

中国人口预测方法及 未来人口政策

王广州*

内容提要 人口预测包括人口总量和人口结构预测, 本文从人口总量冈波斯 (Gompertz) 和逻辑斯蒂 (Logistic) 基本预测模型出发, 介绍了经典的队列要素人口预测方法和孩次递进生育人口预测模型。详细推导了孩次递进人口预测模型中的递进生育率参数和参数反推方法, 以 1990 年、2000 年和 2010 年全国人口普查数据为例, 检验孩次递进生育预测方法在人口结构预测中的可靠性。依据 2010 年人口普查数据, 采用孩次递进生育预测方法对未来中国人口发展过程进行情景模拟。研究结果表明, 全面二孩生育政策调整不足以实现总和生育率稳定在 1.8 左右的人口发展战略目标。抑制生育的因素不断强化, 人口长期快速减少和快速老龄化趋势不可避免。低生育率陷阱、劳动力持续老化和人口快速老龄化等人口结构性失衡问题凸显。

关键词 人口预测方法; 队列要素模型; 孩次递进模型; 递进生育率

JEL 分类号 C54; J24

预测未来不仅是学术研究的课题, 也是涉及日常生活的基本命题。人口预测就是对未来人口总量、结构进行数学推算。人口预测的基本逻辑是: 现在是过去变化的结果, 未来可以从过去和现在推算。引起人口总量和结构变化的因素包括生育、死亡和迁移, 即人口平衡方程:

$$P_{(t+1)} = P_{(t)} + B_{(t,t+1)} - D_{(t,t+1)} + I_{(t,t+1)} - E_{(t,t+1)}$$

$P_{(t+1)}$ 为时点 t+1 总人口, $P_{(t)}$ 为时点 t 总人口, $B_{(t,t+1)}$ 为时点 t 至 t+1 出生人口, $D_{(t,t+1)}$ 为时点 t 至 t+1 死亡人口, $I_{(t,t+1)}$ 为时点 t 至 t+1 迁入人口,

* 王广州: 中国社会科学院人口与劳动经济研究所 北京市朝阳区曙光西路 28 号中冶大厦 10 层 100028 电子邮箱: wanggz@cass.org.cn。

基金项目: 国家社科基金重大项目“人口统计调查的国际前沿理论及其在中国的应用”(项目编号: 16ZDA090)。

$E_{(t,t+1)}$ 为时点 t 至 $t+1$ 迁出人口。 $B_{(t,t+1)}$ 和 $D_{(t,t+1)}$ 是人口 $P_{(t)}$ 自身变化引起的,称为人口的自然变动。 $I_{(t,t+1)}$ 和 $E_{(t,t+1)}$ 是人口 $P_{(t)}$ 以外变动引起的,也称为人口的机械变动。研究时间序列的 $P_{(t+1)}$,就需要对 $B_{(t,t+1)}$ 、 $D_{(t,t+1)}$ 、 $I_{(t,t+1)}$ 和 $E_{(t,t+1)}$ 进行深入的研究。

从人口平衡方程来看,引起人口总量和结构变化的因素包括生育、死亡和迁移三个方面。人口平衡方程是人口总量、结构预测最基本的出发点。通过人口平衡方程进行数学推算,是从已知推算未知。人口预测万变不离其宗,都是以人口平衡方程为最基本的出发点,预测人口总量、结构的变化过程和特征。

人口预测的科学性与可靠性取决于基数、参数和模型。人口预测的基数、参数和模型的复杂程度取决于对人口预测结果的需求和要求。基础数据、参数分析和采用的模型密不可分,相互依存。为满足研究的各种需要,经常需要对 $P_{(t+1)}$ 进行各种条件下的细分,这就使得人口预测变得非常复杂。

凡事预则立不预则废。对不确定性未来的强烈好奇、兴趣和渴望不仅是大到国家大事,也是小到百姓日常生活决策需要面对的。人口预测是社会经济发展趋势的基本预测,是最具有定量研究和科学规律的预测。人口预测不仅对人口研究本身具有重要的意义,而且其预测结果也得到广泛的应用。

一、人口总量模型

人口总量预测是人口预测中比较基础和粗略的需求,仅对未来人口总量进行预测的人口模型通常其基础数据和参数估计都比较简单。对于人口变化的过程经常采用最基本的冈波斯(Gompertz)或逻辑斯蒂(Logistic)模型。

冈波斯(Gompertz)模型表达为:

$$Y = k \times \alpha^{b^x}$$

对冈波斯模型取对数:

$$\log(Y) = \log(k) + \log(\alpha) \times b^x$$

逻辑斯蒂(Logistic)模型表达为:

$$Y = \frac{1}{K + a \times b^x}$$

对 Logistic 模型进行转换：

$$\frac{1}{Y} = k + a \times b^x$$

对模型进行转换的目的是便于模型参数估计。有很多方法可以拟合冈波斯（Gompertz）和逻辑斯蒂（Logistic）模型，其中一个比较简便、经典的方法是采用部分和（Partial Sum）的方法（Croxtton et al., 1968）。计算机技术的普及极大地提高了研究者做人口总量预测的计算能力，越来越多的复杂数学方法在人口预测中应用。

二、队列要素人口预测模型

年龄结构是人口变动的基本特征，它不仅记录了人口发展的重要历程，而且对未来发展至关重要。因此，对人口总量、结构，特别是年龄结构预测具有重大意义，这也是单纯的人口总量预测方法所无法实现的。队列要素人口预测是人口总量、年龄结构预测最经典的方法。1895 年坎南（Edwin Cannan）提出队列一要素人口预测模型，该方法后来在 1924 年被波利（Arthur Bowley）拓展并用于英国人口预测。1928 年维普顿（Pascal Whelpton）再次独立发现该预测方法，并用于美国国家层面的人口预测。目前的队列一要素人口预测方法非常经典、成熟，其基本框架与三位先驱创立的模型变化不大（Siegel and Swanson, 2004）。

队列要素人口预测目的是通过对年龄别人口状况的研究、模拟和分析，了解人口系统的发展过程。队列要素人口预测方法实际上是采用系统仿真的思想，是结构功能模型，可以对人口年龄结构的动态变化趋势和过程进行模拟分析，其预测结果可以进行很多二次开发或应用分析。该模型既可以根据需要给定不同的参数进行人口预测，也可以进行政策分析和仿真实验研究。

（一）预测算法

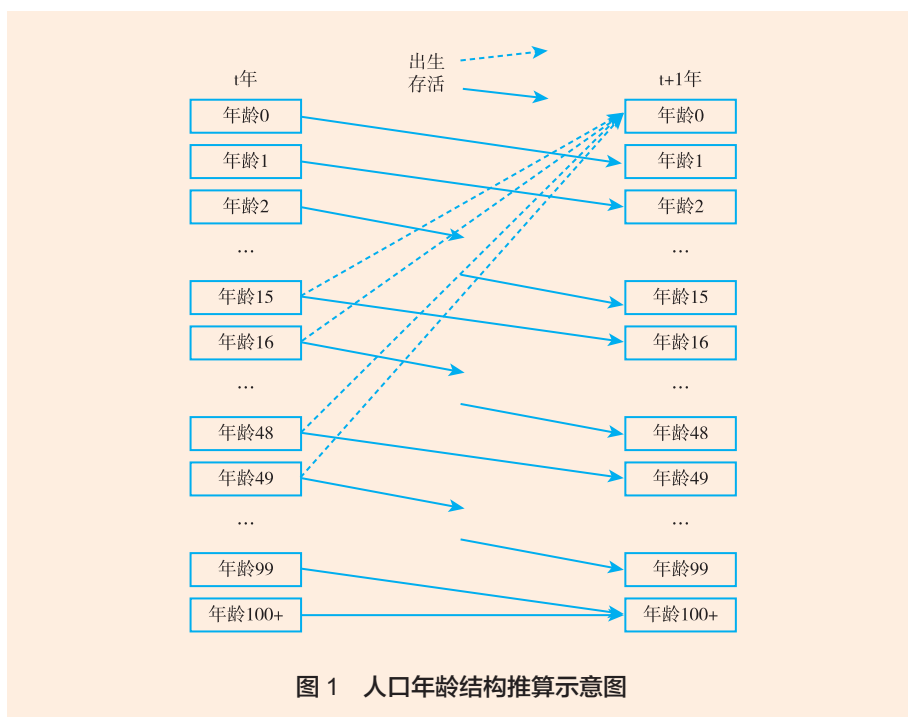
队列要素人口预测方法的基本原理是将人口群体划分为几个部分（见图 1），根据人口变动规律和预测周期反复递推不同年龄别的存活人口剩多少、存活人口生多少和存活人口死亡多少。从图 1 展示队列要素人口预测的基本

