

二级市场股票价格影响因素辨析 ——基于R的多元线性回归

谢淳钰 万文君(通讯作者)

(武汉工程大学管理学院 湖北武汉 430205)

摘要: 本文对可降解塑料行业55家A股上市公司财报中的截面数据进行多元线性回归分析,发现:每股净资产和每股收益等财务指标对二级市场行业板块股票定价有显著影响。多元线性回归分析是最普遍应用的实证分析方法,运用此法可以加深对定性分析结论的认识,得出各种要素间的数量依存关系,从而揭示各要素间内在的规律,将其应用到股票定价因素辨析,可使初识财会知识的人也能准确把握当前市场为行业板块股票定价的基本逻辑。同时,这把尺子具有较强的时效性,且不同的市场趋势下并不稳定。

关键词: 可降解塑料行业; 财务指标; 股价; R语言; 多元线性回归; 多重共线性检验

本文索引: 谢淳钰,万文君.二级市场股票价格影响因素辨析[J].中国商论, 2022(15):099-102.

中图分类号: F831.5

文献标识码: A

文章编号: 2096-0298(2022)08(a)-099-04

多元回归分析方法应用于股票定价逻辑剖析的思路,源于传统定价理论中的多因素定价模型,该模型在发展过程中多是基于实证观点建立起来的。

实证分析研究经济状态“是什么”,它侧重研究经济系统如何运作,分析经济活动的过程、结果和发展方向。实证分析在一定的假定条件下,描述、解释或说明已观察到的事实,对将会出现的情况或现象进行预测。客观事实是检验由实证分析方法得出结论的唯一标准。

有学者指出,实证分析的基本特征:以对经济现象分析归纳出的基本前提假设为起点,并基于这些假设进行逻辑演绎,推导出一系列结论,再逐步放松一些假设,使结论更接近具体事实。

1 研究方法

1.1 从企业价值研究开始

上市公司的整体价值由其股权资本价值和债务价值两部分组成。为了方便,我们将占比较小的优先股暂时忽略,以普通股价值作为企业资本的衡量标准;债务违约风险一般较小,其市场价格的变动也不会很大。由此,在企业资本结构不变的前提下,企业的整体价值越大,其股权资本的价值就越高,股票价格也就越高。

一般投资者都认可股价是由企业的内在价值决定的,但“内在价值”是财报中的哪些指标及怎样的权重决定的?如果这种定价逻辑在二级市场的行业板块中是普遍适用的,那么我们可以用截面数据拟合出定价的回归方程,从而揭示二级市场行业板块的定价逻辑。

1.2 多元线性回归模型

多元线性回归模型(multivariable linear regression model)中,当预测对象 y 同时受到多个解释变量 x_1, x_2, \dots, x_m 影响,且各个解释变量都与预测对象有线性相关时,可以建立多元线性回归模型进行预测和分析,模型为:

$$\hat{y} = \hat{\beta}_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_m x_m + \varepsilon_i$$

对 $i(i=1,2,\dots,n)$ 个样本均可写出:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_m x_m + \varepsilon_i, i=1,2,\dots,n \quad (1)$$

其中: m 为解释变量的数目; β_m 为回归系数(regression coefficient); ε 为随机干扰误差。式(1)也被称为总体回归函数的随机表达式。

由公式可以看出,解释变量有多个,我们构造模型时,应选取对被解释变量影响较大的几个解释变量,并且确定对应系数,再判断模型能否较好地拟合实测数据,才能进一步通过模型来预测。

在使用多元线性回归模型时,有四个基本假定:

假设一:各解释变量之间互不相关。

假设二:随机误差项是一个平均值或期望值为零的随机变量。

假设三:解释变量与随机干扰项不相关。

假设四:随机干扰项满足正态分布。

1.3 使用多元线性回归模型的前提条件

首先,行业板块概念清晰,定价逻辑成熟,即全新的概念板块可能定价逻辑尚在磨合中,对当前的股价没有指导意义。

作者简介: 谢淳钰(1998-),男,汉族,天津人,研究生,主要从事财务管理方面的研究。

通讯作者: 万文君(1972-),男,汉族,湖北武汉人,副教授,硕士,主要从事商务智能方面的研究。

其次,被评估企业所处的资本市场应有效或接近有效,上市公司在同等条件下相较于非上市企业在接受资本市场评估时更易于获得有效价格。

再次,企业不能过于特殊,同一行业中寡头企业也许在收入模式上与其他企业有明显差距,所以应该尽量避免寡头垄断企业数据对行业板块数据的影响。

最后,多元线性回归模型要求解释变量之间互不相关,即不存在多重共线性。

1.4 数据来源和指标选择

数据来源为同花顺股票交易软件中的财报数据,易查找、更官方;截面数据时间选为2021年3月31日,股价选择20日均价;行业板块选为可降解塑料;指标选择综合考虑盈利能力、偿债能力、营运能力和增长能力等。

