

# 性别比对女性婚姻决策的精准研究-来自 CFPS2010 年的数据

卓宇阳 李雪

(湖南大学经济管理研究中心, 湖南省、长沙市, 410000)

**摘要:** 本文利用 2010 年 CFPS 的数据采用多元线性 OLS 回归模型和 Probit 回归模型, 实证地研究了婚姻挤压即性别比对中国女性初婚决策的影响, 估算出地市级 20-34 岁年龄段的年轻人间的性别比对女性成婚概率的具体影响, 并重点关注了高学历女性的婚姻挤压问题。研究发现, 低学历女性的初婚决策易受总体性别比影响, 并推测出高学历女性的初婚决策易受高学历人群性别比的影响。其中一个可能的解释是, 受传统婚配模式影响, 就学历水平而言, 女性存在向上或者同质匹配倾向。因此我们预测出高学历女性面临的婚姻挤压程度要远远大于低学历女性面临的婚姻挤压程度。

**关键词:** 婚姻挤压; 性别比; 女性成婚概率

**中图分类号:** F069.9      **文献标识码:** A

## 1 引言

近年来, “剩男”和“剩女”问题成为一个热门话题。类似“剩女不下嫁, 农村剩男怎么办”等报道引起社会各界的广泛关注。一方面, 作为“重男轻女”思想的重灾区, 农村尤其是贫穷偏远农村男性过多, 女性供给不足, 造成性别比严重失衡, 导致了男性成婚困难。与此同时, 另一方面, 本文所关注的“剩女”问题则刚好相反, 城市中大龄“剩女”人数却在逐渐增多。“剩女”通常指“已经过了社会一般所认为的适婚年龄, 但是仍然未结婚的女性”, 一般指那些高学历、高收入的都市女性<sup>[1]</sup>。“剩女”一般是女性主动选择的结果, 说到底是因为最高一级婚姻市场优质女过多、优质男供给不足, 因找不到合适人选, 宁愿先剩着所造成的<sup>[2]</sup>。由此可见, 在农村和城市两个婚姻市场上, 婚姻挤压分别对男性初婚和女性初婚均产生了影响, 但是背后的作用机理却不完全一样。农村“剩男”存在是由于农村适婚年龄女性人数绝对量的不足, 而城市“剩女”的存在更可能是因为城市优质男性人数相对量的不足。

在目前关于“剩女”的研究中, 已有文献要么没有运用性别比进行直接的回归检验或者没有对研究对象的教育类别进行分类探讨, 同时在匹配性别比的时候单纯地以数据当年的性别比作为解释变量, 忽略了个体在结婚当年或相近几年的性别比的更为直接的影响。因此, 本文克服了以上研究局限, 贡献在于, 首先, 本文利用 2010 年 CFPS 的微观数据, 在回归中直接采用适婚年龄段的性别比数据以及受访女性出生性别比数据, 来研究其对受访女性初婚年龄和成婚概率的影响。其次, 通过在回归模型中加入性别比与教育类别的叉乘, 初步探究不同学历水平女性的初婚年龄和成婚概率对男女性别比的敏感程度, 从而对高学历女性的婚姻推迟作出相关预测。

## 2 理论综述和研究假设

### 2.1 理论综述

许多工业化国家和发展中国家的经验研究表明, 近几十年来, 女性初婚年龄呈现了明显

的上升趋势<sup>[3][4]</sup>。对于女性初婚年龄的推迟,国外的研究一般有三种解释。首先“经济独立假说”认为,女性日益增长的经济独立性减少了女性从婚姻中获得的收益,从而降低了女性结婚的激励,并且导致女性结婚年龄的推迟<sup>[5]</sup>。Blossfeld (1995) 在一些跨国的经验研究中发现,在性别角色严重分化的社会中,女性很难平衡工作和家庭,随着女性经济地位的提高,女性更有可能推迟结婚年龄<sup>[6]</sup>。其次,“职业生涯周期假说”认为,随着女性受教育程度的普遍提高,女性很难平衡学习和家庭两方面的角色,在校就读时间的延长在客观上推迟了初婚年龄<sup>[7]</sup>。最后,女性初婚年龄的推迟可能是因为婚姻市场的结构失衡导致搜寻潜在配偶的成本增加,产生婚姻挤压。另外,在微观层面上,国外也有一些学者研究证实了家庭背景和社会经济水平等其它因素,比如父母的受教育水平,父亲的职业,家庭收入,兄弟姐妹数和出生的地理位置、种族、是否服兵役等均对初婚年龄有影响<sup>[8][9]</sup>。

