

对 2000 年人口普查出生性别比 的分层模型分析^{*}

郭志刚

【内容摘要】本文对以往出生性别比研究中的若干观点进行了学术评论,并讨论了一个个别经验统计分析中的方法问题。在此基础上,本文应用非线性分层模型将 2000 年全国人口普查 1%数据和 1999 年的地区级生育政策数据结合起来对出生性别比失调的影响因素进行了初步分析。结果表明,育龄妇女现有子女数量及性别和一些社会特征对出生性别比存在单独的影响,并且肯定了生育政策既存在对出生性别比的直接影响,也通过其与妇女以往生育结果以及其他社会特征之间的交互效应来影响出生性别比。

关键词: 出生性别比;生育政策;分层模型分析

【作者简介】郭志刚,北京大学中国社会发展研究中心研究员、社会学系教授。北京:100871

1 问题的提出

我国出生性别比从 1980 年代以来出现偏高失调,并且愈演愈烈。早期研究中,还有相当多的人认为这种失调主要是由于出生女婴瞒报漏报的原因,后来随着不断采用新的调查数据对以往统计结果的核对,大部分人都已经认识到出生性别比失调并不是统计虚幻,而是社会现实。2000 年全国人口普查数据表明,五普前一年内的全国出生性别比已经高达 119.92,而国家统计局发布的 2006 年国民经济和社会发展统计公报提供的出生人口性别比为 119.25。

近年来对出生性别比失调原因的研究很多,争论也很大。周云、任强(2004)曾对这方面的研究文献做过归纳和梳理,涉及到的原因方方面面,即包括育龄夫妇本人的生理状况、以往生育史,也包括个人社会经济特征(教育、职业、户籍类型,流动迁移状况等等),还包括所在的家庭类型,并涉及到居住地的自然和社会地理特征(地形地势和城乡),以及更为宏观层次的社会经济状况。^①至于社会经济特征如何能影响到出生子女的性别,主流认识是某些特征人群有更大的倾向做产前胎儿性别鉴定,然后做选择性的人工流产。^②

一个争论热点是计划生育政策或计划生育工作是不是出生性别比失调的原因。认为计划生育与出生性别比失调有关的看法由来已久,但是最初由于这个话题太敏感,公开发表的文献几乎找不到,近年来持这种认识的文献大量增加。但是,主要停留在理论推理的认识阶段,以实际数据分析结果来举证的研究很少。

顾宝昌和罗伊(1996)曾用韩国和中国台湾地区无强制性计划生育,但也存在出生性别比偏高的

* 本研究受教育部人文社会科学重点研究基地课题“家庭代际关系的人口社会学研究”(05JJD840002)和国家社会科学基金课题“人口学方法论研究”(05BRK007)的资助,在此鸣谢。

① 有关具体文献索引可参见周云、任强(2004),这里不再重复罗列。

② 陈卫的专著(2005)利用所收集的各种汇总统计和调查数据对中国的人工流产进行了比较全面的研究与分析。

现象,来说明中国的计划生育不是主要原因。这种学术观点及其论证逻辑的影响很大,广为引用。就其论证逻辑上的问题,我同意乔晓春(2004)的评论,即这种举证“只能说明对出生性别比的偏高不受计划生育的作用是不可能的,但并不能证明它一定不受计划生育的作用”。也就是说,作为这种观点的支持证据,上述现象可作为必要条件,但尚不充分。

并且,我认为还有必要指出,上述推理的缺陷也同样普遍地发生于观点截然相反的研究中。问题在于,不论想支持哪种观点,举证少数例证,都不足以作为充分支持。很多社会经济影响因素并不是以确定性方式、而是以或然性方式发生作用的。因此,不服从一般规律的特例总是存在,但是并不能作为充分的反证而否定一般规律的普遍存在。所以,总是可以比较容易地找到一些特殊案例。但是,这种特例没有普遍性,所以依据它们的理论推理即使在逻辑上成立也不能代表现实中的一般规律。

此外,在出生性别比研究中还有一种较有影响的说法也值得推敲。比如,“中国出生性别比偏高和推行计划生育政策并没直接关系,二者不是因果关系”,并同时认为“计划生育政策促成了中国低生育水平的早日到来,加速人口转变的实现,间接地影响了出生性别比偏高”(原新、石海龙,2005)。应当指出:第一,没有直接影响不一定意味着没有因果关系,因为间接影响也是因果关系。这如同计划生育政策和个人社会经济变量都只能通过中间变量来影响个人生育是同一个道理。通径分析模型是一种因果分析统计模型,其主要特征就是不仅分析直接影响,而且也分析间接影响。第二,不可以简单地将直接影响等同于主要影响,而将间接影响等同于次要影响。在通径分析研究中经常可以看到,一个外生变量对结果变量的间接影响可能是大于直接影响的,并且间接影响的方向甚至可能会与直接影响的方向相反,结果是间接影响的方向决定了总影响的方向。

下面再简评一下四个将生育政策直接与出生性别比联系起来的定量研究。

陈友华(2006)利用广东省1997年曾大规模收紧生育政策这样一个条件,对政策调整前后的孩次递进生育结构和孩次别出生性别比进行了对比,得出了生育政策变化对出生性别比的影响。他的研究视角独特,相当于将广东省的政策调整当作一次“自然”实验(艾尔·巴比,2000:上卷315页),用“实验”前后的变化来反映不同生育政策的影响。但是,这种方法很难应用于一般的横贯数据分析。

张二力(2005)用地区级的政策生育率(即生育政策内在要求的平均生育水平)与地区级出生性别比做了相关分析,肯定了两者之间存在正相关。这一研究从方法论上还有一定缺陷,因为生育政策是一种大环境影响,处于宏观层次,然而夫妇的产前性别选择则是一种微观行为。严格地说,从宏观单位的变量相关推论微观单位的变量也相关,要冒生态学谬误的风险。

