

动态户主率方法与一对夫妻户的定量分析*

~~~~~ 叶小叶 黄颖 宋严 ~~~~~

本文推广运用了I.E. Corner<sup>①</sup>文章中的识别方法,定义了动态户主率及动态户主递进率,对传统的家庭人口学研究中的户主率方法作了补充。本文还引入了总和户主率和总和户主递进率的概念,对家庭的形成进行动态的描述。运用这些方法分析计算了一对夫妻户(一对夫妻户的概念为只由一对夫妻和与其同居同住同经济核算的未婚子女组成的家庭),并对中国一对夫妻户进行了预测。

### 一、家庭定量分析的户主率方法及其推广

目前户主率方法是测算家庭户数量的最基本方法。户主率最简单的定义就是在一个给定的人群中作为家庭户代表的人所占的比例。

户主率方法最大的优点是比较简便易行,所需数据较少且都可从人口普查中获得,具有足够的“鲁棒”性,即对数据中的误差及干扰噪音比较“不敏感”。利用户主率方法来预测未来年代的家庭户数量的作法是,用户主率乘上相对应的人口数量再加总即得总户数。户主率方法的缺点是难以进行家庭结构变动分析,也不能反映人口要素如生育率、死亡率、婚姻率等的变化与家庭户类型与规模变化的关系。近些年来,户主率方法有了许多扩展和改进,但并没有克服上述缺点,因而也没有对户主率方法做出实质性的推广。

户主率方法的实质是利用家庭人口学的某个明确的唯一性的特征来建立家庭户与人群的映象。每一个家庭,有此人群中唯一一人与之对应,称此人为户主或家庭户标记者。此人群中任何一人也都具有或者是户主(家庭户标记),或者是非户主(非标记者)身份。

设t年度年龄为x岁的某类人群的人口数量为 $P(t, x)$ ,当我们用户主率方法确定了家庭户与人口后,设t年度年龄为x岁的家庭户标记者数量为 $H(t, x)$ ,则t年度年龄为x岁的家庭户标记者的户主率定义为:

$$R(t, x) = \frac{H(t, x)}{P(t, x)} \quad (1.1)$$

定义非户主率为:

$$R_o(t, x) = \frac{P(t, x) - H(t, x)}{P(t, x)} = 1 - R(t, x) \quad (1.2)$$

同样,如果我们要讨论N类家庭户,设t年度年龄为x岁的N类家庭户标记者数量分别为 $H_i(t, x)$ ,  $i=1, 2, \dots, N$ 。则可定义这N类家庭户标记者的户主率和非户主率为

$$R_i(t, x) = \frac{H_i(t, x)}{P(t, x)} \quad (i=1, 2, \dots, N) \quad (1.3)$$

$$R_o(t, x) = 1 - \sum_{i=1}^N R_i(t, x) \quad (1.4)$$

这样定义出的户主率是一种静态性指标,只刻划出t年代家庭户的现实状况。设t年度x岁的非家庭户标记者中平均在未来一年中转化成为家庭户标记者的人数为 $D(t, x)$ ,则由此引入动态户主递进率 $t_{OR}(t, x)$ 为:

\* 本研究得到中国科学院管理、决策与信息系统开发研究实验室帮助。

① [英] I.E. Corner: Household Projection Methods J. Forecasting Vol6. (1987)

$$t_{OR}(t, x) = \frac{D(t, x)}{P(t, x) - H(t, x)} \quad (1.5)$$

同样, 如果要探讨N类家庭户和非家庭户标记者的相互转化比率, 则可引入动态户主递进率 $t_{ij}(t, x)$ 如下。令t年度x岁的第j类家庭户标记者中平均在未来一年中转化为第i类家庭户标记者的人数为 $D_{ij}(t, x)$ ,  $i, j=0, 1, \dots, N$ 。当 $i, j$ 为零时表示非家庭户标记者, 则:

$$t_{ij}(t, x) = \frac{D_{ij}(t, x)}{H_j(t, x)} \quad (i, j=0, 1, \dots, N) \quad (1.6)$$

那么按照这种定义应有关系式:

$$\sum_{i=0}^N t_{ij}(t, x) + \mu(t, x) \equiv 1 \quad (j=1, 2, \dots, N) \quad (1.7)$$

式中 $\mu(t, x)$ 为t年代x岁人口的死亡概率。

我们还可引入动态户主率的概念来较全面地描述家庭户变动与人口的关系。动态户主率 $S_r(t, x)$ 的定义为:

$$S_r(t, x) = \frac{D(t, x)}{P(t, x)} \quad (1.8)$$

即t年度x岁全部人口中平均在未来一年中转化为家庭户标记者的比率。动态户主率与动态户主递进率的关系式为:

$$S_r(t, x) = t_{OR}(t, x) \cdot R_0(t, x) \quad (1.9)$$

N类家庭户的动态户主转移率为:

$$S_{ij}(t, x) = \frac{D_{ij}(t, x)}{P(t, x)} \quad (i, j=0, 1, \dots, N) \quad (1.10)$$

动态户主转移率与动态户主递进转移率之间的关系式如下:

$$S_{ij}(t, x) = t_{ij}(t, x) \cdot R_j(t, x) \quad (i, j=0, 1, \dots, N) \quad (1.11)$$

根据这些定义公式和关系式, 本文随后还要建立一些家庭户指标来进一步描述家庭人口状态。

上面引入的这些家庭状态描述参数方程, 如果有相对应口径的家庭人口统计数据, 是可以计算出的, 而且从理论上说, 只要有了明确的家庭户划分标准, 相应的统计工作并不比统计生育状况更困难。但问题是相应的统计数据只有一些家庭户现状的统计数据, 要详细统计动态参数至少要面临人力财力增加的困难。因此确实还需要一些方法来识别出动态参数。下面我们介绍一个方法<sup>①</sup>。

设有N类家庭户, 户主人数分别为 $H_i(t, x)$ ,  $i=0, 1, \dots, N$ 。 $H_0(t, x)$ 为t年代x岁的非户主人数。从t年度到t+1年度, 第i类家庭户户主人数的变化决定于从其它类家庭户户主和非户主转移来的人数、本类户主转移到其它类的人数和死亡人数。于是有方程:

$$H_i(t+1, x+1) = \sum_{j=1}^N t_{ij}(t, x) H_j(t, x) \quad (i=1, 2, \dots, N) \quad (1.12)$$

其中动态户主递进率 $t_{ij}(t, x)$ 满足公式(1.7)。方程(1.7)和(1.12)共有方程 $2 \times N$ 个, 但有未知参数 $N \times N$ 个。有无穷多组解案组合。因此要从此方程中识别出诸动态递进率 $t_{ij}(t, x)$ 还需要补充一些条件。具有明显实际意义的补充条件是假定 $t_{ij}(t, x)$ 是年龄x的连续函数且有充分的光滑性。第二个补充条件为假定 $t_{ij}(t, x)$ 是某种指数型逻辑曲线, 即假定:

$$t_{ij}(t, x) = \frac{1}{1 + e^{-Q_{ij}(x)}} \quad (i, j=1, \dots, N) \quad (1.13)$$

其中 $Q_{ij}(x)$ 是x的多项式函数, 具有未知系数 $a_0, a_1, \dots, a_M$ 。M是多项式的次数。

① 同第29页注①。

$$Q_{ij}(x) = a_0 + a_1x + a_2x^2 + \dots + a_Mx^M \quad (1.14)$$

对应每一对 $j_i$ ，此式中 $M$ 和系数 $a_0, \dots, a_M$ 可能是不同的。将这些公式代入方程(1.12)和(1.7)中后求解这些方程的解 $t_{ij}$ ，就化为求出诸多多项式的次数和系数 $M, \{a_i\}$ 的问题。

