

对伤残死亡独生子女母亲人数的初步测算^{*}

王广州 郭志刚 郭震威

【摘要】 文章在对独生子女、独生子女家庭和独生子女伤残死亡家庭概念进行界定后,以孩次递进人口预测方法为基础,在特定生育间隔分布假定条件下,构造独生子女及其母亲二维年龄结构分布,并对实行计划生育以来伤残死亡独生子女母亲人数进行估算。

【关键词】 独生子女 伤残 死亡 母亲

【作者】 王广州 中国社会科学院人口与劳动经济研究所,副研究员;郭志刚 北京大学社会学系,教授;郭震威 国家人口计生委发展规划司,副司长。

一、研究背景

自20世纪70年代实行计划生育政策以来,中国在抑制人口过快增长方面取得了明显的效果,少生了约3亿多人。为了响应政府号召,执行基本国策,无论城市还是农村许多家庭做出只生一个孩子的生育选择。研究表明(杨书章、王广州,2007),预计2007年底,中国独生子女总量累计已有1亿多人。正是由于计划生育夫妇生育独生子女的选择,一方面为国家迅速抑制人口过快增长、缓解人口压力做出了巨大贡献;另一方面也比多生育子女的夫妇承担了更大的无子女的风险。据全国第五次人口普查资料建立生命表推算,大约有3.91%的人活不到18岁,有5.1%的人活不到30岁,其中1.17%的人是在18~29岁死亡的。在18~29岁年龄段人群中,死亡率为1.1%。根据2006年第二次全国残疾人抽样调查,全国现有残疾人口8296万,占总人口的6.34%(第二次全国残疾人口抽样调查办公室,2007),但低年龄组人口残疾比例要远远低于这个数字。

随着中国计划生育政策的全面实施和生育水平的持续下降,独生子女总量呈现持续快速增长的态势,由于总量规模的持续增加,独生子女及其相关问题也越来越受到研究者的重视。因此,长期以来从不同研究角度对独生子女问题进行了许多研究,如独生子女教育、健康、心理和社会适应等(风笑天,2006)。与相关学科相比,人口学对独生子女问题的研究主要集中在独生子女总量结构(杨书章、郭震威,2000;郭志刚,2001;宋健,2006;杨书章、王广州,2007)、独生子女生育政策和独生子女家庭结构(郭志刚等,2002;丁仁船等,2007)等。而现有对独生子女的总量和结构的研究,则主要集中在对现有独生子女总量的估计和对未来人口发展影响方面上。从人口学研究领域来看,上述相关研究主要是回答现存独生子女的总量结构问题。虽然对独生子女及其风险也有一些研究(潘金洪、姜继红,2007),但对实行计划生育政策以来全国死亡或伤残独生子女总量结构方面还没有公开发表的研究成果。

* 国家社会科学基金资助项目(批准号:06BRK010)阶段性成果之一。

2006年底发布的《中共中央国务院关于全面加强人口和计划生育工作统筹解决人口问题的决定》明确指出,要积极探索建立独生子女伤残死亡家庭扶助制度。前不久,有关政府部门决定在全国部分地区开展试点^①,对独生子女死亡伤残家庭给予必要的扶助。这不仅体现了政府对独生子女父母响应国家政策的利益补偿,而且体现了对这些家庭的人文关怀,更体现了社会和谐和公平正义的基本原则。本研究试图对死亡伤残独生子女母亲总量和结构进行初步测算,希望能够为这项制度在全国推行提供一些参考依据。

二、基本概念界定

(一) 独生子女与独生子女家庭

独生子女可定义为一对夫妇终身只生的一个孩子,判断的法定依据是以领取独生子女证为标志。但实际上由于全国各地持证的情况存在很大差别,不少地区独生子女父母因各种原因,没有领证或没有领到证;还有少数领取了独生子女证的父母并未停止生育。按照定义,只有结束生育才能认定终身只生一个孩子的妇女,如果非要等到生育期结束才确定其独生身份,就太迟了,也无法落实独生子女政策,达不到政策需要收获的社会效果(杨书章、王广州,2007)。

一个孩子是否是独生子女取决于父母的生育行为。从生育的角度看,独生子女与独生子女亲生父母存在一一对应关系。独生子女家庭以独生子女母亲(妇女)为代表。因为是否是独生子女不取决于独生子女本人,而是由其生母的生育行为决定的,因此我们考察的焦点是40岁以上现有一孩且仅有一孩和曾生一孩现无孩的妇女。