笔者(2005)则是将地区级各类生育政策的人口覆盖比例信息与2000年普查的1%样本数据联系起来,根据普查户记录中的市镇县特征按梯次地将较严政策的人口比例随机分配给各户,然后对这样划分的不同生育政策人口进行了探测性分析,也得出了不同生育政策下的出生性别比有显著差别的结论。但是,以这样的方式将两套不同来源、不同层次的数据生硬地结合起来,肯定会有分配错误,所以其分析的有效性较差,结果只能大致看看,聊补此类信息空白。

此外,上述这三个实证研究还有一个共同的短处,就是只考虑了个别人口特征(如出生孩次、母亲的生育年龄),对其他社会经济变量的控制不够。也就是说,就算以上指出的那些方法缺陷都不存在,也不能排除其研究结果受到了育龄妇女其他人口、社会、经济特征干扰。

杨菊华(2006a)用“中国健康与营养调查”的纵向调查数据,以社区为分析单位,应用常规回归证明了生育政策对社区男孩比例存在影响,并应用固定效应模型表明生育政策类型和推行手段的变化也导致社区男孩比例的相应变化。这一研究在方法上已经比较讲究,并且分析时对子女结构和其他社会经济变量都进行了统计控制。但是,社区虽然距离个人很近,但毕竟还不是生育行为的直接发生者,所以推理上仍然有生态学谬误的风险。另外,这种将个人数据汇总为社区层次变量再做回归有两

个统计缺陷:一是个人信息被损失了,二是回归反映出来的其实是社区变量之间的关系,而这种关系与上层变量与下层变量(比如生育政策与个人生育)的实际关系可能并不一致。

建立一个好的理论框架是做好研究分析工作的基础。乔晓春(2004)曾提出了一个出生性别比的研究框架(请参看他原文中的示意图),区分了4种不同的变量:确定性变量(如生育政策、男孩偏好),条件变量(B超技术可得性),中间变量(如选择性人工流产),结果变量(生育数量、性别、时间)。

根据实际研究条件,上述框架既可以扩展,也可以省略。比如,育龄妇女个人的各种社会经济特征会影响个人的男孩偏好强度,而各种社会调查都有这种信息,所以即使没有男孩偏好信息仍然可以做生育结果研究。当没有条件变量和中间变量时,那么它们也可以省略,就像乔晓春在同一篇论文中以及许多其他人所做的一样。

此外,用什么模型来操作上述理论框架也是很重要的问题。由于生育政策数据的不可得,大家可能会忽略这个问题。比如,笔者(2005)所用生育政策信息并不是个人特征,而是以行政地区为单位口径的各类生育政策实际覆盖的人口比例。又如,中国健康与营养调查的生育政策数据(杨菊华,2006a)也是当地干部提供的社区层次数据。统计原理早就证明,无论是将地区级的生育政策变量分别赋值给个人后再进行回归分析,还是将个人信息先按各地区统计出汇总变量,然后在地区级与生育政策进行回归分析,都是不对的。在面对这种分层数据结构的情况下,1990年代以来日趋成熟的分层模型才是一种恰当的分析模型。

本文将尝试应用分层模型来检验生育政策是否影响出生性别比。这一尝试不仅在于取得具有理论和实际意义的结果,而且也有学习和推广先进分析方法的学术意义。

2 分层模型方法

有关分层模型的原理简介请参见笔者《分析单位、分层结构、分层模型》一文,其中说明了什么是分层数据结构以及为什么对于分层数据结构既不能将上层的变量值赋予下一层案例然后在下层做回归分析,也不能将下层数据汇总到上层单位再做回归。该文还简介了分层模型近年来的新扩展。分层模型研究的中文论文还可参见杨菊华《多层模型在社会科学领域中的应用》和郭志刚、李剑钊《农村二孩生育间隔的分层模型研究》。更全面系统的原理和应用可参见 Raudenbush 和 Bryk(2002)以及 Goldstein(1995)的权威英文专著。

这里不准备再对分层模型的原理加以详细讨论与介绍。笔者将 Kreft 和 de Leeuw(1998)在介绍分层模型的书中所用的示意图加以改造,画出常规回归研究框架简图和本研究的分层模型研究框架简图,这样可以形象化地比较其差别。

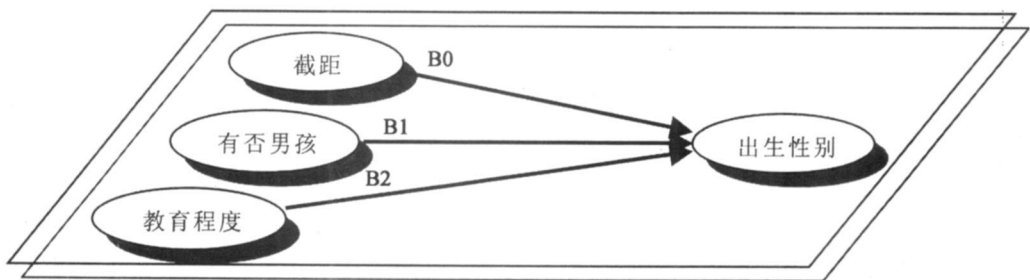


图1 个人层次的常规回归方程

$$\text{出生性别} = B_0 + B_1 \times \text{有否男孩} + B_2 \times \text{教育程度}$$

如果本研究只用某年有生育的育龄妇女的某两个方面的个人特征作为其出生子女性别的解释变量,并不考虑将地区层次的生育政策作为出生性别的解释变量,那么这涉及同一层次的数据,因此可以用常规回归来分析。图1示意了这一常规回归的因果关系,它可以画在一个平面上。

说明:

①图1中的椭圆代表变量,箭头线代表变量之间的因果关系。各B系数为待定估计,代表个人层次自变量X对因变量Y的作用方向及强度。

②将回归截距也作为一个变量对待(取值常数1),这与一般通径图画法略有不同,因为当另外两个真正的自变量采取不同形式时,截距回归系数 B_0 便取得不同的估计值和实际意义。

