

县尺度下的出生人口性别比空间分布研究

——以山东省为例

张 昆 张松林

【摘 要】 文章以山东省 1982 年的县级人口普查数据为基础,采用局部空间自相关技术(LISA)对显著的空间聚集进行探测。结合自 2003 年起在部分试点县市推行的“关爱女孩”行动,文章指出选取 LISA 分析中的“热点”作为试点县市,会使“关爱女孩”行动取得更好的效果。

【关键词】 出生性别比 山东省 LISA “关爱女孩”行动 试点县

【作 者】 张 昆 华东师范大学地理信息科学教育部重点实验室,讲师;张松林 同济大学测量与国土信息工程系,讲师。

一、引 言

中国 1982、1990 和 2000 年的人口普查数据显示出生性别比严重失衡,且呈明显的上升趋势。目前学术界普遍认同的观点是:采用超声波技术进行胎儿性别鉴定,以及随后的选择性人工引产是出生性别比失衡的主要原因。我们注意到,大多数研究没有考虑出生性别比的空间分布,而空间分布特征对揭示这种现象的规律也是很重要的。

图 1 是中国 2000 年各省出生性别比的专题图,它按出生性别比的高低将全国各省份分为四类:第一类包括台湾、香港和澳门,其出生性别比数据是未知的;第二类包括新疆、西藏和贵州,其出生性别比小于 107;第三、四类是出生性别比失衡的省份,取值范围分别为 107~120 和高于 120。海南省居第四类的榜首,出生性别比高达 135.64。从图 1 可以看出,新疆、西藏是西部边陲省份;贵州属于出生性别比正常的省份,但被出生性别比失衡的省份包围;属于第三类的省份连成一片,第四类的省区也是如此。面对这样一种空间分布格局,我们不禁会问:为什么西部边境省份的出生性别比正常?为什么贵州省不受其近邻的影响?海南省名列前茅的出生性别比是否和它独特的地理位置有密切关系呢?为什么出生性别比高的省份呈现聚集状态?这种聚集是由邻近省份相互影响产生的吗?尽管图 1 本身并不能回答这些问题,



图 1 中国 2000 年出生性别比省级分布

但能启发我们从空间的角度看问题。

在认识到出生性别比失衡将带来严重后果的情况下,中国政府于 2003 年开始实施“关爱女孩”行动,在全国范围内选择试点县市开展“关爱女孩”的宣传和教育工作。被选为试点的县市都是 2000 年出生性别比高的地区,如福建省的安溪县为 133.24,辽宁省的大连市为 117.6,湖北省的黄岗市为 142。然而政府在选择试点县市时似乎仅注意到出生性别比的高低,忽略了出生性别比失衡县市之间的空间关系。

二、县级数据和 LISA 分析

本文采用的数据来自于《山东省人口统计资料(1984)》。该书汇总了山东省 1982 年的人口普查数据和 1984 年的抽样调查数据。本文采用的是 1982 年的数据,因为该数据覆盖了全省所有县市,非常适合做空间分布的研究。

山东是人口大省,第三次人口普查时山东省总人口排第三位,仅次于四川、河南省;到 2000 年人口普查,总人口的排名山东省上升至第二位,仅次于河南省。1982 年的山东省下辖 111 个县市。各县市的出生性别比如图 2 所示,按比值的高低分为小于 107、107~120、120~130 和大于 130 四类。