国内关于女性初婚年龄的研究相对较少,早期的研究围绕女性初婚初育年龄的变动情况及其基本分析展开<sup>[10][11]</sup>,从宏观因素,如人口政策与结构、社会经济发展、婚姻观念变化、社会法律法规等因素分析中国女性初婚模式的变迁。后来一些学者从微观层面研究发现,外出经历、社会网络、教育、职业、家庭的社会经济特征、户籍制度与初婚年龄均有一定的联系。研究表明,教育对于推迟初婚年龄具有显著影响,而且受教育程度对女性初婚年龄的推迟效应大于男性<sup>[12][13]</sup>。还有很大一部分学者专门研究中国的婚姻市场,认为性别比的失衡,必然引发婚姻市场的失衡,婚姻市场挤压会造成初婚年龄的增加与推迟,例如中国持续的高性别比加重了我国男性婚姻市场的挤压程度,从而进一步推动男性初婚年龄的增加,甚至有部分男性会终身不婚<sup>[14][15][16]</sup>。

## 2.2 研究假设

在一夫一妻的婚姻制度下,由于婚姻市场供需失衡,即婚姻市场可供选择的男性和可供选择的女性人数相差较大、比例失调,由此导致了男性或女性不能按传统的偏好择偶,婚姻行为发生了较大的变化,这一现象即是婚姻挤压(Marriage Squeeze)。“婚姻挤压”是超常规的性别比最直接导致的后果。Bennett, Neil G., David E. Bloom and Patricia H. Craig (1989)<sup>[17]</sup>利用1985年的Current Population Survey (CPS)数据发现白人和黑人妇女结婚率的下降通常归因于婚姻的挤压。由于婚姻挤压,女性寻找潜在配偶的时间会延长,导致初婚年龄增加。郭志刚,邓国胜<sup>[14]</sup>(2000)利用婚配性别比方法对我国1965年婚姻拥挤的测度表明,1965年我国婚配性别比为109.91,男性处于较为严重的婚姻拥挤之中。而且从分年龄婚配性别比来看,22—36岁年龄段的男性全部处于婚姻拥挤状况,其中26—32岁男性婚配性别比高达110~116,而20—34岁女性则均处于短缺之中。1988年全国2%生育节育抽样调查资料表明,我国上世纪50年代初婚男性平均初婚年龄为22.71岁,60年代初婚男性平均初婚年龄为23.32岁,男性平均初婚年龄上升了0.61岁。虽然建国初期人口文化程度的提高和工业化程度的提高等因素是男性平均初婚年龄上升的主要原因,三年自然灾害经济困难可能也进一步促进了初婚年龄的上升,但60年代初婚男性的婚姻拥挤因素也起到了一定的推动作用。基于上述理论推理,本文提出如下假设:  
假设:男女性别比会影响女性的初婚年龄。同等条件下性别比越高,女性初婚年龄越小,成婚概率越大;反之,女性初婚年龄越大,越有可能不婚。

### 3 性别比对女性成婚概率的实证研究

#### 3.1 数据来源

本文采用 2010 年中国家庭追踪调查 (China Family Panel Studies, CFPS) 的微观数据进行实证研究。CFPS 调查通过跟踪收集个体、家庭、社区三个层次的数据, 反映中国社会、经济、人口、婚姻、教育和健康的变迁, 为学术研究和公共政策分析提供数据基础。CFPS2010 年样本覆盖 25 个省/市/自治区, 代表了中国 95% 的人口, 2010 年基线调查共采访约 14960 户家庭、42590 个人, 共 162 个地级市, 是一项全国性、大规模、多学科的社会跟踪调查项目, 具有很强的代表性, 其中成人数据库样本量为 33600 个。利用 2010 年 CFPS 数据本文可以研究在地级市的水平上, 受访女性出生性别比对受访女性初婚年龄的影响以及 20—34 岁年龄段性别比对 20—34 岁女性成婚概率的影响。

