

“四二一”家庭微观仿真模型与应用

齐险峰 郭震威

【内容摘要】 本文将“四二一”家庭和广义“四二一”家庭区分开来,并基于 2000 年第五次全国人口普查 1‰ 抽样个案数据,建立了直接测算未来“四二一”家庭和广义“四二一”家庭数量的微观仿真模型。仿真结果显示,现行生育政策保持不变时,未来全国“四二一”家庭及广义“四二一”家庭数量并不多,且主要集中在非农业人口。

关键词: 四二一家庭;独生子女

【作者简介】 齐险峰,清华大学经管学院;郭震威,国家人口计生委发展规划司。

1 问题的提出

近年来,随着实行计划生育基本国策以来产生的独生子女开始进入婚育期,我国“四二一”家庭结构问题引起了学术界和实际工作部门的高度关注。其实,“四二一”家庭在每个历史时期都存在,目前之所以备受关注,主要是因为我国的计划生育政策形成了历史上从未有过的大量独生子女,中国的独生子女成为与计划生育政策有关的一代特殊的人口,独生子女户也成为符合政策生育或响应政策号召生育的一类特别的家庭(宋健,2005)。正是由于中国独生子女的特殊产生背景,人们对这一群体及其相应家庭的关注一直没有间断过,研究的领域涉及到教育学、心理学、社会学、人口学、体育学等,研究的主题则包括独生子女的个性特征、人格特点、社会交往、社会化发展、家庭教育、与非独生子女的差异、独生子女家庭等等(风笑天,2002)。但是,“实际上关于独生子女数量的信息却极为贫乏,也因此相关这方面的宏观定量分析也是凤毛麟角。这与众多的定性研究形成了强烈的反差”(郭志刚,2001)。这些为数不多的定量研究,也仅仅局限于推算独生子女与独生子女结婚的概率。2000 年,杨书章、郭震威比较系统研究了中国独生子女现状、“双独”婚配概率及其对未来人口发展的影响。当前这方面最为深入的研究应该是郭志刚、刘金塘、宋健 2002 年的研究,他们通过取消“四二一”家庭定义中的一些限制条件,将三代人简化为二代人,测算了“X 二”家庭(即第三代和第一代存活人数未知)出现的概率。同时对三代人进行跟踪分析,属于完全意义上的“四二一”家庭模型尚未见诸已出版的文献资料。

对“四二一”家庭的定量研究,存在两个方面的难题。

第一个难题是概念问题。有关专家普遍认为“四二一”结构涉及三代人,但对具体含义方面尚存在两种不同的观点(宋健,2000):一种观点把独生子女夫妇作为家庭结构的中心,即“四二一”中的中间代(“二”)来看待,重点讨论独生子女一代及其所产生的社会问题。如郭志刚(1995)认为,独生子女之间结婚,就会形成典型的“四二一”家庭。另一种观点把独生子女作为“四二一”中的第三代(“一”)来看待。如丁士贤(1989)将“四二一”家庭结构定义为祖父母一代的 4 个老人、2 个父母、1 个独生子女。

本文采用上述第一种观点,即把“四二一”中的中间代(“二”)视为一对独生子女夫妇。本文研究的对象主要是实行计划生育政策以来尤其是 20 世纪 80 年代生育政策收紧后,产生的独生子女这一代人。其祖父母辈大都生育了多个子女,无论在城市还是在农村,老有所养的问题不会普遍存在。独

生子女这一代人没有兄弟姐妹,既要赡养父母一代,又要抚养下一代,其抚养负担相对而言就比较沉重,这正是本文关注的重点问题。据此,严格意义上的“四二一”家庭涉及3代人、3对夫妇、共计7个人,他们只需要具有这种直接亲属关系结构,而不需要居住在一个家庭,其形成的4个必要条件是(郭志刚、刘金塘、宋健,2002):“(1)中间一代为性别不同的两个独生子女;(2)这两个独生子女之间存在婚姻关系;(3)这两个独生子女结婚后只生育一个子女;(4)独生子女夫妇双方的父母仍然同时生存”。本文将严格符合上述4个条件者,定义为“狭义‘四二一’家庭”。与这种严格定义相联系,根据独生子女夫妇双方及其父母的存活状况,可以定义“三二一”家庭、“二二一”家庭、“一二一”家庭、“四一一”家庭、“四〇一”家庭、“三一—”家庭、“三〇一”家庭、“二一一”家庭、“二〇一”家庭、“一一一”家庭、“一〇一”家庭等11种家庭结构。同样重要的是,根据现行的生育政策规定,独生子女结合可以生育两个孩子。因此,我们还必须关注“四二二”家庭结构问题,其中第二代独生子女抚养负担比严格的“四二一”家庭还要沉重。因此,本文将由独生子女夫妇及其双方父母、子女共同组成的群体组合,统一定义为“广义‘四二一’家庭”。由于在某个时点上,一对独生子女夫妇可能存活0~2人、双方父母可能有1~4人、子女可有1~2人(我们不研究双方父母均已死亡或没有存活子女的情况,此时,事实上只存在相邻的两代人),广义“四二一”家庭包括了24种组合, $[4(\text{父母人数}) \times 3(\text{独生子女夫妇人数}) \times 2(\text{子女人数})]$ 。本文同时研究狭义“四二一”家庭结构和广义“四二一”家庭结构。为与通常的研究相一致,为便于理解,本文凡不特别说明的地方,均指狭义“四二一”家庭结构。

