

China Family Panel Studies



中国家庭追踪调查

技术报告系列: CFPS-37

系列编辑: 谢宇 责任编辑: 赵逸文

中国家庭追踪调查  
家庭社会经济地位综合变量: 父亲和母亲的  
出生年、最高学历、政治面貌和主要职业

张春泥、叶华、李汪洋、马超、戴利红、胡婧炜、王祎睿、谢宇

2020.11

父母的受教育程度和职业地位反映了家庭社会经济地位与子女的成长环境,是研究社会地位的获得与代际传递、社会流动、个人生命历程等命题不可或缺的变量。通常,社会科学调查会以受访成年子女代答的方式来采集其父母的受教育程度、职业背景、政治面貌等信息,例如“您父亲的最高受教育程度是(包括目前正在读的)”。大多数情况下,受访人都能回答出自己父母的基本信息,缺失值较少。而且,子女作为受访人代答的父母受教育程度、职业、政治面貌等作为本人的背景性变量,在以受访人为研究对象的统计分析中使用起来较为便利。

2010年CFPS基线调查采用了与以往调查不同的形式采集父母的人口及社会经济地位信息:在受访家庭中选定一名成员回答每一位家庭基因成员的父母基本信息(记录在T1表或T3表),并将信息保存在家庭关系库中。研究者需要匹配个人库与家庭关系库才能将父母的基本信息与作为子女的个人数据相联系。理论上,这一采集方式突破了父母不健在、父母未受访、以及父母不同住对信息采集的限制。但是,在实际调查过程中,由于选定的代答人可能不了解其他家庭成员的父母情况,或是受访人或访员对回答要求的理解有偏差,导致漏答或错答不同住或不健在父母的基本信息,造成父母学历、职业等变量取值存在严重缺失,无法满足研究需求。

为此,在2012年的追访中,我们采用成年子女受访人代答的方式,在个人问卷中重新采集了父母的出生年、最高学历、职业和政治面貌。我们将新采集的数据与原有数据相比较、整合,生成父母的出生年(fbirthy, mbirthy)、父母受教育程度(feduc, meduc)、父母政治面貌(fparty, mparty)、父母主要职业(foccup\*, moccup\*)的新综合变量。本报告将介绍这些综合变量的生成规则,以供用户参考。

## 一、基本生成规则

由2010年的基线调查和2012年的追访共提供了父母信息的三个来源(表1)。第一个来源是2010年家庭成员问卷中,选定受访人回答的家中所有基因成员的父亲和母亲的姓名、出生日期(/年龄/属相)、性别、是否健在、最高学历、主要职业、婚姻状况、现居地等信息。所有的变量信息保存在2010年家庭关系库。设计上,这一来源应该采集所有基因家庭

成员的父母信息，但实际操作中，代答信息的缺失情况较严重，其中已经去世的父母信息缺失尤其严重。第二个来源是 2012 年成人问卷 V 部分对受访人采集其父母的出生年（/年龄/属相）、14 岁时父母的最高学历、主要职业和政治面貌。这些信息保存在 2012 年成人库。在一般情况下，2012 子女代答自己父母的信息应该不会比 2010 任意家庭成员代答的信息质量差，但是，如果子女没有接受 2012 年访问，我们就无法重新采集其父母的信息。而且，2012 年问卷在重新采集父母信息时没有加载父母的姓名，这意味着，如果同时存在生父/母、继父/母、或养父/母的情况，子女代答的父/母有可能跟 2010 年家庭关系库中采集的父/母不是同一人。第三个来源是个人问卷。设计上，CFPS 对家庭的所有成员都要进行个人访问，如果受访人的父母在居住上和经济上与受访人没有分开，那么他们也会回答 2010 年和 2012 年的个人问卷，其教育经历、政治组织的参与情况、职业和工作等信息都会被详尽采集。但是，父母与受访人同住且接受了个人访问的情况很有限，在家庭关系库中仅占 38%(父)和 40%(母)。因此，用父母自答的个人信息也很难补全、更正第一、第二来源代答的父母信息的缺失和错误。

