

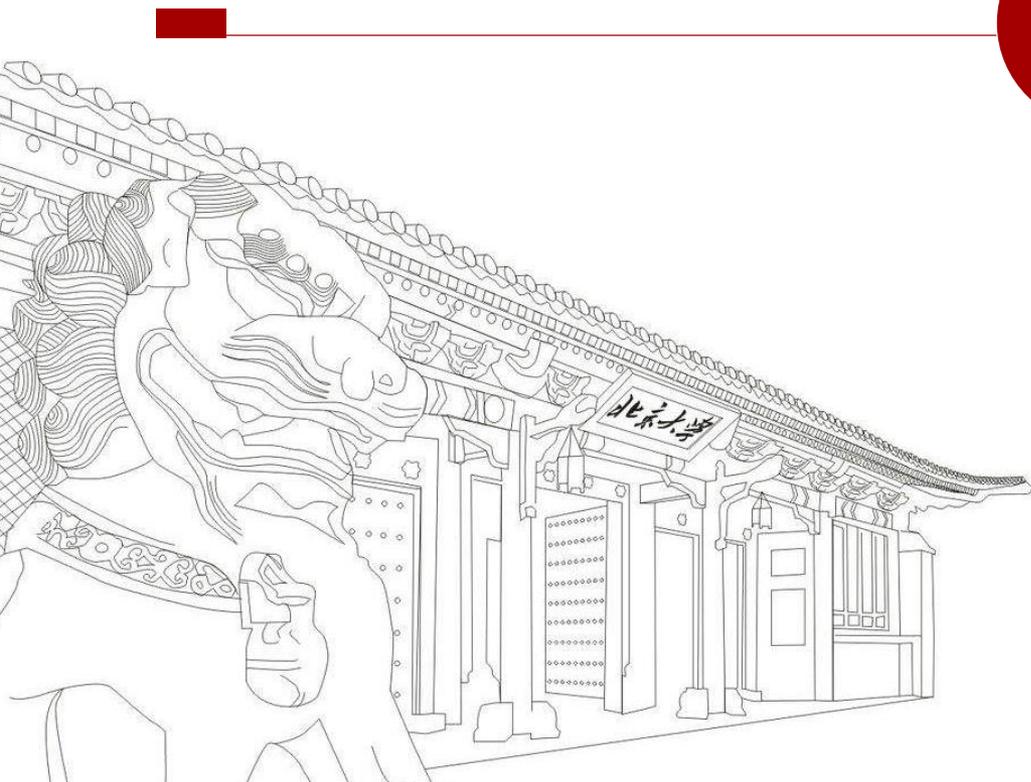


中国家庭追踪调查 China Family Panel Studies (CFPS)

北京大学
中国社会科学调查中心

吴琼

CFPS项目办公室



项目进展

项目时间轴

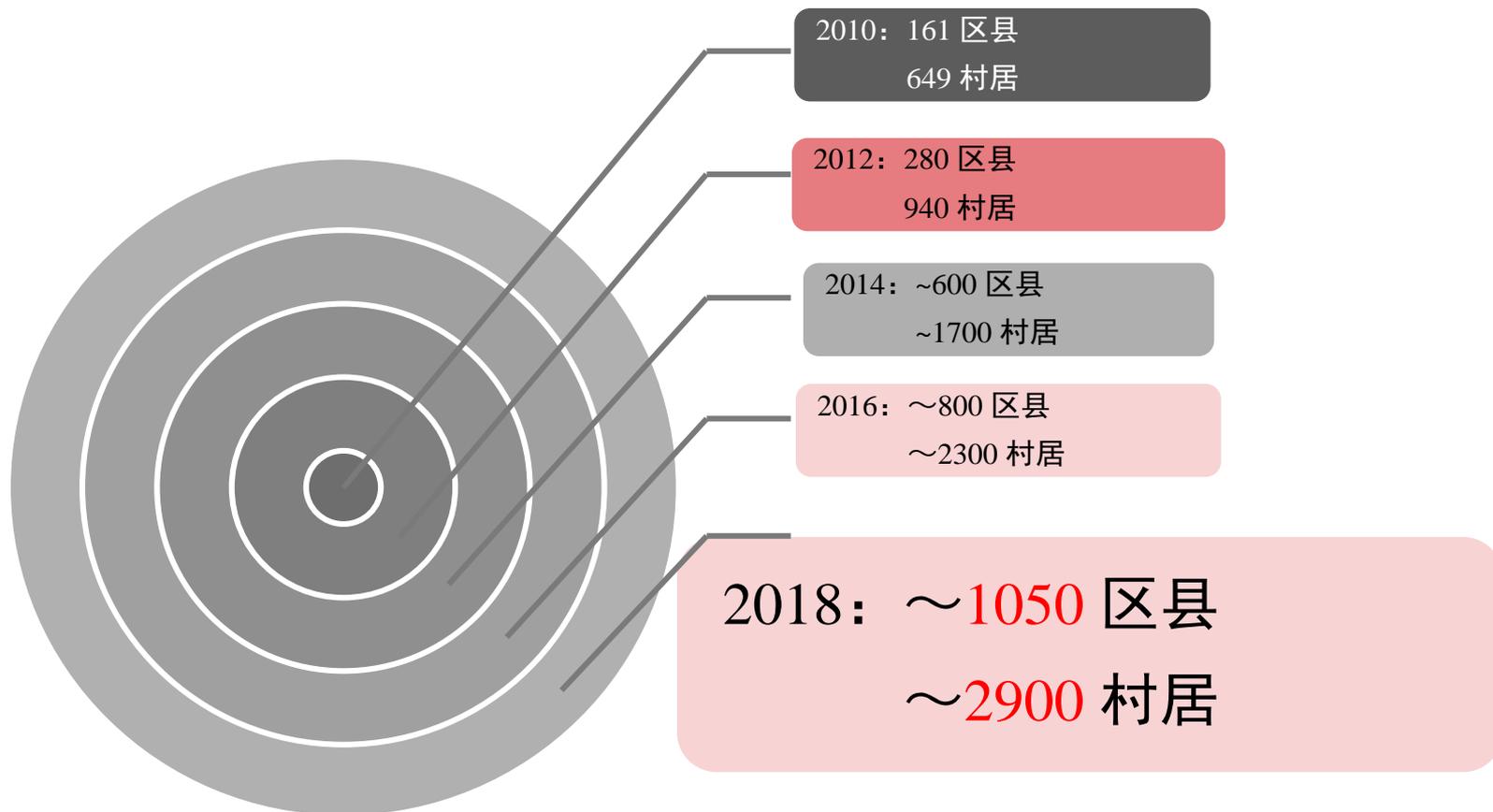


调查规模

| | 2010 | 2012 | 2014 | 2016 | 2018 |
|---------|--------|--------|--------|--------|---------|
| 完成访问样本量 | | | | | |
| 家庭层面 | 14,960 | 13,459 | 14,237 | 14,763 | 14,241 |
| 个人层面 | 42,590 | 42,125 | 42,067 | 42,698 | 38,640* |
| 社区层面 | 635 | NA | 621 | NA | NA |
| 跨轮追踪率 | | | | | |
| 家庭层面 | - | 85% | 89% | 89% | 86% |
| 个人层面 | - | 80% | 83% | 82% | 81% |
| 基线基因应答率 | 74% | 74% | 72% | 69% | 68% |

*暂不含个人代答样本

样本逐年扩散



电访比例逐年升高

| 问卷类型 | 2012 | 2014 | 2016 | 2018 |
|------|------|------|------|------|
| 家庭关系 | 1% | 8% | 26% | 28% |
| 家庭经济 | 1% | 4% | 17% | 21% |
| 成人 | 3% | 6% | 19% | 21% |
| 少儿 | 0.5% | 9% | 11% | |
| 总体 | 2% | 7% | 19% | 24% |

