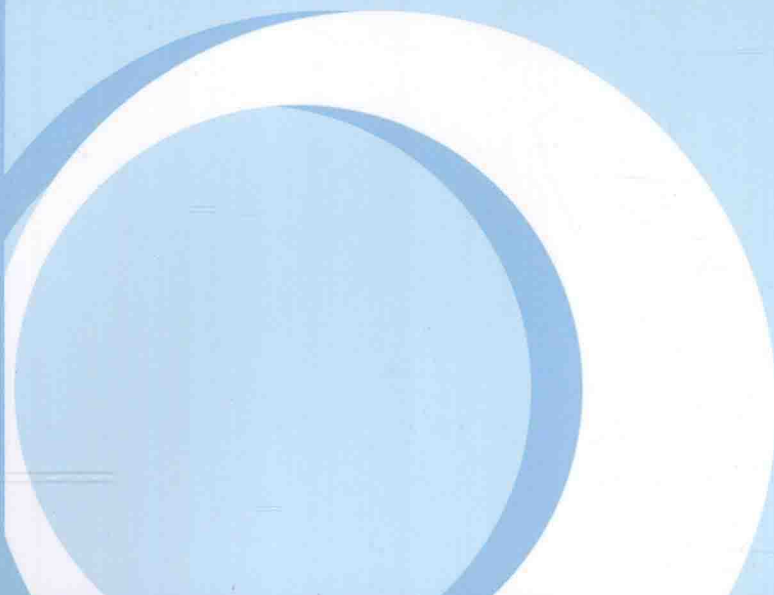


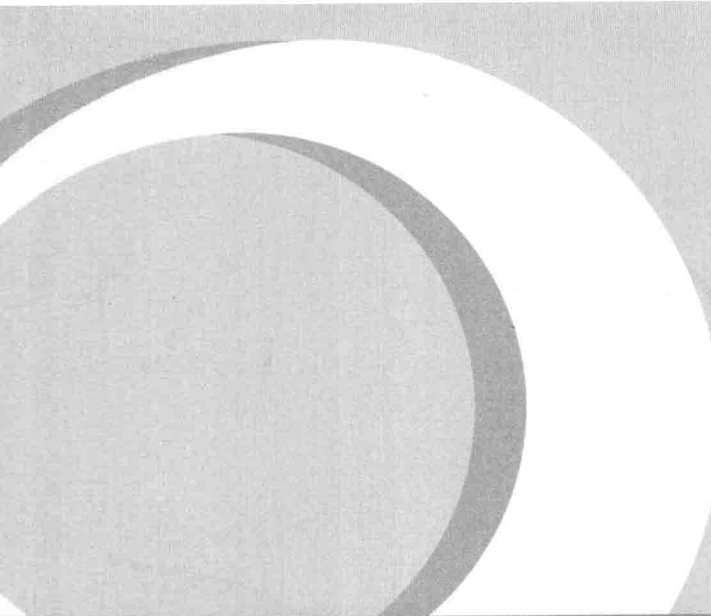


发展中的中国人口

2010年全国人口普查研究课题论文集

上册

国务院人口普查办公室
国家统计局人口和就业统计司 编



发展中的中国人口

2010年全国人口普查研究课题论文集

上册

国务院人口普查办公室 编
国家统计局人口和就业统计司

图书在版编目(CIP)数据

发展中的中国人口：2010年全国人口普查研究课题论文集 /
国务院人口普查办公室 国家统计局人口和就业统计司编.

——北京：中国统计出版社，2014.4

IISBN 978—7—5037—7053—1

I. ①发… II. ①国… III. ①人口普查—中国—文集

IV. ①C924.25—53

中国版本图书馆 CIP 数据核字(2014)第 020810 号

发展中的中国人口：2010年全国人口普查研究课题论文集(上、中、下)

作者/国务院人口普查办公室 国家统计局人口和就业统计司

责任编辑/徐 涛

封面设计/李雪燕

出版发行/中国统计出版社

通信地址/北京市丰台区西三环南路甲 6 号 邮政编码/100073

电 话/邮购(010)63376909 书店(010)68783171

网 址/<http://csp.stats.gov.cn>

印 刷/河北天普润印刷厂

经 销/新华书店

开 本/880×1230mm 1/16

字 数/3940 千字

印 张/136.5

版 别/2014 年 4 月第 1 版

版 次/2014 年 4 月第 1 次印刷

定 价/380.00 元

版权所有。未经许可,本书的任何部分不得以任何方式在
世界任何地区以任何文字翻印、拷贝、仿制或转载。
如有印装差错,由本社发行部调换。

编委会和编辑人员

编委会

顾 问：马建堂

主 任：张为民

副主任：冯乃林

编 委：彭永涛 张志斌 胡 英

编辑人员

总 编 辑：冯乃林

副总编辑：武 洁 孟灿文

编 辑：杨建春 肖 宁 李 睿 吴 珊

责任编辑：徐 涛

前 言

2010年,我国开展了第六次全国人口普查。在党中央、国务院的正确领导下,国务院有关部门密切配合,地方各级党委政府大力支持,人民群众积极参与,第六次全国人口普查获得圆满成功,取得丰硕成果,为深入研究我国人口的特征及其发展变化、人口和经济社会、人口和资源环境等问题提供了大量丰富翔实的数据资料。