以上的介绍反映出三个数据来源各有优劣，将三者整合在一起，可以相互补充、校验。在整合数据及生成综合变量时，我们制定的基本规则是：首先，以家庭关系库的代答信息为基础，因为该来源涵盖的成员最广。接着，用子女代答的信息来补充家庭关系库的缺失值和替换不一致的取值，我们假定子女代答其父母的信息比任意家庭成员的代答信息更准确。然后，用父母回答的 2010 或 2012 年个人问卷信息来补充前两个来源的缺失值和替换不一致的取值，我们假定父母本人自答的信息比子女或任何人代答的信息更准确。

需要注意的是，虽然我们通过三个来源的数据生成综合变量，但仍可能存在缺陷。缺陷主要有以下几种：

1. 三个来源的数据均缺失；
2. 当子女同时有亲生父母、继父母或养父母时，子女代答的父/母与家庭关系库中的父/母可能不是同一人；
3. 如果受访人有同是家庭基因成员的同父/同母的兄弟姐妹，但这些兄弟姐妹由于种种

原因没有接受个人访问，而且其父/母也没有个人问卷，则这些兄弟姐妹在家庭关系库中的父/母信息没有与该受访人的父/母信息同时补充或更正。

4. 由于瞒报、误报，有可能出现父母自答的信息不如代答信息真实、准确的情况，在这些情况下我们设置的不一致取值的替换规则未必是最优的。

对于第一、第三种缺陷，用户可以用多重插补处理缺失值。对于第二、第四种缺陷，用户可以自行进行数据清理。

接下来，我们将分别介绍父母的出生年、最高学历、政治面貌和主要职业的综合变量生成过程及结果。

表 1: CFPS2010~2012 父母出生年、最高学历、政治面貌和主要职业的数据来源及比较

数据来源	问卷位置	数据库变量名	优点	缺点
成员问卷代答	2010 年家庭成员问卷 T1 和 T3 表	出生年/年龄: tb1y_a_f 或 tb1b_a_f (父); tb1y_a_m 或 tb1b_a_m (母) 最高学历: tb4_a_f(父); tb4_a_m (母) 主要职业: tb5_code_a_f (父); tb5_code_a_m (母)	收集所有基因家庭成员的父母基本信息。设计上不局限于父母是否同住、是否健在, 和个人问卷的拒答。	- 如父母已去世, 该变量缺失值多; - 代答人无法回答或错答其他成员父母的情况。
子女代答	2012 年成人问卷 V 部分	出生年: qv101a 或 qv101c (父); qv201y 或 qv201b (母) 最高学历: qv102(父); qv202(母) 政治面貌: qv104(父); qv204(母) 职业: qv103 (父); qv203 (母)	- 子女代答自己父母的情况很可能比其他家庭成员的代答准确。 - 不局限于父母是否健在、父母是否同住。	- 如果子女受访者没有接受 2012 年个人问卷的访问, 则该信息缺失。 - 如果某受访者有同是基因成员的同父/同母的兄弟姐妹, 而这些兄弟姐妹的个人问卷缺失, 则这些兄弟姐妹的父/母信息难以与该受访者的父/母信息同时更正。 - 如果存在生父/母、继父/母、养父/母的情况, 子女代答的父/母有可能跟家庭关系库中的父/母不是同一人。
父母本人自答	2010 年成人问卷、2012 年成人问卷	2010 年最高学历综合变量: cfps2010edu_best 政治面貌: qn401_s_1~14(2012); qa7_s_1~14(2010); sn401(2012) 职业: qg307code (2010); qg411code_a_1~10、qg510code_a_1~10、qg609code_a_1~10 (2012)	- 自答信息比代答信息更准确。	- 父/母不是同住家庭成员 (包括去世), 或者没有完成个人问卷的访问就没有该信息。 - 父/母的最高学历和政治面貌在被访者 14 岁后有可能发生变化, 因此自答与子女代答的经历所对应的时间可能不一致。