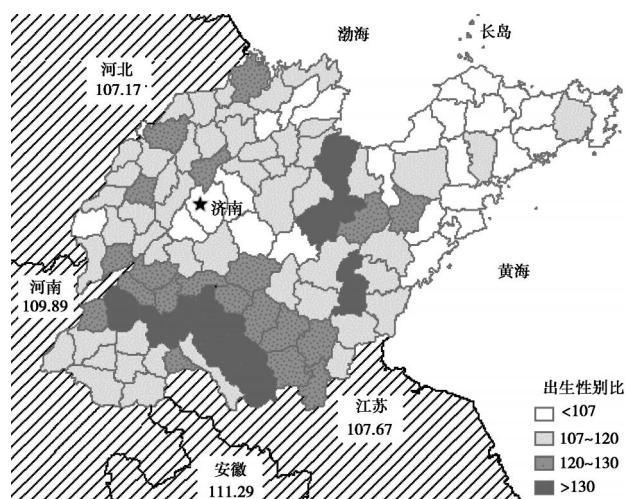


图 2 山东省 1982 年出生性别比分布

从图 2 可看出,取 107 作为出生性别比正常的界限,山东半岛大部分县市的出生性别比都低于 107,我们推测这种差异可能是由山东半岛特殊的地理位置造成的。内陆地区除少数几个县(如济南市及其相邻县)外,绝大多数县的出生性别比都超出了正常值范围。从图 2 还可以看出:大部分出生性别比特别高(大于 130)的县市聚集在山东省南部,第三类(120~130)县市也聚集在南部地区。根据 1982 年人口普查数据,山东的几个相邻省份的出生性别比也高出了正常值:河北为 107.17、河南为 109.89、安徽为 111.29、

江苏为 107.67。这些现象似乎表明相邻的空间单元存在着较强的关联,那么如何度量或是否有必要度量这种空间关联呢?图 2 似乎已经把出生性别比的空间分布特征描述得很清楚,但图 2 中的分类方式只是许多可能性中的一种,改变类别的数量或分类方法(如等间距法、自然分级法等)将生成另一幅不同的专题图,很可能会使我们对相同现象得出不同的结论。Monmonier 在《如何用地图撒谎》一书中指出:“由于众多功能强大、界面友好的计算机制图软件的出现,使得地图制作变得非常容易,在制图过程中出现不经意的自欺行为是难免的”。Tufte(1997)也表达了同样的担忧:“现在可以非常容易地对成百上千种图形和统计结果进行选择,从中挑出与自己观点一致的数据,从而使最终公布的结果带有强烈的主观性”。为了尽量避免制图过程中的主观性及引入定量计算方法,本文采用了空间统计学中的空间关联局部指标(LISA)技术。

Anselin(1995)定义的 LISA 必须满足两个条件:(1)能给出衡量每个空间单元自相关显著性程度的指标;(2)所有空间单元的指标之和与全局空间自相关指标成比例。LISA 已广泛应

用于空间经济学领域 (Ying, 2000; Anselin, 1988; Erica, 2006; Jeremy, 2002; Michael, 1999; Nadege, 2005)。我们认为 LISA 也同样适用于对本文数据的分析, 因为出生性别比失衡与地区之间的联系程度有很大的关系 (如医疗技术的推广、社会道德风气的传播等)。同时, 出生性别比失衡也与经济行为密切相关, 正如 Zeng (1993) 所指出的那样: “性别鉴定已成为一种有利可图的行当”。本文采用 Geoda 软件计算 LISA, Geoda 用局部 Moran 指数计算空间单元 i 的自相关程度, 计算公式为:

$$I_i = \frac{(x_i - \bar{x})}{S^2} \sum_{j=1}^n C_{ij} (x_j - \bar{x}), S^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}, \bar{x} = \frac{1}{n} \sum_{j=1}^n x_j \quad (1)$$