本研究将“独生子女”父母严格界定为已度过生育期、不再继续生育的只有一个孩子的妇女(“独生子女户”)及其丈夫。20世纪90年代以来,中国妇女在35岁以后生育的比例很小(见表1),可以忽略不计。为保险起见,本研究中将40岁作为妇女终止生育的年龄。

表1 35岁以上不同队列妇女的平均曾生子女数

年龄(1997年)	1997年	2001年	年龄(1990年)	1990年	2000年
35	2.00	2.02	35	2.25	2.20
36	1.91	1.93	36	2.35	2.26
37	1.94	1.94	37	2.49	2.36
38	2.05	2.09	38	2.62	2.46
39	2.15	2.14	39	2.76	2.56
40	2.14	2.15	40	2.93	2.79
41	2.14	2.18	41	3.08	3.10*
42	2.28	2.26	42	3.23	3.15*
43	2.30	2.33			
44	2.53	2.42			
45	2.56	2.49			

注:(1)1990年普查与2000年普查调查时点相差4个月,所以队列口径不完全一致,这是导致队列曾生子女数有缩减情况的原因之一。(2)1997年、2001年数据分别为1997年全国人口与生殖健康调查、2001年全国计划生育/生殖健康调查结果。1990年、2000年数据分别为“四普”、“五普”公布数,其中带“*”号的数据为1997年全国人口与生殖健康调查结果。

(二) 独生子女死亡、伤残概念

独生子女死亡事件比较容易界定,为了具有可操作性和明确独生子女对象,本项研究将0~30岁死亡独生子女定义为独生子女死亡事件发生。独生子女伤残采用中国残疾人联合会

① 详见《人口计生委财政部关于印发全国独生子女伤残死亡家庭扶助制度试点方案的通知》(<http://www.nxrk.gov.cn/info/zwgk/083626855.html>)。

认定标准,以三级及以上残疾作为伤残标识。由于伤残也和死亡一样,与年龄有很大关系,而独生子女死亡、伤残的认定涉及到两代人的生育、死亡或伤残,为了简化,同样将伤残独生子女年龄范围界定在0~30岁。

(三) 独生子女死亡、伤残家庭

按独生子女的母亲年龄状况和计划生育政策实施的历史,采用以下几种统计口径:(1)1933年以后出生;(2)年龄在49岁以上(或45岁以上、或40岁以上);(3)只生育过一个子女;(4)现无存活子女,或独生子女存活但为三级以上残疾。在现有数据条件下,实施上述条件的操作存在3个主要问题。一是对已经死亡的妇女,其独生子女夭折、伤残户就被漏统了。二是包括了曾经违反生育政策、但目前无子女或只有一个残疾子女的情况,导致存在高估的可能性。三是没有考虑妇女及其丈夫婚姻变动、子女收养、符合政策生育两个及以上孩子但目前无孩或只有一个残疾子女等特殊状况。所幸这些因素涉及的人群并不大,且有些因素的影响可在一定程度上相互抵消。

三、基础数据

由于本项研究的最终结果是对独生子女死亡伤残家庭人群进行测算,人口属性的基本划分维度是以农业户口和非农业户口、独生子女与非独生子女状况为划分标准,因此所需基础数据必须满足按农业或非农业、曾生子女孩次结构及分地区的交叉分类要求,而现有汇总数据无法满足测算所需基础数据要求。

鉴于研究方法和研究目标所需基础数据的复杂性,本研究必须使用原始数据进行重新汇总。从现有抽样调查和人口普查原始数据来看,人口抽样调查无法满足大数据量和覆盖全国的基本需求,因此可供选择的原始数据只有1982年“三普”、1990年“四普”和2000年“五普”资料。人口普查数据质量研究和普查事后数据质量抽样调查表明,上述原始数据的数据质量差异较大。从数据的完整性和准确性来看,“三普”数据质量最好,“四普”次之,“五普”最差;而从数据的现时性来看,“五普”最好,“四普”次之,“三普”最差。因此需要在数据质量、研究目标和数据现时性之间进行权衡并做出选择。本研究采用的基础数据以“四普”为主,“五普”为辅,同时以其他来源的数据作为必要的参考和补充,如2001年全国计划生育/生殖健康调查数据、2005年全国1%人口抽样调查数据。