原理可以看到：

第一， $t+1$ 年的 1 岁及以上存活人口是 t 年年龄别人口经过一个年度的死亡剩余的存活人口，而且是人口构成主体。

第二， $t+1$ 年的 0 岁人口是 t 年 15 ~ 49 岁育龄妇女在 $[t, t+1]$ 之间生育的结果。

由于持续不断地生育，即新增 0 岁人口，才能使人口的递推过程得以延续，人口群体的构成得以更新。需要注意的是，图 1 只是人口过程变化主体部分的简单示意图，没有考虑更复杂的迁移人口模型的情况。



对于图 1 队列要素人口预测计算过程的数学表达方式可以采用 Leslie 矩阵的形式。为了便于理解，本文采用简单的分部分推算的表示方法，这样做的目的是更清晰地表达队列要素的含义。具体的人口预测模型计算可分为四个部分。

(1) 建立生命表

根据当前人口状况建立生命表或无完整、可靠基础数据时参考模型生命

表。建立生命表的目的在于通过生命表方法，对当前年龄别人口的存活概率予以估计。当然，由于人口规模较小，统计指标可能面临不稳定问题，或者统计数据质量较差，在这种情况下，往往需要数据质量较高的模型生命表的帮助。

(2) 构造存活转移矩阵

$${}_n P_{t_2}(x+n) = {}_n P_{t_1}(x) \times [{}_n L(x+n)/{}_n L(x)]$$

其中，式中 x 的取值范围是 0 岁到 100 岁； ${}_n P_{t_1}(x)$ 是在 t_1 时刻年龄在 x 岁至 $x+n$ 岁的人口数； ${}_n P_{t_2}(x+n)$ 是在 t_2 时刻年龄在 $x+n$ 岁至 $x+2n$ 岁的人口数； ${}_n L(x)$ 是确切年龄在 x 至 $x+n$ 队列存活人年数； ${}_n L(x+n)$ 是确切年龄在 $x+n$ 至 $x+2n$ 队列存活人年数。

(3) 建立生育模型

$$P_{t_2}(0) = [L(0)/2] \times \left\{ \sum [{}_n P_{f_{t_1}}(x) \times {}_n F(x) + {}_n P_{f_{t_1}}(x) \times {}_n F(x+n)] \times {}_n L(x+n)/{}_n L(x) \right\}$$

其中，式中 $P_{t_2}(0)$ 是在 t_2 时刻年龄为 0 岁人口数； ${}_n P_{f_{t_1}}(x)$ 是在 t_1 时刻年龄在 x 岁至 $x+n$ 岁的妇女人口数； x 取值范围是 15 岁至 49 岁； ${}_n F(x)$ 为年龄在 $(x, x+n)$ 之间的育龄妇女生育率； ${}_n F(x+n)$ 为年龄在 $(x+n, x+2n)$ 之间的育龄妇女生育率。

(4) 总人口

$$T_{pop_{t_2}} = \sum P_{t_2}(x)$$

其中，式中 $T_{pop_{t_2}}$ 为 t_2 时刻的总人口数； ${}_n P_{t_2}(x)$ 是在 t_2 时刻年龄在 x 岁至 $x+n$ 岁的人口数。

(二) 预测参数假定与估计方法

任何预测和分析模型都是基于特定的假设条件进行，人口预测主要是回答“如果……，那么……”的问题。人口预测模型也不例外。人口预测参数主要有两种假设方式，第一种假设方式是根据类似的历史发展轨迹，人为设定参数的大小，也就是选择一个类似的参照或目标，进行人为的主观设定；第二种假设方式是收集参数变化或相关的历史数据，采用统计模型或类似统

计模型的方法，预测参数变化的历史趋势和具体数值。对于不考虑人口迁移流动的队列要素人口预测模型来说，根据模型结构可知，参数假定主要是死亡水平和生育水平，死亡水平就是对平均预期寿命假设，生育水平就对总和生育率进行假定。无论平均预期寿命还是总和生育率假定的目的都是在模型中最终转化为年龄别的指标，即年龄别存活概率和年龄别生育率。模型需要对预测年限内各年度年龄别人口存活概率和育龄妇女生育率进行估计。

1. 存活概率

(1) 通过平均预期寿命参数推算存活概率

平均预期寿命是人口预测过程中经常使用和设定的综合指标。在人口预测过程中，需要对人口的预期寿命进行预测或假定。经常使用的方法是基于假定死亡模式不变，死亡水平降低引起平均预期寿命的提高，从而达到对人口变动中存活或死亡参数予以控制的目的。其具体算法如下：

第一，对当前存活函数进行 **logit** 变换。

假定用 $p(a)$ 表示一个人能够从出生（确切年龄为 0 岁）存活至确切年龄 a 岁的存活概率，即 $p(a) = l(a)/l(0)$ 。定义 $p(a)$ 的 **logit** 变换为：

$$y(a) = \text{logit}[p(a)] = (1/2) \times \ln \{ [1 - p(a)] / p(a) \}$$

第二，确定 a 值。

人口学家威廉·布拉斯首次将概率统计中的 **logit** 变换引入人口学研究。布拉斯等人的研究表明，两个不同生命表中的 $p(a)$ 通过 **logit** 函数变换后存在明显的线性关系， $y(a) = a + b \times y_s(a)$ ；式中 a 和 b 为两个待定参数。 $y(a)$ 和 $y_s(a)$ 为两个不同生命表中，从出生存活至确切年龄 a 岁时的存活概率的 **logit** 函数值。

根据上述模型，如果选定了一个标准生命表的存活概率 $y_s(a)$ ，那么新生命表的存活概率 $y(a)$ 的值可根据上述线性关系中的 a 和 b 确定。也就是说， $y(a)$ 的水平将主要由参数 a 的数值决定； $y(a)$ 的模式将主要由参数 b 的数值决定。特别是，当 $b = 1$ 时， $y(a)$ 的模式与 $y_s(a)$ 曲线的模式基本相同。假定在预测期内人口的死亡模式不变且忽略参数 b 对 $y(a)$ 水平的影响，只考虑其对 $y(a)$ 模式的作用，即假定 $b = 1$ ，从而得到新生成的生命表与原始生命表有相

似的模式:

$$y(a) = a + y_s(a);$$

由于人口的死亡模式在短期内具有相对稳定的特征,因此,如果人口预测时期不是很长,通过上述合理假定,可将模型 $y(a) = a + b \times y_s(a)$ 简化为一个单参数 a 的模式。由此可得到:

$$p(a) = ps(a) / \{ ps(a) + [1 - ps(a)] \times \exp(2a) \};$$

上式给出了 $y(a)$ 与参数 a 之间的解析关系,是用来编制新生命表的基本方程。新生命表与当前生命表之间关系的建立,可以得到不同期望寿命下的存活曲线,为人口预测的死亡参数确定奠定基础。此外,从上式可以看出,当 $y_s(a)$ 固定时, $y(a)$ 是关于参数 a 的严格单调递减的连续函数。当 $a < 0$ 时,有 $y(a) > y_s(a)$; 当 $a = 0$ 时,有 $y(a) = y_s(a)$; 当 $a > 0$ 时,有 $y(a) < y_s(a)$ 。显然, $0 < y(a) < 1$, 并且当 a 趋于正时, $y(a)$ 趋于 0; 当 a 趋于负时, $y(a)$ 趋于 1。

人口预测关心 0 岁平均预期寿命的变化。因此,给定 0 岁平均预期寿命确定 a 值是获得存活概率的重要方法。给定 0 岁平均预期寿命确定 a 值方法如下:

由于

$$\begin{aligned} e(x) &= \frac{T(x)}{l(x)} = \frac{1}{l(x)} \int_x^\infty p(a) da \\ &= \int_x^\infty \frac{p_s(a) \{ p_s(x) + [1 - p_s(x)] \times \exp(2a) \}}{p_s(x) \{ p_s(a) + [1 - p_s(a)] \times \exp(2a) \}} da \end{aligned}$$

