

转型期的中国人口

China's population during Transition time

2000年人口普查国家级重点课题

国务院人口普查办公室 编
国家统计局人口和社会科技统计司

 中国统计出版社
China Statistics Press

转型期的中国人口

2000 年人口普查国家级重点课题

国务院人口普查办公室 编
国家统计局人口和社会科技统计司

 中国统计出版社
China Statistics Press

(京)新登字 041 号

图书在版编目 (CIP) 数据

转型期的中国人口:2000 年人口普查国家级重点课题
/ 国务院人口普查办公室, 国家统计局人口和社会科技统计司编.
- 北京: 中国统计出版社, 2005. 6
ISBN 7-5037-4597-5

I. 转…
II. ①国… ②国…
III. 人口普查-中国-2000-文集
IV. C924.25-53

中国版本图书馆 CIP 数据核字 (2005) 第 067180 号

转型期的中国人口

作 者 / 国务院人口普查办公室
/ 国家统计局人口和社会科技司
责任编辑 / 刘勇利 吕 军
E - m a i l / cbsebs@stats.gov.cn
封面设计 / 艺编广告 · 张 冰
出版发行 / 中国统计出版社
通信地址 / 北京市西城区月坛南街 75 号
邮政编码 / 100826
办公地址 / 北京市丰台区西三环南路甲 6 号
电 话 / (010) 63459084 63266600-22500 (发行部)
印 刷 / 北京顺义振华印刷厂印刷
经 销 / 新华书店
开 本 / 880 × 1230 毫米 1/16
字 数 / 1900 千字
印 张 / 63.5
版 别 / 2005 年 8 月第 1 版
版 次 / 2005 年 8 月北京第 1 次印刷
书 号 / ISBN 7-5037-4597-5/C · 2079
定 价 / 450.00 元

版权所有。未经许可, 本书的任何部分不准以任何方式在
世界任何地区以任何文字翻印、拷贝、仿制或转载。
中国统计版图书, 如有印装错误, 本社发行部负责调换。

前言

2000年第五次全国人口普查是世纪之交的人口普查，也是历史上规模最大的人口普查。这次普查收集的各种资料和数据是改革开放以来我国经济和社会发展在人口方面的具体体现，反映了我国转型时期的基本国情。

二十多年来，中国经历了经济体制从计划经济到市场经济、人口生育从高水平到低水平、人口年龄结构从成年型到老年型、人们从安土重迁到大规模流动、城市化从停滞不前到快速发展等重大变化。因此，充分开发利用好2000年人口普查资料，不但可以对人口与经济社会、人口与资源环境进行战略性研究，为各级政府科学决策提供依据，而且还将最大限度地发挥人口普查的社会效益，提高人们的人口意识、人均意识和可持续发展意识，真正实现规划未来、造福人民的目的。

为此国务院人口普查领导小组下发了《关于进一步加强人口普查资料开发利用工作的通知》（国人组字[2001]1号）。国务院人口普查办公室设立了《第五次全国人口普查资料研究国家级重点课题》，面向全国各研究部门、大专院校、人口研究机构及相关单位公开招标，旨在充分调动我国的科研力量，最大限度地开发利用人口普查资料。

第五次全国人口普查资料研究国家级重点课题的招标题目，由国务院人口普查办公室和资料开发利用专家小组根据我国现阶段人口、经济、社会发展的需要，着重从可持续发展的角度出发，重点选择了当前各级政府和社会关注的、有重要影响的人口及相关问题。

资料开发利用专家小组对100多份有效课题申请报告进行了认真审

前言

查。经过反复讨论，通过无记名投票，确定了42个课题予以立项。

展现在读者面前的这本论文集，汇集了我国人口、经济、社会各界专家对有关人口问题的最新研究成果。这些成果围绕人口结构、人口素质与人力资本、人口流动与城市化、家庭、婚姻、生育、死亡、老龄化与社会保障、人口可持续发展等当前各级政府和社会关注的人口、社会及相关问题，进行了深入的研究与探讨，具有较高的学术价值和实践指导意义。

由于篇幅的限制，本书只摘编了42项研究成果的精华部分，研究成果的全部内容，将制作成光盘出版。

全书在编辑过程中得到了社会各界、尤其是课题承担单位的的大力支持，在此我们表示衷心地感谢。编辑过程中不足之处在所难免，恳请读者批评指正。

编者

2005年2月

目 录

1. 生育

- 中国90年代以来生育水平研究 中国人口信息研究中心 (3)
- 转变时期生育率变化特征研究 航天部710所 (22)
- 2000年人口普查按生育政策类型的人口分析 北京大学社会学系 (37)

2. 死亡

- 中国人口的死亡和健康 北京大学人口研究所 (63)
- 90年代以来及今后中国人口死亡率和生育率演化过程研究
..... 中科院数学与系统科学研究院 (86)
- 中国2000年第五次人口普查死亡研究报告 西安交通大学人口所 (94)

3. 性别

- 出生性别比研究 北医公共卫生学院妇幼卫生系 (159)

4. 老龄化

- 中国人口老龄化国际比较研究 中国人民大学人口研究所 (177)
- 人口老龄化与养老保障研究 辽宁大学人口研究所 (191)
- 人口老龄化研究 中国老龄科学研究中心 (211)

