

China Family Panel Studies



中国家庭追踪调查

技术报告系列: CFPS-30

系列编辑: 谢宇 责任编辑: 张聪

中国家庭追踪调查
2012年当前主要职业综合变量的建构

李汪洋 胡婧炜 谢宇 吴琼

2014.12.19

CFPS 在不同年份调查中关于工作模块的设计思路和问题并不一致。2010 年和 2011 年直接提问了受访者当前最主要的一份工作；在 2012 年，CFPS 不仅打算收集调查期间全部工作的情况，而且针对每一类工作设计了个性化的问题。这一设计虽然为研究者提供了尽可能详尽的工作信息，但也给用户的使用带来了一定的困难。最主要的问题是难以确定受访者当前最主要的工作，并生成一个独立的主要职业变量。同时，这也导致无法实现不同年份之间职业流动的比较和研究。因此，本报告的目的是介绍 CFPS2012 年工作模块的设计和生成一个独立的职业变量。简要来说，CFPS 2012 是通过具体的工作单位名称来标识每一份工作的，在生成主要职业变量时有两个步骤：首先，确定受访者当前主要工作单位名称；其次，根据主要工作单位名称来定位到该项工作的职业名称及代码。

一、CFPS 不同年份工作模块不相匹配

正如前文所述，CFPS 2012 在工作模块的设计上与以往年份不同，既难以确定受访者的主要工作，也无法实现工作的历时性比较。在 2010 年和 2011 年，CFPS 直接提问了受访者当前最主要工作的信息。如在 2010 年调查问卷共用模块中，受访者的职业问题如下：

G306 您的职业是？

访员注意：（1）如果受访者有多份工作，请询问占用时间最多的工作；

（2）请详细记录受访者的主要工作，填写具体内容：工作部门+工作职责/工作内容+工作岗位或工种名称。

可见，CFPS 在 2010 年主要工作的界定有三点：（1）涵盖务农和非农职业；（2）目前正在从事；（3）占用时间最多。然而，CFPS 2012 工作模块的基本思路是：首先依次提问受访者过去一年是否从事自家农业生产经营、是否从事农业打工工作、是否从事非农受雇工作、是否从事个体/私营经济、是否从事帮工活动；若受访者在任一类类型下回答“是”，接下来会提问该类工作的份数，然后对每份工作启动该类型的工作问题。CFPS2012 主要职业的数据来源见表 1。

表 1. CFPS 2012 主要职业的数据来源

数据来源	问卷位置	职业领域	题目	数据库变量名
成人自答	共用模块	自家农业生产活动	G201 过去一年，您有没有为自家从事农业生产经营活动？ G202 过去一年里，您有几个月为自家从事农业生产经营活动？ G203 过去一年里，农忙的时候，您一般每月有几天为自家从事农业生产经营活动？ G204 过去一年里，农忙的时候，您一般每天有几个小时为自家从事农业生产经营活动？ G205 过去一年里，非农忙期间，您一般每月有几天为自家从事农业生产经营活动？ G206 过去一年里，非农忙期间，您一般每天有几个小时为自家从事农业生产经营活动？	是否为自家从事农业生产： qg201 从事农业生产的月数：qg202 农忙时每月农业生产天数： qg203 农忙时每天农业生产时长： qg204 非农忙时农业生产天数：qg205 非农忙时农业生产时长：qg206
10-15 岁 少儿自答		农业打工	G301 过去一年里，您有没有至少 10 天在为其它农户做农活/打散工挣钱？ G303 过去一年里，您为其它农户做农活/打散工挣钱总共多少天？ G304 过去一年里，您为其它农户做农活/打散工挣钱时，平均每天工作多少小时？	是否为其他农户打散工赚钱： qg301 为其他农户打工天数：qg303 每天打工时间：qg304
		受雇	G401 过去一年，您是否从事过挣工资的非农工作？ G402 过去一年，您总共从事过几份挣工资的非农工作？ G403 您这几份工作单位名称分别是？ G404 我们记录您上次调查时最主要的工作是否正确？ G405 您上次调查时最主要的主要工作单位名称？ G406 除当时的那份工作外，您总共还从事过几份挣工资的非农工作？ G407 您这几份工作单位名称分别是？	正在从事受雇工作总数： jobbcn 受雇工作单位名称： jobbname_a_1...N 受雇工作持续时间： jobblastdate_a_1...N 受雇工作的月工作天数： qg413_a_1...N 每天工作时间：qg414_a_1...N