从上式很难确定参数 a 与 $e(x)$ 之间的解析关系,但可以通过迭代的方法求解 a 的近似值。对 a 的近似计算可以采用近似计算的二分迭代法。通过二分迭代法,根据给定的 $e(x)$ 的值估计及相应的 a 值。通过二分迭代法,根据给定的 $e(x)$ 的值估计及相应的 a 值(路磊, 1992)。当根据某个死亡水平估计出相应的 a 值后,将该值代入 $p(a) = ps(a) / \{ ps(a) + [1 - ps(a)] \times \exp(2a) \}$ 式中,即可得到一组存活概率 $\{p(a)\}$,从而可以编制出新的生命表。当然,对二分迭代法只是求解方法之一,对于 a 值的估计也可以采用其他优化或数值算法。

(2) 通过模型生命表法推算存活概率

除了可以通过上述方法对未来人口的存活状况予以估计外,还可以通过

对死亡水平和死亡模式的研究，根据已有模型生命表给定年龄别人口的存活概率，从而达到进行人口动态模拟分析的目的。

2. 生育率

与平均预期寿命参数设定相比年龄别生育指标要简单得多。从当前育龄妇女的生育频率分布曲线模式 $g(x)$ (生育模式) 出发，再假定总和生育率和生育模式以后，根据 $F(x) = g(x)/TFR$ ，得到预测中所需要的 $F(x)$ ，即年龄别生育率。

总之，队列要素人口预测模型是人口总量、结构预测的基本模型，是对现实复杂情况的抽象和简化，是人口预测最本质、最核心和最容易把握的部分。上述模型参数估计的基本假定是未来人口的生育、死亡模式保持不变，且所研究的人口为封闭人口。当然，随着基础信息的丰富和可靠性提高，对人口预测参数也可以采用更丰富或更完善的方法进行预测，比如预期寿命预测的 Lee-Carter 模型或模型生命表等。

三、孩次递进人口预测

孩次递进人口预测模型与队列要素人口预测模型的思路一样，只是对育龄妇女的生育过程进行了更为科学和精细的测量。

(一) 孩次递进生育模型

尽管孩次递进生育模型与队列要素人口预测模型有许多相似之处，比如生命表的构建、平均预期寿命参数推算，但孩次递进生育模型与通常的总和生育率模型最大的不同是对育龄妇女孩次结构的记录和推算，在整个育龄妇女递进生育过程中，最关键的部分是育龄妇女年龄别孩次构成和孩次递进比例推算。孩次递进生育模型可以划分为几个部分，第一部分是育龄妇女分年龄、分孩次递进；第二部分是未递进妇女的存活人数和新增低龄 0 孩妇女；第三部分是育龄妇女孩次结构更新和递推，具体算法如下。

1. 育龄妇女年龄别、孩次别递进生育子女数：

$$\text{new_born_Women}[i+1, j+1] = \text{women}[i, j] \times (\text{wLx}[i+1] / \text{wLx}[i]) \times \text{pro_}$$

$ht [i, j+1]$

其中: i 为年龄, j 为孩次, $i=[15, 49]$, $j=[0, 4+]$; $new_born_Women [i+1, j+1]$ 为 i 岁妇女生育 $j+1$ 子女, $women [i, j]$ 为 i 岁、 j 孩育龄妇女, $wLx [i]$ 为妇女年龄为 i 岁至 $i+1$ 岁存活人年数, $pro_ht [i, j+1]$ 为 i 岁、 $j+1$ 孩育龄妇女递进比。

2. 育龄妇女年龄别、孩次别没有递进生育:

$Non_born_Women [i+1, j] = women [i, j] \times (wLx [i+1] / wLx [i]) \times (1 - pro_ht [i, j+1])$

$Non_born_Women [15, 0] = female [14] \times (wLx [i+1] / wLx [i])$;

其中: $Non_born_Women [i+1, j]$ 为 i 岁且未发生递进生育的 j 孩妇女;
 $female [14]$ 为 14 岁妇女人数;

3. 年龄别、孩次别妇女构成生育更新:

$women [i, j] = Non_born_Women [i, j] + new_born_women [i, j]$

经过一个年度的递进生育后, 当前妇女的分年龄孩次结构可以分为两种情况。第一种情况, 发生递进生育的育龄妇女的年龄和孩次都发生变化; 第二种情况, 未发生递进生育的妇女只是年龄增长, 孩次没有发生变化。记录育龄妇女的年龄、孩次构成, 目的是进行下一次递进生育和状态更新, 这样既有低龄无孩妇女源源不断地进入递进生育, 也有超出育龄期妇女不断退出。

(二) 递进生育率与递进生育模式

通过基础数据计算各孩次的年龄别递进生育率和递进生育模式, 目的是设定分孩次总和递进生育率参数时, 根据新的参数计算分年龄、分孩次递进比。分孩次递进生育率与递进生育模式计算方法简单总结如下。

1. 对于 $0 \rightarrow 1$ 孩年龄别递进生育率:

$$aspfr(i, 1) = h(i, 1) \times \prod_{k=15}^{i-1} [1 - h(k - 1, 1)]$$

其中, i 、 k 为年龄, $aspfr(i, 1)$ 中为 $0 \rightarrow 1$ 孩 i 岁 1 孩递进生育率, 1 代表 $0 \rightarrow 1$ 孩递进; $h(i, 1)$ 为 i 岁 1 孩递进比, 规定 $h(14, 1) = 0$, 那么, $1 - h(14, 1) =$

1, 也就是说, 认为 14 岁没有初育行为; $\prod_{k=15}^{i-1} [1 - h(k-1, 1)]$ 为 14 岁开始到 $i-1$ 岁连续未递进比例的乘积。

0 → 1 孩总和递进生育率为:

$$TPFR_1 = \sum_{15}^{49} aspfr(i, 1)$$

2. 1 → 2 孩年龄别递进生育率:

对于不同生育间隔的 1 → 2 孩年龄别递进生育率:

$$sub_aspfr(i, p, 2) = aspfr(i - p, 1) \times h(i, 2) \times \prod_{k=i-p+1}^{i-1} [1 - h(k, 2)]$$

其中, i, k 为年龄且 $i, k \geq 16$, p 为 1 → 2 孩递进的时间间隔, 取值 1 ~ 34;

1 → 2 孩年龄别递进生育率为该年龄 1 → 2 孩各个生育间隔递进生育率的合计:

$$aspfr(i, 2) = \sum_{p=1}^{i-15} sub_aspfr(i, p, 2)$$

2 孩总和递进生育率:

$$TPFR_2 = \sum_{15}^{49} aspfr(i, 2)$$

TPFR₃ 等大于 2 孩的递进生育率和总和递进生育率与 1 → 2 孩的算法相同。

3. 总和递进生育率:

$$TPFR = TPFR_1 + TPFR_2 + \dots + TPFR_n, \quad n \text{ 为生育子女总数。}$$

4. 递进生育模式:

$$PG1(x) = ASPFR1(x) / TPFR_1$$

$$PG2(x) = ASPFR2(x) / TPFR_2$$

$$PGn(x) = ASPFRn(x) / TPFR_n$$

经过上述算法得到孩次别递进生育模式为孩次递进预测参数奠定基础。

(三) 分孩次年龄别递进比参数

根据给定的分孩次递进生育模式和分孩次总和递进生育率得到年龄别孩次递进生育率。然后, 根据年龄别孩次递进生育率推算分孩次年龄别递进比参数, 运算的过程是第(二)部分的逆运算, 具体算法如下。

1. 年龄别孩次递进生育率

根据现有育龄妇女孩次结构和递进生育数据得到：

$$\begin{aligned} PG1(x) &= ASPFR1(x) / TPFRR_1 \\ PG2(x) &= ASPFR2(x) / TPFRR_2 \\ PGn(x) &= ASPFRn(x) / TPFRR_n \end{aligned}$$

假定 $TPFR' [j]$ 为已知或设定参数 $TPFR' [j]$ ，可以推算递进生育模式不变且假定参数条件下，年龄别各孩次的递进生育率：

$$aspfr' [i, j] = pg [i, j] \times TPFRR' [j]$$

当然，也可以推算生育模式转变条件下的年龄别孩次递进生育率，但需要注意的前提条件是递进生育模式的转变要与递进生育模式转变的实际规律相符，这样推算出来的参数才有意义，但有些人设定的参数有可能是不符合实际规律的，即有可能与实际规律相左，计算的结果可能会出现负值的不符合实际人口含义的结果，这里是要有人口学常识和知识，并需要特别注意的。

