

DOI: 10.13718/j.cnki.xdzk.2015.03.018

中国县级单位的城乡二元反差： 基于分位数回归的实证^①

许秀川， 张卫国

西南大学 经济管理学院，农业教育发展研究中心，重庆 400715

摘要：构建了影响中国县级单位的城乡二元反差系数的理论模型，并将其扩展为实证计量模型，采用分位数回归方法，实证分析了不同城乡收入差距水平的县级单位城乡二元反差的主要影响因素，根据实证结果给出了相关政策含义。

关键词：城乡二元反差；分位数回归；城镇化

中图分类号：O212.2；F207

文献标志码：A

文章编号：1673-9868(2015)03-0106-06

中共十八大提出了加快城乡一体化发展的战略，而在此之前十六大则已提出中国应该走符合国情的大中小城市和小城镇协调发展的多样化城镇化道路。新型城镇化包容性发展已经成为当今世界城镇化发展的主流方向，无论新型城镇化、包容性发展的内涵多么丰富，在中国，城乡统筹发展、协调发展仍然是其根本内核和基础。然而中国各个层级的城乡差距并没有随着国家战略的提出而得到有效遏制，十六大至今十多年过去了，城乡收入差距仍然在高位徘徊，并略有上升趋势，城乡统筹发展依然任重而道远。

研究城乡二元结构、城乡收入差距的文献已经十分浩瀚，但以县级单位为基础的城乡二元经济研究，特别是城乡二元反差与城乡收入差距领域的实证研究还相对较少。现有文献大多把问题限定于对某省市或自治区下的个别县进行案例研究，或对县域经济做了一般性的发展战略、制度、规划、法律等方面的研究。文献[1]分析了县(市)域城乡统筹的规划实施方案；文献[2]对我国东部地区县域城乡统筹发展的典型案例进行了研究；文献[3]探讨了我国县域经济发展中城乡统筹的主要内容；文献[4]以浙江省乐清市为例对城乡统筹背景下县域城镇化问题进行了研究；文献[5]运用 AHP 及空间计量模型方法，对重庆市县域经济城乡统筹进程的空间差异做了实证研究；文献[6]则对我国县域经济贫困地区的发展问题进行了探讨。还有很多相关研究无法一一赘述。总体上现有研究中理论研究较多，实证研究较少，案例研究较多，全国范围的大样本研究较少。

本研究以构建城乡二元经济模型为基础，以城乡二元经济结构反差系数(即城镇与农村的劳动生产率差异)为主要指标衡量城乡统筹发展水平，对全国县级单位进行实证研究，分析城乡二元反差系数的主要影响因素。本研究采用全国大样本的截面数据，使用分位数回归的方法，分析被解释变量在不同分位数下解释变量对其影响的变化，在稳健估计的同时获得更多有用信息。

① 收稿日期：2013-12-03

基金项目：国家社科基金重大项目(12&ZD100)；重庆市社会科学规划项目(2010YBJJ13)；西南大学博士基金项目(SWU1209303)。

作者简介：许秀川(1980-)，男，广东茂名，副教授，主要从事农业经济管理研究。

通信作者：张卫国，教授，博士研究生导师。

1 模型基础与实证研究

1.1 模型构建

假设县级单位中,农村部门的收入由柯布—道格拉斯函数决定:

$$y_1 = \frac{Y_1}{L_1} = A_1 \text{Cro}^\alpha$$

其中: y_1 为劳均收入, Y_1 为第一产业增加值, L_1 为第一产业劳动力数量, A_1 为农业技术因素(由于采用截面数据,假定技术不变,为常数), Cro 为农业总产出(根据数据的可得性,本研究以粮食总产出作为代理变量), $\alpha \in (0, 1)$ 为农村部门的收入弹性.

类似地,城镇部门的收入决定函数形式为:

$$y_2 = \frac{Y_2}{L_2} = A_2 \text{Ind}^\beta$$

其中: y_2 为劳均收入, Y_2 为第二产业增加值, L_2 为第二产业劳动力数量, A_2 为工业技术因素(性质假设同 A_1), Ind 为工业总产值(由数据的可得性,本研究以规模以上工业总产值为代理变量), $\beta \in (0, 1)$ 为城镇部门的收入弹性.

