

# 独生属性与婚姻匹配研究\*

## ——对“随机婚配”假定的检验

郭志刚 许琪

**【摘要】**文章使用 2010 年中国家庭追踪调查数据,对独生子女的婚姻匹配问题进行了深入研究,发现独生子女更可能与独生子女结婚,且非独生子女也更可能与非独生子女结婚。这表明,一方面独生子女身份本身可能构成了一种择偶标准,另一方面人们择偶时对城乡户籍、出生地、年龄和教育程度的多重选择也同样会导致独生子女更可能与独生子女结婚。然而,以往对生育政策调整的人口预测模拟多是基于“随机婚配”假定,因此关于婚配结果分布的估计存在两种误差:一是低估“双独”和“双非”夫妇的比例和规模;二是高估“单独”夫妇的比例和规模。因此,对“双独”和“单独”夫妇数量的估计应当充分考虑婚姻匹配过程的选择性,从而提高预测结果的准确性。

**【关键词】**独生子女 婚姻匹配 “随机婚配”假定

**【作者】**郭志刚 北京大学中国社会发展研究中心研究员、社会学系教授;许琪 南京大学社会学院,助理研究员。

### 一、问题缘起

自 1973 年全国性的计划生育政策实施以来,特别是 1980 年中共中央“提倡一对夫妇只生育一个孩子”以来,中国的独生子女数量迅速增加。杨书章、王广州(2007)的估计结果显示,中国 2007 年末 0~17 岁独生子女总量为 1.14 亿,18~28 岁独生子女总量为 3 640 万,30 岁以下独生子女总量合计已超过 1.5 亿。随着年龄的增长,这些独生子女都将面临婚育问题。如果其配偶也是独生子女,就会形成“双独”夫妇;如果其配偶是非独生子女,就会形成“单独”夫妇。目前,学术界已对“双独”和“单独”夫妇进行了很多研究。就“双独”夫妇而言,由于他们是形成“四二一”家庭结构的一个必要条件,而且全国大多数省份都允许“双独”夫妇生育二孩,所以“双独”夫妇的养老问题及其生育意愿与生育行为都备受关注(宋

\* 本文为教育部人文社会科学重点研究基地重大项目“中国低生育率研究”(项目批准号:12JJD840005)的阶段性成果。

健,2001、2010;风笑天,2006;马小红,2004;包蕾萍、陈建强,2005)。就“单独”夫妇而言,十八届三中全会提出了一方是独生子女的夫妇可生育两个孩子的“单独二胎”政策,作为这一政策的目标群体,“单独”夫妇的生育意愿和生育行为近来也已成为社会各界广泛讨论的热点(张丽萍、王广州,2014;乔晓春,2014)。

我们认为,在上述问题中搞清年轻男女中按独生属性婚配分布的实际情况,尤其是“双独”和“单独”夫妇的数量、结构及其在未来的发展趋势是最根本的。虽然已有研究较多(郭志刚等,2002;齐险峰、郭震威,2007;郭震威、齐险峰,2008;王广州,2012;张丽萍、王广州,2014),但均基于年轻男女婚配时对另一方是否为独生子女并没有明显偏好的假定。虽然在没有充分数据资料的条件下,假定独生子女与非独生子女之间随机婚配(以下简称“随机婚配”假定)是当时唯一可行的选择;但必须注意由此可能产生的低估“双独”夫妇的数量和比例等问题(郭志刚等,2002;齐险峰、郭震威,2007)。

郭震威、齐险峰(2008)曾利用2006年全国人口和计划生育抽样调查数据对该假定进行过检验。他们发现,独生子女愿意与独生子女结婚的比例为19.2%,而非独生子女愿意与独生子女结婚的比例为11.1%。虽然二者相差8个百分点,但在统计上并不显著,所以他们认为可以使用“随机婚配”假定进行预测。因为意愿不能等同于真实的行为,况且接近75%的受访者没有明确表示他们是更愿意与独生子女结婚还是更愿意与非独生子女结婚,由于回答“无所谓”和“说不清”的比例高,因此必须对研究结论持谨慎态度。

本文将用实际婚姻匹配的调查数据对“随机婚配”假定进行更加严格的检验,并探讨和分析这种非随机婚配现象产生的原因,讨论以往简单化假定导致既有研究中估计“双独”和“单独”夫妇的规模和比例时存在的问题,并提出可能的改进方案。本研究的发现对深入认识人口实际有重要意义,对正确评估以往有关人口预测结果也有新的启示。