## 二、父亲和母亲的出生年

父/母的出生年有三个数据来源：2010 年家庭问卷中由指定成员代答的所有家庭成员的父亲和母亲的出生年和年龄；如果父/母为同住家庭成员，且接受了 2010 年或 2012 年个人问卷的访问，则 2010 年或 2012 年的成人库中有父母作为受访人回答的本人出生年；2012 年调查对成人受访者提问了其父亲和母亲的出生年、属相和年龄。

根据上述数据及生成规则，我们按如下步骤生成父亲和母亲出生年的综合变量并在表 2 中展示各步骤完成之后的基本情况：

(1) 以 2010 年家庭关系库的父/母出生年 (tbly\_a\_f 和 tbly\_a\_m) 为原始值，当该变量取值为“不适用/不知道/拒答”时，用 2010 年减去父/母年龄计算出的出生年替代，如果父/母年龄取值也为“不适用/不知道/拒答”，则设为缺失。

(2) 用个人编码 pid 将 2010 年家庭关系库与 2012 年成人库匹配，用 2012 年个人问卷访问时受访子女代答的父母出生年 qv101a 和 qv201y 来补充(1)中的缺失值及替换不一致的取值。在这一步，我们假定子女代答的信息要比任意家庭成员代答信息更准确。如果 qv101a 和 qv201y 缺失，我们则用 2012 年减去子女代答父/母的年龄(qv101c 和 qv201b) 来推算父/母出生年。

(3) 用经 2010 和 2012 年成人库数据修正过的本人最佳出生年 (qa1y\_best 和 cfps2012\_birthy) 来补充、替换本人在家庭关系库中作为他人父/母时的出生年。这一步，我们假定父/母本人的自答信息比子女或其他人代答的信息更准确。

表 2 比较了 2010 年家庭关系库中父/母出生年的原始值(tbly\_a\_f 和 tbly\_a\_m) 和综合变量取值 (fbirthy 和 mbirthy) 的差别。经补充和修正，父亲出生年的缺失比例从 49.4%降至 14.3%，母亲出生年的缺失比例从 49%降至 13.8%。修正后父/母出生年的均值和中位数均明显降低，这反映出原始变量的取值缺失很大程度上是由于父母过世而导致的漏报。

表 2：父母出生年修正前后各步骤缺失样本及描述性统计值 (N= 57,155)

	父亲出生年			
	[tb1y_a_f]	STEP1	STEP2	STEP3: [fbirthy]
无缺失的样本数	28915	44261	49006	49008
均值	1957	1949	1947	1947
标准差	18.3	21.0	21.5	21.5
中位数	1962	1951	1949	1949
	母亲出生年			
	[tb1y_a_m]	STEP1	STEP2	STEP3: [mbirthy]
无缺失的样本数	29161	44065	49282	49288
均值	1958	1952	1949	1949
标准差	18.1	20.5	21.0	21.0
中位数	1963	1953	1951	1951

### 三、父亲和母亲的最高学历

父/母的最高学历有三个数据来源：2010 年家庭问卷由指定成员代答的所有家庭成员的父亲和母亲的最高学历；如果父/母为同住家庭成员，且接受了 2010 年或 2012 年个人问卷的访问，则 2010 年或 2012 年的成人库中有父母作为受访人自答的最高学历或离校阶段；2012 年调查对成人受访者提问了其 14 岁时父亲和母亲的最高学历。

据此，我们以如下步骤生成父/母最高学历的综合变量：

(1) 以 2010 年家庭关系库的 tb4\_a\_f 和 tb4\_a\_m 为原始值，这两个变量取值为“不适用/不知道/拒答”时视作缺失。

(2) 用个人编码 pid 将 2010 年家庭关系库与 2012 年成人库匹配，用 2012 年调查子女代答的父母最高学历 qv102 和 qv202 来补充 tb4\_a\_f 和 tb4\_a\_m 的缺失及替换不一致的取值。在这一步，我们假定子女代答的信息要比任意家庭成员代答信息更准确。

