

China Family Panel Studies

CFPS

中国家庭动态跟踪调查

技术报告系列：CFPS-5

系列编辑：谢宇 责任编辑：胡婧炜

# 中国家庭动态跟踪调查 2010年基线调查样本联系情况<sup>1</sup>

孙妍

2012.12.20

---

<sup>1</sup> 美国密歇根大学助理教授闫婷及博士后徐宏伟在报告撰写过程中提出了许多宝贵的意见和建议，在此衷心感谢他们为完善报告所作的贡献。

# 1. 项目背景介绍

中国家庭动态跟踪调查（CFPS）基线调查的目标样本是中国 25 个省/市/自治区（除香港、澳门、台湾以及新疆维吾尔自治区、西藏自治区、青海省、内蒙古自治区、宁夏回族自治区、海南省）的居住在传统居民住宅内且满足项目访问条件的家庭户。为了保证末端抽样的精确度，CFPS 基线调查采用地图地址法构建末端抽样框。<sup>2</sup> 在实地调查开始前，北京大学中国社会科学调查中心（以下简称“中心”）组织绘图人员到各个样本社区绘制住宅分布图，构建末端地址抽样框。末端地址抽样框构建完成后，中心专业抽样人员采用随机抽样原则从中抽取一定数量的备访地址，每一个地址即为一个抽样单元。

虽然在末端样本框构建阶段绘图人员通过实地走访、询问邻居、与居委会人员确认等多种方式尽力将社区内空置房屋和非住宅排除，但由于实际操作难度仍无法保证末端抽样框百分百准确。因此，项目组设计了“住宅过滤模块”用于记录访员寻找地址及判定房屋类型的结果。仅当访员确认项目组提供的地址正确，并确认房屋类型为传统居民住宅方可继续访问，否则终止访问。

确认备访地址上的建筑为传统居民住宅后，访员还需要邀请住宅内的住户完成一套住户过滤问卷以确认住宅内是否有符合 CFPS 访问条件的样本。CFPS 基线调查的目标样本必须同时满足以下两个条件<sup>3</sup>：(1)经济独立的家庭户；(2)家里至少一名家庭成员具有中国国籍。<sup>4</sup> 若该地址上不存在满足条件的住户，则终止访问；若仅存在一户满足条件的住户，该住户自动成为 CFPS 基线调查的目标家户；若存在超过一户满足条件的住户，访问系统将随机抽选其中一户作为目标家户。

确定目标家户后，访员从收集家庭成员关系结构的家庭成员问卷入手开始主体调查。目标家户除了完成家庭成员问卷外，还需要完成一套家庭问卷；家庭内的每名家庭成员则需要分别完成一套个人问卷。

概而言之，CFPS 基线调查是通过地址过滤、住宅过滤、住户过滤三重过滤对抽样单元

---

<sup>2</sup> CFPS 末端抽样框构建方式参见丁华，2012，《中国家庭动态跟踪调查 2010 年基线调查末端抽样框制作（CFPS-2）》。

<sup>3</sup> 最初住户过滤的设计还要求受访家庭户中至少有一名成员在抽样社区居住时间满 6 个月，但在执行过程中这一条件被取消。实际被这一条件过滤掉的仅有极少数家户。

