

基于多元线性回归分析方法的物流流通水平研究

潘新宇¹, 薛 洋²

(1. 唐山学院 经济管理系, 河北 唐山 063000; 2. 天津大学 管理与经济学部, 天津 300072)

摘要:根据 2013 年全国 31 个省市的物流水平相关指标, 分别以经营单位所在地分的进出口总额、按境内目的地和货源地分的商品进出口总额作为因变量对物流流通水平进行线性回归分析, 得到了多元线性回归模型, 由此找出了流通水平的影响因素, 并有针对性地提出了对策建议。

关键词:多元线性回归; 物流; 流通水平

中图分类号:F25 **文献标志码:**A **文章编号:**1672-349X(2016)06-0072-06

DOI:10.16160/j.cnki.tsxyxb.2016.06.019

A Research on Logistical Circulation Efficiency Based on Multivariate Linear Regression Analysis

PAN Xin-yu¹, XUE Yang²

(1. Department of Economics and Management, Tangshan University, Tangshan 063000, China; 2. College of Management and Economics, Tianjin University, Tianjin 300072, China)

Abstract: By studying the logistics-related indexes of the 31 provinces in 2013, the authors of this paper have carried out a linear regression analysis of the logistical distribution in these places, with the total value of imports and exports by business location and the total amount of imports and exports of goods by destination and source location as the dependent variable, obtained the multiple linear regression model, discovered the influencing factors in the circulation efficiency, and put forward the corresponding suggestions.

Key Words: multivariate linear regression analysis; logistics; circulation efficiency

0 引言

流通业的物流构成消费品供应链的终端环节, 也是形成商品使用价值的最重要环节, 涵盖从商品采购到将商品销售给消费者的整个物流全过程, 具体包括商品采购、进货运输、商品储存、加工配送、门店物流以及为各环节提供配套服务的信息系统等。中国流通产业竞争力较低一直是比较突出的问题。产业规模通常被用作衡量流通产业竞争力的基础性指标, 它既反映目前流通产业的总量规模, 又为未来的产业扩张提供平台基础^[1]。《中国连锁零售企业经营状况分析报告》采用流通产业增加值、平均资产

规模和社会消费品零售总额等指标, 综合评价了中国各省区的流通产业规模竞争力。

物流是一个国家实现流通现代化的重要途径, 也是衡量一个国家和地区流通现代化程度的重要标志。发展现代物流, 实现流通现代化, 是提升我国经济素质和效益、降低成本、服务民生的迫切需要。物流在流通业的表现决定了流通业服务水平的高低和商品的最终价格。本文在宏观上通过研究全国各地区 2013 年 31 个省市的物流水平相关指标, 对物流流通水平进行了线性回归分析, 其中能反映物流流通水平的相关因素, 包括国内生产总值、3 种进出口

总额(按经营单位所在地分、按境内目的地分和货源地分的外商投资企业货物)、连锁零售企业通体配送商品购进额、亿元以上商品交易市场数量、货运量、运输业人员数量(铁路、道路、水上、航空、装卸搬运和运输代理、邮政业、仓储业)、居民消费水平和各地区的人口数。

1 现有研究成果及相关理论

1.1 物流流通的概念

对于流通,文献[2]指出,流通衔接生产和消费,包括两个过程:人的转移和商品的实质转移。田岛义博^[3]亦持类似的观点,他认为商品流通就是商品的转移过程,通常包含人格性转移和物理性转移。铃木武^[4]则宏观地将流通界定为“制造商与消费者之间的架桥活动”。宋则^[5]从微观层面界定了流通,即“在实体经济范畴内,由商品流通直接引起或与商品流通直接有关的商流、物流、信息流和资金流的总和”。对于流通产业,郭冬乐、宋则^[6]指出,流通产业是指与商品流通和商业直接相关联,或为商品流通和商业提供必要条件的各种投资领域的总称,它主要包括农产品、工业消费品和工业生产资料等商品的购销体系、商业设施、仓储业、运输业、包装装卸、流通加工、流通技术以及与此相关的信息产业、服务业等。夏春玉^[7]认为,流通就是专门以商品流通为经营内容的赢利性事业。