③在后面的实际操作中,有否男孩、教育程度、出生性别都是代表二分类的虚拟变量。虚拟变量作为回归自变量没有问题,但是作为因变量时违反常规回归假定条件,所以实际分析中采用logistic回归,那时本研究回归中的真正因变量会转换为出生性别比的对数。这里为了简明,先不谈这些细节。

图2提供了加入地区层次生育政策变量的分层模型框架。在这个分析框架中,地区层次变量构成了上一层的平面,其中含有上层截距和生育政策变量。

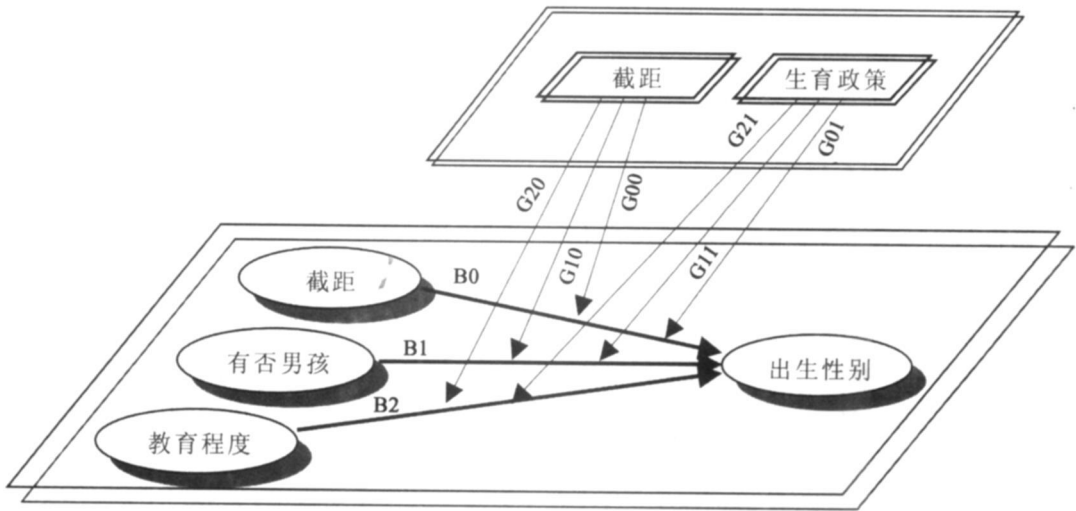


图2 分层模型的分析框架

说明:

①上层平面中的平行四边形代表地区层次的变量。其实,分层模型允许在上层包括更多自变量。

②上层变量指向下层变量的箭头线代表层际之间的作用关系。各G系数为待定估计,代表作用方向及强度。

③分层模型框架与一般通径分析图的不同之处主要在于,层际作用关系并不表现为上层变量值影响下层变量值,而是体现于上层变量影响着下层自变量与下层因变量(它也是整个分层模型的结果变量, outcome variable)之间的关系,即影响各B系数的取值大小和方向。

④图2假设上层变量对下层模型中的所有回归系数都有影响,也就是说研究者有兴趣对它们一一检验。其实,可以根据理论检验需要或数据限制条件将部分层际影响G系数设为0,也就是说从分层模型图中将对应的作用通径删除。当所有层际影响都删除时,分层模型就变回在下层的常规回归了。

⑤图2模型也可以用公式表达如下(其中 U_0 为地区层次截距方程的随机误差):

个人层次方程 出生性别 = $B_0 + B_1 \times \text{有否男孩} + B_2 \times \text{教育程度}$

地区层次方程 $B_0 = G_{00} + G_{01} \times \text{生育政策} + U_0$

$B_1 = G_{10} + G_{11} \times \text{生育政策}$

$B_2 = G_{20} + G_{21} \times \text{生育政策}$

将地区级方程分别代入个人方程,便得到了组合方程:

出生性别 = $G_{00} + G_{01} \times \text{生育政策} + U_0$

+ $G_{10} \times \text{有否男孩} + G_{11} \times \text{生育政策} \times \text{有否男孩}$

+ $G_{20} \times \text{教育程度} + G_{21} \times \text{生育政策} \times \text{教育程度}$

在回归分析中两个变量的乘积可以视为一个交互变量, 它的回归系数就是交互效应。所以, G_{01} 、 G_{11} 、 G_{21} 就是层际交互效应, 也就是生育政策通过与个人层次自变量的互动对最终结果出生性别产生的影响。所以, 本研究自然十分关注各 G_{p1} 的统计显著性以及它们的方向与强度。而 G_{p0} 类系数则是个人层次自变量对最终结果的单独影响。

当某个交互效应 G_{p1} 统计性显著时, 有 $B_p = G_{p0} + G_{p1} \times \text{生育政策}$, 说明生育政策的不同取值会改变各地区中个人层次第 p 个自变量对最终结果的影响系数 (B_p 值)。并且, 由于 $G_{p1} \times \text{生育政策}$ 的存在, 还会影响到 G_{p0} (即个人层次自变量对最终结果的影响系数) 的估计值。当某个交互效应 G_{p1} 统计性不显著时, 即说明这一条层际影响途径可以忽略不计, 于是有 $B_p = G_{p0}$, 那么个人层次第 p 个自变量对结果变量的影响系数就成了一个固定常数, 不受上层变量的影响。