## 二、数据资料

要对“随机婚配”假定进行检验必须同时知道夫妇双方的独生子女属性,但满足这一基本要求的数据资料非常匮乏。例如,历次人口普查并未登记调查对象的兄弟姐妹信息;各种专项抽样调查或者没有询问受访者的兄弟姐妹数量,或者只询问了夫妇其中的一方,所以不能满足研究的基本条件,导致长期以来独生子女是否更可能与独生子女结婚一直是一个悬而未决的问题。

本文使用的2010年中国家庭追踪调查数据不仅数据质量较高,而且在很大程度上克服了其他调查的上述缺陷。该调查采用分层、多阶段、与人口规模成比例的概率抽样方法,样本覆盖了除香港特别行政区、澳门特别行政区、台湾、新疆、西藏、青海、宁夏、内蒙古和海南之外的25个省份,覆盖面约占总人口的95%。谢宇(2012)曾对该调查数据质量进行评估,发现该调查的性别、年龄分布、婚姻状况和受教育程度等基本特征与2010年第六次人口普查非常接近。

就本研究问题而言,中国家庭追踪调查数据具有较多优势。一是调查了同一户中的所有人口,通过家庭关系,可以非常方便地将夫妻匹配起来进行分析。二是询问了所有受访者的兄弟姐妹数量,这样就能同时了解夫妻双方是否为独生子女。三是搜集了夫妇双方在社会、经济和人口等方面的很多信息,可以对独生子女婚姻匹配问题进行很多深入研究和探讨。本文的研究对象限定为年龄在40岁以下的已婚妇女,因为40岁及以上已婚妇女中独生子女比例很低,而且她们已度过了生育最旺盛的年龄,分析其意义不大。样本中满足这一年龄要求的已婚妇女总数为4 612人。通过家庭关系数据库,匹配成功的夫妇有3 559对,成功率为77.2%;匹配失败的主要原因是夫妇两地分居和丈夫拒访。在去除缺失后,进入实际分析的夫妇有3 516对。

三、分析结果

(一) 对“随机婚配”假定进行检验

本文首先对“随机婚配”假定进行了检验。如表1所示,根据夫妻是否独生子女交互分类后,主对角线上的两个单元格的观测频数都明显大于期望频数<sup>①</sup>,非主对角线上两个单元格的观测频数都小于期望频数。这说明,独生子女更可能与独生子女结婚,且非独生子女也更可能与非独生子女结婚,而独生子女与非独生子女结婚的可能性则相对较小。经计算,独生子中与独生女结婚的比例占26.6%,而非独生子中与独生女结婚的比例仅为5.1%,二者相差21.5个百分点。对该比例差的t检验结果显著,所以二者的差异真实存在。此外,本文还对该列联表的独立性进行了卡方检验,结果也是显著的,这就再次拒绝了独生与非独生之间随机婚配的原假设。根据表1计算结果,相对于非独生子,独生子与独生女结婚的优势比为6.69<sup>②</sup>。可见,夫妇双方是否为独生子女的关联性很强,“随机婚配”假定并不符合实际。

表1 丈夫是否独生子女与妻子是否独生子女的关联性

| 丈夫是否<br>独生子女 | 妻子是否独生子女   |              | 合计         |
|--------------|------------|--------------|------------|
|              | 是          | 否            |            |
| 是            | 106(30.1)  | 293(368.9)   | 399(399)   |
| 否            | 159(235.1) | 2958(2881.9) | 3117(3117) |
| 合计           | 265(265)   | 3251(3251)   | 3516(3516) |

注:表中数据为40岁以下已婚妇女。括号内数据为期望频数。表中结果已加权(由于中国家庭追踪调查对上海、辽宁、河南、甘肃和广东5个省进行了超额抽样,所以在对全国总体进行推断时需要加权)。

分城乡(见表2)和年龄(见表3)来看,我们也能得到类似的结论。城市的优势比为5.85,农村为5.57;在25岁以下、25~29岁、30~34岁和35~39岁4个年龄组中,优势比分别为3.88、5.74、4.40和14.13。统计检验结果显示,无论在农村还

① 期望频数反映在随机婚配条件下应有的结果。  
② 软件报告的优势比为6.69,但根据表1计算的结果是6.73。不一致的原因是表1在加权后,观测频数都带有小数,报告时进行了四舍五入,这导致二者存在细微的差异。

是城市,无论妻子的年龄大小,均呈现出独生子女更可能与独生子女结婚,而非独生子女更可能与非独生子女结婚的现象。

(二) 独生子女更可能与独生子女结婚的原因分析

上面的分析结果充分说明“随机婚配”假定并不成立,而且从全国和分城乡、年龄组的优势比的大小看,夫妇双方是否为独生子女的关联性很强。那么,什么原因导致独生子女更可能与独生子女结婚呢?我们认为,要了解这一现象背后的原因,必须从择偶行为的微观机制入手。一方面,在当前特定社会条件下,独生子女属性已成为一个明确的社会择偶标准当然是非常可能的。另一方面,即使独生子女属性并未成为一个独立择偶标准,其他社会选择过程也非常可能催生独生子女婚配的现象。

