



HAL
open science

NEEMSI-2 Survey Manual

Christophe Jalil Nordman, Govindan Venkatasubramanian, Isabelle Guérin,
Sébastien Michiels, Cécile Mouchel, Arnaud Natal, Mary Di Santolo

► **To cite this version:**

Christophe Jalil Nordman, Govindan Venkatasubramanian, Isabelle Guérin, Sébastien Michiels, Cécile Mouchel, et al.. NEEMSI-2 Survey Manual. French Institute of Pondicherry; IRD Institut de recherche pour le developpement. 2021, pp.55. hal-04207258

HAL Id: hal-04207258

<https://hal.science/hal-04207258>

Submitted on 14 Sep 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License



NEEMISIS

NETWORK, EMPLOYMENT, DEBT, MOBILITIES AND
SKILLS IN INDIA SURVEY

<https://neemsis.hypotheses.org>

SECOND WAVE

Survey Manual

2020-21





NEEMISIS-2 core team:

Dr Christophe Jalil Nordman (IRD, LEDa-DIAL and IFP), Dr Isabelle Guérin (IRD, CESSMA and IFP), Dr Govindan Venkatasubramanian (IFP), Dr Sébastien Michiels (IFP), Cécile Mouchel (LEDa-DIAL and CESSMA), Arnaud Natal (Univ. Bordeaux and IFP), Mary Di Santolo (LEDa-DIAL).

Survey coordinator: Dr Christophe Jalil Nordman (jalil.nordman@ird.fr)

This data collection was carried out within the
Observatory of Rural Dynamics and Inequalities in South India.



<https://odriis.hypotheses.org/>

Contents

1	Introduction of the Network, Employment, dEbt, Mobilities and Skills in India Survey 2 (NEEMSI-2)	7
2	Survey methodology	9
2.1	Coverage and sampling	9
2.2	Breadth of topics and in-depth contextual measures	11
3	Files description	13
3.1	Content	13
3.2	Merging files	14
3.3	Data format	14
4	File content and variables	15
4.1	Household	15
4.2	Occupations	24
4.3	Migrations	26
4.4	Loans	28
4.5	Ego	29
4.6	Alter	38
4.7	Household and Ego tracking	40
4.8	Tracking occupations	43
4.9	Alter tracking	44
4.10	Migration path	46
5	Fieldwork procedure	47
5.1	Tasks for enumerators	47
5.2	Tasks for the team leader	52

1 | Introduction of the Network, Employment, dEbt, Mobilities and Skills in India Survey 2 (NEEMISIS-2)

THE *Networks, Employment, dEbt, Mobilities and Skills in India Survey 2* (NEEMISIS-2), conducted in 2020-21 from December 2020 to October 2021, is the third wave of a multi-topic household and individual survey covering 3647 individuals, 632 households and 10 villages in-between Cuddalore and Kallakurichi districts in rural Tamil Nadu (South India). Most of these households (485) had been interviewed for the NEEMISIS-1 wave, conducted in 2016-17.

This NEEMISIS constitutes an extension and follow-up of an original baseline household survey that was implemented among the same 10 villages in 2010 (gathering 388 households) in the framework of the *Rural Microfinance and Employment* (RUME) survey, a project led by Isabelle Guérin (IRD, CESSMA) and Marc Roesch (CIRAD). The three waves (RUME 2010, NEEMISIS-1 2016-17, and NEEMISIS-2 2020-21) thereby constitute a three-year panel of households and individuals.

The NEEMISIS is organised within the French Institute of Pondicherry (IFP) and is coordinated from 2016 by Christophe Jalil Nordman (IRD, LEDa-DIAL) and Venkatasubramanian G. (IFP). The survey is part of a broader programme including quantitative and qualitative data collection tools, that form an *Observatory of Rural Dynamics and Inequalities in South India* (ODRIIS), led by Isabelle Guérin (IRD), Christophe Jalil Nordman (IRD) and Venkatasubramanian G. (IFP).

One of NEEMISIS-2's specificity is its timing: an external shock, the COVID-19 lockdown which directly affected livelihoods of surveyed households, took place a few months before the beginning of the survey.

On the evening of March 24, 2020, the Government of India ordered a nationwide lockdown for 21 days, starting on March 25, strictly limiting the movement of the entire population of India as a preventive measure against the COVID-19 international pandemic. However, India and especially Tamil Nadu experienced the peak in disease in April 2021. Therefore, it pushed the Tamil Nadu government to order a second lockdown to counter the second wave, a wave that was more intense in terms of infections and death. By contract, this second lockdown was less strict than the first in terms of movement constraints. The second lockdown started on April 5

and ended on June 15, during the data collection that had started in November 2020. Almost 60% of the households were interviewed before the second lockdown, 20% during and 20% after. NEEMSI-2 took advantage of this context to try and address its effects on rural households.

Three-hour interviews in each household covered employment, migration, education, debt, financial practices, agriculture, marriage and schemes. Two-hour individual interviews in each household covered labour force participation, social networks, cognitive and non-cognitive questions.

NEEMSI-2 was conducted by a team of LEDa-DIAL and CESSMA research units in Paris, both under the IRD support, a team coordinated by Christophe Jalil Nordman (IRD, LEDa-DIAL). This second NEEMSI wave was funded by the CHALLINEQ project (ANR-18-EQUI-0003), a 36 month Indo-European project (2018-2021) led by the Centre of Human Sciences (CSH) in Delhi, India. The survey was implemented within the IFP in close cooperation with the IRD, and feeds the research analyses of two research programmes of the IFP Social Sciences Departement: “Labour, Skills, Social networks and Mobilities in India (LAKSMI)”, a programme coordinated by Christophe Jalil Nordman; the “Labour, Finance and Social Dynamics”, a programme coordinated by Isabelle Guérin. These two programmes are the pillars of the *Observatory of Rural Dynamics and Inequalities in South India* (ODRIIS), and NEEMSI constitutes therefore its main quantitative data collection tool.

