Since early April, Dalberg has been leading a rapid, multi-round study on the efficacy of government entitlements in helping below-poverty line families navigate the financial impacts of COVID-19. The study is among the largest such efforts underway in India, covering 47,000 low-income households across 15 states, with support from Omidyar Network India and Rohini Nilekani Philanthropies. Our motivation is to provide real-time, actionable intelligence to help policymakers make informed decisions during the crisis.

Objective: The objective of our study is to inform policymakers of dynamics on the ground among low-income populations across India. We wanted to optimize for broad coverage across the country, as well as to sample with sufficient depth to enable meaningful discussion at a state level in the 15 states covered by the survey.

Sample inputs: To fulfil this objective, we sought a sampling methodology that would enable us to capture perspectives of low-income populations during the height of the initial COVID-19 crisis. We created a sampling frame using a database developed prior to the COVID-19 crisis by Kantar Public, a division of the social research and consulting firm Kantar. Their underlying database includes 500,000 phone numbers collected over the past five years.

1. **This database was constructed from two key sources:**
    a. **Kantar-led syndication** of phone numbers (95%), wherein Kantar enumerators invested time to source information from potential interviewees for future studies.
    b. Phone numbers sourced from **their own prior studies** (5%) run by Kantar group companies (not those paid for by external clients).

    The footprint of this database spans across India, and demographic characteristics such as geography and gender broadly align with the 2011 Census distribution, except in the cases of some of the smallest states/UT. Kantar has the express consent of all individuals in the database to be used for surveys like our own.

2. **We drew from this database for our study on COVID-19 entitlements, beginning by filtering** the 500,000 phone numbers for two characteristics:
    a. Respondents who self-reported as **low income**, defined as below-poverty line (BPL).[1]
    b. Respondents with phone numbers collected or updated within the **past three years**.

3. **Following this, we drew a sample of phone numbers on a randomized basis.** Our randomization process was conducted within the framework of quotas aligned with population characteristics we cared about, given the objectives of our study. These included the following:
    a. **Rural/Urban:** We sought a rural urban split of 70% rural, 30% urban in order to reflect India's population distribution (69% rural : 31% urban according to 2017-18 PLFS sample).
    b. **Gender:** Most of our data is collected on a household basis rather than individual level. However, we specified a minimum of 30% women respondents in order to ensure coverage of women's voices, albeit lower (32%) than is observed in the population. It was not feasible to achieve 50:50 representation of respondents due to common hesitancy of women respondents to participate when male heads of household were present in the house - a common occurrence

[1] We included all families who self-reported as BPL, because we wanted to also include families who fall below the poverty line but do not have BPL/AAY cards. We did this in order to capture the experiences of families who need support, but do not meet official eligibility criteria for various government schemes. After interviews were conducted, we validated self-reported BPL status by applying two filters: (i) possession of a BPL or AAY ration card or registration on BPL/AAY list; OR (ii) self-reported monthly household income of INR 30,000 or lower. Families that fulfilled either of these criteria were retained in the sample. Not all the retained households have incomes that fall below official poverty line thresholds (in our sample, 20% households report monthly income below INR 5,000, 71% below INR 10,000, and 95% below INR 20,000). Therefore, we refer to our sample as low-income and not BPL

during the lockdown. We are conducting a separate gender focused survey to analyze the gendered impacts of the COVID-19 crisis.

c. **Broad sampling across states:** Although we did not aim to be nationally representative, we sought to cover the vast majority low-income populations across the country. As such, we selected 10 states covered in our April study to cover 60% of India's overall population, and 15 states in our May survey to cover 80%.

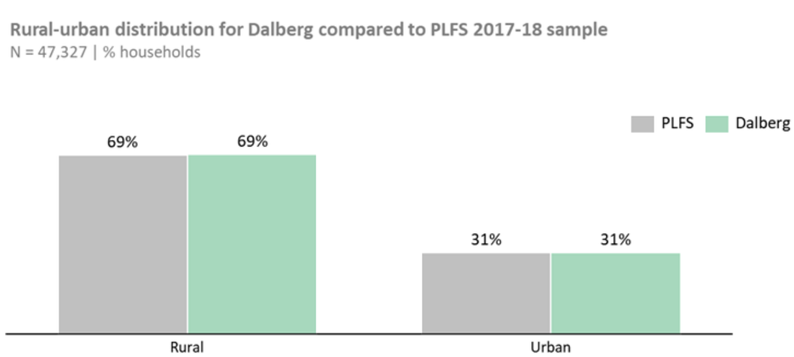
d. **Geographic distribution across districts within each state:** We targeted coverage across most or all districts in each state we covered. In 5 of our 15 states, we achieved 95%+ coverage; we achieved 80% district coverage overall (408 out of 498 total districts in the 15 states) given samples in some districts were reduced in the filtering stage when we removed numbers above our income threshold and numbers more than 3 years old.