独生子女死亡伤残家庭数量测算的运算过程是以育龄妇女的孩次、年龄结构和生育孩子孩次、年龄结构为基础进行分析、判断和推算的,即通过育龄妇女生育孩子的人口属性来进行推算的,是以育龄妇女的孩子状态作为标识和依据研究家庭户(父母)状况,运算所需基础数据涉及各年龄组的数据质量。因此,最终结果受人口的年龄结构数据质量影响很大。整个测算误差将会因数据按属性匹配、对应误差和本身误差双重放大或折中。所以,在考虑数据质量时,女性年龄结构及相关特征数据质量评判是关系到整个运算结果科学性关键问题。

测算将根据较可靠的20世纪90年代生育水平作为预测参数进行“打靶”,先预测至2000年,保证使其预测总人口基本上吻合2000年公布数。然后,以此性别、年龄结构和育龄妇女孩次结构为基数,继续完成全部预测。

四、模型、算法及参数设定

独生子女死亡、伤残家庭的认定涉及到独生子女父母和独生子女两代人,其测算方法不同

于现有奖励扶助目标人群测算等任何人口模型(“中国农村部分计划生育家庭奖励扶助制度”课题组,2003)。测算的目标人群是以独生子女父母的生育历史为基础、以孩子的属性为条件的人口群体估计方法。

测算研究存在3个难点:一是如何在人口预测中建立独生子女矩阵与其父母矩阵的密切联系;二是如何估算独生子女的伤残分布;三是将独生子女伤残分布与其父母联系起来。这在以往的人口预测中均没有现成、可靠的处理技术和估算模型,需要通过模型研制和调试结果逐步形成和完善,以得到比较合理、可信的测算结果。

(一) 独生子女及其母亲二维年龄结构分布

常规人口预测方法不能胜任此类独生子女及其母亲年龄结构二维分布研究的需要,即使是分孩次的年龄别生育预测也存在一定的问题(Feeney, 1985; 马瀛通等, 1986; Feeney等, 1987; Ni Bhrolchain, 1992; Siu-Man Ng, 1992)。主要是因为这类方法的生育预测是以时期年龄别妇女合计人口作为生育预测的基数,完全不考虑该年龄组业已形成的生育孩次结构。采用孩次递进预测方法则可以适应本研究所需要的主要预测工作。孩次递进预测是在条件概率生育的基础上进行分孩次的预测,在预测中,只有尚未生育的妇女才能生育一孩;只有仅生育了一个孩子的妇女才能生育二孩,等等。因此,孩次递进预测可以克服常规人口预测的上述缺陷。并且,由于采用预测概率进行生育预测,预测年份育龄妇女的生育数量实际上是分别按育龄妇女的年龄和孩次类型进行预测的,所以在生育预测完成后便可以进一步严格地根据各交互类别育龄妇女的生育数量来更新育龄妇女的孩次类型分布。

由于我们需要得到独生子女—母亲年龄结构二维分布,而从年龄—孩次递进预测模型可以看到,需要分孩次把子女年龄和其所对应的母亲年龄同时记录下来,才能得到独生子女—母亲年龄结构二维分布数据结果。因此,需要对上述模型中妇女按孩次、孩子年龄、妇女年龄分类,同时,还需要将孩子按孩次、年龄和母亲年龄进行记录。由于需要考察独生子女及其母亲的二维年龄分布,因此,只需要把一孩(无亲生兄弟姐妹)按年龄和母亲年龄进行记录,即需要形成各年度一孩按孩子年龄—母亲年龄存储的数据矩阵,可以用二维变量 $B_a^{(1)}$ 直观表示。这里 i 表示孩子年龄、 a 表示母亲年龄,上标(1)表示一孩。由于一孩(无亲生兄弟姐妹)的母亲与孩子存在一一对应关系,因此,也可以对一孩妇女进行类似的记录和表示,即可以用二维变量 $W_{ia}^{(1)}$ 直观表示, i 表示孩子年龄、 a 表示妇女年龄,上标(1)表示一孩。通过对 i 求和,将 $W_{ia}^{(1)}$ 与模型 $W_{a+1}^{(1)} = W_a^{(1)} - B_a^{(2)} + B_a^{(1)}$ 中描述的 $W_a^{(1)}$ 相对应;也可以通过 a 求和得到孩子的年龄分布 $B_i^{(1)}$ 。