为求解这个问题可应用最小二乘方法。在方程中为简便省略 $t$ ，并设 $H_i(x+1)$ 代替 $\bar{H}_i(t+1, x+1)$ ，则可定义为：

$$J_i(\alpha_i) = \sum_{k=x_0}^{x_1} \left[ \bar{H}_i(k+1) - \sum_{j=1}^N t_{ij}(k) H_j(k) \right]^2 \quad (i=1, 2, \dots, N) \quad (1.15)$$

式中 $\alpha_i$ 为所有函数 $t_{ij}$ 中多项式的系数集合。最小二乘法求解出诸函数 $t_{ij}$ 即为寻找出诸多多项式 $Q_{ij}(x)$ 的系数 $\{a_i\}$ ，使得 $J_i(\alpha_i)$ 最小( $i=1, 2, \dots, N$ )。

应用此公式求得诸函数记为 $\bar{t}_{ij}$ ， $i, j=1, 2, \dots, N$ 。而后还需要对这些函数进行规范化处理，以使得它们满足条件(1.7)。规范化条件为：

$$t_{ij}(x) = \frac{\bar{t}_{ij}(x)}{\sum_{i=1}^N \bar{t}_{ij}(x)} \cdot (1 - \mu(x)) \quad \begin{matrix} (i=1, \dots, N) \\ (j=1, \dots, N) \end{matrix} \quad (1.16)$$

这样得出的诸函数 $t_{ij}(x)$ 即为在最佳逼近意义下问题的解。

## 二、一对夫妻户分析

在社会各种家庭类型中，如果不把单身户家庭计算在内，除去一些极个别的特例外，各类家庭一般都是由一对夫妻户家庭发展演化而成。研究家庭演化现象离不开对一对夫妻户家庭的发展规律的认识。无论从定性还是定量研究的角度，对一对夫妻户家庭的发展演化规律的认识，对家庭人口学起着基础性的作用。

一对夫妻户的概念最基本特征为一个家庭户中只有一对夫妻。这个概念比较窄，是比核心家庭更为基本的单元，较便于理论分析。按照现行统计口径，一对夫妻户的概念还可包括除一对夫妻及其子女外的亲属及非亲属若干。这些亲属非亲属的情况可能比较复杂，例如可能包括有一对夫妻的父或母，或更高辈份的人。但其基本特点仍是组成这一多代家庭的人中只有一对夫妻关系。

一对夫妻户的家庭生命周期相对比较简单。一对夫妻结婚并独立居住，一般是一对夫妻户家庭形成的开端。然后生育子女，进入家庭的扩展和稳定阶段。随着生育的结束与子女长大成人并陆续独自成家离开父母，家庭规模逐渐缩小。按一对夫妻户定义家庭没有空巢阶段。当这对夫妻离婚、丧偶或有子女结婚后仍然同居住同经济核算，以及其它一些原因，使此家庭主体不成其为一对夫妻，则家庭解体或转化成其它类型家庭。这是最基本最典型的一种一对夫妻户家庭演化历程。

按一对夫妻户中妻子年龄增长的时间顺序，我们再讨论几种一对夫妻户家庭演化历程。

在上述演化历程中，当处于扩展或收缩阶段中有一个或多个子女结婚后没有独立分居，仍同父母同一居住，则形成一种直系或联合家庭。经过一段时间后，已婚子女分离出去或处于离婚、丧偶状态，使得这个家庭在父母年龄较高时又重新形成了一对夫妻户，直至解体。

在直系或联合家庭中可能老两口先进入离婚、丧偶状态，剩下的已婚子女们分家而又形成一个或数个一对夫妻户家庭，此时一对夫妻们年龄较轻，家庭演化历程再进入基本过程的扩展或收缩阶段。这三种一对夫妻户家庭演化过程如图1所示。这些过程说明，在一对夫妻户妻子的任何年龄组中，都发生一对夫妻户家庭的生成、转化、解体等现象。在较年轻的年龄区间上，一对夫妻户家庭的形成可能以新结婚并独立居住形式为主。在较大的年龄区间中，则可能以其它家庭形式转化成一队夫妻户家庭为主。此处我们只说明上述几种最简单的演化类型，目的在于理解我们下文给出的动态户主率的计算数据。当然一对夫妻户的演化方式远远不止这三种，而且可能会很复杂。