		<p>G4122 您这份工作持续到什么时候？</p> <p>G413 过去一年里，您这份工作一般每月工作几天？</p> <p>G414 过去一年里，您这份工作一般工作几个小时？</p> <p>G411 您在这份工作中具体做什么事情？</p>	<p>职业描述： qg411_a_1...N</p> <p>职业编码： qg411code_a_1...N</p>
	非农自雇	<p>G501 我们记录您上次调查时候最主要的工作是从事自己经营的工作，您当时的工作属于以下哪一类？</p> <p>G502 过去一年，您是否从事过个体或私营经济活动？</p> <p>G503 过去一年，您从事了几项个体或私营经济活动？</p> <p>G504 您从事的这几项个体或私营经济活动的公司/生意名称分别是？</p> <p>G505 您那份工作的公司/生意名称是？</p> <p>G506 除当时的那份工作外，从上次调查到现在，您总共还从事过个体或私营经济活动？</p> <p>G507 您从事的这几项个体或私营经济活动的公司/生意名称分别是？</p> <p>G5102 您这项生意持续到什么时候？</p> <p>G511 过去一年里，您这项个体或私营经济活动干了几个月？</p> <p>G512 过去一年里，在您从事这项个体或私营经济活动时，平均一个月工作几天？</p> <p>G513 过去一年里，在您从事这项个体或私营经济活动时，一般每天工作多少个小时？</p> <p>G510 您具体从事什么工作？</p>	<p>正在从事非农自雇总数： jobc1cn</p> <p>非农自雇工作单位名称： jobc1name_a_1...N</p> <p>非农自雇持续时间： jobc1lastdate_a_1...N</p> <p>过去一年从事非农自雇月数： qg511_a_1...N</p> <p>非农自雇的月工作天数： qg512_a_1...N</p> <p>非农自雇的日工作时长： qg513_a_1...N</p> <p>职业描述： qg510_a_1...N</p> <p>职业编码： qg510code_a_1...N</p>

		家庭帮工	<p>G601 过去一年，您是否不拿工资为家庭经营活动 帮工？</p> <p>G602 过去一年，您不拿工资为几项家庭经营活动帮工过？</p> <p>G604 您上次调查时最主要的工作为家庭经营活动帮工，拿分工作的公司/生意名称是？</p> <p>G605 除当时那份工作外，从上次调查时到现在，您总共还从事过几份家庭经营帮工活动？</p> <p>G6102 您参与这项生意持续到什么时候？</p> <p>G609 您具体从事什么工作？</p>	<p>正在从事家庭帮工总数： jobc2cn</p> <p>家庭帮工持续时间： jobc2lastdate_a_1...N</p> <p>职业描述： qg609_a_1...N</p> <p>职业编码： qg609code_a_1...N</p>
--	--	------	--	---

此外，CFPS2012 还设有 G7 最主要工作部分，在调查时生成了“主要工作单位名称 (job2012mn)”的变量。生成规则如下：(1) 若受访者正在从事的受雇工作、非农自雇工作、家庭帮工的总数为 0 ($JobBCN+JobC1CN+JobC2CN=0$)，则 job2012mn 为最近结束一份工作的工作单位名称 (JobMRName)；(2) 若受访者正在从事的受雇工作、非农自雇工作、家庭帮工的总数为 1 ($JobBCN+JobC1CN+JobC2CN=1$)，则 job2012mn 为当前正在从事的该项工作对应的工作单位名称 (JobBnameX 或 JobC1nameX 或 JobC2nameX)；(3) 若受访者正在从事的受雇工作、非农自雇工作、家庭帮工的工作总数为 2 份及以上 ($JobBCN+JobC1CN+JobC2CN \geq 2$)，则直接提问受访者最主要的工作单位，job2012mn 为该项工作的单位名称 (qg702)。也就是说，2012 年调查时生成的“主要工作单位名称” (job2012mn)，是在受雇、非农自雇和不拿工资为家庭经营活动帮工 3 类非农工作中当前正在从事或最近结束的一份工作，为字符型变量，其值为该项工作的单位名称。