因为人们在择偶时总会根据一定的标准对潜在的结婚对象进行筛选,如对方的年龄、家庭背景、社会经济地位等(李煜、徐安琪,2004;徐安琪,2000)。其结果往往与自己具有相似特征的人结婚(张翼,2003;李煜、陆新超,2008;Schwartz 等,2005),即婚姻同质性匹配。在中国,户籍、籍贯、年龄和受教育程度是非常重要的4个择偶标准,而这四重标准都与是否为独生子女密切相关。

首先,中国长期存在的“城乡二元分割”导致城市人在收入水平、职业地位、社会保障水平和文化观念等方面与农村人有较大的差异(Wu 等,2004),由此产生的结构性壁垒导致城乡之间通婚的比例一直不高。中国的计划生育政策在城乡之间存在明显差异,城市自1980年以后基本实行的是“一刀切”的独生子女政策,而农村的生育政策则复杂多样,至少存在

表 2 丈夫是否独生子女与妻子是否独生子女的  
关联性(分城乡) 人

| 居住地 | 丈夫是否<br>独生子女 | 妻子是否独生子女 |              | 合计         |
|-----|--------------|----------|--------------|------------|
|     |              | 是        | 否            |            |
| 农村  | 是            | 19(4.8)  | 104(118.2)   | 123(123)   |
|     | 否            | 54(68.2) | 1676(1661.8) | 1730(1730) |
|     | 合计           | 73(73)   | 1780(1780)   | 1853(1853) |
| 城市  | 是            | 80(28)   | 180(232)     | 260(260)   |
|     | 否            | 99(151)  | 1304(1252)   | 1403(1403) |
|     | 合计           | 179(179) | 1484(1484)   | 1663(1663) |

注:同表 1。

表 3 丈夫是否独生子女与妻子是否独生子女的  
关联性(分年龄) 人

| 妻子年龄    | 丈夫是否<br>独生子女 | 妻子是否独生子女 |              | 合计         |
|---------|--------------|----------|--------------|------------|
|         |              | 是        | 否            |            |
| 25 岁以下  | 是            | 12(4.7)  | 46(53.3)     | 58(58)     |
|         | 否            | 25(32.3) | 372(364.7)   | 397(397)   |
|         | 合计           | 37(37)   | 418(418)     | 455(455)   |
| 25~29 岁 | 是            | 51(20.3) | 114(144.7)   | 165(165)   |
|         | 否            | 44(74.7) | 564(533.3)   | 608(608)   |
|         | 合计           | 95(95)   | 678(678)     | 773(773)   |
| 30~34 岁 | 是            | 21(7.2)  | 80(93.8)     | 101(101)   |
|         | 否            | 49(62.8) | 838(824.2)   | 887(887)   |
|         | 合计           | 70(70)   | 918(918)     | 988(988)   |
| 35~39 岁 | 是            | 17(2.2)  | 37(51.8)     | 54(54)     |
|         | 否            | 37(51.8) | 1209(1194.2) | 1246(1246) |
|         | 合计           | 54(54)   | 1246(1246)   | 1300(1300) |

注:同表 1。



“一孩”、“一孩半”和“二孩”3种不同的政策类型。因此,不仅城市人是独生子女的可能性较大,而且如果他们想找城市人结婚,其配偶很有可能也是独生子女。

其次,中国是一个幅员非常辽阔的国家,各地的风土人情和生活习惯各不相同,再加上地缘关系一直是中国传统社会一种非常重要的关系类型,这导致籍贯也往往成为人们在择偶时需要加以考虑的一个重要因素。试想一个上海本地人若想找一个上海本地人结婚,那么除了独生子女之外,已经几乎没有其他选择。

第三,择偶时双方的年龄也需要匹配。有研究表明,中国人的婚龄差通常集中在夫小于妻1岁到夫大于妻4岁之间,婚龄差超过5岁的婚姻相对较少。同龄人结婚很有可能会增加独生子女与独生子女结婚的概率,因为独生子女在不同年龄组的分布是不均匀的,总体而言,低龄组独生子女的比例高,而高龄组独生子女的比例低(王广州,2009)。年龄的限制降低了低龄组独生子女与高龄组非独生子女婚配的可能性,所以会出现大量独生子女与独生子女结婚的现象。

