**Perception of Tax and Subsidies in Nigeria**

**Field Technical and Methodological Report**

***Prepared for:***

**Nigerian Economic Summit Group (NESG)**

***By:***

**Practical Sampling International (PSI)**

**August 2018**

**TABLE OF CONTENTS**

**Introduction 3**

**Project Overview 3**

**Qualitative 3**

*Focus Group Discussion* 3

*In-depth Interviews (IDI)* 4

**Fieldwork and Challenges during the Qualitative Interview 5**

*Fieldwork*  5

*Quality Control* 5

*Challenges during Fieldwork*  5

**Quantitative 5**

*Sample Distribution* 5

**Selection of Sampling Units/Enumeration Areas/LGAs 8**

*Household Survey* **8**

*Firm Survey* **9**

**Interviewer Selection 10**

*How interviews were selected for training and fieldwork* 10

*Total number of interviewers and supervisors recruited for the training*  10

*Total number of interviewers and supervisors selected after training* 10

**Enumerators Training 11**

*Mock Sessions*  11

*Actual Fieldwork Duration* 11

**Household Survey 11**

*Sampling Procedure* 11

*Allocation of Interviewers/Walk Pattern* 12

*Selection of Dwelling Structure* 12

*Selection of Household* 12

*Selection of Respondent* 12

*Call back/Substitution* 12

*Questionnaire Translation* 13

*Coverage/Spread*  13

**Firm Survey**

*Sampling Approach* 13

*Coverage/Spread* 13

**Enforcement of Quality Controls** 13

**Cases and Issues during fieldwork 20**

**Appendix 21**

**Weighting** 22

*Household Survey*  22

*Firm Survey*  22

# INTRODUCTION

The purpose of this survey is to collect views of ordinary Nigerian citizens on their perception regarding taxes, subsidies and expenditure.

**PROJECT OVERVIEW**

Basically, the study was aimed at understanding the general views of citizen in Nigeria regarding perceptions on tax, subsidy and expenditure system in Nigeria.

The study was segmented into two Categories; Qualitative and Quantitative.

Quantitative Segment comprises of both the household and firm survey and was conducted across the thirty-six states (36) including Abuja, the Federal Capital Territory.

Qualitative Segment comprises of Focus Group Discussions (FGD) and In Depth Interviews (IDIs) across the 6 geo- political regions.

**PRETEST**

Before the commencement of fieldwork, a pretest /pilot activities was conducted to fine tune the research instruments (qualitative and quantitative) and understand the project intricacies. The pretest activities was spread over a nine day period and includes presentation of the study overview and piloting of the research questionnaires and discussion guides in selected urban and rural areas in Lagos. The questionnaire/discussion guides were check for consistency, ambiguity in wordings, and Length.

The quantitative research questionnaire was translated to the 4 main languages spoken in the Nigeria (i.e Yoruba, Igbo, Hausa, and Pidgin) apart English. The different versions of the questionnaire were piloted to check for consistency and effectiveness.

**QUALITATIVE**

***Focus Group Discussion (FGD*)**

A total of 48 FGDs were conducted across the 6 geo- political zone of the country with 24 among households (12 each for tax and subsidies) and 24 among firms. The discussions were held in one urban and one rural location in each of the selected state.

Respondents for the household FGDs were males and Females (18+ years) who have lived within the community for at least 2 years and are well informed and familiar with major issues to be discussed, while respondents for the firm FGDs were owners and managers of businesses between 5 to 49 employees.

The Focus Group Discussions were conducted in two selected states from the 6 geo- political zones. The selected states were Lagos, Ogun, Edo, Rivers, Abia, Enugu, Plateau, Nasarawa, Kaduna, Kano, Bauchi and Adamawa State.

**FGD Structure**

The FGD structure is indicated below;

**Household**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Region** | **State** | **No of FGDs** | **Household Tax** | **Household Subsidies** |
| **Urban** | **Rural** | **Urban** | **Rural** |
| **M** | **F** | **M** | **F** | **M** | **F** | **M** | **F** |
| South West | Ogun | 2 | 1 | - | - | - | - | 1 | - | - |
| Lagos | 2 | - | - | - | 1 | 1 | - | - | - |
| South East | Abia | 2 | - | 1 | 1 | - | - | - | - | - |
| Enugu | 2 | - | - | - | - | 1 | - | - | 1 |
| South South | Port Harcourt | 2 | - | 1 | - | - | - | - | 1 | - |
| Edo | 2 | 1 | - | - | - | - | 1 | - | - |
| North East | Nassarawa | 2 | 1 | - | - | - | - | 1 | - | - |
| Plateau | 2 | - | - | - | 1 | 1 | - | - | - |
| North West | Kano | 2 | 1 | - | - | - | - | 1 | - | - |
| Kaduna | 2 | - | 1 | - | - | - | - | 1 | - |
| North Central | Bauchi | 2 | - | 1 | 1 | - | - | - | - | - |
| Adamawa | 2 | - | - | - | - | 1 | 1 | - | - |
| **Total**  | **24** | **4** | **4** | **2** | **2** | **4** | **5** | **2** | **1** |

**Firm**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Region** | **State** | **No of FGDs** | **Urban** | **Rural** |
| South West | Ogun | 2 | 1 | 1 |
| Lagos | 2 | 2 | - |
| South East | Abia | 2 | 1 | 1 |
| Enugu | 2 | 1 | 1 |
| South South | Port Harcourt | 2 | 2 | - |
| Edo | 2 | 1 | 1 |
| North East | Nassarawa | 2 | 1 | 1 |
| Plateau | 2 | 2 | - |
| North West | Kano | 2 | 1 | 1 |
| Kaduna | 2 | 2 | - |
| North Central | Bauchi | 2 | 1 | 1 |
| Adamawa | 2 | 2 | - |
| **Total**  | **24** | **17** | **7** |

***In Depth Interviews (IDI)***

In addition to the FGDs, 17 In Depth Interviews were conducted across the 6 geo- political zones. The In- Depth interviews were conducted with senior representative of tax authority at the state and LGA levels in each of the six geo – political zones.

Specifically, the In- Depth interviews were conducted across the following locations; Ogun, Plateau, Bauchi, Kano, Rivers and Abia states.

In each state, one senior member of the state internal revenue service was interviewed along with one senior member in the LGA revenue service in a randomly selected urban LGA and one senior member in the LGA revenue service in a randomly selected rural

**IDI Structure**

The IDI structure is indicated below;

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Region** | **State** | **Quota** | **State Internal Revenue Service** | **Urban LGA Revenue Service** | **Rural LGA Revenue Service** | **Total** |
| South West | Ogun | 2 | 1 | 1 | - | 2 |
| South East | Abia | 3 | 1 | 1 | 1 | 3 |
| South South | Rivers | 3 | 1 | 1 | 1 | 3 |
| North East | Nassarawa | 3 | 1 | 1 | 1 | 3 |
| North West | Kano | 3 | 1 | 1 | 1 | 3 |
| North Central | Bauchi | 3 | 1 | 1 | 1 | 3 |
| **Total** | **17** | **6** | **6** | **5** | **17** |

**Fieldwork and Challenges during the Qualitative Interviews**

**Fieldwork**

The FGDs and IDIs interviews were conducted by experienced moderators who had worked on similar project in the past

The interview sessions were recorded on tapes so as to fully capture on all the information from the interview. The recorded information was fully transcribed for use during analysis and reporting.