The ODRIIS, a long-term observatory of rural life in South India, hosted at the IFP and supported by the IRD, analyzes how structural changes in contemporary India are reshaping the organization of work, social hierarchies, and household livelihoods.

Finally, NEEMSI-2 collects data that aim to understand the linkages between household and individuals’ labour, financial practices, socio transfers, skills, social networks and social and spatial mobilities. This includes the investigation of various forces at play, spanning from the role of social structure (norms and institutions), the development and use of interpersonal networks, to the formation of cognitive and non-cognitive skills.

2 | Survey methodology

2.1 Coverage and sampling

NEEMSI-2 collects data from 632 households and 3647 individuals. Out of 632 households, 485 are NEEMSI-1 wave recovered households and, among them, 388 households are RUME survey recovered households. 147 are newly selected households. Also, NEEMSI-2 consists of the “tracking” of 82 individuals who have migrated between 2016-17 NEEMSI-1 wave and 2020-21 NEEMSI-2 wave.

The sampling method follows that of the 2016-17 NEEMSI-1 wave. RUME, NEEMSI-1, and NEEMSI-2’s main data collection take place in 10 rural villages in Tamil Nadu, located at the border between Kallakurichi and Cuddalore districts (Manappakam, Semakottai, Manamthavizhthaputhur, Natham, Korattore, Karumbur, Oraiyur, Govulapuram, Elamthampattu, and Kuvagam). Villages were selected depending on land characteristics - half are irrigated villages, the other half have dry lands- and in terms of accessibility and distance to main roads and to small towns (Panruti, Villupuram, and Cuddalore) in the area.

Inside villages, caste was the main criterion to choose households in the first place: since Dalits (lower castes) are more numerous than middle and upper castes households, more upper castes and middle castes households have been selected in order to better observe cross castes dynamics. Therefore, the number of households selected per caste was not exactly proportional to their actual number in the villages.

For more information about sampling, see the baseline survey documentation.¹

2.1.1 Panel data

A longitudinal study is particularly interesting in the moving context of rural India which experiences both fast changes, but also keeps some rigid social structure. One objective of this quantitative survey was thus to recover all the 2016-17 households, including households who are seasonally migrating for work-related reasons and who moved to other close cities.

NEEMSI-2 wave is conducted on the same NEEMSI-1 492 households, and RUME 405 households in order to constitute panel data. 485 households, and 2628 individuals were

¹Guérin, I., Venkatasubramanian, G., Michiels, M., Roesch, & Natal, A. (2023). *RUME (2010): Survey Manual* (Tech. Rep. No. hal-04153130). IRD-IFP. <https://hal.science/hal-04153130>

recovered, i.e., 1.4 percent attrition rate, from NEEMSI-1. Among them, 382 households, and 1783 individuals were recovered from RUME.

2.1.2 Refresh sample

In addition to the 485 households feeding the panel dataset, around 10 households by village were added to the final sample to increase the sampled population. Since NEEMSI-2 attrition rate of the panel households between the two last waves is low, new families were not selected to have the same characteristics as attritioned households. In each village, around five households were selected in “Colony” area where mostly lower castes live, and around five households in “UR” where middle and upper castes live.

Households in each village area were randomly chosen: every five houses, enumerators asked to answer the questionnaires.

2.1.3 Questionnaire administration

The household questionnaire is answered by one household member, usually the head, about all household members so we have information on each household member for all these modules. Only agriculture, consumption and some schemes are household level questions. Three household members are directly addressed individual questionnaires: the respondent of the household questionnaire (called “ego 1”) and two younger household members (“ego 2” and “ego 3”).² NEEMSI-2 has 1693 “ego” individuals, and among them, 835 are recovered from NEEMSI-1.

NEEMSI-2, like NEEMSI-1, used tablets for data collection, relying on the Survey CTO software. This tool allowed to increase the quality of the data collected, because it is meant to check quality at each stage of the data entry process (missing observations, constraints on answers), and also to reduce the cost, time and errors associated with data entry as this is done instantaneously on the field.

2.1.4 Overview of the NEEMSI-2 sample

The following table summarises NEEMSI-2 household and individual total sample and sub samples.

	Cross sectional setting			Panel setting		
	RUME 2010	NEEMSI-1 2016-17	NEEMSI-2 2020-21	2010 / 2016-17	2016-17 / 2020-21	2010 / 2016-17 2020-21
Number of households	405	492	632	388	485	382
Number of individuals	1928	2696	3647	1826	2628	1783
Egos	0	953	1693	0	835	0

²Members of the household aged between 18 and 25 years old, if no one is available, members aged between 26 and 35 and if no one is available, members aged over 35.

2.2 Breadth of topics and in-depth contextual measures

2.2.1 Household questionnaire

NEEMSI-2 focuses on all aspects of rural life, including in-depth modules on the households' financial practices, social transfers, migrations, remittances, labour, agricultural activities.

NEEMSI-2 questionnaire includes all 2016-17 NEEMSI-1 questionnaire modules, which includes 2010 RUME household questionnaire, on employment, migration and remittances, financial practices (loans, savings, lending practices, gold), agriculture, consumption and housing. NEEMSI-2 kept the same variables in order to observe their variation between the three waves.

