CHAPTER 3

# RESEARCH METHODOLOGY, SAMPLE DESIGN & SAMPLE COVERAGE

## 3.1 RESEARCH METHODOLOGY

This study utilizes qualitative and quantitative methodologies for estimating the economic and social impact of road crash outcomes. Using purposive sampling, a multi-pronged approach was adopted to assess the impact of road traffic injuries and fatalities on victims.

Based on the objectives of the study, the research methodology is divided into two phases:

### PHASE-I: EXPLORATORY RESEARCH

#### A) DESK RESEARCH

Desk research was undertaken to review the various provisions in MVAA 2019 regarding road safety provisions and compensation to road crash victims. During this stage, various data sets were explored with an effort to create a database that would be most suitable to generate contact details of road crash victims – both from urban and rural areas.

#### B) PREPARATION OF SURVEY INSTRUMENTS

Draft survey instruments were prepared for each respondent category separately. Survey instruments (questionnaires and guidelines for FGDs and IDIs) were created and later fine-tuned post the feedback from the pilot.

#### C) PILOT SURVEY

Due to the National Lockdown mandated by the Central

Government due to COVID-19, the pilot survey was conducted telephonically. The sampling design and survey instruments were tested with a small sample of 30 interviews. The pilot survey was carried out from May 7, 2020 -May 13, 2020.

#### D) POST-PILOT FINE-TUNING OF SURVEY INSTRUMENTS

As part of the revised sampling strategy, more detailed questionnaires were retained for face to face interviews and a slightly shorter version of the questionnaire was used for telephonic surveys. These survey instruments were translated into regional languages to enhance their comprehensibility.

### PHASE-II: DESCRIPTIVE RESEARCH

The field survey was conducted between 19th June, 2020 – July 23, 2020. The following stages were covered in this phase of the study:

#### A) QUANTITATIVE RESEARCH:

The survey was conducted among respondent categories using structured questionnaires administered by experienced investigators. The respondents were approached via two sampling strategies:

Additionally, awareness levels about motor insurance, compensation and other provisions of the Motor Vehicles (Amendment) Act, 2019 (MVAA, 2019) was also tested among truck drivers.

#### B) QUALITATIVE RESEARCH - FGDs AND IDIs

Qualitative surveys were used to understand the perspective of the victim/victim's family member regarding the impact of fatality/crash on their social condition as well as their emotional health. In-depth interviews were conducted among adolescents (aged 14-18) who have been survivors of a road crash or have lost a family member in a road crash. These interviews were conducted through a mix of video calls and face-to-face interviews.

Similarly, FGDs were conducted among women and men with participants who had either been involved in a road crash or their immediate family member was in a road crash. These FGDs were conducted in Patna (Bihar) and Lucknow (UP) by an experienced researcher and moderator with adequate precautions and adherence to social distancing protocols.

## 3.2 SAMPLING DESIGN

A multi-stage sampling procedure was adopted to select the target respondents for this study. The first unit of sampling was the State. Looking at the number of crashes during the last 15 years and the geographical location of states, the top 4 states were selected for conducting this survey, i.e. Uttar Pradesh in North India, Maharashtra in West India, Tamil Nadu in South India and Bihar in East India. Additionally, Maharashtra and Tamil Nadu represent High Capacity States (HCS) demonstrating a higher economic growth and better performance in Human Development indicators while Uttar Pradesh and Bihar represent Low Capacity States (LCS) demonstrating a sluggish economy, higher poverty rates and low levels of social and administrative progress.

# RESEARCH METHODOLOGY, SAMPLE DESIGN & SAMPLE COVERAGE

In the next stage of sampling, in each of the selected states, 4 districts were selected to find out relevant respondents. These districts were also selected based on the number of crashes and geographical location (diversity) in the state.

The quantitative survey aimed at providing state-level point estimates to key aspects of economic and financial impacts on the families of road crash victims. Hence, at a 95% confidence interval and a 5% margin of error in the sample size was calculated using the following formula -

$$SS = \frac{Z^2 * p*(1-p)}{C^2}$$

where:
Z = Z value (e.g. 1.96 at 95% confidence level)
p = percentage picking a choice expressed as a decimal (since we do not have estimates of the p-value, we have taken it as 0.5 to maximize the sample to avoid any under-sampling)
C = confidence interval, expressed as decimal (e.g., .05 = ±5)

Using the above formula, the population, the number of road crash in each state during 2005-18 and other aspects of the target group, the sample size was calculated for each state.

During the field survey, purposive and snowball sampling was used to select an adequate number of respondents from each state. After identification of respondents, a mix of telephonic and face to face surveys was conducted among road crash victims or their family members.

## 3.3 SAMPLE COVERAGE

A total number of 2,499 interviews (against the target sample of 2,400) were conducted across 4 states. Category wise, 1647 LIH, 432 HIH and 420 truck drivers were interviewed across 4 states[7]. In addition to the above, 3 FGDs and 8 IDIs were conducted among road crash victims/their family members. The summary table 3A.1 in the annex provides a snapshot of descriptive statistics about the participants. Statistically, a 95% confidence interval and a 5% margin of error in a sample size of 384 is sufficient to conduct multi variate analysis of the data. Thus, statistically each category (LIH, HIH and truck drivers) has an adequate sample size and the comparison between LIH and HIH is justified and stands valid for deriving conclusions. Further, for comparison purposes, the proportion figures have been quoted instead of absolute numbers.

## 3.4 PROFILE OF LOW INCOME AND HIGH-INCOME HOUSEHOLDS

The study was conducted among low -income and high-income households in 4 selected states across urban and

7. Out of total, 986 (47%) interviews among victims/ family members were conducted through face to face mode while remaining 1093 (53%) interviews were conducted telephonically. All interviews among truck drivers were conducted through face to face mode.

rural areas in 4 zones (East – Bihar, West – Maharashtra, North – Uttar Pradesh, South – Tamil Nadu).