为了做好第六次全国人口普查资料开发利用工作,最大限度发挥人口普查资料的社会效益,国务院第六次全国人口普查领导小组办公室于2011年9月组织了第六次全国人口普查研究课题公开招标,并成立第六次全国人口普查研究课题专家组,负责招标课题的选题、评审、鉴定工作。这次招标面向全国各级政府部门、大专院校及相关的研究机构和单位。专家组根据我国现阶段人口、经济和社会发展的需要,选择当前各级政府和社会关注的与人口相关的重点、难点问题,确定了14个课题研究方向,向社会公开发布。课题招标共收到竞标申请书237份,专家组根据《第六次全国人口普查研究课题招标管理办法》的规定,本着“公开、公平、公正”的原则,经过独立评分、集体评议、择优立项等环节,最终确定了中标课题。

第六次全国人口普查招标课题研究工作在全体课题研究人员辛勤努力下按时完成,全部研究成果通过专家组评审,准予结题。为满足社会各界需求,现将课题成果报告精编成《发展中的中国人口:第六次全国人口普查研究论文集》出版,论文集共收集85篇课题主要成果,按课题研究方向分为14个部分,涵盖了人口的生育、死亡、性别年龄结构、老龄化、婚姻家庭、迁移流动、城镇化、素质、民族、住房、就业失业、地区分布以及人口和经济发展、人口和资源环境发展等内容。论文集从多角度对普查数据进行深入挖掘和分析,全面展示了我国人口的发展变化特征,客观分析了人口与社会、经济、资源环境可持续发展的现状和前景,并在分析研究的基础上提出了相应的政策建议,具有较高的学术价值和实践指导意义。

本书在编辑过程中得到了课题承担单位的大力支持,在此我们表示衷心感谢。由于时间仓促,编辑工作难免有不妥之处,恳请广大读者给予批评指正。

编 者

2014年4月

目 录

上 册

一、人口生育水平

中国生育水平研究	3
2010 年全国人口普查的生育数据分析	30
基于 2010 年普查补充汇总数据的分省生育率估计	75
中国人口惯性百年:1950—2050	103
我国婚姻状态和育龄妇女生育率分析及放宽二胎政策的最佳时机探讨	120

二、人口死亡水平

中国人口死亡水平研究	153
中国人口死亡水平及影响因素研究	205

三、人口性别和年龄结构

中国人口性别结构研究	251
中国出生性别比孩次性别递进过程及影响因素研究	310
中国人口性别结构的演变及其后果研究	338
中国女性人口状况研究	365
中国儿童人口发展状况	392

四、人口老龄化

中国老龄化的特征和养老负担研究	423
人口老龄化与老龄问题研究	463
中国人口老龄化和老年人口问题研究	508
人口老龄化对中国养老保险制度的影响及对策研究	527
人口外流、老龄化与农村发展研究	553
人口老龄化对我国经济发展的影响研究	568
中国失能老人状况及演变趋势分析	592
以家庭和社区为基础的全方位老年护理服务体系研究	608



中 册

五、婚姻和家庭

中国家庭模式变迁的现状与趋势	633
中国婚姻家庭的变动及其经济社会影响研究	704
中国城乡家庭结构最新变动、特征和影响因素分析	735
中国家庭生命周期结构及其政策涵义	762
性别失衡对婚姻家庭的影响研究	775

六、人口迁移和流动

农民工问题研究	807
中国城乡人口迁移流动的规律和最新趋势	834
人口迁移、户籍制度改革与城市化发展	861
当前我国流动人口面临的主要问题和对策	872
中国境内外籍人口分析报告	902
我国省际人口迁移及其动力机制与经济效应	926
中国人口迁移流动的新特点及对城镇化发展影响的研究	946
流动人口的现状与变迁研究	965
人口迁移流动条件下的户籍管理制度的改革	986
省际人口流动对地区公共财政负担影响与中央转移支付调节方案研究	1004
跨省迁移与中国城市体系优化	1022

七、城镇化

中国人口城镇化:动态、结构与趋势	1037
中国城镇化进程中的人口与环境关系研究	1059
基于 GIS 与人口普查数据的中国未来城镇化情景分析研究	1084
中国城市化态势下的小城镇人口集聚能效分析	1098
中国城镇化的区域差异及其对消费需求的影响	1116