第二个难题是方法问题。分析、预测未来我国“四二一”家庭数量变动情况,属于人口预测的范畴。人口预测方法通常有人口宏观预测和人口微观预测两类方法。分要素预测方法作为人口宏观仿真方法的主流技术,已经在我国人口研究中得到广泛深入的应用。人口宏观仿真中,需要把总体按照不同的属性组合,细分为不同的子人口。如果有 k 个属性,取值个数分别为 $M_i(i=1, \dots, k)$,那么,在宏观层面,状态空间包含 $(M_1 \times M_2 \times \dots \times M_k)$ 个单元格,也就是说,必须构造如此大规模的矩阵才能完整地描述总体的特征。“四二一”家庭是涉及三代人的复杂组合,如果使用人口宏观仿真方法来进行分析预测,至少需要按照个体的以下属性将总人口划分为若干子人口:

(1)年龄。取值可为0, 1, ..., 100+岁,共101类;(2)性别。分男性、女性,共2类;(3)孩次。仅对女性人口有效,取值可为0~12之间的整数,共13类;(4)独生子女属性。每个人可能是独生子女,也可能是非独生子女,还可能不可辨识。共分3类;(5)婚姻状况。只考虑未婚、已婚两种情况,共2类;(6)婚姻属性。仅对已婚人口有效,分为双独、单独、双非、不可辨识等4类;(7)户口性质。分农业户口、非农业户口,共2类;(8)第一代年龄。取值可为0, 1, ..., 100+岁,共101类;(9)第二代年龄。取值可为0, 1, ..., 100+岁,共101类;(10)第三代年龄。取值可为0, 1, ..., 100+岁,共101类;(11)第四代年龄。取值可为0, 1, ..., 100+岁,共101类;(12)第五代年龄。取值可为0, 1, ..., 100+岁,共101类;

那么,人口宏观仿真模型中,需要构造的矩阵规模至少应达:(1)对男性人口: $101 \times 3 \times 2 \times 4 \times 2 \times 101 \times 101 \times 101 \times 101 \times 101 = 50,952,967,228,848$;(2)对女性人口: $101 \times 13 \times 3 \times 2 \times 4 \times 2 \times 101 \times 101 \times 101 \times 101 = 662,388,573,975,024$

如此庞大的状态空间,显然已经超过人口宏观模型所能处理的能力。因此,运用人口宏观仿真方法,难以处理这类“四二一”家庭的建模问题。

本文将从定量分析的角度,对今后一个时期我国“四二一”家庭结构进行较为系统、深入的探索与研究,具有一定的学术价值和现实意义。一方面,本研究力图建立适合我国特点和数据情况的微观人口仿真模型,探索从微观基础出发、揭示宏观人口现象的定量研究方法;另一方面,本研究试图通过测算未来可能出现的“四二一”家庭数量,可望为生育政策、养老政策等公共政策的调整与完善提供更为

科学、合理的依据。

2 “四二一”家庭微观仿真模型的建立

微观仿真(microsimulation)模型的概念最早是由美国耶鲁大学 Orcutt 教授于 1957 年提出的,并在 1961 年与他人合作初步实现了这一模型。微观仿真,不是直接将目标系统的总体作为研究对象,而是从一种微观的角度出发,模拟组成目标系统的各个微观个体的行为和它们之间的相互作用过程,并对其行为产生的结果进行加总而得到所需要的宏观水平的变量。它特别适合于研究组成系统的微观个体是决策的主体并且个体之间的交互作用非常复杂因而难以通过其他的数学方法获得确定解的系统。这种方法在社会经济系统的解释方面也具有独到的优势,通过相对简单的微观个体行为,突现出宏观层面的复杂行为,以此理解社会和经济过程,更符合经济社会系统的内在本质。40 多年来,尤其是 20 世纪 90 年代以来,微观仿真模型在各主要西方发达国家得到不断完善与发展(Zaidi & Rake, 2001),在人口、税收、社会福利、卫生、教育、分配等方面发挥着越来越重要的作用。但是,正如曾毅(2004)指出:“〔微观仿真〕模型非常复杂,计算量极大。……使不少研究者特别是发展中国家的研究者望而却步”。

人口微观仿真是微观仿真方法的一个特例,是微观仿真方法在人口预测领域的具体应用。其基本思路是从人口总体中抽取一定比例的代表样本,根据事件按一定概率分布随即发生的原理,利用计算机仿真技术对抽取样本中的每个人的生育、死亡、婚姻、迁移、家庭状态变化等的复杂过程一一进行模拟,然后予以汇总得出人口和家庭的整体特征和分布(Keilman, 1988; Reeves, 1994; Smith, 1994; von Imhoff & Post, 1998; Zaidi & Rake, 2001)。人口微观仿真能较好地考虑个体间的交互作用。在宏观模型中,个体之间的联系信息通常不被考虑。而在微观模型中,只需简单地在数据记录中增加一个指针,就很容易建立个体之间的联系。为此,本研究决定采用人口微观仿真方法,来预测未来“四二一”家庭的有关情况。基本模型框架见图 1。一个好的模型,应该是在能满足研究目的的前提下,最为简化(即涉及非相关变量和数据最少)、最为经济(即建模工作量和数据运算量最少)、具有扩展性(即为未来模型扩展和解决更为复杂的问题预留充分的弹性)的模型。本文就试图建立一个符合上述标准的模型。现将所建立的“四二一”家庭微观仿真模型的有关问题说明如下:

(1) 预测起始年份(基年)设定为 2000 年。我国政府于 2000 年进行了第五次全国人口普查,国家统计局于 2002 年向有关政府部门和科研单位提供了 1% 的随机抽样数据。这是本研究所能得到的、能够公开使用的权威数据。通过 10 重匹配,将普查登记中以户主为纽带建立的亲属关系结构,改变为以两两个人之间夫妇、父(母)子(女)为纽带建立的亲属关系结构,以此作为微观仿真的起点数据。尽管普查登记数据并没有包含所有个人之间的全部亲属关系,但是,仍然可以辨识出上世纪 80 年代我国提倡独生子女政策以来大部分新增人口的独生子女属性以及 1971~1980 年出生队列中一半左

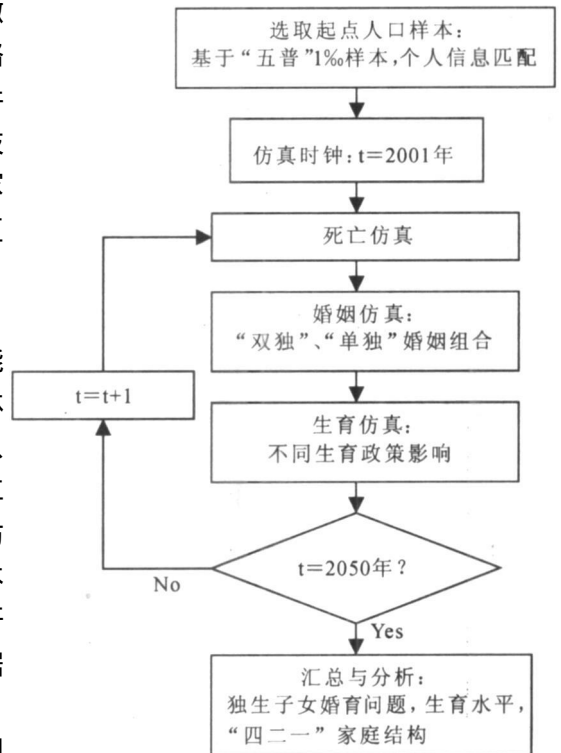


图 1 “四二一”家庭仿真模型框架

右的独生子女,可以在相当程度上反映全面推行计划生育基本国策对独生子女数量和“四二一”家庭结构的影响。本研究以此为基年,依次向后进行人口仿真。

(2) 预测终止年份设定为 2050 年。本研究将预计出 2001~ 2050 年各年度的人口变动尤其是“四二一”家庭数量变动情况。向后预测 50 年,基本上能够满足政策仿真的目的。

(3) 该模型将全国人口视为封闭人口。与全国 13 亿人口相比,我国每年的国际迁移人口量是微不足道的,完全可以忽略不计。我国现有人口预测方案,几乎都做了与此相同的假定。在封闭人口的情况下,我们需要处理好“两性问题”之类的一致性问题。

(4) 该模型属于离散仿真。仿真时钟以日历年度为单位向前递进。这与当前我国各种人口宏观预测方案高度一致,也能满足实际人口规划工作的具体要求。

(5) 该模型内部的变量均为纯人口变量。该模型采用了与一般总和生育率等分要素人口预测(即传统人口宏观仿真)方法相同的策略。模型内部不考虑教育程度、经济收入等经济社会变量,而只涉及人口婚姻、出生、死亡以及由此导致的人口自然增长等变量。婚姻水平、生育水平、死亡水平被设定为外生变量(预测参数),由外部经济社会因素所决定,因此,没有被内化在该模型之中。

(6) 关于该模型的状态空间。人口微观仿真,要求详细跟踪、仿真、记录每个人在生命周期过程中发生的每一个事件(仅限于研究相关事件),因此,在模型中必须对每个人赋予一个唯一的、终身不变的、不重复使用的标识号码(相当于现实生活中的公民身份号码或某些西方国家的社会保障号码),并记录其相关属性值。本研究所建立的模型,除了记录其唯一的标识号码(在仿真过程中用此号码代表这个人)、出生年份、性别、初婚年份、婚姻属性、现存死亡年份等必需的数据项以外,还记录了其所在的省份、民族、农转非(由农业人口转变为非农业人口)年份等信息,这就保证了模型较强的可扩展性,为将来考虑各省人口及政策差异、民族差异、农村人口迁入城镇等问题,留下了必要的扩展空间。

研究“四二一”家庭结构,在人口微观仿真模型中的处理十分方便。只需在个人信息中增加父亲、母亲、配偶、子女的标识号码,就意味着业已建立起夫妇之间、上代人与下代人之间的链接关系乃至所有的家族成员(亲属)关系网络(示意图见图 2,以男性“个人”为例)。由此,就可以从一个人的属性向量出发,很方便地找到其父母、配偶、子女的属性向量;通过查询其父母的子女信息,就可以判断其兄弟姐妹数量及其独生子女属性;通过查询其配偶的父母信息,就可以判断其岳父母(或公公婆婆)的属性;等等。