1.2 物流流通水平的影响因素

流通产业规模:产业规模是衡量流通产业发展水平高低的基石。

流通结构:主要反映流通产业的要素内部结构状况。包括人员组成、零售额比重,等等。

流通产业效率:反映流通产业运行效率和发展状况的核心指标。包括交易市场等信息。

流通产业贡献:改革开放以来,流通产业逐渐成为支撑国民经济或区域经济运行的基础部门,它决定着其他产业部门经济活动的发展水平。包括进出口额等。

2 数据收集和解析

2.1 宏观统计数据收集和解析

通过《中国统计年鉴》中的数据,搜集到了2013年31个省市的生产总值、货物进出口总额、外商投

资企业货物进出口总额、连锁零售企业基本情况、亿元以上商品交易市场基本情况、货运量、交通运输业人员数、仓储业人员数、邮政业人员数、各地区居民消费水平、各地区人口数等数据。

将所有数据按顺序整理到一个大表格中,用于之后的数据分析。数据解析如下:

生产总值:各地区经济情况的度量,是指在一定时期内(一个季度或一年),一个地区的经济中所生产出的全部最终产品和劳务的价值。

货物进出口总额:各地区实际进出我国国境的货物总金额。

外商投资企业货物进出口总额:各地区外商投资企业实际进出我国国境的货物总金额。

连锁零售企业基本情况:各地区连锁零售企业数量、营业额等。

亿元以上商品交易市场基本情况:各地区消费品综合市场、农产品市场、工业消费品市场、生产资料综合市场、工业生产资料市场、农业生产资料市场的总数量。

货运量:各地区运输企业在一定的时期内实际运送的货物数量。

交通运输:各地区海陆空运输业、仓储和邮政业人员数量。

居民消费水平:各地区居民在物质产品和劳务的消费过程中,对满足人们生存、发展和享受需要方面所达到的程度。

人口数:各地区人口数量。

2.2 各地区物流流通水平多元线性回归分析

2.2.1 模型的建立

进出口是体现流通水平的一个重要方面,所以在模型中,分别将按经营单位所在地分的进出口总额、按境内目的地和货源地分的进出口总额,作为因变量进行线性回归分析。

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \beta_6 X_6 + \beta_7 X_7 + \beta_8 X_8 + \beta_9 X_9 + \beta_{10} X_{10} + \beta_{11} X_{11} + \beta_{12} X_{12} + \beta_{13} X_{13} + \beta_{14} X_{14}。$$

变量的定义如下:

Y1:按经营单位所在地分的进出口总额(万美元);

Y2:按境内目的地和货源地分的进出口总额(万美元);

X₁:外商投资企业货物进出口总额(万美元);

X_2 :连锁零售企业通体配送商品购进额(亿元);

X_3 :亿元以上商品交易市场数量(个);

X_4 :货运量(万吨);

X_5 :铁路运输业人员数量(人);

X_6 :道路运输业人员数量(人);

X_7 :水上运输业人员数量(人);

X_8 :航空运输业人员数量(人);

X_9 :装卸搬运和运输代理业人员数量(人);

X_{10} :邮政业人员数量(人);

X_{11} :仓储业人员数量(人);

X_{12} :国内生产总值(亿元);

X_{13} :居民消费水平(元);

X_{14} :人口数(人)。

2.2.2 Y1 的多元线性回归分析

2.2.2.1 参数估计

采用强行进入法进行参数的最小二乘估计,求得回归系数如表 1 所示。

表 1 回归系数

模型	未标准化系数		标准化系数	t	Sig.	共线性统计	
	B	标准误差	Beta			公差	VIF
(常量)	9 087 629.563	5 240 169.275		1.734	0.105		
X_1	0.891	0.268	0.488	3.322	0.005	0.035	28.319
X_2	-177.301	3 883.824	-0.008	-0.046	0.964	0.026	37.761
X_3	23 513.380	10 313.018	0.181	2.280	0.039	0.121	8.278
X_4	15.816	14.185	0.062	1.115	0.284	0.244	4.102
X_5	-30.797	33.172	-0.047	-0.928	0.369	0.297	3.371
X_6	30.555	27.802	0.111	1.099	0.290	0.074	13.424
X_7	-183.836	103.214	-0.173	-1.781	0.097	0.080	12.424
X_8	183.065	145.347	0.185	1.259	0.228	0.036	28.149
X_9	270.562	150.406	0.201	1.799	0.094	0.061	16.317
X_{10}	89.168	111.631	0.121	0.799	0.438	0.033	30.166
X_{11}	300.212	205.105	0.109	1.464	0.165	0.139	7.214
X_{12}	620.642	320.744	0.409	1.935	0.073	0.017	58.459
X_{13}	-789.551	365.929	-0.242	-2.158	0.049	0.061	16.452
X_{14}	-526.122	150.194	-0.505	-3.503	0.004	0.037	27.277