#### 3.2 初婚性别比对已婚女性初婚年龄影响的实证研究

##### 3.2.1 模型设定与变量说明

利用 2010 年 CFPS 的数据, 研究受访时已婚的女性样本, 包括离异、丧偶等。经过对 2010 年 CFPS 样本的筛选, 排除受访女性出生地与现在所在地不一致的情况, 最后获得有效样本量 12270 个。设定在地级市  $j$  内女性  $i$  的初婚年龄为  $AFM_{ij}$ , 2010 年该受访女性的年龄为  $age_{ij}$ , 受教育年限为  $education_{ij}$ , 该已婚女性所在地级市  $j$  内同一年出生的男女性别比为  $birth-ratio_{ij}$ 。根据女性  $i$  的学历把教育水平  $category$  分两类: 当女性  $i$  的最高学历为专科及以上时,  $category = 1$ ; 当女性  $i$  的最高学历为高中及以下时,  $category = 0$ 。该女性 12 岁时的户口状况为  $Hukou_{ij}$ , 若为非农户口, 则  $Hukou_{ij} = 1$ ; 若为农业户口, 则  $Hukou_{ij} = 0$ 。我们利用 OLS 方法, 回归影响女性初婚年龄的各因素, 具体模型设定为:

$$AFM_{ij} = \alpha + \beta_1 birthratio_{ij} + \beta_2 category_{ij} \times birthratio_{ij} + \beta_3 age_{ij} + \beta_4 age_{ij}^2 + \beta_5 education_{ij} + \beta_6 Hukou_{ij} + \beta_7 \chi_{ij} + \varepsilon_{ij} \quad (1)$$

其中,  $AFM_{ij}$  是已婚女性初婚年龄,  $birthratio_{ij}$  是该已婚女性所在地级市  $j$  内同一年出生的男女性别比,  $category_{ij} \times birthratio_{ij}$  是教育类别和出生性别比的叉乘,  $age_{ij}^2$  是年龄的平方,  $education_{ij}$  是受访女性受教育年限,  $Hukou_{ij}$  是样本女性 12 岁时的户口状况。  $\chi_{ij}$  为其它控制变量, 包括家庭兄弟姐妹数, 父母最高受教育水平, 样本女性所在区域 (东部、西部、中部),  $\varepsilon_{ij}$  是误差项。

##### 3.2.2 相关变量描述性统计

表 1 是相关变量的描述性统计分析, 其中  $AFM$  是受访女性的初婚年龄,  $birthratio$  是样本女性所在地级市同一年出生的男女性别比。  $category \times birthratio$  是教育类别与出生性别比的叉乘, 教育类别分为 0/1 变量, 样本专科及以上 (高学历样本) 学历则定义为 1; 否则为 0。  $age$  是 2010 年样本女性受访时的年龄,  $agesq$  是年龄的平方。  $education$  是样本女性受教育年限。  $Siblings$  是受访女性家里的兄弟姐妹数, 如果样本女性家里没有兄弟姐妹, 则兄弟姐妹数定义为 0。  $pareduc$  代表父母的最高受教育水平, 本文对父母最高受教育程度的设定依照王鹏和吴愈晓<sup>[13]</sup> (2013) “初婚年龄的影响因素分析” 这篇文章对父母最高受教育程度的测量方法, 其中父母有一方学历在高中或者以上则  $pareduc = 1$ , 父母双方最高学

历均为高中以下则  $pareduc = 0$ 。Hukou 定义样本女性 12 岁时的户口状况，非农户口定义为 1，农业户口定义为 0。此外，按照受访样本所在 25 个省份划为东部、中部和西部三个地区。东部地区包括北京、天津、河北、辽宁、上海、江苏、浙江、福建、山东、广东 10 个省份；中部地区包括山西、吉林、黑龙江、安徽、江西、河南、湖北、湖南 8 个省份；西部地区包括四川、重庆、贵州、云南、陕西、甘肃、广西 7 个省份；其中，East、Central 和 West 分别代表受访女性所在地区为东部、中部和西部地区。

表 1 相关变量描述性统计  
Tab.1 Descriptive statistics

变量	观测数	均值	标准差
AFM	12270	22.05	3.09
sexratio	12270	94.16	39.84
category×birthratio	12270	4.59	21.91
age	12270	46.62	13.31
agesq	12270	2350.64	1287.83
education	12266	5.25	4.85
siblings	12270	3.15	1.88
pareduc	12270	0.28	0.45
Hukou	12194	0.15	0.35
East	12270	0.41	0.49
Central	12270	0.28	0.45
West	12270	0.31	0.46