其中 C 是空间权矩阵, C_{ij} 表示空间单元 i 和 j 之间的邻接关系。正的 I 值表示相似值 (高值或低值) 存在空间聚集现象; 负的 I 值表示不相似值 (高值和低值或低值和高值) 存在空间聚集。空间聚集的显著性水平可通过计算 I 的标准化值, 再与标准正态表里对应的值进行比较得到。空间权矩阵的选择一直是最困难的, 也是最有争议的问题 (Anselin, 1988), Julie (2003) 建议采用 K 近邻权作为研究欧洲经济状况的空间权矩阵, 因为采用 K 近邻权后, 一些岛屿 (如西西里岛、撒丁岛) 就成了欧洲大陆的邻居, 避免了权矩阵中与这些岛屿相关的行和列取值为 0 的问题。同时还保证了隶属于不同国家的空间单元之间的邻接关系。但对于本文的研究而言, 采用简单的邻接权矩阵可能更合适, 因为一个县市的出生性别比往往与其邻县的出生性别比关系最密切, 而且山东省行政单元的空间布局比欧洲紧凑得多。邻接权矩阵中, 若 i 和 j 单元存在公共边界, 则 $C_{ij} = 1$, 否则 $C_{ij} = 0$ 。有两种构造邻接权矩阵的准则, 称为 Rook 和 Queen 准则, 前者认为具有公共边的空间单元是邻居, 而后者认为具有公共边或公共点的空间单元都是邻居。本文采用 Rook 权构造邻接权矩阵。根据公共边的准则, 岛屿县没有邻居, 为了保证计算的可靠性, 需要为岛屿县设置邻居。Geoda 采用文本文件存储空间权矩阵, 因此手工修改很容易。长岛是山东省唯一的岛屿县, 本文将离其最近的两个县 (蓬莱、龙口) 设为长岛的邻居; 反之, 蓬莱和龙口县也把长岛设为邻居。

采用 Geoda 的 LISA 分析可得到显著性图、聚集图、方盒图和 Moran 散点图。其中最实用的是聚集图, 它用不同的颜色渲染不同的空间自相关类别。空间自相关是指某空间变量 (这里是出生性别比) 在空间上的相关性, 在 Geoda 中分为: (1) “高 - 高”型, 俗称“热点”, 代表出生性别比高, 且被出生性别比高的县市包围的区域; (2) “低 - 低”型, 俗称“冷点”, 代表出生性别比低, 但被出生性别比低的县市包围的区域; (3) “低 - 高”型, 代表出生性别比低, 但被出生性别比高的县市包围的区域; (4) “高 - 低”型, 代表出生性别比高, 但被出生性别比低的县市包围的区域; (5) 不显著, 代表出生性别比不存在显著空间自相关, 呈空间随机分布的区域。

三、计算结果及其稳定性分析

图 3 是 LISA 分析得到的聚集图, 采用 9999 次置换生成 (置换次数越多, 结果越稳定)。由于 Geoda 中生成的彩图在灰度模式下不易区分, 本文将结果转换到 Arc GIS 中, 采用另外的图例重绘了结果。表 1 列出了存在显著空间自相关的县市的出生性别比统计。

从图 3 可以看到, 除利津 ($id = 7$) 和东营 ($id = 1$) 外, 大多数属于“低 - 低”聚集的县市位于山东半岛。需要注意的是: “低 - 低”中的“低”表示出生性别比小于平均值 (116.75), 因此“低”并不一定意味着性别比正常。例如属于“低 - 低”聚集的县市中, 东营 ($id = 1$)、长岛 ($id = 2$) 和莱阳 ($id = 31$) 的出生性别比偏高, 分别为 112.25、115.77 和 108.94。属于“高 - 高”聚集的所

表 1 四类空间自相关县市的出生性别统计

id	县市	出生男婴数	出生女婴数	出生性别比	id	县市	出生男婴数	出生女婴数	出生性别比
低 - 低					高 - 高				
1	东营	2163	1927	112.25	52	昌乐	4366	3064	142.49
2	长岛	257	222	115.77	55	安丘	7646	6210	123.12
4	蓬莱	2551	2481	102.82	80	梁山	7166	5852	122.45
7	利津	1383	1315	105.17	81	汶上	4558	3618	125.98
8	烟台	4169	4021	103.68	83	泗水	5010	4076	122.91
11	荣成	3666	3505	104.59	84	曲阜	4748	3501	135.62
13	文登	3880	3625	107.03	87	平邑	5988	4893	122.38
14	牟平	3020	2879	104.90	88	兖州	4346	3230	134.55
16	栖霞	3846	3606	106.66	90	费县	7161	5739	124.78
17	招远	3103	2920	106.27	91	嘉祥	5751	3927	146.45
28	乳山	3634	3454	105.21	92	济宁	5968	4267	139.86
29	海阳	3824	3754	101.86	94	邹县	6289	4029	156.09
31	莱阳	5230	4801	108.94	98	滕州	10046	5902	170.21
51	即墨	3819	3711	102.91	106	枣庄	14201	10360	137.08
62	青岛	12706	12199	104.16	高 - 低				
低 - 高					26	博兴	1933	1655	116.80
47	青州	4953	4413	112.24	40	平度	5741	4822	119.06
100	微山	4969	4406	112.78					