因变量:按经营单位所在地分进出口总额(万美元)。

初步得到本问题的多元线性回归模型为:

$$Y1 = 9\,087\,629.563 + 0.891X_1 - 177.301X_2 + 23\,513.380X_3 + 15.816X_4 - 30.797X_5 + 30.555X_6 - 183.836X_7 + 183.065X_8 + 270.562X_9 + 89.168X_{10} + 300.212X_{11} + 620.642X_{12} - 789.551X_{13} - 526.122X_{14}。$$

2.2.2.2 三种检验

检验结果如表 2 所示。因文章篇幅关系,相关检验图表省略。

回归方程的拟合度检验:复相关系数 R 、判定系数、调整判定系数、估计值误差和 DW ,这些值反映了自变量解释因变量的能力。由表 2 可知 R 的平方是 0.989,调整后的 R 的平方等于 0.979,即判定

系数和调整后的判定系数都很接近于 1,故认定其通过了拟合度检验,且模型拟合度很好。

回归关系的整体性检验(F 检验):检验因变量的方差来源、方差平方和、自由度、均方、 F 检验统计量的观测值和显著性水平。得出 $F = 92.605$,且显著性水平是 0.000。因为本模型的 F 统计量服从 $F(14,16)$ 分布,由此查表得临界值 $F(14,16) \approx 3.4$,而本模型的 F 值 92.605 大于此临界值,且显著性水平 0.000 小于 0.05,故拒绝原假设,则表明此回归模型的整体回归模式显著,具有统计上的意义。

回归系数的显著性检验(T 检验):本模型中样本数是 31 个,自变量数是 14 个,则本模型各回归系

数的 T 统计量应服从 $T(16)$ 的分布,由检验结果得到其临界值约为 2.12,而本模型的 15 个回归系数的 T 值有的大于此临界值,有的则小于临界值,所以没有全部通过 T 检验。

2.2.2.3 多重共线性分析

通过自变量的方差膨胀因子分析的方法判断自变量间是否存在共线性。

容许度在 0~1 之间变化,其值越接近于 0,说明自变量间的共线性越强;其值越接近于 1,说明自变量间的共线性越弱。所以,若方差膨胀因子 VIF 接近于 1,则表明自变量间的共线性较弱;VIF 大于 10,则表明自变量之间存在多重共线性。

由检验结果可知,其方差膨胀因子值有的大于 10,有的小于 10,故本模型中的自变量之间存在多

重共线性。

2.2.2.4 残差的回归假设检验

(1)残差的自相关分析。表 2 为 Y1 的模型摘要。由表 2 可知,此回归模型的 DW 值为 1.962。而其样本个数是 31,自变量个数是 14,显著水平为 0.05。由于解释变量数目过多,查不到相关数据。但是可见 DW 值接近 2,因此可以认为没有自相关性。

(2)残差值统计量。表 3 为残差统计。从表 3 中可以得到此回归模型残差值的描述性统计量,包括预测值、残差、标准预测值和标准残差等。从中可以看出,最小和最大残差为 $-5.491\ 445\ 500\ 000\ 0E6$ 和 $4.623\ 251\ 500\ 000\ 0E6$,且标准化后,其均在 3 倍的标准差以内,故此模型残差值的正态分布性较好。

表 2 Y1 的模型摘要

模型	R	R 平方	校正的 R 平方	估计量的 标准误差	变化统计				Sig. F 变化值	Durbin-Watson
					R 平方 变化值	F 变 化值	df1	df2		
1	0.995	0.989	0.979	3.444 080 603 918 6E6	0.989	92.605	14	14	0.000	1.962

