

# 工资溢价与集聚经济 ——基于中国城市面板数据的实证分析

徐小靓<sup>1</sup>, 田相辉<sup>2</sup>

(1. 武汉大学 中国中部发展研究院 湖北 武汉 430072;  
2. 青岛农业大学 经济与管理学院 山东 青岛 266109)

**摘 要:** 基于新经济地理学模型, 利用中国城市面板数据, 运用系统 GMM 方法有效处理集聚经济识别与估计过程中的内生性问题, 检验集聚经济与工资溢价之间的内在联系。实证结果发现, 核心变量市场潜能与城市职工平均工资显著正相关, 表明劳动力在市场潜能更高的城市会获得更高的能力, 或者说市场潜能更高的城市有能力支付给劳动力更高的工资, 从而揭示了生产要素纷纷流向大城市和较发达地区的主要原因。

**关键词:** 新经济地理学模型; 工资方程; 内生性; 系统 GMM

**基金项目:** 青岛市社会科学规划研究项目 (QDSKL150483); 青岛农业大学高层次人才科研基金项目 (632014); 青岛农业大学人文社会科学研究基金项目 (2314129)

**作者简介:** 徐小靓 (1980 -), 女, 山东青岛人, 博士研究生, 主要从事区域经济研究; 田相辉 (1981 -), 男, 山东临沂人, 博士, 主要从事空间经济学研究。

**中图分类号:** F061.5      **文献标识码:** A      **文章编号:** 1006 - 1096 (2015) 05 - 0014 - 05      **收稿日期:** 2014 - 01 - 22

DOI:10.15931/j.cnki.1006-1096.2015.05.003

## 一、文献评述

为什么大量劳动力流向大城市就业? 工资水平较高是一个显而易见的原因。在 21 世纪以来的中国更是如此。(副) 省级城市要比地级市的职工平均工资高 47% 左右, 东部地区要比中西部地区城市的职工平均工资高 27% 左右。这种现象可以概括为“大城市工资溢价”, 对其解释主要有两条线索 (Clark et al 2003): 一是城市内在的不同特征导致了不同城市的工资差异; 另外一个解释为劳动力被吸引到不同规模的城市是由于集聚经济提高了劳动力的收益。

新经济地理学从空间需求联系的视角考察集聚经济与工资溢价之间的内在联系。Redding 等 (2004) 将这种空间需求联系界定为市场获得 (market access), 并发现其与人均收入显著正相关。而 Hanson (2005) 则将这种空间需求联系界定为市场潜能 (market potential), 发现其与工资增长也显著

正相关。由于循环累积因果机制存在, 内生性问题是集聚经济实证分析必须要解决的难题。为了实现一致估计, 通常需要同时采用多种处理内生性问题的方法, 以保证结论的稳健性。Rosenthal 等 (2008) 利用差分 and 工具变量法 (GMM 估计) 来处理工资和集聚经济间的内生性问题。Martin 等 (2011) 利用固定效应和广义矩估计方法解决了集聚经济估计中经常遇到的不可观测异质性和同步偏误等估计问题。借鉴上述研究文献, 本文将采用新经济地理学模型, 利用中国城市面板数据, 运用系统 GMM 法处理集聚经济识别与估计过程中的内生性问题, 实现城市集聚经济的有效识别和估计。

## 二、理论模型和实证分析框架

本文借鉴 Combes 等 (2008) 新经济地理学的简化模型来构造工资方程, 并导出相应的集聚经济估计的计量经济模型。在垄断竞争的市场环境下, 区域  $r$  内的厂商  $i$  在区域  $s$  实现均衡时的经营利润 (需

要扣除生产的固定成本) 为

$$\pi_{rs}^* = (p_r^* - mc_r) \tau_{rs} x_{rs}^* = mc_r \frac{\tau_{rs} x_{rs}^*}{\sigma - 1} \quad (1)$$