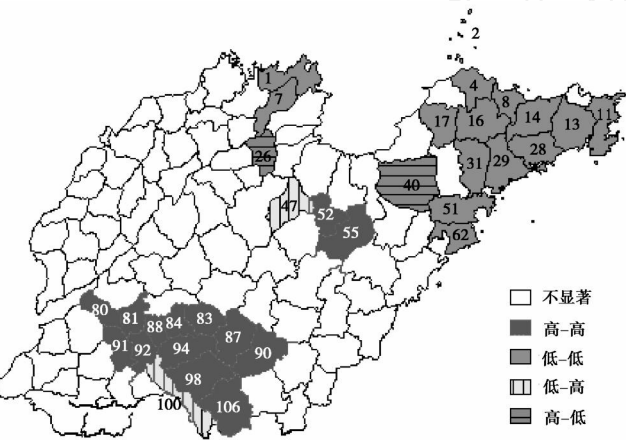


图 3 LISA 聚集图(9999 次置换)

注:图中数字是每个县的编号,与表 1 中的“id”对应。

有县市,其出生性别比无一例外地都偏高,其中平邑县(id = 87)的最小,为122.38,仍远大于正常的出生性别比。青州(id = 47)和微山(id = 100)属于“低 - 高”聚集,他们都和“高 - 高”聚集的县市相邻;属于“高 - 低”聚集的包括博兴(id = 26)和平度(id = 40),它们都和“低 - 低”聚集的县市相邻。从局部看,一些县的性别比非常高(如滕州等),这是由于县的样本量较小造成的偏差。因此,这里的计算结果与大样本量(全国或省级数据)计算出来的性别比不具可比性,但对于本文的分析技术是适用的。

南部的“热点”由 12 个县市组成,是应该重点关注的地区。如图 4 所示,这些县市聚成一团,该地区无疑是女性胎儿最不易生存的地方。根据局部 Moran 指数的计算公式及热点的定义,热点区域还应该包括“热点”周边的空间单元,从图 4 可以看到,“高 - 高”区域周边县市的出生性别比也普遍偏高。

为了检验计算结果的稳定性,本文引入了概率转移矩阵(Julie, 2003)。共计算了 3 个概率转移矩阵,每个矩阵反映的是用一阶 Rook 权和高阶(二至四阶)Rook 权计算后,空间单元显著性水平的转移情况。如果某个空间单元用一阶 Rook 权计算后是“高 - 高”聚集,用高阶 Rook 权计算后仍是“高 - 高”聚集,说明该空间单元的状态是稳定的。高阶 Rook 权采用累加权,即包括了所有的低阶邻居。表 2 列出了计算结果。从表 2 可以看出,在一阶 Rook 权下“不显著”的县