资料来源：根据 2010CFPS 数据计算得到。

### 3.2.3 实证结果及分析

表 2 是用 STATA 12.0 软件对模型进行回归的结果，其中标准差是经过 robust standard error 处理的。回归（1）选取的控制变量包括与样本女性同一年出生的人群间的性别比、2010 年受访时的年龄和年龄的平方、样本女性最高受教育年限。回归（2）在回归（1）的基础上又增加了出生性别比和教育类别的叉乘。回归（3）在回归（2）的基础上又控制了家庭兄弟姐妹数和父母的最高受教育水平。回归（4）是控制所有变量的一个回归。虽然回归（1）的性别比系数不显著，但是符号为负，说明一般情况下出生性别比越高，样本女性初婚年龄越小，这与文章假设相符。从回归（2）到回归（4）当我们加入出生性别比和教育类别的叉乘后，我们发现，高学历即专科及以上学历的女性出生性别比的系数要小于低学历即高中及以下学历的女性的出生性别比系数，甚至高学历女性的出生性别比系数为正，这个现象有违常理值得深入探讨。

表 2 初婚性别比对样本女性初婚年龄的影响

Tab.2 The effect of birth-ratio on females' age at first marriage in prefectural level

	(1)	(2)	(3)	(4)
变量	初婚年龄	初婚年龄	初婚年龄	初婚年龄
birthratio	-0.000597	-0.00102	-0.000993	-0.000660

	(0.000652)	(0.000656)	(0.000656)	(0.000650)
category×birthratio		0.00777***	0.00737***	0.00612***
		(0.00115)	(0.00115)	(0.00113)
age	0.248***	0.252***	0.259***	0.230***
	(0.0111)	(0.0111)	(0.0118)	(0.0118)
agesq	-0.00250***	-0.00254***	-0.00261***	-0.00245***
	(0.000120)	(0.000120)	(0.000127)	(0.000127)
education	0.224***	0.210***	0.208***	0.148***
	(0.00583)	(0.00644)	(0.00649)	(0.00678)
siblings			-0.0216	0.00368
			(0.0153)	(0.0151)
pareduc			0.156***	0.0535
			(0.0597)	(0.0585)
Hukou				1.341***
				(0.0878)
East				0.795***
				(0.0584)
West				-0.160**
				(0.0674)
Constant	15.23***	15.23***	15.12***	15.82***
	(0.252)	(0.252)	(0.257)	(0.259)
Observations	12,266	12,266	12,266	12,190
R-squared	0.146	0.148	0.149	0.187

注：括号中为稳健标准误，\*、\*\*和\*\*\*分别表示在 10%、5%、1%水平上显著。

#### 4 各市县 20-34 岁年龄段性别比对女性婚姻状态影响（2010 年）的实证研究

##### 4.1 模型设定与变量说明

在以上研究出生性别比对样本女性初婚年龄影响的基础上，我们采用 CFPS 数据继续研究地级市层面 20-34 岁年龄段适婚人群间的性别比对女性成婚概率的影响。设定地级市  $k$  的某样本女性  $i$  2010 年受访时婚姻状态为  $Y_{ik}$ ，未婚（未婚、同居）则定义为 0，已婚（已婚、离异、丧偶等）则定义为 1。2010 年该受访女性的年龄为  $age_{ik}$ ，受教育年限为  $education_{ik}$ 。该女性所在地级市  $k$  相同年龄段的男女性别比为  $sexratio_{ik}$ 。该女性 12 岁时的户口状况为  $Hukou_{ij}$ ，若为非农户口，则  $Hukou_{ij}=1$ ；若为农业户口，则  $Hukou_{ij}=0$ 。本文利用多元 OLS 回归和 Probit 回归方法，回归影响女性是否成婚的各因素，回归影响女性初婚年龄的各因素，具体模型设定为： OLS 具体模型设定为：

$$Y_{ik} = \alpha + \beta_1 sexratio_{ik} + \beta_2 age_{ik} + \beta_3 age_{ik}^2 + \beta_4 education_{ik} + \beta_5 Hukou_{ik} + \beta_6 \chi_{ik} + \varepsilon_{ik} \quad (2)$$