表1所列财务指标都是经过理论和实践长期检验的,因此首选将其列为备选指标。

表1 备选解释变量一览

指标分类	指标名称	指标说明
企业规模	总资产	企业拥有或控制的能以货币计量的经济资源,如土地、房产、机械设备等
	每股净资产	每股股票所含的资产现值,一般来说,每股净资产越高,公司的经济实力就越强
经营能力	资产周转率	体现企业经营期间资产从投入到产出的流转速度,反映企业资产的管理质量和利用效率,数值越高表明资产利用效率就越高
	每股现金流量	企业在不动用外部筹资的情况下,用自身经营活动产生的现金偿还贷款、维持生产、支付股利及对外投资的能力,它是一个评估每股收益“含金量”的重要指标
偿债能力	流动比率	流动资产是企业短期内可以运用和变现的资产,企业一般用流动资产来偿还短期负债,所以流动资产相对于流动负债的规模越大,企业短期偿债能力就越强,但流动比率并不是越高越好,大量资金处于闲置状态,则资金效率低下
	速动比率	反映的是企业流动资金中流动性较强的一部分资产,也就是速动资产的短期偿债能力,分析时可以与流动比率相结合
盈利能力	营业收入	指在一定时期内,商业企业销售商品或提供劳务所获得的货币收入
	净利润	是评价企业盈利能力、管理绩效以至偿债能力的一个基本工具,是反映和分析企业多方面情况的综合指标,净利润是企业经营的最终成果,净利润多,企业的经营效益就好
	每股收益	反映企业的经营成果,是债权人、股东等信息使用者评价企业经营能力、预测企业成长能力、做出相关经济决策的重要财务指标之一
	销售毛利率	反映商品经过生产转换后增值的部分,增值越多毛利自然就越多
	净资产收益率	反映股东权益的收益水平,可以用来评估企业资本运用的效率,投资带来的收益越高该指标就会越高,通常负债增加会导致净资产收益率的上升
	每股主营业务收入	主营业务收入销售收入除以股数
成长能力	净利润同比增长率	企业当期净利润比上期净利润的增长幅度,企业盈利能力越强数值越大
流通能力	换手率	指在一定时间内市场中股票转手买卖的频率,是反映股票流通性强弱的指标之一

2 实例分析

2.1 建立多元线性回归模型

2.1.1 绘制散点图矩阵

散点图矩阵如图1所示。

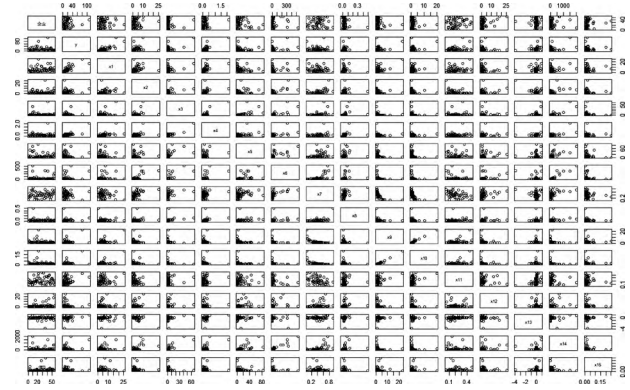


图1 散点图矩阵

2.1.2 拟合多元线性回归方程

以前述股票数据中的“股价”为因变量,拟合各财务指标的多元线性回归方程,运行R语言summary()函数返回的信息如下:

Call:

```
lm(formula=newdata$股价~newdata$每股净资产+newdata$每股主营业务收入+newdata$净利润+newdata$每股收益+newdata$主营业务收入+newdata$资产周转率+newdata$净资产收益率+newdata$流动比率+newdata$速动比率+newdata$销售毛利率+newdata$净利润同比增长率+newdata$每股经营现金流+newdata$总资产+newdata$换手率, data=newdata)
```