八、人口素质

中国人口素质和人力资本研究报告	1135
中国人口健康状况分析及健康促进研究	1192
基于受教育年限的全国及各地区人力资本测算	1215
我国人口受教育状况的趋势和队列分析	1236
中国人口文化素质特点与趋势分析(2000—2010 年)	1252
我国人口素质及其提高路径研究	1271
农村留守人口人力资本状况评价及投资路径研究	1281

九、民族人口

中国少数民族人口的规模、结构、分布、再生产类型的现状与变动分析研究	1297
我国各民族人口科学文化素质、就业水平的现状及进步程度	1326
民族人口现代化进程的族际比较研究	1350
中国少数民族人口分布的地域性研究	1372

下 册

十、人口居住状况

中国家庭户住房需求分析与预测研究	1405
中国人口、家庭与住房状况研究	1448
人口转变、人口流动与城乡居民住房状况	1475
城镇家庭住房:现状、特点与差异	1493
流动人口的流迁模式与住房保障研究	1514
城市人口居住状况、需求特征与住房保障政策关键参数研究	1534
人口、家庭与住房占有关系	1568
中国城市家庭住房状况的时空变动及社会分化研究	1588

十一、就业和失业

中国就业与失业问题研究	1603
中国的就业、失业与劳动参与的主要特征、突出问题与对策研究	1657
中国职业阶层结构的变迁 1990—2010	1695
我国劳动力市场中的就业与失业:现状、问题、趋势及对策研究	1712
我国职业结构变迁研究	1725
中国城市群中农村劳动人口特征比较与就业转移对策研究	1747
中国劳动力素质与产业升级	1764
沿海地区城镇人口蓝领化趋势变化分析与对策研究	1783
劳动力市场分割、搜寻匹配低效与劳动力市场结构性矛盾	1805

十二、人口地区分布

人口地区分布与空间布局优化研究	1831
中国人口分布新形势及其社会经济效应研究	1844
人口迁移流动条件下的城市规模和结构发展趋势研究	1860
区域人口发展模式的比较研究	1881

十三、人口和经济发展

人口转变、人口红利与中国的经济增长	1913
-------------------------	------



农村劳动力转移、粮食安全与区域经济发展	1930
我国区域人力资本与经济协调发展问题研究	1950
我国人口与经济发展研究报告	1967
中国人力资源现状与发展战略研究	1996
劳动力供需缺口现状与特征分析	2015
中国经济是否超越刘易斯拐点	2037
经济发展方式转变下的人力资本需求研究	2046

十四、人口、资源和环境可持续发展

我国人口空间分布、合理人口容量与区域经济资源环境协调发展研究	2063
西部地区的人口迁移与环境变化互动及作用机制研究	2103
我国人口发展与植被覆盖变化关系研究	2124
人口发展对我国碳排放影响的统计研究	2146

一、人口生育水平

中国生育水平研究

课题承担单位:中国人民大学人口与发展研究中心

课题负责人:翟振武

课题组成员:杨凡 盛亦男 张现苓 靳永爱 张浣琨

赵梦晗 伍海诚

【摘要】在中国各种不同来源的统计数据中,许多都有直接或间接与历年出生人口数量相关的信息。但是,这些统计数据中所反映的出生人口数量或者生育水平并不一致。由于生育水平指标的基础性和重要性,如果对当前生育水平无法做出比较准确的判断,就无法进行正确的重大决策,甚至一旦决策失误会对中国人口发展的方向和态势产生灾难性的影响。所以,弄清中国的生育水平,是具有重大实践意义和决策意义的选题,它将会为政府部门准确判断人口形势,科学进行人口决策提供坚实的基础和有效的数据支撑。

本研究对第六次人口普查的数据进行评估,通过对不同来源的数据(包括普查、抽样调查、教育数据、户籍登记数据等)进行比较、分析,寻找各种统计数据之间的差异性和规律性,研究各种数据所存在的问题、影响和纠正方法,对1990年以来中国出生人口数量进行了研究,还原1990年以来历年中国人口的真实生育水平。

首先,本研究对普查数据进行了质量评估。无论是1990年人口普查,还是2000年人口普查,在低年龄组上都存在一定程度的漏报,而且漏报程度不低,远高于国家统计局所公布的漏登率。但是,通过对六普数据和五普数据的比对,只能够对五普数据的漏报情况进行估计,并不能对六普数据的漏报情况进行估计。所以,必须借助独立于普查数据的其它来源的数据进行分析。而对其它来源数据的评估表明,教育统计数据的质量较高。它不但在数据收集过程中受到人为干扰较少,数据本身也具备很高的成熟性和稳定性。公安数据的质量在近些年来有了很大程度的提高,但是公安数据的低年龄组还存在一定程度的漏报。