则定义城乡二元反差系数为城乡劳均比较劳动生产率,由城乡劳均收入的比值(rato)决定:

$$\text{rato} = \frac{y_2}{y_1} = \frac{A_2 \text{Ind}^\beta}{A_1 \text{Cro}^\alpha}$$

对上式两边取对数,则有:

$$\ln \text{rato} = \ln A_2 - \ln A_1 + \beta \ln \text{Ind} - \alpha \ln \text{Cro}$$

令 $C = \ln A_2 - \ln A_1$, 考虑到城乡两部门的收入同时受到固定资产投资(Inv)及政府财政支出(Gep)的影响,加入 $\ln \text{Inv}$ 与 $\ln \text{Gep}$ 作为控制变量,则最终的实证模型为:

$$\ln \text{rato} = C + \beta \ln \text{Ind} - \alpha \ln \text{Cro} + \gamma \ln \text{Inv} + \omega \ln \text{Gep} + \varepsilon$$

其中: C 为常数项; $\ln \text{Inv}$ 为全县固定资产投资额的对数(不含农户投资); $\ln \text{Gep}$ 为县政府一般性财政支出总额的对数; $\alpha, \beta \in (0, 1)$; γ, ω 大小及方向不确定; ε 为随机误差项.

1.2 数据来源与描述性统计

由《中国区域经济统计年鉴 2012》得到 2011 年相关变量的原始数据,对没给出粮食总产出的相应省份以及存在缺失值的县级单位作删除处理,共得全国 1340 个县级单位的完整数据,取对数后变量的描述性统计见表 1.

表 1 变量的描述性统计

变量	样本数	平均值	标准差	最小值	最大值
$\ln \text{rato}$	1 340	2.819	0.972	-3.019	6.027
$\ln \text{Ind}$	1 340	13.235	1.668	7.343	18.104
$\ln \text{Cro}$	1 340	11.841	1.157	4.419	14.949
$\ln \text{Inv}$	1 340	12.895	0.984	8.071	15.778
$\ln \text{Gep}$	1 340	11.926	0.525	9.216	14.373

对二元反差系数进行正态检验,结果见表 2:

表 2 城乡二元反差系数偏度与峰度正态检验

变量	样本数	Pr(Skewness)	Pr(Kurtosis)	联合检验 adj chi2(2)	Prob>chi2
$\ln \text{rato}$	1 340	0.000 0	0.000 0	.	0.000 0

由表 2 可知, 偏度和峰度检验的显著性均超过 1%, 显著地拒绝了二元反差系数对数是正态分布的假设, 联合检验自然地也拒绝了正态分布的假设。

1.3 模型的分位数回归分析

由于中国县级城乡二元反差系数为非正态分布, 进一步做核密度估计发现数据存在左偏, 并且在中心位置高度集中(为简约起见, 此处省略了核密度估计的图形和结果)。如果采用 OLS 进行回归, 一方面正态假设无法得到满足; 另一方面, 即使大样本条件下 OLS 正态渐近性得到满足, 其基于条件期望的假设使研究只能得到各个解释变量的作用在被解释变量处于平均中心位置时对被解释变量的影响, 无法观察各控制变量在不同分位点时对解释变量的影响, 这是重要的信息损失。虽然 OLS 估计存在诸多的不足, 但 OLS 常作为评价其他估计的参照, 比如, 矩估计类的工具变量法(IV)、二段和三段最小二乘法(2SLS, 3SLS)、广义矩估计(GMM)、极大似然估计(MLE)等。有关分位数估计的教材和论著多以 OLS 估计作为分位数估计的参照^[7-9], 因此本文在分位数回归的基础上, 同时给出 OLS 估计作为参照。

OLS 采用的方法为:

$$\min \sum_i (\mathbf{Y}_i - \mathbf{X}_i \boldsymbol{\beta})' (\mathbf{Y}_i - \mathbf{X}_i \boldsymbol{\beta})$$

区别于 OLS, 分位数回归采用:

$$\min \{ p \sum_{\mathbf{Y}_i \geq \mathbf{X}_i \boldsymbol{\beta}^{(p)}} |\mathbf{Y}_i - \mathbf{X}_i \boldsymbol{\beta}^{(p)}| + (1-p) \sum_{\mathbf{Y}_i < \mathbf{X}_i \boldsymbol{\beta}^{(p)}} |\mathbf{Y}_i - \mathbf{X}_i \boldsymbol{\beta}^{(p)}| \}$$