NEEMSI-2 supplemented existing modules with new questions in order to have better measures and economic indicators. For instance:

- Marriage module has more detailed questions about the decision (arranged marriage, blood relation, etc.) and the cost (dowry, engagement, marriage, gifts, etc.).
- New questions on mobile finance.

2.2.2 Individual questionnaire

Individual questionnaires provide a range of information on labour force participation and outcomes (including wages and earnings), social networks related to community involvement, interpersonal employment networks, perception of community level employment networks, proximity networks in the village, and cognitive (numeracy, literacy, and Raven's test) and personality traits assessments (Big Five, Grit).

NEEMSI-2 also supplemented existing modules of individual questionnaire with new questions on:

- The locus of control personality trait.
- Decision-making power within the household, in terms of education and employment.
- Working conditions and discrimination at work.

2.2.3 COVID-19 variables

NEEMSI-2 team worked at improving the questionnaires by adding questions in order to understand the effect of the COVID-19 lockdowns, on labour, financial practices, and agriculture. These variables have the prefix *cov-* and concern all households since our entire sample was surveyed after the first lockdown of March 2020.

2.2.4 New features of NEEMSI-2

As the NEEMSI survey is already time consuming (around 3 hours for the household questionnaire and 2 hours per individual questionnaire), We have pre-entered information on the tablets for some households and individuals in order to save time during the interview: socio-demographic characteristics of the household members (checked then during the interview, but not re-entered), education of the adult members, etc.

Thus, questions that are time-invariant (or fixed over time) were not asked. For instance, the **Education** module was not asked for individuals higher than 28 years old. Also, questions related to religion, castes, sex, age, name, were not asked.

2.2.5 NEEMSI-2 migrant tracking questionnaire

In addition to household and individual level questionnaires, NEEMSI-2 wave collects data on individuals who have migrated between 2016-17 NEEMSI-1 wave and 2020-21 NEEMSI-2 wave, for work-related reasons. This survey takes place between July 2022 and September 2022.

The NEEMSI-2 migrant tracking survey seeks to collect data on the characteristics of migrants (by relying on **NEEMSI2-HH** and **NEEMSI2-ego** databases) as well as their environment (characteristics of household members, housing, agriculture, etc.) and their migration path (destination, satisfaction, network, decision).

These data, combined with 2016-17 NEEMSI-1 wave and 2020-21 NEEMSI-2 wave, provide unique information on the interconnection of migration strategies, indebtedness, rural employment and networks while taking into account individual skills.

3 | Files description

3.1 Content

3.1.1 Household questionnaire data

NEEMSI2 household questionnaire includes 5 different databases:

1. `NEEMSI2-HH.dta`
2. `NEEMSI2-occupations.dta`
3. `NEEMSI2-migrations.dta`
4. `NEEMSI2-loans_mainloans.dta`
5. `NEEMSI2-GPS.dta`

The first file, `NEEMSI2-HH.dta`, is the main file. It includes all household questionnaire variables except occupations, migrations, and loans which are sub-modules and kept in separate `.dta` files since they contain more than one observation by individual. `NEEMSI2-GPS.dta` is an auxiliary database containing the geographical coordinates of the dwelling of each household surveyed as part of NEEMSI2.

3.1.2 Individual “Ego” questionnaire data

NEEMSI2 individual questionnaire include 2 databases:

1. `NEEMSI2-ego.dta`
2. `NEEMSI2-alters.dta`

`NEEMSI2-ego.dta` is the main file. `NEEMSI2-alters.dta` contains the list of alters cited following the variable that opens the name generator.

3.1.3 Migrant tracking data

NEEMSI2 migrant tracking questionnaire includes 4 different databases:

1. `NEEMSI2-tracking.dta`
2. `NEEMSI2-tracking_occupations.dta`
3. `NEEMSI2-tracking_alters.dta`
4. `NEEMSI2-tracking_migpath.dta`

The first file, `NEEMSI2-tracking.dta`, is the main file. It includes all tracking questionnaire variables except occupations, alters, and migration step modules, which are sub-modules and kept in separate `.dta` files since they contain more than one observation by individual. This database is composed by variables from `NEEMSI2-HH` database and `NEEMSI2-ego` database.

3.2 Merging files

All files, except `tracking`, can be merged together using `HHID2020` and `INDID2020`.

3.3 Data format

The third column on the pdf questionnaire shows the lists of choices for categorical and dummy variables. In Stata, dummy variables and single choice categorical variables are numerical with labels. Multiple choice categorical variables are strings. Other variables are either integer or decimal (numerical) or text (string).

- . *Missing values.* Values are missing in early questionnaire versions because some questions were addressed later in the survey.
- 66 *Not applicable.* The question is irrelevant in view of the previous question(s). Irrelevant questions are skipped on the tablet (used to appear as “missing”, recoded as “66”).
- 77 *Other.* For categorical variables with “Other” option, new categories were added when needed without deleting information contained in the next question “If ‘Other’, please specify”.
- 88 *Don’t know.* The respondent doesn’t have the answer to the question.
- 99 *No response.* The respondent doesn’t want to reply.

4 | File content and variables

4.1 Household

Head questionnaire:	Household questionnaire
File:	NEEMSIS2-HH.dta
ID:	HHID2020, INDID2020, egoid
Unit:	one observation is one household member

Household database is the main database of NEEMSIS-2 since it includes most variables of the household questionnaire. One observation corresponds to one household member, that is identified with the combination of *HHID2020* and *INDID2020*. Indeed, *HHID2020* represents the unique identifier of a household in NEEMSIS-2, and *INDID2020* represents, within a household, a unique individual identifier.