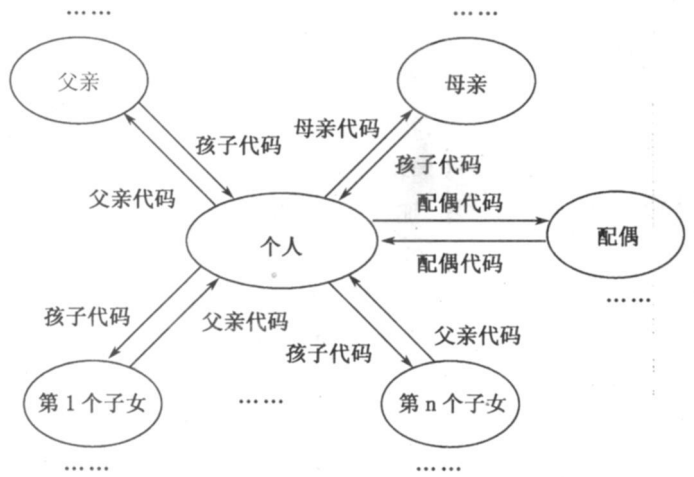


图 2 家族成员(亲属)的链接关系

本文所建立的“四二一”家庭微观仿真模型中,每个人的属性共有 26 项。基年(2000 年)的人数为 1,180,116 人,即使未来 50 年新增 1 倍人口,模型的状态空间最大也不会超过 $1,180,116 \times 2 \times 26 = 61,366,032$,远远低于对应的宏观仿真模型的状态空间数。

(7) 该模型考虑了农业人口和非农业人口的差异,但没有考虑城乡人口迁移。我国的经济社会呈现典型的城乡二元分立结构,农业人口与非农业人口在生育政策、经济社会政策、生活质量、发展机会等方面存在巨大的差别,为此,我们根据人口的户口性质,在模型中将农业人口和非农业人口区分开

来,分别同时进行仿真。当前,我国正在经历着世界上规模最大的城乡人口迁移,每年从农村到城镇净流入 1000 万人以上,且双向流动的趋势十分明显,对经济社会发展已经产生并将继续产生巨大、深远的影响。但是,人口迁移本质上属于经济社会现象,由于在社会转型时期,城乡人口迁移模式变化较快,且相关数据难以获得;同时,当一个人从农村转移到城镇时,其对父母和子女的抚养关系并不随之改变,“四二一”家庭结构也不随之改变;综合考虑这两个因素,我们在建模时暂未考虑农业人口与非农业人口的相互转换问题。

(8) 该模型使用面向个体行为(事件)的模拟方式。在每一个仿真年度,严格按照“死亡 初婚 生育 1 孩 生育 2 孩 生育 3 孩 生育 4 孩 生育 5 孩”的事件发生顺序,对每一个人可能经历的人口事件进行仿真;直到所有人的所有可能经历的事件都仿真完毕,然后再将仿真时钟递进到下一年,再依次按照上述人口事件发生的逻辑次序,对所有个人进行新一轮的仿真。为简单起见,对婚姻的仿真,只考虑初婚、丧偶的情况,暂未考虑离婚、再婚、复婚等情况;也不考虑初婚后当年生育 1 孩、生育 i 孩后当年生育 $(i+1)$ 孩、双胞胎或多胎生育等情况。同时,从我国的现实情况出发,可以合理地假定婚前生育可以忽略不计,生育现象只发生在结婚之后(此时,婚前生育也可理解为事实婚姻条件下的生育)。

计算任一预测年份(2001~2050年)的分年龄初婚概率,本研究使用以女性人口为主导的总和递进初婚率模型。初婚仿真中还有 4 个假定:一是城乡不通婚,婚姻只发生在具有相同户口性质的人群之中。这个假定既是为了计算的方便,也是基本符合实际情况的。二是适婚人口在选择配偶时不特意追求对方的独生子女属性,即独生子女与独生子女之间、独生子女与非独生子女之间的婚姻组合完全是随机的。目前还没有令人信服的统计资料表明适婚男女在选择配偶时在对方是否独生子女。三是考虑我国妇女终身不婚的比例很低,且多数在育龄期结婚,因此,我们有理由将女性结婚年龄限制在 15~49 岁之间。四是夫妇婚龄差不超过 10 岁。同第一个假定一样,做出这个假定既是为了计算的方便,也是基本符合实际情况的。

在预测年份 t , 初婚事件的仿真过程大致如下:

顺序搜索样本,找出第一个适婚年龄(15~49岁)未婚存活妇女;

进行蒙特卡罗随机试验,产生一个(0, 1)区间均匀分布的随机数。如果这个随机数小于该年龄妇女的递进初婚概率,那么这名妇女被仿真为 t 年应该初婚。但这并不意味着这名妇女就能够真正在 t 年初婚,只有存在符合条件的男性人口时,这名妇女才能被真正仿真为在 t 年初婚。

按照夫妇年龄差分布,查找并定位该妇女的丈夫(解决“两性”问题)。这又分为以下几个步骤:

第一步,进行蒙特卡罗随机试验,产生一个 $[-10, 10]$ 区间均匀分布的整数型随机数,作为男女双方的年龄差。

第二步,将妇女年龄和男女双方的年龄差相加,得到“候选”丈夫的年龄。

第三步,将所有符合条件的男性人口都找出来,均作为该妇女的配偶候选人。这里共有 5 个条件:性别为男性、 t 年仍存活、 t 年婚姻状态为未婚、 t 年年龄等于第二步中求得的候选丈夫年龄、户口与该妇女相同。

