

# The Influence Factors of Regional GDP Are Discussed by Using Multi-Level Model

Jingyu Bai

Yunnan University of Finance and Economics, Kunming Yunnan  
Email: 767436049@qq.com

Received: Jul. 19<sup>th</sup>, 2018; accepted: Aug. 1<sup>st</sup>, 2018; published: Aug. 8<sup>th</sup>, 2018

---

## Abstract

The influence factors of regional GDP are discussed by using multi-level model. Methods: The research objects are the major national economic indicators of 11 cities of 132 counties (cities) that come from the Hebei Statistic Almanac of 2015. Taking 132 counties (cities) as the first level and 11 prefecture-level cities as the second level, the two-level model was fitted to analyze the influencing factors of regional GDP. Results: The data were clustered in prefecture-level cities. After applying the multi-level model, it can be concluded that the total industrial output value has a great influence on the GDP of each county (city) in Hebei province. Assets investment in the whole society, local public finance budget revenue, town on-the-job worker total wages, animal husbandry fishery output value, total retail sales of social consumer goods, urban and rural residents savings at the end of the balance of regional GDP has a certain degree of positive influence. All variables are fixed effects. Discussion: The multi-level model is applicable to hierarchical data and has many advantages. It is more effective to control regional GDP from the perspective of prefecture-level cities.

## Keywords

Multi-Level Model, Regional GDP, Impact Factors

---

# 利用多水平模型讨论地区生产总值影响因素

白婧毓

云南财经大学, 云南 昆明  
Email: 767436049@qq.com

收稿日期: 2018年7月19日; 录用日期: 2018年8月1日; 发布日期: 2018年8月8日

---

## 摘要

利用多水平模型讨论地区生产总值影响因素。方法: 以河北省15年统计年鉴中11市的132个县(市)的主

要国民经济指标研究对象,以132个县(市)为第一水平,以11个地级市为第二水平,拟合两水平模型进行地区生产总值影响因素的分析。结果:该数据在地级市间存在聚集性,应用多水平模型后得到,工业总产值对河北省各县(市)生产总值有较大影响,全社会资产投资额、地方公共财政预算收入、城镇在岗职工工资总额、农林牧渔业总产值、社会消费品零售总额、城乡居民储蓄存款年末余额对地区生产总值也有一定程度的正面影响。其各变量皆为固定效应。讨论:多水平模型适用于层次结构数据,具有众多优点。从地级市角度来控制地区生产总值更有效。

## 关键词

多水平模型,地区生产总值,影响因素

Copyright © 2018 by author and Hans Publishers Inc.

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

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



Open Access

## 1. 问题的提出

在许多社会经济领域,我们得到的数据都具有层次或聚类结构。例如在分析地区经济发展的研究中,城市嵌套于省,省嵌套于国家或地区,城市的经济发展与时间有关(时间层),也形成了一个多个层次的数据结构。这种分层数据常见于人口普查、经济普查、抽样调查及跨地区跨文化的研究中。对层次或聚类数据,同一层次的数据具有较高的相似性,不同层次的数据具有较强的差异性(或异质性),传统的最小二乘(ordinary least squares, OLS)估计的假设(如同方差及独立性)不再适合这类数据的要求,使用传统的线性回归方法将会产生较大的估计误差,并得出不正确的推断结果[1]。这时就需要用多水平模型来分析。多水平模型(Multilevel Models)又称随机效应模型(Random Effect Models),它是在二十世纪八十年代,由英美教育统计学家基于方差成分分析而提出的统计模型,用于研究具有层次结构或嵌套式结构的数据,如分层抽样或整群抽样的数据[2]。生产总值是衡量一个地区经济发展的重要依据,研究生产总值的影响因素,是为了以更加科学、更加可持续地方式促进地区生产总值增长,从而更好地推动人民生活水平的提高。由于一个市内的发展战略、基础设施建设、扶持政策相类似,所以同市的各县生产总值会存在聚集性。本研究考虑了河北省同一个市内的各县不独立的特性,采用多水平模型来分析处理数据。

## 2. 生产总值和多水平模型的相关文献

### 2.1. 生产总值实证分析

李丽敏(2010) [3]用多元线性回归的方程,选取固定资产投资、消费、进出口、第一产业、工业、建筑业、第三产业和税收等影响因子,对吉林省GDP增长因素进行分析。得到了吉林省目前固定资产投资过量,已经造成了投资囤积,从机会成本来看,已经阻碍了吉林省的经济增长的结论。文静(2011) [4]通过对国内生产总值的变动进行多因素分析,建立以国内生产总值为因变量,以利用外资情况、国家财政支出为自变量的多元线性回归模型,并利用模型对国内生产总值进行数量化分析,就有关国内生产总值进行深入分析。冶涛(2012) [5]对影响新疆GDP增长因素进行实证分析,并以全社会固定资产投资总额、财政收入、出口总额、居民消费总额、能源生产总量、工业生产总值六个代表性的解释变量建立影响GDP的多元回归模型,以阐明影响新疆GDP的主要因素。