Residuals:

Min	1Q	Median	3Q	Max
-9.7294	-3.9276	0.1963	2.7864	16.8312

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	-3.50317	5.29626	-0.661	0.512121
newdata\$每股净资产	1.37113	0.41189	3.329	0.001880**
newdata\$每股主营业务收入	-1.37848	0.70218	-1.963	0.056610.
newdata\$净利润	-0.17363	0.45587	-0.381	0.705307
newdata\$每股收益	52.66254	13.20849	3.987	0.000277***
newdata\$主营业务收入	0.04515	0.06328	0.713	0.479712
newdata\$资产周转率	15.74758	8.63897	1.823	0.075806.
newdata\$净资产收益率	-28.40762	16.14301	-1.760	0.086097.
newdata\$流动比率	-1.17506	1.73180	-0.679	0.501349
newdata\$速动比率	2.86244	1.93679	1.478	0.147258

```
newdata$ 销售毛
利率      0.06329    10.61732    0.006    0.995274
newdata$ 净利润
同比增长率 -0.54154    0.19454    -2.784    0.008169**
newdata$ 每股经
营现金流   -2.06474    2.35375    -0.877    0.385610
newdata$ 总资产   -0.01055    0.01846    -0.572    0.570624
newdata$ 换手率    5.34480    20.06902    0.266    0.791360
```

```
---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
Residual standard error: 6.215 on 40 degrees of freedom
Multiple R-squared: 0.9148, Adjusted R-squared:
0.8849
```

F-statistic: 30.66 on 14 and 40 DF, p-value: < 2.2e-16

从回归参数的显著性可以看到:每股净资产、每股收益、净利润同比增长率通过t检验,拟合优度值达0.8849,多元线性回归方程整体通过F检验。

2.2 优化多元线性回归方程

以通过t检验的回归参数每股净资产、每股收益、净利润同比增长率,重新优化拟合回归方程,运行R语言summary()函数返回的信息如下:

Call:

```
lm(formula=newdata$股价~newdata$每股净资产+newdata$每股收益+newdata$净利润同比增长率,
data=newdata)
```

Residuals:

Min	1Q	Median	3Q	Max
-16.7502	-3.3802	-0.5693	2.9710	19.7372

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	0.5432	1.6251	0.334	0.73956
newdata\$ 每股净资产	1.9018	0.2371	8.022	1.34e-10***
newdata\$ 每股收益	34.4049	3.8201	9.006	4.04e-12***
newdata\$ 净利润同比增长率	-0.5417	0.1920	-2.822	0.00678**

```
---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
Residual standard error: 6.884 on 51 degrees of freedom
Multiple R-squared: 0.8667, Adjusted R-squared:
0.8588
```

F-statistic: 110.5 on 3 and 51 DF, p-value: < 2.2e-16

各自变量的p值为0.05水平上显著不为零,通过显著性检验;多元线性方程的拟合优度达到0.8588,通过F检验。

2.3 异方差检验

运行R语言bptest()函数,返回的信息如下:

BP=2.9039, df=3, p-value=0.4067

通过异方差检验。

2.4 自相关检验

由于此次回归分析中各自变量的数值为各股的财务指

标,数值之间没有时间顺序上的相关性,因此自相关性检验在本例中没有意义。

2.5 多重共线性检验

运行R语言vif()函数,返回信息如下:

newdata\$每股净资产 1.449929

newdata\$每股收益 1.454376

newdata\$净利润同比增长率 1.033444

各回归参数的vif值均小于10,通过多重共线性检验。

2.6 建立多元线性回归方程

我们获得了关于股价的多元线性回归方程:

股价=0.5432+1.9018每股净资产+34.4049每股收益-0.5417净利润同比增长率

2.7 用回归方程计算拟合值

运用关于股价的多元线性回归方程,可计算得到个股的拟合价格,用拟合价格减实际价格的偏差计算偏差率,其偏差率数列的描述统计如表2所示。

表2 个股拟合价格偏差率数列的描述统计

指标	值	指标	值
平均	0.092614002	峰度	0.071636284
标准误差	0.067492199	偏度	0.467366321
中位数	0.078515879	区域	2.205450315
众数	#N/A	最小值	-0.83241558
标准差	0.500535541	最大值	1.373034735
方差	0.250535827	观测数	55