运用家庭投影方法，可以把一对夫妻户家庭与女性人口相联系。在一对夫妻户中选择妻子作为家庭的标志者。这个家庭中妻子可能不是辈份最高的女性，但却是唯一能体现家庭户特征的女性，称为“户主”。户

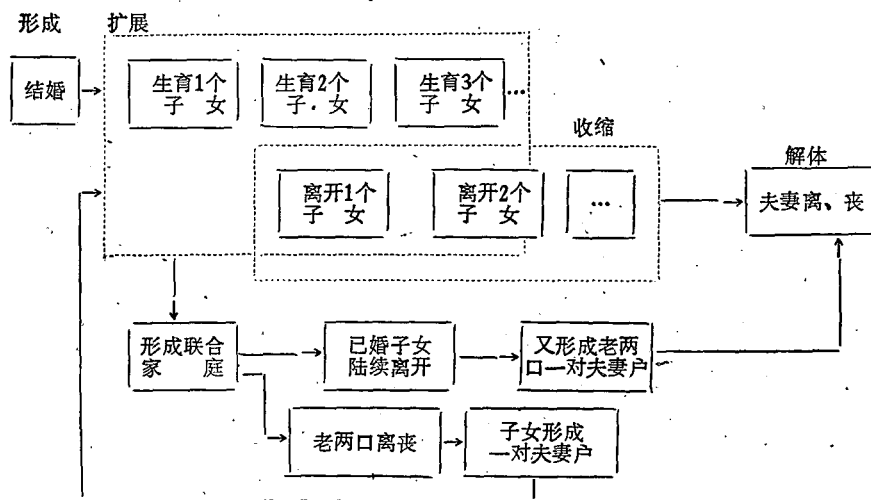


图1 一对夫妻户家庭的三种最简单演化历程

主在同年龄组全体妇女中的比率称为户主率。在不特别讨论其它家庭类型的情况下,对其它妇女统统赋与非家庭户标记,称之为“非户主”。(当我们要进一步讨论其它类型家庭时,可把其它类型家庭的标记者再从非户主中划分出来)。这样,我们把进入家庭状态的所有妇女划分为户主和非户主两类。

上文引入了动态户主率和动态户主递进率的概念。一对夫妻户的动态户主递进率的含义为 $t$ 年度 $k$ 年龄组的非户主妇女中在未来一年中成为户主的妇女所占的比率。因此,在较年轻的年龄区间上,此时一对夫妻户家庭的形成以妇女新婚并离开父母独立居住为主要方式,动态户主递进率体现了年轻妇女组成自己独立家庭的比率。子女离开父母建立自己独立的家庭是一般家庭演化历程的开端。因此对于家庭人口学研究尤其是定量研究来说,动态户主率和动态户主递进率这两个要素非常重要。

类似于对生育率和递进生育率的研究,我们在本文中也引入了总和户主率和总和户主递进率的概念,用以对子女离开父母建立自己独立家庭的动态特征进行描述。

设 $t_{10}(t, x)$ 为 $t$ 年度妻子年龄为 $x$ 岁的一对夫妻户的动态户主递进率。 $S_1(t, x)$ 为 $t$ 年代妻子年龄为 $x$ 岁的一对夫妻户的动态户主率。定义 $t$ 年代一对夫妻户的总和户主率 $TS_1(t)$ 为:

$$TS_1(t) = \sum_{x=x_0}^{x_1} S_1(t, x) \quad (2.1)$$

式中 $x_0, x_1$ 为妇女具有家庭户标记的年龄区间,定义 $t$ 年代一对夫妻户的总和户主递进率 $Tt_1(t)$ 为:

$$Tt_1(t) = \sum_{x=x_0}^{x_1} t_{10}(t, x) \quad (2.2)$$

令 $Sc(x), tc(x)$ 分别为

$$Sc(x) = \frac{S_1(t, x)}{TS_1(t)} \quad (x=x_0, x_0+1, \dots, x_1) \quad (2.3)$$

$$tc_0(x) = \frac{t_{10}(t, x)}{Tt_1(t)} \quad (x=x_0, x_0+1, \dots, x_1) \quad (2.4)$$

$Sc(x)$ 和 $tc_0(x)$ 分别称为一对夫妻户的形成模式和形成递进模式。它们的含义分别为平均每个妇女或每个非户主妇女转化成户主的年龄加权因素。总和户主率 $TS_1(t)$ 的含义为 $t$ 年代每个妇女平均一生中成为一对夫妻户户主的次数。按照上节讨论的一对夫妻户的一般演化历程,一个妇女一生中可能发生多次由非户主

标记向户主标记以及由户主标记向非户主标记的转化。这样,总和户主率 $TS_2$ 的数量应该大于1。

国家统计局1987年全国1%人口抽样调查数据中,按户主年龄分布给出的家庭类型只有一对夫妻户(单身户除外)。该数据是按一对夫妻户的妻子女年龄分组(5岁为一组)的一对夫妻、一对夫妻及其子女、一对夫妻及其子女以及亲属非亲属划分的家庭户数量。其中带有亲属及非亲属的一对夫妻户数量比例较小。

根据这些数据我们应用前述模型方法来选择一对夫妻户妻子作为户主并用单性别模型来描述妇女的家庭状态。除了户主外,其它类型家庭的妇女以及不是户主的妇女等都具有非户主标记,这样就把所有妇女赋与了家庭状态。 $H_1(t, x)$ 为t年代x岁妇女中的户主人数,  $H_0(t, x)$ 为t年代x岁妇女中的非户主人数。这样,方程(1.12)与(1.7)为:

$$H_1(t+1, x+1) = t_{11}(t, x)H_1(t, x) + t_{10}(t, x)H_0(t, x) \quad (2.5)$$

$$H_0(t+1, x+1) = t_{01}(t, x)H_1(t, x) + t_{00}(t, x)H_0(t, x) \quad (2.6)$$

$$t_{01}(t, x) + t_{11}(t, x) + \mu(t, x) \equiv 1 \quad (2.7)$$

$$t_{00}(t, x) + t_{10}(t, x) + \mu(t, x) \equiv 1 \quad (2.8)$$

应用前述识别方法得出上面4个动态递进转移率 $t_{01}$ ,  $t_{11}$ ,  $t_{00}$ ,  $t_{10}$ 。下面只说明 $t_{11}$ 和 $t_{10}$ 。

$$Tt_1(t) = \sum_{x=x_0}^{x_1} t_{10}(t, x), \quad Tt_2(t) = \sum_{x=x_0}^{x_1} t_{11}(t, x)$$

$Tt_1(t)$ 和 $Tt_2(t)$ 分别为非户主转移成户主和户主转移成户主的总和递进转移率。

$$tc_0(x) = \frac{t_{10}(t, x)}{Tt_1(t)}, \quad tc_1(x) = \frac{t_{11}(t, x)}{Tt_2(t)}$$

$tc_0(x)$ 和 $tc_1(x)$ 分别为相应的转移模式。于是 $Tt_1$ 和 $tc_0$ 即为一对夫妻户的总和户主递进率和形成递进模式,表1为15岁到64岁的 $tc_0(x)$ 、 $tc_1(x)$ 以及一对夫妻户的形成模式 $Sc(x)$ 。

图2分别为 $tc_0(x)$ 、 $tc_1(x)$ 和 $Sc(x)$ 的曲线。

从表1的一对夫妻户的年龄形成递进模式 $tc_0(x)$ 来看,在非户主妇女30岁时为向户主转移的第一个峰值年龄。此时平均约有14.15%的非户主妇女在未来一年中转变成为户主。形成递进模式曲线经历一下降后再度上升,一直上升到64岁,说明在较高年龄段上,按照前文所讨论的一对夫妻户的形成转化方式,非户主有较

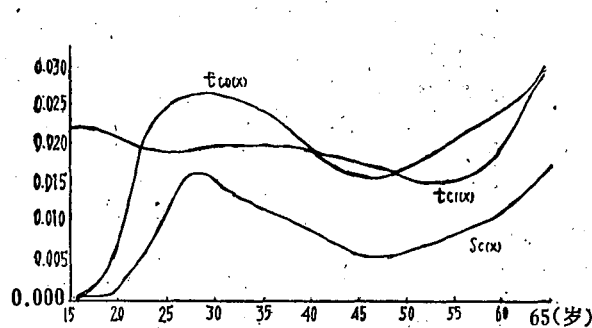


图2 1987年一对夫妻户中女性户主转移模式

高的转化成为户主的可能性。

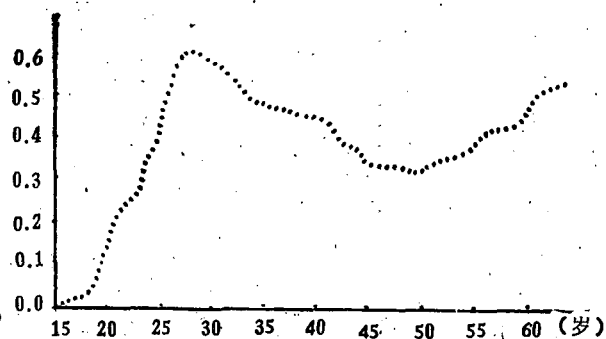


图3 1987年一对夫妻户女性单性别户主率曲线

图3为用1987年一对夫妻户实际户数按年龄分组的女性单性别户主率曲线经过适当平滑后的曲线(这种平滑是为把统计数据的5岁年龄组分成单岁年龄组而作的微小调整)。

表2则是应用本文得出的动态户主递进率函数对未来50年间中国一对夫妻户情况的预测<sup>①</sup>(预测中假定了这些家庭人口要素,包括生育率、婚姻率、户主递进率等在未来年代仍为1987年的模式而不改变),从表2可看出将来拥有1个或2个子女的一对夫妻户为主要家庭类型。

① 参见叶小叶、黄颖的“家庭人口多状态发展方程及其应用”。《中国人口科学》1990,1。

表1

1987—一对夫妻户中不同年龄女性户主转移率

| 年龄 | $t_{c1}(x)$ | $t_{co}(x)$ | $Sc(x)$   |
|----|-------------|-------------|-----------|
| 15 | .02209854   | .00005791   | .00000000 |
| 16 | .02207400   | .00019376   | .00000025 |
| 17 | .02201302   | .00057005   | .00000183 |