因此，我们需要生成 2012 年主要职业的综合变量，不仅要包括农业工作，还要考虑到不同类型工作的特征和问题的设计，使其尽可能与 2010 年可比。

二、CFPS 2012 当前主要职业综合变量的生成规则

1. 当前最佳主要工作

在生成 CFPS 2012 当前主要职业变量时，我们首先要确定受访者当前主要工作单位名称。这是因为在 CFPS 中受访者的工作是通过工作单位名称来定位的。为了在规则上与长时期的主要工作生成方式保持一致，CFPS 2012 特意生成一个新的变量“当前最佳主要工作 (job2012mn_best)”，以工作时间为标准将农业生产整合进来。也就是说，“当前最佳主要工作”是在自家农业生产、农业打工、受雇、非农自雇和不拿工资为家庭经营活动帮工 5 类中正在从事的一份工作，为字符型变量，其值为该项工作的单位名称。这一变量可供 CFPS2014 加载使用。

(1) 生成规则

以下是 2012 年“当前最佳主要工作”的生成规则。所需的变量列表见附表 1。

■ 当 $JobBCN+JobC1CN > 0$ 时，受访者当前从事受雇、自雇类中的不少于一份工作，且已知主要非农工作变量 job2012mn。

1) 若 job2012mn 对应的是受雇工作 jobbname_a_1-10 且每周工作时间不少于 20 小时，则 $job2012mn_best=job2012mn$ ；若 job2012mn 对应的是非农自雇工作 jobc1name_a_1-10 且至少工作 6 个月 ($qg511_a_1-10 \geq 6$)，同时每周工作时间大于或等于 20 小时，则 $job2012mn_best=job2012mn$ ；

2) 对于受雇工作每周工作时间小于 20 小时、非农自雇工作小于 6 个月或每周工作时间小于 20 小时的情况:

a. 若过去一年自家农业生产至少有 3 个月 ($qg202 \geq 3$) 且农忙时每周工作时间不少于 20 小时, 则 job2012mn_best 赋值为“自家农业”;

b. 对于过去一年自家农业生产小于 3 个月或农忙时每周工作时间少于 20 小时的情况, 若过去一年农业打工至少有 90 天 ($qg303 \geq 90$), 且每天工作时间不少于 4 小时 ($qg304 \geq 4$), 则 job2012mn_best 赋值为“农业打工”;

c. 若以上条件都不满足, 则 job2012mn_best = -8。

■ JobBCN+JobC1CN=0 时, 受访者当前没有从事除家庭帮工以外的非农工作。

1) 若过去一年自家农业生产至少有 3 个月 ($qg202 \geq 3$) 且农忙时每周工作时间不少于 20 小时, 则 job2012mn_best 赋值为“自家农业”;

2) 对于过去一年自家农业生产小于 3 个月或农忙时每天工作时间小于 20 小时的情况, 若过去一年农业打工至少有 90 天 ($qg303 \geq 90$), 且每天工作时间不少于 4 小时 ($qg304 \geq 4$), 则 job2012mn_best 赋值为“农业打工”;

3) 若以上条件都不满足, 则 job2012mn_best = -8。

“当前最佳主要工作”生成规则的基本原则有两点: (1) 以受访者主观认定的为最重要; (2) 以工作时间为标准, 将农业整合进去, 从而选取出主要工作。

(2) 工作时间的计算

农业生产、农业打工、受雇和非农自雇的工作时间要求满足以下所列举的标准。

■ 农业生产

1) 至少工作 3 个月 ($qg202 \geq 3$);