其中, 市场  $s$  的均衡价格  $p_{rs}^* = \tau_{rs} p_r^* = \tau_{rs} mc_r \frac{\sigma}{\sigma - 1} p_r$ , 表示区域  $r$  厂商  $i$  所生产商品的出厂价格,  $mc_r$  为该厂商的边际生产成本,  $\tau_{rs}$  表示把货物从区域  $r$  运输到区域  $s$  所耗费的“冰山”类型的运输成本,  $x_{rs}$  表示该厂商在区域  $s$  的产品销售数量, 其和“冰山成本”的乘积表示有一部分产品在运输过程中“融化”了,  $\sigma$  表示不同类型商品的替代弹性。

在短期情况下, 厂商数量是外生的, 而且厂商的利润为正, 因此  $x_{rs}$  由如下不变替代弹性 (constant elasticity of substitution, CES) 类型的需求函数而定:

$$x_{rs}^* = (p_r^* \tau_{rs})^{-\sigma} \delta_s Y_s P_s^{\sigma-1} \quad (2)$$

其中  $P_s = \left[ \sum_r n_r (p_r^* \tau_{rs})^{-(\sigma-1)} \right]^{-1/\sigma-1}$  为地区  $s$  的 CES 价格指数,  $Y_s$  为该地区的收入水平,  $\delta_s$  为产品在地区  $s$  总消费中所占的份额, 进一步得出位于区域  $r$  的厂商总利润方程

$$\Pi_r^* = \sum_s \pi_{rs}^* - F_r = cmc_r^{-(\sigma-1)} RMP_r - F_r \quad (3)$$

其中  $RMP_r = \sum_s \tau_{rs}^{-(\sigma-1)} \delta_s Y_s P_s^{\sigma-1} c = \sigma^{-\sigma} / (\sigma - 1)^{-(\sigma-1)} F_r$  为厂商的固定成本。

根据零利润条件可得

$$mc_r = (cRMP_r / F_r)^{1/(\sigma-1)} \quad (4)$$

假设厂商的生产成本结构主要有劳动力, 也包括其他生产要素和中间投入, 并由如下柯布 - 道格拉斯函数给出:

$$\ln w_r = \frac{1}{\sigma} \ln RMP_r - \frac{1}{\sigma} \ln \left( \frac{\alpha}{c} \right) \quad (5)$$

其中  $\alpha, \beta, \gamma$  三个参数之和为 1。根据 (4) 和 (5) 可得

$$\ln w_r = \frac{1}{\alpha\sigma} \ln RMP_r - \frac{\gamma}{\alpha} \ln p_r - \frac{\beta}{\alpha} \ln q_r - \frac{1}{\alpha\sigma} \ln \left( \frac{\alpha}{c} \right) \quad (6)$$

如果进一步假定劳动力是唯一的生产要素, 即  $\beta = \gamma = 0$ , 那么工资方程 (6) 简化为 (Hanson 2005)

$$\ln w_r = \frac{1}{\sigma} \ln RMP_r - \frac{1}{\sigma} \ln \left( \frac{\alpha}{c} \right) \quad (7)$$

工资方程 (7) 表明, 区域  $r$  的均衡工资是真实市场潜能 (Real Market Potential, RMP) 的函数。真实市场潜能 RMP 捕捉了空间联系因素, 其中  $\delta_s Y_s$  表明区域  $r$  所表现的吸引力来源于地区的相对市场规

模, 而  $\tau_{rs}^{-(\sigma-1)}$  则表现为区域间贸易的交易成本, 降低了区域  $r$  的 RMP, 价格指数  $P_s$  是 RMP 与哈里斯市场潜能 MP 区别的根本标志。由新经济地理学的工资方程可知, 地区均衡工资与其他地区的工资水平以及空间接近程度密切相关。