**Quality Control**

Experienced field recruiters that have recruited on several similar projects in the past were engaged to contact and recruit the respondents

To reduce chances of wrong recruitments, the field supervisors and quality controllers randomly call on the recruited respondents independently before the interview session to verify authenticity of the contact and adherence to specified respondent’s profile.

**Challenges during Fieldwork**

Generally, there was no major challenge in conducting the FGD sessions across locations. For the IDI interviews, some state tax offices were slow in giving requested directive permitting tax officials to be interviewed. In Ogun state, permission was not granted by the state revenue service to interview the rural revenue service officer.

**QUANTITATIVE**

The quantitative survey comprises both household interviews and firm interviews. The household survey included **10,000** main interviews and **6,000** over sample interviews. The Firm survey also included **5,000** main interviews and **3,000** over sample interviews.

#### Sample Distribution

The main samples (both household and firm) were distributed proportionally to the 36 states and FCT using information from the 2006 census figure and National Population Commission while the over sample was equally distributed across 6 selected states. The allocation of primary sampling unit/ enumeration areas to rural and urban was not proportional to the population of the stratum in each state. In order to ensure sufficient number of respondents that have some experience of the tax system in the sample, 70% of the interview per state was conducted in the urban and 30% in rural. The sample distribution is indicated below;

|  |
| --- |
| **MAIN SAMPLE HOUSEHOLD** |
| **Region** | **States** | **Grand Total by State** |
| NORTH CENTRAL | BENUE | 310 |
| FCT ABUJA | 100 |
| KOGI | 240 |
| KWARA | 170 |
| NASARAWA | 140 |
| NIGER | 290 |
| PLATEAU | 230 |
| **TOTAL** |  | **1480** |
| NORTH EAST | ADAMAWA | 235 |
| BAUCHI | 341 |
| BORNO | 300 |
| GOMBE | 172 |
| TARABA | 170 |
| YOBE | 170 |
| **TOTAL** |  | **1388** |
| NORTH WEST | JIGAWA | 310 |
| KADUNA | 444 |
| KANO | 670 |
| KATSINA | 420 |
| KEBBI | 231 |
| SOKOTO | 273 |
| ZAMFARA | 241 |
| **TOTAL** |  | **2589** |
| SOUTH EAST | ABIA | 211 |
| ANAMBRA | 301 |
| EBONYI | 164 |
| ENUGU | 241 |
| IMO | 283 |
| **TOTAL** |  | **1200** |
| SOUTH SOUTH | AKWA IBOM | 282 |
| BAYELSA | 121 |
| CROSS RIVER | 210 |
| DELTA | 302 |
| EDO | 240 |
| RIVERS | 383 |
| **TOTAL** |  | **1538** |
| SOUTH WEST | EKITI | 180 |
| LAGOS | 651 |
| OGUN | 275 |
| ONDO | 250 |
| OSUN | 251 |
| OYO | 400 |
| **TOTAL** |  | **2007** |
| **Grand Total** |   | **10202** |
|  |  |  |
| **OVER SAMPLE HOUSEHOLD** |
| **Region** | **States** | **Grand Total** |
| North Central | Nasarawa | 1005 |
| North East | Bauchi | 1002 |
| North West | Kano | 1005 |
| South East | Abia | 1006 |
| South South | Rivers | 1002 |
| South West | Ogun | 1006 |
| **Grand Total** |  | **6026** |

|  |
| --- |
| **MAIN SAMPLE - FIRM** |
| **REGION** | **STATE** | **Grand Total BY State** |
| **NORTH CENTRAL** | BENUE | 158 |
| FCT ABUJA | 52 |
| KOGI | 127 |
| KWARA | 89 |
| NASARAWA | 71 |
| NIGER | 159 |
| PLATEAU | 118 |
| **TOTAL** |  | **774** |
| **NORTH EAST** | ADAMAWA | 127 |
| BAUCHI | 182 |
| BORNO | 152 |
| GOMBE | 89 |
| TARABA | 89 |
| YOBE | 95 |
| **TOTAL** |  | **734** |
| **NORTH WEST** | JIGAWA | 161 |
| KADUNA | 221 |
| KANO | 341 |
| KATSINA | 219 |
| KEBBI | 121 |
| SOKOTO | 137 |
| ZAMFARA | 120 |
| **TOTAL** |  | **1320** |
| **SOUTH EAST** | ABIA | 109 |
| ANAMBRA | 155 |
| EBONYI | 84 |
| ENUGU | 121 |
| IMO | 149 |
| **TOTAL** |  | **618** |
| **SOUTH SOUTH** | AKWA IBOM | 143 |
| BAYELSA | 63 |
| CROSS RIVER | 105 |
| DELTA | 145 |
| EDO | 120 |
| RIVERS | 189 |
| **TOTAL** |  | **765** |
| **SOUTH WEST** | EKITI | 87 |
| LAGOS | 326 |
| OGUN | 131 |
| ONDO | 123 |
| OSUN | 125 |
| OYO | 207 |
| **TOTAL** |  | **999** |
|   | **Grand Total** | **5210** |

|  |
| --- |
| **OVER SAMPLE - FIRM** |
| **REGION** | **STATE** | **Grand Total** |
| NORTH CENTRAL | NASARAWA | 503 |
| NORTH EAST | BAUCHI | 502 |
| NORTH WEST | KANO | 512 |
| SOUTH EAST | ABIA | 531 |
| SOUTH SOUTH | RIVERS | 518 |
| SOUTH WEST | OGUN | 507 |
|   | **Grand Total** | **3073** |

\*Note: Extra Interviews were done to ensure adequate representation of interviews in some EAs

#### Selection of Sampling Units/Enumeration Areas/LGA

#### The random selection of sampling units/enumeration areas/LGAs for the survey was conducted by the National Population commission. A total of 1,616 enumeration areas were selected for the household survey. These EAs fall within 1,200 locality and 576 LGAs. The firm interviews were conducted at LGA level. These were the same LGA randomly selected for the household survey.