2) 农忙时每周工作时间不少于 20 小时, 即农忙时每月工作天数*每天工作小时数/4 周:
 $agri_time = qg203 * qg204 / 4$ 。

■ 农业打工

1) 至少工作 90 天 ($qg303 \geq 90$);

2) 每天工作时间不少于 4 小时 ($qg304 \geq 4$)。

■ 受雇: 每周工作时间不少于 20 小时, 即每月工作天数*每天工作小时数/4 周:

$jobbtime_a_1-10 = qg413_a_1-10 * qg414_a_1-10 / 4$ 。

■ 非农自雇

1) 至少工作 6 个月 ($qg511_a_1-10 \geq 6$) ;

2) 每周工作时间不少于 20 小时, 即每月工作天数*每天工作小时数/4 周:

$jobc1time_a_1-10 = qg512_a_1-10 * qg513_a_1-10 / 4$ 。

由于“当前最佳主要工作 (job2012mn_best)”的值是具体的工作单位名称, 涉及到受访者的隐私, 只作为生成当前主要职业变量的依据, 不在公开数据库中出现。

2. CFPS2012 当前主要职业

这一步的目的是根据已经生成的“当前最佳主要工作”定位到具体职业, 并赋予其相应的职业代码, 生成“当前主要职业”变量 (job2012mn_occu)。生成规则如下:

■ 若当前最佳主要工作 job2012mn_best 等于受雇工作 jobbname_a_1-10 或受雇工作具体内容 qg411_a_1-10 时, 则当前主要职业 job2012mn_occu 等于对应的受雇职业代码 qg411code_a_1-10;

■ 若当前最佳主要工作 job2012mn_best 等于非农自雇工作 jobc1name_a_1-10 或非农自雇工作具体内容 qg510_a_1-10 时, 则当前主要职业 job2012mn_occu 等于对应的受雇职业代码 qg510code_a_1-10;

■ 若当前最佳主要工作 job2012mn_best 为“自家农业”或“农业打工”时, 则当前主要职业 job2012mn_occu 等于 50000, 即“农、林、牧、渔、水利业生产人员”。