其中， $\chi_{ik}$  为其它控制变量，包括家庭兄弟姐妹数，父母最高受教育水平，样本女性所在区域（东部、西部、中部）， $\varepsilon_{ik}$  是误差项。

Probit 的模型设定如下, 其中  $X_i$  为影响女性初婚年龄的各因素:

$$P(Y_i = 1 | X_i) = \Phi(X_i' \beta) = \int_{-\infty}^{X_i' \beta} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right) dz \quad (3)$$

## 4.2 主要变量的描述性统计

表 3 是主要变量的描述性统计结果:

表 3 主要变量描述性统计

Tab.3 Descriptive statistics

变量	观测数	均值	标准差
Marristas (婚姻状态)	3504	0.81	0.39
sexratio	3504	86.50	21.22
category × sexratio	3504	14.05	33.50
age	3504	27.36	4.28
agesq	3504	766.70	233.95
education	3504	8.34	4.62
siblings	3504	1.77	1.34
pareduc	3504	0.29	0.45
Hukou	3487	0.16	0.37
East	3504	0.41	0.49
West	3504	0.29	0.45
Central	3504	0.30	0.46

资料来源: 根据 2010 年 CFPS 数据计算得到。

表 3 具体列出样本相关变量的定义选取和描述性统计。列述的变量包括样本女性的婚姻状态, 2010 年所在地级市 20-34 岁年龄段的男女性别比, 教育类别与初婚性别比叉乘, 2010 年受访时的年龄和年龄的平方, 受教育年限, 家庭兄弟姐妹数, 父母最高受教育水平, 12 岁时的户口状况, 以及受访女性所在区域。由于计划生育政策和 90 年代超生设备的普及, 基于性别的选择性流产造成中国新生人口的性别结构失衡非常严重, 通常情况下出生性别比是大于 100 的[18]。而根据 CFPS 数据统计可知 20-34 岁年龄段即 1976-1990 年出生的各地级市男女比均值为 86.497, 这并非与我国性别结构失衡的状况有出入。事实上, 随着年龄增长男性人口所占比重逐渐下降至与女性相等甚至低于女性比例[19], 另一方面调查中可能由于很多男性外出务工或其他原因(服刑或参军等)不在调查地导致总体样本男性少于女性, 也使得 20-34 岁年龄段的男女比小于 100, 因此统计的各地区相同教育水平下的男女比均值也小于 100 (注: 根据 2010 年 CFPS 基线调查样本联系情况的数据报告显示 57155 名基因家庭成员中有 11.43% (6535 人) 的样本因外出不符合访问条件, 最主要的外出原因是外出工作 (5135 人占总样本的 8.98%), 其次是外出读书和探亲访友, 分别各占 1.70% 和 0.43%)。因变量婚姻状态 (marristas) 均值显示样本已婚率达 80.8%, 样本平均年龄为 27.36 岁, 受访女性的平均受教育年限为 8.34 年, 样本平均受教育程度在小学以上。父母最高的受教育水平 (pareduc) 显示, 父母一方最高学历在高中及以上的女性样本数仅为 28.7%, 说明父母一辈受教育水平普遍较低。样本的兄弟姐妹数 (nsiblings) 显示, 受访样本平均至少有一个兄弟姐妹, 平均兄弟姐妹数量为 1.77 个。根据访问时受访女性所在省份, 将受访女性所在地区划分为东部、中部和西部三个地区, 统计结果显示 40.9% 的女性样本处于东部地

区，28.8%的女性处于西部地区，30.3%的女性样本处于中部地区。

### 4.3 实证结果及分析

表4和表5是用STATA 12.0软件对模型分别进行OLS和Probit回归的结果，其中标准差是经过robust standard error处理的。Probit模型的回归结果与OLS回归结果高度一致，在关键变量上的回归系数也十分显著。为了方便分析，我们对表4的OLS回归结果进行简要分析阐述。