<sup>4</sup> CFPS 住户过滤流程参见孙妍、严洁等，2011，《中国家庭动态跟踪调查（2010）访员培训手册》，北京大学出版社。

是否符合访问条件进行甄别。在确认存在符合条件的目标家户后，分别采集家庭和个人两个层面上的数据，具体访问流程参见图 1。

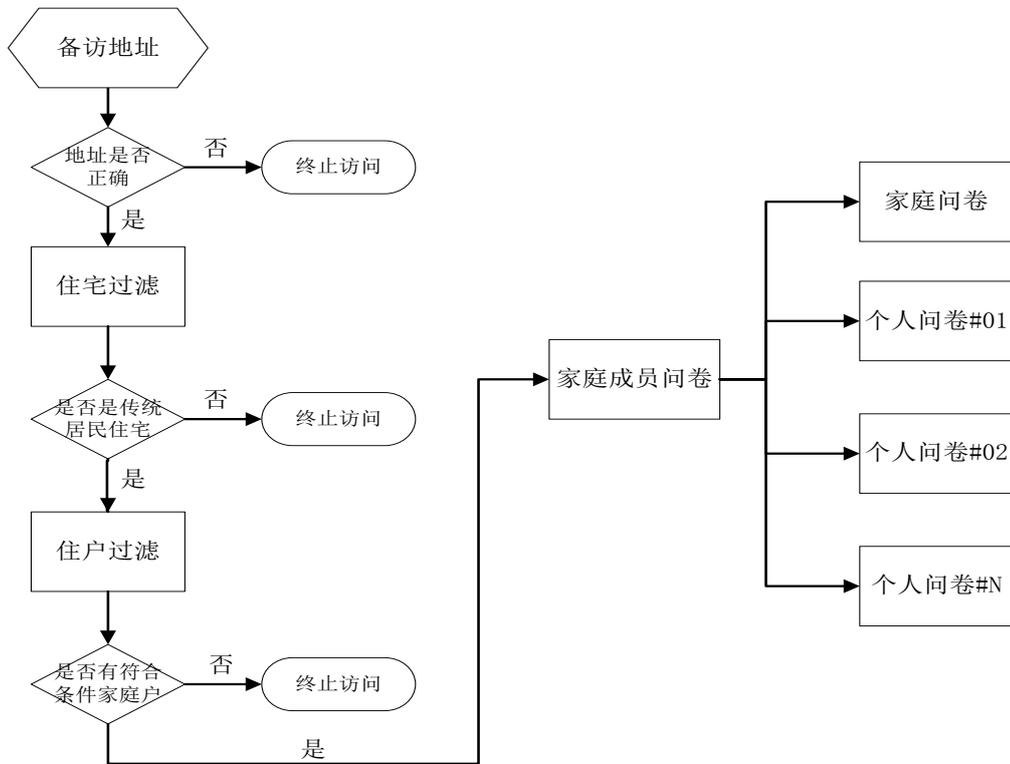


图 1. CFPS 基线调查访问流程图

本报告将参照美国公众民意调查协会(American Association for Public Opinion Research, AAPOR)<sup>5</sup> 给出的计算标准，从抽样单元层面及个人层面分别计算 CFPS 基线调查的各类执行率<sup>6</sup>。

## 2. 抽样单元层面执行率计算

### 2.1 抽样单元最终状态分类及定义

将是否符合访问条件作为分类标准，CFPS 基线调查抽样单元访问状态可分为符合访问条件、不符合访问条件、不确定是否符合访问条件三大类。

<sup>5</sup> The American Association for Public Opinion Research. 2011. *Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys. 7th edition*. AAPOR.

<sup>6</sup> 本报告计算的各类执行率均未加权，详细权数可参见吕萍、谢宇，2012，《中国家庭动态跟踪调查 2010 年基线调查权数计算 (CFPS-17)》。

### 2.1.1 符合访问条件

符合访问条件指的是该抽样单元已经完成住户过滤，且确认住宅内至少存在一户符合CFPS访问条件的家庭户。符合访问条件的样本根据家庭成员问卷完成情况可以继续分为完访和未完访两类。

#### (1) 完访

考虑到家庭成员问卷在本次调查中重要性及不可分割性，CFPS基线调查将家庭成员问卷全部完成作为抽样单元层面完访的标准，不再设定部分完访的标准。

#### (2) 未完访

未完访是指已确定抽样单元内存在符合访问条件的家庭户，但因各种原因无法继续完成家庭成员问卷，具体可分为拒访、未联系上、其他原因无法完成访问三类。

##### ① 拒访

目标家户拒绝参与家庭成员问卷访问或在访问过程中中途退出都被视为“拒访”。此外，参照AAPOR样本最终状态分类建议，本报告中将访员曾经联系上并没有明确表示拒访，但项目执行季结束时仍无法完成访问的家户也归为“拒访”。

##### ② 未联系上

未联系上指的是在确定目标家户后，因各种原因无法与其取得联系。这其中包括住户过滤问卷完成后无法再次联系上目标家户的情形，也包括从未与目标家户取得联系的情形。CFPS实地执行过程中，住户过滤问卷可以由住宅内的任何一个住户完成，当遇到一宅多户的情形时，住户过滤问卷回答人所在的家户与目标家户可能不一致。因此，存在从未与目标家户取得联系的情况。

##### ③ 其他原因无法接受访问

其他原因无法接受访问是指目标家户因方言问题、身体原因或其他原因致使家庭成员问卷没有回答或没有完整回答。

### 2.1.2 不符合访问条件

正如前文所提，CFPS 在主体访问开始前需要完成地址、住宅、住户三重过滤，因此不符合访问条件的类型可以分为地址错误、建筑物类型不符合条件和住户过滤不合格三个层次：