当个人层次因变量为出生性别这种虚拟变量时, 上述示意图及各种系数的意义并没有本质上的变化, 但是个人层次不能再用常规回归, 必须要改用 logistic 回归方程, 即:

$$\log[P/(1-P)] = B_0 + B_1 \times \text{有否男孩} + B_2 \times \text{教育程度}$$

其中, P 为原虚拟编码因变量 Y 取值 1 时所对应的那一性别所占的概率。如果将生男编码为 1、生女编码为 0, 模型中的 P 就是生男孩的概率, 于是, 这时的发生比 $P/(1-P)$ 的意义其实就是条件出生性别比。所谓条件出生性别比指控制了有否男孩和教育程度影响条件下的出生性别比, 具体条件还取决于这两个自变量是如何取值或编码的。

由于 logistic 回归中的因变量本来是表示性别的 0、1 编码, 只代表类别不同, 不具备尺度信息, 因此 logistic 回归系数也就不具有确定尺度信息。也就是说, logistic 回归系数提供的是自变量影响的相对幅度, 而不是绝对幅度。这时通常采用优势比 (odds ratio) 来表示两个类别出生性别比之间的比。当以前有男孩编码为 0 时, 这一类就是参照类, 这个自变量的优势比就是, 在控制了教育程度影响的条件下, 没有男孩这类妇女的出生性别比相比已经有男孩这类妇女的出生性别比的倍数。优势比定义为 $\exp(B_p)$: 即当 $B_p > 0$ 时有 $\exp(B_p) > 1$, 代表生男概率大于生女概率; 当 $B_p < 0$ 时有 $\exp(B_p) < 1$, 代表生男概率低于生女概率。尽管 logistic 回归不能提供各自自变量影响的绝对幅度, 但是它所提供的相对影响幅度仍然是十分重要的信息。并且, 出生性别比本身已经就是相对数指标, 而我们恰恰关注的是不同条件下出生性别比会发生哪些变化。

由以上对分层模型的说明可知, 当 logistic 回归作为个人层次模型纳入分层模型以后, 这一条件出生性别比的意义就不同于一般的 logistic 回归了, 因为它不仅控制了个人自变量的影响, 也控制了地区自变量与个人自变量之间的交互效应, 并且这一条件出生性别比本身还包含着地区自变量的直接影响。

本研究采用 HLM 6.02 软件 (Raudenbush et al. 2004) 进行分层模型分析, 它可以允许个人层次采用 logistic 回归 (或其他非线性回归), 而地区层次自变量对个人层次自变量系数的影响则采用线性回归。

这样一来, 采用非线性分层模型对五普的出生性别比进行研究, 既可以在地区层次借助其他来源的生育政策人口比例的信息, 又可以避免生硬地对每个妇女做生育政策的生硬“分配”。^① 并且, 分层模型可以将地区层次的生育政策变量和妇女个人层次的其他变量结合在一起, 来解释出生性别比的失调。此外, 分层模型还可以根据样本规模和模型设置检验各层解释变量效应的统计显著性。

3 数据处理

本研究数据处理的具体操作步骤如下:

^① 实际上这是我近年来努力钻研分层模型方法的初衷。在进行本研究之前, 我还曾收集了大量其他的地区级社会经济统计数据, 但是由于有相当数量的地区数据缺失, 所以没有在本研究中包括这些社会经济变量。

(1) 个人层次数据

从五普数据中选择在普查前一年有生育的育龄妇女(11735人)。变量包括此次生育子女的性别、此次生育的孩次、此次生育前是否尚无男孩、教育程度、户籍性质、民族成份、城乡居住地、人口流动状态。这些变量都是分类变量,因此先对各变量的类别做了适当合并,以减少模型中自变量的数目。

数据中存在少量一年中有两次生育的育龄妇女,只有123例(出生数为246)。从两个孩子的生育月份的交互分析发现绝大部分都是同月出生,因此可认为是双胞胎。数据检查发现两个孩子性别相同的占62%。这些双胞胎中男孩为144人,女孩为102人,总出生性别比为141.1,前一孩的出生性别比为192.9,后一孩性别比为 $63/60 \times 100 = 105.0$ 。因为这些双胞胎占出生总量比例很小,且后一孩的出生性别比并不失调。所以为了数据处理方便,我们将忽略双胞胎中的第二个孩子。

数据中还有63人的户口性质未知,所占比例极小,就作为农业户口处理。

关于此次生育前是否尚无男孩是根据育龄妇女申报的曾生男孩数与本年生育子女性别联系起来进行判断的。

是否流动人口按本人户籍登记状况来划分,将五款人中的第二、三、四类(共同特征为户籍均不在本乡镇街道)作为流动人口。

有关个人变量的定义、编码方式和描述性统计,详见以下两个模型相应的变量信息表。

(2) 地区层次变量

地区级变量采用国家计生委1999年统计的各地区中不同生育政策覆盖的人口比例,^①共有345个地区,平均每个地区只有34人。由于每个地区平均拥有人数并不多,就意味着无论是个人层次还是地区层次的方程中都不可能容纳较多的自变量。我们将一孩生育政策的人口比例和一孩半生育政策的人口比例合并,称之为较严政策覆盖比例。

(3) 将个人层次数据与地区层次数据联接起来的识别码(ID)

2000年全国人口普查1%数据样本提供了每个人的国标地区识别码。

但是,由于人口流动现象很多,实际上个人所执行的生育政策并不是居住地实行的生育政策,而是自己户籍地实行的生育政策。因此,对生活在户籍地所在地区以内的妇女,直接采用样本提供的地区地址码;而对那些不在其户籍所在地区的妇女,其地区识别码换用个人对五普问题“从何地来本乡镇街道居住”(r10)回答时所申报的原籍地区地址码。

再将国家计生委调查的各地区生育政策人口比例按地区赋予五普采用的相应国标地区码,就能够在分层模型软件HLM6.02中使妇女个人数据与其户籍地区的生育政策信息联接起来了。

4 分析结果与讨论

(1) 对五普前一年内所有孩次出生性别比的模型分析

第一个分层模型对五普前一年内所有孩次的出生性别比进行分析。因变量是代表育龄妇女在五普前一年内生育子女性别的虚拟变量,其他个人层次和地区层次解释变量的说明和描述性统计在表1中提供。