表4中回归(1)选取的控制变量包括样本女性2010年所在地级市20-34岁年龄段的男女性别比，2010年受访时的年龄和年龄的平方，受教育年限。为了考察相同年龄段性别比对不同受教育水平(高学历、低学历)女性婚姻状态的斜率影响，回归(2)在回归(1)控制变量的基础上，又增加了教育类别与性别比叉乘这个控制变量。为了考察家庭的社会经济特征比如家庭兄弟姐妹数和父母最高受教育水平对样本女性初婚年龄的影响，回归(3)增加了家庭兄弟姐妹数和父母最高受教育水平这两个控制变量。回归(4)是加入所有变量的一个全回归模型，在回归(3)的基础上新增的变量包括样本12岁时的户口状况和样本所在的区域特征。

表4 各市县20-34岁年龄段性别比对女性婚姻状态的影响

Tab.4 The effect of sexratio on females' marriage status between the age of 20-34 in prefectural level

	(1)	(2)	(3)	(4)
变量	婚姻状态 (OLS)	婚姻状态 (OLS)	婚姻状态 (OLS)	婚姻状态 (OLS)
sexratio	0.000331 (0.000269)	0.000645** (0.000266)	0.000635** (0.000267)	0.000697*** (0.000270)
category*sexratio		-0.00211*** (0.000233)	-0.00207*** (0.000236)	-0.00190*** (0.000242)
age	0.314*** (0.0193)	0.329*** (0.0192)	0.332*** (0.0192)	0.337*** (0.0193)
agesq	-0.00502*** (0.000341)	-0.00529*** (0.000339)	-0.00534*** (0.000340)	-0.00542*** (0.000341)
education	-0.0210*** (0.00127)	-0.0119*** (0.00139)	-0.0112*** (0.00146)	-0.0102*** (0.00161)
siblings			0.00526 (0.00429)	0.000663 (0.00427)
pareduc			-0.0129 (0.0130)	-0.00816 (0.0131)
Hukou				-0.0599*** (0.0187)
Central				0.0301** (0.0135)
West				0.0217 (0.0139)

Constant	-3.779*** (0.270)	-4.070*** (0.267)	-4.115*** (0.268)	-4.211*** (0.269)
Observations	3,524	3,524	3,504	3,487
R-squared	0.308	0.328	0.329	0.334

注：括号中为稳健标准误，\*、\*\*和\*\*\*分别表示在 10%、5%、1% 水平上显著。

表 5 各市县 20-34 岁年龄段性别比对女性婚姻状态的影响

Tab.5 The effect of sexratio on females' marriage status between the age of 20-34 in prefectural-level cities

变量	(1) 婚姻状态 (Probit)	(2) 婚姻状态 (Probit)	(3) 婚姻状态 (Probit)	(4) 婚姻状态 (Probit)
sexratio	0.000729 (0.00140)	0.00259* (0.00144)	0.00259* (0.00145)	0.00264* (0.00145)
category*sexratio		-0.00712*** (0.00107)	-0.00683*** (0.00108)	-0.00615*** (0.00110)
age	0.761*** (0.110)	0.821*** (0.111)	0.836*** (0.112)	0.869*** (0.113)
agesq	-0.0104*** (0.00210)	-0.0114*** (0.00211)	-0.0117*** (0.00213)	-0.0122*** (0.00214)
eduy	-0.120*** (0.00745)	-0.0774*** (0.00953)	-0.0726*** (0.00984)	-0.0676*** (0.0102)
nsiblings			0.0429 (0.0274)	0.00908 (0.0285)
pareduc			-0.0742 (0.0666)	-0.0489 (0.0682)
Hukou				-0.276*** (0.0897)
Central				0.134* (0.0724)
West				0.134* (0.0808)
Constant	-10.62*** (1.419)	-11.90*** (1.440)	-12.18*** (1.449)	-12.72*** (1.465)
Observations	3,524 (1)	3,524 (2)	3,504 (3)	3,487 (4)

注：括号中为稳健标准误，\*、\*\*和\*\*\*分别表示在 10%、5%、1% 水平上显著。

表 4 的回归结果与第一个模型研究出生性别比对女性初婚年龄影响的回归结果具有一致性。很显然，回归 (1) 到 (4) 20-34 岁年龄段的人群性别比与此年龄段低学历女性的