5. 家庭与住房

- 家庭结构与住房 天津社会科学院社会学所 (235)
- 中国人口和家庭户变动与住房状况研究 北大人口研究所 (268)

6. 人口流动

- 城市外来人口的现状和管理对策 清华人文社会学院社会学系 (309)
- 中国人口迁移流动与人口分布研究 中国人民大学人口所 (325)
- 城市流动人口的现状与管理方法研究 上海社科院人口与发展研究所 (337)
- 外来人口聚集地户籍制度改革研究 中山大学人口研究所 (360)
- 90年代中国人口分布变动和迁移流动的新形式研究 华东师大人口研究所 (376)
- 中国省际人口迁移与区域经济发展互动关系研究 复旦人口研究所 (392)
- 人口迁移、劳动力流动与人力资本转移研究 中南财经政法大学 (413)
- 农村人口迁入大中城市户籍制度创新研究 河北省人口研究中心 (428)

中国省际人口迁移流动态势分析 广东省社科院社会学与人口学研究所 (444)

7. 城镇化

中国城镇化水平的修补和城市规模分布研究 北大城市与环境学系 (467)
流动人口对三大都市圈发展影响研究 社科院城市发展与环境研究中心 (493)
东北地区人口城市化与非农产业集聚研究 东北师范大学 (513)
人口迁移流动对城市化进程的影响 中南民族大学 (532)
中国流动人口的行为特征与城市化 ... 中国人民大学区域经济与城市管理研究所 (550)

8. 教育

21 世纪中国人口变动与教育现代化目标预测论证和规划建议
..... 浙江大学可持续发展研究中心 (571)
中国地区国民素质国际竞争力研究 中国人大应用统计科学研究中心 (612)
中国人口文化素质分析报告 教育部教育发展研究中心 (629)
欠发达地区农村劳动力素质的现状及其成因与对策 湖北省黄冈师范学院 (650)

9. 就业

城镇就业与失业问题研究 社科院人口与劳动研究所 (661)
中国劳动力人口职业、行业结构变迁与经济发展区域比较研究
..... 南开大学人口与发展研究所 (723)
中国人口迁移与收入差距 北大经济研究中心 (747)
劳动参与率、劳动生产率与经济增长 东北财经大学 (773)
转型期我国行业变迁与人力资本配置 国家发展与改革委员会经济研究所 (794)
全球化条件下中国就业政策研究 劳动部国际劳工与信息研究所 (813)
1990—2000年我国人口流动、劳动就业及其相关性研究 四川大学人口研究所 (829)

10. 民族

20世纪90 年代中国各民族人口变化特点研究 首都经贸大学 (869)

11. 可持续发展

中国的职业结构变化与中国社会阶层结构演变 中国社会科学院社会学所 (887)
我国未来人口分布和城市化趋势的区域研究 复旦大学人口所 (912)
西部大开发中的人口问题 西安财经学院 (950)
21世纪初期 (2001-2020年) 中国人口与资源、环境综合协调及其经济可持续发展
系统模式识别与评估 厦门大学系统工程研究所人力资源环境系统研究室 (968)

附录

附件一：第五次全国人口普查资料研究课题招标公告 (990)
附件二：第五次全国人口普查资料研究课题招标管理办法 (1000)
附件三：课题指南 (1003)
附件四：第五次人口普查国家级重点课题立项分类标准 (1005)
附件五：第五次全国人口普查国家级重点课题项目立项及数据使用协议书 (1006)
附件六：第五次全国人口普查国家级研究课题工作总结 (1007)

1

生 育

中国 90 年代以来生育水平研究^①

本文分为五个部分：对 90 年代以来生育水平研究的综述和评价、对第五次人口普查公布的人口总数的评估、对第五次人口普查数据重报问题的分

析、对第五次人口普查年龄结构数据的评估与调整、对总和生育率的估计。

一、对 90 年代以来生育水平研究的综述和评价

90 年代以来，各年的全国人口抽样变动调查都有人口出生率指标，我们可以从国家统计局历年出版的《中国统计年鉴》和每年的《国民经济和社会发展统计公报》中获得。从统计数据来看，90 年代以来，中国人口出生率的下降趋势非常显著，从 1990 年的 21.06‰，下降到 2000 年的 14.03‰，10

年间大致下降了 7 个千分点。相应地，人口自然增长率也从 1990 年的 14.39‰，下降到了 2000 年的 7.58‰。尤其值得一提的是，1998 年的人口自然增长率首次在非天灾人祸的条件下降低到了 10‰ 以下的低水平，取得了历史性的突破，实现了老一辈国家领导人的夙愿。



图 1 90 年代以来出生率、死亡率、自然增长率变化情况

① 除非特殊说明外，本报告所涉及的“年代”，都是指 20 世纪。

然而,人口出生率的下降只能从一个侧面粗略地说明 90 年代中国妇女的生育水平确有下降,但是由于受人口结构的影响,人口出生率不能准确地说明 90 年代以来各个年度生育水平变化的真实情况。90 年代以来,多次大型的全国抽样调查结果都显示,中国的妇女生育水平已经在冲破了 80 年代的徘徊局面,总和生育率已经降低到了更替水平