然后,本研究利用独立于普查数据的其它来源的数据来推算了1990年以来,特别是2000年以来的生育水平,主要思路就是把公安数据中出生人口数量的发展趋势和教育数据中出生人口数量的水平相结合。根据“是否直接推算总和生育率”这个标准,推算方法可以分为直接推算法和间接推算法。在直接推算法中,本研究根据2011公安数据推算的1990—2010年出生人数计算了每年的总和生育率,然后用历年教育数据推算的1990—2000年出生人数计算了相应年份的总和生育率。然后拟合这两套数据的线性关系,并估算了教育数据2000年以后的总和生育率,以此作为对总和生育率的一种估计。在间接推算法中,本研究利用最小二乘回归的方法对公安数据和教育数据所推算出的出生人口数量之间的关系进行线性拟合,并根据回归公式进行外推,对2000—2010年的出生人口数进行估算。然后根据各年龄组妇女人数及生育的年龄分布模式数据估算总和生育率。两种方法得到的结果都表明,我国2000年以来的生育水平在1.6—1.7左右。

本课题的研究有以下几点发现:一是人口普查数据的低年龄组确实存在比较严重的漏报现象,这对我国真实生育水平的估计产生了很大的干扰。二是这种漏报现象是可以进行估计和纠正的。可以利用出生同批人的规律,用下一次普查对应的较高年龄组数据进行检验和纠正,也可以利用独立于统计数据之外的其他来源数据辅助判断。三是通过对多种来源数据的评估、分析和比对,本研究认为1990—2000年间中国的生育水平呈现出持续下降的趋势,大约从更替水平下降到1.7左右,而2000年—2010年间生育水平保持在1.6—1.7左右,呈现出波动中缓慢下降的趋势。

【关键词】人口普查;数据质量评估;总和生育率;出生人数



一、研究背景和意义

对中国人口生育水平的准确判断,始终是国家进行重大决策和长远战略性规划的基础性依据,也是政府各个部门制定相关政策的重要参考。但是,由于计划生育指标是一项对各级领导考核具有“否决”意义的指标,同时中国又有一项社会抚养费征收制度。这些政策导致出生的孩子数量与政府官员的政绩以及老百姓的切身利益都存在着密切关系。所以,无论是计划生育部门的报表统计,还是人口普查以及人口抽样调查,甚至其它有关人口的各种调查,只要涉及出生婴儿,数据质量就会由于利益相关方的种种原因而大打折扣,而且基本上是漏报、瞒报。由于婴儿的漏报、瞒报,直接根据婴儿数量和育龄妇女数量所计算出来总和生育率水平明显存在偏差,在此基础上进行的我国人口形势的判断也必然会产生偏差。

目前,在中国各种不同来源的统计数据中都有直接或间接与历年出生人口数量相关的信息。但是,这些统计数据中所反映的出生人口数量或者生育水平存在许多不一致的地方,以致莫衷一是、无法使用。例如,2000年全国人口普查直接调查的总和生育率仅为1.22,2005年小普查的总和生育率上升到1.33,2010年普查的直接调查结果显示总和生育率为1.18。但是,教育部统计的在校小学生数据却表明,2007年7—9岁小学生人数比2000年0—2岁的小学生人数多出900多万人。而国家统计局连续多年的年度的人口变动抽样调查又显示出,同一个队列的人数(5岁以下低年龄组)随着年龄越来越大,人数也越来越多。显然,各种有关人口的调查所显示的生育率水平高低不一,根本令人摸不着规律。“中国目前生育水平究竟有多高?”甚至已经成为各级政府部门和全社会关注和的焦点和激烈争论的热点之一。

由于生育水平指标的极其重要性,如果对当前生育水平无法做出比较准确的判断,就无法进行正确的重大决策。更严重的是,如果对生育水平判断失误(无论是过高还是过低),会导致决策的重大失误,这种失误会对中国人口发展的方向和态势产生灾难性的影响。

因此,利用科学方法,对各种来源(包括普查、抽样调查、教育数据、户籍登记数据、计划生育数据、计划免疫数据等)的调查资料和数据进行深入的分析、比对,全面系统地对1990年以来中国人口的生育水平进行准确的评估和判断,力求解开当前的“生育率之谜”,认清中国的人口形势,是具有重大实践意义和决策意义的选题。研究结果将会为政府部门准确判断人口形势,科学进行人口决策提供坚实的基础和有效的数据支撑。

二、研究目的和内容

本课题的研究目的是,通过对第六次人口普查的数据进行评估和对不同来源的数据进行比较、分析,寻找各种统计数据之间的差异性和规律性,研究各种数据所存在的问题、影响和纠正方法,对1990年以来中国出生人口数量进行研究,还原1990年以来历年中国人口的真实生育水平。

本课题研究的主要内容可以分为以下三个方面:

第一,通过对1990年第四次全国人口普查、2000年第五次全国人口普查和2010年第六次全国人口普查三次普查数据的分析和相互比对,对这三次普查数据的质量特别是出生人口数量数据的质量进行评估。