2. 从年龄别孩次递进生育率推算年龄别孩次递进比

从年龄别孩次递进生育率推算年龄别孩次递进比是一个从孩次递进比计算孩次递进生育率的逆运算（郭志刚，2004）。由于是一个递推的逆运算，所以可以连续递推，原因是假定已知分孩次年龄别递进率且起始年龄孩次递进生育率与起始年龄孩次递进比相等。

一孩递进的情况比较简单， $0 \rightarrow 1$ 年龄别孩次递进比的推算方法如下。

由于

$$aspfr' (i, 1) = h(i, 1) \times \prod_{k=15}^{i-1} (1 - h(k - 1, 1))$$

其中， $aspfr' [i, 1]$ 已知，而且 $h [15, 1] = aspfr' [15, 1]$ ，当 $i > 15$ 时，递推各年龄 1 孩年龄别递进比的方法如下。

$$h[i, 1] = \frac{aspfr' [i, 1]}{\prod_{k=16}^{i-1} (1 - h[k - 1, 1])}$$

二孩及以上的情况比较复杂， $1 \rightarrow 2$ 年龄别孩次递进比的推算方法如下：

已知 $aspfr' [i, 1]$ 和 $aspfr' [i, 2]$ ，根据

$$sub_aspfr(i, p, 2) = aspfr(i - p, 1) \times h(i, 2) \times \prod_{k=i-p+1}^{i-1} [1 - h(k, 2)];$$

$$aspfr(i, 2) = \sum_{p=1}^{i-15} sub_aspfr(i, p, 2);$$

即:

$$aspfr(i, 2) = \sum_{p=1}^{i-15} \{aspfr(i - p, 1) \times h(i, 2) \times \prod_{k=i-p+1}^{i-1} [1 - h(k, 2)]\};$$

推算:

$$pro_ht [16, 2] = aspfr' [16, 2] / aspfr' [15, 1] ;$$

$$\text{对于 2 孩及以上递进: } pro_ht [16, j] = aspfr' [16, j] / aspfr' [15, j-1] ;$$

j 为孩次。

16 岁及以上 1 孩间隔递进:

$$ht [16, 2] = aspfr' [16, 2] / aspfr' [15, 1] ;$$

$$duration[i, p, 2] = aspfr[i - p, 1] \times \prod_{k=i-p+1}^{i-1} (1 - ht[k, 2]); p = 1 \dots 34; i = 16 \dots 49;$$

$$h(i, 2) = \frac{aspfr(i, 2)}{\sum_{p=1}^{i-15} duration[i, p, 2]}$$

根据不同假定的孩次总和递进生育率和年龄别孩次递进模式推算新的年龄别孩次递进比参数, 比如, 根据 1990 年全国人口普查育龄妇女的曾生孩次结构和 1990 年度分年龄分孩次生育数据计算 0 → 1 孩、1 → 2 孩和 2 → 3 孩年龄别孩次递进比和总和递进生育率为基础。以此为例, 1990 年 0 → 1 孩、1 → 2 孩和 2 → 3 孩总和递进生育率分别为 0.992217、0.713059 和 0.284747, 设定新的 0 → 1 孩、1 → 2 孩和 2 → 3 孩总和递进生育率分别为 0.95、0.6 和 0.2, 如果递进生育模式不变, 那么, 可以推算出与此孩次递进生育率相对应的年龄别孩次递进比并用于未来的人口预测。

四、人口预测方法检验

人口总量和结构预测的目的是把握未来人口变化趋势和相互关系, 通过人口预测数据结果的对比分析, 研究人口变化的基本规律和面临的问题。队列要素人口预测与孩次递进人口预测各有各的优势, 队列要素人口预测方法的基础数据要求相对比较简单, 但与计划生育政策等结合不够紧密, 而孩次

递进人口预测不仅具有队列要素人口预测的功能，而且可以获得孩次结构，因此在子女构成的分析中具有重要意义。以 1990 年全国人口普查数据为基础，对 2000 年和 2010 年中国人口的年龄结构等进行预测，并与相关人口普查数据进行比较，以此来展示基础数据、预测模型、预测参数和预测结果的可靠性。根据 2000 年和 2010 年全国人口普查数据推算相应的预测参数，预测参数见表 1，假定死亡模式和孩次递进生育模式不变，普查间参数采用线性插值。

		1990 年	2000 年	2010 年
预期寿命	男	68.41	70.81	75.77
	女	71.76	74.51	80.60
递进生育率	0→1 孩	0.9922	0.9727	0.9726
	1→2 孩	0.7131	0.3329	0.3581
	2→3 孩	0.2847	0.0267	0.0427
	3→4+ 孩	0.1355	0.0025	0.0058
出生性别比		110.92	116.86	117.96

（一）年龄结构预测

年龄结构是人口变化的基础，对人口年龄结构的预测是人口预测最常规的内容之一。对年龄结构的预测结果检验也是对人口结构性预测能力的考验。

1. 2000 年预测结果检验

首先，从总人口来看，2000 年总人口预测结果为 124729.50 万人，比 2000 年人口普查的 124261.22 万人多 468.28 万人，预测人数比普查人数多了 3.77‰。

其次，从年龄构成来看，将 2000 年人口预测年龄结构与 2000 年人口普查数据进行比较发现（见图 2），2000 年人口预测与 2000 年人口普查的年龄结构之间的差距主要是低龄人口（0 ~ 9 岁）。预测 2000 年 0 ~ 9 岁人口总数为 18628.23 万人，比 2000 年人口普查调查人口的 15913.10 万多了 2715.13 万人，平均每个年龄组多 270 万人左右，多了 17.06%。预测 2000 年 10 岁及以上人口为 106101.26 万人，比 2000 年人口普查调查数的 108348.13 万人少了

2246.87 万人，少了 2.07%。

总之，尽管根据 1990 年基础数据和 2000 年普查推算参数得到预测人口与人口普查人口总量非常接近，但各个年龄之间的差距还是有一些不同的，误差主要来源于低龄人口。



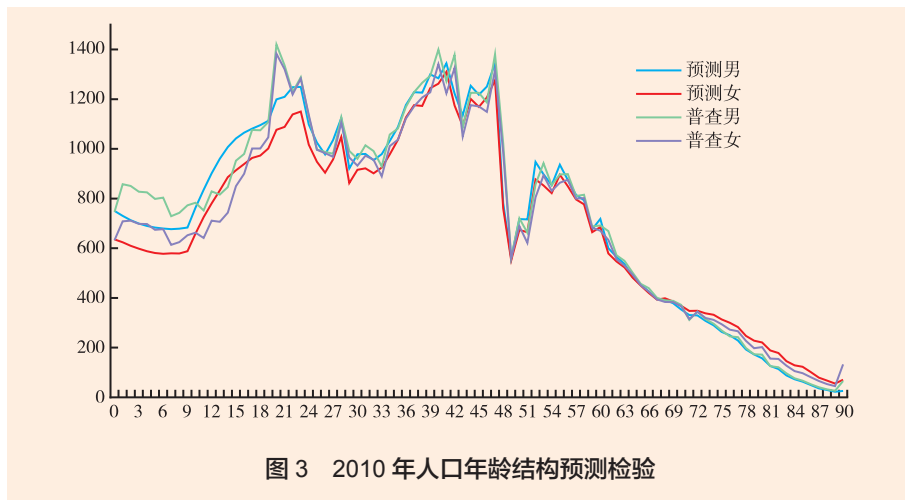
2. 2010 年预测结果检验

同样，从总人口的预测结果来看，2010 年总人口预测结果为 130378.35 万人，比 2010 年人口普查的 133281.09 万人少 2902.74 万人，预测人数比普查人数少了 2.18%。

然而，年龄构成的预测结果却有更大不同。2010 年预测年龄结构与 2010 年人口普查数据（见图 3）之间的差距仍然主要是低龄人口，只是偏差的方向发生了改变，0 ~ 9 岁与 10 ~ 19 岁不同。2000 年 0 ~ 9 岁人口预测结果与普查数据相比高估了，2010 年 0 ~ 9 岁人口预测结果与普查相比却低估了。具体的情况是 2010 年 0 ~ 9 岁预测人口总数为 12944.64 万人，比 2010 年人口普查调查人口 14641.42 万人少了 1696.78 万人，平均每个年龄组少 170 万人左右，少了 11.59%。2010 年预测 10 ~ 19 岁人口为 18536.74 万人，比 2010 年人口普查调查数 17479.76 万人多了 1056.98 万人，多了 6.05%，比 2000 年原来多 17.06% 的情况大大下降。预测 2010 年 20 岁及以上人口为 98896.97 万人，