市在二阶权下还有 70.5 %保持为“不显著”状态,到三阶、四阶时降为 52.6 %和 37.2 %。属于“高 - 高”和“低 - 低”空间聚集模式的县市的转移概率最小:一阶 Rook 权下为“高 - 高”聚集的县市在高阶 Rook 权下有 85.7 %仍然为“高 - 高”聚集;一阶 Rook 权下为“低 - 低”聚集的县市在高阶权下有 93.3 %保持了“低 - 低”状态。属于“高 - 低”和“低 - 高”空间聚集模式的县市表现得不稳定,可能与属于该聚集模式的样本数量太少有关。仅一半的“低 - 高”聚集县市在高阶权下仍保持为“低 - 高”;一阶 Rook 权下表现为“高 - 低”的县市在高阶权(三阶和四阶)下全部转移到了“不显著”状态。有趣的是,与 4 种(高 - 高、低 - 低、高 - 低、低 - 高)模式对应的聚集模子矩阵为对角阵,表明在一阶权下存在空间自相关的县市在高阶权下仅向“不显著”状态转移。由以上检验可知:LISA 分析采用一阶 Rook 权计算的结果中,关于“高 - 高”和“低 - 低”聚集状态的判断是非常稳定的。换句话说,分析得到的“热点”和“冷点”区域是可信的。

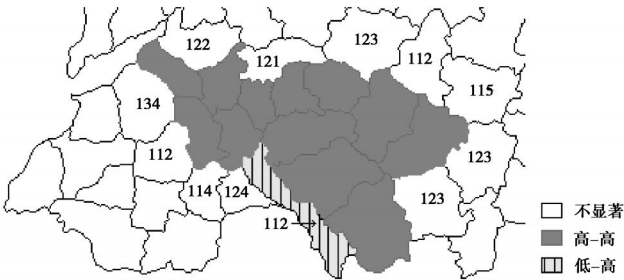


图 4 “热点”周边县市的出生性别比

表 2 稳定性检验的概率转移矩阵

		总数	不显著	高 - 高	低 - 低	低 - 高	高 - 低
一阶 二阶	不显著	78	0.705	0.077	0.128	0.064	0.026
	高 - 高	15	0.143	0.857	0	0	0
	低 - 低	14	0.067	0	0.933	0	0
	低 - 高	2	0.500	0	0	0.500	0
	高 - 低	2	0	0	0	0	1.000
	一阶 三阶						
一阶 三阶	不显著	78	0.526	0.115	0.128	0.167	0.064
	高 - 高	15	0.143	0.857	0	0	0
	低 - 低	14	0.067	0	0.933	0	0
	低 - 高	2	0.500	0	0	0.500	0
	高 - 低	2	1.000	0	0	0	0
	一阶 四阶						
一阶 四阶	不显著	78	0.372	0.153	0.141	0.282	0.051
	高 - 高	15	0.143	0.857	0	0	0
	低 - 低	14	0.067	0	0.933	0	0
	低 - 高	2	0.500	0	0	0.500	0
	高 - 低	2	1.000	0	0	0	0

四、“关爱女孩”行动试点县的选择

目前“关爱女孩”行动在被选中的试点县市中陆续开展。2004 年 3 月山东省启动了“关爱女孩”行动,8 个县被选为试点,即枣庄市薛城区、莱阳、昌乐、微山、平邑、乐陵、临清和郓城。其位置分布如图 5 所示。

表 3 给出了这些试点县市 1982 年的出生人口统计数据。由于没有 2000 年县级人口统计数据,我们根据 2000 年山东省的出生性别比(112.17)推测,所选中的 8 个试点县的出生性别比在 2000 年时应高于平均值 112.17。从表 3 可以看出,只有乐陵和莱阳的出生性别比在 1982 年是低于 112.17 的,分别为 108.48 和 108.94。可以断定,2000 年乐陵和莱阳的出生性别比升高,且超出了 112.17。

“关爱女孩”行动试点县的选择非常重要,试点县应该选择在最具决定意义的地方,我们认为,图 3 所示的“热点”(高 - 高聚集区)就是具有决定意义的目标。同时“关爱女孩”行动应把力量集中在“热点”区域,这样才会取得更有效的成果。图 3 所示的“热点”有两处,通常应选取