#### (1) 地址错误

地址错误是指访员根据中心提供的地址无法找到相应建筑物的情形。地址不存在，或者地址不详细无法准确找到对应建筑物均属于地址错误的范畴。

#### (2) 建筑物类型不符合条件

##### ① 非住宅

非住宅是相对于传统居民住宅而言的，具体包含商用建筑和非传统住宅两类。

##### A. 商用建筑

商用建筑是指纯用作商业用途的建筑物，例如小卖铺、游戏机厅、台球厅等。

##### B. 非传统住宅

非传统住宅是指供人居住但不具备传统住宅特征的房屋或场所，包括工棚、旅馆、军营、福利院、养老院、学生宿舍等。

##### ② 空置房屋

空置房屋是指达到住人和使用标准，但目前没有投入使用的建筑物。

#### (3) 住户过滤不合格

住户过滤不合格是指完成住户过滤问卷，但住宅内不存在满足 CFPS 访问条件的住户。

### 2.1.3 不确定是否符合访问条件

该类型包含因各种原因未能完成住户过滤的所有样本，具体包括未接触样本、正确地址无法确定房屋类型、住宅无法完成住户过滤三大类。

#### (1) 未接触样本

未接触样本是指在调查执行季内访员未尝试联系的抽样单元,包括未分配样本及无法到达访问地点两类。

##### ① 未分配样本

未分配样本是指项目组因特殊原因没有发放给访员的样本。

##### ② 无法到达访问地点

无法到达访问地点是指访员因天气原因或其他外界干扰因素(如保安、门禁等)无法抵达抽样单元所在的地点。

#### (2) 正确地址无法确定房屋类型

正确地址无法确定房屋类型是指访员找到备访地址但没有能够接触到住宅内的住户或其他信息人,从而无法准确确定建筑物的使用情况。

#### (3) 住宅无法完成住户过滤

住宅无法完成住户过滤是指在访员判定建筑物为传统居民住宅后,因住户拒访、方言问题、身体原因或其他原因致使住户过滤问题没有回答或没有完整回答,因此无法判断住宅内是否存在符合条件的家庭户。

### 2.1.4 小结

为了后期计算方便,笔者将三种类型 13 类的最终状态进一步归纳为判断访问条件、联系目标家户以及尝试完成访问三个阶段 6 个子类,分别为:完访(I)、拒访(R)、其他原因无法完成访问(O)、未联系上(NC)、不符合访问条件(NE)、不确定是否符合访问条件(UE)。具体分类情况参见图 2。为了便于读者理解,将抽样单元的 6 类最终状态产生的环节与访问流程图相结合,详见图 3。