年龄—孩次递进预测模型在妇女递进生育过程中,可以直接记录妇女的生育年龄和孩次,但模型 $W_{a+1}^{(1)} = W_a^{(1)} - B_a^{(2)} + B_a^{(1)}$ 中 $W_a^{(1)}$ 所包含的孩子年龄是被合并的,因此为了满足独生子女—母亲年龄结构二维分布变动过程记录,需要通过以下步骤完成模型的扩展:(1)通过母子匹配的方法从原始数据中获得一孩母亲和孩子的二维年龄分布,将 $W_a^{(1)}$ 转换为 $W_{ia}^{(1)}$ 或 $B_{ia}^{(1)}$ 。(2)将新递进的 $B_{ia}^{(1)}$ 记录为 $B_{0a}^{(1)}$ 或 $W_{0a}^{(1)}$,即为 a 岁妇女生育的 0 岁一孩。(3)将 $B_a^{(2)}$ 根据生育二孩间隔分布,转换成 $B_a^{(2)}$ 。(4)经过一轮递进后,将未递进的妇女和孩子及存活妇女—孩子年龄结构矩阵 a 和 i 分别增加 1 岁。

(二) 死亡独生子女及其母亲二维年龄结构分布

年龄—孩次递进预测模型经改造和拓展后,可以得到各个年度的独生子女的年龄结构,因此,可以通过生命表推算的、按母亲年龄分类的不同年龄死亡独生子女人数。由于独生子女和

母亲是一一对应的,母亲有独生子女死亡的算法可表示为: $W_{i,a+1}^{(1)} = W_{ia}^{(1)} \cdot (1 - \frac{L_{i+1}}{L_i}) \cdot \frac{L_{a+1}}{L_a}$;
式中 L_i 为 i 岁独生子女存活人年数; L_a 为 a 岁妇女存活人年数。

由于死亡概率和年龄密切相关,因此,这里独生子女年龄最大记录到 30 岁,而 $W_{ia}^{(1)}$ 所记录的数据一直是独生子女母亲,对应的孩子死亡,应该认为是独生子女死亡。死亡一孩母亲转化为生育二孩母亲的比例和规模取决于二孩递进概率。

虽然独生子女死亡记录将会受到独生子女母亲和独生子女本人死亡而发生目标人群“丢失”的现象,尤其是年龄较大的独生子女和独生子女母亲死亡的概率相对会大一些,但由于目标人群“丢失”的概率是两个死亡概率的乘积,因此,独生子女母亲和独生子女同时死亡发生目标人群“丢失”的可能性较小,最大“丢失”人数可以表示为: $W_{i,a+1}^{(1)} = W_{ia}^{(1)} \cdot (1 - \frac{L_{i+1}}{L_i}) \cdot (1 - \frac{L_{a+1}}{L_a})$ 。

(三) 伤残独生子女及其母亲二维年龄结构分布

伤残是对健康状态描述。从概率事件和实际观察的角度看,伤残不仅有发生和不发生的问题,还有什么时候被认定或发现的问题,同时还有严重程度的界定标准问题。因此,伤残与死亡事件明显不同。死亡事件只需要确定发生事件,而发生事件的结果是很容易确定和比较明确的,且具有单向不可逆的特点。相比之下,确定独生子女伤残总量和结构就显得更加复杂。

虽然伤残与死亡事件不同,但伤残和死亡都与年龄密切相关,或者说,经过认定的年龄别伤残发生的比例与年龄密切相关。一般而言,随着年龄的增长,伤残比例不断提高。本研究使用年龄别伤残比例,基于模型生成的各年度独生子女—母亲年龄二维分布,计算各年度独生子女目前处于残疾状态的总量和结构。从残疾发生的比例确定各年度独生子女伤残总量和结构需要两个基本前提和假设:(1)假定伤残发生的状态不可逆;(2)假定目前观察到的年龄别伤残状态的累计分布是比较稳定可靠的。

(四) 死亡伤残独生子女及其父亲二维年龄结构分布

虽然死亡伤残独生子女及其父亲二维年龄分布也可以和母亲一样分别建立模型进行推算,但由于生育史方面的基础数据是按妇女统计登记的,因此独生子女及其父亲的二维年龄结构只能通过独生子女及其母亲二维年龄分布进行推算。这样就需要建立父亲和母亲之间的联系。父亲和母亲之间的联系涉及对婚姻状态的描述。从婚姻关系构成来看,婚姻状态包括未婚、初婚有配偶、再婚有配偶、离婚和丧偶 5 种状态。对于独生子女母亲来说,可能的婚姻状态包括除未婚外的 4 种状态,父母联系在一起就可能出现 12 种状态。