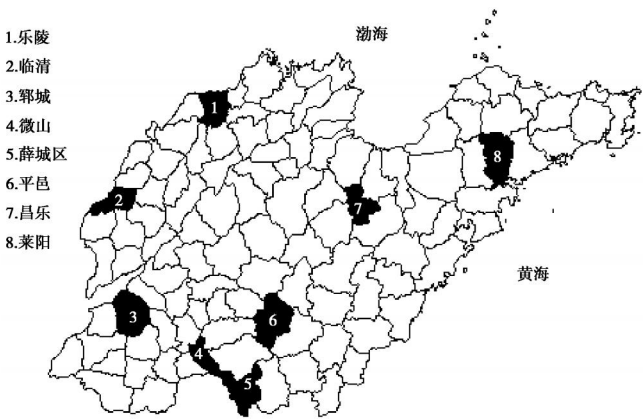


图 5 山东省“关爱女孩”行动试点县分布

表 3 1982 年 8 个试点县市的出生人口

县市	出生男婴数(人)	出生女婴数(人)	出生性别比
乐陵	3964	3654	108.48
临清	4653	4147	112.20
郓城	5592	4170	134.10
微山	4969	4406	112.78
薛城区	14201	10360	137.08
平邑	5988	4893	122.38
昌乐	4366	3064	142.49
莱阳	5230	4801	108.94

面积大的“热点”作为首选对象,即南面由 12 个县市组成的区域。尽管 2004 年出生性别比的“热点”位置我们无从得知,但图 5 显示,所选取的 8 个试点县分散在山东省境内,没有集中于一处。分散策略尽管也有优点,如扩大影响,但面对根深蒂固的“重男轻女”观念,如果措施的力度不够集中,不够深入,很可能达不到期望的效果。根据 2005 年的 1%抽样数据,中国出生性别比是 118.58,高于 2000 年的 116.86,从这个角度看,“关爱女孩”行动在降低出生性别比方面做得还不够成功,相关部门应考虑采取新的策略来帮助解决出生性别比失衡问题。

五、结 语

本文从空间分布的角度对性别比失衡进行了探讨,基于山东省县级出生人口数据,引入 LISA 技术探测了出生性别比在空间上的自相关性。分析时采用

了空间邻接权矩阵,对探测出的显著空间聚集区域,采用概率转移矩阵检验了结果的稳定性。检验结果表明,探测出的“热点”和“冷点”都是非常可信的。本文中的“热点”是指出生性别比高且趋于空间聚集的县市,而“冷点”是出生性别比低且趋于空间聚集的县市。我们确信,“热点”地区普遍存在“重男轻女”的思想,并可能存在非法鉴别胎儿性别的地下市场。“关爱女孩”行动是中国发起的针对出生性别比失衡的一项有力的举措,通过在试点县开展“关爱女孩”行动,营造有利于女婴生存和发展的环境。我们认为,在试点县的选取上可以有改进的地方,与当前分散的试点县相比,将试点县放在“热点”区域,即出生性别比聚集的地方,可以取得更有效的成果。

参考文献:

1. Anselin, L (1988) ,Spatial Econometrics. Methods and Models ,Dordrecht Kluwer Academic Publishers.
2. Erica F (2006) ,Patterns of Spatial Autocorrelation of Assemblages of Birds ,Floristics ,Physiognomy ,and Primary Productivity in the Central Great Basin ,USA. Diversity and Distributions 12 ,236-243.
3. Julie L (2003) ,Exploratory Spatial Data Analysis of the Distribution of Regional Per Capita GDP in Europe , 1980 ~ 1995. Regional Science 82 :175-201.
4. Michael S (1999) ,Spatial Autocorrelation of Cancer in Western Europe. European Journal of Epidemiology 15 :15-22.
5. Nadege C (2005) ,Spatial Autocorrelation as a Tool for Identifying the Geographical Patterns of Aphid Annual Abundance. Agricultural and Forest Entomology 7 ,31-43.
6. Ying L G (2000) ,Measuring the Spillover Effects :Some Chinese Evidence. Regional Science 79 :75-89.