首先计算了一个完全无条件模型(即个人和地区两个层次的模型中都没有自变量):

$$\text{个人层次模型: } \log\{P/(1-P)\} = B_0$$

$$\text{地区层次模型: } B_0 = G_{00} + U_0$$

由于因变量取值1时代表生男孩,所以模型中的P就是生男孩的概率。而这时的发生比 $P/(1-P)$ 的意义其实就是出生性别比。

^① 对这一生育政策数据的专门分析参见郭志刚、张二力、顾宝昌、王丰(2003)的论文。

表 1 第一个分层模型中两层解释变量的定义与描述性统计

变量	定义	比例或 平均值	标准差
生育子女性别	育龄妇女在五普前一年内生育的性别 虚拟变量: 1= 男孩, 0= 女孩	0.54	0.50
个人特征	个人案例数: 11735		
无男孩	此次生育前尚无男孩: 是= 1, 否= 0	0.88	0.33
孩次 1	此次生育是否为孩次 1: 是= 1, 否= 0	0.69	0.46
孩次 2	此次生育是否为孩次 2: 是= 1, 否= 0	0.26	0.44
农业户口	是否为农业户口: 是= 1, 否= 0	0.79	0.41
低教育	是否教育程度为初中及以下: 是= 1, 否= 0	0.83	0.38
汉族	是否为汉族: 是= 1, 否= 0	0.87	0.34
城乡	是否居住于城镇: 是= 1, 否= 0	0.34	0.47
流动	是否户籍不在本地区: 是= 1, 否= 0	0.07	0.25
地区特征	地区单位数: 345		
较严政策比例	一孩生育政策与一孩半生育政策 在本地区的合计人口覆盖比例	0.80	0.33

完全无条件模型得到的截距 $B_0 = G_{00} = 0.175779$, 其优势比 $\exp(0.175779) = 1.192$ 。其实这个数值就是五普样本以小数表达的总出生性别比 119.2, 分层模型输出的相应 95% 的置信区间为 (1.147, 1.239)。

下面我们将在个人层次模型加入上述 8 个解释变量及层 - 1 截距, 并在地区模型中分别用生育政策变量加一个层 - 2 截距来参与拟合层 - 1 的 9 个参数估计。通过这种方式, 就可以将个人层次和地区层次的数据信息结合起来分析, 不但可以检查个人层次变量对出生性别比是否存在显著影响, 而且还能检查生育政策是否影响出生性别比, 以及它是如何影响出生性别比的。

在模型的两层都分别加入自变量以后, 个人层次模型中截距 B_0 就是条件出生性别比。我们先假设各地区的生育政策(较严政策覆盖的人口比例)不同会使各地区出生性别比偏离这一参照水平, 所以在地区层次的截距模型(求解 B_0 的模型)中纳入了关于生育政策的宏观自变量, 并且还加上了随机波动项。对于其他个人层次自变量, 也同样都假设生育政策对它们与出生性别比之间的斜率关系有影响, 但我们假定这种影响是固定的, 所以不再设随机波动项。最后, 我们将根据模型的拟合结果检验生育政策是否对出生性别比失调有影响, 并可检查它的影响体现在什么方面。

由于这一研究的初始模型所含参数很多, 然而绝大多数参数估计的统计性都不显著, 所以不再专门列表提供。初始模型中只有 3 个参数估计在 0.05 水平显著, 它们是: 尚无男孩与生育政策的交互效应, 孩次 1 的单独效应, 孩次 1 与生育政策的交互效应。实际上, 初始模型结果已经可以肯定生育政策与子女性别构成以及孩次之间的交互效应对出生性别比有显著影响了。

至于初始模型中其他各项效应的统计性不显著, 有可能是因为这些自变量之间相关程度太高而产生了多元共线性问题。所以, 我们通过依次删除模型中效应最不显著的变量来寻找最终得以通过显著性检验的那些变量。^①

表 2 中提供了最终模型的系数估计、按稳健标准误计算出来的显著度、优势比的估计值及其 95%

① 需要注意的是, 社会变量之间普遍存在相关, 所以模型中显著的自变量总是不可避免地在一定程度上“代表”着其他未列入模型的变量, 有时其模型解释能力中的大部分都只不过是模型中并未明列的其他变量那里“借”来的。所以, 要特别注意最终模型保留的那些自变量影响的方向是否与初始模型保持相同。

置信区间。我们发现,除了在初始模型中就已经显著的那 3 个变量以外(回归系数方向未变),又多了 4 项得以显著的参数估计(系数方向也未变),即层 - 1 截距项中的生育政策效应(G01),孩次 2 与生育政策的交互效应(G31),以及农业户口的单独效应(G40)及其与生育政策的交互效应(G41)。下面对分析结果做一些说明。

最终模型的截距 B0 是条件出生性别比,即在控制了生育政策影响条件下、个人层次模型参照类的出生性别比。本模型的总参照类(所有个人自变量全都取 0 值)就是:已经有过男孩、此次生育为三孩及以上孩次、非农业户口的育龄妇女。由于现在个人层次回归方程的截距值还取决于地区级生育政策(即 $B0 = G01 \times$ 较严政策比例),已经不再是一个常数。所以,即使其他各自变量都完全没有作用,仅此一项的统计性显著已经肯定,一个地区较严政策人口比例越大其出生性别比越高。G01 对应的优势比 $\exp(0.407259) = 1.503$ 的实际意义是,当较严政策的人口比例变化 1 个单位(即从 0 变为 100%时),它导致参照类出生性别比变化 1.503 倍!这是对各地区生育政策对出生性别比直接影响的一种统计归纳。^①