第二,通过实地调研和数据梳理分析,评估与人口有关的抽样调查数据、小学生教育数据、户籍登记数据等的出生人口数量或者分年龄人口数报告质量,以及误差的方向、程度、原因、影响等。

第三,深入分析这些数据反映的1990年以来生育水平的变化情况。对上述各套数据进行回推或时间调整,在共同的时间点上进行比对,在对各种数据比较和评估的基础上,对1990年以来的真正生育水平(包括出生人数和总和生育率)进行科学的判断和论证。在评估和分析的基础上,全面重构自1990年以来,特别是2000年以来中国生育水平的变化轨迹。

三、已有研究综述

中国的真实生育水平一直是学术界争论的焦

点,目前对于生育水平的研究非常丰富,不仅估计方法多样、参考数据多样,而且在结论方面也出现了百家争鸣的局面。

由于普查数据,尤其是低龄组普查数据存在着严重的漏报问题,学者们采用了多种方法估计普查漏报、推算生育水平。已有研究多集中在利用2000年普查数据估计1990—2000年的总和生育率。大多数研究认为普查数据中的出生人数和低龄组人数存在大量的漏报(崔红艳、张为民,2002;崔红艳、张为民,2004;于学军,2002;王金营,2003;于学军、王广州,2004;Robert D. Retherford等,2004;陈卫、翟振武,2007),中国的实际生育水平应该上调。

估计生育水平最常用的方法是利用普查数据进行生命表回推,推算出生人数,从而估算生育水平。如于学军、王广州(2004)以三次普查数据为基础,对现存人口进行队列存活分析,发现调整后2000年0—9岁各年龄组人口比调整前的同批人多出3005万人;王金营(2003)通过年龄移算法回推1990—2000年的出生人口数,估计1990年总和生育率为2.46,到1992年降到更替水平以下,1994—2000年之间基本稳定在1.72—1.76。王谦(2008)则根据1955—1959年、1960—1964年、1965—1969年三个队列累计生育率的持续下降,推断出过去的一、二十年间,总和生育率的总趋势也是持续下降的。而2005年队列累计生育率降到1.67可以表明,在2000年以前中国的总和生育率就已经降到了1.67左右。

第二种比较常用的方法是借用其他来源的数据(如教育数据、公安数据)对人口年龄性别比结构进行修正,回推出生人数,估计生育水平。张为民、崔红艳(2003)利用全国小学生入学人数反推历年出生人数的方法为生育研究提供了一个很有价值的新思路。翟振武、陈卫(2007)利用中国一直独立存在的严密而周全的教育统计数据估计2000年人口普查中的漏报,指出我国生育率于1991年达到更替水平,之后进一步下降,低于更替水平,90年代后期我国的总和生育率在1.73左右。考虑到教育数据中在校小学生人数存在的多报问题,蔡泳(2009)对实际在校和公布在校学生的比例进行调整,计算了90年代生育水平的上限:我国的生育水平在1992年下降到更替水平之下以后,其下降的趋势非常明显,在1998年下降到了TFR为1.58的水平。周长洪等人

以2006年公安部户籍管理中的农业和非农业人口比例、少数民族分布状况及计划生育管理实践中的有关数据,推算出中国计划生育政策下预期终身生育率上限为1.45(周长洪、潘金洪,2010)。还有学者利用普查数据进行母子匹配和生育史重构来研究中国的生育水平。如Robert D. Retherford等(2004)运用“亲生子女法(OWCH)”和“生育史重构法(BHR)”对生育水平进行重构,结果表明由年龄别生育率计算的总和生育率按亲生子女法估算值应该上调约16%,从而使2000年的1.36的原始估算上调到1.58,1996—2000年的总和生育率需要上调10%,使1.35的原始估算值上调至1.48。郭志刚(2004)根据2000年人口普查数据对0—10岁少儿人口进行户内母子匹配,按这种方法重建的1990—2000年各年份总和生育率分别为:2.73, 1.80, 1.68, 1.57, 1.47, 1.48, 1.36, 1.31, 1.31, 1.23, 1.23。估算出的2000年总和生育率与根据五普数据样本中育龄妇女本人申报的1999年11月至2000年10月期间的生育情况所计算出来的总和生育率1.229非常接近。另外,郭志刚(2000)引进去进度效应总和生育率研究中国的生育水平,认为中国的时期生育率比较低,虽然存在一定的漏报问题,但不是主要因素,时期生育率有其内在缺陷,会低估终身生育水平。他利用1997年生殖健康调查数据计算的总和生育率为1.3,而去进度效应总和生育率却为1.7。丁俊峰(2003)分析了中国90年代生育模式的变化,通过计算1991—2000年中国去进度效应总和生育率,分析认为初婚年龄和拉大一孩至二孩生育间隔造成的生育模式变化是中国这一时期总和生育率偏低的重要原因。他也认为,去进度效应总和生育率更接近中国妇女的实际终身生育水平。