但是由上可知,国内对国内生产总值或地区生产总值影响因素的分析,大多都建立在多元线性回归

模型的基础上。但事实上,一个县(市)的生产总值,脱离不了它所属地级市的政策指引,而属于不同地级市的县(市)的生产总值又有很大差异。如果脱离这种分层关系,强制认为各县生产发展独立,可能会使研究结果不准确。在这种具有分层现象的情况下,我们可以利用多水平模型来研究县(市)生产总值。

## 2.2. 多水平模型

### 2.2.1. 多水平模型的理论研究

王艳梅等(2007) [6]应用多水平模型分析山东省和河南省不同级别的45家医院8个年度的业务收入水平。以两省间医院业务收入水平及其增长速度和增长加速度的差别无统计学意义,而省内不同级别医院间的差别具有统计学意义的结果,得到多水平模型可以有效地分析具有层次结构的、含有缺失值的纵向数据资料的结论。张旭等(2010) [7]对两水平模型与静态面板数据模型进行对比分析,面板数据可以看成是具有截面水平与时间水平的两层数据,两水平模型也能对面板数据进行分析,在一定条件下具有一定的相似性。因此,提出多水平的静态面板数据模型,为分析具有多个层次结构的面板数据提供分析工具。

### 2.2.2. 多水平模型的实际应用

近年来多水平在经济、医疗及教育等各方面都有新的发展和應用。向其凤[8]等基于云南红河州农村住户调查的微观数据,使用多水平模型分析了多种因素对农户消费支出的影响,结果发现:消费环境影响农户的自发性消费,但不影响农户的边际消费倾向;消费环境中,农户居住地的地势、交通状况、到乡镇一级市场的距离以及民族文化传统对农户的生活消费支出均有显著的影响。刘彩(2014) [9]构建两水平模型,探讨农村成人乙肝疫苗接种现状及影响因素。得到结果显示,受教育程度、乙肝知识得分是促进因素,年龄是阻碍因素。认为多水平模型适用于层次结构数据的成人乙肝疫苗接种分析,应重视村级卫生组织建设,提高村级医务人员素质,加强健康教育。

## 3. 理论模型与数据

假设数据具有两个层次, $y_{ij}$ 表示第*i*个个体(第二层次)的第*j*次(第一层次)观测变量,此时*j*表示2水平,而*i*代表1水平。1水平表示个体单位,2水平表示组单位。

### 3.1. 无条件两水平模型[10]

考虑最简单的无条件两水平模型,又称为截距模型(Intercept-Only Model)或空模型(Empty Model),是两水平模型建模的基础。其形式为

$$\text{水平 1: } y_{ij} = \alpha_i + e_{ij} \quad (3.1.1)$$

$$\text{水平 2: } \alpha_i = \alpha_0 + u_{0i} \quad (3.1.2)$$

$i = 1, \dots, n; j = 1, \dots, m$ 。将式(3.1.2)代入式(3.1.1)可得总模型为

$$y_{ij} = \alpha_0 + u_{0i} + e_{ij} \quad (3.1.3)$$

可用ICC值衡量组间、组内变异,范围在0到1之间。当ICC值趋于1时表示组间方差相对于组内非常大;当ICC值趋于0时表示没有组群效应,此时两水平模型可简化为固定效应模型。ICC值若统计性不显著,则进行多元回归模型分析,而无需两水平模型分析。

### 3.2. 条件两水平模型

条件两水平模型是在截距模型中加入了解释变量,其中既包括1水平解释变量也可能包括2水平解释变量。设*y*为因变量,*x*为1水平解释变量,*w*为2水平解释变量,且均为线性函数形式的关系(可以

具有其他函数形式的关系)。

当只有 1 水平解释变量时为如下模型

$$\text{水平 1: } y_{ij} = \alpha_i + \beta_i x_{ij} + e_{ij} \quad (3.2.1)$$

$$\text{水平 2: } \alpha_i = \alpha_0 + u_{0i}, \quad \beta_i = \beta_0 + u_{1i} \quad (3.2.2)$$