生育政策严与尚无男孩之间形成的层际交互效应(G11)的统计性极为显著(0.000),因为 $G11 = 0.769$,是个正值,表明在其他情况相同时,较严生育政策的覆盖面越大,以前没男孩的育龄妇女的出生性别比就越偏高。这里并不是说生育政策就是男孩偏好的原因,而是说在存在男孩偏好为既成事实的社会条件下,生育政策严时个人就很难通过扩大生育数量的途径来满足这一偏好,所以会采取其他方式(如选择性人工流产)来达到目的,以避免直接与生育政策对抗。

表 2 所有孩次出生性别比的分层模型分析的最终结果

固定效应估计	系数估计	显著度	优势比	95%置信区间
层 - 1 条件 SRB 截距: B0				
较严政策比例, G01	0.407259	0.005	1.502694	(1.136, 1.988)
层 - 1 无男孩 斜率: B1				
较严政策比例, G11	0.769265	0.000	2.158179	(1.779, 2.618)
层 - 1 孩次 1 斜率: B2				
层 - 2 截距, G20	-0.290634	0.020	0.747789	(0.586, 0.955)
较严政策比例, G21	-0.748402	0.000	0.473122	(0.322, 0.695)
层 - 1 孩次 2 斜率: B3				
较严政策比例, G31	-0.361970	0.003	0.696303	(0.550, 0.882)
层 - 1 农业户口 斜率: B4				
层 - 2 截距, G40	0.216998	0.007	1.242342	(1.064, 1.451)
较严政策比例, G41	-0.289131	0.004	0.748914	(0.618, 0.907)

孩次 1 的单独影响系数(G20)是负值,表明孩次 1 的出生性别比显著低于孩次 3+(孩次的参照类)时的相应水平。这与实际经验是一致的。

孩次 1 与生育政策之间的交互效应(G21)以及孩次 2 与生育政策的交互效应(G31)都是负值,表示在其他情况相同时,生育政策越严,一孩和二孩相对于多孩在出生性别比上的比值会下降。换句话说,就是生育政策越严,多孩性别失调严重化的速度会比一、二孩次更快。

农业户口的一般效应和与生育政策的交互效应统计性都是显著的。一般效应为正值,其优势比

^① 从统计分析而言,较严生育政策覆盖比例可以从小数形式更换为百分点形式,那么 $\exp(G01)$ 表达的便是这一比例提高一个百分点时的出生性别比变化率。自然也能够计算出若干个百分点提高导致的出生性别比变化率。就本研究而言, $\exp(G11/10)$ 为 10 个百分点导致的变化率,代入上述 G11 值,得到的相应变化率为 1.042 倍。

表明在其他条件不变时农业户口育龄妇女的出生性别比是非农户口的相应水平的 1.24 倍。而农业户口与政策的交互效应则是负值,说明较严政策比例越高,两类出生性别比的比值反而越低。

总的来说,以上分层模型结果不但肯定了生育政策对出生性别比失调有显著影响,而且显示了它是以哪些途径对个人生育施加其影响的。在个人层次,显著性的变量有以往生育的数量和性别两个生育史因素,^①而个人其他社会特征中只有户口性质的影响得以肯定。

在回归分析中,一些自变量不显著,并不一定表明它们真的没有实际作用,可能存在其他原因,不要轻易放弃。

回归系数统计性不显著的一个原因是前面已经提到的自变量之间的多元共线性问题,它不仅发生于定距自变量之间,也同样发生于标识编码自变量之间。^②除了扩大样本规模(通常做不到)以外,一个办法是事先合并类别,另一个办法是从模型中删除不显著的自变量,两种办法都是为了减少多元共线性问题的影响。

就本研究具体情况而言,还可能存在另一个原因,就是样本中一孩出生所占的比例很大(69.3%),而样本的一孩出生性别比(106)其实并不失调,也就是说它几乎是天然决定。所以,当一孩出生混在分析中时,反而会干扰寻找和检验出生性别比失调的解释变量。根据这个思路,下面我们将选择样本中有二孩及以上孩次的出生的妇女案例再做一次分层分析。

(2) 对五普前一年内的二孩及以上孩次出生性别比的分析

对二孩及以上孩次生育的分析模型所有的变量其实与前面的分析相同,只不过不再包括孩次 1 变量,表 3 提供了这个孩次选择样本中各变量描述性统计。

表 3 第二个分层模型中两层解释变量的定义与描述性统计

变量	定义	比例或 平均值	标准差
生育子女性别	育龄妇女在五普前一年内生育的性别 虚拟变量: 1= 男孩, 0= 女孩	0.61	0.49
个人特征	个人案例数: 3603		
无男孩	此次生育前尚无男孩 是= 1, 否= 0	0.61	0.49
孩次 2	此次生育是否为孩次 2 是= 1, 否= 0	0.84	0.36
农业户口	是否为农业户口; 是= 1, 否= 0	0.95	0.22
低教育	是否教育程度为初中及以下; 是= 1, 否= 0	0.95	0.21
汉族	是否为汉族; 是= 1, 否= 0	0.82	0.39
城乡	是否居住于城镇; 是= 1, 否= 0	0.19	0.39
流动	是否户籍不在本地区 是= 1, 否= 0	0.05	0.22
地区特征	地区单位数: 326		
较严政策比例	一孩生育政策与一孩半生育政策 在本地区的合计人口覆盖比例	0.81	0.33

个人层次模型的因变量为表示所生二孩及以上孩次的子女性别的虚拟变量,其中男孩所占比例 61%。也就是说,二孩及以上孩次的无条件的发生比为 $P/(1-P) = 0.61/(1-0.61) \times 100 = 156$, 可

① 这与乔晓春(2004)的 logistic 回归结果是一致的。

② 认为代表同一分类变量不同类别的虚拟变量之间相互独立是不对的。虚拟变量(组)只能保证每个案例唯一确定性地属于某一类,但是同组之间的虚拟变量之间肯定相关,因为起码参照类所有案例的所有虚拟变量值都是 0。至于不同组的虚拟变量之间的相关也普遍存在,比如代表城乡居住地和户口性质这两个虚拟变量之间的相关程度肯定非常高。