Below is the number of randomly selected enumeration areas and LGA per state for the household and firm survey;

 **Household Survey**

|  |  |  |  |
| --- | --- | --- | --- |
| **State** | **Number of LGA** | **Number of Locality** | **Number of Enumeration Area** |
| Abia | 17 | 51 | 121 |
| Adamawa | 17 | 21 | 23 |
| Akwa Ibom | 11 | 18 | 28 |
| Anambra | 18 | 30 | 30 |
| Bauchi | 19 | 54 | 134 |
| Bayelsa | 7 | 12 | 12 |
| Benue | 15 | 17 | 31 |
| Borno | 16 | 28 | 30 |
| Cross River | 9 | 20 | 21 |
| Delta | 16 | 30 | 30 |
| Ebonyi | 8 | 16 | 16 |
| Edo | 14 | 23 | 24 |
| Ekiti | 12 | 17 | 18 |
| Enugu | 17 | 24 | 24 |
| Fct Abuja | 5 | 10 | 10 |
| Gombe | 7 | 16 | 17 |
| Imo | 20 | 27 | 28 |
| Jigawa | 17 | 20 | 31 |
| Kaduna | 22 | 42 | 44 |
| Kano | 40 | 146 | 167 |
| Katsina | 25 | 39 | 42 |
| Kebbi | 12 | 19 | 23 |
| Kogi | 14 | 23 | 24 |
| Kwara | 11 | 11 | 17 |
| Lagos | 20 | 63 | 65 |
| Nasarawa | 14 | 50 | 114 |
| Niger | 13 | 24 | 29 |
| Ogun | 20 | 79 | 127 |
| Ondo | 15 | 19 | 25 |
| Osun | 22 | 22 | 25 |
| Oyo | 26 | 40 | 40 |
| Plateau | 10 | 21 | 23 |
| Rivers | 22 | 105 | 138 |
| Sokoto | 14 | 19 | 27 |
| Taraba | 11 | 13 | 17 |
| Yobe | 10 | 12 | 17 |
| Zamfara | 10 | 19 | 24 |
| **Grand Total** | **576** | **1200** | **1616** |

 **Firm Survey**

|  |  |
| --- | --- |
| **State** | **Number Of LGA** |
| Abia | 17 |
| Adamawa | 17 |
| Akwa Ibom | 11 |
| Anambra | 18 |
| Bauchi | 19 |
| Bayelsa | 7 |
| Benue | 15 |
| Borno | 16 |
| Cross River | 9 |
| Delta | 16 |
| Ebonyi | 8 |
| Edo | 14 |
| Ekiti | 12 |
| Enugu | 17 |
| Fct Abuja | 5 |
| Gombe | 7 |
| Imo | 20 |
| Jigawa | 17 |
| Kaduna | 22 |
| Kano | 40 |
| Katsina | 25 |
| Kebbi | 12 |
| Kogi | 14 |
| Kwara | 11 |
| Lagos | 20 |
| Nasarawa | 14 |
| Niger | 13 |
| Ogun | 20 |
| Ondo | 15 |
| Osun | 22 |
| Oyo | 26 |
| Plateau | 10 |
| Rivers | 22 |
| Sokoto | 14 |
| Taraba | 11 |
| Yobe | 10 |
| Zamfara | 10 |
| **Grand Total** | **576** |

#### Interviewer Selection

**How interviewers were selected for Training and Fieldwork.**

PSI has interviewers and supervisors in all the 36 states of the federation and the Federal Capital Territory that always work on all the agency studies on a regular basis. For this survey, PSI recruited those that conducted the survey from its pool of experienced supervisors and interviewers.

Criteria for recruitment include;

* Academic qualifications-All the interviewers have minimum of Ordinary National Diploma while majority are undergraduates with over three (3) years field interviewing experience within their Region.
* Location based- All the interviewers used in each of the locations were based in the location.
* Language of Interviewers – Only interviewers that understands the local languages, orientation, traditions and culture of each location were recruited for the study
* Age/ Maturity- All the interviewers recruited in each of the study location are matured in terms of age, behavior. None of the interviewer is less than 23 years of age.

**Total Number of Interviewers and Supervisors Recruited for the training**

The total number of interviewers and supervisors that were recruited for the training is indicated below;

|  |  |  |
| --- | --- | --- |
| **Regions** | **Household Survey** | **Firm Survey**  |
| **Main Sample** | **Over Sample** | **Main Sample** | **Over Sample** |
|  | **Int** | **Sup** | **Int** | **Sup** | **Int** | **Sup** | **Int** | **Sup** |
| South West  | 52 | 12 | 21 | 5 | 51 | 12 | 21 | 6 |
| South South | 37 | 13 | 19 | 7 | 40 | 13 | 20 | 5 |
| South East  | 41 | 14 | 24 | 6 | 31 | 7 | 23 | 5 |
| North Central  | 40 | 16 | 23 | 5 | 46 | 13 | 18 | 5 |
| North West  | 73 | 26 | 21 | 8 | 62 | 14 | 17 | 4 |
| North East  | 41 | 13 | 25 | 6 | 39 | 14 | 15 | 6 |
| **Total**  | **284** | **94** | **133** | **37** | **269** | **73** | **114** | **31** |

**Total Number of Interviewers and Supervisors Selected after training:**

Final selection of interviewers and supervisors that participated in the fieldwork was done after the training and only those that were adjudged to be competent were selected

The total number of interviewers and supervisors selected is broken down below;

|  |  |  |
| --- | --- | --- |
| **Regions** | **Household Survey** | **Firm Survey** |
| **Main Sample** | **`Over Sample** | **Main Sample** | **Over Sample** |
|  | **Int** | **Sup** | **Int** | **Sup** | **Int** | **Sup** | **Int** | **Sup** |
| South west  | 36 | 9 | 12 | 2 | 38 | 5 | 11 | 2 |
| South South | 31 | 7 | 12 | 2 | 33 | 7 | 15 | 2 |
| South East  | 20 | 5 | 12 | 2 | 29 | 5 | 10 | 2 |
| North central  | 32 | 7 | 12 | 2 | 36 | 7 | 12 | 2 |
| North west  | 52 | 12 | 10 | 2 | 50 | 10 | 9 | 2 |
| North East  | 26 | 7 | 11 | 2 | 29 | 6 | 10 | 2 |
| **Total**  | **197** | **47** | **69** | **12** | **215** | **40** | **67** | **12** |

#### *Enumerators Training*

The training of the field staff (interviewers, supervisors, and quality control officers) was conducted in Lagos between 29th of June and 1st of July 2018. While other location, outside Lagos had their training from 1st of July and 3rd of July respectively.

The following salient points were fully discussed and explained during the training;

* The project objective
* Sampling methodologies to be adopted (stratified multi-stage)
* Procedure on questionnaire administration
* Interviewer’s role
* Supervisor’s role
* Quality Control Officer’s role
* The need to adopt the Group Interview technique
* Full explanations of technical terms
* No of calls allowed per Enumeration area
* The use of stimuli materials like deck to identify each rotation to be read to each of the respondent.

**Mock Sessions**

After training the field team on the research instrument on both paper and electronic device, interviewers broke into pairs and conducted mock interviews (A role –playing class sessions). The Field and Regional Managers supervised this exercise. This was done to assess the interviewers’ understanding of the project techniques and questionnaire administration.

**Actual Fieldwork Duration**

The actual fieldwork on household and firm component of this study commenced in Lagos and other study locations from 3rdof July 2018 and ended on 2nd of August 2018.

**Use of Electronic Device (Smart Phone/Tablets) for Data Collection*.* .**

In order to strengthen the quality control measures adopted for the project, electronic data collection device were used to capture data for this survey. This gave room for accurate data collection and reduces interviewer’s recording errors. It also allow for GPS capturing.