将式(3.2.2)代入式(3.2.1)可得总模型为

$$y_{ij} = \alpha_0 + \beta_0 x_{ij} + u_{0i} + u_{1i} x_{ij} + e_{ij} \quad (3.2.3)$$

### 3.3. 数据结构

分层数据经常出现于社会问题中, 这些数据具有层次嵌套结构, 如学生嵌套于学校, 时间嵌套于个体。在许多经济数据的分析中, 层次结构的数据也经常出现如宏观经济测量数据中, 城市嵌套于省区, 乡镇嵌套于县市, 不同省区或县市的测量指标数据的差异是非常明显的在微观经济数据中, 个体嵌套于乡村, 乡村嵌套于县市, 时间测量变量嵌套于个体等。传统的最小二乘法以高斯假设为基础, 不考虑数据的层次结构, 忽略了层次上个体的差异, 在这种同质性的假设下, 必然带来较大的估计误差。多水平模型是近年提出的一种研究具有层次结构数据的统计模型, 能够较好的处理数据中的组内同质或组间异质问题, 从而保证了用模型估计参数进行统计推论的准确性[11]。

### 3.4. 数据来源和变量

#### 3.4.1. 数据来源

数据来源于河北省统计局 15 年经济年鉴, 包含 14 年河北省 11 市所辖的 132 个县(除市区)的主要国民经济指标, 主要摘录了如下指标作为我们关注的变量: 工业总产值、城镇在岗职工工资总额、全社会固定资产投资额、地方公共财政预算收入、农林牧渔业总产值、出口总额、社会消费品零售总额、城乡居民储蓄存款年末余额。

#### 3.4.2. 变量介绍

本数据中选取的各变量的变异程度和单位有所不同, 而相差较大的变异, 会使不同变量的关系系数比重相差很大。为了消除量纲、变量自身变异大小和数值大小的影响, 故将数据标准化。 $y$  为因变量,  $x$  为 1 水平解释变量, 为线性函数形式的关系, 无 2 水平解释变量。各变量定义见表 1:

Table 1. Definition of variables

表 1. 变量定义

类别	变量名	变量说明	
决策变量	$y$	Z 地区生产总值(万元) (标准化)	
	$x_1$	Z 工业总产值(万元)标准化	
	$x_2$	Z 城镇在岗职工工资总额(万元) (标准化)	
	$x_3$	Z 全社会固定资产投资额(万元) (标准化)	
	1) 水平解释变量	$x_4$	Z 地方公共财政预算收入(万元) (标准化)
		$x_5$	Z 农林牧渔业总产值(万元) (标准化)
		$x_6$	Z 社会消费品零售总额(万元) (标准化)
		$x_7$	Z 出口总额(万美元) (标准化)
2) 水平解释变量	$x_8$	Z 城乡居民储蓄存款年末余额(万元) (标准化)	
	$w_1$	Z 市生产总值(万元) (标准化)	

## 4. 实证分析

### 4.1. 数据层次结构的检验

本研究以河北省 11 个地级市作为第二水平，132 个县(市)作为第一水平，拟合河北省地区生产总值的影响因素。

截距模型(空模型)，由 SPSS 软件分析，得到结果表 2。

由表 2 的结果可知，组间方差  $\sigma_{u_0}^2$  的估计值为 0.648623，组内方差  $\sigma^2$  的 0.578757， $P$  值分别 0.000 和 0.045， $P < 0.05$  具有统计意义，则组内相关系数

$$ICC = \frac{\sigma_{u_0}^2}{\sigma_{u_0}^2 + \sigma^2} = \frac{0.648623}{0.648623 + 0.578757} = 0.52846$$

这表明在地区生产总值约有 52.846% 的变异是由所属地级市不同而起的，河北省各个地级市间的地区生产总值存在差异，同一个地级市的县(市)地区生产总值有相似性。因此，该数据须用多水平模型。

### 4.2. 多水平模型的建立、估计和检验

本节将截距模型(空模型)纳入第一水平解释变量，在截距模型中引入性工业总产值、城镇在岗职工工资总额、全社会固定资产投资额、地方公共财政预算收入、农林牧渔业总产值、出口总额、社会消费品零售总额、城乡居民储蓄存款年末余额等变量。

将以上变量全部假设为固定效应，同时考虑多水平结构的随机(效应)截距方差，SPSS 软件分析，得到结果表 3。

**Table 2.** Covariance parameter estimation of county (city) GDP

**表 2.** 县(市)生产总值的协方差参数估计

参数	估计	标准误差	Z	显著性
残差 $\sigma^2$	0.578757	0.074547	7.764	0.000
截距 $\sigma_{u_0}^2$	0.648623	0.323475	2.005	0.045