这些研究虽然利用了不同来源的数据、不同的方法,但总的结论大致可以分为两类:

一是根据普查上报的前一年出生人数计算的总和生育率低估了中国的实际生育水平,考虑到大量的出生漏报和低龄组漏报,中国的总和生育率应该高于普查公布数(于学军,2002;于学军、王广州,2004;Robert D. Retherford等,2004),90年代的生育水平在1.7左右(翟振武、陈卫,2007;王金营,2003)。翟振武(2005)利用教育统计数据对人口普查0—9岁人口进行调整,并推算出1991—2000年



历年的出生人数,同时利用2000年人口普查的生育年龄模式和15—58岁妇女人数推算出1991—2000年历年育龄妇女分年龄人数和分年龄出生人数,并估算出历年分年龄生育率和总和生育率。研究发现,我国1991年生育率达到更替水平,并在之后继续下降,20世纪90年代末综合生育率水平估计定位在1.73左右。

另外一种观点则认为即便考虑出生人口的漏报和瞒报,中国的生育水平实际上已经低到了一定水平(郭志刚,2004;马小红、孙超,2011)。如郭志刚认为中国的总和生育率已经降到极低生育水平了,通过母子匹配法估算的2000年中国生育水平为1.23;同时还引进去进度效应总和生育率证明中国的时期生育率较低主要是因为生育年龄推迟的作用。张广宇等(2004)认为没有足够证据证明90年代中国存在大规模和持续出生漏报,相反,有理由认为各次全国人口调查中的低生育率是实际生育水平大大降低的反映,出生漏报水平可能被高估。丁峻峰(2003)根据2000年人口普查得到的全国0—9岁分年龄人口数,利用生育基数法推算的1999年和2000年TFR分别为1.11和1.35。

崔红艳,张为民(2002)利用人口分析技术和其他调查及统计数据对比的方法,对2000年人口普查的人口总量进行了评价,估计了人口统计漏登的水平。他们利用了户籍管理数据、小学入学教育数据,以及第四次人口普查资料对比,分析了五普人口漏登情况。他们认为,20岁以下,特别是0—9岁儿童漏报是第五次人口普查漏登人口的主要因素,且估计漏登人口在2000万以上。

受数据限制,对2000年以来的生育水平实证研究比较少,多为经验性的判断,目前有两类比较典型的观点。一是认为目前的生育水平在1.7左右,如美国人口咨询局认为,2010年我国妇女总和生育率的估计值为1.5,联合国人口司对2005—2010年间196个国家或地区生育水平的估计,中国为1.77。其他无论是用国家统计局公布的出生人口数还是用教育部公布的历年小学入学人数回推六七年前的出生人口数,大致可以估算出2000年的总和生育率为1.7左右(于学军,2011)。于学军(2002)分析了第五次人口普查数据中总人口和年龄结构之间出现不吻合现象。并根据1996—2000年各年份出生的人

口数量和近年来的相关抽样调查结果,以及有关部门的抽查和清理信息,认为目前中国妇女的总和生育率大致在1.6—1.8之间,低于政策要求的生育水平。他对2000年以来人口生育水平进行分析,认为2005年总和生育率在1.7左右(于学军,2005)。对比世界人口生育水平变化趋势,于学军认为,中国人口生育水平的下降是正常的,只是速度略快于世界人口生育水平的变化趋势。马瀛通(2005)认为,生育率水平低于或大大低于1.8的任何数据都难以信服,在21世纪初最接近实际的总和生育率估计约为2.0左右。

另一类则强调目前中国的生育率已经在较低的水平上了。有学者认为2006年人口和计划生育调查的生育率严重偏高,调整后2005年总和生育率为1.31,而同年的一孩总和生育率仅为0.86(郭志刚,2009)。郭志刚(2000)利用1997年中国人口与生殖健康抽样调查的生育率统计结果与1992年中国生育率抽样调查的统计结果进行了一致性比较检验,应用三个不同生育指标对90年代生育水平及其变化原因进行了分析。根据1997年调查结果,我国生育水平于1991年达到更替水平以下。同时还表明,20世纪90年代的TFR并不是单调下降,而是在1994年达到最低点,然后便保持在这一水平,甚至还略有回升。

任强(2005)将不同机构、学者对2000年生育水平的估计进行了归纳,发现学者们得出2000年总和生育率在1.2—2.3之间,并没有达成共识。他认为现有的数据推断均难以信服,未来还需完善民事登记和生命统计制度,掌握目前的生育水平。

除此之外,还有学者对不同数据质量以及可能造成的影响进行了分析。如张广宇,原新(2004)回顾了对90年代出生漏报和生育水平估计的研究,认为在育龄群众的生育观念转变和生育水平已经下降到一定水平时,瞒报和出生漏报不应成为解释生育率的持续下降的唯一原因。他们分析了几种人口统计数据的出生漏报情况。第一种为公安户口统计数据。公安户口统计从20世纪80年代初以来数据质量有所下降,但大部分超生的夫妇也想方设法的给孩子上户口。因此,他们认为户口统计中漏报的主要为超生儿童的看法比较片面。第二种为计划生育统计数据质量分析。他们认为计生统计中不准确的