(3) 用经更正过的 2010 年个人最高学历来补充、替换此人在家庭关系库中作为他人父/母时的最高学历。这一步，我们假定父/母的自答信息比子女或其他人代答的信息更准确。个人最高学历的更正方案请参见黄国英等（2014）的技术报告。

经上述步骤，我们生成了父亲最高学历的综合变量 feduc 和母亲最高学历的综合变量 meduc。

我们在表 3 中比较了新生成的综合变量和家庭关系库中父母最高学历原始变量 (tb4\_a\_f 和 tb4\_a\_m) 取值的差异。表 3 显示，父母最高学历原始变量的缺失值的比例（包括不适用、拒答、不知道）在四成左右。修正后的综合变量的缺失值比例则降至 16.8%（父）和 13.4%（母）。原始变量的绝大多数缺失值分布在低教育程度上，这也反映出取值缺失很大程度上是由于对过世成员信息的漏报所致。

**表 3：父母最高学历变量修正前后取值对比（N= 57,155）（单位：%）**

	父亲最高学历			
	[tb4_a_f]	STEP1	STEP2	[feduc]
缺失(=.)		43.7	16.9	16.8
- 不适用(=-8)	41.9			
- 拒绝回答(=-1)	0.0			
- 不知道(=-2)	1.8			
文盲/半文盲(=1)	14.4	14.4	30.2	31.4
小学(=2)	14.6	14.6	21.8	21.8
初中(=3)	17.6	17.6	19.8	18.9
高中(=4)	7.0	7.0	8.3	8.2
大专(=5)	1.6	1.6	1.7	1.8
大学本科(=6)	1.0	1.0	1.2	1.2
硕士(=7)	0.1	0.1	0.1	0.1
博士(=8)	0.0	0.0	0.0	0.0
合计	100.0	100.0	100.0	100.0
	母亲最高学历			
	[tb4_a_m]	STEP1	STEP2	[meduc]
缺失(=.)		37.0	13.5	13.4
- 不适用(=-8)	35.6			
- 拒绝回答(=-1)	0.0			
- 不知道(=-2)	1.4			
文盲/半文盲(=1)	29.3	47.9	47.9	50.5
小学(=2)	14.2	18.0	18.0	16.8
初中(=3)	13.4	14.0	14.0	12.8
高中(=4)	4.4	4.8	4.8	4.7
大专(=5)	1.1	1.1	1.1	1.1
大学本科(=6)	0.6	0.6	0.6	0.6
硕士(=7)	0.0	0.0	0.0	0.0
博士(=8)	0.0	0.0	0.0	0.0
合计	100.0	100.0	100.0	100.0

## 四、父亲和母亲的政治面貌

2010 年家庭问卷没有采集父母的政治面貌，因此，父/母的政治面貌只有两个数据来源：如果父/母为同住家庭成员，且接受了 2010 年或 2012 年个人问卷的访问，则 2010 年或 2012 年的成人库中有父母作为受访人自答的调查当时的社会组织参与情况；2012 年调查对成人受访者直接提问了其 14 岁时父亲和母亲的政治面貌。

据此，我们以如下步骤生成父/母政治面貌的综合变量：

(1) 用个人编码 pid 将 2010 年家庭关系库与 2012 年成人库匹配，用 2012 年调查子女代答的其 14 岁时父/母的政治面貌，即 qv104 和 qv204，作为父/母政治面貌的初始取值。取值有四个：1=共产党员；2=民主党派；3=共青团员；4=群众。

(2) 利用 2010 年和 2012 年成人问卷社会组织参与情况（“您参加了以下哪些组织”）生成受访人的政治面貌变量。该变量的取值及编码与(1)中的定义相同。其中，不属于共产党员、民主党派或共青团员者被定义为“群众”。由于社会组织的参与是多选题，可能有人回答既加入了共产党、也加入了民主党派或共青团。我们规定，这些人被认定为共产党员优先于被认定为民主党派和共青团员，被认定为民主党派优先于被认定为共青团员。