问卷进一步整合

面访/电访整合

| CFPS2016之前 | CFPS2016及之后 |
|--------------------------------------|--|
| <input type="checkbox"/> 面访问卷长于电访问卷 | <input type="checkbox"/> 面访与电访几乎相同（认知模块除外） |
| <input type="checkbox"/> 面访的题干比电访更详细 | <input type="checkbox"/> 面访电访题干相同 |
| <input type="checkbox"/> 部分题借助于卡片 | <input type="checkbox"/> 取消卡片 |

问卷进一步整合

成人/少儿问卷整合

| CFPS2018之前 | CFPS2018 |
|--|---|
| <ul style="list-style-type: none">□ 成人和少儿使用两套问卷 | <ul style="list-style-type: none">□ 10岁及以上个人使用同一套问卷，由年龄筛选模块 |
| <ul style="list-style-type: none">□ 成人和少儿形成单独数据集 | <ul style="list-style-type: none">□ 10岁以上个人在一个数据集中 |

CFPS2018新增模块

第一份工作

慈善行为

风险偏好

15岁以上人群的人格测试

10-15岁青少年的问题行为

发布数据

公开数据集（目标：调查总体结束后一年之内发布）

| 名称 | 2010 | 2012 | 2014 | 2016 | 2018 |
|-------|------|------|------|------|------|
| 家庭关系库 | √ | √ | √ | √ | ? |
| 家庭库 | √ | √ | √ | √ | √ |
| 成人库 | √ | √ | √ | √ | √ |
| 少儿库 | √ | √ | √ | √ | |
| 社区库 | √ | NA | √ | NA | NA |
| 跨年id库 | NA | √ | √ | √ | ? |

其它数据

其它公开数据集

- 2011年维护调查（家庭库以及18岁以下个人问卷）
- 2008、2009年（三省试调查及其追踪调查）

限制使用数据集（restricted use data）

- 区县数据库（不含具体区县名）
- 方言库
- 部分未发布变量（如访问的具体日期等）

保密机

地址数据

发布库中地址信息

- 具体省和直辖市名称
- 区县级、村居级以CFPS自编代码表示

“我能否申请具体区县名称？”

- “不能”