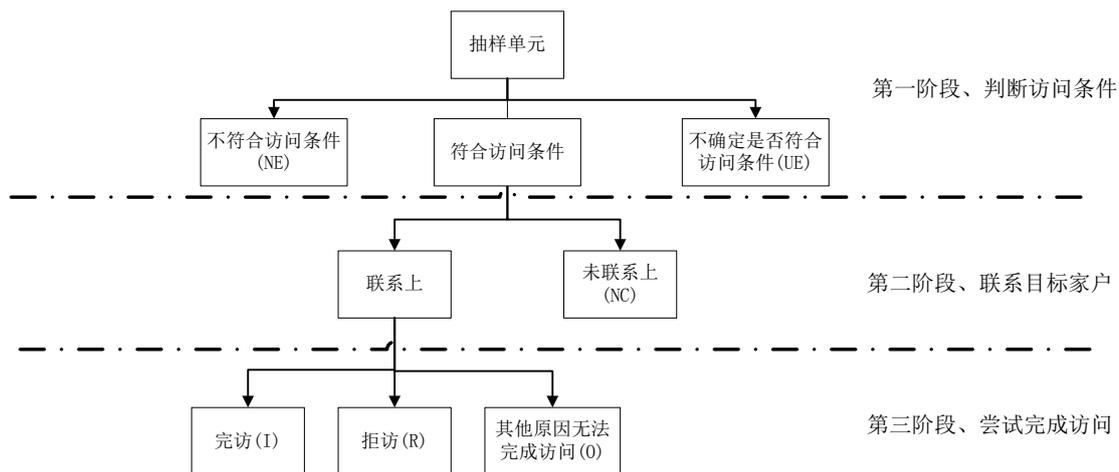


图 2. CFPS 基线调查抽样单元最终状态层次图

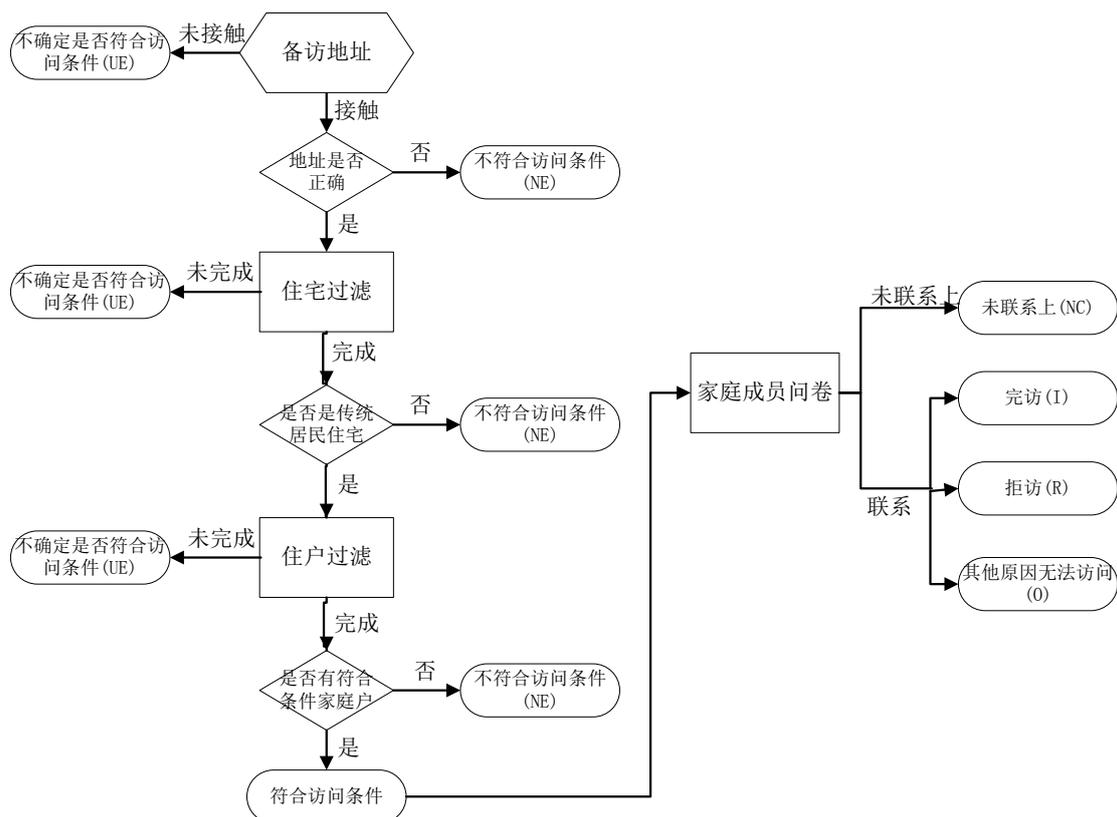


图 3. CFPS 基线调查抽样单元最终状态生成流程图

## 2.2 抽样单元最终状态分布

### 2.2.1 符合访问条件样本

表1展示了CFPS2010基线调查中19,986个抽样单元最终状态在6个子类的分布情况。在所有19,986个抽样单元中一共过滤出15,543个住址符合CFPS访问条件，其中，14,960个样本被成功转化为完访样本；拒访则是符合访问条件样本无法成功转化为完访样本的最主要原因，占有未完访样本的84.22%。从社区类型角度分析，居委会访问难度明显高于村委会，体现为较低的完访率和较高的拒绝率。若综合考虑CFPS基线调查末端抽样扩大比例抽样的设计，居委会和村委会符合条件样本的最终完访比例大体相当。<sup>7</sup>

### 2.2.2 不符合访问条件样本

不符合访问条件的1,464个样本中，空置房屋占71.04%，住户过滤不合格的占14%，非住宅和错误地址分别占9.15%和5.81%，详见表2。抽样单元中空户比例较高主要有两方面原因：其一、末端抽样框的精度受到一些客观因素的制约；其二、末端抽样框构建与入户调查之前有3-5个月的时间差。从数值上看，村委会地区抽样单元的空户率及不符合访问条件的比例高于居委会，而居委会中的抽样单元因非住宅和错误地址而被排除的比例较高。

### 2.2.3 不确定是否符合访问条件样本

住户拒访是导致无法确定抽样单元是否满足访问条件的首要原因，其次是无法接触到住址内的住户，因其他原因无法完成住户过滤的样本仅占不到7%。在拒访样本中，约90%的样本是在住户过滤访问开始前就直接拒访。从表3中可以看出，导致住户过滤问卷无法完成的原因在村、居有所不同。拒访问题在居委会地区表现尤为突出，村委会地区的样本则更多是因为无法联系到住户而导致无法进行住户过滤问卷访问。

### 2.2.4 小结

整体而言，CFPS基线调查样本的损耗主要集中在第一阶段判断访问条件的住户过滤环节，该环节损耗了2,967个样本，占样本总体的14.85%。影响访问项目执行的两大因素是

---

<sup>7</sup> 详细末端抽样扩大比例可参见谢宇、邱泽奇、吕萍，2012，《中国家庭动态跟踪调查抽样设计（CFPS-1）》

拒访和无联系。在居委会所在的城市地区住户拒访成为调查执行中遇到的最棘手难题；而在农村地区，则需要克服人户分离、空置房屋、难以寻找住户的困难。

**表 1. CFPS 基线调查抽样单元最终状态分布**

抽样单元最终状态		居委会		村委会		整体	
		数量	百分比	数量	百分比	数量	百分比
符合访问条件	完访(I)	5,081	63.09	9,879	82.79	14,960	74.85
	拒访(R)	348	4.32	143	1.20	491	2.46
	其他原因未完成访问(O)	14	0.17	25	0.21	39	0.20
	未联系(NC)	17	0.21	36	0.30	53	0.27
不符合访问条件(NE)		652	8.10	812	6.80	1,464	7.33
不确定是否符合访问条件(UE)		1,941	24.10	1,038	8.70	2,979	14.91
合计		8,053	100	11,933	100	19,986	100