***Household Survey***

The basic methodology employed for data collection was in-home, face-to-face personal interview using a ***clustered, stratified, multi-stage random selection procedure*** to achieve a nationally representative sample.

Respondents for this study were adult Nigerian males and females aged, eighteen years and above. They were also expected to have stayed in the selected household for a period of not less than six months. The survey excluded none Nigerian citizens, people aged less than eighteen years and people living in institutionalized settings.

**Sampling Procedure**

Respondents’ selection followed a clustered, stratified, multi stage random selection process thus:

* Random selection of sampling start points
* Random selection of households
* Random selection of eligible respondents.

Within each Primary Sampling Unit or Enumeration Area, the team randomly selected sampling Start Point (SSP). This enabled the team to know where to start the random walk pattern within the (PSU). In each of the enumeration areas, Group Interviewing Technique (GIT) was adopted. By this design, a team of interviewers and one supervisor moved as a group to each selected EA, this design afforded the supervisors the opportunity to closely monitor the interviewers under them.

**Allocation of Interviewers/Walk pattern**

Each field team structure consists of four interviewers and one supervisor. Interviewers were allocated to different route walk pattern as described below:

*Interviewers 1 walked towards the sun, interviewer 2 walked away from the sun. Interviewer 3 walked at right angle to interviewer 1 and interviewer 4 walked at right angle to interviewer 2.*

**Selection of Dwelling Structure**

In each of the EAs, the “Days Code” was used to randomly determine each interviewer starting point i.e. (the 1st house or dwelling structure to enter).

A dwelling structure is defined as a floor of a distinct residential building within a sector of a town/village. Where only one household occupies a multi-storey building, the entire building (and not the floor) constituted a dwelling structure. Where it is a multi storey building with multiple occupants, counting of floors was carried out consistently from the upper floor to the ground floor in an unbroken chain from floor to floor. A fixed sampling gap of one in three (1:3) and one in five (1:5) were observed after each successful call in low and high density areas respectively.

**Selection of Household**

On entering a selected dwelling structure, the interviewer determined the number of household within the structure. Having done that, the interviewer then established the number of household with the use of the data collection electronic device / tablet to randomly select the household where the interview would take place. A household is defined as people living under the same roof and having a common arrangement for feeding. However, members of the household were also expected to have stayed together for a period of not less than 6 months. In line with this definition, household does not include servant and house helps as well as family member who currently lives elsewhere for the purpose of work or schooling and are only visiting.

**Selection of Respondent**

All the adults, males and females (depending on the sex to be interviewed) aged, 18 years and above, in the selected household were listed by name on the respondents’ selection grid table. Members of the selected household not at home at the first time of call were also listed. Then the device/Tablet was used to select one member of the household automatically at random.

**Call back/Substitution**

In a case where the randomly selected adult was not available at the time of call, interviewers were instructed to enquire about the whereabouts of the selected respondent (they may perhaps be at work) and if nearby, the interviewer walk to that place to conduct the interview. Where the selected respondent was not at home, the interviewer booked an appointment to meet them later in the day.

Two call backs were made to ensure that the selected respondent was interviewed. Where the call backs were still unsuccessful, such a case was regarded as a non-response situation or in effective call.

Also, where the selected respondent was at home but refused to be interviewed, such respondent was replaced in another dwelling structure/household, using consistent selection procedure.

Where the selected room was empty and the interviewer could not ascertain the possibility of meeting anybody there. The household was replaced with the next household. Respondent substitution in the same household was not allowed. Household substitution in the same dwelling structure was done once.

**Questionnaires Translation**

The questionnaire was translated to **Yoruba, Hausa, Igbo** and **Pidgin** as these were the local languages spoken predominantly in the various study locations. This was done to ensure proper administration of questionnaire.

***Firm Survey***

The Firm surveyed were small business that is not household enterprises and has identifiable permanent or fixed location out of home. It excluded any outlet or business enterprise that were mobile or that are run from homes. Qualified must have up to 5 employees and less than 50 employees. The universe of the firm Survey included both registered and un-registered SME accordingly.

The primary sampling unit of the firm survey was the Local Government Area.

**Sampling Approach**

The sampling procedure for the firm survey followed the stages below:

Sample per state where proportionally allocated to the selected LGA using PSI data base generated from previous study on population of firms in LGAs across Nigeria

On reaching the LGA, the field team first canvas for and identify the main commercial area where the survey was conducted.

A listing of all firms within the selected commercial area was done while establishing the number of employees. As the universe of the Firm Survey puts into consideration all firms registered and un-registered that have between 5 to 49 employees, any firm with less than 5 employees or more than 49 employees was excluded from the list.

The team then randomly selected the required allocated sample to the LGA from the list

If a selected firm happens to be a branch office, and the head office is outside the LGA, efforts and arrangement were made to conduct the interview with the head office.

#### *Enforcement of Quality Control*

PSI implemented rigorous standards during fieldwork to ensure quality data enhancement. Hence, Quality control was a high priority during the survey and numerous quality control measures were implemented at various stage of fieldwork.

To ensure that no interviewer has the ability to bias the results of the survey by producing false results, individual interviewer conducted not more than 5% of the total number of interviews. To ensure proper completion of data in the field, the team supervisors checked all devices to ensure that no interview was left uncompleted before leaving the Local Government area and the EAs .This was done each evening of fieldwork to identify and correct any potential issues as they occurred. During fieldwork, team members met every evening to discuss their experiences and any problems they faced during that particular day, and worked together to build strategies to overcome future problems.

Team supervisors were required to accompany a minimum of 25% of the interviews conducted by each interviewer, checking that the correct instructions and procedures were being followed and the interviewing was of a high standard. Interviewer accompaniment took place predominantly at the beginning of fieldwork so that problems could be identified early on, and learning could be shared with the rest of the team.

Team supervisors were also required to back-check approximately 25% of all interviews conducted by each interviewer. Back-checking includes contacting the respondent directly in person to ensure that the interview was done and checking the length of interview.

Below is the number of interviewed accompanied and back check by Supervisors.