(责任编辑:朱 犁)

The Change in Dysfunctional Pregnant Outcomes in China and its Implications to Population Health

Zheng Xiaoying and Others ·48 ·

Based on data of surveys conducted by the National Population and Family Planning Commission in 1988, 1997 and 2001, this paper studies the changing trend of dysfunctional pregnant outcomes and the regional variation and age variation of the outcomes. The results show an increase tendency of dysfunctional pregnancy's proportion in recent years: the birth ratio of dysfunctional pregnancy of the low age cohort (15 ~ 24) and the high age cohort (35 ~ 49) tend to be higher than that of the middle age cohort (25 ~ 34), and the birth ratios of dysfunctional pregnant outcomes in all the age cohorts have a growing tendency. Analysis of regions' classification based on GDP figures indicates different models in different regions.

New Rural Cooperative Medical System and Rural Residents' Medical Security

Chen Zaiyu Kuai Xuguang ·55 ·

Using the data of China Health and Nutrition Survey (CHNS) in 2004, this article analyzes the influence of new rural cooperative medical system on rural residents' security to explain the difficulties faced by the rural cooperative medical system. Two decision making models for rural residents' choices to participate in the rural cooperative medical system and medical expenditures are constructed. The results show that the peasants with best and worst health in self-evaluation are not willing to join the rural cooperative medical system, indicating the system encounters adverse selection and the ability to disperse the risk of medical services is much lower. The medical expenditure model suggests that the probability and quantity of medical expenditure for rural residents depend mainly on the severity of illness and are not affected by enrolling in the cooperative medical system. Therefore, we conclude that the membership of new rural cooperative medical system has not provided enough medical security.

College Students' Employment Decision-making and Career Expectations in China :Evidence from Zhejiang Province

Han Yixiang Weng Jie Zhou Biyu ·63 ·

This paper studies the effects of college students' human capital and socio-economic backgrounds on their employment decision-making and career expectations. The results show that college students' human capital plays a decisive role in their employment decision-making and career expectations. College students with more human capital stock plan their careers earlier. Although these students wish to work in big cities, they don't hope to enter labor market too early. Upgrading the economic condition of students families will weaken their employment intension. Students in better off families not only raise positive claims for selecting careers and jobsites, but also expect for higher salaries. These findings suggest that college students' current employment decision-making and career expectations not only reinforce labor market dualistic structure, but also enlarge social inequality by intergeneration transmission.

The Impulse Response Analysis on the Relation between Labour Productivity and Wage :The Case of Guangdong Province

Ding Yuan ·72 ·

The article studies the long-term relationship between labour productivity and wage in Guangdong province by the Cointegration approach, and builds the model of VAR to conduct the Impulse response analysis and Variance decomposition. Results indicate that the fluctuation of labour productivity in Guangdong province is mainly attributed to the shock itself, and this shock has a large intensity and a long-term effect on the fluctuation of average wage in different ownership units, which is much stronger than the effect of the mutual shocks among average wages.

Research on Handling Sex Ratio Unbalance at Birth in China :Stage, Characteristics, Mechanism and Prospects

Mu Guangzong Yu Liming Yang Yuezhong ·81 ·

Based on investigations on "Pilots of Caring for Girls" in Pengshui County of Chongqing, Huize County of Yunnan province, and Shaoxin city of Zhejiang province, this paper discusses the development stage, the characteristics, the mechanism and prospects of controlling Sex Ratio Unbalance at birth, and explores the regularity of governing population sex ratio issues. This paper presents three stage assumption for controlling Sex Ratio Unbalance at birth, and summarizes the mechanics, prospects and the strategic framework of governing Sex Ratio Unbalance at birth in China.

Spatial Distribution Analysis on Sex Ratio at Birth in Shandong Province

Zhang Kun Zhang Songlin ·89 ·

Based on birth data at county's level, this paper investigated the distribution of SRB in 1982 in Shandong province. The technique of Local indicator of spatial association (LISA) was employed to detect significant spatial clustering. The findings are facilitative to policy making in the "Care for Girls" campaign conducted since 2003 and aimed to reduce the SRB. This paper claims that selecting "hot spots" in LISA analysis as experiment counties and cities may lead to better results in the "caring for girls" campaign.