第四,受教育程度也是一个非常重要的择偶标准。研究表明,中国人倾向于同自己受教育程度相同的人结婚,尤其是1980年以来,教育的同质性匹配在不断增强(李煜,2008)。另外,受教育水平与兄弟姐妹数量密切相关。由于兄弟姐妹数量越多,平均每个子女能够分到的教育资源就越少,这在西方国家和中国都已得到证明(Blake,1989;Lu等,2008)。所以,如果受教育程度较高的人更可能与受教育程度较高的人结婚,那么独生子女与独生子女结婚的可能性也会随之增加。

综上所述,至少存在4种可能导致独生子女更可能与独生子女结婚的微观选择机制。为了对上述假设进行检验,本文首先分“双独”、“单独”和“双非”<sup>①</sup>3种类型对夫妇双方3岁时的户籍、出生地、年龄和受教育程度进行描述(见表4)。从表4可以发现,“双独”夫妇3岁时是非农户籍的比例明显大于“单独”和“双非”夫妇;而且“双独”夫妇出生在北京、天津、上海、四川这几个计划生育政策较严的省份的比例也非常高。此外,“双独”夫妇的年龄普遍较小,且受教育程度最高,这均与我们的预期一致。特别值得注意的是,这些特征在夫妻双方的分布中具有较强的对称性。例如,“双独”夫妇中丈夫在3岁时是非农户籍的比例为62.2%,而妻子为62.0%,二者非常接近,这很可能是城市人之间相互通婚的结果。从夫妇的出生地、年龄和受教育程度,也能发现类似的对称性。总之,表4的数据在一定程度上验证了本文的理论假设,即人们择偶过程中的选择性增加了独生子女与独生子女结婚的概率。

择偶过程中的选择性在多大程度上能够解释夫妇在独生子女属性上的关联性?为了回答这个问题,我们以丈夫是否是独生子女为因变量,以妻子是否是独生子女为主要的预测变量建立Logistic回归模型,结果如表5所示。

表5中模型1仅纳入了妻子是否独生子女一个自变量,该变量的回归系数为1.901,换算

①“双非”指夫妇双方都不是独生子女。

成优势比后为 6.693，这与表 1 中的描述性统计结果完全一致。自模型 2 开始，逐步加入夫妇 3 岁时的户籍、出生地、年龄和受教育程度 4 个控制变量。可以发现，在逐步纳入控制变量以后，妻子是否是独生子女的回归系数有明显下降，从模型 1 的 1.901 下降到模型 5 的 0.943，下降幅度为 50.4%；如果换算成优势比，则从 6.693 下降到 2.568，下降幅度为 72.5%。妻子是否独生子女这个自变量的回归系数值在各模型中不断减少（更接近于 0）的变化过程反映出，最初未被纳入模型的其他自变量由于同时与该自变量和因变量都有较强相关，因而它们对因变量的影响都被附加到该自变量上，所以该自变量的真正影响其实被高估了。然而，当其他自变量被一一加入回归模型后，该自变量上原来被附加的其他变量影响就被