|  |
| --- |
| **Interview Quality Control (Accompanied and Back checked) -Across the study location** |
| **Int. ID** | **# of Ints** | **% of Total** | **# Accompanied** | **% Accompanied** | **# Back-Checked** | **% Back-Checked** |
| 1201 | 104 | 0.6 | 26 | 25 | 26 | 25 |
| 1202 | 57 | 0.4 | 14 | 25 | 14 | 25 |
| 1203 | 77 | 0.5 | 19 | 25 | 19 | 25 |
| 1204 | 72 | 0.4 | 18 | 25 | 18 | 25 |
| 3201 | 16 | 0.1 | 4 | 25 | 4 | 25 |
| 3202 | 24 | 0.1 | 6 | 25 | 6 | 25 |
| 3203 | 26 | 0.2 | 7 | 25 | 7 | 25 |
| 3204 | 34 | 0.2 | 9 | 25 | 9 | 25 |
| 6201 | 62 | 0.4 | 16 | 25 | 16 | 25 |
| 6202 | 61 | 0.4 | 15 | 25 | 15 | 25 |
| 6203 | 56 | 0.3 | 14 | 25 | 14 | 25 |
| 6204 | 61 | 0.4 | 15 | 25 | 15 | 25 |
| 7201 | 56 | 0.3 | 14 | 25 | 14 | 25 |
| 7202 | 33 | 0.2 | 8 | 25 | 8 | 25 |
| 7203 | 38 | 0.2 | 10 | 25 | 10 | 25 |
| 7204 | 43 | 0.3 | 11 | 25 | 11 | 25 |
| 2201 | 39 | 0.2 | 10 | 25 | 10 | 25 |
| 2202 | 32 | 0.2 | 8 | 25 | 8 | 25 |
| 2203 | 30 | 0.2 | 8 | 25 | 8 | 25 |
| 2204 | 39 | 0.2 | 10 | 25 | 10 | 25 |
| 6401 | 82 | 0.5 | 21 | 25 | 21 | 25 |
| 6402 | 84 | 0.5 | 21 | 25 | 21 | 25 |
| 6403 | 77 | 0.5 | 19 | 25 | 19 | 25 |
| 6404 | 71 | 0.4 | 18 | 25 | 18 | 25 |
| 6405 | 85 | 0.5 | 21 | 25 | 21 | 25 |
| 6406 | 96 | 0.6 | 24 | 25 | 24 | 25 |
| 1111 | 83 | 0.5 | 21 | 25 | 21 | 25 |
| 1112 | 82 | 0.5 | 21 | 25 | 21 | 25 |
| 1113 | 88 | 0.5 | 22 | 25 | 22 | 25 |
| 1114 | 88 | 0.5 | 22 | 25 | 22 | 25 |
| 1115 | 83 | 0.5 | 21 | 25 | 21 | 25 |
| 1116 | 86 | 0.5 | 22 | 25 | 22 | 25 |
| 4201 | 55 | 0.3 | 14 | 25 | 14 | 25 |
| 4202 | 55 | 0.3 | 14 | 25 | 14 | 25 |
| 4203 | 50 | 0.3 | 13 | 25 | 13 | 25 |
| 4204 | 50 | 0.3 | 13 | 25 | 13 | 25 |
| 4205 | 20 | 0.1 | 5 | 25 | 5 | 25 |
| 4206 | 20 | 0.1 | 5 | 25 | 5 | 25 |
| 4207 | 19 | 0.1 | 5 | 25 | 5 | 25 |
| 4209 | 21 | 0.1 | 5 | 25 | 5 | 25 |
| 5201 | 63 | 0.4 | 16 | 25 | 16 | 25 |
| 5202 | 64 | 0.4 | 16 | 25 | 16 | 25 |
| 5203 | 53 | 0.3 | 13 | 25 | 13 | 25 |
| 5204 | 50 | 0.3 | 13 | 25 | 13 | 25 |
| 7501 | 57 | 0.4 | 14 | 25 | 14 | 25 |
| 7502 | 58 | 0.4 | 15 | 25 | 15 | 25 |
| 7503 | 51 | 0.3 | 13 | 25 | 13 | 25 |
| 7504 | 69 | 0.4 | 17 | 25 | 17 | 25 |
| 8403 | 92 | 0.6 | 23 | 25 | 23 | 25 |
| 8404 | 85 | 0.5 | 21 | 25 | 21 | 25 |
| 8401 | 112 | 0.7 | 28 | 25 | 28 | 25 |
| 8402 | 97 | 0.6 | 24 | 25 | 24 | 25 |
| 8405 | 68 | 0.4 | 17 | 25 | 17 | 25 |
| 8406 | 48 | 0.3 | 12 | 25 | 12 | 25 |
| 1311 | 94 | 0.6 | 24 | 25 | 24 | 25 |
| 1312 | 75 | 0.5 | 19 | 25 | 19 | 25 |
| 1313 | 96 | 0.6 | 24 | 25 | 24 | 25 |
| 1314 | 86 | 0.5 | 22 | 25 | 22 | 25 |
| 1315 | 101 | 0.6 | 25 | 25 | 25 | 25 |
| 1316 | 50 | 0.3 | 13 | 25 | 13 | 25 |
| 63001 | 27 | 0.2 | 7 | 25 | 7 | 25 |
| 63002 | 26 | 0.2 | 7 | 25 | 7 | 25 |
| 63003 | 15 | 0.1 | 4 | 25 | 4 | 25 |
| 63004 | 22 | 0.1 | 6 | 25 | 6 | 25 |
| 6301 | 52 | 0.3 | 13 | 25 | 13 | 25 |
| 6302 | 71 | 0.4 | 18 | 25 | 18 | 25 |
| 6303 | 63 | 0.4 | 16 | 25 | 16 | 25 |
| 6304 | 63 | 0.4 | 16 | 25 | 16 | 25 |
| 1401 | 77 | 0.5 | 19 | 25 | 19 | 25 |
| 1402 | 70 | 0.4 | 18 | 25 | 18 | 25 |
| 1403 | 75 | 0.5 | 19 | 25 | 19 | 25 |
| 1404 | 78 | 0.5 | 20 | 25 | 20 | 25 |
| 7301 | 47 | 0.3 | 12 | 25 | 12 | 25 |
| 7302 | 44 | 0.3 | 11 | 25 | 11 | 25 |
| 7303 | 39 | 0.2 | 10 | 25 | 10 | 25 |
| 7304 | 42 | 0.3 | 11 | 25 | 11 | 25 |
| 83001 | 45 | 0.3 | 11 | 25 | 11 | 25 |
| 83002 | 46 | 0.3 | 12 | 25 | 12 | 25 |
| 83003 | 39 | 0.2 | 10 | 25 | 10 | 25 |
| 83004 | 40 | 0.2 | 10 | 25 | 10 | 25 |
| 2401 | 44 | 0.3 | 11 | 25 | 11 | 25 |
| 2402 | 34 | 0.2 | 9 | 25 | 9 | 25 |
| 2403 | 42 | 0.3 | 11 | 25 | 11 | 25 |
| 2404 | 50 | 0.3 | 13 | 25 | 13 | 25 |
| 2301 | 75 | 0.5 | 19 | 25 | 19 | 25 |