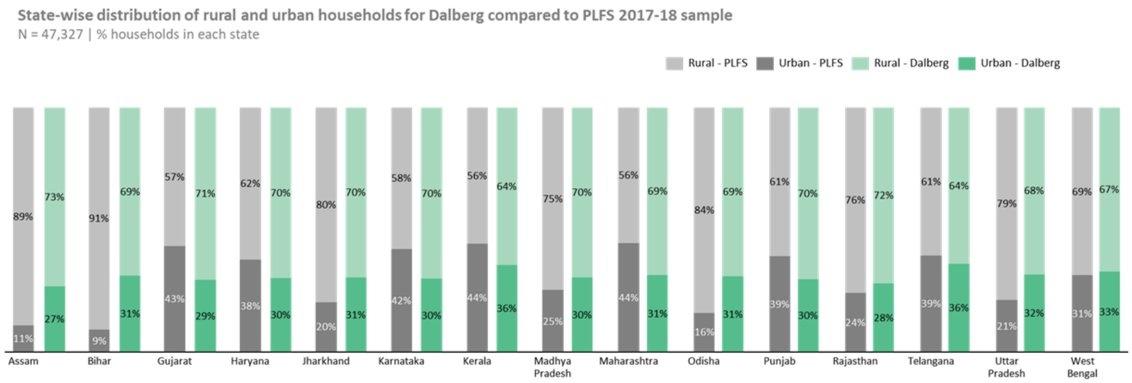
Ultimately, Kantar's database provided 80% of our final sample. To ensure coverage of our quotas above, we relied on snowballing (6%) and random digit dialing (14%) to fill the remainder of the sample.

4. **Finally, we applied population weights** when conducting analyses on the total sample across all 15 states, to appropriately reflect the contribution of each state across the pool of 15 states surveyed. As such, our findings are not nationally representative, but rather representative across the 15 states surveyed. The weights we used are based on the number of households registered for PHH / AAY ration cards in the NFSA database, since our study focuses on the BPL population. As the primary objective of this exercise was to gather reliable findings at an aggregate level for the 15 states that we studied, we have only applied state-level weights from NFSA data to estimate aggregate numbers. Within each state, we rely upon quota setting, as described above, to enable proper sample distribution. Small deviation of these quotas from the population can be expected to have negligible impact on our findings.

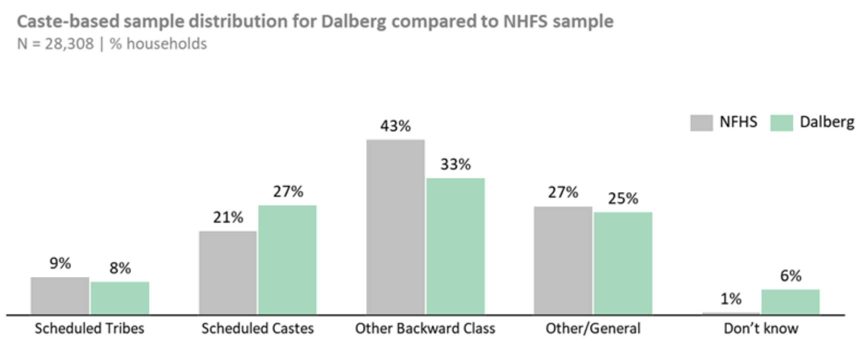
Outputs: We feel confident about the results of our study. They broadly represent the key characteristics of the populations our study focused on in the real world. To ensure this and to identify any in-built biases, we triangulated our sample demographics against existing public data sources that are validated as nationally representative, namely the Periodic Labour Force Survey (PLFS) 2017-18, the National Family Health Survey (NFHS) 2015-16, both government-led, large-scale, multi-round surveys conducted in a representative sample of households throughout India; and the India Human Development Survey (IHDS) 2011-12 and the Centre for Monitoring Indian Economy survey (CMIE) 2019, both nationally representative, multi-topic panel survey of 41K+ households and 174K+ households respectively.

- **Rural - Urban distribution**: Given the breadth of Kantar's database, we successfully controlled rural urban distribution to match the PLFS 2017-18 rural:urban ratio at a national level. Some variance at the state level was observed.