记该妇女的配偶候选人共有 n 人。如果经过查找,发现 $n=0$,即没有符合条件的候选丈夫,那么,返回第一步计算下一个年龄差,重新查找与该年龄差相对应的配偶候选人。如果将所有年龄差(21个)均搜索完毕,仍然没有找到符合条件的配偶候选人,那么,只能假定这个妇女在 t 年找不到配偶。此时,我们只能继续搜索样本,找出下一个适婚年龄(15~49岁)未婚存活妇女,重复进行 及以下的

各个仿真过程。

第四步,进行蒙特卡罗随机试验,产生一个 $[1, n]$ 区间均匀分布的整数型随机数 r ,将全部候选丈夫(共 n 人)中的第 r 个人作为该妇女的实际丈夫。

修改该新婚夫妇的属性值。主要有:

第一步,将夫妇初婚年份修改为 t 年(初婚年份为 -1 时表示未婚)。

第二步,将双方的“配偶标识号码”分别修改为对方的标识号码。

第三步,将双方的现有子女数设定为 0 ,为仿真其生育行为作准备。

推算并设定该夫妇的婚姻属性。

至此,对该妇女的婚姻仿真过程结束。然后继续搜索样本,找出下一个适婚年龄(15~49岁)未婚存活妇女,重复上述~过程,直到每一个妇女的婚姻事件均被仿真一遍。

下面简单介绍一下生育仿真。本研究采用了女性人口主导的年龄-孩次递进预测模型。生育仿真也比较复杂,既涉及对妇女本人生育情况的模拟,又要模拟其生育子女的性别及存活状况,还要考虑生育政策的影响。在每一预测年度,对每个妇女的生育行为进行仿真,需要做3次蒙特卡罗随机试验。具体过程从略。

(9)该模型大量使用了事件之间相互独立的假定。假定死亡概率与婚姻状况、生育状况等因素无关,仅仅是年龄、性别的函数(当然,也要受预测年份和户口性质的影响,这是通过给定外生变量——平均期望寿命来体现的)。假定生育概率与结婚时间无关,已婚女性人口的生育概率仅仅是年龄、现有孩子数(孩次)、政策因素(如“单独”、“双独”等)的函数(当然,也要受经济社会因素的影响,这是通过给定外生变量——预测年度的生育水平来体现的)。应该说明的是,这些假定符合人口的宏观统计行为,也是人口宏观仿真方法普遍采用的假设条件,据此得出的预测数据是比较合理、可信的。

(10)关于仿真参数的选择。综合各方面的数据分析,当前妇女实际生育水平应在 1.8 左右(蒋正华,2006),据此设定仿真时期农业和非农业妇女的总和递进生育率:对政策不允许生育两个孩子的妇女,设为 2.08 和 1.19 ;对政策允许生育两个孩子的妇女,设为 2.27 和 1.98 。假定在仿真时期农业和非农业妇女的总和递进初婚率、夫妇婚龄差分布、出生性别比等参数,均与2000年相同;2050年男、女性农业人口的平均期望寿命为 76.27 岁和 80.99 岁,非农业人口的数值为 77.88 岁和 82.59 岁。

(11)通过模型的重复运行,推算各指标的平均值和估计误差。选用不同的独立随机数序列,采用相同的参数和初始条件,重复运行仿真模型。由于每次运行是相对独立的,因此可以认为每次仿真运行结果 X_1, X_2, \dots, X_n 是独立且同分布的随机变量。仿真结果 X 在置信区间 $1-\alpha$ 下的估计值为:

$$\mu = \bar{X} \pm t_{n-1, \frac{\alpha}{2}} \sqrt{\frac{S^2}{n}}$$

式中:

$$\text{样本均值 } \bar{X} = \frac{1}{n} \sum_{i=1}^n X_i; \text{ 样本方差 } S^2 = \frac{1}{n-1} \sum_{i=1}^n (\bar{X} - X_i)^2$$

根据中心极限定理,若产生的样本点 X_i 越多,即仿真运行的次数越多,则 \bar{X} 越接近于正态分布,因此,重复运行次数 n 不能选取得太小。

3 仿真结果的分析

我们经过半年的努力,基于Visual C++平台实现了上述模型的计算机化,并将模型运行了20

这里隐含的假定是,夫妇婚后保持其结婚后的婚姻属性(双独、单独、双非或不可辨识)不变。严格地讲,该夫妇的兄弟姐妹数量还可能继续变化,从而影响其婚姻属性。但是,实际政策操作中,即使该夫妇双方或一方的兄弟姐妹随后死亡而变成事实上的独生子女,也不会将该夫妇视为双独(或单独)婚姻;同时,该夫妇双方的父母继续生育的比例很小。从这些方面考虑,本研究作出上述假定,是符合实际的。