| 2302 | 75 | 0.5 | 19 | 25 | 19 | 25 |
| 2303 | 80 | 0.5 | 20 | 25 | 20 | 25 |
| 2304 | 80 | 0.5 | 20 | 25 | 20 | 25 |
| 8201 | 56 | 0.3 | 14 | 25 | 14 | 25 |
| 8202 | 55 | 0.3 | 14 | 25 | 14 | 25 |
| 8203 | 54 | 0.3 | 14 | 25 | 14 | 25 |
| 8204 | 51 | 0.3 | 13 | 25 | 13 | 25 |
| 4111 | 53 | 0.3 | 13 | 25 | 13 | 25 |
| 4112 | 60 | 0.4 | 15 | 25 | 15 | 25 |
| 4113 | 60 | 0.4 | 15 | 25 | 15 | 25 |
| 4114 | 55 | 0.3 | 14 | 25 | 14 | 25 |
| 7401 | 101 | 0.6 | 25 | 25 | 25 | 25 |
| 7402 | 102 | 0.6 | 26 | 25 | 26 | 25 |
| 7403 | 101 | 0.6 | 25 | 25 | 25 | 25 |
| 7404 | 99 | 0.6 | 25 | 25 | 25 | 25 |
| 7405 | 96 | 0.6 | 24 | 25 | 24 | 25 |
| 9201 | 84 | 0.5 | 21 | 25 | 21 | 25 |
| 9202 | 85 | 0.5 | 21 | 25 | 21 | 25 |
| 9203 | 85 | 0.5 | 21 | 25 | 21 | 25 |
| 9204 | 85 | 0.5 | 21 | 25 | 21 | 25 |
| 9801 | 78 | 0.5 | 20 | 25 | 20 | 25 |
| 9802 | 84 | 0.5 | 21 | 25 | 21 | 25 |
| 9803 | 87 | 0.5 | 22 | 25 | 22 | 25 |
| 9804 | 80 | 0.5 | 20 | 25 | 20 | 25 |
| 1211 | 103 | 0.6 | 26 | 25 | 26 | 25 |
| 1212 | 101 | 0.6 | 25 | 25 | 25 | 25 |
| 1213 | 100 | 0.6 | 25 | 25 | 25 | 25 |
| 1214 | 98 | 0.6 | 25 | 25 | 25 | 25 |
| 1215 | 106 | 0.7 | 27 | 25 | 27 | 25 |
| 1305 | 54 | 0.3 | 14 | 25 | 14 | 25 |
| 1306 | 56 | 0.3 | 14 | 25 | 14 | 25 |
| 1307 | 55 | 0.3 | 14 | 25 | 14 | 25 |
| 1308 | 55 | 0.3 | 14 | 25 | 14 | 25 |
| 1301 | 50 | 0.3 | 13 | 25 | 13 | 25 |
| 1302 | 50 | 0.3 | 13 | 25 | 13 | 25 |
| 1303 | 50 | 0.3 | 13 | 25 | 13 | 25 |
| 1304 | 50 | 0.3 | 13 | 25 | 13 | 25 |
| 4301 | 56 | 0.3 | 14 | 25 | 14 | 25 |
| 4302 | 53 | 0.3 | 13 | 25 | 13 | 25 |
| 4303 | 60 | 0.4 | 15 | 25 | 15 | 25 |
| 4304 | 62 | 0.4 | 16 | 25 | 16 | 25 |
| 501 | 4 | 0.0 | 1 | 25 | 1 | 25 |
| 502 | 5 | 0.0 | 1 | 25 | 1 | 25 |
| 503 | 4 | 0.0 | 1 | 25 | 1 | 25 |
| 504 | 4 | 0.0 | 1 | 25 | 1 | 25 |
| 509 | 8 | 0.0 | 2 | 25 | 2 | 25 |
| 510 | 3 | 0.0 | 1 | 25 | 1 | 25 |
| 3301 | 59 | 0.4 | 15 | 25 | 15 | 25 |
| 3302 | 65 | 0.4 | 16 | 25 | 16 | 25 |
| 3303 | 54 | 0.3 | 14 | 25 | 14 | 25 |
| 3304 | 43 | 0.3 | 11 | 25 | 11 | 25 |
| 601 | 6 | 0.0 | 2 | 25 | 2 | 25 |
| 602 | 8 | 0.0 | 2 | 25 | 2 | 25 |
| 603 | 5 | 0.0 | 1 | 25 | 1 | 25 |
| 604 | 5 | 0.0 | 1 | 25 | 1 | 25 |
| 5301 | 63 | 0.4 | 16 | 25 | 16 | 25 |
| 5302 | 66 | 0.4 | 17 | 25 | 17 | 25 |
| 5303 | 61 | 0.4 | 15 | 25 | 15 | 25 |
| 5304 | 51 | 0.3 | 13 | 25 | 13 | 25 |
| 8901 | 76 | 0.5 | 19 | 25 | 19 | 25 |
| 8902 | 76 | 0.5 | 19 | 25 | 19 | 25 |
| 8903 | 78 | 0.5 | 20 | 25 | 20 | 25 |
| 8904 | 91 | 0.6 | 23 | 25 | 23 | 25 |
| 8905 | 41 | 0.3 | 10 | 25 | 10 | 25 |
| 8906 | 106 | 0.7 | 27 | 25 | 27 | 25 |
| 4401 | 83 | 0.5 | 21 | 25 | 21 | 25 |
| 4402 | 97 | 0.6 | 24 | 25 | 24 | 25 |
| 4403 | 86 | 0.5 | 22 | 25 | 22 | 25 |
| 4404 | 89 | 0.5 | 22 | 25 | 22 | 25 |
| 4405 | 76 | 0.5 | 19 | 25 | 19 | 25 |
| 4406 | 107 | 0.7 | 27 | 25 | 27 | 25 |
| 8101 | 51 | 0.3 | 13 | 25 | 13 | 25 |
| 8102 | 55 | 0.3 | 14 | 25 | 14 | 25 |
| 8103 | 55 | 0.3 | 14 | 25 | 14 | 25 |
| 8104 | 50 | 0.3 | 13 | 25 | 13 | 25 |
| 5101 | 75 | 0.5 | 19 | 25 | 19 | 25 |
| 5102 | 77 | 0.5 | 19 | 25 | 19 | 25 |
| 5103 | 82 | 0.5 | 21 | 25 | 21 | 25 |
| 5104 | 67 | 0.4 | 17 | 25 | 17 | 25 |
| 7101 | 42 | 0.3 | 11 | 25 | 11 | 25 |
| 7102 | 39 | 0.2 | 10 | 25 | 10 | 25 |
| 7103 | 40 | 0.2 | 10 | 25 | 10 | 25 |
| 7104 | 43 | 0.3 | 11 | 25 | 11 | 25 |
| 6101 | 65 | 0.4 | 16 | 25 | 16 | 25 |
| 6102 | 65 | 0.4 | 16 | 25 | 16 | 25 |
| 6103 | 56 | 0.3 | 14 | 25 | 14 | 25 |
| 6104 | 55 | 0.3 | 14 | 25 | 14 | 25 |
| 9101 | 69 | 0.4 | 17 | 25 | 17 | 25 |
| 9102 | 70 | 0.4 | 18 | 25 | 18 | 25 |
| 9103 | 72 | 0.4 | 18 | 25 | 18 | 25 |