以下。但是,生育水平到底比更替水平低多少?直到今天也没有一个准确而权威的说法。由于各个抽样调查的目的、调查规模、调查对象各有不同,所得到的结果也不尽相同。90 年代以来进行的这些抽样调查所显示的总和生育率有一个共同的特点:调查的妇女总和生育率大大低于实际生育水平。

表 1 不同机构对 20 世纪 90 年代中国人口总和生育率的估计

调查或估计单位	1991	1992	1993	1994	1995	1996	1997	1998
中国人口信息研究中心	2.03	1.82	1.81	1.81	1.78	1.80	1.82	1.85
“38 万”抽样调查	1.65	1.52						
张二力等	2.00	1.84	1.83	1.81	1.78			
国家计生委	2.01	1.84	1.83	1.81	1.78	1.81	1.82	1.82
中国人民大学人口所	2.16	2.00	1.98	1.94	1.87	1.86	1.82	1.76
美国人口咨询局	2.2	1.9	2.0	1.9	1.8	1.8	1.8	1.8
亚太经社会	2.2	2.2	2.0	1.9	1.8	1.8	1.8	1.8

资料来源:中国人口信息研究中心,中国人民大学人口所的估计结果来自国家计划生育委员会政策法规司《未来人口发展与生育政策》研究课题报告。张二力等人的结果来自《全国和分地区人口预测》,中国人口出版社,1998 年。国家计生委的估计来自于《计划生育统计公报》(2001 年 1 号);PRB 的结果来自《世界人口数据表,1991—99》。亚太经社会的结果来自《亚太经社会人口数据表,1991—99》。

例如,第四次全国人口普查结束两年后,1992 年国家计划生育委员会进行了“38 万人全国生育抽样调查”。该调查公布的全国 1991 年和 1992 年总和生育率分别为 1.65 和 1.52,大大低于总和生育率为 2.1 的替代水平。1995 年由国家统计局组织的全国 1%人口抽样调查直接得到的总和生育率为 1.46,农村妇女的总和生育率只有 1.56,甚至比政策总和生育率还低 0.2。2000 年,人们终于迎来了企盼已久的第五次人口普查。但是,最终的结果比以往更令人难以接受:2000 年第五次全国人口普查直接统计的总和生育率只为 1.22。

面对上述调查和研究结果,如何全面、科学地分析我国 90 年代以来生育水平的变化过程并得到科学的、令人信服的研究结果,成为现阶段我国生育水研究过程中的重大问题。基于上述目的本研究主要利用 90 年代以来历年国家统计局公布的有

关数据,参照第三次人口普查、第四次人口普查有关数据,提出调整和修正第五次人口普查有关数据的建议,并在此基础上,研究 90 年代以来全国生育水平的变化。重点在于对 1990 年以来各年份总和生育率的估计,难点在于对人口年龄结构的调整。经过仔细研究和分析上述问题,我们的研究分以下几个步骤进行:

- 第一:评估第五次人口普查公布的人口总数。
- 第二:在肯定第五次人口普查公布的人口总数的基础上,估计重报和漏报人口数量,为理顺年龄结构作准备。
- 第三:在甄别重报和漏报问题的基础上,修正年龄结构。
- 第四:在修过的人口年龄结构的基础上,估计 90 年代以来的生育水平。

二、对第五次人口普查公布的人口总数的评估

我们注意到,2000 年第五次人口普查以后,国家统计局同样对 1990 年以来的数据进行了相应的调整。与 1990 年第四次人口普查时调查人口数少于历年公布人口数的情况相反,这次是普查公布数

多于历年公布的人口数,因此只是调整了 1998 年和 1999 年两年的人口总数,即把 1998 年总人口从 124810 万人下调到 124761 万人,减少 49 万人;把 1999 年总人口从 125909 万人下调到 125786 万

人,减少 123 万人,两年合计下调 172 万人,其他年份没有变化。与 1990 年第四次人口普查相比,调整的幅度是很小的。这是否意味着国家统计局对历年人口变动抽样调查结果和公布的结果比较有把握?从而对 2000 年人口普查的总人口数量比较满意?我们的研究也正是要回答这一问题。

2001 年 3 月公布的《2000 年第五次人口普查主要数据公报》第一号显示,祖国大陆 31 个省、自治区、直辖市和现役军人的人口为 126583 万人。漏登率为 1.81%。也就是说,实际普查登记只有 124337 万人,比公布人数少 2246 万人。公布的人口总数是实际普查到的人口与漏报人口之和。问题的焦点就在于在调查到的 124337 万人基础上,加入 2246 万人是多了,还少了,还是差不多?我们不妨从以下几个信息渠道考察一下。

1. 从国家统计局历年的人口变动抽样调查结果看人口总数

无论是历年公布的人口总数,还是 2000 年普

查公布的人口总数都是经过人为“处理”后的数据。如上所述,“五普”公布的人口总数加入了 2246 万人,历年公布的人口总数又是如何得到的?表 2 告诉了我们历年来各省市自治区抽样变动调查的结果与国家统计局经过调整后正式的公布结果的差异。从表中我们可以看到,即使国家统计局每年组织各省市自治区进行的人口变动抽样调查,所得到的出生人口数量的结果也是偏低的,比如 1993 年和 1994 年调查得到的出生率分别为 15.58‰和 15.32‰,实际公布的结果上调为 18.09‰和 17.70‰。仅 1993 年和 1994 年两个年份就将各省市自治区报表的出生人口数量向上调整了 583 万。1990 年后,国家统计局每年将人口出生率加 2—3 个千分点,即平均每年加 275 万人口,10 年累积“加水”2750 万。这一数字高于第五次人口普查漏报的 2246 万人。鉴于这种“冒水”现象,国家统计局在“五普”后,向下调整了过去公布的 1998 和 1999 年的人口总数。