见这时性别失调表现得极为严重。从表 3 还可看出,生育二孩及以上孩次的妇女中在此次生育前有 61%还没有男孩,有 95%为初中及以下的教育程度,有 95%为农业户口,等等。这些信息本身就很有启发性。

完全无条件模型得到了更为精确的截距计算值为 $B_0 = G_{00} = 0.455262$, 其优势比 $\exp(B_0) = 1.577$ 。这个结果说明五普样本的二孩及以上的出生性别比为 157.7, 模型输出的 95%置信区间为 (1.455, 1.708)。

由于研究的主要关注是,在控制个人特征条件下地区生育政策不同对二孩及以上孩次出生性别比失调的影响,因此还是从假设生育政策对所有个人层次回归系数都有影响的初始模型着手,然后根据每次拟合结果将那些不显著的效应从模型中删去,最后得到最终解释模型。最终模型中所有保留变量的影响方向都仍与初始模型相同,拟合结果列在表 4 中。

这一最终模型的总参照类是:已经有男孩、此次生育为三孩及以上孩次、非农业户口、教育程度在高中及以上的育龄妇女。最终模型有 $B_0 = G_{01} \times$ 较严政策比例。值得注意的是,这时的 G_{01} 系数的优势比(1.584)高于以上对所有孩次分析时的相应水平(1.503)。

表 4 二孩及以上孩次出生性别比的分层模型分析的最终结果

固定效应估计	系数估计	显著度	优势比	95%置信区间
层-1 条件 SRB 截距: B_0				
较严政策比例, G_{01}	0.459669	0.046	1.583550	(1.008, 2.488)
层-1 无男孩 斜率: B_1				
较严政策比例, G_{11}	0.896163	0.000	2.450183	(2.019, 2.973)
层-1 孩次 2 斜率: B_2				
较严政策比例, G_{21}	-0.411082	0.001	0.662933	(0.526, 0.836)
层-1 农业户口 斜率: B_3				
较严政策比例, G_{31}	-0.480408	0.012	0.618531	(0.426, 0.899)
层-1 低教育 斜率: B_4				
层-2 截距, G_{40}	0.297463	0.001	1.346439	(1.138, 1.593)

在对二孩及以上孩次分析的最终模型中,尚无男孩与生育政策的交互效应(G_{11})仍然十分显著,优势比达到 2.45 倍,也大于以上对所有孩次出生分析时的相应倍数(2.16 倍)。

孩次 2 与生育政策的交互效应系数(G_{21})为负值,它代表二孩的条件出生性别比低于更高孩次,也就是说更高孩次的性别比失调程度更为严重。

不同于以上对所有孩次的模型分析,在对二孩及以上孩次的分析中,低教育程度这个社会变量也得到了统计性显著的结果。其单独影响的系数(G_{40})为正值,反映出提高教育程度对缓解出生性别比失调是有作用的。

在本模型中,农业户口的单独作用由于统计性不显著已经被删除了,但是它与生育政策的交互效应系数(G_{31})仍然得以保留。其系数仍然还是取负值。也就是说,在较严政策人口比例和其他条件相同时,农业户口的条件出生性别比要低于非农业户口。这种结果似乎不太容易理解。其实我们在前面对所有孩次进行分析时,就遇到了类似结果,但是前面只用统计学描述一笔带过,而没有加以解释。

为了能够更清楚地说明这一结果的实际意义,这里专门对五普 1%样本做了进一步的分析。将各地区的较严政策比例值按样本中所有育龄妇女户籍所在地区匹配后附加到个人记录上以后,再将较严政策比例这个定距变量划分为三类:一类为比例极高,一类为比例相对较高,一类为比例相对较低。由于较严政策覆盖的地区比例和人口比例都很高,在确定分界点时考虑了使三类中的出生数大致平衡。

结果如表 5 所示,确实反映出在较严政策比例极高和较高两类中,非农户口育龄妇女的二孩及以上出生数虽然极少,但的确在出生性别比上显著高于农业户口妇女。^①只是在较严政策覆盖面较低的类别中,非农户口的出生性别比低于农业户口的相应值,但这类出生所占权重较小,自然对回归系数估计影响较小。因此,在生育政策更严的地区中,非农户口的二孩及以上出生性别比甚至高于农业户口妇女的水平这一统计现象,可能反映了在其他条件不变时,非农户口的夫妇一旦决定要生育二孩或多孩便更有能力设法实现产前性别选择或其他手段。并且从表 5 计算的农业户口比非农户口的出生性别比优势比按较严政策人口比例扩大的方向看,至少从两端优势比值相比确实也是大大减小了。当然,与分层模型分析结果相比,表 5 这种交互表分析很粗糙,本身受分类数与分类界限影响较大,而且并未控制其他个人变量的影响。

总之,专门分析二孩及以上的出生时,尽管样本规模大大缩小了,分层模型分析仍然能够肯定生育政策和个人生育史变量对出生性别比影响的存在,而且有更多的社会特征变量取得了统计性显著的结果。

表 5 二孩及以上孩次按所户籍性质和户籍所在地区政策类型交互划分的出生性别比

户籍性质	1. 较严政策 $\geq 99.5\%$		2. 较严政策为 77.5~99.5%		3. 较严政策 $< 77.5\%$	
	出生性别比	案例数	出生性别比	案例数	出生性别比	案例数
农业户口	178.8	1344	156.6	1088	130.0	989
非农户口	216.0	79	250.0	70	65.2	38
优势比:农:非	0.828		0.626		1.993	