表 4 “双独”、“单独”和“双非”夫妇特征描述

|           | 丈夫   |      |      |      | 妻子   |      |      |      |
|-----------|------|------|------|------|------|------|------|------|
|           | 双独   | 单独   | 双非   | 合计   | 双独   | 单独   | 双非   | 合计   |
| 3 岁时户籍(%) |      |      |      |      |      |      |      |      |
| 农业        | 37.8 | 65.0 | 89.8 | 85.0 | 38.0 | 69.9 | 90.6 | 86.3 |
| 非农        | 62.2 | 35.0 | 10.2 | 15.0 | 62.0 | 30.1 | 9.4  | 13.7 |
| 出生地(%)    |      |      |      |      |      |      |      |      |
| 北京        | 13.1 | 1.4  | 0.3  | 0.8  | 8.2  | 1.2  | 0.1  | 0.5  |
| 天津        | 5.6  | 3.5  | 0.8  | 1.3  | 2.6  | 3.0  | 0.7  | 1.1  |
| 河北        | 3.6  | 4.5  | 8.6  | 7.9  | 3.6  | 5.7  | 8.7  | 8.0  |
| 山西        | 2.8  | 7.4  | 7.5  | 7.3  | 2.8  | 7.5  | 7.4  | 7.3  |
| 辽宁        | 9.0  | 7.6  | 2.3  | 3.2  | 9.6  | 7.3  | 2.3  | 3.2  |
| 吉林        | 6.1  | 4.4  | 2.5  | 2.8  | 6.1  | 4.4  | 2.5  | 2.8  |
| 黑龙江       | 6.6  | 9.6  | 6.1  | 6.6  | 9.7  | 8.4  | 6.3  | 6.6  |
| 上海        | 9.9  | 3.1  | 0.5  | 1.1  | 9.8  | 2.0  | 0.2  | 0.7  |
| 江苏        | 2.7  | 4.2  | 2.4  | 2.6  | 3.1  | 4.2  | 2.2  | 2.5  |
| 浙江        | 1.9  | 4.0  | 1.6  | 1.9  | 2.2  | 2.7  | 1.4  | 1.6  |
| 安徽        | 1.3  | 0.2  | 2.6  | 2.3  | 0.7  | 0.3  | 2.4  | 2.1  |
| 福建        | 0.1  | 1.4  | 1.8  | 1.7  | 0.0  | 0.9  | 1.8  | 1.7  |
| 江西        | 0.1  | 1.3  | 2.4  | 2.2  | 0.6  | 1.2  | 2.5  | 2.3  |
| 山东        | 1.3  | 12.4 | 5.3  | 6.1  | 1.3  | 13.0 | 5.3  | 6.2  |
| 河南        | 2.7  | 4.2  | 7.8  | 7.1  | 4.9  | 3.6  | 7.8  | 7.2  |
| 湖北        | 6.4  | 4.3  | 2.1  | 2.5  | 7.1  | 4.7  | 2.3  | 2.8  |
| 湖南        | 1.3  | 3.0  | 4.5  | 4.2  | 1.3  | 2.6  | 4.3  | 4.0  |
| 广东        | 3.7  | 3.6  | 7.8  | 7.1  | 3.7  | 4.7  | 7.4  | 7.0  |
| 广西        | 0.0  | 1.5  | 3.9  | 3.5  | 0.0  | 1.8  | 4.5  | 4.0  |
| 重庆        | 0.6  | 0.2  | 0.9  | 0.8  | 1.6  | 0.5  | 0.9  | 0.9  |
| 四川        | 18.7 | 11.0 | 8.9  | 9.6  | 18.6 | 11.4 | 8.3  | 8.7  |
| 贵州        | 0.2  | 2.6  | 7.7  | 6.8  | 0.2  | 3.0  | 7.8  | 7.0  |
| 云南        | 1.0  | 1.4  | 6.7  | 5.9  | 1.2  | 1.9  | 7.3  | 6.5  |
| 陕西        | 0.0  | 1.0  | 2.7  | 2.4  | 0.4  | 1.3  | 2.8  | 2.6  |
| 甘肃        | 0.5  | 1.5  | 2.0  | 1.9  | 0.5  | 1.8  | 2.1  | 2.0  |
| 其他        | 0.8  | 0.7  | 0.3  | 0.4  | 0.2  | 0.9  | 0.7  | 0.7  |
| 年龄(岁)     |      |      |      |      |      |      |      |      |
| 均值        | 30.4 | 31.0 | 33.6 | 33.2 | 28.4 | 28.9 | 31.3 | 30.9 |
| 标准差       | 4.9  | 5.2  | 6.0  | 6.0  | 4.6  | 4.8  | 5.6  | 5.5  |
| 受教育程度(%)  |      |      |      |      |      |      |      |      |
| 文盲        | 5.3  | 4.9  | 12.3 | 11.2 | 4.5  | 6.8  | 18.6 | 16.7 |
| 小学        | 9.4  | 11.4 | 22.3 | 20.5 | 6.6  | 14.6 | 24.1 | 22.4 |
| 初中        | 24.5 | 41.4 | 40.7 | 40.3 | 27.6 | 36.9 | 38.2 | 37.6 |
| 高中        | 18.5 | 18.4 | 14.5 | 15.1 | 16.2 | 18.0 | 10.8 | 11.9 |
| 大专        | 23.2 | 14.3 | 6.2  | 7.7  | 22.6 | 13.2 | 4.9  | 6.5  |
| 本科及以上     | 19.1 | 9.6  | 4.0  | 5.2  | 22.5 | 10.5 | 3.4  | 4.9  |
| 样本量(人)    | 106  | 452  | 2958 | 3516 | 106  | 452  | 2958 | 3516 |