次。在不运行其他应用软件的情况下,该模型在 IBM ThinkPad T 43 笔记本上运行一次需要 8 小时左右的时间。

为了检验微观仿真模型的有效性,我们将仿真结果与相应的人口宏观仿真模型进行了比较。国家人口发展战略研究设立了未来人口发展预测的子课题(郭志刚,即将出版),对未来人口发展进行了 38 个方案的预测,其中的第 13 套方案(编号为 PDP_0BH)与本研究具有较强的可比性。比较分析表明,无论是人口总量指标还是人口结构指标,本研究所建立的人口微观仿真模型均能给出较为理想的预测结果(图 3、图 4)。出生人口预测值,2006~2020 年平均相差 4.90%,2021~2050 年平均相差 0.03%。每年总人口预测值相差不到 0.7%,微观仿真模型测算的人口峰值年份为 2034 年、峰值人口为 14.645 亿人,宏观仿真模型给出的人口峰值年份为 2032 年、峰值人口为 14.629 亿人。15~64 岁劳动年龄人口比例的峰值年份,微观、宏观仿真结果分别为 2008 年和 2009 年,尤其是本世纪 30 年代以后,两种方法关于人口年龄结构的预测结果基本相同。

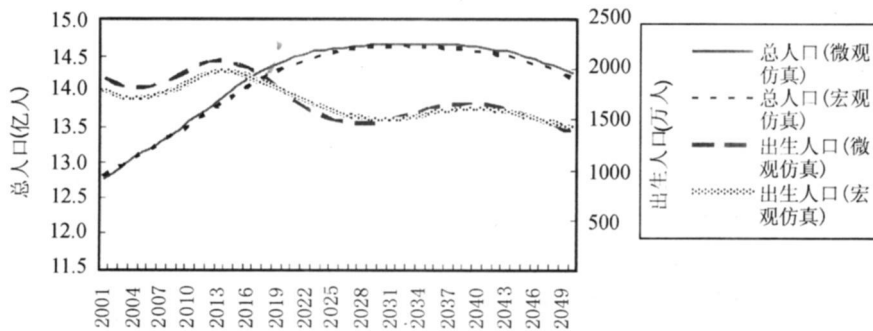


图3 微观仿真与宏观仿真结果比较:总人口与出生人口

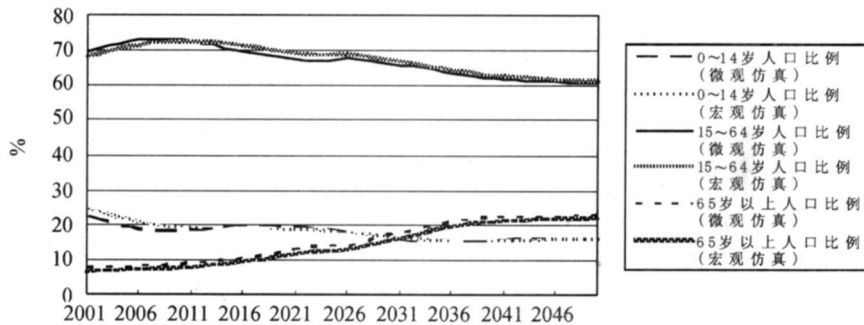


图4 微观仿真与宏观仿真结果比较:人口年龄结构构成(%)

以下是“四二一”家庭微观仿真的主要结果(详见下表):

(1) “四二一”家庭及广义“四二一”家庭数量的预测值估计误差很小。各仿真年度农业人口、非农业人口和全国人口“四二一”家庭及广义“四二一”家庭数量的预测值的估计误差在±6万之间。

(2) “四二一”家庭及广义“四二一”家庭数量不会无限增长下去。“四二一”家庭及广义“四二一”家庭数量在很大程度上取决于每年新婚夫妇双独数量的累计数。仿真结果表明,年新婚夫妇双独数量会将在2020年代达峰值8%左右,这也是广义“四二一”家庭比例的上限值。年新婚夫妇双独数量增加时,“四二一”家庭数量会以正的加速度递增。年新婚夫妇双独数量达到高峰值后下降时,在其他条件相同时,“四二一”家庭及广义“四二一”家庭数量增加的速度也会逐步下降。同时,随着时间的推移,第一代(即“四”所对应的那代人)逐步死亡,这将使“四二一”家庭及广义“四二一”家庭数量趋于下降(假定其他条件相同)。这两种因素共同作用,将最终导致“四二一”家庭及广义“四二一”家庭数量在某一时期到达峰值后逐步下降。当然,“四二一”家庭及广义“四二一”家庭数量的峰值年份将滞

后于年新婚夫妇双独数量的峰值年份。

(3)从全国范围看,“四二一”家庭及广义“四二一”家庭数量并不多。2020年分别为152万个左右和550万个以内,峰值年份分别为2036年和2047年,峰值分别为188万个和1672万个。广义“四二一”家庭数量远远多于“四二一”家庭数量。妇女的总和生育率,2033~2037年间达峰值2.01左右,比2005年提高0.17。总人口将于2020年达14.4亿人,2034年达峰值14.6亿人。

表1 未来“四二一”家庭和广义“四二一”家庭数量的变动预测

(万个)