**表 2. CFPS 基线调查抽样单元不符合访问条件类型分布**

不符合访问条件类型	居委会		村委会		整体	
	数量	百分比	数量	百分比	数量	百分比
错误地址	62	9.51	23	2.83	85	5.81
非住宅	97	14.88	37	4.56	134	9.15
空置房屋	412	63.19	628	77.34	1,040	71.04
住户过滤不合格	81	12.42	124	15.27	205	14.00
合计	652	100	812	100	1,464	100

**表 3. CFPS2010 抽样单元不确定是否符合访问条件类型分布**

不确定是否符合访问条件类型	居委会		村委会		整体	
	数量	百分比	数量	百分比	数量	百分比
未接触地址	7	0.36	5	0.48	12	0.40
正确地址无法确定房屋类型	662	34.11	484	46.63	1,146	38.47
因住户拒访无法完成住户过滤	1,194	61.51	430	41.43	1,624	54.51
其他原因无法进行住户过滤	78	4.02	119	11.46	197	6.61
合计	1,941	100	1,038	100	2,979	100

## 2.3 抽样单元层面各类执行率计算

本报告将引用 AAPOR 标准计算 CFPS2010 基线调查的抽样单元层面的应答率 (Response Rate, 缩写为 RR)、合作率 (Cooperation Rate, 缩写为 COOP)、拒绝率 (Refusal Rate, 缩写为 REF) 及联系率 (Contact Rate, 缩写为 CON)。

### 2.3.1 应答率

应答率是指符合访问条件的样本中最终转化为完访或部分完访样本所占的比例。由于 CFPS2010 基线调查中未设立部分完访的概念, 因此在 CFPS 基线调查中抽样单元的应答率可以表述为完成家庭成员问卷家庭占有所有符合访问条件家庭的比例。本报告将分别运用 RR1、RR3、RR5<sup>8</sup>三种方式计算 CFPS 基线调查应答率, 具体公式如下:

$$RR1 = \frac{I}{I+R+NC+O+UE} \dots\dots\dots \text{公式(1)}$$

$$RR3 = \frac{I}{I+R+NC+O+eUE} \dots\dots\dots \text{公式(2)}$$

$$RR5 = \frac{I}{I+R+NC+O} \dots\dots\dots \text{公式(3)}$$

其中, 公式(2)中的 e 代表已完成访问资格筛选的样本中符合访问条件样本所占的比例。在 CFPS 符合条件有三层含义: 首先, 地址应该是个正确地址; 其次, 地址上的建筑功能应该是传统居民住宅; 最后住宅内的住户应该符合项目设定的访问条件。因此, 分别对应正确地址率、住宅率及住户符合条件率, 方便起见分别记为 e1, e2, e3。各类符合条件率计算公式及结果如下:

$$e1 = \frac{\text{正确地址数}}{\text{抽样数} - \text{未接触地址数}} = \left(1 - \frac{85}{19986 - 12}\right) * 100\% = 99.57\%$$

$$e2 = \frac{\text{住宅数}}{\text{正确地址数} - \text{未确定房屋数}} = \left(1 - \frac{134 + 1040}{19889 - 1146}\right) * 100\% = 93.74\%$$

$$e3 = \frac{\text{符合条件样本数}}{\text{住宅数} - \text{未确定住户数}} = \left[1 - \frac{205}{17569 - (1624 + 197)}\right] * 100\% = 98.70\%$$

<sup>8</sup> 文章所引用公式如无特殊注明, 均来自 AAPOR. 2011. "Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys. 7th edition."

根据以上计算的符合条件率估算各类地区符合条件抽样单元数，结果如下：

$$eUE_{\text{整体}} = \text{未接触地址数} * e1 * e2 * e3 + \text{未确定房屋数} * e2 * e3 \\ + \text{未确定住户数} * e3 \approx 2868$$

运用同样的方法计算居委会地区及村委会地区符合条件抽样单元估计数，结果如下：

$$eUE_{\text{居委会}} = 7 * 0.9923 * 0.9305 * 0.9854 + 662 * 0.9305 * 0.9854 + 1272 * 0.9854 \\ \approx 1867$$

$$eUE_{\text{村委会}} = 5 * 0.9981 * 0.9418 * 0.9879 + 484 * 0.9418 * 0.9879 + 549 * 0.9879 \\ \approx 997$$