| 18 | .02188440   | .00148974   | .00001666 |
| 19 | .02165378   | .00344965   | .00010861 |
| 20 | .02130289   | .00693170   | .00124376 |
| 21 | .02085150   | .01175596   | .00254214 |
| 22 | .02036200   | .01674963   | .00425169 |
| 23 | .01991624   | .02070756   | .00558080 |
| 24 | .01958084   | .02331256   | .00794899 |
| 25 | .01938639   | .02486865   | .01055668 |
| 26 | .01932841   | .02576672   | .01298886 |
| 27 | .01938021   | .02628093   | .01576187 |
| 28 | .01950633   | .02656704   | .01622739 |
| 29 | .01967157   | .02670258   | .01626831 |
| 30 | .01984551   | .02671889   | .01565901 |
| 31 | .02013061   | .02639862   | .01373878 |
| 32 | .02021344   | .02603681   | .01309709 |
| 33 | .02024680   | .02552129   | .01228196 |
| 34 | .02022889   | .02484599   | .01176496 |
| 35 | .02016094   | .02401814   | .01128463 |
| 36 | .02004618   | .02306089   | .01076224 |
| 37 | .01988885   | .02201245   | .01014469 |
| 38 | .01969343   | .02092187   | .00952055 |
| 39 | .01946406   | .01984275   | .00885566 |
| 40 | .01920420   | .01882668   | .00833638 |
| 41 | .01891659   | .01791833   | .00696613 |
| 42 | .01860340   | .01715275   | .00657974 |
| 43 | .01826661   | .01655504   | .00613068 |
| 44 | .01790859   | .01614119   | .00544420 |
| 45 | .01753274   | .01591972   | .00525840 |
| 46 | .01714427   | .01589268   | .00523575 |
| 47 | .01675089   | .01605620   | .00531782 |
| 48 | .01636356   | .01640000   | .00555167 |
| 49 | .01599701   | .01690646   | .00565236 |
| 50 | .01567017   | .01754999   | .00609511 |
| 51 | .01540632   | .01829741   | .00654476 |
| 52 | .01523327   | .01911027   | .00686705 |