注：中国家庭追踪调查调查仅覆盖了全国 25 个省份的人口，但受访者的出生地可能在其他省份，由于这部分样本的比例很小，这里合并为“其他”。

表 5 对丈夫是否是独生子女的 Logistic 回归分析

|                 | 模型 1              | 模型 2              | 模型 3             | 模型 4              | 模型 5              |
|-----------------|-------------------|-------------------|------------------|-------------------|-------------------|
| 妻子是独生子女         | 1.901*** (0.179)  | 1.471*** (0.207)  | 1.266*** (0.240) | 1.026*** (0.281)  | 0.943*** (0.286)  |
| 妻子 3 岁时户籍为非农    |                   | -0.009 (0.235)    | 0.082 (0.263)    | 0.117 (0.279)     | 0.021 (0.279)     |
| 丈夫 3 岁时户籍为非农    |                   | 1.409*** (0.201)  | 1.350*** (0.222) | 1.735*** (0.240)  | 1.517*** (0.245)  |
| 妻子年龄            |                   |                   |                  | 0.013 (0.030)     | 0.011 (0.031)     |
| 丈夫年龄            |                   |                   |                  | -0.152*** (0.030) | -0.150*** (0.031) |
| 妻子受教育程度 (文盲 =0) |                   |                   |                  |                   |                   |
| 小学              |                   |                   |                  |                   | 0.990* (0.442)    |
| 初中              |                   |                   |                  |                   | 1.257** (0.422)   |
| 高中              |                   |                   |                  |                   | 1.418** (0.454)   |
| 大专              |                   |                   |                  |                   | 1.624** (0.503)   |
| 大学及以上           |                   |                   |                  |                   | 2.047*** (0.523)  |
| 丈夫受教育程度 (文盲 =0) |                   |                   |                  |                   |                   |
| 小学              |                   |                   |                  |                   | -0.259 (0.466)    |
| 初中              |                   |                   |                  |                   | -0.280 (0.450)    |
| 高中              |                   |                   |                  |                   | -0.418 (0.479)    |
| 大专              |                   |                   |                  |                   | -0.288 (0.523)    |
| 大学及以上           |                   |                   |                  |                   | -0.421 (0.576)    |
| 截距              | -2.311*** (0.079) | -2.588*** (0.091) | -1.093 (0.743)   | 3.249*** (0.805)  | 2.374*** (0.857)  |
| Wald 卡方         | 112.78***         | 159.35***         | 407.65***        | 417.65***         | 406.74***         |
| 自由度             | 1                 | 3                 | 53               | 55                | 65                |
| 伪 R-square      | 0.065             | 0.112             | 0.212            | 0.275             | 0.289             |
| 样本量             | 3516              | 3516              | 3516             | 3516              | 3516              |

注：妻子和丈夫的出生地以虚拟变量的形式纳入模型，由于篇幅限制，没有列出这些系数。括号内数据为回归系数的标准误。\* $p<0.05$ ，\*\* $p<0.01$ ，\*\*\* $p<0.001$ 。

一一剥离并直接反映为其他自变量的独立影响，于是该自变量逐步减小的回归系数就越来越接近于它本身对因变量的独立影响。另一方面，最终模型结果表明，其他变量的影响确实也很显著。所以，人们在择偶时对户籍、出生地、年龄和受教育程度的选择性确实是导致独生子女更可能与独生子女结婚的重要原因。

不过，从表 5 中的模型 5 可以发现，即使在控制了夫妇 3 岁时的户籍、出生地、年龄和受教育程度以后，妻子是否独生子女依然对丈夫是否独生子女具有显著的解释力。这说明，独生子女本身很有可能也是择偶时需要考虑的一个重要因素。一方面，这可能是由于相似的家庭环境和成长经历造就了独生子女在性格特征、生活方式等方面的相似性，而且人们总是希望寻找与自己性格相近的人结婚，导致独生子女更可能与独生子女走到一起。另一方面，中国大多数省份在政策上都允许“双独”夫妇生育二孩，这在一定程度上也会导致有二孩生育意愿的独生子女在择偶时刻意选择独生子女。

总之,上述分析结果充分说明,择偶与“随机摇号”不同,它是一个充满选择性的过程。无论是因为户籍、出生地、年龄、受教育程度等原因被动地选择独生子女,还是由于性格特征、二孩生育等原因主动地选择独生子女,都会导致独生子女更可能与独生子女结婚的客观事实。所以“随机婚配”假定严重脱离实际。

#### 四、对人口预测的启示

上述分析结果可以表明,独生子女与非独生子女之间“随机婚配”的假定与人口现实的差距十分明显。那么,由于迄今为止几乎所有对“双独”和“单独”夫妇所做的估计和预测都建立在这个假定上,便一定会影响“双独”和“单独”夫妇估计和预测结果的准确性。具体来说,这种影响体现在两个方面。