一个重要原因是由于婴儿的母亲没有在计生统计的范围内。随着流动人口规模的增加,流入地与流出地登记责任不明确,加大了对妇女统计的可能性。第三种情况是抽样调查数据的偏差来源于抽样框的不完整。第四种情况是流动人口规模的增多,加大了人口普查中出生漏报出现的可能。郭志刚(2010)根据2005年全国1%人口抽样调查的样本数据,分析了流动人口对当前生育水平的影响。通过计算和比较流动人口和非流动人口在生育率水平与孩次年龄模式方面以及户籍性质方面的差别,表明人口流动显著地降低了农业户籍人口的生育水平,在全国层面显著地降低了生育率。曾毅(1995)认为我国1991—1992年生育率在2.1左右,未低于替代水平。与1989—1990年相比,我国1991—1992年真实生育水平大幅度下降的原因是已婚生育率的下降以及初婚年龄的上升。

已有研究对评估数据质量、估算生育水平提供了大量有价值的方法和具有启发性的结论,为后续研究打下了重要基础。而第六次人口普查又为估计和评价2005年以来的生育水平提供了最全面和最详尽数据基础。由于这次普查的高质量,利用这次人口普查数据,对近些年的生育水平进行比较准确的评估和确定,具有很高的可靠性和坚实的基础。此外,我们还得到了教育数据、公安数据等多途径的、可信度较高的数据。方法的成熟和数据的可靠性为本研究估计生育水平奠定了坚实的基础。

四、对普查数据的评估

数据是研究各种社会现象的重要媒介,通过各种相关调查数据的分析,可以将抽象的社会事件与现象具体化,从中归纳演绎现象背后的动因与规律。在利用各种数据进行统计分析时,需要注意两方面的问题:一是统计分析方法是否适用,二是数据是否真实可靠。

在进行人口统计分析时,同样需要对数据的信度与效度进行评估。真实的人口数据对于判断我国未来的人口形势具有重要的意义,是探寻人口规律、制定人口政策的有力证据。我国关于人口的专门性调查非常多,包括十年一次的全国人口普查、两次普查之间的全国1%人口抽样调查、全国1‰生育抽样调查

等,通过这些普查或抽样性调查,可以获得大量的人口数据,对于认识我国人口问题十分有益。但是,历次人口普查数据之间的一致性检验、普查数据的内部核查以及与其他调查数据的对比分析发现,普查数据往往存在一些问题。在对普查数据进行使用和分析之前,数据的准确性检验成为不可迈过的门槛。

回顾以往研究,学者们利用不同的方法对我国历次人口普查数据做出了评价与结构调整。第一次人口普查和第二次人口普查登记项目较少,数据信息量较小,而且时隔久远,对分析我国现在的人口问题与现状作用相对较小,但是数据质量相对较高,因而针对这两次人口普查进行的评估文献也较少。学者们更为关注距今较近的第三次人口普查、第四次人口普查以及第五次人口普查的评价与调整。不同学者利用不同的分析方法对普查数据质量进行细致的评估,得出相对一致的观点:三普和四普数据整体质量较五普数据要好,但是四普数据在低龄组人口仍存在一定的漏报现象;五普数据存在较多问题,主要表现在0—9岁低龄组漏报,青壮年人口漏报与重报并存。

乔晓春(1992)利用存活分析法分析了三普和四普相同队列人口的年龄结构和性别结构的一致性,发现两次普查数据在性别年龄结构上质量较好,但是在以“0”和“5”结尾的年龄上存在一定的堆积现象。涂平(1992)利用四普数据公布的0岁死亡人口及0岁尚存人口,估计预期出生人数,并与普查公布的出生人数相比,发现四普数据中出生人口存在较大漏报。周皓(2003)基于五普数据对四普数据质量进行了再检验,他认为四普数据的漏报率远高于国家统计局公布的水平,应该在1.866%左右,其中低龄组人口漏报在50%以上。李树苗、姜全保等(2006)分析发现五普数据0—9岁低龄组人口漏报了3743万人,10—60岁年龄组人口重报与漏报等现象共同存在,相互抵消后表现出重报1134万人,并计算出五普漏报率是2.05%。王广州(2004)在四普数据质量很高的前提假设下,得出结论,五普数据0—14岁少儿人口至少漏报3300万人左右,15—40岁年龄组至少重报1100万人。张为民、崔红艳(2003)发现四普数据中0—9岁人口存在一定的漏报,五普数据中青壮年人口存在轻微的重报,46岁及以上人口吻合度较高,数据质量良好。