Z score: 生产总值。

**Table 3.** Fixed effect parameter estimation of county (city) GDP

**表 3.** 县(市)生产总值的固定效应参数估计

参数	估计	标准误差	统计量值	显著性
固定效应				
$\alpha_0$	0.021480	0.047886	0.449	0.665
$\beta_1$ (工业)	0.616566	0.020245	30.456	0.000
$\beta_2$ (工资)	0.058989	0.016524	3.570	0.001
$\beta_3$ (投资)	0.063012	0.018562	3.395	0.001
$\beta_4$ (预收)	0.089262	0.016879	5.288	0.000
$\beta_5$ (农林)	0.070929	0.013319	5.326	0.000
$\beta_6$ (出口)	-0.010933	0.010283	-1.063	0.290
$\beta_7$ (消费)	0.077871	0.023297	3.343	0.001
$\beta_8$ (储蓄)	0.078108	0.027562	2.834	0.005
随机效应				
残差 $\sigma^2$	0.008306	0.001110	7.481	0.000
截距 $\sigma_{u_0}^2$	0.024295	0.012020	2.021	0.043

a. 因变量; Z score: 生产总值。注: 固定效应部分的统计量值为  $t$  值, 随机效应部分为  $z$  值。

如表 3 结果显示,截距项不显著( $P > 0.05$ ),其它固定效应变量工业总产值、城镇在岗职工工资总额、全社会固定资产投资额、地方公共财政预算收入、农林牧渔业总产值、社会消费品零售总额、城乡居民储蓄存款年末余额( $P < 0.05$ )具有统计意义,出口总额( $P > 0.05$ )无统计学意义。截距方差  $\sigma_{u_0}^2$  的估计值为 0.024295,  $P$  值  $< 0.001$ , 有  $P < 0.05$  具有显著统计意义,这表明市与市之间的截距有显著差异。因此具有截距项的随机效应  $u_{0i}$ 。

以上是把第一水平的解释变量全部假设为固定效应变量,即认为解释变量对结局变量的效应在组间是无差异的,但是在实际中有些解释变量对结局变量的影响是会随着组单位的不同而变化的,即还需要确定第一水平解释变量是否有随机效应[12]。

将以上变量全部假设为随机效应,同时考虑多水平结构的随机(效应)截距方差,SPSS 软件分析,得到结果显示若变量作为随机效应都不显著,即都有  $P > 0.05$ ,所以确认第一水平解释变量全部为固定效应。

SPSS 软件结果显示(由于不显著表格略),截距方差以及各随机效应变量的方差都不显著( $P > 0.05$ ),故认为解释变量对结局变量的效应在组间是无差异的。

### 4.3. 多水平模型的结论

结果显示,工业总产值( $x_1$ )、城镇在岗职工工资总额( $x_2$ )、全社会固定资产投资额( $x_3$ )、地方公共财政预算收入( $x_4$ )、农林牧渔业总产值( $x_5$ )、社会消费品零售总额( $x_7$ )、城乡居民储蓄存款年末余额( $x_8$ )共 7 个因素对地区生产总值( $y$ )都有影响。其中每增加 100 万元的工业总产值,会同比增加 61.66 万元的正产总值,即工业产值对河北地区生产总值影响最大。剩余 6 个因素对河北省地区生产总值产生不同程度正的影响。模型大致为  $y_{ij} = \alpha_0 + \beta_0 x_{ij} + u_{0i} + e_{ij}$ 。

## 5. 结论与建议

### 5.1. 河北省生产总值影响因素结论分析

研究结果显示河北省各县(市)生产总值在其所属的地级市里存在聚集性,属于相同地级市的县(市)的生产总值较相近,不同地级市的县(市)生产总值差异较大。比如唐山市,在国家倡导可持续发展政策之前,唐山长久地大力发展工业,其所辖的县(市)依照唐山市统一部署的发展战略,定位在重工业,着力于开采煤矿、铁矿等自然资源以及锻造钢铁。由多水平模型中得到的结论,工业产值对地区生产总值的影响占极大的比重,所以唐山各县生产总值居高就得到了解释。另外唐山是地震之后建造的新兴城市,三十多年来,伴随城市的新生,也提供了大量的就业岗位,由于城镇在岗职工工资总额、全社会固定资产投资额、地方公共财政预算收入也对生产总值有着不同程度的影响,所以唐山地区有着普遍较高的生产总值。