其中: $\mathbf{Y}_i, \mathbf{X}_i, \boldsymbol{\beta}$ 分别为被解释变量、解释变量及相应参数的向量; p 为分位数。可见, OLS 的估计是基于残差平方和最小的估计, 而分位数估计则基于相应分位数下的最小化加权残差绝对值。由于绝对值函数不可导, 分位数回归估计实质上采用解线性规划的外点法(exterior-point algorithms)求解, 相应的参数标准差则需要采用自抽样(Bootstrap)模拟获得, 计算过程虽相对复杂耗时, 但在计算机上用十几分钟即能完成上万次模拟, 获得较可靠结果。由于分位数回归可以很好地处理不同分位点下解释变量对被解释变量的影响, 并且由于分位数回归是基于顺序统计方法的估计, 对异方差不敏感, 对数据分布的位置、尺度的变化的影响能做出相应的估计, 因而可获得更稳健(robust)的估计结果, 同时得到更多有用的信息。实证模型的估计结果见表 3。

表 3 分位数估计与 OLS 估计结果

分位数	ln Ind	ln Cro	ln Inv	ln Gep	C	判定系数
0.1	0.424*** (0.044)	0.027 (0.043)	0.024 (0.073)	-0.316*** (0.122)	-0.595 (1.032)	0.191
0.25	0.357*** (0.035)	-0.036 (0.028)	0.004 (0.049)	-0.131** (0.07)	-0.421 (0.641)	0.164
0.5	0.278*** (0.031)	-0.080*** (0.032)	0.045 (0.062)	-0.051 (0.077)	0.157 (0.654)	0.139
0.75	0.283*** (0.032)	-0.174*** (0.045)	0.082 (0.059)	-0.06 (0.077)	1.291** (0.657)	0.147
0.9	0.307*** (0.046)	-0.347*** (0.075)	0.099 (0.09)	0.089 (0.128)	1.481 (0.998)	0.139
OLS	0.291*** (0.021)	-0.069*** (0.024)	0.079** (0.039)	-0.118** (0.064)	0.170 (0.553)	0.244

注: 分位数回归括号内为 1 万次 Bootstrap 模拟标准差, OLS 则为普通标准差; *, **, *** 分别表示 5% 和 1% 显著, 最后一列判定系数中, 分位数回归为伪 R^2 , OLS 为调整后的 R^2 。

由表 3 的 OLS 估计结果可知,所有变量均在 1% 或 5% 的水平显著,并且显著的符号与理论假设完全一致.工业总产值对城乡二元反差有正的影响,农业产出则对城乡二元反差有负的影响.全社会固定资产投资对城乡二元反差有正的影响,而政府的一般性财政支出对城乡二元反差有负的影响.这似乎支持了在县级单位,固定资产投资主要组成服务于城镇和工业,而政府一般性财政支出更多地支持了农村和农业的假设.但这只是就全国各县平均而言得出的结论,并且忽略了总体分布的情况及异方差的影响.

对分位数的结果进行分析可以看出:工业总产值在各个百分位上均对城乡二元反差系数存在 1% 显著的正影响;但农业总产值只在 50% 分位数及以上的情况才对城乡二元反差系数存在 1% 的显著负影响,在 10% 和 25% 分位数上,其影响并不显著.可见,城乡收入差距越大的县,农业产出对收入差距的缩减作用越强,在收入差距较小的县,农业产出对收入差距的影响力则变得不明显.全社会固定资产投资在 10% 至 90% 的 5 个分位点上,系数估计均不显著,可见在县级层面,全社会固定资产投资对工业与农业的收入差距影响并不显著.最后,在县级政府一般性财政支出方面,在较低的分位点 10% 和 25% 上,分别是 1% 和 5% 显著,而在 50% 及以上分位点上则不显著.一般情况下,分位点较低表示城乡二元反差较小,即城乡收入差距较小,这种情况通常发生在经济较发达的县,县级政府的财力较为充裕,因此有更多的财政用于农业支出,对减少城乡收入差距作用显著,实证结果得到了较好的解释.