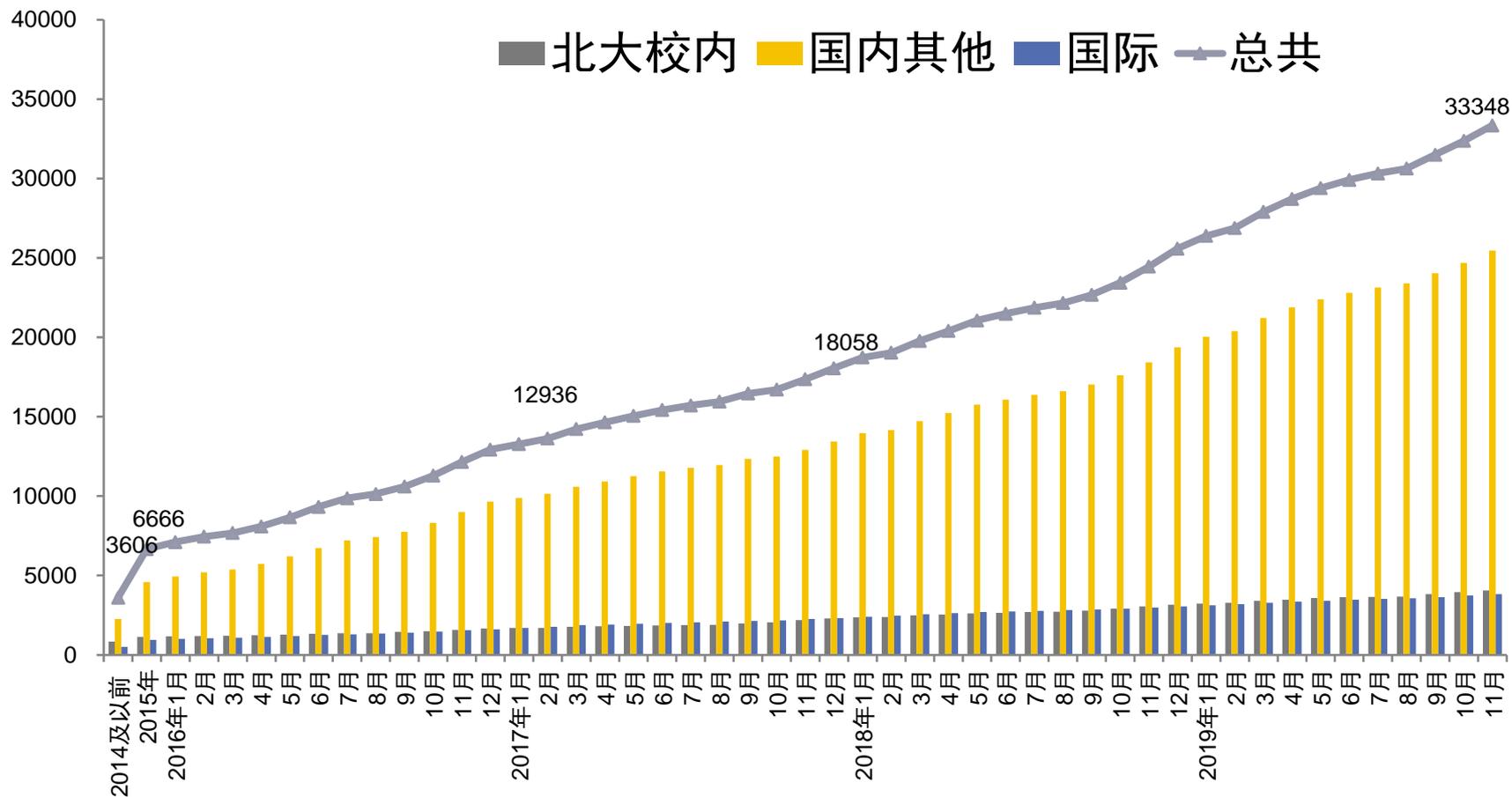
地址数据

你可以-

- 申请“区县数据库”（含2010年区县宏观指标）
- 申请使用“保密机”
- 申请附加服务

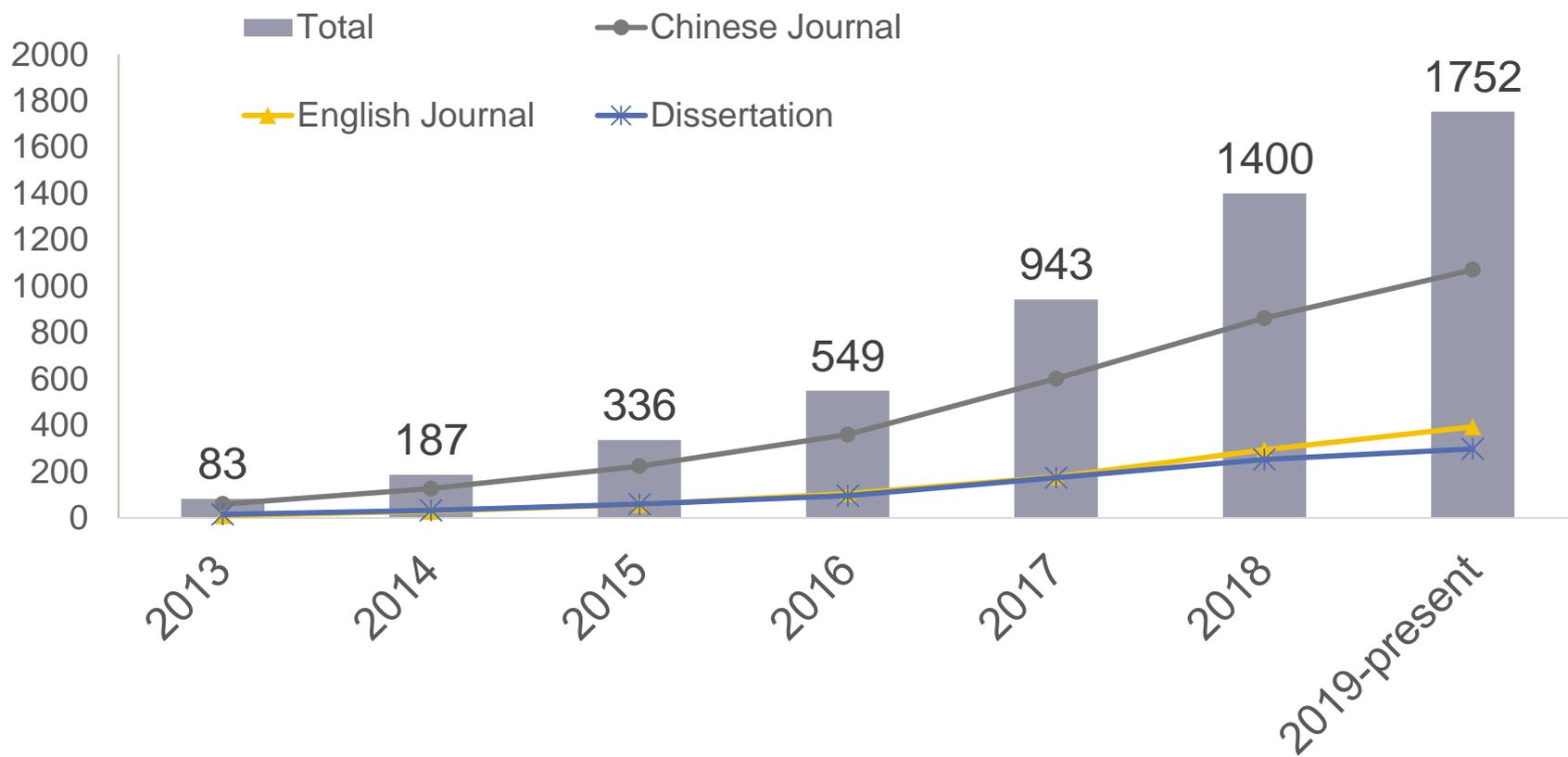


用户统计



用户成果

CFPS用户发表历年累积量



学科分布

| 学科 | 用户分布 | 成果分布 |
|-------|------|-------|
| 经济学 | 45% | 31.0% |
| 社会学 | 15% | 13.9% |
| 教育学 | 2% | 3.4% |
| 数学和统计 | 4% | 2.7% |
| 管理学 | 19% | 2.7% |
| 公共卫生 | 1% | 2.7% |
| 环境和资源 | 1% | 1.0% |
| 心理学 | 1% | 0.9% |
| 政治 | 2% | 0.7% |
| 生物 | <1% | 0.6% |

CFPS2018测试版

发布平台

网站首页 | 关于CFPS | 相关新闻 | 文档中心 | **数据中心** | 常见问题 | 学术研究 | 招聘信息 | 联系我们

当前位置： 首页 > 数据中心 > 公开数据

栏目导航

- 公开数据
- 限制数据

CFPS公开数据

请访问[北京大学开放研究数据平台](#)，注册成为“**高级用户**”并点击“**申请**”按钮，加入相关用户组，审核通过后方可下载数据(请使用Stata14打开Stata数据集，使用SAS9.4打开SAS数据集)。审核工作**需要1至2天**，感谢您对CFPS的支持与信赖！

发布内容

CFPS2018家庭库

CFPS2018个人自答库的大部分模块

- (10岁以上所有个人)

CFPS2018少儿的家长代答库

- 0-15岁少儿的家长代答问卷

2018年中文版问卷

- “文档中心” -> “调查问卷”

家庭库、个人自答、少儿代答的codebook

- “文档中心” -> “数据文档”

测试版不包括

家庭关系库

个人代答

自答库中的部分模块

- 迁移模块
- 权数
- 省级以下地址等

测试版使用事项：个人自答

CFPS2016及以前

- 分少儿（16岁以下）和成人问卷（16岁及以上）数据集
- 同一个变量在少儿和成人数据中有可能变量名不同

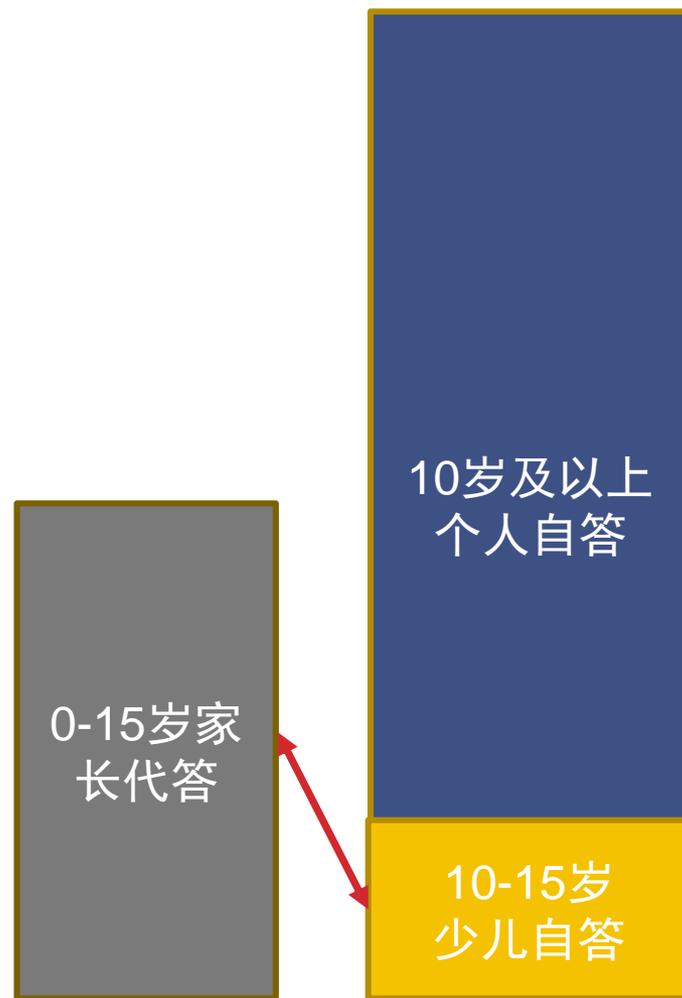
CFPS2018开始

- 10岁及以上的个人全部在个人自答数据集中
- 0-15岁少儿父母/监护人的代答数据单独放在家长代答数据中

以前的少儿库呢？



2010-2016少儿库



2018

测试版使用注意事项

样本代表性问题

- 测试版没有权数，用subsample=1得到近似全国代表性样本

地址信息

- 测试版的个人自答和家庭库只包括省级信息
- 少儿的家长代答库包括省级信息和区县顺序码
- 村居编码等到正式版发布
- urban18变量是根据村居所在地的国统局定义赋值的