对于实证分析而言, 最主要的难题在于方程 (7) 的准确估计。一般的做法是直接估计

$$\ln w_r = \beta_1 + \beta_2 \ln MP_r + \varepsilon_r \quad (8)$$

其中  $MP_r = \sum_s \frac{\delta_s Y_s}{d_{rs}}$  为哈里斯市场潜能,  $d_{rs} = 1/\tau_{rs}^{-(\sigma-1)}$ 。方程 (8) 两边的变量均是对数形式, 从而回归系数可以解读为弹性, 而且使回归残差接近于正态分布, 有益于统计检验的准确性。

### 三、动态面板数据模型设定和描述性统计分析

#### (一) 动态面板数据模型的设定

采用面板数据进行集聚经济识别与估计, 既可以控制城市固定效应, 也可以同时控制时间固定效应, 从而在一定程度上减缓遗漏变量带来的内生性问题。但由于模型核心变量和其他控制变量均可能存在与模型随机干扰项相关的情况, 这仍然需要利用工具变量法等方法来处理。根据理论模型和实证分析框架, 本文将构建如下动态面板数据模型:

$$\ln w_{it} = \varphi_1 \ln w_{i,t-1} + \dots + \varphi_p \ln w_{i,t-p} + \beta_2 \ln MP_{it} + \beta_3 \ln Con_{it} + \eta_c + \varepsilon_{it} \quad (9)$$

其中  $\ln w$  为各城市市辖区的职工平均工资水平, 下标  $i$  代表各个地区,  $t$  代表年份, 市场潜能 ( $MP$ ) 是本文关注的核心变量, 扰动项由城市固定效应  $\eta_c$  和随机扰动项  $\varepsilon_{it}$  两部分构成。Con 为城市层面的相关控制变量, 主要包括: 一是城市人力资本, 分别用教育 (Edu) 和医疗水平 (Bedspace) 代表。其中, 教育水平采用中高等学校学生占潜在就业人数比表示, 医疗水平采用人均床位数表示; 二是城市政府经济行为, 分别用政府财政支出占地区生产总值的比重 (Budget) 和固定资产投资占地区生产总值的比重 (Invest) 代表; 三是制度变量, 分别用私营和个体就业人数占总就业人数 (Ownership) 的比重和 FDI 占地区生产总值的比重 (Fdi) 代表。

借鉴 Arellano 等 (1991)、Blundell 等 (1998) 所提出系统广义矩估计方法 (System Generalized Method of Moments, System GMM), 充分利用水平变化和差分变化的信息作为工具变量, 以解决动态面板数

据模型(9)的内生性问题。

## (二) 数据来源与描述性统计

本文数据来自于1998年~2011年《中国城市统计年鉴》。表1列举了城市面板数据变量的描述性统计指标,主要有均值(Mean)、标准差(Std. Dev.)、最小值(Min)和最大值(Max)。宏观城市数据标准差较大,比如城市职工平均工资(Wage)的最小值和最大值相差约200多倍,而且少数变量存在离群值。所以,本文将方程两边变量均采用对数线性化处理,以减缓异方差带来的不利影响。

表1 城市面板数据的描述性统计

变量	Mean	Std. Dev.	Min	Max
Wage	17028.12	10146.53	675	141387
Mp	175.350	133.534	14.04	1128.49
Mp0	236.119	192.596	25.8	2071.25
Area	1949.924	2345.92	50	26041
Area0	82.427	114.034	5	1350
Div	2.323	1.082	0.65	10.76
Spe	4.015	4.76	1.17	112.85
Edu	0.261	0.155	0.0002	1
Bedspace	0.005	0.002	0	0.025
Budget	0.105	0.065	0	1.06
Invest	0.470	0.253	0	2.15
Ownership	0.144	0.228	0	1
Fdi	0.059	0.112	0	1.35

从总体上来看,全国城市职工平均工资水平和市场潜能较高的城市均明显集中在东部沿海地区。为了考察市场潜能和城市职工平均工资之间的联系,本文描绘了两者的散点图,并进行了线性拟合。如图1所示,市场潜能与城市职工平均工资之间呈明显的正相关关系,即市场潜能越大的地区,职工的平均工资水平越高。