- **Social category distribution**: Our representation of general social categories[2] is roughly the same as those measured in the NFHS database. Our representation of the Other/General category is aligned with NFHS. Within more marginalized non-Other/General categories, we see some variation, but in sum are roughly aligned (totalling 73% within NFHS vs. 68% within the Dalberg sample); the difference is attributable to those responding "I don't know" in the Dalberg sample.



- **Occupation breakdown**: Comparing occupational representation across datasets is fraught with challenges given differing categorization processes, as well as ongoing changes in the occupational makeup of the wider population (especially during times of economic upheaval such as the period of our study). However, to shine a light on any foundational differences, we compared our raw sample against the PLFS 2017-18 database after aligning the occupation categories in the two datasets.[3] We observe a bias toward construction and casual labour in our sample[4]; we expect this may be a result of exclusively targeting low-income households. However, agricultural households appear to be underrepresented in our sample compared to PLFS 2017-18.

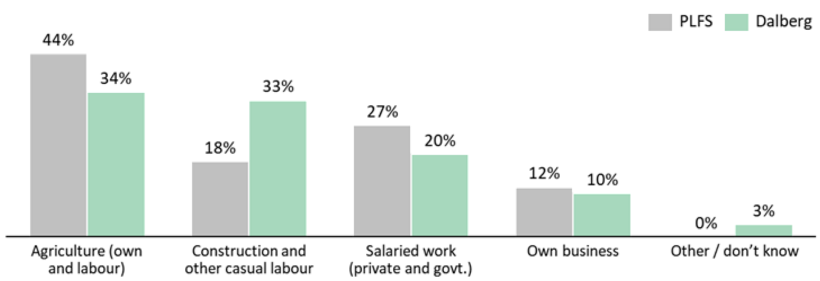
---

[2] Note, this question was only posed in the May round of our study, thus has a lower N of 28k.

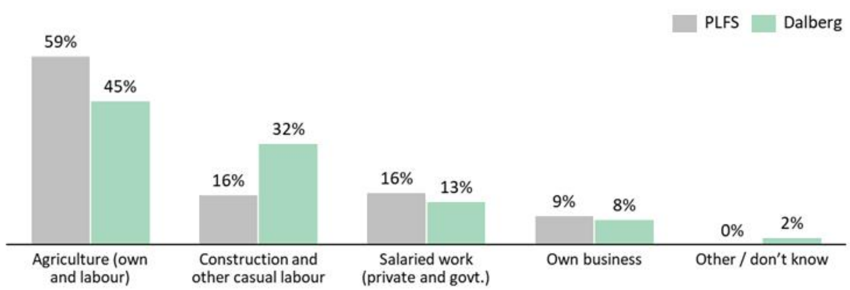
[3] PLFS 2017-18 uses ten occupation divisions (1, 2, 3, 4, 5, 6, 7, 8, 9, x) from the National Classification of Occupations (2004). We have mapped this to our occupation labels in the following manner: Divisions 1-5, or legislators, senior officials, managers, professionals, technicians and associate professionals, clerks, and service / sales workers in the organised sector have been mapped to our 'Salaried' category (both private and govt.); Division 6 and Division 92, or market oriented and subsistence agriculture work, as well as agriculture labour, have been mapped against our clubbed categories of 'Own agriculture' and 'Agricultural labour', since the distinction of land ownership is unclear in the NCO classification; Division 7, or craft and related trade workers (e.g. mechanics, electricians, tailors) have been mapped to our 'Own shop / business' category; Division 8, 91, and 93, or plant and machine operators and assemblers, as well as elementary sales and service occupations (e.g. hawker, milkman) and labourers in mining, construction, manufacturing, and transport, have been mapped to the clubbed category 'Construction work' and 'Other Casual labour', since Division 93 does not distinguish between these two; Division x, or new workers seeking employment, workers not reporting their occupation or reporting unidentifiable / inadequately described occupations, have been mapped to our 'Others' and 'Don't know' categories.

[4] Note: We excluded earners who are students, retired, or unemployed, thus have a lower total sample size for this analysis, at 40k.