图 1 中水平虚线为 OLS 估计值,折线为分位数回归系数变化,阴影部分为分位数回归 95% 的置信区间.由图 1 可以看出,对于城乡二元反差系数处于大约 50% 分位点以下及 75% 分位点以上部分的区县,工业总产值对城乡收入差距的影响力大于基于 OLS 回归的平均值,特别是在 10% 及 25% 等较低的分位点上差距较大,弹性系数相差达 0.1 个单位.而城乡二元反差系数在 50% 至 75% 分位点部分的县,工业总产值对城乡收入差距的影响力低于基于 OLS 回归的平均值.这表明在工业较发达,城乡差距相对较小的县,工业对城乡收入差距的影响力较大.

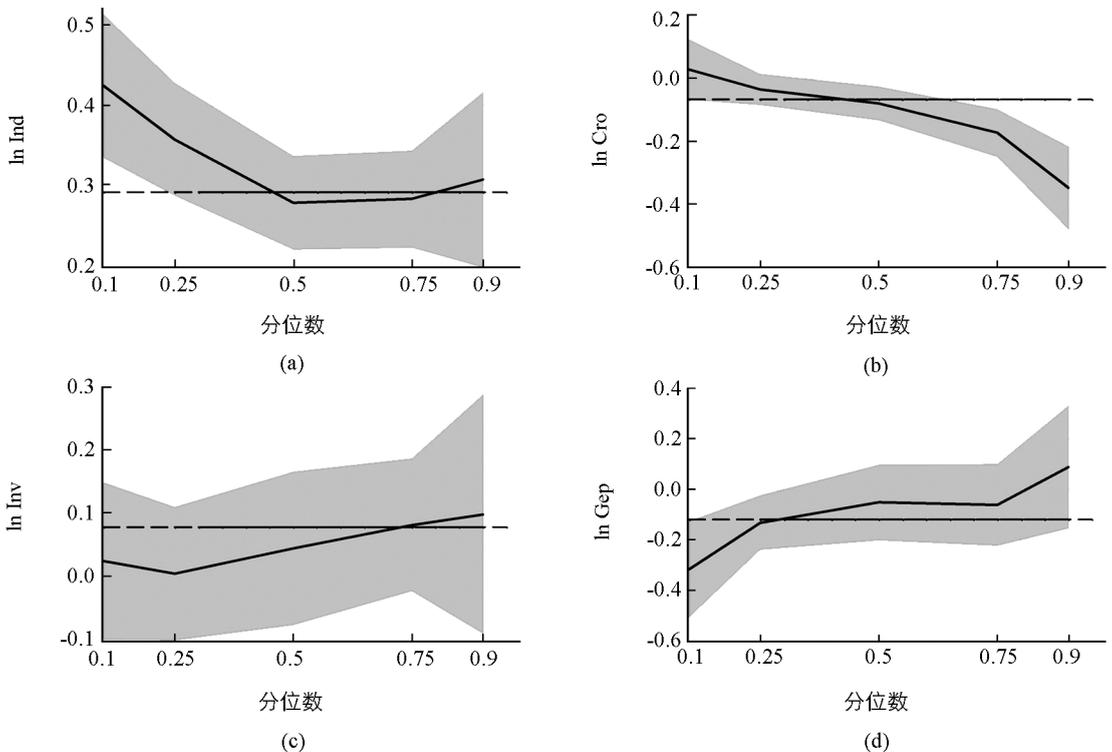


图 1 分位数回归系数估计、置信区间与 OLS 估计的比较

在农业总产值方面,二元反差系数约处于 50% 分位点以上的分位数估计较 OLS 估计的绝对值大,在

50%分位点以下则较 OLS 估计的绝对值小,这表明在城乡收入差距较大的县,农业对缩小城乡收入差距的贡献也较大。

在一般性财政支出方面,城乡二元反差系数在 25%分位以下的县,财政支出对减少收入差距有显著的作用,并且在 10%分位数和 25%分位数的估计绝对值大于 OLS 估计,而在 25%分位数至 75%分位数的估计绝对值小于 OLS 估计,在 75%分位则大于 0. 表明较富裕的县财政支农对缩小城乡收入差距有较明显的作用,而较穷的县则作用不显著。