表 2 国家统计局数据分析

年 份	国家公布 人口总数 (万人)	省市区报 表出生率 (‰)	国家公 布出生率 (‰)	省市区报 表出生 (万人)	省市区公 布出生数 (万人)	国家公布 出生数 (万人)	省市区报表出 生数和国家公 布出生数之差 (万人)
1990	114333	20.20	21.06	2308.86	2309	2407.16	98.30
1991	115823	18.32	19.68	2121.46	2121	2278.94	157.49
1992	117171	16.35	18.24	1916.22	1916	2137.73	221.51
1993	118517	15.58	18.09	1846.23	1932	2143.67	297.43
1994	119850	15.32	17.70	1835.34	1906	2120.46	285.12
1995	121121	14.42	17.12	1746.26	1833	2073.23	326.97
1996	122389	14.32	16.98	1752.77	1815	2078.35	325.58
1997	123626	13.47	16.57	1664.89	1763	2048.05	383.16
1998	124810	13.28	16.03	1657.34	1726	2000.54	343.20
1999	125909	12.70	15.23	1598.93	1662	1917.46	318.53

资料来源:中国人口信息研究中心,《人口与计划生育常用数据手册》。

从这个角度看,第五次人口普查公布的人口总数已经充分地考虑了各年各省市自治区上报的出生人口和人口总数的水分。国家统计局向上调整出生人口数的依据是抽查的漏报率,现在的问题是,各省市自治区是否在上报出生人口和人口总数时也作了向上的调整?倘若各省市自治区也如国家统计局一样,在汇总各地市的数据后,也作了相应的调整,那么,国家统计局所掌握的历年出生人口数和人口总数或许会高于实际的数字。这种担心不是没有道

理的,表 2 告诉我们,各省市自治区上报的出生人口数与其公布的数字并不一致,同样作了一定幅度的上调。从表 1 我们可以看到,1990—1993 年,各省市自治区的报表出生人口数与公布的出生人口数完全一致,此后公布的出生人口数均大于报表出生人口数。到 1999 年,累计多出 535 万人。

总之,我们认为,第五次人口普查公布的人口总数没有超出经过“加水”后的历年人口抽样变动的调查结果,或者说,第五次人口普查的人口漏报

率与国家统计局所掌握的历年各省市上报的人口漏报率大体相当。从这个角度说,第五次人口普查公布的人口总数是可信的。

2. 从国家计生委掌握的出生报表看人口总数

除了国家统计局外,掌握出生人口信息较全面的机构就是国家计生委。由于比较严格的计划生

育政策使得人们的生育意愿不能得到实现,于是超生后瞒报现象比较严重是众所周知的事实。问题在于,国家计生委统计的出生人口漏报率到底有多高? 以下几个信息或许会帮助我们回答这一问题,见表 3:

表 3 国家统计局和国家计生委出生人口数据比较分析

年份	SFPC 报表	SSB 公布	SFPC+30%	SFPC+35%	SFPC+40%
1992	1596	2137.73	2074.80	2154.60	2234.40
1993	1568	2143.67	2038.40	2116.80	2195.20
1994	1574	2120.46	2046.20	2124.90	2203.60
1995	1521	2073.23	1977.30	2053.35	2129.40
1996	1453	2078.35	1888.90	1961.55	2034.20
1997	1388	2048.05	1804.40	1873.80	1943.20
1998	1383	2000.54	1797.90	1867.05	1936.20
1999	1287	1917.46	1673.10	1737.45	1801.80
2000	1292	1772.00	1679.60	1744.20	1808.80
2001	1240	1702.00	1612.00	1674.00	1736.00

