. *Construction Management and Economics*, 5, 21-44. <https://doi.org/10.1080/01446198700000004>
- [4] Fellows, R.E. (1988) Escalation Management. Ph.D. Thesis, Department of Construction Management, University of Reading, UK.
- [5] Skitmore, R.M. and Marsden, D.E. (1988) Which Procurement System? Towards a Universal Procurement Selection Technique. *Construction Management and Economics*, 6, 71-89. <https://doi.org/10.1080/01446198800000008>
- [6] Berk, J.M. and Bikker, J.A. (1995) International Interdependence of Business Cycles in the Manufacturing Industry: The Use of Leading Indicators for Forecasting and Analysis. *Journal of Forecasting*, 14, 1-23. <https://doi.org/10.1002/for.3980140102>
- [7] Akintoye, A., Bowen, P. and Hardeastle, C. (1998) Macro-Economic Leading Indicators of Construction Contract Prices. *Construction Management and Economics*, 16, 159-175. <https://doi.org/10.1080/014461998372466>
- [8] 吴韵侠, 吴汉雄, 樊国勋. 工程造价影响因素分析[J]. 黄河水利技术学院学报, 2002, 4(14): 41-42.
- [9] 陈光林. 影响公路项目造价的因素及对策[J]. 公路与汽车, 2004, 6(3): 116-118.
- [10] 郭计云, 师林生. 煤炭建设工程造价的影响因素及对策[J]. 煤炭经济研究, 1998(10): 35-36.
- [11] 尚梅, 周明. 建筑产品价格探讨[J]. 基建优化, 2004, 4(25): 24-26.
- [12] 尚梅, 张金锁. 中国建筑产品价格先导变量研究[J]. 西安科技大学学报, 2008, 28(1): 163-183.
- [13] 王振鑫. 灰色 GM(1,1)模型在电力建设工程造价预测中的应用[J]. 内蒙古电力技术, 1996(5): 46-50.
- [14] 钱永峰. 城镇建造价格预测模型的研究[J]. 云南工业大学学报, 1995, 11(1): 62-70.
- [15] 王桂松, 史建红, 等. 线性模型引论[M]. 北京: 科学出版社, 2004: 147-197.