Other household questionnaire modules (*occupations*, *migrations* and *loans*) are part of separate databases where *HHID2020* and *INDID2020* are no more the unique identifiers since individuals have several occupations, migration activities, and loans. In those sub modules, *occupationid*, *migrationjobid* or *loanid*, combined with *HHID2020* and *INDID2020* are unique identifiers.

Each household belongs to one village *villageid* and one village area *villagearea* (“Ur” or “Colony”). Because of the length of the questionnaires, a financial compensation may have been given to the household. *compensation* identifies whether the household receive a compensation, or not. *compensationamount* gives the amount, in Indian rupees (INR, ₹).

New households New households are identified with *dummynewHH=Yes*. NEEMSIS-2 tried as much as possible to obtain new households linked to households already present in NEEMSIS-1 wave. For instance, a daughter in 2016-17, who got married and is part of a new household in 2020-21.

Four questions allow us to identify if the new household is linked, by parents, to another one already in the data.

1. *newfrompanel*. Are the parents of the household head part of our sample?

2. *villagenameparent*. In which village do his/her parents live?
3. *villagearearent*. In which village area do his/her parents live?
4. *householdidparent*. Who are his/her parents?

4.1.1 General information

interviewplace. Variable informs about the interview place.

address. Key variable in order to draw the household migration trajectory at least in three points in time (2010/2016-17/2020-21).

caste, *religion*, and *comefrom*. The same as in 2016-17.

numfamily. Number of household members, exclude those who have permanently left the household with *dumylefthousehold=1*, but includes members who have permanently left the household with *livinghome=3, 4*.

Household member module

The respondent is not necessarily the household head in the **Household member module** (*relationshiptohead*).

All 2016-17 and new 2020-2021 household members are listed in this module. Each has an ID (*INDID2020*) and name (*name*).

dumylefthousehold is a key variable of NEEMSSIS-2 as it tells us whether the household member under consideration has left permanently the household (*Yes*) or not (*No*), between 2016-17 NEEMSSIS-1 wave and 2020-21 NEEMSSIS-2 wave.

reasonlefthome and *reasonlefthomeother* precise the reason.

In the same vein, *livinghome* tells us about each members's status in the household. Individuals in the first category *Yes, share meals or/and income* are considered as living home because they are coming back home daily and sharing meals. NEEMSSIS-2 attempts to quantify individual migration using two other categories for *livinghome*: *temporary* migration and *permanent* migration. *Temporary* migrants' home is the same as other household members, they mostly left home temporarily for work or studies and share systematically their income with household members who stayed back home. Their income is pooled with other household members' income. For that reason, they are still considered as living home. *Permanent* migrants, on the other hand, got married and settled in order to start a new household or have left home for more than one year and don't plan to come back living home in the short run. Their income is not pooled with other household members' income. Permanent migrants informations are not asked in **Occupation** and **Migration** modules but in **General Information** module only (*permanent* migrants job). Their monetary contribution to the household is accounted as remittances.

dummynewmember. Is the member under consideration new?

dummyformermember. Reversely compared to *dummynewmember*: Is the member under consideration former?

lefthomedurationlessoneyear, *lefthomedurationmoreoneyear*. In months if left less than one year ago (*temporary* migrants), in years if left more than one year ago (*permanent* migrants).

lefthomedestination, *lefthomereason*. Individual migrants', whether temporary or permanent, destination and reason for migrating.

relationshiptohead. Household questionnaire respondent is not necessarily the household head. The head might be different from the 2016-17 head (for instance if the 2016-17 head died or is much older). The status as "head" doesn't depend on age but is left to the respondent's appreciation: who is considered as the head of this household.

jatis is equivalent to *caste* in the questionnaire. We complete this variable with *caste*, which constitutes a simpler variable to analyse castes: 1. *Dalits*, 2. *Middle castes*, 3. *Upper castes*.

4.1.2 Education

Education questions are repeated for each new household member, and former household member below 28 years old. For individuals over 28, we assume that the education level has not changed and therefore use the preloaded data from 2016-17 NEEMSI-1 wave.

If individuals have ever attended school (*everattendedschool=1*), detailed questions follow about education level, the kind of education, and scholarship. If they haven't, we ask why (*reasonneverattendedschool*). If they are not currently at school (*currentlyatschool=0*), we also ask why (*reasondropping*) and who takes the decision to drop (*decisiondropping*); if they are (*currentlyatschool=1*), we want to know about their education expenses (total, school fees, materials, transportation).

4.1.3 Employment

dummyworkedpastyear. Variable repeated for each household member who is living at home (*livinghome=1* or *2*) and who is more than 10 years old. It includes unpaid workers, excluding houseworkers. Even though he/she might not be working on the interview date, the answer is *Yes* if she/he worked over the previous year. When the answer is *Yes*, occupation module has data about this individual's occupation(s). See [NEEMSI2-occupations](#) database.

stoppedworking. If hasn't worked over the previous year, informs if it is permanent because of health issues/accident. Variable existing in the first and second wave and particularly relevant for RUME-NEEMSI context.