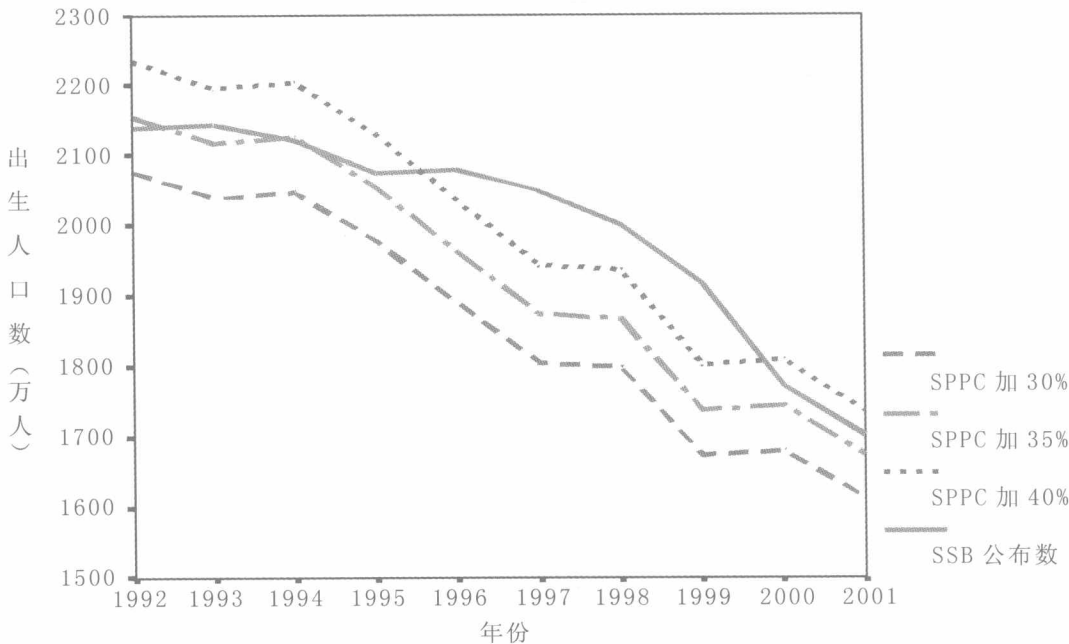


图 2 调整后国家计生委的数据与国家统计局公布的数据比较

第一:1992 年时,国家统计局报表和国家计划生育委员会出生报表的差距有 320 万左右;1999 年,这一差距为 310 万。1992—1999 年期间二者差距有些波动,平均每年相差大约 280 万人,二者之间的差距似乎很有规律。

第二:将国家计生委报表出生数上调 35%—

40%,与国家统计局的公布出生数非常接近。

第三:2000 年,国家计生委根据报表掌握的总人口数为 12.35 亿,当年出生人口为 1292 万人^①。根据报表出生人口数推算的总和生育率大约为 1.25,比第五次人口普查直接调查的总和生育率略高。

① 《中国计划生育年鉴》2001 年卷,第 448 页。

3. 从漏报和重报的对象看人口总数

2000 年人口普查结果中,最让人们心里没底的有三部分人口:低龄人口、流动人口和死亡人口。一般说来,0—10 岁的低龄人口漏报的可能性最大,而重报的可能性微乎其微。死亡人口的主要问题是漏报,而流动人口兼有重报和漏报的双重可能性。流动人口的漏报主要是受经济利益驱动的影响,用工单位故意少报外来人口或拒绝调查,主要缘于逃避劳动部门关于企业临时用工的政策,以及公安、工商、税务部门的管理。此外,由于对流动人口的各种管理收费太多,有些流动人口也不愿如实申报。武洁的研究发现,从“五普”事后质量抽查结果看,在漏报的总人口中,有近 80% 的人口是由于暂住人口漏报造成的(武洁,2002 年)。遗憾的是,文中没有进一步说明漏报的暂住人口年龄构成。流动人口的重报主要是由于离开户口所在地的流动人口既在流入地进行了登记,也在流出地进行了登记。由于信息的不对称,这部分重登人口数量无法估计。这或许是国家统计局没有公布 2000 年人口普查重报率的原因。1982 年第三次人口普查公报指出:“重报人口占 0.71‰;漏报人口占 0.56‰,重漏相抵,净差人口占 0.15‰”。1990 年第四次人口普查公报所公布的普查登记质量抽查结果是重登率为 0.1‰,漏登率为 0.7‰,重报和漏报相抵消后,净误差率为 0.6‰。但是,第五次人口普查只是简单说“普查登记结束后,全国统一抽取 602 个调查小区进行了登记质量的抽样调查。抽查结果,人口漏登率为 1.81‰。”。而对其他误差情况没有任何说明。是没有重报吗?不可能。与往次人口普查相比,这次人口普查面临的最大问题是流动人

口大量增加,人户分离现象严重,重报现象肯定会比 10 年前增加。如果重报问题存在的话,漏报率肯定会大于公布的 1.81%。

4. 小结

第一. 1990 年后,国家统计局将每年的人口抽样变动调查得到的人口出生率加 2—3 个百分点,即平均每年加出生人口 275 万,10 年累积“加水”大约 2750 万。第五次人口普查公布的人口总数没有超出经过“加水”后的历年人口抽样变动的调查结果,或者说,第五次人口普查的人口漏报率与国家统计局所掌握的历年各省市上报的人口漏报率大体相当。从这个角度说,第五次人口普查公布的人口总数是可信的。

第二. 只有把国家计生委报表出生人口数上调 35%—40%,或者把国家计生委报表的计划生育率由 94.64%下调到 70%左右之后,国家计生委掌握的出生人口数才与国家统计局公布的出生人口数大体一致。从历年国家计生委组织的突击抽查结果看,2000 年的报表出生人口水分高达 40%,计划生育率只有 70%的假设似乎不太可能成立。从这一点上看,国家统计局历年公布的出生人口数和人口总数应该是估计的最大值。