对于固定资产投资系数的估计,OLS 估计值完全落入了各个分位数估计的置信区间内,且各分位点上固定资产投资对城乡收入差距的作用并不显著。

2 结论与政策含义

2.1 结论

本研究构建了影响中国县级单位的城乡二元反差系数的理论模型,并将其扩展为实证计量模型,采用分位数回归方法进行了实证分析. 研究表明,在各个分位点上,工业总产值对城乡收入差距有显著的正影响,而全社会固定资产投资在各个分位点上均不显著. 农业总产出则在 50%及以上分位点上对城乡收入差距有显著的负影响,而在低于 50%的分位点上则不显著. 相反,财政支出在低于 50%的分位点上对城乡收入差距有显著的负影响,在 50%及以上分位点上则不显著。

2.2 政策含义

1) 工业总产出对城乡二元结构,或城乡收入差距的影响是决定性的. 在中国的县级单位层面,城乡收入差距较小的县,即工业化程度和收入较高的县,工业产出对城乡收入差距的影响力远大于工业化程度和收入较低的县. 因此,对于已跨越经济起飞阶段,进入工业化加速阶段的地区而言,缩小城乡收入差距或缓解城乡二元结构,可以在加速工业化的同时推进城镇化,加快农村劳动力和人口的转移,以提高整体收入水平,减少城乡收入差距。

2) 农业产出对城乡收入差距的减少作用在城乡收入差距较大的县显著,在城乡收入差距较小的县则不显著,表明农业在收入较低、较落后的县显得尤为重要. 在这些县应加大对农业的支持力度,对贫困县应坚持开发式扶贫,以缩小城乡差距,实现城乡协调发展。

3) 政府的财政支出则在城乡收入差距较小的县对缩小城乡收入差距较为明显. 说明在较富裕的县,政府更有能力和财力对农业农村进行支持,较富的县财政支出应更注重公平,而较穷的县财政支出应更关注发展与效率. 应加快经济增长与发展步伐,增加县域经济的财政收入,推动工业化跨越式发展,最终实现工业反哺农业。

参考文献:

- [1] 张伟,徐海贤. 县(市)域城乡统筹规划的实施方案探讨 [J]. 城市规划, 2005, 29(11): 75-79.
- [2] 张晴,高明杰,罗其友. 我国东部地区县域城乡统筹发展模式典型案例探究 [J]. 中国农业资源与区划, 2010, 31(6): 73-78.
- [3] 丁萃华,高焕喜. 我国县域经济发展中城乡统筹的主要内容 [J]. 理论学刊, 2010, (11): 49-53.
- [4] 谢忠辉. 城乡统筹背景下县域城镇化问题研究:以浙江省乐清市为例 [D]. 上海:复旦大学, 2007.
- [5] 杨孟禹,杨刚,张引. 重庆市县域城乡统筹进程空间差异实证研究 [J]. 重庆与世界, 2013, 30(1): 9-15.
- [6] 马新民. 试论我国贫困地区县域经济的发展问题 [J]. 农业经济, 2009(2): 25-27.

- [7] ANGRIST J, PISCHKE J. *Mostly Harmless Econometrics: An Empiricist's Companion* [M]. Princeton: Princeton University Press, 2009: 205–206.
- [8] 郝令昕, 奈曼. 分位数回归模型 [M]. 肖东亮, 译. 上海: 格致出版社, 2012: 145–146.
- [9] 李子奈, 叶阿忠. 高级应用计量经济学 [M]. 北京: 清华大学出版社, 2012: 63–64.

Urban and Rural Economic Disparity of Chinese Counties: An Empirical Analysis Based on Quantile Regression

XU Xiu-chuan, ZHANG Wei-guo

*School of Economics and Management, Agricultural Education Development Research Center,
Southwest University, Chongqing 400715, China*

Abstract: In this paper, a theoretical model about the urban-rural dual disparity coefficient at the county level in China is constructed, and then it is further extended into an econometric model. With the quantile regression method, the main factors influencing the dual urban-rural disparity of Chinese counties with different income gap levels are empirically analyzed. Finally, the paper gives the related policy meanings for the empirical results.

Key words: disparity of urban and rural areas; quantile regression; urbanization

责任编辑 张 桉