测试版使用注意事项

综合变量

- 收入：家庭总收入、家庭人均收入、分项收入

| | |
|----------------|-----------------------|
| fwage_1 | 工资性收入 (调整后) |
| fwage_2 | 工资性收入 (与2010年可比) |
| felse_1 | 其他收入 |
| felse_2 | 其他收入 (与2010年可比) |
| fincome1 | 全部家庭纯收入 |
| fincome2 | 全部家庭纯收入 (与2010年可比) |
| fincome1_per | 人均家庭纯收入 |
| fincome1_per_p | 人均家庭纯收入分位数 |
| fincome2_per | 人均家庭纯收入 (与2010年可比) |
| fincome2_per_p | 人均家庭纯收入分位数 (与2010年可比) |
| foperate_1 | 经营性收入 |
| foperate_2 | 经营性收入 (与2010年可比) |
| fproperty_1 | 财产性收入 |
| fproperty_2 | 财产性收入 (与2010年可比) |
| ftransfer_1 | 转移性收入 |

测试版使用注意事项

综合变量

- 支出

| | |
|---------|------------|
| daily | 家庭设备及日用品支出 |
| dress | 衣着鞋帽支出 |
| eec | 文教娱乐支出 |
| eptran | 转移性支出 |
| epwelf | 福利性支出 |
| expense | 家庭总支出 |
| food | 食品支出 |
| house | 居住支出 |
| med | 医疗保健支出 |
| mortage | 房贷支出 |
| other | 其他消费性支出 |
| pce | 居民消费性支出-加总 |
| trco | 交通通讯支出 |

测试版使用注意事项

教育

| | |
|--------------|--------------------------|
| cfps2018sch | CFPS2018个人问卷受访者离校/上学阶段 |
| cfps2018edu | CFPS2018个人问卷受访者已完成的最高学历 |
| cfps2018eduy | CFPS2018个人问卷受访者已完成的受教育年限 |

认知测试

| | |
|----------------|---------------------|
| wordtest18_sc2 | 18年词组测试题得分：18年问卷算法 |
| wordtest18 | 18年词组测试题得分：与10年可比算法 |
| mathtest18_sc2 | 18年数学测试题得分：18年问卷算法 |
| mathtest18 | 18年数学测试题得分：与10年可比算法 |

测试版使用事项

加载变量

| | |
|----------|-----------------------|
| iintervy | 加载变量: 最近一次调查个人自管长卷完访年 |
| gene | 加载变量: 是否属于基因成员 |
| ibirthy | 加载变量: 出生年份 |
| gender | 加载变量: 受访者性别 |
| minzu | 加载变量: 民族数据是否有数据 |
| birthw | 加载变量: 出生时体重是否有数据 |
| birthp | 加载变量: 出生时家庭住址 |
| a12p | 加载变量: 12岁时家庭住址 |
| a12hk | 加载变量: 12岁时户口 |
| military | 加载变量: 是否是退伍军人 |
| edu_last | 加载变量: 最近一次调查最高学历 |
| r1_last | 加载变量: 最近一次调查上学阶段 |
| train | 加载变量: 是否乘坐过火车 |
| plane | 加载变量: 是否乘坐过飞机 |
| m1m | 加载变量: M1M主观态度题是否有数据 |
| m401 | 加载变量: M401主观态度题是否有数据 |

ibirthy 加载变量: 出生年份

ibirthy_update 出生年份更新

数据质量评估：收入

| | 均值 | 中位数 |
|----|----------|----------|
| 全国 | 28685.11 | 16666.67 |
| 上海 | 70646.33 | 42000.00 |
| 辽宁 | 30223.04 | 20000.00 |
| 河南 | 22480.32 | 14274.11 |
| 甘肃 | 24547.65 | 11666.67 |
| 广东 | 29373.17 | 17500.00 |

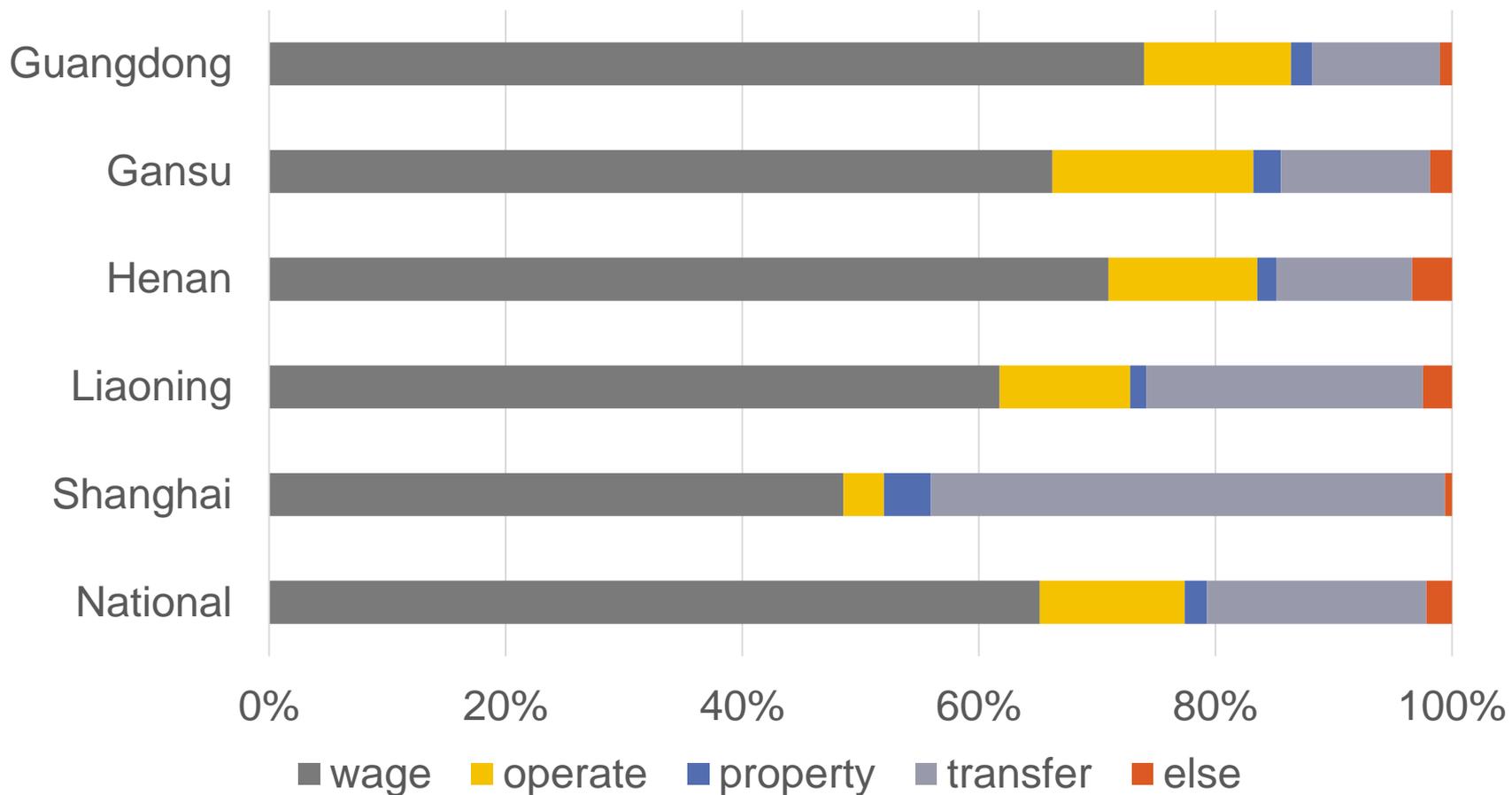
数据质量评估：收入跨轮变化（截面）

| | | CFPS | | |
|----|----|-------|-------|-------|
| | | 2016 | 2018 | ratio |
| 全国 | 城镇 | 30937 | 37846 | 1.22 |
| | 农村 | 14980 | 17838 | 1.19 |
| 上海 | 城镇 | 59456 | 73488 | 1.24 |
| | 农村 | 27000 | 36137 | 1.17 |
| 辽宁 | 城镇 | 31011 | 39053 | 1.26 |
| | 农村 | 13404 | 19012 | 1.42 |
| 河南 | 城镇 | 23057 | 29821 | 1.29 |
| | 农村 | 12367 | 14760 | 1.19 |
| 甘肃 | 城镇 | 28338 | 43974 | 1.55 |
| | 农村 | 12145 | 13632 | 1.12 |
| 广东 | 城镇 | 33645 | 38311 | 1.14 |
| | 农村 | 14288 | 15243 | 1.07 |