(3) 用(2)中生成的个人政治面貌来补充、替换此人在家庭关系库中作为他人父/母时的政治面貌。这一步，我们假定父/母的自答信息比子女代答的信息更准确。需要说明的是，(2)中生成的个人政治面貌是调查当时的情况，而(1)中子女填报的是他们 14 岁时父母的政治面貌。由于人在中年以后的政治组织参与仍可能发生变化，因此有可能父/母在其子女 14 岁之后政治面貌还发生过改变，对此，我们用父/母作为受访人时回答的加入共产党的年份对数据进行了调整。如果父母的入党年龄是在 45 岁以后，我们猜测这些父母在其子女 14 岁时可能是“群众”，因此，将这些父母在综合变量中的取值改为“群众”。

表 4 展示了利用个人库匹配父母后得到的父母政治面貌和 14 岁时父亲和母亲政

治面貌的综合变量（fparty 和 mparty）的分布。由于我们只有两个可以利用的数据来源，在生成综合变量之后，父/母政治面貌仍分别有 24.5%和 22.6%的缺失。对于这些缺失值，我们建议用户可以将之认定为“群众”或者做多重插补。

**表 4：14 岁时父/母政治面貌 单位：%**

	父亲政治面貌			
	个人库 (N= 35729)	STEP1 (N= 57,155)	STEP2(CFPS2010、 CFPS2012 个人数据) (N= 42926)	[fparty] (N= 57,155)
缺失(=.)	66.0	46.2	6.8	24.5
共产党员(=1)	3.7	7.3	0.1	9.6
民主党派(=2)	0.0	0.1	15.7	0.1
共青团员(=3)	3.4	0.3	72.1	3.6
群众(=4)	26.8	46.2	5.3	62.2
合计	100.0	100.0	100.0	100.0
	母亲政治面貌			
	个人库 (N= 35729)	STEP1 (N= 57,155)	STEP2(CFPS2010、 CFPS2012 个人数据) (N= 42926)	[mparty] (N= 57,155)
缺失(=.)	63.4	45.5	6.8	22.6
共产党员(=1)	0.9	1.1	0.1	1.6
民主党派(=2)	0.0	0.0	15.7	0.1
共青团员(=3)	3.3	0.2	72.1	3.4
群众(=4)	32.4	53.2	5.3	72.3
合计	100.0	100.0	100.0	100.0

## 五、父亲和母亲的主要职业

父/母的主要职业有三个数据来源：2010 年家庭问卷由指定成员代答的所有家庭成员的父亲和母亲的主要职业；如果父/母为同住家庭成员，且接受了 2010 年或 2012 年个人问卷的访问，则 2010 年或 2012 年的成人库中有父母作为受访者自答的主要职业；2012 年调查对成人受访者提问了其 14 岁时父亲和母亲的主要职业。

据此，我们以如下步骤生成父/母主要职业的综合变量：

根据《1999年中国标准职业分类与代码》(GB/T6565-1999)，我们首先对2012年调查子女代答的其14岁时父/母的主要职业信息(qv103、qv203)进行编码，生成父/母主要职业的代码，分别是qv103code(父)和qv203code(母)，作为父/母主要职业的初始取值。

其中，部分受访者14岁时父/母主要职业的原始回答过于笼统、模糊，例如“工人”、“技术工”，而无法判断出其所属的行业和职业类别，难以进行职业分类和编码。这部分的样本量分别是563(父)、443(母)。<sup>1</sup>在编码中，我们赋予它们一个临时代码，分别是99998(工人)、99999(技术工人、技工、技术员、操作工)。

对于(1)中存在缺失的情况，我们采用个人编码pid将2010年成人库与2012年成人库匹配的方法，用2010年成人库中父/母作为受访人自答的主要职业qg307code来补充2012年调查子女代答的其14岁时父/母的主要职业。<sup>2</sup>