由表2可以看出,有少数股票的拟合股价正向偏差过大,牵动行业板块整体拟合股价正向偏移。

2.8 分析拟合股价的决策价值

我们以拟合股价正向偏差大于90%的4只个股(苏美达、紫江企业、湖北宜化、双星新材)为例,观察2021年3月31日—2021年11月30日的股价波动,发现各股20日均价的最大涨幅依次为33.95%、110.70%、685.95%、152.08%,而同期可降解塑料指数(代码885901)的最大涨幅仅为56.13%。

由此发现,拟合股价正向偏差较大的个股(实际股价远不及按行业定价逻辑拟合的股价)均在此后约半年的时间中有强于板块的补涨冲劲。

3 多元线性回归分析的特点

3.1 优点

3.1.1 降低了研究者的进入门槛

与通过财务报表对股价进行分析相比,数理统计分析方法无需分析人员深刻把握宏观经济、产业发展趋势、市场竞争条件、企业经营状况等,仅需掌握数理统计分析技术和初步的财务知识即可。

3.1.2 有效防范研究者的主观臆断

多元线性回归分析仅依据行业和个股财报数据作出分析结果,排除了分析人员的主观因素,也可以降低因分析人员经验不足导致的分析偏差。

3.2 缺点

3.2.1 需要数据量庞大

在数据量较少时,回归分析更容易得到显著结论。一般认为,样本数量大于30时,分析结论才更有保障,同时因为要在众多的自变量中筛选具有显著性的自变量,所以有大量的数据收集和整理工作。

3.2.2 忽略了非上市公司在行业中的影响

此次行业股价回归分析中,由于非上市企业的财报不公开披露,同时非上市企业才是行业市场中的多数,所以抽样范围受限造成结论偏差在所难免,但没有影响我们探索市场定价的基本逻辑。

4 结语

相较传统的企业价值评估,数理统计分析的特点在于客观性、时效性、普遍性,只要有足够的数据,就可以充分揭示复杂经济现象下深刻的内在规律。

此次对可降解塑料行业二级市场定价影响因素的辨

析,我们发现:多元线性回归分析对于行业板块股价的定价机制揭示准确、可靠,可以作为一种投资参考工具。

随着市场周期性的波动,在市场不同趋势阶段,二级市场股票定价的影响因素是否保持稳定?会不会在市场盘整趋势和下降趋势中显现出其他影响因素的显著作用?值得我们进一步研究和揭示。

参考文献

- [1] 王青.计量经济学模型及R语言的应用[M].北京:经济科学出版社,2021.
- [2] 吴喜之.应用回归及分析:基于R[M].北京:中国人民大学出版社,2016.
- [3] 何晓群.应用回归分析[M].北京:电子工业出版社,2017.
- [4] 丁少敏.我国上市公司财务信息对股票价格影响的研究[D].南京:南京师范大学,2012.
- [5] 李潇宁.多元线性回归与时间序列模型在股票预测中的应用[J].科技创业月刊,2019(2):153-155.
- [6] 朱泓屹.基于多元线性回归的股票收益率影响因素研究[J].中国新通信,2018.
- [7] 赵理东.创业板上市公司财务信息对股票价格影响的实证研究[D].青岛:青岛理工大学,2018.
- [8] 撒媛媛.我国上市公司财务指标与股票价格之间的相关性研究[D].济南:山东大学,2018.

Analysis on Influencing Factors of Stock Price in Secondary Market

—— Multiple Linear Regression Based on R

School of Management, Wuhan Institute of Technology Wuhan, Hubei 430205

XIE Chunyu WAN Wenjun

Abstract: This paper makes multiple linear regression analysis on the cross-sectional data of the financial reports of 55 A-share listed companies in the degradable plastics industry. It is found that the financial indicators such as net assets per share and earnings per share have a significant impact on the stock pricing of the industry sector in the secondary market. Multiple regression analysis is the most commonly used empirical analysis method. Using this method, we can deepen the understanding of the conclusion of qualitative analysis, draw the quantitative dependence between various elements, and reveal the internal laws among various elements. Applying it to the analysis of stock pricing factors can enable people who are new with financial and accounting knowledge to accurately grasp the basic logic of stock pricing for the industry sector in the current market. Meanwhile, this ruler has strong timeliness, and it is not stable under different market trends.

Keywords: degradable plastics industry; financial indicators; price of stock; R language; multiple linear regression; multicollinearity test