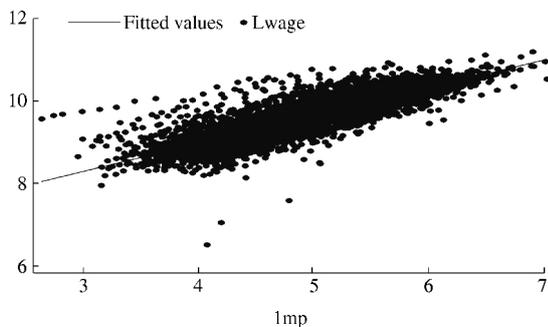


图1 市场潜能与对数平均工资的散点图及线性拟合线

## 四、计量方法和结果分析

在动态面板数据模型中,不仅存在市场潜能与平均工资有双向因果关系的内生性问题,而且被解

释变量的滞后一期也成为被解释变量,这意味着模型的内生性问题更加复杂。基于此,本文采用系统广义矩估计方法(System Generalized Method of Moments, System GMM)对方程进行参数估计。系统GMM可以很好的处理集短面板、动态因变量、固定效应和缺乏良好严格外生工具变量的情况(Roodman 2009)。其既可以利用差分方程来消除固定效应,同时也产生更多的工具变量。此外,系统GMM也有效解决了模型异方差和序列相关问题,在具体估计过程中,本文采用两步系统GMM估计和Windmeijer(2005)纠偏稳健性估计量。

为了保证动态模型设定的合理性和工具变量的有效性,本文对模型残差序列相关性和过度识别约束进行了相关检验(Arellano et al 1991)。由于系统GMM属于典型的工具变量方法,所以设定合适的工具变量是关键。内生变量(endogenous variables)和前定变量(predetermined variables)的设定及其滞后阶数的选择都会造成工具变量矩阵的巨大变化。根据模型设定框架,本文将市场潜能设定为内生变量,反映城市产业层面集聚经济变量的城市化经济和地方化经济将根据模型选择需要设定为内生变量或前定内生变量。由于城市人力资本变量和平均工资之间可能存在双向因果关系,本文也将根据模型选择需要将其设定为前定内生变量或外生变量。控制变量中除反映城市人力资本变量之外的其他变量均视为严格外生变量,它们既是水平方程的工具变量,也是差分方程的工具变量。

表2报告了动态面板模型回归结果。其中,模型(1)和模型(2)采用的是城市市辖区地理单元数据,模型(3)和模型(4)采用的是城市建成区地理单元数据。从工具变量设定的角度来看,模型(1)和模型(3)均将模型核心变量市场潜能设定为内生变量,而将城市化经济、地方化经济和城市人力资本变量设定为前定内生变量;模型(2)和模型(4)均将模型核心变量市场潜能设定为内生变量,同时也将城市化经济、地方化经济和城市人力资本变量设定为内生变量。对于因变量而言,由于模型随机干扰项存在一阶移动平均过程,所以因变量的滞后三阶以上才可以作为差分方程的工具变量。

就回归结果来看,核心变量市场潜能与城市职工平均工资的回归系数在0.20~0.23之间,而且均在1%的显著水平上统计显著,采用城市建成区地理单元数据的模型(3)和模型(4)的回归系数略大

于采用城市市辖区地理单元数据的模型(1)和模型(2) 这表明市场潜能大的区域能够吸引企业和劳动力向本区域集聚,其要素价格(比如工资)也要高于其他区域,从而证实了劳动力在市场潜能更高的城市会获得更高的能力,或者说是市场潜能更高的城市有能力支付给劳动力更高的工资。