第三. 第五次人口普查调查的人口既有漏报,也有重报。公布的漏报率为 1.81%或许是重报和漏报相抵消后的净误差率。

总之,我们认为,第五次人口普查公布的总人口数为 126583 万是可信的,但是人口年龄结构的误差非常严重,需要重新估计漏报率和重报率后加以相应地调整。

三. 对第五次人口普查数据重报问题的分析

第五次人口普查存在比较严重的漏报问题成为一个重要的、不可忽视的问题。尽管第五次人口普查官方公布的漏报率高达 1.81%,然而,由于缺乏比较详尽的漏报人口年龄结构等数据信息,因此对漏报人口的分布和判断成为比较困难的问题。目前比较一致的看法是:漏报人口大体主要落在 20 岁以下,特别是 0—9 岁儿童漏报可能是主要因素(于学军,2002;崔红艳等,2002)。由于人口普查存在比较严重的漏报问题,那就意味着对相关人口状态和过程的推断将会存在很大的偏差。人口数据分析、调整和间接估计等任务将异常繁重,尤其

是对人口数量、质量、结构、水平、过程和发展趋势等一系列问题的研究和评判将直接影响到人口与计划生育政策的制定和实施,影响对未来中国人口发展过程和发展趋势判断。同时,由于人口过程具有长周期、不可逆和滞后性,因此任何决策失误都会对未来中国社会的发展导致不堪设想的严重后果。

除了漏报问题以外,“五普”是否也存在比较严重的重报问题?遗憾的是“五普”并没有给出可供参考的数据(乔晓春,2002),这就给正确评价“五普”数据带来了一定的困难。可是,为了弥补人口普查数据的缺陷就必须对人口普查数据的问题所

在进行全面、详细的分析和评价,然后根据已有的可靠信息对“五普”数据进行重新修正,以期达到对中国人口发展状况、发展过程的全面认识和准确把握的目的,以防错误信息对人口问题判断的误导。在此我们试图对第五次人口普查存在的重报问题予以论述,以便为今后“五普”数据年龄结构调整、生育水平的判定以及死亡等问题的研究提供参考。

由于我国由计划经济向市场经济体制的转换,社会、经济、文化构成向多元化方向发展,受户籍制度改革和社会结构转型的影响,人口信息的收集遇到前所未有的困难,这不仅是数据量增加带来的困难,更是由人口结构的复杂性产生的巨大影响,因此高质量数据获得难度加大。纵观我国大规模人口调查的历史,我国已进行了五次人口普查和多次大规模的人口抽样调查,人口数据信息日益丰富,尤其是1982年人口普查数据质量之高是前所未有的。这对分析和判断以前和以后的人口普查与人口调查起到重要的作用。1990年人口普查也是一次较高质量的人口普查,因此上述数据对分析、研究第五次人口普查数据的质量,如数据漏报和重报问题具有重要意义。本研究所采用的数据为《中国1982年人口普查资料(电子计算机汇总)》,《中国1990年人口普查资料》以及《中国2000年人口普查资料》数据。在进行数据质量分析过程中所使用的数据不包括香港、澳门和台湾。

为了对第五次人口普查存在的重报问题予以分析,本项研究从年龄结构分析出发,采用的方法是根据已有的高质量人口数据推算2000年人口状况,将推算结果与2000年实际普查数据进行比较分析,以期达到对“五普”数据的数据质量进行全面衡量和认识的目的。在分析方法上采用存活分析法、对普查时点进行调整和数据汇总比较三个步骤。

首先,存活分析方法的应用可以达到分析2000年人口普查中相应人口数据的准确程度的目的。存活分析的基本表达式为:

$${}_nP_{t_2}(x+n) = {}_nP_{t_1}(x) \cdot [{}_nL(x+n)/{}_nL(x)]$$

式中 x 的取值范围是0岁到90岁; ${}_nP_{t_1}(x)$ 是在 t_1 时刻年龄在 x 岁至 $x+n$ 岁的人口数; ${}_nP_{t_2}(x+n)$ 是在 t_2 时刻年龄在 $x+n$ 岁至 $x+2n$ 岁的人口数; ${}_nL(x)$ 确切年龄在 x 至 $x+n$ 队列存活人年数; ${}_nL(x+n)$ 确切年龄在 $x+n$ 至 $x+2n$ 队列存活人年数。

第二,由于第五次人口普查与第三次人口普查和第四次人口普查的普查时点不同,为了使存活分

析更接近第五次人口普查时的人口状态,相应地将存活分析的时点调整到与第五次人口普查相同,即11月1日。各年龄组数据时点人口计算方法是:

$${}_nP'_{t_{11}}(x) = {}_nP_{t_1}(x) + ({}_nP_{t_2}(x+n) - {}_nP_{t_1}(x)) \cdot 0.25$$

式中 ${}_nP'_{t_{11}}(x)$ 是在2000年11月1日时年龄在 x 岁至 $x+n$ 岁的人口数; ${}_nP_{t_1}(x)$ 是在2000年7月1日年龄在 x 岁至 $x+n$ 岁的人口数; ${}_nP_{t_2}(x+n)$ 是在2001年7月1日时年龄在 $x+n$ 岁至 $x+2n$ 岁的人口数。

第三,通过总量和分组的方法将推测数据与“五普”数据进行分析和比较。

最后需要说明的是由于第三次人口普查现役军人年龄结构是按5岁分组,缺少按1岁分组的数据资料,但为了尽可能准确地反映普查数据间的数量关系,对第三次人口普查现役军人采取按“四普”公布现役军人年龄结构进行分解的方式得到1982年1岁组现役军人的年龄结构。由于各年龄组现役军人的相对数量较少,且现役军人年龄结构变化不会很大,因此,现役军人年龄结构按1岁分解后的年龄别人口数不会对总人口的年龄结构带来很大的分解误差。