As per the latest estimates (WB, 2015), the world's 73.6 crore people are living in extreme poverty (less than US$1.90 a day) with a poverty rate of 10%.

In the case of India (Census of India, 2011) about 21.9 per cent population of India is extremely poor and lives under the poverty line (as per Tendulkar committee). The poverty rate in rural areas is higher (25.7%) than the urban population (13.7%).

Further, 41% of road crashes were recorded in urban areas and 59% in rural areas. The proportion of persons killed in urban areas and rural areas was 34% and 66% respectively

For this study, LIH and HIH were defined based on income and ownership. The upper threshold pre-tax income of a LIH sample was taken as INR 13,450 per month per adult[8]. Similarly, for HIH category, the lower threshold pre-tax income of INR 50,000/- per month per adult was considered.

The Low Income Households' test sample, comprising the bottom 40% of the population was selected keeping in mind standard models of population representation used globally by the United Nations (UN) and World Bank in its poverty related research. UN SDG target 10.1 aims for the income of the bottom 40 percent to be growing faster than the national average by 2030. Progress is measured by the difference between growth in the consumption or income of the bottom 40 percent and growth in the consumption or income of the mean of the population as a whole.

After carefully analyzing the World Inequality Database (that aims to provide open and convenient access to the most extensive available database on the historical evolution of the world distribution of income and wealth, both within countries and between countries), the control sample of HIH was defined as comprising of the top 10% of the population in terms of income. Across the world, the income earned by the top 10 percent is often larger, sometimes much larger, than the share earned by the bottom 40. The ratio of the share of income between the top 10 percent and bottom 40 percent is known as the Palma ratio.

To monitor progress against its goal of boosting shared prosperity, the World Bank tracks growth in the consumption or income of the poorest 40 percent of the population in each country—the bottom 40 percent. Shared prosperity focuses on the poorest 40 percent of the population in each economy (the bottom 40) and is defined as the annualized growth rate of their mean household per capita income or their consumption.

This 40% and 10% population ratio is thus an important and useful development indicator for the World Bank to measure socio-economic impact in any given country, especially developing countries. It helps to juxtapose

8. https://wid.world/country/india/

# RESEARCH METHODOLOGY, SAMPLE DESIGN & SAMPLE COVERAGE

various socio-economic realities and highlights the widening gap amongst sections of the population more clearly. It also helps to make the data in this report more globally comparable with other nations.

On comparing the income profile of LIH and HIH categories at a global level while keeping all factors the same (World Inequality Database); LIH population of India that earn a pre-tax income of up to INR 13,500 per month per adult would be among the bottom 19% population globally (refer to Table 3A.2). Similarly, HIH population of India that earn pre-tax Income of INR 50,000 or above would be among the top 42% globally[9].

71% of the LIH respondents were from rural areas, while for HIH, it was opposite as major proportion (82%) were from urban areas (Refer to Table 3A.4). This is very close to the actual distribution of poor population in rural (about 20%) and urban (about 80%) areas in India (as per Census 2011).

Based on income and self-declaration, every 2nd LIH surveyed was found below the poverty line. In each LIH, there were average 5.6 (median value = 5) members while the average size of HIH was 4.8 (median value = 4). Cognizance was taken of the fact that respondents might not disclose their real income during the surveys. To overcome this limitation, the surveyors visited the locality for more than half of the respondents surveyed and verified their range of income through these field visits. In case of telephonic surveys (53% of the total sample), the tallying was done with a verifiable database. Additionally, respondents were asked about their expenses at various levels of the survey and their responses gave a fair indication of their household income. The framing of short pointed questions on whether the household had to borrow money/ sell or mortgage valuables and whether anyone in the family had to quit study/relocate, etc. helped in verifying their income-brackets.

A pre-testing of the methodology was conducted to finetune the approach ensure rigrous data collection.

Among the LIH crash victims, 86.2% were male while 13.8% were female. Among HIH crash victims, such a proportion of males and females was 78.7% and 21.3% respectively.

Similarly, age-group wise, about every 2nd road crash victim from LIH was in the 26-45 yrs age bracket, whereas among HIH such proportion was about 64%. Across households, more than half of the all victims (54% for LIH and 64% for HIH belonged to the productive age group of 26-45 years. A probable reason is that most commute/ travel is work related and hence the working age population is more prone to road crashes.

In terms of educational qualification, with ample facilities and resources for the HIH category, they were found to be more educated than the LIH victims. About two-third of HIH victims had at least completed graduation. While among LIH it was opposite, as about two-thirds of victims were educated only up to 12th standard and only one-fourth were either graduate or above. The link between educational qualifications and recovery to work has not been captured/ established directly in this study. Though it can be said that

9. For more details see Annexure.

over 40% of the LIH victims were either unemployed or engaged in the agriculture or informal sector as unskilled labourers before the crash while 83% of the HIH victims were either self-employed or doing business or working in the formal sector as salaried employees. A combination of factors like higher educational levels, family savings to dip into and better social status did facilitate their early transition back to the workforce and previous learning levels. However, disability adds another dimension/layer of analysis that delays the process of resuming work or finding meaningful and well-paid work across households after the crash.

Most of the victims in LIH were engaged as unskilled labour/ farmers in the agriculture sector or shop owner/ petty traders and about 4 out of 10 were engaged in businesses/ self-employed or working as salaried employees. On the other hand, over 8 out of 10 victims in HIH were engaged in businesses/ self-employed or working as salaried employees.