Sampling and populations surveyed

November 19, 2020

1 Sample description

The dataset contains 2690 observations and was collected during the period of 24th July till 10 September 2020. The sample is subset of a larger ongoing project in Sheikhupura and Nankana. There are two types of sample:

1. Survey Sample (sampling strategy described below)

2. Community Policing Forums: This was a randomly drawn sample from citizens living in the same neighborhoods of the survey sample but self selected themselves to attend the intervention meetings

2 Sampling for Survey based Sample

Our study site will be Sheikhupura and Nankana districts in Sheikhupura Region, Punjab. Both districts have a combined population size of 4.6 million people. These two districts consist of 27 police stations and 151 beats consisting of 1053 villages and 516 urban neighborhoods. Each Police stations in these three districts have 3-9 beats. There is great variation in the size of beats, which can consist of between 1-26 villages. Each beat has 2-3 police officers with a patrol car at their disposal, and they are the main unit of community interface and patrol in rural areas.

2.1 Unit of Analysis

We choose beat because a "beat" is the smallest administrative unit of police. Beats are assigned to Assistant Sub-Inspectors or Sub-Inspectors for patrolling, surveillance and collection of intelligence.

2.2 Treatment Assignment

Randomization will be carried out within police stations and the randomization would be a beat. Within a police station, beats are randomly assigned such that roughly 3 beats in each of the 27 stations is assigned to C-POP, 3 to CPOP-G, and 3 to control. This gives a total of 36 CPOP beats, 36 CPOP-G and 36 Control beats.

We have employed two-stage sampling method:

- S1: In each police station randomly assign one beat to each condition (CPOP-G[1], C-POP[2], Control) = 81 beats

- S2: PPS sample of 27 additional beats, then randomize into 3 conditions (CPOP-G, C-POP, Control)

2.2.1 Sampling and Subjects

We use stratified random sampling to survey 32 individuals per beat giving a total sample of 3,456 respondents in 108 beats. We use AsiaPop data which enabled probability proportional to size (PPS) sampling technique. Based on a validation exercise AsiaPop appears to provide better estimates in lower population areas compared to other gridded population data.

Based on PPS we choose 4 grid cells in each beat. In each grid we used GIS tools to identify 8 random main target survey points making a total of 32 surveys from each beat along with random replacement points. Surveyors reach these points with the help of major landmarks in the grid identified by the survey team. Each surveyor will be provided maps of the target grid area and 3 adjacent grids. The maps will be given to provide them an understanding of the exact location and boundaries of their assigned grids.

The points given to the surveyors are numbered 1 to 8. The surveyor uses point 1 to begin sample selection irrespective of whether they are residential or commercial, before moving to the next point. If a point is located in an area where no surveys can be conducted (e.g. because there is no population) the surveyor replaces this point with a first point on the replacement list and goes down the list if needed.

Respondents are identified by following the a "left hand rule" at each designated point. This means that when facing property number 1, the surveyor lists the properties to his left and numbers them in numeric order. The surveyor selects the respondents to be surveyed by going down this list. After each successful interview, the surveyor draws a sample by skipping two properties on the list. If a street ends the surveyor chooses a property that is on the left and adjacent to the previous property and continues selecting respondents from properties on the left. In case the left hand rule cannot be applied because of lack of properties or reaching the grid boundary the surveyor uses a "right hand rule" for sampling.

---

[1]Intervention for Male and Female community members

[2]Intervention for Male community members only

3   Sample for the vaccine data

From the above two sample types, we intended to survey at 10 people who self-selected into attending the forums and 10 Households who were randomly selected for our survey panel. For our survey sample, we intended to survey both the male and female members of the households. The total intended sample was 3240. However given the data was collected through phone based survey, we experienced a significant drop in response rates and were only able to collect data from 2690. The distribution of the final sample reached is as follows: 1056 (835 Males 217 Females[3] individuals from community policing forums, 817 HH (817 Males 816 Females) from survey sample

---

[3]Females were required to participate only in CPOP-G beats (36 out of 72 Treatment beats)