年份	“四二一”家庭			广义“四二一”家庭		
	全国	农业人口	非农业人口	全国	农业人口	非农业人口
2011	56.2 ± 1.1	6.5 ± 0.4	49.7 ± 1.1	124.7 ± 1.3	15.6 ± 0.5	109.1 ± 1.2
2012	69.0 ± 1.2	8.2 ± 0.5	60.8 ± 1.2	160.2 ± 1.6	20.1 ± 0.7	140.0 ± 1.5
2013	82.6 ± 1.5	10.2 ± 0.6	72.5 ± 1.5	201.1 ± 2.1	25.5 ± 0.8	175.6 ± 1.9
2014	96.4 ± 1.1	11.8 ± 0.6	84.5 ± 1.2	244.7 ± 2.0	30.8 ± 0.8	214.0 ± 1.8
2015	108.3 ± 1.3	13.9 ± 0.7	94.4 ± 1.2	291.1 ± 2.0	37.0 ± 0.9	254.0 ± 1.8
2016	120.0 ± 1.3	16.1 ± 0.6	103.9 ± 1.3	339.5 ± 2.1	44.0 ± 1.1	295.4 ± 2.0
2017	129.9 ± 1.3	18.2 ± 0.6	111.7 ± 1.1	390.3 ± 2.2	51.3 ± 1.0	339.0 ± 1.9
2018	138.5 ± 1.3	20.2 ± 0.7	118.3 ± 1.3	441.7 ± 2.7	58.8 ± 1.2	382.9 ± 2.2
2019	146.0 ± 1.7	21.6 ± 0.7	124.5 ± 1.8	494.1 ± 2.6	66.1 ± 1.3	428.0 ± 2.2
2020	151.8 ± 1.5	22.9 ± 0.8	128.9 ± 1.8	546.3 ± 3.1	73.4 ± 1.3	473.0 ± 2.6
2021	155.0 ± 1.8	24.0 ± 0.8	130.9 ± 2.1	597.7 ± 3.4	80.6 ± 1.4	517.1 ± 3.0
2022	157.7 ± 1.5	24.8 ± 1.0	132.9 ± 1.8	649.2 ± 3.6	87.7 ± 1.4	561.5 ± 3.2
2023	158.8 ± 1.6	25.4 ± 0.9	133.4 ± 1.6	699.9 ± 3.8	94.7 ± 1.5	605.2 ± 3.2
2024	158.9 ± 1.7	25.3 ± 0.7	133.5 ± 1.6	750.6 ± 3.8	101.0 ± 1.5	649.6 ± 3.4
2025	158.7 ± 2.2	24.9 ± 0.8	133.8 ± 2.0	801.0 ± 4.1	106.7 ± 1.6	694.2 ± 3.7
2026	157.3 ± 1.7	24.4 ± 0.8	133.0 ± 1.6	852.2 ± 3.9	112.4 ± 1.6	739.8 ± 3.6
2027	158.4 ± 1.5	24.1 ± 0.9	134.2 ± 1.4	905.6 ± 3.5	118.1 ± 1.5	787.5 ± 3.2
2028	159.5 ± 1.8	24.0 ± 1.1	135.5 ± 1.6	959.3 ± 3.7	123.4 ± 1.7	835.9 ± 3.3
2029	162.8 ± 2.0	24.0 ± 1.2	138.8 ± 1.7	1013.8 ± 3.9	128.0 ± 1.9	885.8 ± 3.3
2030	166.8 ± 2.1	24.2 ± 0.9	142.5 ± 2.0	1070.6 ± 4.0	132.9 ± 1.8	937.7 ± 3.4
2031	170.9 ± 1.4	24.3 ± 0.8	146.6 ± 1.6	1127.3 ± 3.9	137.8 ± 1.8	989.4 ± 3.4
2032	175.2 ± 1.9	24.8 ± 0.5	150.4 ± 1.9	1183.7 ± 4.2	143.1 ± 1.8	1040.6 ± 3.7
2033	179.9 ± 2.3	26.1 ± 0.6	153.8 ± 2.1	1240.1 ± 4.9	149.5 ± 1.8	1090.6 ± 4.2
2034	183.7 ± 2.4	27.4 ± 0.6	156.2 ± 2.3	1295.1 ± 4.6	155.8 ± 1.7	1139.3 ± 4.2
2035	185.3 ± 2.5	28.7 ± 0.7	156.6 ± 2.1	1347.1 ± 4.8	162.2 ± 1.8	1184.9 ± 4.6
2036	187.8 ± 2.4	31.2 ± 0.7	156.6 ± 2.1	1398.3 ± 5.2	169.9 ± 1.9	1228.5 ± 4.8
2037	186.7 ± 2.3	32.7 ± 0.7	154.0 ± 2.1	1444.8 ± 5.7	176.9 ± 2.1	1267.9 ± 5.1
2038	183.8 ± 1.9	34.3 ± 0.8	149.5 ± 1.9	1488.4 ± 5.4	184.1 ± 2.0	1304.3 ± 4.6
2039	179.5 ± 1.9	35.3 ± 0.8	144.1 ± 1.7	1528.4 ± 5.0	191.2 ± 2.0	1337.2 ± 4.2
2040	173.1 ± 1.3	35.7 ± 0.8	137.4 ± 1.2	1562.8 ± 5.3	197.4 ± 2.1	1365.4 ± 4.5
2041	165.7 ± 1.5	35.8 ± 0.7	129.9 ± 1.4	1592.4 ± 5.3	203.2 ± 1.9	1389.2 ± 4.6
2042	157.8 ± 1.4	35.7 ± 0.7	122.0 ± 1.4	1617.6 ± 5.5	208.8 ± 2.0	1408.8 ± 4.7
2043	150.2 ± 1.6	35.1 ± 0.7	115.1 ± 1.5	1638.5 ± 5.5	213.5 ± 2.1	1425.0 ± 4.6
2044	140.8 ± 1.6	34.0 ± 0.8	106.8 ± 1.5	1653.9 ± 5.5	217.9 ± 2.1	1436.0 ± 4.6
2045	131.9 ± 1.7	32.9 ± 0.8	99.0 ± 1.6	1664.5 ± 5.5	221.2 ± 1.9	1443.3 ± 4.8
2046	123.7 ± 1.9	31.3 ± 1.0	92.4 ± 1.6	1670.8 ± 5.6	224.2 ± 2.0	1446.6 ± 4.8
2047	114.8 ± 1.7	30.3 ± 0.9	84.6 ± 1.2	1671.8 ± 5.4	226.7 ± 2.0	1445.1 ± 4.6
2048	106.7 ± 1.4	29.1 ± 0.8	77.6 ± 0.9	1668.5 ± 5.6	229.1 ± 2.1	1439.5 ± 4.7
2049	98.7 ± 1.5	28.2 ± 0.9	70.6 ± 1.0	1660.5 ± 5.6	230.8 ± 2.2	1429.7 ± 4.5
2050	92.5 ± 1.4	27.4 ± 0.9	65.1 ± 1.1	1649.4 ± 6.0	232.7 ± 2.3	1416.7 ± 4.6