对人口普查数据质量的检验事关重大,为了表明推算方法的有效性和研究问题的需要,使用高质量数据进行检验就显得尤为重要,这不仅关系到对推算准确性的评价,而且是涉及到对事后质量检验的可靠性评价。由于可获得的高质量数据只有“三普”和“四普”数据,因此,本文采用“三普”和“四普”数据进行分析方法的有效性检验。在用“三普”数据对“四普”8—90岁人口进行存活分析时,考虑到“四普”死亡漏报问题,因此将男、女平均预期寿命分别设定68.3岁和71.3岁,对1982与1990年间各年度的男性和女性预期寿命采取线性差值的方法得到。根据上述假定的推测结果见表5。从表5可以看到,根据“三普”数据推测1990年8—90岁男性人口为49355.6万、女性人口为46641.3万,男女合计为95996.9万;相应地,1990年实际普查数男性为49294.9、女性为46481.0,男女合计为95775.9万。推测值与实际调查的差值为221万人,即推测数据比实际普查数多了220多万人。根据1990年质量抽查漏报率为0.7‰,重报率为0.1‰,0—100岁及以上的人数净误差为0.6‰,即推算可能漏报79.1万人,净误差为67.8万。因此从推算结果可以断定1990年人口普查漏报的发生,并与事后抽样结果的结论是一致的。此外,根据推测可以看到18—60岁数据质量相对较高,净误差

为 196.0 万人,而 60 岁以上则相对较差。由于“三普”和“四普”调查时点相同,不涉及时点调整问题,因此从推算符合精度上来看相对较高。总之,数据检验表明上述分析方法可以比较确切地反映问题所在。

表 5 1990 年人口普查数据与 1982 推测数据比较表
单位:万人

1990 年人口数		1990 年普查		1982 年普查推算	
		男	女	男	女
8—90 岁		49294.9	46481.0	49355.6	46641.3
合计		95775.9		95996.9	
其中	18—90 岁	38620.7	36453.0	38873.4	36799.6
	合计	75073.7		75673.0	
	18—60 岁	34415.5	31762.0	34470.0	31903.5
	合计	66177.5		66373.5	
	61—90	4205.2	4691.0	4403.4	4896.1
	合计	8896.2		9299.5	

由于我们可以获得的准确和比较准确的近期人口普查数据为 1982 年人口普查和 1990 年人口普查数据,虽然从时间的有效性的角度看,1990 年

人口普查数据好于 1982 年人口普查,但由于 1990 年普查也存在一定的人口漏报问题,漏报率达 0.7‰,相比之下 1982 年人口普查漏报率仅 0.56‰,因此从数据的有效性和完整性上来看,1982 年人口普查是一个难得的准确数据标准。此外,由于死亡水平变化相对比较缓慢且方向确定,而且对总人口的变化影响相对比较确切,因此通过恰当的假定对现存人口的预测误差不会很大。正是出于上述考虑,为了比较全面验证和推断 2000 年人口普查数据存在的主要问题,下面研究采取 1982 年人口普查为主、1990 年人口普查为辅,结合部分 2000 年人口普查数据的分析策略来分析“五普”的重报问题。

根据 2000 年人口普查数据计算男性和女性人口的预期寿命分别为 70.61 和 74.45 岁,考虑到死亡漏报和男性和女性预期寿命的差距问题,因此将 2000 年男性和女性预期寿命估计为 70 和 73 岁,对 1982 或 1990 年与 2000 年间各年度的男性和女性预期寿命同样采取线性差值的方法得到。基于上述数据、算法和基本思路,得到 2000 年 11 月 1 日中国人口 18—90 岁推测值,具体数值见表 6。

表 6 2000 年人口普查数据与人口推测值比较表

单位:万人

		2000 年普查		1982 年普查推算		1990 年普查推算	
		男	女	男	女	男	女
10—90 岁		55649.0	52880.5	—	—	54384.0	51472.5
合计		108529.5		—		105856.5	
其中	18—90 岁	45978.0	43928.7	45412.2	43256.1	45281.6	43164.2
	合计	89906.7		88668.3		88445.8	
	18—60 岁	40136.5	37755.1	39466.0	36942.3	39546.1	37019.8
	合计	77891.6		76408.3		76565.9	
	61—90 岁	5841.5	6173.6	5946.2	6313.9	5735.5	6144.4
	合计	12015.1		12260.1		11879.9	

从表 6 可以看到,与 1982 年的推算的结果相比,2000 年人口普查得到 18—90 岁的人口数量多了 1238.4 万,其中,18—60 岁多了 1483.3 万,而 61—90 岁则少了 244.9 万。如果不考虑时点转换问题,直接用 1982 年数据推算到 2000 年 7 月 1 日,18—90 岁推算结果是男性 45525.9 万人,女性 43353.7 万人,合计 88879.6 万人,相应地,与 2000 年人口普查的误差为 1027 万人。然而,事后人口普查数据质量抽样结果认为,2000 年人口普查漏报率为 1.8%,即漏报 2246 万人,比较“四普”漏报