| 9104 | 72 | 0.4 | 18 | 25 | 18 | 25 |
| 8001 | 60 | 0.4 | 15 | 25 | 15 | 25 |
| 8002 | 73 | 0.4 | 18 | 25 | 18 | 25 |
| 8003 | 82 | 0.5 | 21 | 25 | 21 | 25 |
| 8004 | 67 | 0.4 | 17 | 25 | 17 | 25 |
| 2101 | 30 | 0.2 | 8 | 25 | 8 | 25 |
| 2102 | 29 | 0.2 | 7 | 25 | 7 | 25 |
| 2103 | 30 | 0.2 | 8 | 25 | 8 | 25 |
| 2104 | 32 | 0.2 | 8 | 25 | 8 | 25 |
| 10001 | 52 | 0.3 | 13 | 25 | 13 | 25 |
| 10002 | 42 | 0.3 | 11 | 25 | 11 | 25 |
| 10003 | 58 | 0.4 | 15 | 25 | 15 | 25 |
| 10004 | 58 | 0.4 | 15 | 25 | 15 | 25 |
| 4101 | 74 | 0.5 | 19 | 25 | 19 | 25 |
| 4102 | 77 | 0.5 | 19 | 25 | 19 | 25 |
| 4103 | 75 | 0.5 | 19 | 25 | 19 | 25 |
| 4104 | 76 | 0.5 | 19 | 25 | 19 | 25 |
| 3101 | 71 | 0.4 | 18 | 25 | 18 | 25 |
| 3102 | 72 | 0.4 | 18 | 25 | 18 | 25 |
| 3103 | 44 | 0.3 | 11 | 25 | 11 | 25 |
| 3104 | 49 | 0.3 | 12 | 25 | 12 | 25 |
| 3105 | 4 | 0.0 | 1 | 25 | 1 | 25 |
| 5111 | 35 | 0.2 | 9 | 25 | 9 | 25 |
| 5112 | 41 | 0.3 | 10 | 25 | 10 | 25 |
| 5113 | 30 | 0.2 | 8 | 25 | 8 | 25 |
| 5114 | 34 | 0.2 | 9 | 25 | 9 | 25 |
| 1101 | 71 | 0.4 | 18 | 25 | 18 | 25 |
| 1102 | 59 | 0.4 | 15 | 25 | 15 | 25 |
| 1103 | 21 | 0.1 | 5 | 25 | 5 | 25 |
| 1104 | 27 | 0.2 | 7 | 25 | 7 | 25 |
| 1105 | 10 | 0.1 | 3 | 25 | 3 | 25 |
| 1106 | 55 | 0.3 | 14 | 25 | 14 | 25 |
| 9901 | 91 | 0.6 | 23 | 25 | 23 | 25 |
| 9902 | 92 | 0.6 | 23 | 25 | 23 | 25 |
| 9903 | 96 | 0.6 | 24 | 25 | 24 | 25 |
| 9904 | 90 | 0.6 | 23 | 25 | 23 | 25 |
| 9905 | 44 | 0.3 | 11 | 25 | 11 | 25 |
| 9906 | 36 | 0.2 | 9 | 25 | 9 | 25 |
| 5401 | 111 | 0.7 | 28 | 25 | 28 | 25 |
| 5402 | 100 | 0.6 | 25 | 25 | 25 | 25 |
| 5403 | 105 | 0.6 | 26 | 25 | 26 | 25 |
| 5404 | 109 | 0.7 | 27 | 25 | 27 | 25 |
| 5405 | 75 | 0.5 | 19 | 25 | 19 | 25 |
| 5406 | 58 | 0.4 | 15 | 25 | 15 | 25 |
| 6001 | 45 | 0.3 | 11 | 25 | 11 | 25 |
| 6002 | 45 | 0.3 | 11 | 25 | 11 | 25 |
| 6003 | 45 | 0.3 | 11 | 25 | 11 | 25 |
| 6004 | 45 | 0.3 | 11 | 25 | 11 | 25 |
| 2001 | 88 | 0.5 | 22 | 25 | 22 | 25 |
| 2002 | 87 | 0.5 | 22 | 25 | 22 | 25 |
| 2003 | 75 | 0.5 | 19 | 25 | 19 | 25 |
| 2004 | 74 | 0.5 | 19 | 25 | 19 | 25 |
| 1001 | 94 | 0.6 | 24 | 25 | 24 | 25 |
| 1002 | 68 | 0.4 | 17 | 25 | 17 | 25 |
| 1003 | 71 | 0.4 | 18 | 25 | 18 | 25 |
| 1004 | 93 | 0.6 | 23 | 25 | 23 | 25 |
| 7901 | 81 | 0.5 | 20 | 25 | 20 | 25 |
| 7902 | 81 | 0.5 | 20 | 25 | 20 | 25 |
| 7903 | 79 | 0.5 | 20 | 25 | 20 | 25 |
| 7904 | 84 | 0.5 | 21 | 25 | 21 | 25 |
| 7905 | 82 | 0.5 | 21 | 25 | 21 | 25 |
| 7906 | 86 | 0.5 | 22 | 25 | 22 | 25 |
| 3401 | 125 | 0.8 | 31 | 25 | 31 | 25 |
| 3402 | 87 | 0.5 | 22 | 25 | 22 | 25 |
| 3403 | 77 | 0.5 | 19 | 25 | 19 | 25 |
| 3404 | 115 | 0.7 | 29 | 25 | 29 | 25 |
| 3405 | 91 | 0.6 | 23 | 25 | 23 | 25 |
| 3406 | 19 | 0.1 | 5 | 25 | 5 | 25 |
| 5001 | 45 | 0.3 | 11 | 25 | 11 | 25 |
| 5002 | 45 | 0.3 | 11 | 25 | 11 | 25 |
| 5003 | 45 | 0.3 | 11 | 25 | 11 | 25 |
| 5004 | 46 | 0.3 | 12 | 25 | 12 | 25 |
| 5005 | 44 | 0.3 | 11 | 25 | 11 | 25 |
| 5006 | 45 | 0.3 | 11 | 25 | 11 | 25 |
| 7001 | 60 | 0.4 | 15 | 25 | 15 | 25 |
| 7002 | 60 | 0.4 | 15 | 25 | 15 | 25 |
| 7003 | 65 | 0.4 | 16 | 25 | 16 | 25 |
| 7004 | 65 | 0.4 | 16 | 25 | 16 | 25 |
| 4001 | 55 | 0.3 | 14 | 25 | 14 | 25 |
| 4002 | 61 | 0.4 | 15 | 25 | 15 | 25 |
| 4003 | 60 | 0.4 | 15 | 25 | 15 | 25 |
| 4004 | 65 | 0.4 | 16 | 25 | 16 | 25 |
| 6111 | 5 | 0.0 | 1 | 25 | 1 | 25 |
| 6112 | 5 | 0.0 | 1 | 25 | 1 | 25 |
| 3001 | 91 | 0.6 | 23 | 25 | 23 | 25 |
| 3002 | 72 | 0.4 | 18 | 25 | 18 | 25 |
| 3003 | 87 | 0.5 | 22 | 25 | 22 | 25 |
| 3004 | 84 | 0.5 | 21 | 25 | 21 | 25 |
| 3005 | 32 | 0.2 | 8 | 25 | 8 | 25 |
| 3006 | 34 | 0.2 | 9 | 25 | 9 | 25 |
|  | **16,228** |  |  |  |  |  |