三种方式计算所得的 CFPS 基线调查应答率如表 4 所示，整体应答率的上限为 96.25%，应答率的下限为 80.77%。不同方式计算的区别在于分母中对于不确定是否符合访问条件样本（以下简称为 UE 样本）的处理。RR1 的处理方式是将所有 UE 样本全部视为符合条件样本；RR5 则完全不考虑 UE 样本；RR3 则采用折中的方式根据实际符合条件率推断 UE 中符合条件的样本数。结合前文分析不难发现，CFPS 基线调查中第二重住宅类型过滤符合条件率仅为 93.74%，若采用 RR1 全部认定的方式，则低估了应答率。此外，对于采用三重过滤的 CFPS 基线调查而言，访问难度主要集中于住户过滤环节，在这个过程中会出现大量的拒访或无联系样本，若采用 RR5 不考虑 UE 的方式，则高估了应答率。结合 CFPS 基线调查的特点，参考 Robert M. Groves 等人提供的应答率计算方式，<sup>9</sup> 建议选用 RR3 为应答率计算的标准。即 CFPS 基线调查整体应答率为 81.25%，居委会样本的应答率为 69.35%，村委会样本的应答率为 89.16%。

**表 4. CFPS 基线调查抽样单元应答率计算结果 (%)**

应答率类型	居委会	村委会	整体
RR1=I/(I+R+NC+O+UE)	68.65	88.83	80.77
RR3=I/(I+R+NC+O+eUE)	69.35	89.16	81.25
RR5=I/(I+R+NC+O)	93.06	97.98	96.25

<sup>9</sup> Robert M. Groves et al. 2009. *Survey Methodology (2nd edition)*. New York: Wiley.

### 2.3.2 累积应答率

介于 CFPS 基线调查采用多阶段过滤的设计模式，抽样单元层面应答率的另一种计算方式是综合考虑住户过滤阶段及家庭成员问卷访问阶段的执行情况，生成累积应答率，具体公式如下：

$$RR_{\text{住户过滤}} = \frac{I + R + O + NC + NE_{\text{住户过滤}}}{\text{住宅数} + eUE_{\text{住宅}}}$$

$$\text{住宅数} = I + R + O + NC + NE_{\text{住户过滤}} + UE_{\text{住户过滤}} + eUE_{\text{住宅}}$$

$$eUE_{\text{住宅}} = \text{未接触地址数} * e1 * e2 + \text{未确定房屋数} * e2$$

$$RR_{\text{家庭成员}} = RR_5 = \frac{I}{I + R + O + NC}$$

$$RR_{\text{累积}} = RR_{\text{住户过滤}} * RR_{\text{家庭成员}}$$

由于住户过滤阶段样本不符合条件的比例极低（2%以下），因此累积应答率的计算结果与 RR3 的结果基本一致。CFPS 基线调抽样单元累积应答率为 81.25%，居委会样本的应答率为 69.35%，村委会样本的应答率为 89.16%。

表 5. CFPS 基线调查抽样单元累积应答率计算结果（%）

应答率类型	居委会	村委会	整体
住户过滤应答率	74.52	91.00	84.42
家庭成员应答率	93.06	97.98	96.25
累积应答率	69.35	89.16	81.25

### 2.3.3 合作率

AAPOR 给出两类合作率的定义：第一、合作率是指所有符合访问条件且成功接触的样本中转化为完访样本的比例；第二、合作率是指所有符合访问条件且能够接受访问的样本中转化为完访样本的比例。本报告将分别运用 COOP1、COOP3 两类方式计算 CFPS 基线调查的合作率，具体公式如下：

$$COOP1 = \frac{I}{I+R+O} \dots\dots\dots \text{公式(4)}$$

$$COOP3 = \frac{I}{I+R} \dots\dots\dots \text{公式(5)}$$

**表 6. CFPS 基线调查抽样单元合作率计算结果 (%)**

合作率类型	居委会	村委会	整体
COOP1=I/(I+R+O)	93.35	98.33	96.58
COOP3=I/(I+R)	93.59	98.57	96.82

由于 CFPS 基线调查中因其他原因无法访问样本所占的比例非常低，因此两种方式计算结果相当，详见表 6。考虑到无论何种原因导致无法完成访问对项目而言都是一种样本损耗，建议选用较为严格的 COOP1 为合作率的计算标准。即 CFPS 基线调查整体合作率为 96.58%，居委会样本的合作率为 93.35%，村委会样本的合作率为 98.33%。