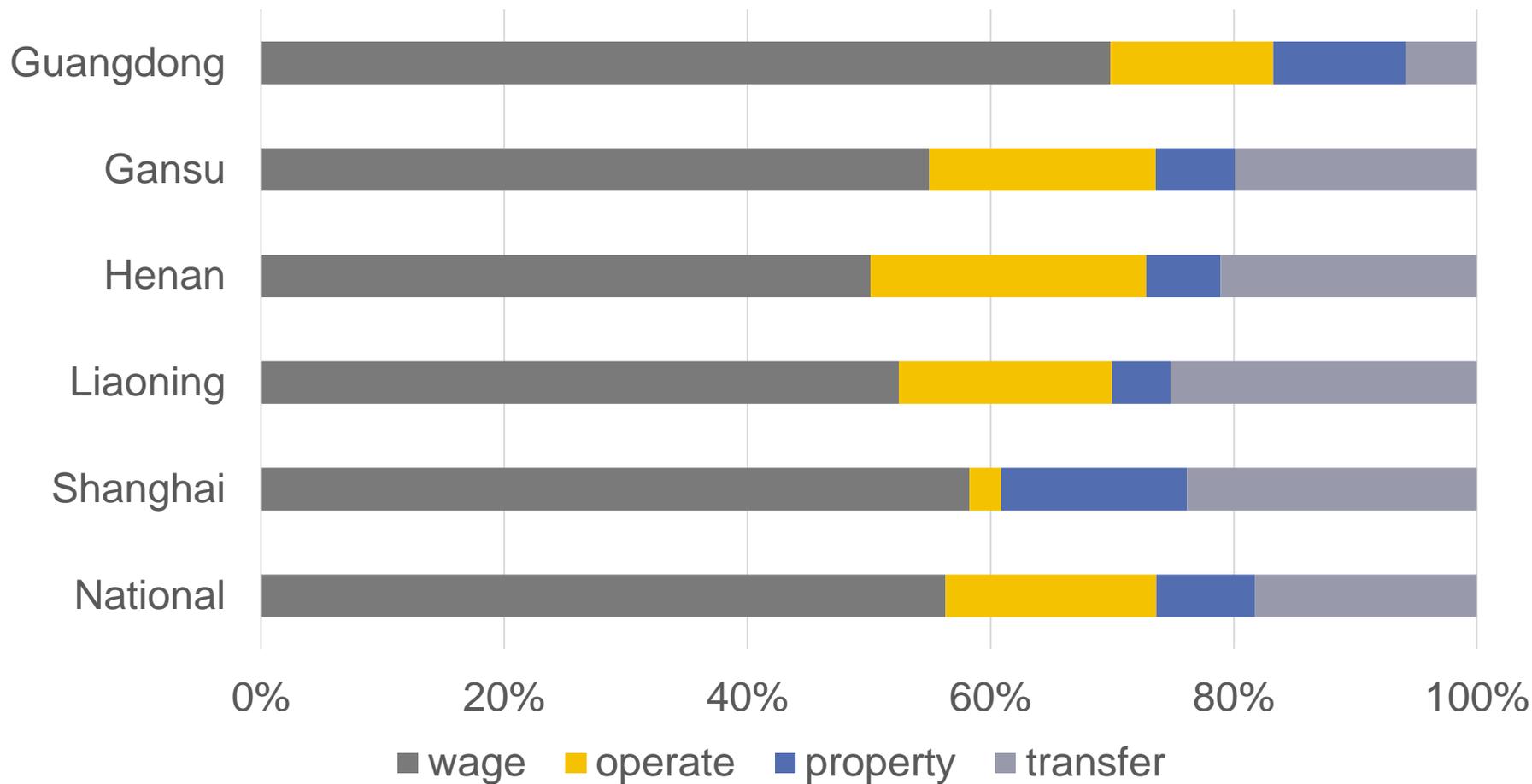
数据质量评估：收入跨轮变化（面板）

| | 2016 | | 2018 | | 2018/2016 ratio | | N |
|----|----------|----------|----------|----------|-----------------|------|------|
| | 均值 | 中位数 | 均值 | 中位数 | 均值 | 中位数 | |
| 全国 | 22598.20 | 13333.33 | 28219.67 | 16500.00 | 1.25 | 1.24 | 7836 |
| 上海 | 64158.02 | 34000.00 | 64918.32 | 40010.00 | 1.01 | 1.18 | 705 |
| 辽宁 | 22643.10 | 15550.00 | 30049.59 | 20000.00 | 1.33 | 1.29 | 1162 |
| 河南 | 16993.49 | 11401.00 | 21917.27 | 14050.40 | 1.29 | 1.23 | 1391 |
| 甘肃 | 15556.23 | 10000.00 | 24644.02 | 11666.67 | 1.58 | 1.17 | 1425 |
| 广东 | 25036.21 | 15459.17 | 29237.17 | 17500.00 | 1.17 | 1.13 | 1018 |