需要说明的是，(1)中调查的是受访者14岁时父/母的主要职业，而2010年成人库中本人的主要职业是调查当时的情况。父/母在其子女14岁之后的职业很有可能发生过改变。(2)的主要目的并非是要识别出具体的职业类别，而是判断其是否是技术工人。我们假定：若父/母作为受访者在2010年调查时的主要职业属于技术工人或管理阶层的范畴，其子女14岁时的主要职业很有可能已经是技术工人；否则为非技术工人。

通过父母作为受访者在2010年时的主要职业识别其子女14岁时的职业时，我们依据的是EGP职业分类代码。具体来说，若EGP的取值小于等于8，则上述因信息模糊而无法分类的14岁时父/母的主要职业被归类为技术工人；若EGP的取值在9-11之间，则归类为非技术工人<sup>3</sup> 据此，我们给出技术工人和非技术工人的特殊职业编码，

---

<sup>1</sup> 这里针对的是受访者给出“工人”、“技术工”等回答的个案数，而不包括80000“无职业者分类及代码”和90000“不便分类的其他从业人员”。

<sup>2</sup> 2012年成人库同样是父/母主要职业的数据来源之一。但是，与2010年调查直接询问受访者主要职业的方式不同，2012年调查将工作分为自家农业、农业打工、受雇、非农自雇和家庭帮工等5类，针对每一类型下的多份工作进行了提问。在此，我们并没有以2012年成人库受访者的职业信息来补充。

<sup>3</sup> EGP职业分类共11类，分别是higher service, lower service, routine clerical/sales, manual foremen, skilled manual, semi-unskilled manual, farm workers, and farmers/farm managers，依次取值是1-11。

分别是 99998 和 99999。

用 2010 年家庭关系库中父/母的主要职业 `tb5_code_a_f` 和 `tb5_code_a_m` 来补充 2012 年调查子女代答的其 14 岁时父/母的主要职业，区分出技术工人和非技术工人。这一步骤的目的和 (2) 相同。<sup>4</sup>

若 (2) 和 (3) 还无法区分技术工人和非技术工人，则根据父/母的最高学历来生成主要职业的综合变量。在这一步，我们以父母的最高学历 `feduc` 和 `meduc` 为基准，若父/母的最高学历高于平均水平（受教育年限的中位数），则归为技术工人；相反，则为非技术工人。<sup>5</sup>

(5) 对于 (4) 之后仍缺失的情况，我们用 2010 年家庭关系库中父母的职业 `tb5_code_a_f` 和 `tb5_code_a_m` 填补。

经过上述步骤，我们生成了子女 14 岁时父亲主要职业的综合变量 `foccupcode` 和母亲主要职业的综合变量 `moccupcode`。这里使用的是基于《1999 年中国标准职业分类与代码》的 CFPS 职业分类代码，同时增加了两个特殊分类和代码：技术工人 (99998) 和非技术工人 (99999)。之后，我们将父母职业的国标码 (CSCO) 转化为 1988 年的国际标准职业编码 ISCO88，得到父母主要职业 ISCO88 综合变量：`foccupisco` 和 `moccupisco`。

表 5 展示了父亲和母亲在家庭关系库中的原始变量 (`tb5_code_a_f` 和 `tb5_code_a_m`) 和经过重新收集数据、编码、用各来源数据填补缺失值后的主要职业国标码的综合变量 (`foccupcode` 和 `moccupcode`) 的分布。经过 2012 年的数据收集和我们的清理，父/母主要职业的缺失值比例下降了将近 50%。