同样道理,石家庄、邯郸以及沧州市自上世纪八十年代起,对自己所辖县区有着相似的部署定位,着力发展工业及矿业,虽然近年来国家推崇绿色发展,提倡节能减排,但他们三十年积累的工业有一个相对较大的发展基数,由工业产值对地区总产值的影响占大比重,所以石家庄、邯郸以及沧州的各县(市)有相似稳定的生产总值。

而承德、张家口市及其县(市),多年政策是保持生态,绿色发展。并没有着力推动工业进步,各县(市)工业产值相对较少;而且承德和张家口所辖县(市),有大概一半是覆盖草原的坝上地区,由于其地理环境、气候特点让其城市化较为缓慢,城镇在岗职工工资总额、全社会固定资产投资额、地方公共财政预算收入、出口总额、社会消费品零售总额、城乡居民储蓄存款年末余额都相对较少,故这两个地区的生产总值处在河北省靠后的位置。当然利弊同在,承德和张家口没有通过重工业发展来提高经济等实力,却也得到了与河北其它市、县大相不同的碧水蓝天。他们现在主打生态发展,有了不一样的综合实力。

## 5.2. 关于提高生产总值的建议

由多水平模型得到, 工业总产值对地区生产总值有较大影响, 城镇在岗职工工资总额、全社会固定资产投资额、地方公共财政预算收入、农林牧渔业总产值、社会消费品零售总额、城乡居民储蓄存款年末余额对地区生产总值有相似比重的影响。所以想要提高生产总值, 要在可持续战略的前提下保证工业稳定地发展。同时又要加快产业结构调整, 逐步解放传统工业, 使工业化稳步、快速地向科技化、信息化转换。农业化向林牧渔化转换, 推动承德、张家口地区草畜牧业发展, 做到生态、绿色、进步。从而创造更多的就业领域及岗位, 增加城镇在岗职工工资总额、社会消费品零售总额和城乡居民储蓄存款年末余额。宏观方面, 加大投资力度, 合理进行全社会固定资产投资以及地方公共财政预算的投入, 要因地制宜, 根据地方特点着重支持, 使得投入有所收获。

## 参考文献

- [1] 吴晓云, 曾庆, 周燕荣. 多水平模型的最新进展[J]. 数理医药学杂志, 2003, 16(2): 152-154.
- [2] 石磊, 向其风, 鲁筠. 多水平模型及其在经济分析中的应用[J]. 数学建模及其应用, 2016, 5(1): 3-8.
- [3] 李丽敏, 王秀波. 吉林省 GDP 增长的影响因素分析[J]. 河北农业科学, 2010, 14(9): 111-113.
- [4] 文静. 影响国内生产总值的因素分析[J]. 商业文化, 2011(5): 109.
- [5] 冶涛. 国内生产总值影响因素实证分析——以新疆 GDP 增长因素为例[J]. 金融经济, 2012(16): 36-38.
- [6] 王艳梅, 王洁贞, 丁守奎, 等. 多水平模型在纵向研究资料中的应用[J]. 山东大学学报(医学版), 2007(45): 658-661.
- [7] 张旭, 石磊. 多水平模型及静态面板数据模型的比较研究[J]. 统计与信息论坛, 2013, 3(25): 22-26.
- [8] 向其风, 马晓兰. 微观消费环境与西部民族地区农户的生活消费支出——基于多水平模型的研究[J]. 云南财经大学学报, 2012(5): 22-26.
- [9] 刘彩. 河北农村成人乙肝疫苗接种多水平模型分析[J]. 中国卫生统计, 2014, 6(31): 1001-1002.
- [10] 石磊, 向其风, 陈飞. 多水平模型及其在经济领域中的应用[M]. 北京: 科学出版社, 2013.
- [11] 向其风, 石磊. 西部民族地区农村劳动力转移的影响因素分析——基于多水平 Logistic 模型的研究[J]. 数理统计与管理, 2012, 31(6): 965-975.
- [12] 骆华萍, 郜艳晖, 张丕德. 多水平模型及其对肝癌患者住院费用影响因素的分析[J]. 数理医药学杂志, 2010, 23(1): 1-4.

### 知网检索的两种方式:

1. 打开知网页面 <http://kns.cnki.net/kns/brief/result.aspx?dbPrefix=WWJD>  
下拉列表框选择: [ISSN], 输入期刊 ISSN: 2169-2556, 即可查询
2. 打开知网首页 <http://cnki.net/>  
左侧“国际文献总库”进入, 输入文章标题, 即可查询

投稿请点击: <http://www.hanspub.org/Submission.aspx>  
期刊邮箱: [ass@hanspub.org](mailto:ass@hanspub.org)