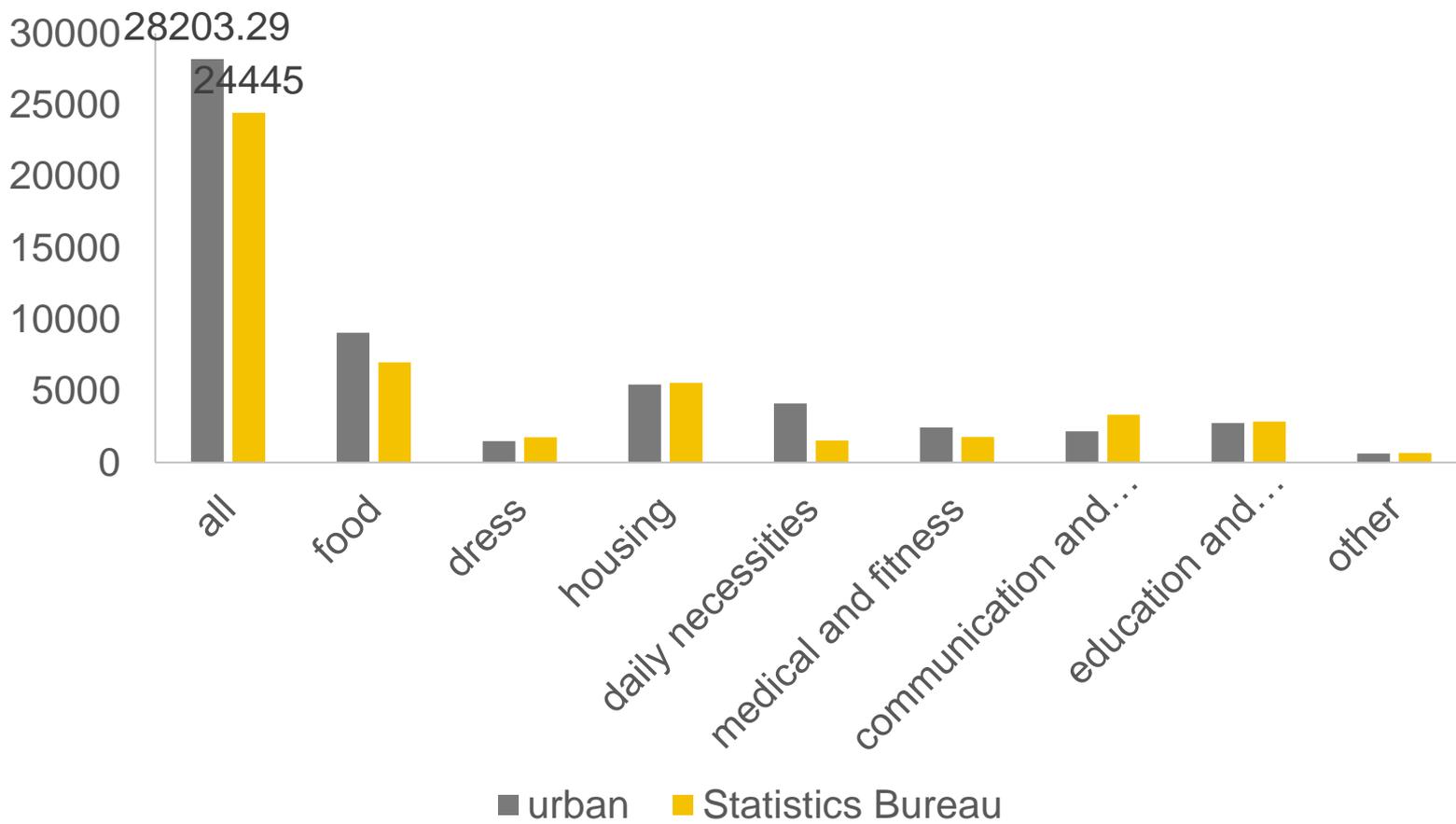
收入结构: CFPS



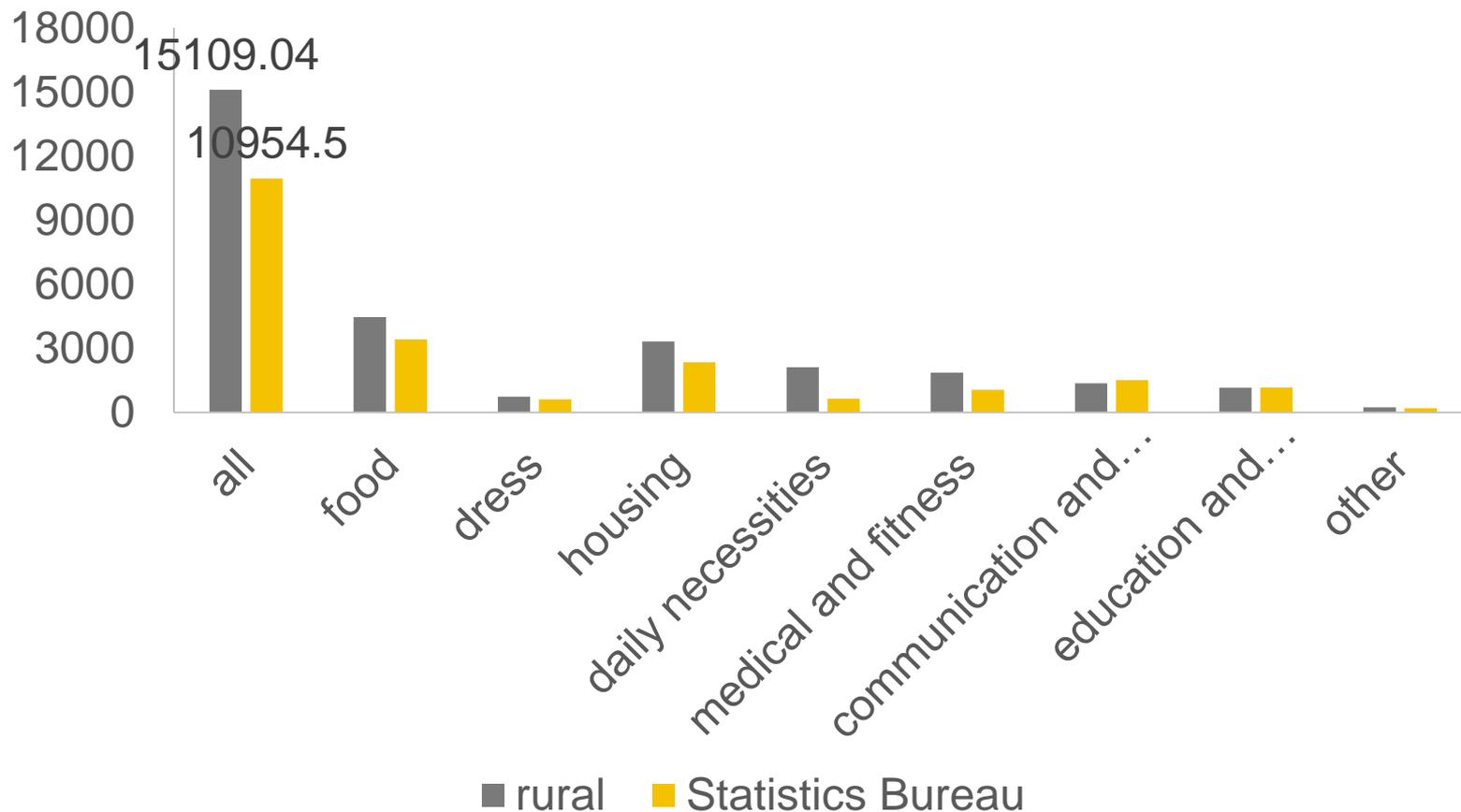
收入结构：国统局



数据质量评估：城镇家庭支出



数据质量评估：农村家庭支出



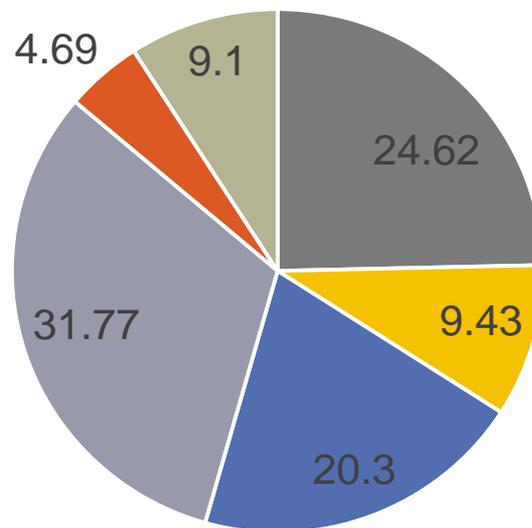
数据评估：第一份工作