4.1.4 Migrations

dummymigration. *Yes* if at least one household member currently migrates for work (household level question). The job has to be outside the village and the household member stays outside more

than one day. If the answer is *Yes*, *migrantlist* gives the list of migrants. In the data, we replace *migrantlist* with *dummymigrantlist* which identify whether the individual under consideration is a migrant (*dummymigrantlist=1*). The list of potential migrants excludes “permanent” migrants since household members and permanent migrant jobs already provided this information. It includes only members living home/temporary migrating (*livinghome=1* or *2*).

See [NEEMIS2-migrations](#) database for detailed variables about migration activities of household members listed in *migrantlist*.

4.1.5 Remittances and gifts

Remittances received

dummyremreceived. *Yes* if at least one household member received money as a remittance over the previous year (household level question).

remreceivedlist. List of household members, who received remittances over the previous year. In the data, we replace *remreceivedlist* with *dummyremreceivedlist* which directly identify whether the individual under consideration received remittances (*dummyremreceivedlist=Yes*) or not (*dummyremreceivedlist=No*).

Remittances questions are repeated for household members who have received remittances and for each remittance source, which can be identified looking at the variables’ prefix *1* to *4* (the data shows maximum four sources), replacing previous ID *remreceivedsourceid*.

remrecipientsourcename1-4 informs about the source of the remittance: $\{\text{name+“household member’s position in the roster”}\}$ or *Someone else*.

If the sender is a household member (*remrecipientsourcename1-4!=Someone else*), *remreceivedsourcerelation1-4* and *remreceivedsourceplace1-4* are skipped since this information are already provided in **General information** module. Other variables apply for all kinds of senders (*remreceivedservices1-4*, *remreceivedmean1-4*, etc.).

remgift1-4 identifies whether the sender also sends gifts over the last year. A set of variables provide information on these gifts.

One variable is related to the effect of COVID-19 lockdown: *covremreceived1-4*.

Remittances sent

dummyremsent. *Yes* if at least one household member sent money as a remittance over the previous year (household level question).

remsenderlist. List of household members, who sent remittances over the previous year. In the data, we replace *remsenderlist* with *dummyremsenderlist* which directly identify whether the individual under consideration sent remittances (*dummyremsenderlist=Yes*) or not (*dummyremsenderlist=No*).

Remittances questions are repeated for household members who have sent remittances and for each remittance recipient, which can be identified looking at the variables' prefix *1* to *5* (the data shows maximum five recipients), replacing previous ID *remsentrecipientid*.

remsentname1-5 informs about the recipient of the remittance: $\{\text{name+“household member’s position in the roster”}\}$ or *Someone else*.

If the recipient is a household member (*remsentname1-5!=Someone else*), *remsentrelation1-5*, *remsentoccup1-5*, and *remsentplace1-5* are skipped since this information are already provided in **General information** module. Other variables apply for all kinds of senders (*remsentervices1-5*, *remsentmean1-5*, etc.).

remsentgift1-5 identifies whether the recipient also received gifts over the last year. A set of variables provide information on these gifts.

One variable is related to the effect of COVID-19 lockdown: *covremsent1-5*.

4.1.6 Financial practices

Most of the financial practices questions build on 2016-17 NEEMSSIS-1 wave, and thus 2010 RUME survey. They include a wide range of semi-formal and informal financial transactions, beyond formal transactions.

Loans

dummyloans. Yes if at least one household member has taken a loan from January 2019 or still has outstanding loans taken before 2018 (household level question).

borrowerlist. List of borrowers. In the data, we replace *borrowerlist* with *dummyborrowerlist* which identify whether the individual under consideration is a borrower (*dummyborrowerlist=1*).

nbloansbyborrower. Number of loans contracted for each borrower.

sumhhloans. Calculate variable: sum of *nbloansbyborrower*. Total number of loans contracted in the household over the period defined above.

loandetails1-79. Catalogue of all members' loans in **NEEMSSIS2-loans_mainloans** database. Variable built in order to select the three main loans of each households.

One question is related to COVID-19 lockdown: *covrefusalloan*.

Main loans

threemainloans. See **NEEMSSIS2-loans_mainloans** database.

Other household incomes

This module asks for other household incomes outside work, land, housing and schemes. It includes incomes from interest, dividends share sale or capital gains (*dummyincomeassets*, *incomeassets*).

Lending money

dummylendingmoney. Yes if at least one household member lent money over the previous year (household level question).

hhlenderlist. List of household members who lent money over the previous year. In the data, we replace *hhlenderlist* with *dummyhhlenderlist* which identify whether the individual under consideration is a borrower (*dummyhhlenderlist=1*).

Two variables are related to COVID-19 lockdown: *covlendrepayment* and *covlending*.

Recommendation for loans, chit funds, savings, gold, insurance

Each of these five modules starts with a dummy variable to know if at least one household member recommended someone (*dummyrecommendgiven*), belongs to a chit fund (*dummychitfund*), has a bank account (*dummysavingaccount*), has gold (*dummygold*) or has a formal insurance (*dummyinsurance*).

The second variable lists household members concerned by the module: *recommendgivenlist*, *chitfundbelongelist*, *savingsownerlist*, *goldownerlist*, *insuranceownerlist*. However, as previously, in the data, we replace *varlist* with *dummyvarlist* which identify whether the individual under consideration is in the list (*dummyvarlist=1*).

Members can belong to several chit funds, have several bank accounts or several insurances (*nbchitfunds*, *nbsavingaccounts*, *nbinsurance*). In that case, variables are duplicated with suffixes *1* to *3* for chit funds, *1* to *4* for bank account, and *1* to *6* for insurances, in order to have the same information for each chit fund, bank account and insurance.