成婚概率具有正向相关关系。性别比越高，女性成婚的可能性越高，反之成婚的可能性越小。假设成立。回归（4）当我们控制所有变量时，尤其是在控制女性教育类别后，性别比对低学历女性的成婚概率影响显著，根据数据对低学历女性来说从同年龄段性别比均值 86.50 到最大值 200，低学历女性成婚的可能性增加约  $(200-86.50) * 0.000697 \approx 7.9\%$ 。对高学历女性来说，回归（4）中同年龄段性别比系数  $0.000697-0.0019=-0.0012$  为负，说明同等情况下，高学历女性的平均成婚率要低于低学历女性的平均成婚率。根据 2010 年的 CFPS 数据，低学历女性的平均结婚率为 85.2%，而高学历女性的平均结婚率为 57.7%，远低于低学历女性的成婚率。高学历女性来说同年龄段性别比系数为负，这个结果和前面研究出生性别比对女性初婚年龄的影响的回归模型高度一致，与常理不相符，对于这个现象，其中一个解释是高学历女性初婚年龄和成婚概率可能更多地受其它因素如同质教育间性别比的影响，而非单纯的性别比的影响。出生性别比或者 20-34 岁年龄段性别比越高，高学历女性初婚年龄越大，成婚率越小，背后的原因可能是高学历女性受教育水平较高自身条件好，因此当面临众多潜在男性配偶时，更有资本挑剔等待直到遇到最适合自己的另一半。而一般意义上的性别比对低学历女性初婚年龄和成婚概率才有着最显著最直接的影响。这个结果是独特的发现，在以后的研究中，我们可以把样本的受教育水平进行具体的分类，利用同等教育水平下的性别比进行回归分析。

## 5 讨论

当考察婚姻市场挤压或者性别比对女性初婚年龄的影响时，国际上通用的做法是研究出生性别比对女性初婚年龄的影响。本文中利用 2010 年 CFPS 数据研究与受访女性同一年出生的人群性别比（出生性别比）对样本女性初婚年龄的影响时，一方面由于数据有限男性样本移民等原因；另一方面，可能由于，在中国社会夫妻婚龄匹配主要是男大女小的模式。数据显示，家庭中丈夫年龄较妻子年龄大 1 至 4 岁是主流<sup>[20]</sup>，使得出生性别比对初婚年龄的回归系数不是很显著。对于第一个原因，我们潜在假设外出务工男性不在本地婚姻市场，如果他们返乡结婚，结果可能会存在偏误。根据已有文献发现，男性选择在婚前外出务工的比例高于女性，而且男性外出务工的次数更多，而返乡的比例却相对较低<sup>[21]</sup>。由此推断我们的假设合理，即男性外出务工或移民等原因不进入本地婚姻市场部分地造成了出生性别比对女性初婚年龄的回归系数统计上不显著的结果。

## 6 结论

女性初婚对于家庭乃至社会的福利有着重要的影响，初婚年龄尤其是女性初婚的提前或者推迟不仅会影响个人在劳动力市场上劳动力的供给和家庭的生育决策，甚至会影响到整个社会人口的结构与再生产。本文就影响女性婚姻决策（初婚年龄和成婚概率）的一个重要因素即男女性别比出发，采用微观数据进行具体的实证研究，探究性别比（婚姻挤压）对女性初婚年龄和成婚概率的定性影响关系。研究结果发现，性别比对低学历女性的初婚年龄和婚姻决策有显著直接的影响，研究结果验证了本文开始的假设，即性别比越高，低学历女性的初婚年龄越小，成婚概率越高；性别比越低，低学历女性的初婚年龄越大，成婚的可能性越低。然而对于高学历女性来说，这个结论趋于反向。考虑到受传统婚配模式的影响，女性都有向上匹配或者同质匹配倾向<sup>[22]</sup>，因此本文认为未来对女性婚姻决策的研究可以围绕把受访样本按照受教育水平分成不同类别，然后分别探究同等教育水平下的性别比对其初婚年龄和成婚概率的影响来展开，从而为我们研究城市大龄高学历“剩女”出现的原因及高学历“剩女”推迟婚姻的程度提供了一个新线索，进而为我们政府或相关部门就如何缓解“剩女”现