比 2010 年人口普查调查数 101159.91 万人少了 2262.94 万人，少了 2.24%。

总之，根据 1990 年基础数据对 2010 年人口年龄结构的预测结果与 2000 年普查既有相同点又有不同点，相同点是低龄人口预测误差相对较大，但各个年龄之间的差距还是有一些不同。不同点是 2010 年人口年龄结构 0 ~ 9 岁预测结果低于 2010 年人口普查调查结果，而 10 ~ 19 岁预测结果高于 2010 年调查，但相比 2000 年的预测结果与调查误差大大降低。综合误差的结果是，2010 年总人口预测结果低于 2010 年人口普查调查人口总数，大体上低 2.18% 左右。

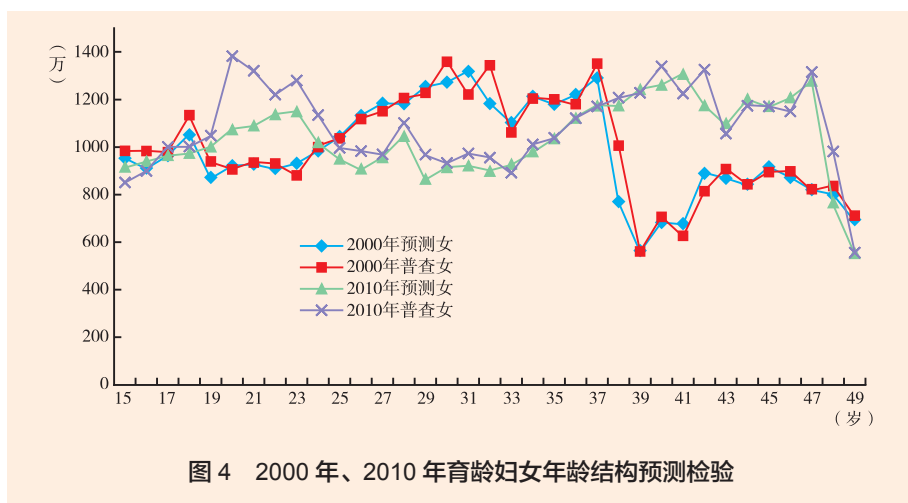


（二）育龄妇女年龄结构预测

育龄妇女总量和生育水平是决定出生人口及低龄人口预测数据可靠性的重要因素。从 2000 年育龄妇女年龄别人口数的预测结果与 2000 年人口普查调查数据的比较来看，2000 年育龄妇女年龄别人口数与 2000 年人口普查调查数据非常吻合（参见图 4），2000 年育龄妇女预测数与普查数分别是 34380.81 万人和 34970.09 万人，预测数比普查数少了 589.28 万人，只少了 1.69%。与 2000 年的情况有所不同，2010 年育龄妇女年龄结构预测与 2010 年人口普查相比相对差距较大（见图 4）。2010 年育龄妇女预测数与普查数分别是 336379.02 万人和 337977.97 万人，预测数比普查数少了 1598.95 万人，少了 4.73%。两者的差

别主要体现在 19 ~ 24 岁。2010 年 19 ~ 24 岁育龄妇女预测数与普查数分别是 6471.50 万人和 7386.80 万人，预测数比普查数少了 915.30 万人，少了 12.39%。

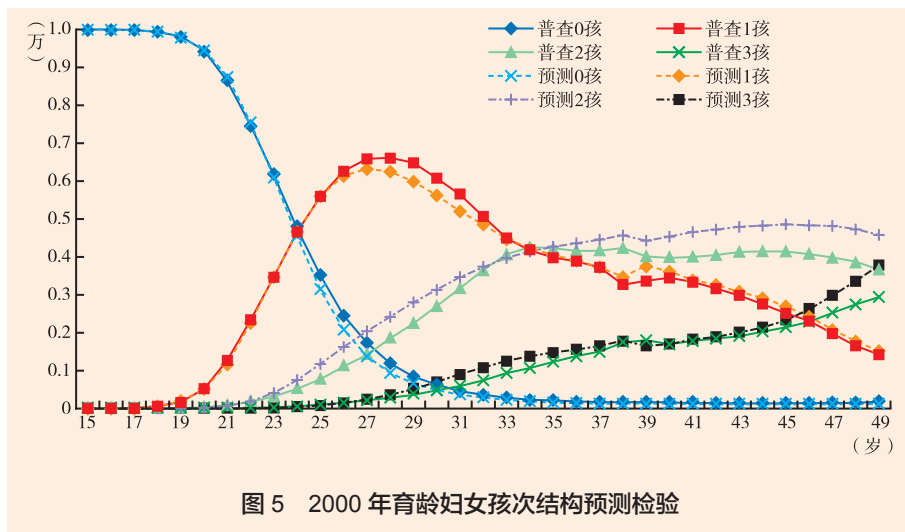
总之，如果假定育龄妇女孩次递进生育水平是准确的，那么，育龄妇女年龄别人口预测误差的大小将影响到出生人口的总数与孩次结构。同时，对比 2000 年和 2010 年育龄妇女总量、结构与人口普查结果的偏差可以看到，预测的结果更多的可能是由于低龄人口数据之间的误差而导致远期人口预测的累积误差增大。



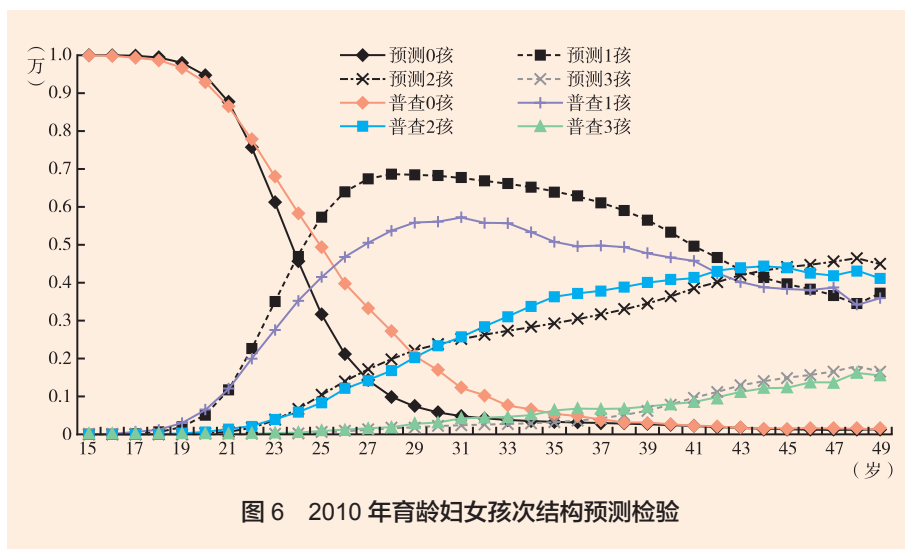
（三）育龄妇女孩次结构预测

如果假定生育模式不变，即 1990 年以后的孩次年龄别生育模式与 1990 年相同，只是不同孩次的总和递进生育率不同。那么，根据 2000 年和 2010 年人口普查数据计算的分孩次年龄别总和递进生育率推算各孩次年龄别递进比，并以此为参数预测 2000 年和 2010 年育龄妇女的孩次结构，预测结果与普查调查数据进行比较（参见图 5 和图 6，实线为普查值、虚线为预测值）。

从对 2000 年育龄妇女孩次结构的预测值来看，2000 年年龄别无孩妇女所占比例的预测值曲线与 2000 年人口普查数据推算值非常接近。1 孩、3 孩次之，2 孩的误差相对大一些（见图 5）。



与 2000 年育龄妇女孩次结构预测的情况不同，2010 年育龄妇女孩次结构的预测值与 2010 年人口普查数据推算值还是有比较大的差别（见图 6）。整体来看，2010 年预测无孩妇女和 1 孩妇女所占的比例与 2010 年人口普查数据调查的比例差别较大，24 ~ 34 岁无孩妇女的比例比普查调查的比例低很多，而 23 ~ 40 岁 1 孩妇女预测值比普查调查的比例又高很多，两者的差距在 5% ~ 15%。



2孩和3孩的情况要好很多，2010年预测2孩妇女和3孩妇女所占的比例与2010年人口普查数据调查的比例差别不大。

通过对育龄妇女孩次结构的预测结果检验可以看到，孩次结构预测是相互联系的，某一个孩次的误差必然会影响到其他孩次，因此，这是一个相互联系而非相互独立的预测，误差的累积会越来越大。离调查数据越近的预测结果与实际调查的误差越小，预测越远的结果由于误差的累积和对未来不确定性的判断差距，预测的误差会越来越大。