采集了近18,000个样本的第一份工作信息

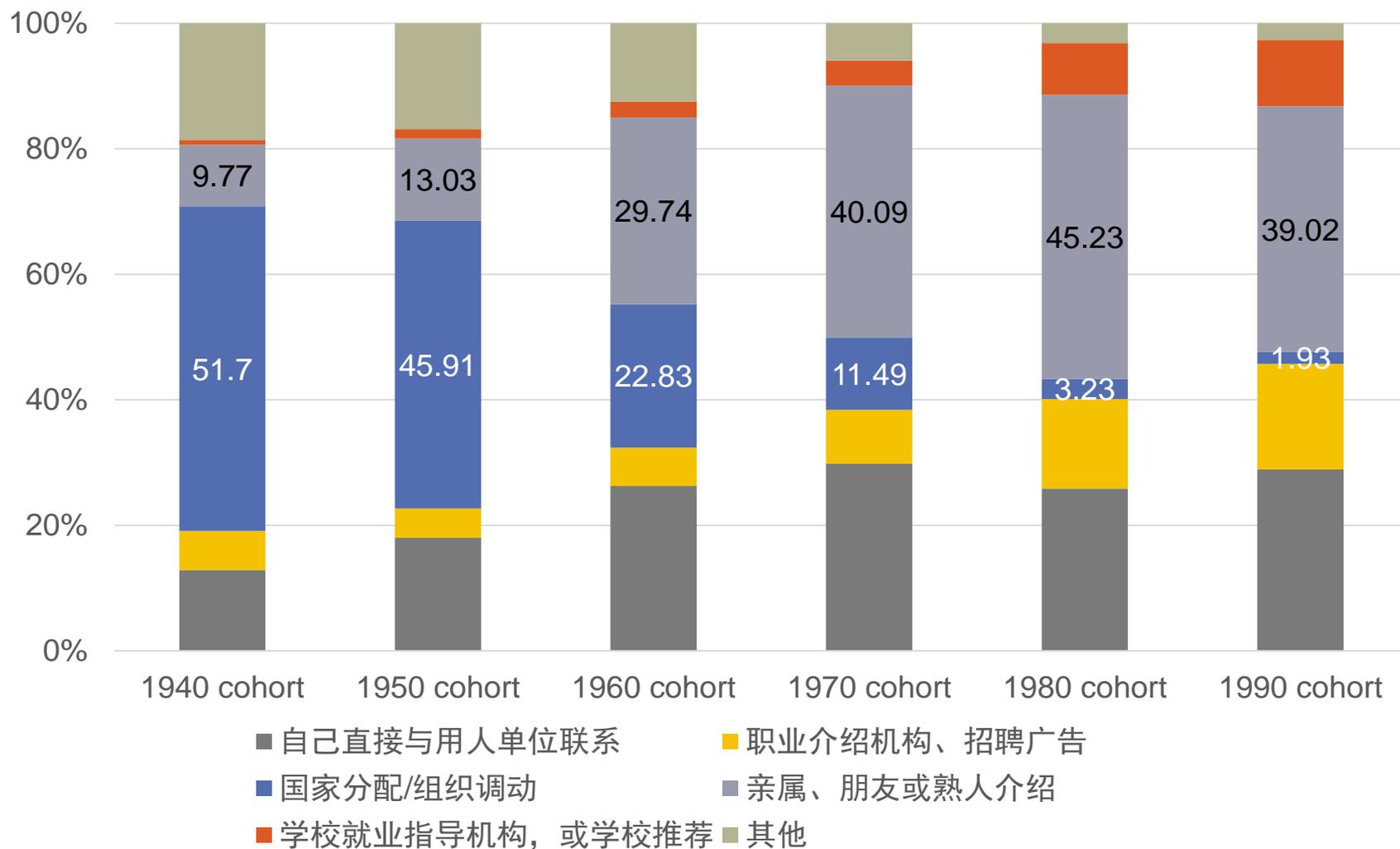
雇主性质、求职渠道、单位所在行业、职业信息、所需教育程度

第一份工作的求职渠道

- 自己直接与用人单位联系
- 职业介绍机构、招聘广告
- 国家分配/组织调动
- 亲属、朋友或熟人介绍
- 学校就业指导机构，或学校推荐

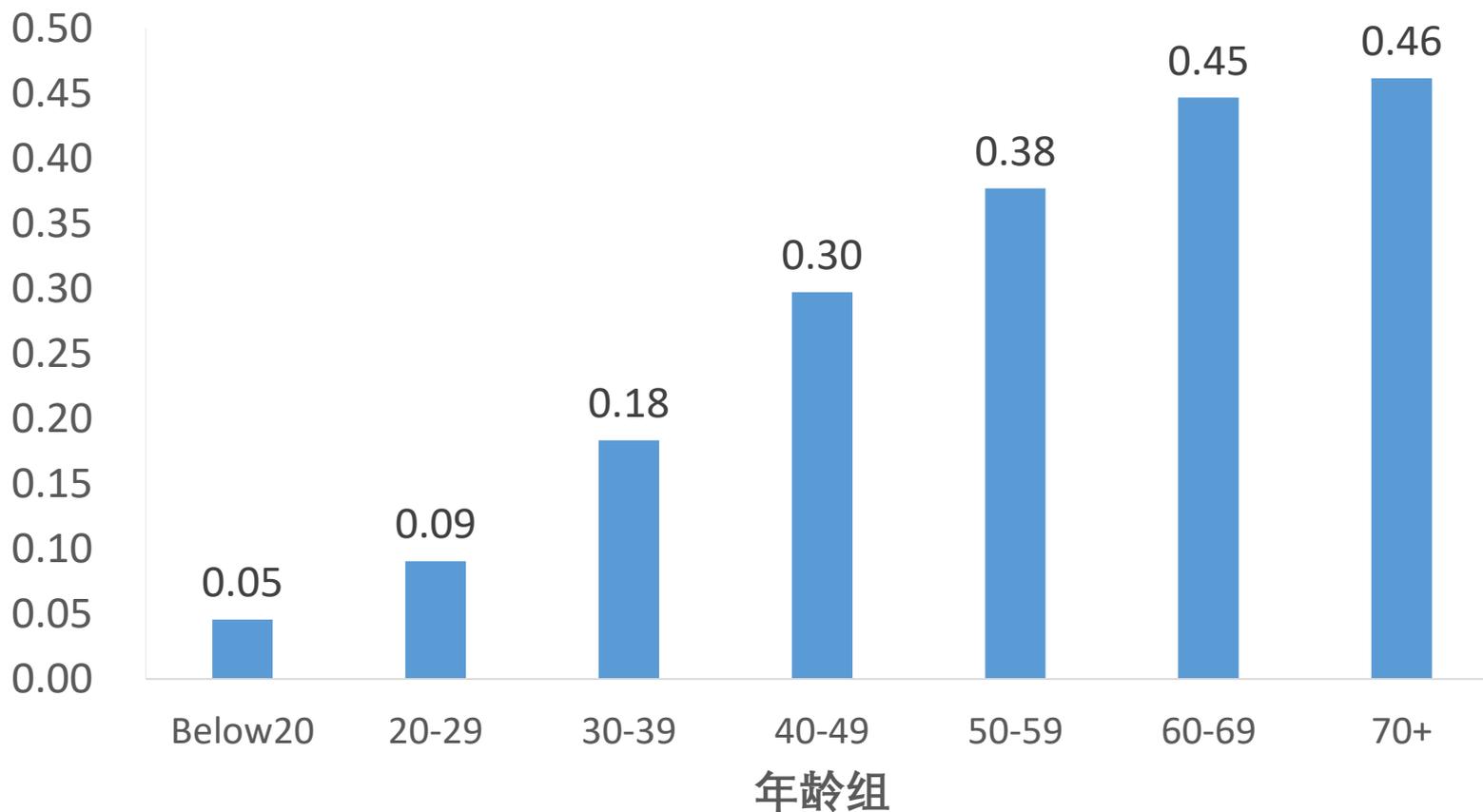