利用妇女的在婚比例(即有配偶的比例)及妇女与丈夫的年龄差分布数据可以计算丈夫的人数(即根据妇女的在婚比例计算),然后再根据妇女与丈夫的年龄差分布,推算丈夫的年龄构成。

在离婚和丧偶率比较低的情况下,初婚有配偶占绝对比例。为了简便和保险起见,将死亡和残疾独生子女—父亲分布简化为死亡和残疾独生子女—母亲年龄二维分布,并假定父母同时存活,这样一方面夸大了独生子女父亲存活的比例,但另一方面却弥补了独生子女母亲死亡的情况。

(五) 预测参数设定

人口预测所需要的数据可分为预测基数和参数两大类。预测基数指预测起始年份的分年龄分性别人口数据,其中,妇女人数还应划分出孩次。预测参数是对预测时期人口出生、死亡、

迁移情况的设定。预测参数^①包括生育、死亡、人口城镇化、年龄别残疾人口比例等方面,而对于残疾发生率参数,由于中国残疾人口调查比较少,1949年以来,中国只有1987年和2006年进行过两次残疾人口全国性抽样调查。残疾人口年龄别发生比例的高低一方面与年龄大小有关,另一方面与农业、非农业性质不同有关。为此,本研究使用按农业和非农业户口划分的年龄别残疾人口比例进行测算。

五、测算结果及讨论

(一) 全国测算结果

基于上述参数和假定条件,可以根据前面的统计口径,估计妇女的总量和结构。按年龄在49岁以上统计口径可以得到2007~2050年独生子女死亡、伤残母亲总量(见图1、图2)。估计2007年49岁以上全国死亡独生子女母亲总量在30万以内。死亡独生子女母亲数量在2038年以前持续增长,2038年以后开始下降。峰值规模在110万左右,那么,死亡独生子女父母人数估计应该在220万以内。估计2007年全国49岁以上伤残独生子女母亲总量在22万左右,独生子女伤残父母人数在44万以内。伤残独生子女母亲目标人群规模将比死亡独生子女母亲数量提前达到峰值,预计49岁以上伤残独生子女母亲数量在2017年以前持续增长,2017年以后开始下降,峰值规模在40万左右,相应的,伤残独生子女父母数量估计应该在80万以内。从图2和图1还可以看出,40岁及以上伤残独生子女母亲数量与49岁以上的规模之差,同两种口径的死亡独生子女母亲数量之差明显不同,这主要是由于规模估计方法、参数设定和死亡等方面的原因引起的。

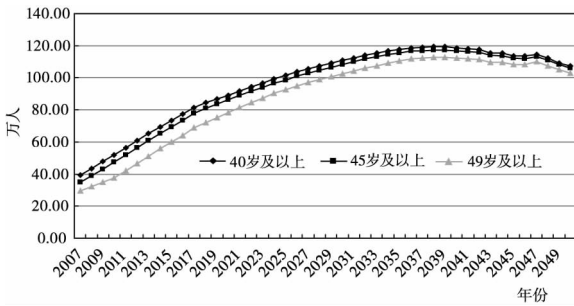


图1 全国死亡独生子女母亲人数估计

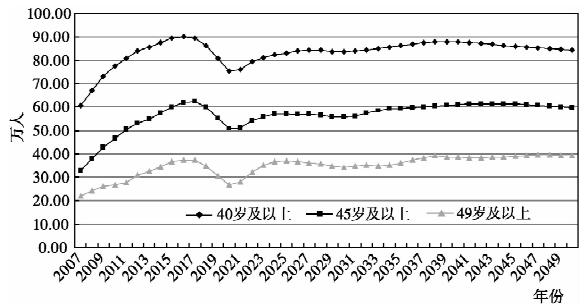


图2 全国伤残独生子女母亲人数估计

(二) 讨论

死亡伤残独生子女家庭数量测算是一项非常复杂的研究,无论测算方法设计还是数据运算规模都是目前人口预测中最为复杂的研究项目之一,需要在方法创新的同时,仔细研究预测参数的设置,尤其是在目前直接可供使用的数据比较缺乏的情况下,进行测算更加困难。考虑到数据来源、统计口径和参数设置等方面的原因,本研究结果与实际情况的可能偏差主要来源于以下几个方面。