总之，除了对人口总量、结构和妇女孩次结构预测以外，还可以对亲子结构等进行预测。随着对人口结构性数据指标需求越来越细致，预测方法必然越来越复杂，基础数据的结构和质量要求也越来越高。人口分组越细致，需要的人口总量规模就越来越大。因此，对于小区域人口预测确实是一个更加严峻的挑战。

五、中国人口预测

从对中国历史人口发展过程的调查数据与预测数据的比较可以看到，短期内人口预测结果与调查结果还是非常接近或吻合的，但长期的预测会由于参数估计问题造成累积误差增大，与实际人口的偏差可能会越来越大。为了解决偏差增大的问题往往采取增大估计区间的方法。然而，如果预测的期限远远超过现有历史数据的积累，历史数据所总结或反映的趋势就很难推广到一个更长的区间范围内。换句话说就是，比较符合实际的参数统计估计几乎是不太可能的。

（一）情景预测参数及主要预测结果

为了反映不同政策的长期趋势和人口后果，本文采用情景预测分析的方法，目的是比较不同生育水平对中国人口总量、结构的长期影响，而非对未来人口总量结构的区间估计。为了简化情景预测条件，反映不同情景生育水平的影响，本文对所有情景的平均预期寿命、出生性别比参数都给定相同的区间，对一孩、三孩和四孩及以上也采用相同的参数，而只对递进生育率

的二孩生育水平设定四种情景水平，情景一是如果保持 2010 年的二孩生育水平即 35.82% 的生育一孩的育龄妇女终身会生育二孩。情景二、情景三、情景四分别为从 2017 年开始 40%、60% 和 80% 的育龄妇女终身生育二孩，2013—2017 年采用由 0.3582 现行插值 60% 的办法，具体情景预测参数设定见表 2。

指标			2010 年	2020 年	2050 年	2100 年
递进生育率	0 → 1 孩		0.9733	0.9733	0.9733	0.9733
	1 → 2 孩	情景一	0.3582	0.3582	0.3582	0.3582
		情景二	0.3582	0.4	0.4	0.4
		情景三	0.3582	0.6	0.6	0.6
		情景四	0.3582	0.8	0.8	0.8
	2 → 3 孩		0.0427	0.0427	0.0427	0.0427
3 → 4+ 孩		0.0058	0.0058	0.0058	0.0058	
平均预期寿命	均值	男	75.77	76.67	79.79	85
		女	80.60	81.37	84.23	89
	下限	男	75.77	76.44	78.90	83
		女	80.60	80.93	82.46	85
	上限	男	75.77	77.00	81.12	88
		女	80.60	81.49	84.68	90
出生性别比	均值		117.96	107	107	107
	下限		117.96	102	102	102
	上限		117.96	110	110	110

根据表 2 的参数设定，得到以下人口模拟结果（见表 3）。

首先，从总人口来看，如果二孩递进生育率一直保持在 2010 年的生育水平，那么，预计到 2100 年中国总人口将比 2010 年减半。如果全面放开二孩生育政策后，只有 40% 的生育一孩的育龄妇女终身生育第二个孩子，那么，2100 年中国总人口比 2010 年总人口也近似减半，人口高峰过后，每年比上年减少总人口的幅度在 9.3% 以上。即使是有 60% 的生育一孩的育龄妇女终身

生育第二个孩子，那么，2100年中国总人口总量只是2010年的65%左右，比2010年减少4.6亿人，人口高峰过后，每年比上一年减少总人口的幅度在6.6‰以上。

表3 二孩生育情景预测结果

指标	二孩递进生育率 (2017—2100)	2010年	2020年	2050年	2100年
总人口(亿)	0.3582	13.33	13.72	12.20	6.58
	0.4	13.33	13.75	12.31	6.84
	0.6	13.33	13.91	13.05	8.72
	0.8	13.33	14.06	13.85	11.01
劳动年龄人口(%) (15~64)	0.3582	74.47	71.48	58.29	50.81
	0.4	74.47	71.33	58.38	51.48
	0.6	74.47	70.50	58.32	53.69
	0.8	74.47	69.75	58.28	55.23
少儿人口(%) (0~14)	0.3582	16.61	14.76	10.37	9.69
	0.4	16.61	14.93	10.72	10.05
	0.6	16.61	15.91	12.51	12.16
	0.8	16.61	16.82	14.23	14.20
老年人口(%) (65+)	0.3582	8.92	13.77	31.35	39.50
	0.4	8.92	13.74	30.90	38.48
	0.6	8.92	13.58	29.17	34.14
	0.8	8.92	13.43	27.49	30.57

其次，从人口年龄构成来看，如果二孩递进生育率一直保持在2010年的生育水平，那么，预计到2100年中国15~64岁劳动年龄人口比例将由2010年的75%左右下降到50%左右（见表3），下降幅度达25%左右。65岁及以上老年人口比例将由2010年的不到9%上升到39.50%左右，上升幅度达30%以上。与此相对应0~14岁少儿人口的比例将由2010年的16.61%下降到10%以内，下降6%以上。如果全面二孩政策实施后的二孩生育水平上升，那么，从不同二孩递进生育率水平对人口年龄结构的影响来看，二孩递进生育

率为 40%、60% 和 80% 对总人口的规模有很大影响，而对劳动年龄人口和少儿人口的比例的影响并不是非常大，比如二孩的递进生育率为 40%，预计到 2100 年劳动年龄人口比例为 51.48%，如果二孩的递进生育率为 60% 对应的劳动年龄人口比例为 53.69%，二孩递进生育率相差 20%，而劳动年龄人口比例只相差 2.21%。同样，对于不同二孩递进生育率少儿人口比例的情况与劳动年龄人口比例的情况类似，两者只相差了 2.11%。与此不同，对于老年人口的影响要更显著一些，比如二孩的生育率为 40%，预计到 2100 年老年人口比例为 38.48%，60% 所对应的老年人口比例为 34.14%，二孩递进生育率相差 20%，而老年人口比例相差 4.34%。

最后，从年龄结构整体形状来看，二孩生育水平的高低直接影响未来人口年龄结构金字塔的整体形状。如果是只有 40% 的一孩育龄妇女生育二孩，那么到 2050 年 0 岁人口规模将在 850 万人以内，不到 2010 年人口普查调查 0 岁人口的 60%，到 2100 年将继续下降到 440 万人以内，不到 2010 年人口普查调查 0 岁人口的 1/3，年龄结构的整体形状也转变为上大下小的“倒梯形”（见图 7 和图 8）。即便是二孩生育水平提高到目前看来不太可能的 80%，人口年龄结构整体形状与 40% 的情况有很大不同，但整体上大下小的“倒梯形”格局不变。