Several questions are related to the COVID-19 lockdown: *covchitfundstop*, *covchitfundreturn*, *covsavinguse*, *covsavinguseamount*, *covsoldgold*, *covsoldgoldquantity*, *covlostgold*, *covgoldpledged*.

Mobile money

A new feature of NEEMIS-2 is the integration of questions related to the mobile finance/money: *usemobilefinance*, *usemobilefinancetype*, and *usemobilefinanceother*.

4.1.7 Agriculture

Agriculture questions mostly build on NEEMIS-1 wave. Agriculture module is household level questions (in **NEEMIS2-HH**, same values will appear for all household members).

Land

dummyeverhadland. Answer is “Yes” if the household owns or leases land from someone else now or/and since August 2016. In that case, following sets of questions, each starting with a dummy, are asked if the dummy is 1:

- Land at present: *ownland*, *leaseland*;
- Land purchase and selling since August 2016: *landpurchased*, *landlost*;
- Own land leasing over the last 3 years: *dummyleasedland*;
- Leasing someone else's land over the last 3 years: *dummyleasingland*.

One question is related to COVID-19 lockdown: *covsellland*.

Cropping

This module opens if the household own or lease land from someone else at present (*dummyeverhadland*=1) and if the household cultivates at least one crop (*productlist* answer is not *None*).

productlist. List of crops. Crops variables recoded by type of crop using a suffix: *_paddy*, *_cotton*, *_sugarcane*, *_savukku*, *_guava*, *_groundnut*, *_millets*, *_cashew*, *_other*.

Besides incomes from cropping, this module also deals with cultivation costs (production cost and labour cost). Since **Occupation** module did not ask those questions when occupation is *Agricultural activity on own household farm* (*kindofwork*=1), NEEMSIS-2 focuses on them here. Several variables are related to the effect of COVID-19 lockdown: *covsubsistence*, *covsubsistencereason*, *covsubsistencesize*, *covsubsistencenext*, *covsubsistencereasonother*, *covharvest*, *covselfconsumption*, *covharvestquantity*, and *covharvestprices*.

Livestock

Questions asked no matter if the household owns and leases land. Livestocks can represent an important share of the household income and is significant of the household's wealth.

livestocklist. List of animals. Livestock variables recoded by kind of livestock using a suffix: *_cow*, *_goat*, *_chicken*, *_bullock*, and *_bullforploughing*.

To account for the flow of livestock over the last 3 years, NEEMSIS-2 has details cattle loss and details cattle sold questions.

One question is related to COVID-19 lockdown: *covselllivestock*.

Farm equipment

It contains owned and borrowed farm equipment.

equipmentlist. List of farm equipment. Equipment variables recoded by kind of equipment using a suffix: *_tractor*, *_bullockcart*, *_plowingmach*.

equipmentborrowedlist identify which equipment is borrowed.

One question is related to COVID-19 lockdown: *covsellequipment*.

4.1.8 Consumption and assets

NEEMIS-2 accounts for consumption practices, consumer durables and goods. All variables are at the household level, meaning that in **NEEMIS2-HH**, same values will appear for all household members.

Expenses

Food expenses are measured on a weekly basis on average. Health, festivals and death expenses are recorded on an annual basis. Festivals expenses include festivals for household members but also festivals outside the household, and exclude marriage expenses accounted for in **Marriage** module.

In NEEMIS-2, a new variable is added to observe who take the decisions in terms of consumption (*decisionconsumption*).

The effect of COVID-19 lockdown policy on consumption is measured through several variables: *covfoodenough*, *covfoodquality*, *covgenexpenses*, *covexpensesdecrease*, *covexpensesincrease*, *covexpensesstable*, *covplacepurchase*, and *covsick*

Good details

listgoods. List of potential consumer durables owned by the household.

Consumer durable variables recoded by kind of consumer durables using a suffix: *_car*, *_bike*, *_fridge*, *_furniture*, *_tailormach*, *_phone*, *_landline*, *_camera*, *_cookgas*, *_computer*, *_antenna*.

Good property is significant of the household's wealth and public schemes distributing free consumer durables, hence questions about the year of purchase and how the good was bought (credit, free, cash, etc.).

Two questions are related to the effect of the COVID-19 lockdown: *covsellgoods*, *covsellgood-sother*.

4.1.9 Marriage

dummymarriage. Yes if at least one household member got married since 2016 (household level question). The objective is to know who got married in the household since 2016-17 wave.

marriedlist. List of household members who got married. In the data, we supplement *marriedlist* with *dummymarriedlist* which identify whether the individual under consideration is get married since 2016 (*dummymarriedlist=1*).

This list includes household members who left home (*livinghome=3*) after marriage (mainly daughters, but also sons and daughters who started a new household elsewhere), who joined the household after marriage (mainly daughters-in-law) or who stayed in the household even after marriage (mainly sons).

Marriage module consists of five sets of questions:

- Marriage type and decision: *marriagetype, marriageblood, marriagearranged, marriagedecision*.
- Marriage cost: *marriagedowry, engagementtotalcost, marriagetotalcost, marriageexpense*,
- Marriage finance: *howpaymarriage*,
- Marriage loans details: variables are available in **NEEMSI2-loan_mainloans** database, identified with *loan_database=marriage*. It includes: *loanamount, loandate, loanlender*, lender characteristics, *loansettled, loanbalance*.
- Marriage gifts: *marriagegiftsource*. List of gift senders for marriage. Gift sender is recoded by kind of gift sender using a suffix: *_WKP* (well-known people), *_rela* (relatives), *_empl* (employer), *_mais* (maistry), *_coll* (colleague), *_frie* (friends), *_SHG* (self-help groups).