分出生队列



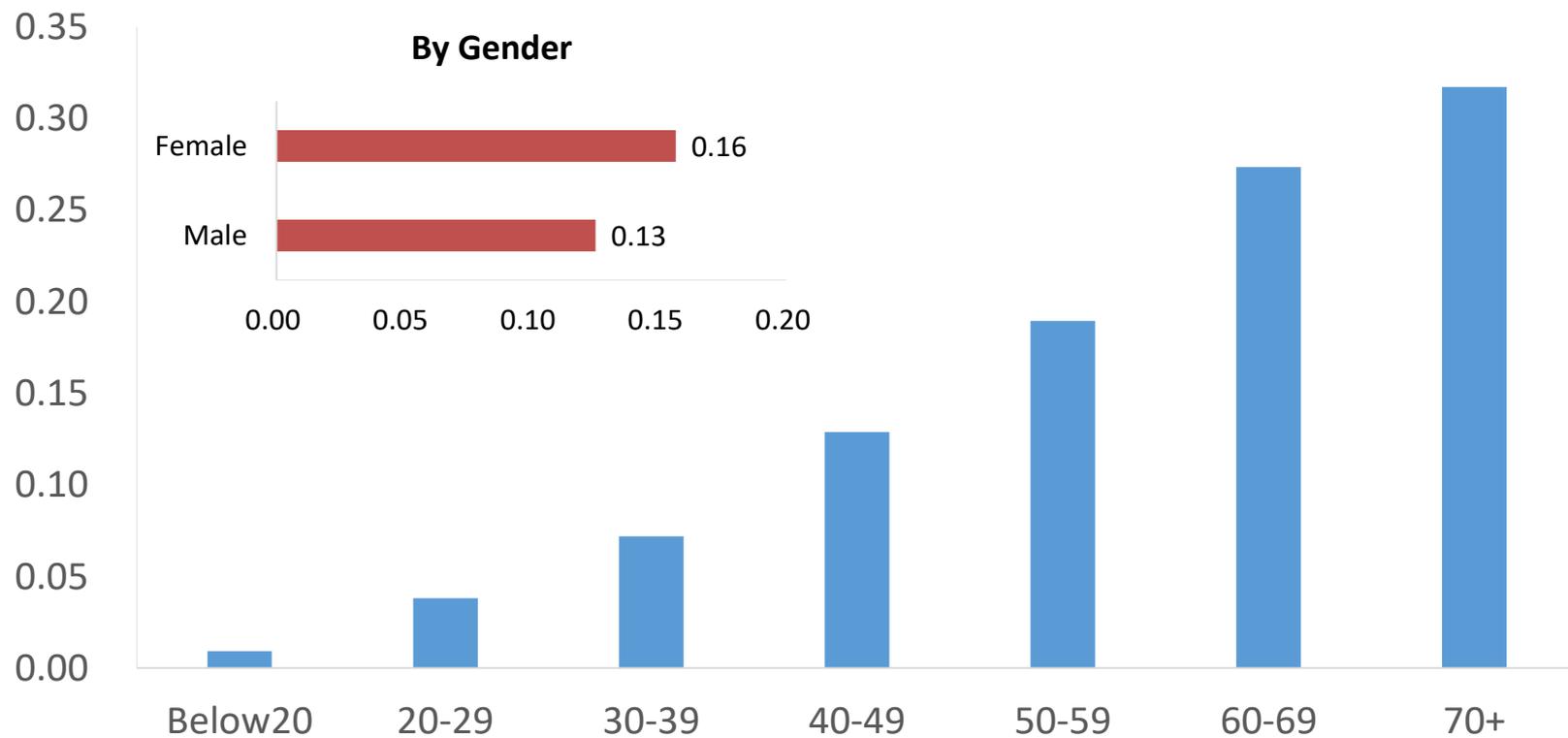
数据质量评估：健康

“不健康” / “一般” 占比

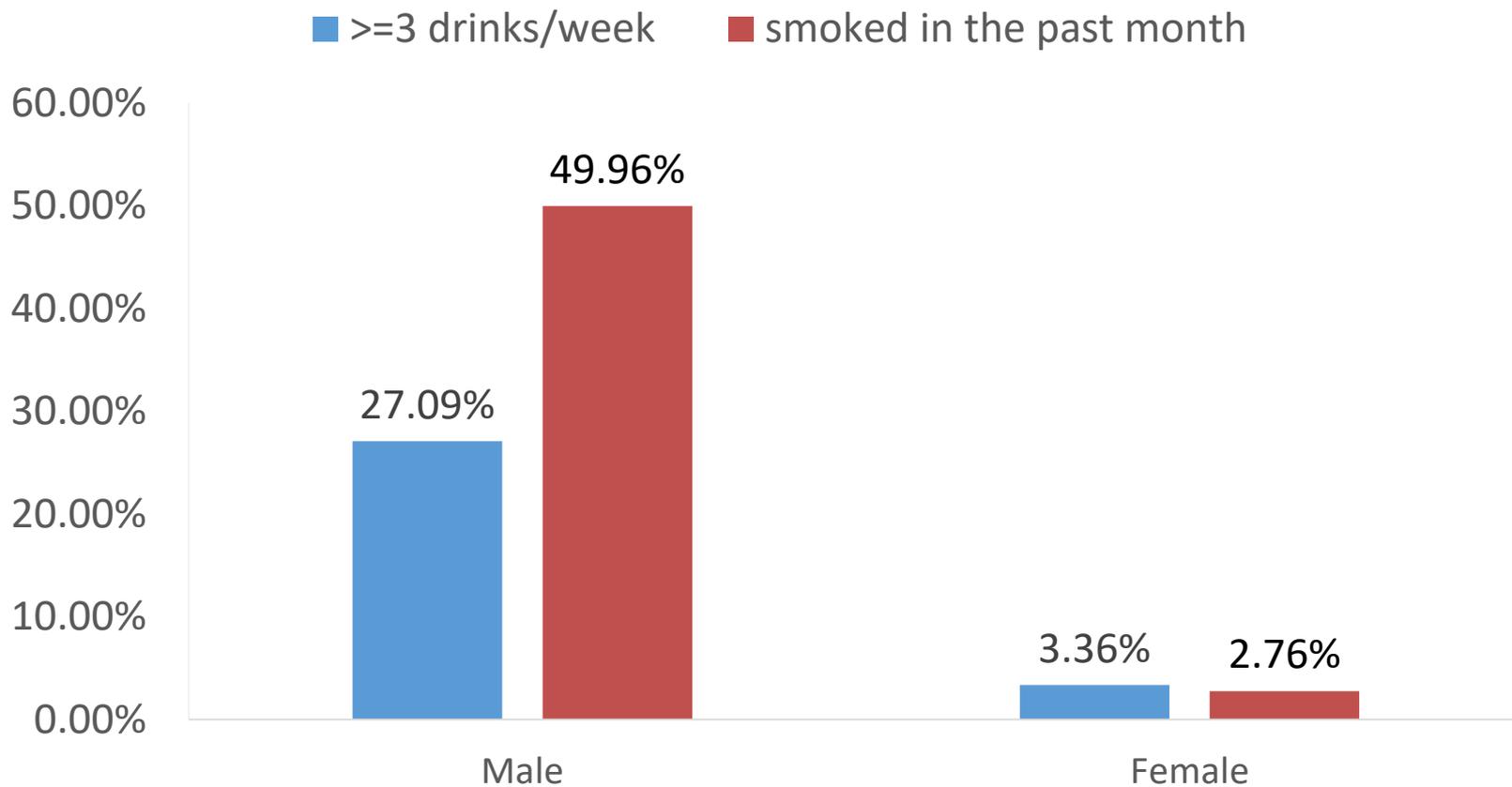


慢性病

过去六个月经医生诊断的慢性病



饮酒和吸烟





联系我们



isss.cfps@pku.edu.cn