生成“当前主要职业”所需的变量列表见附表 2。

3. 数据结果

本部分将根据“当前主要职业 (job2012mn_occu)”来报告 2012 年受访者的职业分布, 并且与 2010 年的职业分布进行比较。¹

¹ 该报告展示的是 CFPS 2012 成人库的数据结果; 少儿库仅有不足 30 例个案需要进行当前主要职业的确认, 故在此不作报告。

■ 职业大类的分布

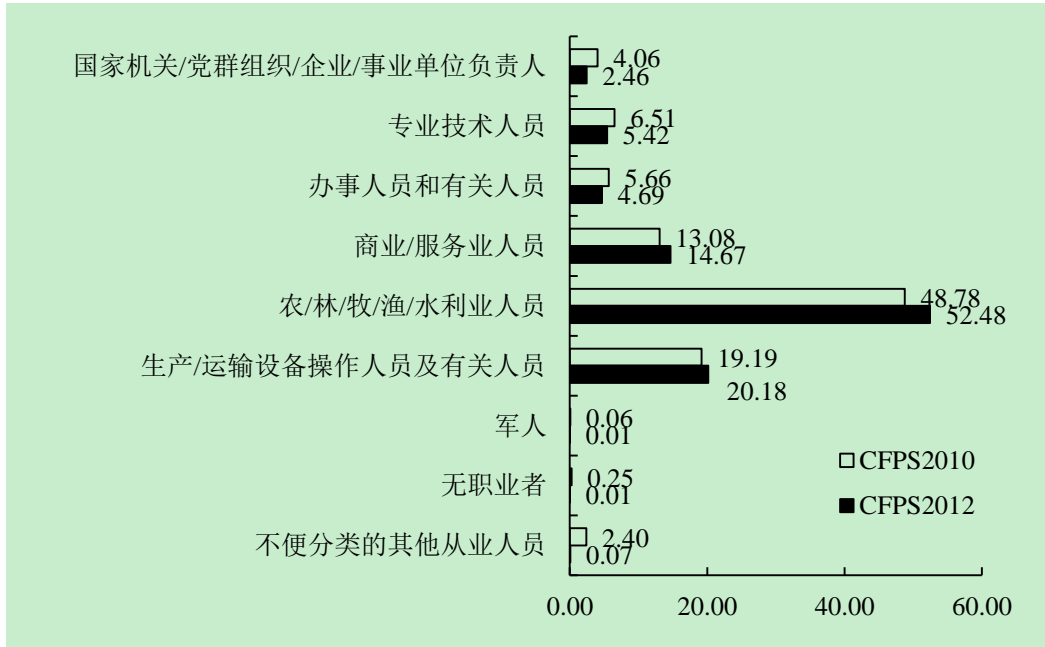


图 1. CFPS 2010 与 2012 的职业分布 (%)

在 CFPS 2010 的就业人口中，有 99.38% 的受访者给出了有效的职业信息，可以进行职业编码；在 2012 年，就业人口中有 59.38% 根据以上步骤生成了职业变量，赋予职业分类和代码。² 也就是说，在 2012 年的就业人口中，有 40% 的受访者并没有当前主要职业的信息，赋值为 -8 “不适用”。这主要是由于 2010 年采取的是直接询问受访者的主要工作的方式，而 2012 年则是根据以上步骤生成的综合变量，损失了一定的信息。比如，尽管受访者当前从事某项工作，但该项工作并不满足工作时长的条件，则当前主要职业变量为不适用。

图 1 是 2010 年和 2012 年职业大类的分布。在 2012 年，农、林、牧、渔、水利业生产人员的比例最大，为 52.48%，其次是生产、运输设备操作人员及有关人员（20.18%），商业、服务业人员（14.67%），三者之和超过总体的 85%。总的来说，这与 2010 年职业大类的分布基本一致。

² 需要说明的是，2010 年和 2012 年调查关于就业的统计口径不一致。2010 年就业的判断标准是目前有工作，或者目前没有工作但是处于参加培训的期间；2012 年就业指的是过去一周从事工作至少一小时，或者过去一周工作未满一小时但处在 6 个月或其他确定时间以内的临时放假、短期离职休假、在职培训等，或者过去一周工作未满一小时但处在生意淡季、农闲季节的暂时歇业状态。