表 3 残差统计

	最小值	最大值	均值	标准偏差	N
预测值	$-1.615\ 732\ 625\ 000E6$	$1.082\ 572\ 800\ 000E8$	$1.423\ 088\ 360\ 345E7$	$2.343\ 561\ 927\ 252\ 9E7$	29
残差	$-5.491\ 445\ 500\ 000\ 0E6$	$4.623\ 251\ 500\ 000\ 0E6$	$0.000\ 000\ 007\ 763\ 7$	$2.435\ 332\ 749\ 983\ 9E6$	29
标准预测值	-0.676	4.012	0.000	1.000	29
标准残差	-1.594	1.342	0.000	0.707	29

因变量:按经营单位所在地分进出口总额(万美元)。

(3)标准化残差直方图。本模型的标准化残差直方图如图 1 所示,这是为了观测模型的残差是否符合正态分布。由图 1 可知,本模型标准化残差的分布结果基本符合正态分布的假设,且残差值都在 3 个标准差内。

2.2.2.5 逐步修正法

因为原模型并不是所有系数都通过了 T 检验,所以对模型进行逐步回归,得到表 4。

得到 3 个回归模型:

$$Y1 = 2\ 638\ 292.88 + 1.753X_1;$$

$$Y1 = -2\ 009\ 306.170 + 1.381X_1 + 194.345X_{10};$$

$$Y1 = -1.072E7 + 1.14X_1 + 218.808X_{10} + 574.69X_{13}。$$

对上述 3 个模型进行检验,结果如表 5,表 6 所示,从中可以看出,模型 1,模型 2,模型 3 都通过了

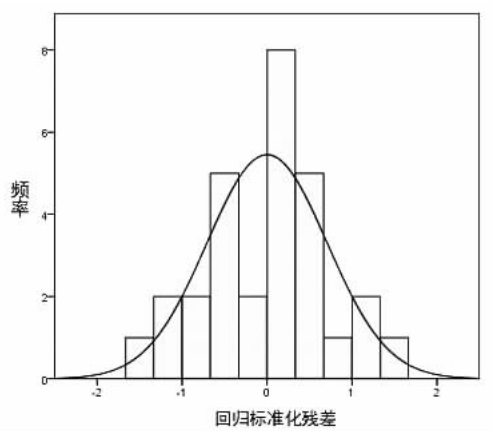


图 1 标准化残差直方图

T 检验和 F 检验。

从表 6 拟合优度 R 平方来看,模型 3 的拟合优度最高,其次是模型 2,模型 1。

表 4 修正后的模型系数

模型		非标准化系数		标准化系数	t	Sig.	共线性统计	
		B	标准误差	Beta			公差	VIF
1	(常量)	2 638 292.888	1 390 637.274		1.897	0.069		
	X_1	1.753	0.097	0.961	18.040	0.000	1.000	1.000
2	(常量)	-2 009 306.170	1 640 433.871		-1.225	0.232		
	X_1	1.381	0.124	0.757	11.168	0.000	0.405	2.471
	X_{10}	194.345	49.883	0.264	3.896	0.001	0.405	2.471
3	(常量)	-1.072E7	2 509 220.148		-4.273	0.000		
	X_1	1.140	0.114	0.625	9.961	0.000	0.296	3.377
	X_{10}	218.808	39.954	0.297	5.476	0.000	0.396	2.528
	X_{13}	574.694	141.602	0.176	4.059	0.000	0.620	1.612

因变量:按经营单位所在地分进出口总额(万美元)。

表 5 Y1 的 ANOVA 模型

模型	平方和	df	均方	F	Sig.
1 回归	1.435E16	1	1.435E16	325.430	0.000
残差	1.191E15	27	4.411E13		
总计	1.554E16	28			
2 回归	1.479E16	2	7.396E15	255.753	0.000
残差	7.519E14	26	2.892E13		
总计	1.554E16	28			
3 回归	1.509E16	3	5.030E15	277.452	0.000
残差	4.533E14	25	1.813E13		
总计	1.554E16	28			

表 6 修正后的 Y1 模型摘要

模型	R	R 平方	校正的 R 平方	估计量的标准误差	变化统计					Durbin-Watson
					R 平方变化值	F 变化值	df1	df2	Sig. F 变化值	
1	0.961	0.923	0.921	6.641 276 859 297 2E6	0.923	325.430	1	27	0.000	
2	0.976	0.952	0.948	5.377 689 356 834 7E6	0.028	15.179	1	26	0.001	
3	0.985	0.971	0.967	4.258 018 080 261 9E6	0.019	16.472	1	25	0.000	1.837