# *Cases and Issues during Fieldwork*

At the end of the fieldwork a total of **Sixteen Thousand Two Hundred and Twenty-Eight(16,228) successful interviews** were achieved on **Household Survey** while a total of **Eight Thousand, Two Hundred and Eighty-Three (8,283) Successful interviews** were achieved on **Firm Survey** across the locations.

Generally, there were no major issues that affected fieldwork across the study locations.

Some LGAs and EAs where replaced because of communal clashes and inaccessibility because of raise in sea level as a result of incessant rainfall.

In Kogi State, Bassa LGA was replaced with Ogori/Mangogo because of communal clash. In Rivers state, EAs in Bonny LGA were replaced with new EAs in Port Harcourt LGA due to inaccessibility based on increased water level of the sea because of incessant rain. The new EAs are Npolu, Elekahia, Npolu –oroworukwo, Borokiri. Also three enumeration Areas at Emohua LGA in Rivers state were replaced due to communal clash. These are Egbeda EA replaced with Oduoha EA, Eganmini EA replaced with Rumuakwande EA and Ibaa EA replaced with Rumuji EA

For the firm survey, the field team had multiple visits to keep appointment as many owners and managers of the businesses have busy schedule.

APPENDIX

# *Weighting*

At the end of the data processing of results from the fieldwork, weights were applied to the survey result to reflect the true representation of the finding across demographic breaks and urbanisation. The following procedure was adopt for the weight,

**Household Survey**

There were two samples drawn for the households survey: a main sample, of 10,000 households; and an oversample of 6000 household (1000 in each of the six oversampled states).

The main sample was split up in proportion to state populations. Within each state, the sample was then split 70% urban; and 30% rural (to reflect the fact that it is much more likely to encounter households with any experience of the taxation system in urban than in rural areas). This gave the target numbers of urban and rural households for the survey in each state. The number for each was then divided by 10 – the number of households in each enumeration area (EA) - to give the number of urban and rural EAs to be selected. The National Population Comission (who hold the national sampling frame) were then asked to select this number of urban and rural EAs, with probability proportional to size (i.e. proportional to the population of the EAs). NPC selected the urban EAs from the full list of urban EAs in the state; they selected the rural EAs from the list of rural EAs in the state – again with probability proportional to size (PPS). Once the urban and rural EAs had been selected, then the survey team selected households randomly from the EA; and then selected a male or female individual randomly from each selected household.

We can write the probability of selection of an individual respondent as:

$$Prob\left(individual selected\right)=\sum\_{i=1}^{EA\_{i rural/urban} }\frac{n\_{EA\_{i}}}{pop\_{EA\_{i}}}.\frac{pop\_{EA\_{i}}}{pop\_{State Rural/Urban}}$$

where

$n\_{EA\_{i}}$= the sample taken in Enumeration Area i

$pop\_{EA\_{i}}$= the population of Enumeration Area i

$pop\_{State Rural/Urban}$= the urban or the rural population of the state

However, the population of the EAi cancels out in the equation above so that the probability becomes:

$$Prob\left(individual selected\right)=\frac{\sum\_{i rural/urban}^{}n\_{EA\_{i}}}{pop\_{State Rural/Urban}}=\frac{n\_{state rural/urban}}{pop\_{State Rural/Urban}}$$

In other words, in each state, there are only two probabilities – one for the probability of selecting an individual from a rural area, which is the total rural sample from the state divided by the rural population of the state; and one is the probability of selection an individual from an urban area, which is the total urban sample from the state divided by the urban population of the state. Because we sample 70:30 urban:rural – the latter probability will be bigger than the former (which means that the weight for urban areas will be smaller than the weight for rural areas, since the rural sample is ‘representing’ more individuals). We know the total rural and urban sample from each state; and we know the state population (split by rural/urban) and hence can calculate the weights.

**Example:**

Say that we have a sample of 200 for a state – 140 urban and 60 rural. Then the probability of selection for the individuals in urban EAs is simply 140/pop(urban) and the probability of selection for those in rural areas is simply 60/pop(rural). This means that the weights for urban individuals in this state are pop(urban)/140; and the weights for rural individuals are pop(rural)/60.

One way in which you know that you have the right probabilities and weights is that the weights in a state should sum up to the total population in each state. We can see this is true above. There are 140 urban individuals each with weight pop(urban)/140 – so the sum of these is pop(urban); and there are 60 rural individuals each with weight pop(rural)/60 – so the sum of these is pop(rural). So together, they sum to pop(urban) + pop(rural) = pop(state).

**Oversample example:**

In 6 states there was an oversample of 1000 households (700 urban; 300 rural). The correct weights to apply to these individuals depends on whether the oversample was an independent random sample, taken separately from the main sample; or whether the sample was simply added onto the main sample for these states and drawn at the same time.

It turns out that the sampling of the EAs for the oversample was done simultaneously with the sampling of EAs for the main sample. Hence, extending our example above, the probability for BOTH the main and oversample individuals (i.e. for the entire survey) in these 6 states should be:

Probability for urban individuals = (700+140 (or whatever the urban main sample was for the state))/pop(urban)

Probability for rural individuals = (300+60 (or whatever the rural main sample was for the state))/pop(rural)

(these are just example figures – the actual figures were used for each state).

**Firm Survey**

The calculation of the weights for the firm survey is complicated by the fact that, even more so than households, firms are very unevenly distributed in states. There were 5000 firms in the main firm sample and, as with the households, this sum was divided proportionately by state (using state population as the metric since we don’t know the number of small firms by state). Again, this was then split 70:30 urban:rural to give the size of the urban and rural samples for each state.

However, at this point the approach differed from the household survey. For the main sample, the household survey had already selected EAs using the method described above. Because the EA is too small a unit in which to find a significant number of firms, the Local Government Authority (LGA) for each EA selected by the household survey was taken. Then the main commercial area in each LGA was selected since this is where the firms are located. To work out the sample size for each LGA in the state, a previous survey done by PSI was used. This survey provided an estimate of the number of establishments per LGA in each state (although this survey included micro businesses and branches). The small firm sample size for urban and rural areas at the state level, was then broken down by the selected LGAs proportionately with the number of establishments found in this previous survey.

In each main commercial area of the selected LGAs, a listing was done of all firms with between 5-49 employees and with a fixed location. Then the calculated sample size for each LGA was selected randomly from the list.

For the oversampled states, far more EAs were selected and therefore, in several instances, the selected EA was in the same LGA as another selected EA. Therefore, for these states, in addition to the main commercial centre, the sampling went out to second or third commercial area in the LGA until the sample was fulfilled (again listing and selecting in each centre).

Because the sample is distributed to the LGA level with probability proportional to size, the weights are easy to calculate. Ideally they would be the number of firms selected divided by the population of small firms in the LGA. Unfortunately we do not have the population of small firms at the LGA level – however, we have a proxy from the previous survey, the number of small establishments (mostly shops). We therefore assume that the number of small firms with 5-49 employees is roughly proportional to the number of small establishments from the previous survey. Thus we calculate the probability of selection as simply the number of firms selected in each LGA divided by the number of small establishments from the previous survey. The weight wt\_firm is the reciprocal of this probability.

For the oversampled states, the main sample and the oversample were combined, so we assume that the combined sample is a random sample of the small firms in the LGA and simply add the numbers of firms in the main and oversample and divide by the number of establishments from the previous survey in the LGA as described above.