4.1.10 Housing and facilities

This module accounts for the household's housing status (owner, tenant) at the date of the interview, and the house facilities (water, electricity).

Two questions are related to COVID-19 lockdown: *covsellhouse, covsellplot*.

4.1.11 Schemes

Most of the rural households in Tamil Nadu benefit from public schemes and pensions: only 2 households of our study area is not benefiting at all from any of the schemes listed in *schemeslist*. Schemes variables were recoded by kind of schemes and pension using a suffix: *_freehouse, _cashfunerals, _freecows, _freelpg, _educfees, _farmequip, _land, _cashmarriage, _goldmarriage, _oldage, _widows, _maternity, _disability, _retirement, _shgloan, _freepatta, _laptop*, and *_girlprotection*.

Ration card has the prefix *rationcard-*.

For NREGS (National Rural Employment Guarantee Schemes), *nreganberdaysworked* and *nregaincome* give values for each individual who declared *Yes* in *dummyschemelist_nrega*.

4.2 Occupations

Head questionnaire:	Household questionnaire
File:	NEEMSYS2-occupations.dta
ID:	HHID2020, INDID2020, occupationid
Unit:	one observation is one occupation

This module gathers information about all occupations during the past year, i.e., the twelve months before the interview date. For each occupation, the interviewees declared the type of occupation (*kindofwork*), an estimated number of days a month (*daysamonth*) and hours a day worked (*hoursaday*) –the number of hours a year calculated (*hoursayear*), an estimated annual income (*annualincome*), and the date they took up the job *datestartoccup*. It includes unpaid workers. Individuals who haven't worked the past year answer questions about the reasons they haven't worked (*reasonnotworkpastyear*, *stoppedworking*).

However these individuals may have worked before. Thus, *everworksalaried* and *kindofworkinactive* provide further information. Also, the Individual questionnaire provides further information about labour force participation (see *NEEMSYS2-ego.dta* database).

The observation unit is the occupation identifier (*occupationid*).

kindofwork. This variable is the main variable of this module. Depending on its value, specific questions are asked about self-business and wage jobs:

1. *Agricultural activity on own household farm* → See **Agriculture** module (*NEEMSYS2-HH.dta*),
2. *Self-employed, own account worker* → See **Self-employment** module (*NEEMSYS2-occupations*),
3. *Agricultural salaried job* → See **Salaried job** module (*NEEMSYS2-occupations*),
4. *Non agricultural salaried job* → See **Salaried job** module (*NEEMSYS2-occupations*),
5. *Unpaid worker in family business (non-agricultural)* → See **Salaried job** module (*NEEMSYS2-occupations*),
6. *Unpaid worker in other business (non-agricultural)* → See **Salaried job** module (*NEEMSYS2-occupations*),
7. *Unpaid worker in own farm* → See **Salaried job** module (*NEEMSYS2-occupations*),
8. *Unpaid worker in another farm* → See **Salaried job** module (*NEEMSYS2-occupations*).

4.2.1 Self-employment

These questions exclude farm businesses since there is a module in household questionnaire dealing specifically with agricultural occupations. See **Agriculture** module in *NEEMSYS2-HH.dta*.

Business labourer

If the business involves several people's labour (*dummybusinesslabourers=1*), *nbbusinesslabourers* gives the number of labourers. Labourers were listed as much as possible during the interview and some businesses involve lots of labourers, hence a high number of observations for some occupations (suffix *1* to *15*).

dummybusinesslabourerhmember1-15. Business labourer is or not a household member. If the answer is *Yes*, following variables are skipped because we already have information about this person in household information and occupation modules.

relationshipbusinesslabourer1-15 variable informs if it is a household member or not. There are several variables about labourers' socioeconomic characteristics and jobs (address, relation, caste, type of job and salary).

businesslabourerdate1-15. This variable is to know the date at which labourer have joined the business.

4.2.2 Salaried jobs

Wage jobs include a wide range of occupations from daily employment to permanent jobs, piece rate to monthly wages, but they have the common characteristics that the worker is under the supervision of an employer. Questions in this module are also asked to unpaid workers who said they worked over the past year (excepting housewives). This section includes farm and non-farm jobs and the term "salaried" job here doesn't mean we talk about formal employment. Wage job questions are related to the workplace, the employer's characteristics and the terms of the employment.

Employer informations

As workers can have several employers (*nbemployer*), *typewageemployer*, *relationemployer*, *casteemployer*, *othercasteemployer*, and *otheremployertype* are repeated for each employer. Prefix from *1* to *9* are used.

4.3 Migrations

Head questionnaire:	Household questionnaire
File:	NEEMISIS2-migrations.dta
ID:	HHID2020, INDID2020, migrationjobid
Unit:	one observation is one occupation

The list of household migrants for work (*dummymigrantlist*) is given in NEEMISIS2-HH, **Migration** module when at least one household member migrates for work, except “permanent” migrants. Several questions were then asked about migrants’ migration jobs. These variables are in NEEMISIS-migration. Each line corresponds to one kind of migration job (brick kiln, sugarcane cutting, farm and non-farm coolie, construction work, work in a private company, self-employment), identified by the variable *migrationjobid*.

Migration jobs recorded in this module are also supposed to appear in NEEMISIS2-occupations database as occupation of respective household members, except for children below 10 years old. *migrationtype*. Individual or familial migration. Familial means with household members listed in the Household members module. It doesn’t include family members outside the household roster.

dummyadvance. Seasonal advance on migrants’ salary that they have to pay back, at best at the end of the season (can take many years to repay the employer).