2.2.3 Y2 的多元线性回归分析

本节根据“2.2.2 Y1 的多元线性回归分析”同理可得 4 个回归模型:

$$Y2=1\ 136\ 849.844+1.986X_1;$$

$$Y2=-1\ 572\ 064.174+1.858X_1+19\ 697.532X_3;$$

$$Y2=-4\ 026\ 313.219+1.673X_1+17\ 061.435X_3+109.158X_{13};$$

$$Y2=-1\ 503\ 281.847+1.726X_1+20\ 772.420X_3+200.505X_6-52.332X_9。$$

检验结果如表 7,表 8 所示。从拟合优度来看,模型 4 的拟合优度最高,其次是模型 3,模型 2,模型 1。

3 结论

(1)从分析结果可以看出,外商投资企业货物进出口总额、邮政业人员数量、居民消费水平这几个指标在 14 个指标中对按经营单位所在地分的进出口总额有较大影响。所以按经营单位所在地划分中,关于居民日常的货物流通较为显著。因此,在进出口时,对于日常类的流通应多加管理和监控,运用物流管理知识合理规划运输和仓储,从而提升流通效率。

(2)外商投资企业货物进出口总额、亿元以上商品交易市场数量、道路运输业人员数量这几个指标在 14 个指标中对按境内目的地分和货源地分的进

出口总额有较大影响。所以按境内目的地和货源地划分中,关于工业、商业等的大中型非日常的流通较

为显著。因此加强各大企业和政府部门的管理和监控是很有必要的。

表7 Y2的ANOVA模型

模型	平方和	df	均方	F	Sig.
1 回归	1.806E16	1	1.806E16	548.429	0.000
残差	8.235E14	25	3.294E13		
总计	1.889E16	26			
2 回归	1.833E16	2	9.164E15	393.019	0.000
残差	5.596E14	24	2.332E13		
总计	1.889E16	26			
3 回归	1.845E16	3	6.151E15	324.472	0.000
残差	4.360E14	23	1.896E13		
总计	1.889E16	26			
4 回归	1.856E16	4	4.639E15	308.188	0.000 ^d
残差	3.312E14	22	1.505E13		
总计	1.889E16	26			

表8 模型摘要

模型	R	R平方	校正的 R平方	估计量的 标准误差	变化统计					Durbin-Watson
					R平方 变化值	F变 化值	df1	df2	Sig. F 变化值	
1	0.978	0.956	0.955	5.739293187326E6	0.956	548.429	1	25	0.000	
2	0.985	0.970	0.968	4.828866024628E6	0.014	11.316	1	24	0.003	
3	0.988	0.977	0.974	4.353891927422E6	0.007	6.522	1	23	0.018	
4	0.991	0.982	0.979	3.879890263974E6	0.006	6.963	1	22	0.015	1.856

(3)在不同的运输方式中,对进出口的影响并不相同,有的呈现正相关,有的呈现负相关。例如,对于按境内目的地和货源地划分的进出口中,装卸搬运和运输代理人员数量和其呈负相关,说明装卸搬运和运输代理人员数量并不能提高此类进出口的效率。但在各类运输方式都是必需的情况下,应合理分配运输人员,以提高流通的整体效率。

参考文献:

- [1] 中国流通产业发展报告课题组.我国流通产业的整体竞争力发展亟待提速[N].经济日报,2005-09-05.
- [2] 刘向阳.20世纪营销学的范式演变与特点[J].科技进

步与对策,2002(7):74-76.

- [3] 田岛义博.流通的活力[M].于淑华,译.北京:中国商业出版社,2000.
- [4] 铃木武.现代流通政策和课题[M].王哲,译.北京:中国商业出版社,1993.
- [5] 宋则.流通现代化及流通业竞争力研究(上)[J].商业时代,2006(4):11-13.
- [6] 郭冬乐,宋则.中国商业理论前沿(二)[M].北京:社会科学文献出版社,2001.
- [7] 夏春玉.商品流通产业的历史发展及其理论解说[J].财经问题研究,1998(2):71-76.

(责任编辑:夏玉玲)