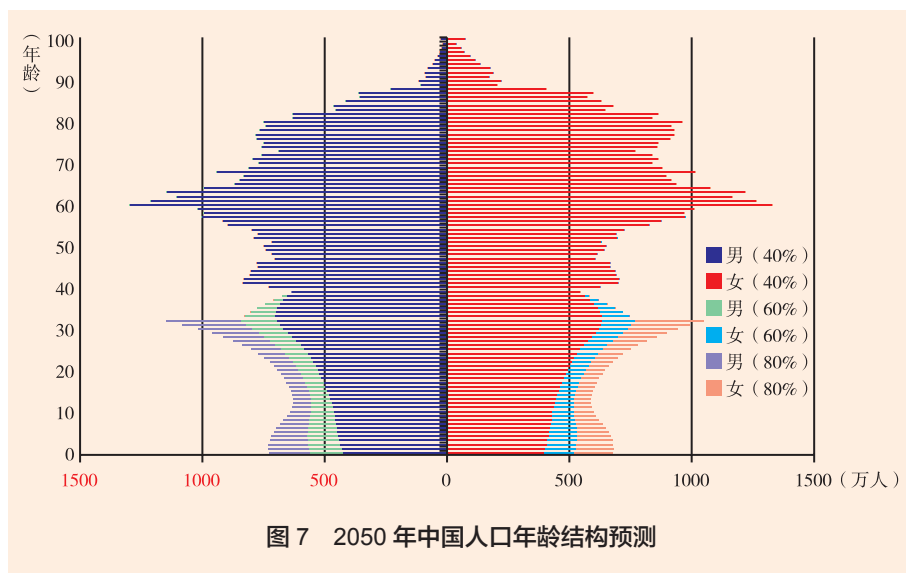
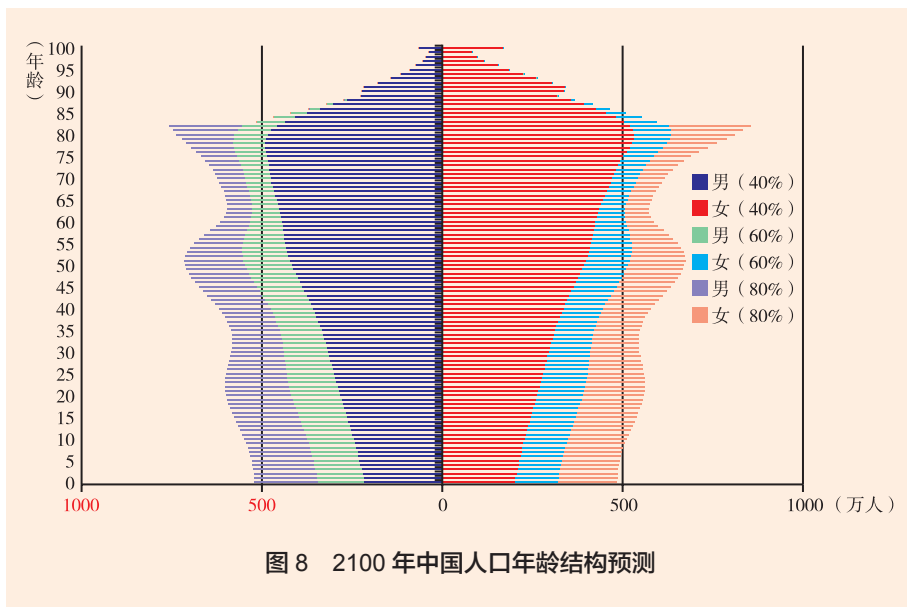


图 7 2050 年中国人口年龄结构预测



（二）人口变动过程情景分析

人口总量和结构的变动过程一方面体现了人口的基本变动特征；另一方面体现人口内在规律，特别是参数长期固定条件下，其实是体现人口系统从远离稳定状态向稳定或静止状态变化的基本规律。

1. 总人口

从以往国家人口发展战略研究提出的总和生育率稳定 1.8 左右的人口长期目标来看，并根据目前二孩生育情况和情景参数预测，中国总人口在达到 1.41 亿人左右的目标后，持续下降的趋势将是历史的必然，只是下降的幅度大小和速度快慢而已（见图 9）。根据目前生育意愿、生育计划和实际生育状况调查结果来看，未来 10 ~ 20 年内中国总人口变动过程处于 40% ~ 60% 生育二孩情景的可能性更大一些，而处于 80% 的可能性很小，因此，总人口的高峰达到或超过 14.5 亿人的可能性很小。如果考虑到目前全面二孩生育政策效果并结合发达国家的生育水平变动规律和基本情况，中国远期生育二孩的比例在 40% 以内的可能性更大，因此，总人口下降迅速的可能性是非常大的。

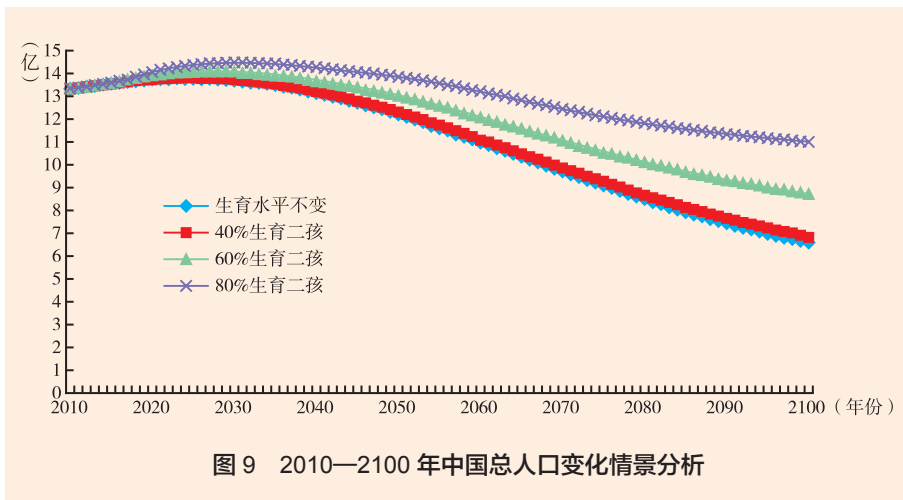


图 9 2010—2100 年中国总人口变化情景分析

2. 劳动年龄人口

无论二孩递进生育率达到多少，2050 年前 15 ~ 64 岁劳动年龄人口比例快速下降的趋势是不可避免的。无论是二孩递进生育率是 40% 还是 80%，未来 30 多年里二孩递进生育水平的不同对劳动年龄人口比例的影响很小。从图 10 可以看到，不同二孩递进生育水平条件下劳动年龄人口比例变化过程可以粗略地划分为两个明显不同的阶段，第一阶段是从 2020 年到 2055 年前后是劳动年龄人口比例迅速下降的阶段。第二阶段是从 2055 年到 2100 年是劳动年龄人口比例缓慢下降或小幅回升后稳定的阶段。由此可见，劳动年龄人口比例快速下降和远期有可能稳定的变化趋势已经很难逆转。

3. 少儿人口

0 ~ 14 岁少儿人口比例的变动趋势和变动过程与劳动年龄人口有很大差别。少儿人口比例的下降幅度或波动情况取决于到底有多大比例的育龄妇女生育二孩。从情景预测的情况来看，如果仅有 40% 的育龄妇女生育二孩，那么，少儿人口比例将迅速下降到 10% 左右，并长期稳定在 10% 的水平上。如果生育二孩的育龄妇女的比例达到 60%，那么，少儿人口比例将下降的趋势不变，只是长期稳定的水平比 40% 的提高 2% 左右，波动中稳定在 12% 左右。只有 80% 育龄妇女终身生育两个孩子才有可能实现短时间少儿人口

比例的明显提高，但长期的发展趋势仍然是下降并稳定在 14% 左右的水平上（见图 11）。

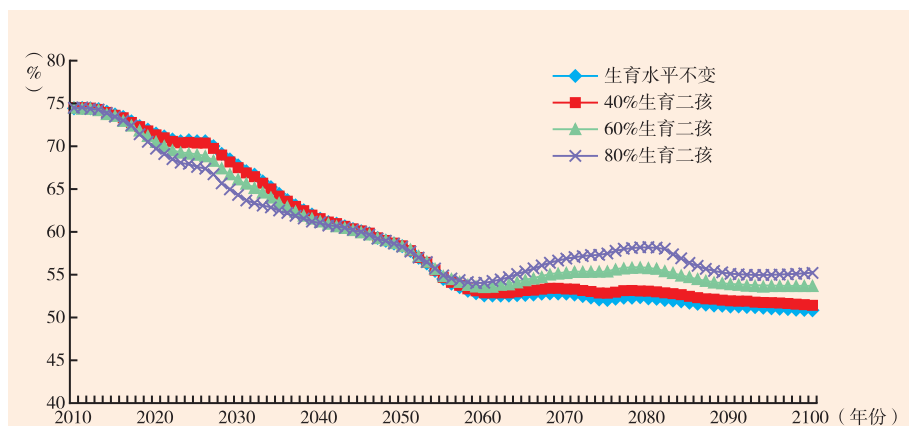


图 10 2010—2100 年中国劳动年龄人口比例变化情景分析

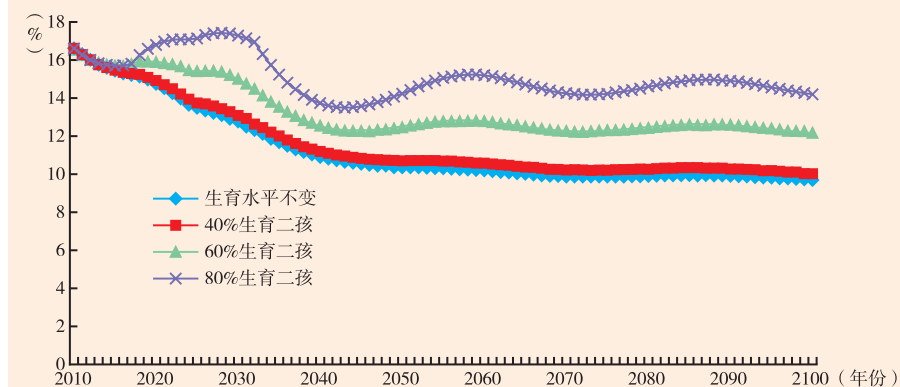


图 11 2010—2100 年中国少儿人口比例变化情景分析

4. 老年人口

老年人口比例持续、快速上升的趋势是不会发生改变的。未来老年人口比例上升的过程将可以划分为两个阶段，第一阶段是快速上升时期；第二个阶段是缓慢上升时期。第一阶段，即预计未来三十年 65 岁及以上老年人口比例迅速增长到 30% 以上，不同二孩递进生育率的差别只是达到 30% 还是 35% 左