(4)“四二一”家庭主要集中在非农业人口。非农业人口的年新婚夫妇双独比例的峰值年份出现

在2029年前后,峰值高达40%以上,2050年仍在14%左右;妇女的总和生育率2031年前后达峰值1.51左右,比2005年提高0.3;“四二一”家庭及广义“四二一”家庭数量,2020年分别为129万个和473万个左右,峰值年份分别在2035~2036年之间和2046年,峰值分别为157万个和1447万个左右。非农业人口“四二一”家庭、广义“四二一”家庭数量占全国的比重分别高达70%~88%和86%以上。

(5) 必须努力使出生性别比尽快恢复正常。与人口宏观仿真方法相比,人口微观仿真方法能够更好地分析婚姻市场上男、女性人口的匹配问题。如果出生性别比长期过高,未来一个时期,20~45岁旺婚年龄男、女性未婚人口之差呈现不断扩大的趋势,2020年、2030年和2050年分别为4290万人左右、6360万人左右和7360万人左右。旺婚年龄的男性“过剩”人口占该年龄段男性人口的比例将逐步攀升,从2010年的约1.2%分别上升到2020年的1.8%、2030年的2.6%和本世纪中叶的3.2%~3.3%。婚姻挤压的问题及其后果,受灾区将主要是农村人口、低收入人口及低素质人口,由此将进一步加剧城乡经济社会发展不协调的问题,并将增加未来养老问题的复杂性。

受篇幅限制,我们将另行撰文对“四二一”家庭微观仿真模型的技术细节、起点数据、预测参数、仿真结果作进一步的讨论。

参考文献:

- 1 王贤. 对城市一代独生子女教育问题的社会性考察. 人口研究, 1989; 5
- 2 风笑天. 中国独生子女研究: 回顾与前瞻. 江海学刊, 2002; 5
- 3 郭志刚. 低生育率下我国家庭模式的发展前景. 人口研究(增刊), 1995
- 4 郭志刚. 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, 2001; 1
- 5 郭志刚, 刘金塘, 宋健. 现行生育政策与未来家庭结构. 中国人口学科, 2002; 1
- 6 郭志刚. 人口发展战略研究——人口预测模拟报告. 国家人口发展战略研究报告(即将出版)
- 7 蒋正华. 新时期人口政策思考. 中国人口科学, 2006; 6
- 8 宋健. “四二一”结构: 形成及其发展趋势. 中国人口科学, 2000; 2
- 9 宋健. 中国的独生子女与独生子女户. 人口研究, 2005; 2
- 10 杨书章, 郭震威. 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, 2000; 4
- 11 曾毅. 一门十分活跃的人口学分支学科——家庭人口学. 中国人口分析. 北京大学出版社, 2004
- 12 N. Keilman (1988). Dynamic Household Models. N. Keilman etc. (ed.), Modelling Household Formation and Dissolution, Clarendon Press: Oxford
- 13 G. Orcutt (1957). A New Type of Socio-Economic System. Review of Economics and Statistics, 58: 773-797
- 14 G. Orcutt, M. Greenberger, J. Korbel, A. Rinlin (1961). Microanalysis of Socio-Economic System: A Simulation Study. New York: Harper & Row
- 15 J. H. Reeves. 亲属人数预测. 曾毅, 郭志刚等译. 家庭人口学: 模型及应用. 北京大学出版社, 1994
- 16 J. E. Smith. 亲属集合的计算机模拟. 曾毅, 郭志刚等译. 家庭人口学: 模型及应用, 北京大学出版社, 1994
- 17 E. van Imhoff, W. Post (1998). Microsimulation Methods for Population Projection. In: Population: An English Selection, Vol. 10, No. 1, New Methodological Approaches in the Social Science, 97-138
- 18 A. Zaidi & K. Rake (2001). Dynamic Microsimulation Models: A Review and Some lessons for SAGE. At: www.lse.ac.uk/depts/sage

(责任编辑: 段成荣 宋 严 收稿时间: 2007-02)