---

<sup>4</sup> 2010 年成人库提供了本人的 EGP 职业分类代码，而 2010 年家庭关系库没有提供父/母主要职业的 EGP 职业分类。EGP 职业分类代码是基于 ISCO88 职业分类所得到的，部分职业还需要考虑雇佣状态和管理地位，但 2010 年家庭问题并没有采集父/母的雇佣状态和管理地位的信息。不过，考虑到在上述因信息模糊而无法分类的 14 岁时父/母的主要职业个案中，仅有涉及父亲主要职业的 3 个个案需要考虑雇佣状态和管理地位，对 EGP 职业分类代码的生成不会产生太大的影响。在此，我们只执行从 ISCO88 到 EGP 的第一步，得到一个不包括自雇类的 EGP 职业分类。

<sup>5</sup> 此处的平均受教育水平是基于 2010 年成人库计算的分性别的受教育年限的中位数。

表 5: 父/母主要职业 单位: %

	父亲主要职业			
	STEP1:qv103 (N=35719)	STEP2:qv103 结合 qg307code (N=35719)	基于 [tb5_code_a_f] (N=57,155)	基于 [foccupcode] (N=57,155)
缺失(=)				23.0
- 不适用(=-8)	19.5	17.6	56.3	
- 无法编码(-7)	0.0		0.0	
- 拒绝回答(=-1)	0.0	0.0	0.0	
- 不知道(=-2)	0.0		0.3	
国家机关、党群组织、 企业、事业单位负责人	5.7	5.8	0.8	4.3
专业技术人员	4.3	4.4	1.3	3.5
办事人员和有关人员	1.4	1.5	1.4	1.8
商业、服务业人员	2.7	2.9	2.4	3.7
农、林、牧、渔、水利 业生产人员	50.7	53.8	20.8	43.0
生产、运输设备操作人 员及有关人员	11.8	12.4	8.4	13.8
军人	0.4	0.4	0.0	0.3
无职业者分类及代码	1.0	1.0	2.9	1.9
不便分类的其他从业人 员	1.7	1.8	5.4	4.4
技术工人(99998)	0.5	0.5		0.3
非技术工人(99999)	0.3	0.3		0.2
合计	100.0	100.0	100.0	100.0
	母亲主要职业			
	STEP1:qv203 (N=35719)	STEP2:qv203 结合 qg307code (N=35719)	基于 [tb5_code_a_m] (N=57,155)	基于 [moccupcode] (N=57,155)
缺失(=)				28.5
- 不适用(=-8)	24.2	19.9	60.1	
- 无法编码(-7)	0.0		0.1	
- 拒绝回答(=-1)	0.0		0.0	
- 不知道(=-2)	0.0		0.3	
国家机关、党群组织、 企业、事业单位负责人	1.7	1.7	0.2	1.2
专业技术人员	1.4	1.5	1.2	1.8
办事人员和有关人员	0.5	0.5	0.5	0.8
商业、服务业人员	1.7	1.9	2.9	3.6
农、林、牧、渔、水利 业生产人员	58.7	62.3	24.3	50.5
生产、运输设备操作人 员及有关人员	3.9	4.1	2.4	4.4

军人	0.0	0.0	0.0	0.0
无职业者分类及代码	6.2	6.2	4.9	6.1
不便分类的其他从业人员	1.1	1.1	3.1	2.7
技术工人(99998)	0.5	0.5		0.3
非技术工人(99999)	0.2	0.2		0.2
合计	100.0	100.0	100.0	100.0

经过 2012 年对受访人父母社会经济地位信息的重新收集、数据清理、整理、各来源数据的比对、相互补充，我们得到如下家庭社会经济地位综合变量（见表 6），并将这些变量均储存于 2010 年家庭关系数据库，以供用户使用。

**表 6：家庭社会经济地位综合变量一览表**

变量名	变量标签
fbirthy	父亲出生年
mbirthy	母亲出生年
feduc	父亲最高学历
meduc	母亲最高学历
fparty	父亲政治面貌
mparty	母亲政治面貌
foccupcode	父亲主要职业编码（国标）
foccupisco	父亲主要职业 ISCO88
moccupcode	母亲主要职业编码（国标）
moccupisco	母亲主要职业 ISCO88