右。第二阶段，即三十多年后，老年人口比例将长期处于一个高比例平台，下降的可能性非常小。具体来看就是，如果只有 40% 的育龄妇女生育二孩，那么，2060 年老年人口比例将由目前的 10% 左右上升到 35% 以上，此后的上升趋势不变只是速度减缓，到 2100 年老年人口比例将接近 40%。即使是有 80% 的育龄妇女生育二孩，那么，2060 年老年人口比例也将由目前的 10% 左右上升到 30% 以上，与 40% 有所不同的是此后的上升趋势将有可能稳定或波动，到 2100 年将稳定在 30% 左右（见图 12）。

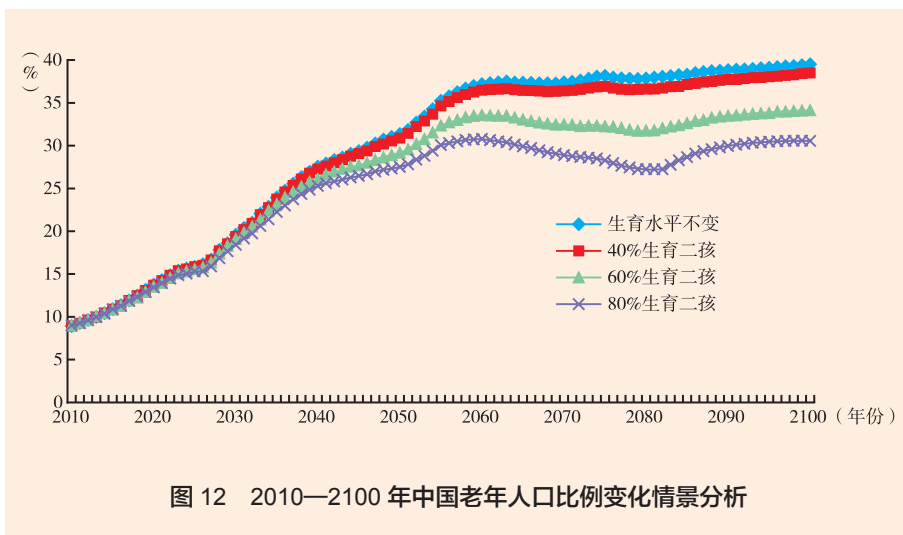


图 12 2010—2100 年中国老年人口比例变化情景分析

总之，长期低生育率带来的人口持续快速老化将对个人、家庭和社会经济带来巨大挑战，面对长期积累的人口问题及风险，生育政策、社会政策和经济政策不断调整也将成为历史的必然。

六、从人口模拟结果看未来人口政策

中国人口发展的历史决定人口发展的未来。对比人口发展情景模拟结果可以看到中国人口发展的内在规律。人口政策特别是生育政策是未来中国人口发展符合人口发展战略要求，进入良性循环轨道的辅助工具。面对中国人口未来三个方面的严峻挑战，人口政策更需要与时俱进。

第一，低生育率陷阱。通过四种二孩生育水平的情景分析可以看到，人口持续、快速下降的趋势已经形成。从目前生育意愿和生育行为调查数据结果来看，二孩递进生育水平能够稳定在 50% 左右或更低的可能性更大一些，达到 60% 或 80% 的可能性不大（王广州，2017；张丽萍、王广州，2015；王军、王广州，2016）。对照中国人口长期发展目标，即总和生育率稳定在 1.8 左右的目标，目前的生育政策调整是无法实现的。同时，当前和未来促进生育率持续下降的因素不断强化，比如，育龄妇女受教育程度不断提高，城镇人口比例持续增加和子女养育成本居高不下，其结果是生育率长期走低的风险远远大于稳定回升的可能，跨越低生育率陷阱的任务非常艰巨。

第二，劳动力持续老化与创新型国家建设。人口快速老化的前提是劳动力快速老化，劳动力快速老化的结果是严重阻碍产业或经济的转型升级。众所周知，科学技术发展核心竞争是创新，经济的快速发展动力也是创新。对绝大多数历史经验丰富但知识、技能陈旧的高龄劳动力的教育、培训是一个非常艰难的系统工程，如何面对持续、快速老化的劳动年龄人口，将是未来创新经济、创新社会和创新国家的严峻挑战。

第三，人口快速老化条件下的人口结构性失衡问题凸显。人口快速老化和老年人口比例长期居高不下，老年人口总量和比例持续增加所形成的人口结构失衡问题凸显，其结果是一方面加大养老金支付系统的压力和在职人口负担，同时长寿因素也是养老系统安全运行的不确定性风险；另一方面家庭养老支持人均照料负担前所未有，如何破解人口快速老化对养老保障和社会支持系统的持续增压，也将是全社会面临的研究课题和不可回避的突出矛盾。

总之，未来中国人口政策将是稳定、协调中国人口总量和结构问题，解决近期和长期的矛盾冲突的重大战略举措之一，也就是解决所谓的时期和队列或代内和代际人口问题重要策略之一。人口问题的解决不能寄希望于通过短期举措就能够解决长期历史积累的问题。因此，不断根据人口形势变化，前瞻性政策调整和干预才是缓解主要矛盾的必要条件。

参考文献

郭志刚：“关于生育政策调整的人口模拟方法探讨”，《中国人口科学》2004年第2期。

路磊：“已知死亡水平和死亡模式时的生命表编制”，翟振武主编，《人口数据分析方法及其应用》，外文出版社1992年。

马瀛通：《数理统计分析人口学》，中国人口出版社2010年。

张丽萍、王广州：“育龄人群二孩生育意愿与生育计划”，《人口与经济》2015年第6期。

王广州：“中国低生育水平下的二孩生育意愿研究”，《青年探索》2017年第5期。

王军、王广州：“中国低生育水平下的二孩生育意愿与生育行为差异研究”，《人口学刊》2016年第2期。

Preston, S.H., Heuveline, P. and M. Guillot, 2001, Demography-Measuring and Modeling Population Processes, Blackwell Publishers Ltd 2001.

David, P. S., 1992, Formal Demography, Plenum Press.

Siege, I J.S and D.A., Swanson, 2004, The Methods and Materials of Demography (Second Edittion), Elsevier Academic Press.

(责任编辑：春 耕)

the related theoretical study and the empirical research. Then we discuss some of the recent development and application of the DID model in the policy evaluation from three aspects: parallel trend assumption, stable unit treatment value assumption and linear hypothesis. Finally, we compare the difference between this method and other policy evaluation methods under a unified framework. The purpose of the paper is to provide some scientific suggestions for researchers who use DID model to evaluate the treatment effect of the public policy.

Key words: Public Policy Evaluation, Difference-in-Difference, Parallel Trend, Stable Unit Treatment Value Assumption, Non-linear Model

JEL: C54, C02

China's Population Forecast Methods and Future Population Policies

WANG Guangzhou (Institute of Population and Labor Economics, CASS, 10028)

Abstract: Population projection can be divided into to two different areas, one is total population projection, another is population structures projection. This paper has started from the fundamental total population projection model of Gompertz and logistic model and introduced the classical cohort-component and parity progression model. And then the reverse deduction method of the parity progression fertility ratio has been discussed for the parity progression population projection model. Taking the census data of 1990, 2000 and 2010, the reliability of the parity progression model for the population structure projection has been double-checked. Based on the census 2010 of China and parity progression method, this paper has made several simulation scenarios and pointed out that under the strengthen condition of inhibition fertility, the population strategy target of stabilization the fertility level at 1.8 can not be reached. The long-term fast population aging cannot be avoided. The low fertility trap, fast aging for labor force and structure imbalanced problems will be seriously and urgent in the future.

Key words: Population Projection, Cohort-component Model, Parity Progression Model, Parity Progression Fertility

JEL: C54, J24