同时,由于普查数据在不同年龄段上存在不同的问题,对相关人口指标的估计也产生明显的影响,如出生性别比、生育水平、死亡水平等。翟振武、陈卫(2007)和张为民等人分别利用不同来源的数据估算了1990年以来的出生性别比,认为我国出生性别比偏高是客观存在的,但存在数据漏报导致的虚高。于学军(2002)利用历年抽样调查数据及四普数据、五普数据对比分析,发现五普数据的人口结构与总量之间存在不吻合现象,在低龄组人口上存在漏报,在对数据调整后他估计我国2000年生育水平在1.6—1.8之间,高于政策生育率。翟振武、陶涛(2010)利用多年教育数据分析,发现五普数据漏报了大量人口,我国生育水平、未来劳动力数量及育龄人群规模被低估,而出生性别比、老龄化等指标被高估。黄荣清对四普数据和五普数据进行再分析,认为普查数据存在死亡漏报,经过对普查数据的调整,他估算出20世纪90年代我国人口的预期寿命要比普查资料公布的结果低1.5岁左右。

2010年我国进行了全国第六次人口普查,为进一步分析我国人口形势和人口变化提供了丰富的数据资源。郭志刚(2010)利用六普数据估算了我国1990—2010年的各项人口指标,认为这段时期我国的生育水平、人口规模被高估,相应地老龄化等指标被低估。

以往针对历次人口普查数据的分析清晰地指出,普查数据存在各种误差,对于拥有庞大人口规模、实施特殊生育政策的我国而言,普查资料的准确性检验是使用普查数据进行分析前不可迈过的重要步骤。

尤其是在对生育水平、出生性别比等直接受到低龄组人口规模和结构影响的指标进行研究分析时,更需要对相关数据的质量进行检验及必要的调整。

(一)六普数据性别年龄结构的一致性检验

2010年第六次人口普查工作的完成,为进一步研究我国21世纪以来的人口新变化提供了丰富的数据信息。但在使用六普数据进行相关分析之前,需要先对六普数据的质量进行检验,利用历次人口普查数据来分析六普数据的结构一致性问题。

首先分析六普数据性别结构的一致性。分别计算第五次人口普查0—90岁及第六次人口普查10—100岁(相同出生队列人口)的分年龄性别比,并对比两套数据的计算结果。从图4-1中的两条曲线可以看出,2000年0—19岁人口的性别比明显高于2010年10—29岁人口的性别比。导致这种现象有两个可能的原因,一是五普数据中0—19岁人口存在漏报,且女性漏报更为严重;二是六普数据10—29岁人口存在重报,且是男性重报严重。但是,以往的研究证明,前一种解释的可能性更大,即2000年0—19岁人口存在漏报,且漏报了较多的女性人口。在其余的年龄组上,五普数据与六普数据相同队列人口的分年龄性别非常吻合,而且波动趋势也非常一致,说明成年人口普查数据质量较高。老年人口性别比出现差异,主要是由于50岁之后,年龄别死亡率开始大幅度上升,且男女死亡率性别差异逐渐拉大,男性死亡更多,导致老年人口中分年龄性别比下降。

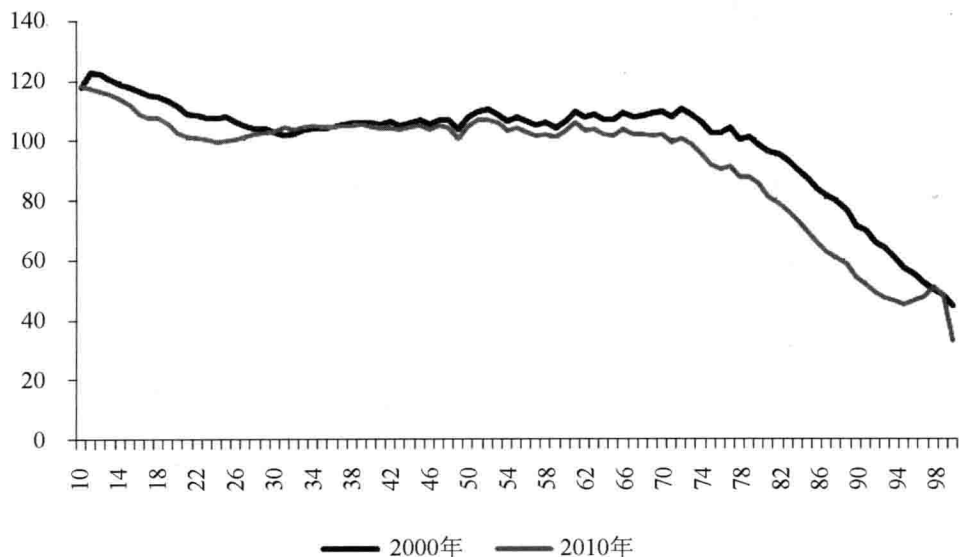


图4-1 五普与六普相同队列人口的分年龄性别比