其一,以往的估计和预测有可能低估了“双独”夫妇的规模 and 比例,从而低估了“四二一”家庭的规模 and 比例及其在未来的发展趋势,也就是说,中国家庭养老的压力可能比以往预期的结果还要严重。目前,学术界对“四二一”家庭仍然存在不同的界定,狭义的“四二一”家庭涉及三代共 7 口人,其中两代人是独生子女;而广义的“四二一”家庭则放松了这 7 口人必须同时存活的条件,并由此衍生出多种可能的家庭类型(齐险峰、郭震威,2007),也有学者将这种广义的“四二一”家庭称为“X-2”家庭(郭志刚等,2002)。但无论如何界定,所有学者都认为,独生子女与独生子女结婚是形成“四二一”家庭的一个必要条件,所以“四二一”家庭的规模 and 比例在很大程度上取决于“双独”夫妇的规模 and 比例。以往的估计和预测都建立在“随机婚配”这一假定之上,而正如表 1 所示,现实生活中“双独”夫妇的实际观测频数要比该假定条件下的期望频数多得多,所以我们可以预期,在放松“随机婚配”这一简化假定以后估计得到的“双独”夫妇的规模 and 比例也会比以往的研究大得多。同时,低估“双独”夫妇的规模 and 比例也意味着过去低估了“双独二孩”这一生育政策调整的目标人群。人口学界关于生育政策调整的争论已经持续了很多年,争论的焦点之一就在于:生育政策的变动会给出生人口带来多大的影响?会造成多大程度的出生人口堆积?会给未来的教育、医疗、就业造成多大程度的冲击?而在放开“双独”夫妇生育二孩导致出生数量及总和生育率水平的预测模拟上,按“随机婚配”假定会由于低估“双独”夫妇的实际规模而导致低估这种“双独”调整对出生数量和生育率的影响。换句话说,放开“双独”二孩生育的影响幅度其实是明显大于基于这一假定的估计结果的。那么,前些年的“双独”政策调整并未形成对人口发展的显著影响,主要是由于相当一部分放开对象自愿推迟、甚至是完全放弃了再次生育的机会。

其二,以往的预测有可能高估了“单独”夫妇的规模 and 比例。也就是说,“单独二孩”政策的目标人群规模并没有以往预期的那么大。仅就这一点而言,即使对目标人群的生育意愿和行为的估计不存在偏差,那么“单独二孩”放开后的出生堆积也要比以往预期的数量少得多。十八届三中全会宣布启动“单独二孩”政策,那么这一政策调整可能产生的出生堆积有



多大,便具有更重要的现实意义。以往学术界在这方面已经进行了较多研究(中国发展研究基金会,2012;乔晓春,2014;张丽萍、王广州,2014),估计了“单独二孩”政策实施以后每年新增的人口数量、生育水平及相应调整目标人群的育龄妇女数量和结构。但这些研究并未充分考虑独生子女和非独生子女之间婚姻匹配的非随机性,很有可能由此高估了“单独二孩”育龄妇女的总量和可能产生的出生堆积。因此,“单独二孩”政策调整对中国人口规模和结构的影响可能远没有以往预期的那么大。

综上所述,研究独生子女的婚姻匹配结果不仅有助于理解现实中的择偶过程,而且具有非常重要的政策含义。众所周知,任何政策的制定和评估都必须建立在数据准确的基础上,所以有必要在放松“随机婚配”假定以后对“双独”和“单独”夫妇的规模和比例进行重新测算。与一般的人口预测不同,对“双独”和“单独”夫妇进行预测必然涉及婚姻的匹配,所以研究者应当充分认识到婚姻匹配过程的复杂性,并在预测模型中将这种复杂性考虑进来,而不能继续想当然地应用不符合实际情况的假定。

人口普查和绝大多数人口专项抽样调查都询问了受访者的户籍、出生地、年龄和受教育程度,而这4个维度也都是非常重要的择偶标准,所以改进预测模型的第一步应当从这4个维度的匹配入手,根据调查数据对人们的择偶过程进行模拟。此外,研究者也要充分考虑在控制这4个因素后,夫妇独生子女属性之间的关联性。当然,这方面还有很多研究工作要做。例如,中国在广大农村地区长期实行的“一孩半”的生育政策和人们对男孩的偏好导致独生子女中独生子的比例通常明显高于独生女,这意味着“单独”夫妇中,丈夫是独生子的情况会明显多于妻子是独生女的一方。这在表1的频数分布中已经清楚地显示,独生子“单独”的频数约为独生女“单独”频数的两倍。此外,除了男女中独生子女比例差异的分布在婚配结果上的影响外,独生子或独生女在择偶标准和行为倾向是否存在显著差异?这两种“单独”夫妇在生育意愿和生育行为上是否存在显著差异?这些都需要继续深入研究才能回答。