### 2.3.4 联系率

联系率是指符合访问条件的样本中成功接触的样本所占的比例。本报告将分别运用 CON1、CON2、CON3 三种方式计算 CFPS 基线调查联系率，具体公式如下：

$$CON1 = \frac{I+R+O}{I+R+NC+O+UE} \dots\dots\dots \text{公式(6)}$$

$$CON2 = \frac{I+R+O}{I+R+NC+O+eUE} \dots\dots\dots \text{公式(7)}$$

$$CON3 = \frac{I+R+O}{I+R+NC+O} \dots\dots\dots \text{公式(8)}$$

**表 7. CFPS 基线调查抽样单元联系率计算结果 (%)**

联系率类型	居委会	村委会	整体
CON1=(I+R+O)/(I+R+NC+O+UE)	73.54	90.34	83.63
CON2=(I+R+O)/(I+R+NC+O+eUE)	74.29	90.68	84.13
CON3=(I+R+O)/(I+R+NC+O)	99.69	99.64	99.66

三种方式计算所得的 CFPS 基线调查联系率如表 7 所示，整体联系率的上限为 99.66%，下限为 83.36%。正如前文所述，CON2 计算方法中对于不确定是否符合访问样本的处理较为符合 CFPS 基线调查的特点，建议选用 CON2 为联系率的计算标准。即 CFPS 基线调查整体联系率为 84.13%，居委会样本的联系率为 74.29%，村委会样本的联系率为 90.68%。

### 2.3.5 拒绝率

拒绝率是指符合访问条件的样本中拒访样本所占的比例。本报告将分别运用 REF1、REF2、REF3、REF4<sup>10</sup>计算 CFPS 基线调查拒绝率，具体公式如下：

$$REF1 = \frac{R}{I+R+NC+O+UE} \dots\dots\dots \text{公式(9)}$$

$$REF2 = \frac{R}{I+R+NC+O+eUE} \dots\dots\dots \text{公式(10)}$$

$$REF3 = \frac{R}{I+R+NC+O} \dots\dots\dots \text{公式(11)}$$

$$REF4 = \frac{R}{I+R} \dots\dots\dots \text{公式(12)}$$

**表 8. CFPS 基线调查抽样单元拒绝率计算结果 (%)**

拒绝率类型	居委会	村委会	整体
REF1=R/(I+R+NC+O+UE)	4.70	1.29	2.65
REF2=R/(I+R+NC+O+eUE)	4.75	1.29	2.67
REF3=R/(I+R+NC+O)	6.37	1.42	3.16
REF4=R/(I+R)	6.41	1.43	3.18

四种方式计算所得的 CFPS 基线调查拒绝率如表 8 所示，整体拒绝率的上限为 3.18%，下限为 2.65%。沿用应答率计算方式选用理由，建议选用 REF2 为拒绝率的计算标准。即 CFPS 基线调查整体拒绝率为 2.67%，居委会样本的拒绝率为 4.75%，村委会样本的拒绝率为 1.29%。

### 2.3.6 小结

CFPS 基线调查抽样单元层面累积应答率为 81.25%，合作率为 96.58%，联系率为 84.13%，拒绝率为 2.67%。村委会地区的执行情况普遍优于居委会地区。

<sup>10</sup> Robert M. Groves et al. 2009. *Survey Methodology (2nd edition)*. New York: Wiley.

表 9. CFPS 基线调查抽样单元各类执行率计算结果 (%)

类型	计算公式	居委会	村委会	整体
应答率	$RR3=I/(I+R+NC+O+eUE)$	69.35	89.16	81.25
累积应答率	$RR_{\text{累积}}=RR_{\text{住户过滤}}*RR_{\text{家庭成员}}$	69.35	89.16	81.25
合作率	$COOP1=I/(I+R+O)$	93.35	98.33	96.58
联系率	$CON2=(I+R+O)/(I+R+NC+O+eUE)$	74.29	90.68	84.13
拒绝率	$REF2=R/(I+R+NC+O+eUE)$	4.75	1.29	2.67

### 3. 个人层面执行率计算

#### 3.1 个人层面最终状态分类及定义

CFPS 基线调查中将以下三类人员认定为家庭成员：(1)与该家庭有血缘/婚姻/领养关系的直系亲属；(2)截至调查时点，在该家庭居住满 3 个月，与该家庭有血缘/婚姻/领养关系的非直系亲属；(3)截至调查时点，在该家庭居住满 6 个月，与该家庭没有血缘/婚姻/领养关系的其他人员。<sup>11</sup> 其中，前两类家庭成员被视为符合受访资格、需要完成个人访问的“基因家庭成员”。