情况,这里推算的结果与实际调查结果存在数据矛盾现象。如果根据 1990 年人口普查数据推算,与推算结果比较,10—90 岁 2000 年普查多报 2673 万人,其中 18—90 岁多报 1461 万人。如上所述,由于根据 1990 年人口普查质量抽查的漏报率 0.7‰推算,1990 年可能漏报 79.1 万人,那么,如果假定 1990 年所有漏报人口都发生在 0—80 岁之间(80 岁以上人口比重相对较小,忽略不计),可以推断 10—90 岁 2000 年普查至少多报 2000 万人左右。鉴于上述推测可以确信 2000 年人口普查存在

比较严重的重报问题。如果不考虑 10—90 岁人口的重漏与漏报的相抵问题,年龄结构数据的差错率可能会相当大。

经过使用 1982 年人口数据和 1990 年人口数据的推算值与 2000 年人口普查比较可以看到人口重报问题的存在。由于 1982 年人口普查数据资料质量相对较高,可以以此作为数据推测与判断的依据,通过比较推测值和实际普查值看到“五普”误差的分布情况(见图 3)。从图 3 可以看到,重报问题

主要发生在 45 岁以下,尤其是 30—40 岁和 20 岁以下。如 30—40 岁重报人口占全部重报人口一半以上,具体情况见表 7。从表 7 可以看到误差最大的年龄组依次为 38、30、31、32,尤其是 37、38、39 和 40 岁连续出现负值,即推算低估现象,而 1982 年对 1990 年进行推测误差却没有发生连续为负值或正值的情况,可以断定 2000 年普查重报现象的发生,而绝非运算方法带来的误差。

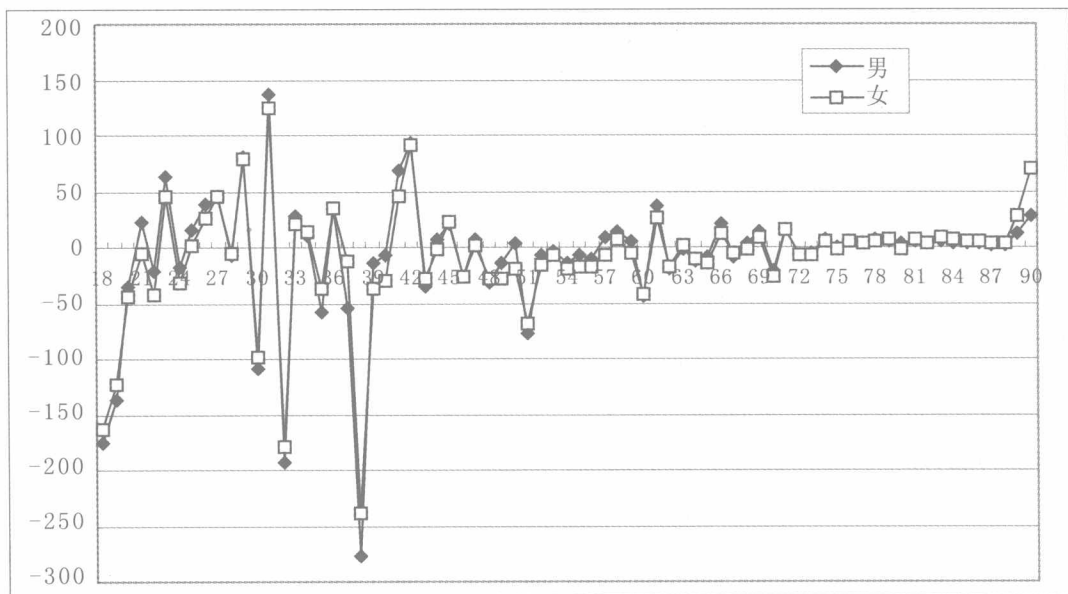


图 3 2000 年人口普查数据与 1982 年普查数据推算值误差分布图①单位:万人

表 7 2000 年人口推测值与人口普查数据误差比较表②

单位:万人

年龄	2000 年普查人口数		2000 年人口推测值		估计误差	
	男	女	男	女	男	女
30	1444.0	1362.0	1334.7	1263.7	-109.3	-98.3
31	1286.0	1220.5	1422.6	1345.5	136.6	125
32	1429.3	1345.9	1237.0	1166.2	-192.3	-179.7
33	1116.3	1060.6	1145.2	1081.2	28.9	20.6
34	1278.3	1208.1	1288.8	1222.8	10.5	14.7
35	1282.9	1200.7	1224.4	1164.5	-58.5	-36.2
36	1241.1	1177.4	1275.6	1212.4	34.5	35
37	1439.8	1350.3	1385.1	1338.7	-54.7	-11.6
38	1081.5	1012.6	804.6	773.4	-276.9	-239.2
39	581.8	561.0	568.5	523.6	-13.3	-37.4
40	762.0	707.7	754.6	678.7	-7.4	-29
合计	12943.0	12206.8	12441.1	11770.7	-501.9	-436.1
	25149.8		24211.8		-938	

① 根据 1982 年人口普查数据推算,误差=1982 年预测值-2000 年普查值

② 同上。