| 53 | .01518339   | .01994924   | .00722683 |
| 54 | .01529376   | .02077967   | .00828768 |
| 55 | .01560618   | .02157718   | .00886664 |
| 56 | .01616672   | .02233177   | .00929540 |
| 57 | .01702342   | .02305029   | .00995842 |
| 58 | .01822023   | .02375800   | .01025024 |
| 59 | .01978429   | .02450013   | .01141411 |
| 60 | .02170481   | .02534380   | .01205913 |
| 61 | .02390752   | .02637630   | .01361058 |
| 62 | .02623481   | .02768812   | .01463923 |
| 63 | .02848995   | .02932245   | .01573391 |
| 64 | .03044915   | .03119352   | .01713528 |

表2 15~64岁女性人口所在的各类一对夫妻户的户数预测 (万户)

| 年份   | 一对夫妻户<br>无子女 | 一对夫妻户<br>1个子女 | 一对夫妻户<br>2个子女 | 一对夫妻户<br>3及3个以上子女 | 一对夫妻户<br>总计 |
|------|--------------|---------------|---------------|-------------------|-------------|
| 1987 | 1 361        | 4 178         | 4 737         | 2 184             | 12 460      |
| 1997 | 1 051        | 5 059         | 5 964         | 1 679             | 13 453      |
| 2007 | 1 230        | 4 881         | 6 186         | 1 633             | 13 929      |
| 2017 | 1 574        | 5 543         | 6 122         | 1 368             | 14 407      |
| 2037 | 1 581        | 5 303         | 6 280         | 1 393             | 14 557      |

(本文责任编辑: 张京华)

(作者工作单位: 叶小叶、黄颖: 北京信息控制研究所

宋严: 中国人民大学人口学系)

(上接第62页)

口档案, 暂住地要建立流入人口档案。两地要经常联系, 及时交流信息。

第四, 各部门密切配合, 实行综合治理。计划生育、公安、工商、税务、妇联、劳动、宣传、共青团等部门要密切配合, 要利用本部门的特点与优势, 把好计划生育关。

第五, 农村计划生育活动, 结扎手术一定突破女儿户, 生两个或两个以上的女儿户, 夫妇双方一定要有一方做结扎手术, 只有突破女儿户才能真正有效控制人口增长。

第六, 各地劳动部门, 要有组织、有计划、有目的地将农村剩余劳动人口向城镇输入, 消除“盲流”现象。加强流动人口的组织性、纪律性, 制止生育上的无政府主义。

第七, 继续开展多种形式的调查研究活动, 摸索流动人口的生活、生育等规律, 以便加强管理。(本文责任编辑: 朱犁)  
(作者工作单位: 安徽省无为县计生委)