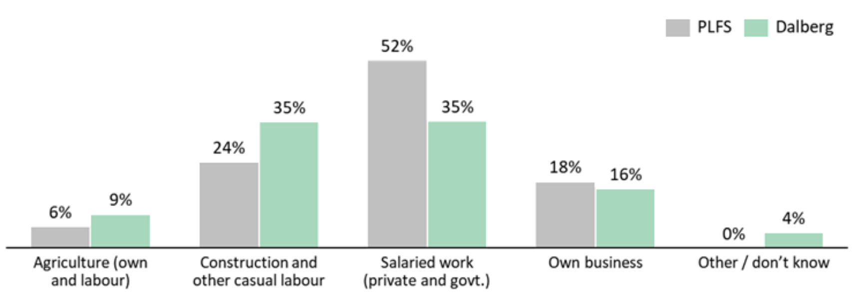
Overall occupation distribution for Dalberg compared to PLFS 2017-18 sample
N = 40,830 | % primary income earners



Rural occupation distribution for Dalberg compared to PLFS 2017-18 sample
N = 28,665 | % rural primary income earners



Urban occupation distribution for Dalberg compared to PLFS 2017-18 sample
N = 12,165 | % urban primary income earners

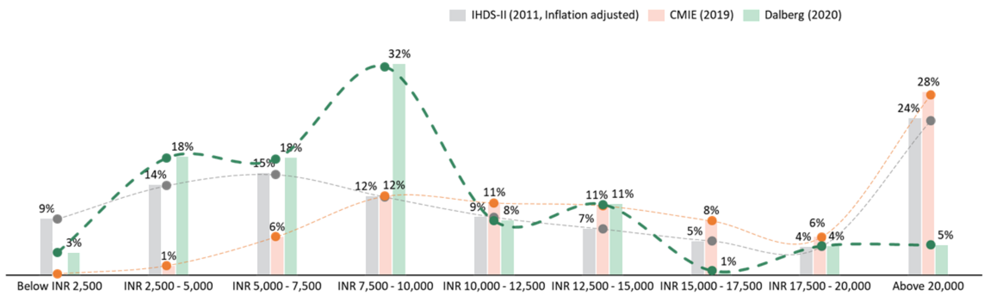


- **Household income breakdown.** Some difference in income breakdown is to be expected between national face-to-face surveys and this study, given that our survey was conducted by telephone and thus excludes those without access to a working telephone - likely the poorest of the poor (see Limitations below). Further, many surveys, such as NSSO, collect household financial information on the basis of expenditure rather than income due to common reluctance from respondents to share accurate income information. Noting these inherent limitations, we have compared the makeup of our sample against other surveys:
  - Overall, 21% of the households in our sample report monthly income of INR 5,000 or less, 71% report INR 10,000 or less, and 95% report INR 20,000 or less.[5]

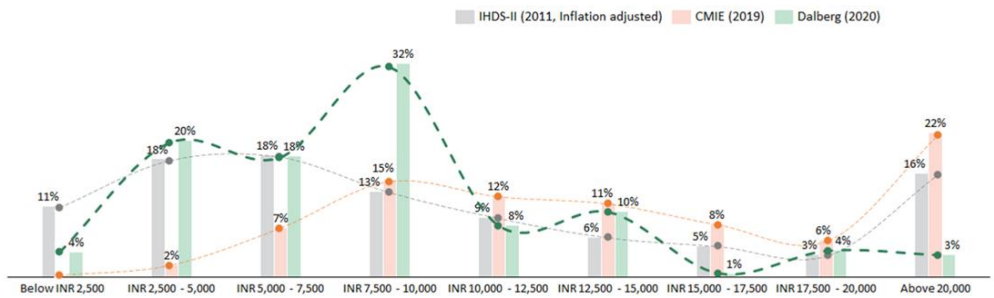
[5] If we filter out those earning > 30k/m, ~3% of respondents in CMIE and 25% in IHDS report monthly income of INR 5,000 or less, 28% and 56% respectively report INR 10,000 or less, and 80% and 87% respectively report INR 20,000 or less.

- Household distribution in our sample represents a significantly higher concentration of households in the <INR 10,000 income range compared to both IHDS (2011, adjusted for inflation[6]) and CMIE (2019), both of which are nationally representative and contain all income segments. This holds for both rural and urban households. This is as expected, given that our survey specifically targets low income populations.