表2 动态面板回归结果

变量	模型(1)	模型(2)	模型(3)	模型(4)
L. Lwage	0.382 *** (37.76)	0.370 *** (36.97)	0.284 *** (28.84)	0.262 *** (28.22)
L2. Lwage	0.148 *** (27.60)	0.150 *** (35.13)	0.160 *** (38.80)	0.156 *** (47.17)
Lmp	0.201 *** (17.21)	0.207 *** (19.53)		
Lmp0			0.216 *** (18.90)	0.225 *** (18.61)
Lspe	0.104 *** (14.39)	0.045 *** (10.53)	0.073 *** (16.34)	0.037 *** (10.05)
L. lspe	-0.058 *** (-8.76)		-0.037 *** (-6.86)	
Ldiv	0.028 *** (6.64)	0.029 *** (8.08)	0.024 *** (5.28)	0.018 *** (4.73)
L. ldiv	-0.003 (-0.51)		-0.013* (-1.76)	
Ledu	0.006 (1.41)	0.001 (0.44)	0.035 *** (7.91)	0.020 *** (5.43)
L. ledu	-0.001 (-0.83)		-0.001 (-1.19)	
Lbedspace	0.073 *** (9.69)	0.034 *** (5.85)	0.126 *** (13.46)	0.058 *** (9.60)
L. lbedspace	-0.044 *** (-7.73)		-0.073 *** (-10.68)	
Lbudget	-0.005* (-1.90)	-0.006** (-2.31)	-0.004 (-1.17)	-0.006* (-1.88)
Linvest	0.005 (1.38)	0.006** (1.98)	0.016 *** (4.65)	0.013 *** (3.67)
Lownership	0.030 *** (11.23)	0.029 *** (12.69)	0.023 *** (9.69)	0.021 *** (8.74)
Lfdi	0.014 *** (8.74)	0.013 *** (8.23)	0.013 *** (7.31)	0.015 *** (8.75)
Larea	0.033 *** (13.03)	0.035 *** (14.13)		
Larea0			0.007** (1.99)	0.004 (1.12)
常数项	3.494 *** (30.19)	3.530 *** (33.70)	4.488 *** (42.10)	4.657 *** (50.30)
时间效应	是	是	是	是
N	2539	2542	2539	2542
Arm1	-2.491	-2.449	-2.509	-2.462
Arm1_p	0.013	0.014	0.012	0.014
Arm2	-0.576	-0.748	-0.935	-1.369
Arm2_p	0.565	0.455	0.350	0.171
Sargan	204.977	212.794	191.144	203.541
Sargan_p	0.232	0.182	0.483	0.323

注:括号中的数字为t值;\* , \*\* , \*\*\* 分别表示显著性水平为10% , 5%和1%。

需要进一步说明的是,上述结论也可以从城市市辖区的土地面积与城市职工平均工资在1%的显著水平上显著正相关得到佐证:在控制市场潜能和其他条件不变的情况下,土地面积较大城市的劳动力工资水平要高于土地面积较小城市的劳动力工资水平。反映地区结构效应的地方化经济(Lspe)和城市化经济(Ldiv)与职工平均工资均在1%的显著水平上显著正相关,而且地方化经济效应明显大于城市化经济效应,这表明生产要素空间集聚效应在区域城市经济发展过程中提高了劳动力的收益,而且城市多样化经济和城市专业化经济在经济发展过程中并不相互排斥,可能同时发生作用,只是对经济发展的影响程度和作用机制不同而已。

## 五、小结

基于新经济地理学模型,本文利用中国城市面板数据,运用系统GMM方法有效处理集聚经济识别与估计过程中的内生性问题,检验集聚经济与城市职工工资溢价之间的内在联系。实证结果表明,核心变量市场潜能与城市职工平均工资的回归系数在0.20~0.23之间,而且在1%的显著水平上统计显著。这一方面实现了城市集聚经济的有效识别,证实了城市集聚经济提高了劳动力的收益;另一方面也表明空间要素是重要的,一个地区的经济发展与其临近地区经济发展水平和经济特征密切相关,分析中国的经济地理必须要考虑到区域之间的交互作用,区域之间的要素合理有序流动,对于我国区域统筹发展意义重大。此外,基于不同地理单元的回归分析进一步证明了结论的稳健性。