■ 职业大类的性别差异

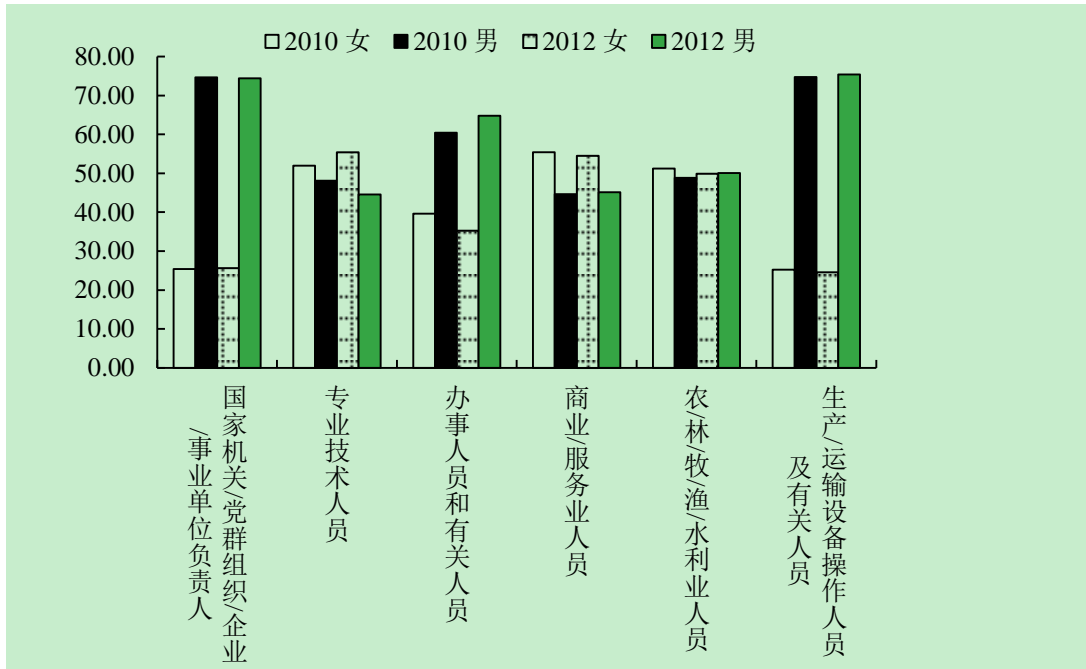


图 2. CFPS 2010 和 2012 职业大类的性别分布³

由图 2 可知，在职业大类中，国家机关、党群组织、企业、事业单位负责人和生产、运输设备操作人员及有关人员的男女构成差异最大，男性比例是女性比例的 3 倍左右；办事人员和有关人员的男性比例也更高一些，为 64.75%，是女性比例的近 2 倍；而在专业技术人员和农、林、牧、渔、水利业人员中，男女比例大致相同。2012 年和 2010 年职业大类的性别分布趋势基本一致。不同的是，从 2010 年到 2012 年，专业技术人员的女性比例略有上升，而办事人员和有关人员的女性比例有所下降。