##### (1) 完访

CFPS 基线调查将整套个人问卷全部完成作为个人层面完访的标准，不再设定部分完访的标准。

##### (2) 未完访

未完访是指因各种原因无法完成个人问卷。根据无法完成的原因，可分为拒访、未联系上、其他原因无法接受访问三类。

##### ① 拒访

目标受访对象拒绝参与个人问卷访问或在访问过程中中途退出都被视为“拒访”。此外，将访员曾经联系上并没有明确表示拒访，但项目执行季结束时仍无法完成访问的个人样本也

<sup>11</sup> CFPS 家庭成员的判定规则参见孙妍、严洁等，2011，《中国家庭动态跟踪调查（2010）访员培训手册》，北京大学出版社。

视为“拒访”。

② 未联系上

未联系上指的是因各种原因未能与目标受访对象取得联系。

③ 其他原因无法访问

其他原因无法访问是指目标受访对象身体原因或其他客观原因致使个人问卷没有回答或没有完整回答。

(3) 不符合访问条件的样本

综合考虑项目执行期、执行难度及执行成本，项目组决定在 CFPS 基线调查中放弃访问目前正在服刑、参军/服役或出境的家庭成员。同时，放弃访问居住在非本区县内、因其他原因外出的家庭成员。本报告中将这部分样本视为是不符合访问条件的样本。

### 3.2 个人样本最终状态分布

#### 3.2.1 整体分布情况

个人样本中有 11.43%（6,535 人）的样本因外出不符合访问条件，符合访问条件的样本数为 50,620。拒访是导致符合访问条件样本无法转化为完访样本的最主要障碍，详见表 10。

表 10. CFPS 基线调查个人样本最终状态分布

样本状态	居委会		村委会		整体	
	数量	百分比	数量	百分比	数量	百分比
未联系	508	3.06	1,163	2.87	1,671	2.92
完访	12,793	77.03	29,797	73.49	42,590	74.52
拒访	1,752	10.55	2,533	6.25	4,285	7.50
其他原因未完成访问	539	3.25	1,535	3.79	2,074	3.63
不符合访问条件	1,015	6.11	5,520	13.61	6,535	11.43
合计	16,607	100	40,548	100	57,155	100

#### 3.2.2 不符合条件样本分布情况

57,155 名基因家庭成员中有 7,000 人不住在家中，占个人样本的 12.25%。最主要的外

出原因是外出工作，其次是外出读出和探亲访友。外出人员中有 465 人当前的居住地属于本区县，符合个人访问条件，详见表 11。

**表 11. CFPS 基线调查不符合访问条件个人样本分布**

不在家的原因	当前居住地属于 CFPS 调查区县		当前居住地不属于 CFPS 调查区县		整体	
	数量	百分比	数量	百分比	数量	百分比
外出读书	126	1.80	972	13.89	1,098	15.69
外出工作	294	4.20	5,135	73.36	5,429	77.56
出家	17	0.24	59	0.84	76	1.09
探亲访友	28	0.40	246	3.51	274	3.91
服刑					19	0.27
参军/服役					75	1.07
出境（包含港、澳、台）					29	0.41
合计	465	6.64	5,931	84.73	7,000	100

### 3.3 个人层面各类执行率计算

本报告中仅限于抽样单元层面完访的家庭，即在至少完成家庭成员问卷的受访家户范围内计算 CFPS 基线调查个人层面执行率。主要计算应答率（Response Rate，缩写为 RR）、合作率（Cooperation Rate，缩写为 COOP）、拒绝率（Refusal Rate，缩写为 REF）及联系率（Contact Rate，缩写为 CON）四个指标。

CFPS 基线调查中执行团队成功与 96.7% 符合条件的个人取得联系，个人层面整体应答率为 84.14%。村委会地区整体完成情况优于居委会地区，青少年样本完成情况优于成人样本。影响个人样本完访情况的最主要因素仍然是拒访，这一问题在居委会地区尤为明显。各类执行率的结果详见表 12。

**表 12. CFPS 基线调查个人层面各类执行率计算结果（%）**

执行率类型	社区类型		年龄段		整体
	居委会	村委会	成人	少儿	
$RR5=I/(I+R+NC+O)$	82.05	85.07	82.52	90.76	84.14
$COOP1=I/(I+R+O)$	84.81	87.99	85.69	92.34	87.01
$CON3=(I+R+O)/(I+R+NC+O)$	96.74	96.68	96.31	98.29	96.70
$REF3=R/(I+R+NC+O)$	11.24	7.23	9.05	6.08	8.47

## 附录：CFPS 基线调查抽样单元最终状态分类表

- 1 符合访问条件
  - 1.1 完访
  - 1.2 未完访
    - 1.2.1 拒访
    - 1.2.2 未联系上
    - 1.2.3 其他原因无法完成访问
  
- 2 不符合访问条件
  - 2.1 地址错误
  - 2.2 建筑物类型不符合条件
    - 2.2.1 非住宅
      - 2.2.1.1 商用建筑
      - 2.2.1.2 非传统住宅
    - 2.2.2 空置房屋
  - 2.3 住户过滤不合格
  
- 3 不确定是否符合访问条件
  - 3.1 未接触样本
    - 3.1.1 未分配样本
    - 3.1.2 无法到达访问地点
  - 3.2 正确地址无法确定房屋类型
  - 3.3 住宅无法完成住户过滤