此外,以往这类预测模拟通常还假定中国是一个普遍结婚的社会,但由于受长期以来出生性别比失衡的影响,很多学者认为在不久的将来,中国将会出现比较严重的对男性的婚姻挤压(李树茁等,2006;潘金洪,2007)。婚配年龄人口性别比的变化是否会对婚姻行为产生影响?对独生子女和非独生子女而言,这种影响是否存在显著差异?需要继续深入研究。

## 五、结 语

本文使用2010年中国家庭追踪调查数据,对独生子女的婚姻匹配问题进行研究,发现独生子女更可能与独生子女结婚,且非独生子女也更可能与非独生子女结婚。一方面人们在择偶时对户籍、出生地、年龄和受教育程度的多重选择会导致独生子女更可能与独生子女结婚;另一方面独生子女本身也是一个非常重要的择偶标准。以往涉及生育政策调整的人口预测中往往采用“随机婚配”假定,虽然操作简单,但这种简化既会导致对“双独”目标

人群规模的低估,又会导致对“单独”目标人群的高估。本研究只是初步的,以期抛砖引玉,推动这一方面的研究继续深入。

### 参考文献:

1. 包蕾萍、陈建强(2005):《中国“独生父母”婚育模式初探:以上海为例》,《人口研究》,第4期。
2. 风笑天(2006):《第一代独生子女婚后居住方式:一项12城市的调查分析》,《人口研究》,第5期。
3. 郭震威、齐险峰(2008):《“四二一”家庭微观仿真模型在生育政策研究中的应用》,《人口研究》,第2期。
4. 郭志刚等(2002):《现行生育政策与未来家庭结构》,《中国人口科学》,第1期。
5. 李树苗等(2006):《中国的男孩偏好和婚姻挤压——初婚与再婚市场的综合分析》,《人口与经济》,第4期。
6. 李煜、徐安琪(2004):《择偶模式和性别偏好研究——西方理论和本土经验资料的解释》,《青年研究》,第10期。
7. 李煜、陆新超(2008):《择偶配对的同质性与变迁——自致性与先赋性的匹配》,《青年研究》,第6期。
8. 李煜(2008):《婚姻的教育匹配:50年来的变迁》,《中国人口科学》,第3期。
9. 马小红(2004):《“双独政策”影响下北京市人口生育水平变动分析》,《人口研究》,第1期。
10. 齐险峰、郭震威(2007):《“四二一”家庭微观仿真模型与应用》,《人口研究》,第3期。
11. 潘金洪(2007):《出生性别比失调对中国未来男性婚姻挤压的影响》,《人口学刊》,第2期。
12. 乔晓春(2014):《“单独二胎”政策下新增人口测算方法及监测系统构建》,《人口与发展》,第1期。
13. 宋健(2001):《“四二一”结构:形成及其发展趋势》,《中国人口科学》,第2期。
14. 宋健(2010):《再论“四二一”结构:定义与研究方法》,《人口学刊》,第3期。
15. 王广州(2009):《中国独生子女总量结构及未来发展趋势估计》,《人口研究》,第1期。
16. 王广州(2012):《“单独”育龄妇女总量、结构及变动趋势研究》,《中国人口科学》,第3期。
17. 谢宇(2012):《中国家庭动态跟踪调查(2010)用户手册》(<http://www.iss.edu.cn/cfps/wd/jsbg/2010jsbg/>)。
18. 徐安琪(2000):《择偶标准:五十年变迁及其原因分析》,《社会学研究》,第6期。
19. 杨书章、王广州(2007):《一种独生子女数量间接估计方法》,《中国人口科学》,第4期。
20. 张丽萍、王广州(2014):《“单独二胎”政策目标人群及相关问题分析》,《社会学研究》,第1期。
21. 张翼(2003):《中国阶层内婚制的延续》,《中国人口科学》,第4期。
22. 中国发展研究基金会(2012):《中国发展报告 2011/2012:人口形势的变化和人口政策的调整》,中国发展出版社。
23. Blake, Judith(1989), *Family Size and Achievement*. Berkeley: University of California Press.
24. Lu, Yao and Donald J. Treiman(2008), The Effect of Sibship Size on Educational Attainment in China: Period Variations. *American Sociological Review*. Vol.73(5), 813-834.
25. Schwartz, Christine R. and Robert D. Mare(2005), Trends in Educational Assortative Marriage from 1940 to 2003. *Demography*. Vol.4(42), 621-646.
26. Wu, Xiaogang and Donald J. Treiman(2004), The Household Registration System and Social Stratification in China: 1955-1996. *Demography*. Vol.2(41), 363-384.

(责任编辑:朱 犁)