³ 军人、无职业分类及代码、不便分类的其他从业人群的个案数很小，各有 2 人、1 人、10 人，故在性别、年龄、教育、收入差异部分不做讨论。

■ 职业大类的年龄差异

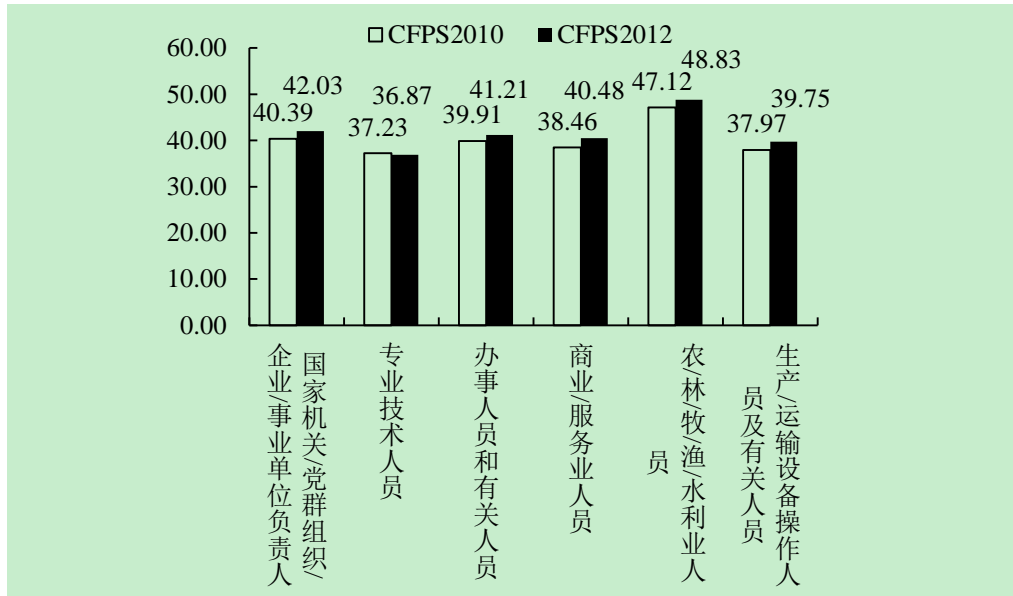


图 3. CFPS 2010 和 2012 职业大类的平均年龄

图 3 反映的是 6 大职业类别的平均年龄。从年龄组来看，不同职业的从业人员集中在 20-59 岁，正是劳动年龄人口。农、林、牧、渔、水利业人员的年龄总体上更大一些，平均年龄最高，2012 年为 49 岁左右；而专业技术人员的平均年龄最低，2012 年约为 37 岁。这些都反映出不同职业的特征。2012 年和 2010 年职业大类的平均年龄大致相同，除专业技术人员外，其他 5 类职业 2012 年的平均年龄略高于 2010 年。

■ 职业大类的教育差异

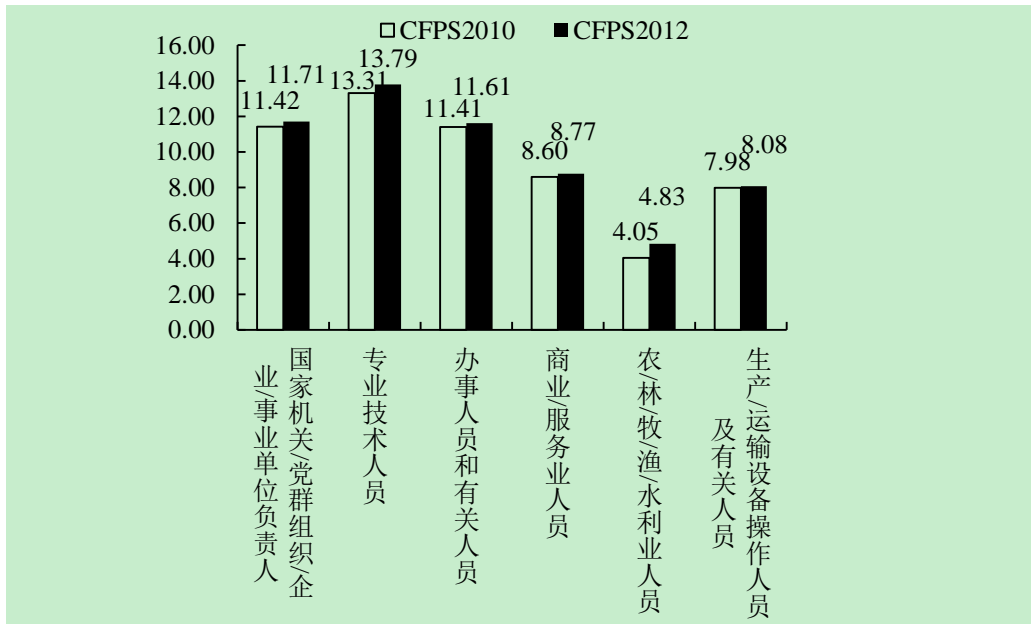


图 4. CFPS 2010 和 2012 职业大类的平均受教育年限

图 4 是 6 大职业类别的平均受教育年限。数据显示，专业技术人员的受教育程度普遍较高，无论是 2010 年还是 2012 年，平均受教育年限均超过 13 年；其次是在国家机关、党群组织、企业事业单位负责人与办事人员和有关人员，平均受教育年限在调查年份均为 11 年左右；商业、服务业人员与生产、运输设备操作人员及有关人员则是以中学水平为主，平均受教育年限约为 8 年；农、林、牧、渔、水利业人员的受教育程度最低。2012 年各个职业的受教育程度均比 2010 年有所提高。