基于工资溢价的集聚经济识别结果证实了我国城市经济发展中存在显著的集聚经济效应,劳动力在市场潜能更高的城市会获得更高的能力,或者说市场潜能更高的城市有能力支付给劳动力更高的工资;在控制市场潜能和其他条件不变的情况下,城市市辖区土地面积较大城市的劳动力工资水平要高于土地面积较小城市的劳动力工资水平,这也就是为什么生产要素纷纷流向大城市和较发达地区的主要原因,即便这些地区的生产和生活成本比较高。如果政府忽视这一事实,不注重深化发达地区的开放程度,逐步消除要素流动的制度障碍和降低交易成本,而是盲目出台有关政策吸引资金和企业落户到集聚经济较弱的地区,那么,政策效果势必事倍功半。所以,如何科学推进我国的城市化进程,无疑是当前我国现代化过程中的一个难题,特别是在城乡二元结构和区域发展极不平衡的现实背景下。

①2000年至2010年(副)省级城市(直辖市)的职工平均工资约为27373.46元,地级市约为18580.85元;东部地区城市的职工平均工资约为22497.67元,中西部地区约为17673.78元。数据来源于《中国城市统计年鉴》。

参考文献:

- ARELLANO M ,BOND S. 1991. Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations [J]. *The Review of Economic Studies* 58 ( 2 ):277 - 297.
- BLUNDELL R , BOND S. 1998. Initial conditions and moment restrictions in dynamic panel data models [J]. *Journal of Econometrics* 87 ( 1 ):115 - 143.
- CLARK G L , GERTLER M S , FELDMAN M P. 2003. *The Oxford handbook of economic geography* [M]. Oxford: Oxford University Press.
- COMBES P P ,MAYER T ,THISSE J F. 2008. *Economic geography: the integration of regions and nations* [M]. Princeton: Princeton University Press.

- HANSON G H. 2005. Market potential ,increasing returns and geographic concentration [J]. *Journal of International Economics* 67 ( 1 ): 1 - 24.
- MARTIN P ,MAYER T ,MAYNERIS F. 2011. Spatial concentration and plant-level productivity in France [J]. *Journal of Urban Economics* 69 ( 2 ):182 - 195.
- REDDING S ,VENABLES A J. 2004. Economic geography and international inequality [J]. *Journal of International Economics* 62( 1 ): 53 - 82.
- ROODMAN D. 2009. A note on the theme of too many instruments [J]. *Oxford Bulletin of Economics and Statistics* 71( 1 ):135 - 158.
- ROSENTHAL S S ,STRANGE W C. 2008. The attenuation of human capital spillovers [J]. *Journal of Urban Economics* 64 ( 2 ):373 - 389.
- WINDMEIJER F. 2005. A finite sample correction for the variance of linear efficient two - step GMM estimators [J]. *Journal of Econometrics* 126 ( 1 ):25 - 51.

( 编校: 家伟)

## Wage Premium and Agglomeration Economies

——Based on Empirical Analysis of Panel Data of Chinese Cities

XU Xiao-jing<sup>1</sup> , TIAN Xiang-hui<sup>2</sup>

( 1. *The Institute for the Development of Central China , Wuhan University , Wuhan 430072 , China;*

2. *School of Economics and Management , Qingdao Agricultural University , Qingdao 266109 , China)*

**Abstract:** Based on new economic geography model , this paper uses panel data of Chinese cities and also uses system Generalized Method of Moment ( GMM) to deal with endogeneity in the process of identification and estimation of urban agglomeration economies to identify the internal relation between urban agglomeration economies and wage premium. The empirical results show that average wages and the market potentials are significantly positively related , and it appears significant at 1% significance level , which explains that labor is able to achieve higher ability in the cities with higher potential markets , or that the cities with higher potential markets are able to pay higher wages to labor. The results reveal the main reason of why labor flow into big cities and more advanced regions.

**Key words:** New Economic Geography Model; Wage Equation; Endogeneity; Generalized Method of Moment