象提供一个参考依据。

### 参考文献

- [1] 宁鸿. "剩女"现象的社会学分析[J]. 理论界, 2008(12):222-223.
- [2] 李树茁, 姜全保, 伊莎贝尔 阿塔尼,等. 中国的男孩偏好和婚姻挤压--初婚与再婚市场的综合分析[J]. 人口与经济, 2006(4):1-8.
- [3] Manning W D. Cohabitation, Marriage, and Entry into Motherhood[J]. Journal of Marriage & Family, 1995, 57(1):191-200.
- [4] Raymo J M. Educational attainment and the transition to first marriage among Japanese women.[J]. Demography, 2003, 40(1):83-103.
- [5] Becker G S. A theory of marriage: part I. J Polit Econ[J]. 1973, 81(4):813-846.
- [6] Blossfeld and H.-P.. Changes in the Process of Family Formation and Women's Growing Economic Independence:A Comparison of Nine Countries[M].1995, Boulder,CO:Westview Press.
- [7] Blossfeld H P, Huinink J. Human Capital Investments or Norms of Role Transition? How Women's Schooling and Career Affect the Process of Family Formation[J]. American Journal of Sociology, 1991, 97(Volume 97, Number 1):143-168.
- [8] Voss P R. Social determinants of age at first marriage in the United States[J]. University of Wisconsin, 1979.
- [9] Bracher M, Santow G. Economic Independence and Union Formation in Sweden[J]. Population Studies, 1998, 52(3):275-294.
- [10] 陈友华. 中国女性初婚、初育年龄变动的基本情况及其分析[J]. 中国人口科学, 1991(5):39-45.
- [11] 郭维明. 20世纪90年代我国婚育模式的初步分析[J]. 人口学刊, 2003(5):18-21.
- [12] 刘利鸽, 靳小怡, 费尔德曼. 婚姻挤压下的中国农村男性[M]. 社会科学文献出版社, 2014.
- [13] 王鹏, 吴愈晓. 初婚年龄的影响因素分析基于CGSS2006的研究[J]. 社会, 2013(3):89-110.
- [14] 郭志刚, 邓国胜. 中国婚姻拥挤研究[J]. 人口与发展, 2000(3):2-18.
- [15] 李树茁, 姜全保, 伊莎贝尔 阿塔尼,等. 中国的男孩偏好和婚姻挤压--初婚与再婚市场的综合分析[J]. 人口与经济, 2006(4):1-8.
- [16] 姜全保, 果臻, 李树茁. 中国未来婚姻挤压研究[J]. 人口与发展, 2010, 16(3): 39-47.
- [17] Bennett N G, Bloom D E, Craig P H. The Divergence of Black and White Marriage Patterns[J]. American Journal of Sociology, 1989, 95(Volume 95, Number 3):692-722.
- [18] Wei S J, Zhang X. The Competitive Saving Motive: Evidence from Rising Sex Ratios and Savings Rates in China[J]. Journal of Political Economy, 2011, 119(Volume 119, Number 3):511-564.
- [19] 陈建伟. 教育的婚姻回报:“学得好”与“嫁得好”[J]. 上海财经大学学报, 2015(6):22-34.
- [20] 李建新, 王小龙. 初婚年龄、婚龄匹配与婚姻稳定--基于CFPS 2010年调查数据[J]. 社会科学, 2014(3):80-88.
- [21] 许琪. 外出务工对农村男女初婚年龄的影响[J]. 人口与经济, 2015(4):39-51.
- [22] 沈新风. 内生家庭谈判力与婚姻匹配[J]. 经济学:季刊, 2011, 10(4):1235-1250.

## **An accurate study of the effect of sexratio on females' marital status-data from CFPS 2010**

Zhuo Yuyang

(Center for Economics, Finance and Management Studies , Chang Sha / Hunan Province, 410000)

**Abstract:** Using CFPS data in 2010, this paper empirically studies the effect of marriage squeeze, that is sex-ratio on female's marriage decision-making, like age at first marriage and the probability of getting married on prefectural level for people aged 20-34 year old, by using multiple OLS regression model and Probit model. On this basis, this paper mainly focus on marriage squeeze for higher educated women. The paper finds out that low educated females' marriage decisions are more affected by overall sex ratios, and then predicts that higher educated females' marriage decisions are more affected by sex-ratios among the people of higher education. For this prediction, one explanation is that influenced by the traditional marriage matching patterns, females tend to match up or have a homogeneous match with males. Thus the paper also predicts that the degree of marriage squeeze higher educated females face may be far greater than that of lower educated females may face.

**Keywords:** marriage squeeze; sex-ratio; female marriage probability