第一,抽样基础数据。虽然2000年第五次人口普查和2005年全国1%人口抽样调查数据已经公布,但由于目前可获得的2000年人口普查原始数据仅为0.95%,而且存在比较严重的

① 生育、死亡和人口城镇化参数参见郭振威等:《2003~2050年农村实行计划生育的老年夫妇人数变动预测》,《人口研究》,2005年第2期。

漏报问题。2005年可供使用的原始数据的抽样比例为2%，数据质量和样本量较小都不适合省级为单位的研究任务，只能作为研究的参考。本研究采用1990年第四次人口普查1%的抽样原始数据，难免由于数据的现实性和抽样问题对一些省级单位存在一定的偏差。残疾发生率方面可利用的数据也非常有限。

第二，二孩递进率与二孩递进间隔分布。二孩递进率的高低一方面影响总体生育水平的高低，另一方面影响独生子女总量结构的大小。二孩递进间隔关系到子女和母亲二维年龄分布。实际上育龄妇女在不同年龄上有着不同的二孩递进间隔分布，然而由于可得数据规模限制及预测程序的简练，本测算对不同子人口采用的是相应人口类型汇总的递进间隔分布。

第三，城镇化水平。本研究根据户口性质将总人口划分为农业人口和非农业人口，并将农业人口视为农村人口，非农业人口视为城镇人口。由于城镇与农村计划生育政策存在明显差别，目前城镇和农村实际生育水平也存在明显的差异，因此，对城镇化水平（即假定人口非农化进程的快慢）和城乡生育水平差别大小的估计等，直接影响对独生子女总量结构的判断，从而进一步影响对目标人群的估计。

此外，本研究只给出了死亡、伤残独生子女的母亲人数；测算父母双方的人数，也仅仅是按照最简单的处理方法，将母亲人数乘以2，而没有更精细地测算父母双方的人数。这样简单处理既存在高估的可能，即妇女平均期望寿命高于男性，年龄较高的妇女（40岁以上）人数因此将多于其丈夫人数；也存在低估的可能，即对于父亲在、但母亲不在（去世或离家）的独生子女，其父亲被部分漏统了。两种因素相抵（不考虑抽样及预测参数等其他影响因素），得到的推算结果应该略高于实际情况，是较为保险的估计。

参考文献：

1. “中国农村部分计划生育家庭奖励扶助制度”课题组编(2003):《农村部分计划生育家庭奖励扶助制度》,《目标人群测算研究报告》,中国人口出版社。
2. 第二次全国残疾人口抽样调查办公室(2007):《第二次全国残疾人口抽样调查主要数据手册》,华夏出版社。
3. 马瀛通等(1986):《递进人口发展模型的提出与总和递进指标体系的确立》,《人口与经济》,第1、2期。
4. 杨书章、郭震威(2000):《中国独生子女现状及对未来人口发展的影响》,《市场与人口分析》,第4期。
5. 郭志刚(2000):《利用人口普查原始数据对独生子女信息的估计》,《市场与人口分析》,第1期。
6. 郭志刚、刘金塘、宋健(2002):《现行生育政策与未来家庭结构》,《中国人口科学》,第1期。
7. 风笑天(2006):《中国独生子女:规模、差异与评价》,《理论月刊》,第4期。
8. 宋健(2006):《中国农村独生子女的数量与分布》,《中国人口科学》,第4期。
9. 杨书章、王广州(2007):《一种独生子女数量间接估计方法》,《中国人口科学》,第4期。
10. 潘金洪、姜继红(2007):《江苏省独生子女数量测算及其风险分析》,《扬州大学学报(人文社会科学版)》,第1期。
11. 丁仁船等(2007):《独生子女比例、婚育意愿变动对未来政策生育率的影响》,《南方人口》,第3期。
12. Feeny, Griffith(1985), Parity Progression Projection. International Population Conference, Florence 1985, vol. 4. International Union for the Scientific Study of Population.
13. Feeny, Griffith and Yu Jingyuan(1987), Period Parity Progression Measures of Fertility in China. *Population Studies*, 41.
14. Ni Bhrolchain, Maire(1992), Period Paramount? A Critique of the Cohort Approach to Fertility. *Population and Development Review*. 18; No. 4.
15. Siu-Man Ng(1992), Age-parity-duration-specific Measures of Fertility; Application to China. PhD. Dissertation. Dec.

(责任编辑:朱犁)