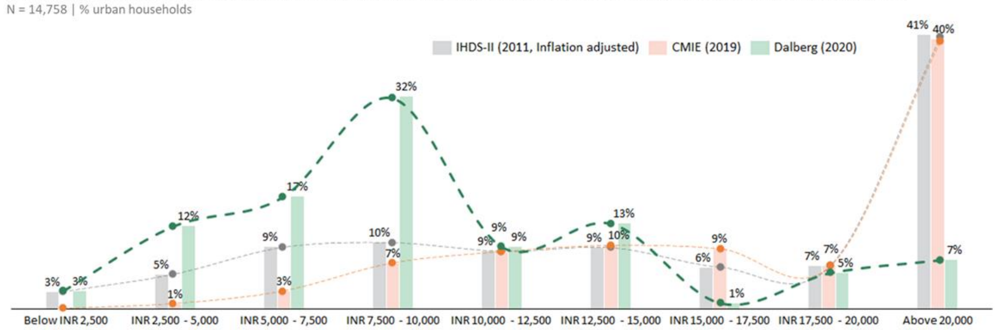
Overall household income distribution in various monthly income brackets of Dalberg sample compared to IHDS 2011-12 and CMIE 2019
N = 47,327 | % households



Rural household income distribution in various monthly income brackets using Dalberg data compared to IHDS 2011-12 and CMIE 2019
N = 32,569 | % rural households



Urban household income distribution in various monthly income brackets using Dalberg data compared to IHDS 2011-12 and CMIE 2019
N = 14,758 | % urban households



Additional limitations: Overall we feel comfortable with our methodology as it aligns with our objective to capture the real time perspectives of low-income populations in India during the initial months of the COVID-19 crisis. However, as with any survey, there are inherent biases in our data, further to those described above, which we feel it important to call out:

1. **Phone survey biases:** Telephonic surveys in general, not just our own, have inherent limitations:

[6] World Bank Open Data. The cumulative inflation factor of 1.8 times was estimated by compounding year-on-year values of annual % inflation (consumer prices) between 2011 and 2019

a. They exclude a number of populations, such as those who don't own a phone, those not able to charge their phone (such as those on the move), those who lack the money to top up their phone credit, and those without network coverage. We found 20% of potential respondents were unreachable, likely due to reasons such as these. As such, any telephonic survey is unable to be fully representative of the country as a whole. Those excluded are likely worse off than those we were able to reach by phone - as such, our findings may be expected to represent a best case scenario.
b. Few phone surveys are as representative as stratified random sampling surveys because the underlying datasets most phone numbers are drawn from are not randomized. An exception to this is random digit dialing (RDD); we chose not to anchor on RDD techniques (other than to fill gaps in our quotas) in order to ensure sufficient representation of the characteristics of interest for our study, such as geographic spread, gender distribution, and urban/rural split.

Beyond these biases, maintaining attention on telephonic interviews is difficult. Therefore, we tried to keep our interview length under 20 minutes. This limited the number of themes we could cover and the number of questions we could ask.

2. **Inbuilt biases in interview-based surveys**:
    a. Surveys may risk self-selection bias due to non-response. Our overall non-response rate was 29%, including those who were unreachable or hung up before the enumerator got a chance to introduce the purpose of the call. Our true non-response rate, or those respondents who declined to take the survey once briefed about its purpose or taken through the consent script, was 9%. This is lower than the typical true non-response rates of ~15-25%, per Kantar's experience with other studies. We believe our lower rate is largely explained by the short length of the survey (<20 minutes), which increased respondent willingness, and the personal relevance of entitlement access, which increased respondent enthusiasm.
    b. Interview-based surveys rely on self-reporting and are limited by what respondents are willing and able to answer. Further, our questions focus on the individual's or family's experience, rather than a systemic view.
    c. Surveys and in-depth interviews are limited by the framing, sequence, translation, communication, and interpretation of questions. We have tried to minimise biases arising from these limitations through a simple questionnaire design and enumerator training, but some will inevitably remain.

3. **Limitations per our focus areas**:
    a. By design, our study covers BPL households only (we allow self-reporting of BPL status since poor households without BPL cards would otherwise be excluded from our study). Many APL households have also experienced extreme economic shocks and may also need financial relief or other assistance - they are not captured by this survey.
    b. Further, our study does not cover all types of government support. We study select entitlements that are offering top-ups or advance payments during the Covid-19 crisis, including: PDS rations, PM-Kisan, PM Ujjwala Yojana, MGNREGS, Jan Dhan Yojana transfers to women-owned accounts, and social pensions. Many individual states have developed their own state-level disbursal mechanisms - we were unable to capture these in detail. We focus on the national schemes above because they have a very broad reach, can be compared across the country, and because both cash and in-kind transfers are represented.

In our view, these limitations are acceptable in light of our objectives to rapidly capture perspectives on the ground of low-income people across the country as the COVID-19 crisis unfolded. We have several follow up surveys on the impacts of the crisis in planning, and welcome questions and ideas on our survey methodology at ImpactsofCovid@dalberg.com.