5 小结

本文对以往出生性别比研究中的某些重要观点做了一些评论,还对研究生育政策与出生性别比之间关系的几个经验统计分析中存在的问题进行了讨论。

在此基础上,本文简要介绍了分层模型分析方法,并应用非线性分层模型将 2000 年全国人口普查 1%数据和 1999 年的地区级生育政策数据结合起来对出生性别比失调的影响因素做了初步的分析。

对所有孩次出生和二孩及以上孩次出生的两个分层模型的结果表明,生育政策不同不仅对各地区的出生性别比有直接影响,而且还通过与育龄妇女个人的生育数量及子女性别结构之间的交互效应以及与个人的户口性质、教育程度之间的交互效应来影响出生性别比。此外,分析还肯定了育龄妇女以往生育结果和某些社会特征对出生性别比还存在着单独的影响。

以上统计结果启示我们,当前出生性别比失调是多种因素作用的结果,所以需要综合对策治理。一方面我们应当正视现行生育政策对出生性别比失调存在影响,同时我们也不能指望调整甚至取消生育政策就能使出生性别比完全恢复平衡,因为妇女在家庭、社会、经济方面的弱势地位尚未改变,养老保障、财产继承、宗亲关系等方面也很难改变,因此既作为文化遗产、又作为社会现实反映意识的男孩偏好也不是那么容易改变的。

尽管当前政府、学界、媒体、公众都对出生性别比失调问题十分关注,但是议论很多、定量研究很少。并且,在定量研究中,描述性分析较多,解释性分析较少,并且对生育政策是否影响出生性别比的研究更少。最后一种状况显然是与得不到生育政策数据有关,导致许多不同观点只能在理论层面争论,而不能结合实际数据进行理论检验。这种状况亟待改善。

本研究应用非线性分层模型将不同层次、不同来源的数据结合在一起分析,能够更有效地开发和

^① 这里提供的汇总结果并不想推断总体,只是为了印证上述分层模型的估计结果。

利用信息。在应用分层模型研究出生性别问题时,下层当然是以育龄妇女为分析单位最好,但是上层以地区为分析单位则是无可奈何的选择,因为只有这一层的生育政策数据可得,并且能够与普查数据中的区划相对应。如果能有县、乡、村作为二层分析单位更好,因为这种脉络层次距离育龄妇女更为接近。此外,由于数据限制,本研究中没有考虑经济等其他方面的变量也是一个缺憾。所以,本研究仍然只是这方面的一个初步分析。

分层模型是一种较新的分析手段,特别是非线性的分层分析,有关应用研究文献很少。本文一方面是为了给出生性别比研究提供一些新的定量分析结果,同时也是为了介绍一种新的研究手段。本研究是笔者首次尝试应用非线性的分层模型分析,在应用和结果解释中可能会存在一些问题和不足,就此机会求教于各位专家和同行。

参考文献:

- 1 Goldstein, H. Multilevel Statistical Models. (2nd ed). New York: John Wiley, 1995
- 2 Kreft, I. and J. de Leeuw. Introducing Multilevel Modeling. London: Sage Publications, Inc. 1998
- 3 Raudenbush, S. W. and A. S. Bryk. Hierarchical Linear Models: Applications and Data Analysis Methods, Second Edition. London: Sage Publications, Inc. 2002
- 4 Raudenbush, S., A. Bryk, Y. F. Cheong, R. Congdon, and M. Du Toit. HLM6: Hierarchical Linear and Nonlinear Modeling. SSI: Scientific Software International Inc. 2004
- 5 艾尔·巴比(邱泽奇译). 社会研究方法. 华夏出版社, 2000
- 6 陈卫. 中国的人工流产——趋势、模式与影响因素. 科学技术文献出版社, 2005
- 7 陈友华. 关于出生性别比的几个问题——以广东省为例. 中国人口科学, 2006; 1
- 8 顾宝昌, 罗伊. 中国大陆、台湾省和韩国出生性别比失调的分析. 人口研究, 1996; 5
- 9 郭志刚, 张二力, 顾宝昌, 王丰. 从政策生育率看中国生育政策的多样性. 人口研究, 2003; 5
- 10 郭志刚. 分析单位、分层结构、分层模型. 北京大学社会学学刊: 第一辑. 北京大学出版社, 2004(本文全文转载于中国社会科学院社会学研究所编. 中国社会学第三卷. 上海人民出版社, 2004)
- 11 郭志刚. 2000年人口普查按生育政策类型的人口分析. 2000年人口普查国家级重点课题研究报告. 中国统计出版社, 2005
- 12 郭志刚, 李剑钊. 农村二孩生育间隔的分层模型研究. 人口研究, 2006; 4
- 13 乔晓春. 性别偏好、性别选择与出生性别比. 中国人口科学, 2004; 1
- 14 杨菊华. 生育政策的地区差异与儿童性别比关系研究. 人口研究, 2006; 3
- 15 杨菊华. 多层模型在社会科学领域的应用. 中国人口科学, 2006; 3
- 16 原新, 石海龙. 中国出生性别比偏高与计划生育政策. 人口研究, 2005; 3
- 17 张二力. 从“五普”地市数据看生育政策对出生性别比和婴幼儿死亡率性别比的影响. 人口研究, 2005; 1
- 18 周云, 任强. 高龄老人1910-1969年间出生子女性别比的研究. 北京大学学报(哲学社会科学版), 2004; 4

Multilevel Analysis of Sex Ratio at Birth in China 2000 Population Census

This paper comments on some studies of sex ratio at birth by discussing methodological issues in these empirical analyses. The major purpose of this paper is, however, to conduct a multilevel investigation of sex ratio at birth using micro sample data from China 2000 population census and data of fertility policy at prefecture level. Results show that fertility history and social characteristics of women significantly affect their sex ratio at birth, while macro level fertility policy also influences significantly sex ratio at birth both directly and through the interaction with individual fertility history and social variables.

Keywords: Sex ratio at birth, Fertility policy, Multilevel modeling

Author: Guo Zhigang is Professor, Department of Sociology, Peking University.