■ 职业大类的收入差异

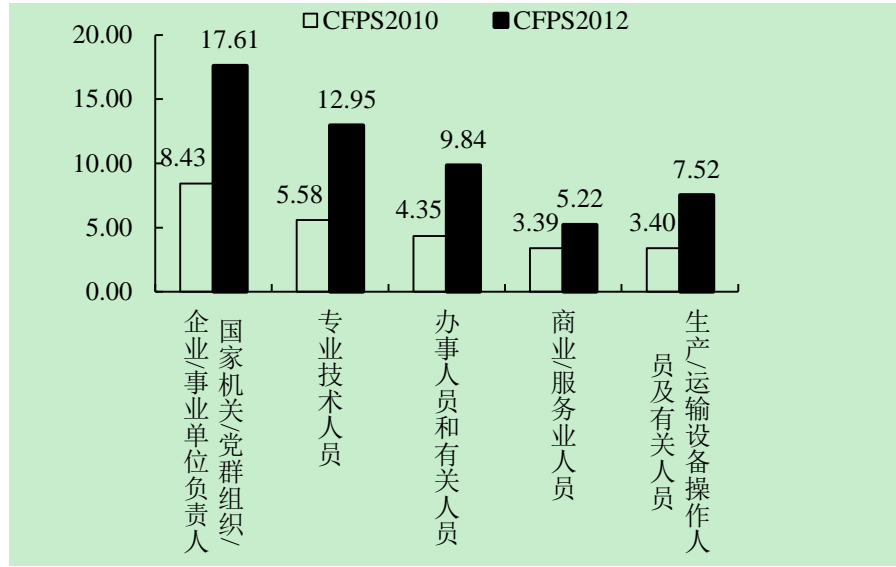


图 5. CFPS 2010 和 2012 非农与农业职业的收入之比

图 5 反映的是 5 个非农职业（国家机关、党群组织、企业、事业单位负责人，专业技术人员，办事人员和有关人员，商业、服务业人员和生产、运输设备操作人员）的个人年收入与农业生产人员的个人年收入之比。与基本认识一致，国家机关、党群组织、企业、事业单位负责人的收入水平最高，2012 年的个人年收入平均值约为 5 万元；专业技术人员的收入位居第二位；其次是办事人员和有关人员，生产、运输设备操作人员及有关人员。此外，2012 年的收入比远远高于 2010 年的收入比。这一方面是因为非农职业的个人年收入普遍提高，比如，2012 年专业技术人员的年收入平均值近 4 万元，比 2010 年增长了约 20%；另一方面，与 2010 年相比，2012 年农业生产人员的年收入近为 3000 元左右，显著降低。

需要说明的是，CFPS 收集了受访者的职业描述，并将其编码为国家标准职业分类（GB/T6565-2009）。因此，当前主要职业（job2012mn_occu）采取的也是国家标准职业分类及其代码。从国家标准职业分类代码匹配至 ISCO-88 的操作步骤可见参技术报告《中国家庭动态跟踪调查职业社会经济地位测量指标构建》（黄国英、谢宇，2012）。

附录

附表 1. 生成“当前最佳主要工作”的变量列表

变量	标签
jobbcn	当前从事受雇工作的份数
job1cn	当前从事非农自雇工作的份数
qg202	从事农业生产的月数
qg203	农忙时每月农业生产天数
qg204	农忙时每天农业生产时长
qg303	为其他农户打工天数
qg304	每天打工时间
jobBlastdate_a_1-10	受雇工作持续时间
jobC1lastdate_a_1-10	非农自雇工作持续时间
qg413_a_1-10	受雇工作的月工作天数
qg414_a_1-10	受雇工作的每天工作时间
qg511_a_1-10	过去一年非农自雇月数
qg512_a_1-10	非农自雇的月工作天数
qg513_a_1-10	非农自雇的日工作时长
jobbname_a_1-10	受雇工作单位名称
jobc1name_a_1-10	非农自雇工作单位名称
job2012mn	主要工作单位名称

附表 2. 生成“当前主要职业”的变量列表

变量	标签
job2012mn_best	当前主要职业
job2012mn_isco	当前主要职业 ISCO88 分类
jobbname_a_1-10	受雇工作单位名称
qg411_a_1-10	受雇工作具体内容
qg411code_a_1-10	受雇职业代码
jobc1name_a_1-10	非农自雇工作单位名称
qg510_a_1-10	非农自雇工作具体内容
qg510code_a_1-10	受雇职业代码