- If the migrant received an advance in this migration job, several questions are asked, for instance: *migrationadvanceamount*, *dummyadvancebalance*.
 - If the advance is still in balance (*dummyadvancebalance=1*), we asked about its amount, potential issues faced by the migrant (*advanceamountbalance*, *advancebalanceproblem*, *advancebalanceperception*).

Following variables are related to the effect of COVID-19 lockdown on advances: *covdeal-covid19*, *covmealemployer*, *covmealemployerfree*, *covpressureadvance*, *covpressuremanage*, *covmigrationagain*, *covmigrationagainreason*, *covpressuremanageother*, *covmigrationagaindate*, *covstayworkagain*, and *covmigrationotherreason*.

migrationjobtype. Degree of stability of migration jobs but permanent and fixed-term jobs do not mean signed contracts (often implicit contract/oral agreement).

migrationsalary. Migration season’s total wage including potential allowances on the migration place but excluding the advance. It should match with *annualincome* variable in NEEMISIS2-occupations database.

migrationpension, *migrationbonus*, *migrationinsurance*. Should similarly match with *salaried-jobpension*, *salariedjobbonus*, *salariedjobinsurance* in NEEMISIS2-occupations database.

migrationmainoccup. Workers' perception about the importance of this migration job compared to other jobs, if any over the last year.

4.4 Loans

Head questionnaire:	Household questionnaire
File:	NEEMISIS2-loans_mainloans.dta
ID:	HHID2020, INDID2020, loanid
Unit:	one observation is one loan

Two sets of variables are available to document the borrowing practices of the households (**Loans** and **Main loans** modules). Both sets are combined in `NEEMISIS2-loans_mainloans` database. In a first stage, the list of borrowers and the number of loans contracted (*borrowerlist* and *nbloansbyborrower*) are recorded in `NEEMISIS2-HH`. Several questions were then asked for each loan contracted by each member. These variables are in `NEEMISIS2-loans_mainloans`.

Each line corresponds to one loan, identified by the variable *loanid*.

Marriage loans are also recorded in this database. Marriage loans are identified with the variable *loan_database*, which takes *marriage* if the loan comes from the **finance marriage** module, and *finance* if the loan comes from the **Finance** module. Thus, *loanid* combined with *loan_database*, are the unique identifier.

Four variables are related to the effect of COVID-19 lockdown: *covfrequencyinterest*, *covamountinterest*, *covfrequencyrepayment*, and *covrepaymentstop*.

Compared to 2016-17 NEEMISIS-1 wave, NEEMISIS-2 has more detailed questions about the lender, in order to complete the **Social Network** module: *snmoneylenderoccup* for instance. As a result, informal lenders also appear in the alters database. See section 4.6.

loaninfo. Concanates three variables for each loan: the borrower's name (*name*), the amount of the loan (*loanamount*) and the reason for borrowing (*loanreasongiven*). Thanks to this variable, all members' loans are catalogued in `NEEMISIS2-HH` (*loandetails1-79*).

In a second stage, three loans per household (*threemainloans*) were selected among the pool of all members' loans (*loandetails1-79*) in `NEEMISIS2-HH`. Denominated in the questionnaire as “main loans”, there are not necessarily the largest loans. Households had to select the loans whose repayments were regarded as the most crucial, with the constraint that one of these main loans had to be a woman's one if such loan existed in the household. For these three loans per household, additional questions were asked and are included in `NEEMISIS-loans_mainloans`. The corresponding variables are recorded on the lines of the loans chosen as main loans, and are consequently missing for the other loans.

4.5 Ego

Head questionnaire:	Individual questionnaire
File:	NEEMISIS2-ego.dta
ID:	HHID2020, INDID2020, egoid
Unit:	one observation is one ego

NEEMISIS-2 includes the individual unit of analysis, using a specific questionnaire explicitly addressed to three household members. Everything has been done so that the ego individuals in NEEMISIS-1 are ego individuals in NEEMISIS-2. For the new egos, three household members were interviewed with the individual questionnaires.

- Ego 1 is the household questionnaire respondent, except in few instances where he/she can't answer the individual questionnaire.
- Ego 2, 3 are randomly selected on a criterion of age since NEEMISIS-2 aims at collecting more detailed information about young people. Two young members between 18 and 25 in the household roster are first selected and interviewed if available. If there is no such individuals in this age range or these persons are not available, we look for other young members between 26 and 35. Once again, if there is no one in that age range or not available to answer, NEEMISIS-2 interview household members above 35.

NEEMISIS-2 has data for 1653 egos. There are two different individuals IDs in this database: *INDID2020* refers to the respondent position in the household roster (same ID used in the household database) and *egoid* informs us if the respondent is Ego 1, Ego 2, or Ego 3.

NEEMISIS-2 addresses exactly the same questions to Ego 1, Ego 2, and Ego 3. A diversity of approaches and tools were used to design the individual questionnaire: it includes modules on labour force participation and professional aspiration (1), formal and informal social network modules (2) and cognitive skills and personality traits inspired by psychology and cognitive sciences (3).

4.5.1 Employment variables

Labour force participation variables complement household questionnaire employment module. While household questionnaire employment questions focus on employment over the previous year, the individual questionnaire goes further into the interviewee's employment history (previous jobs, first job, method for searching jobs) and asks about job aspirations.

workedpastyear. In the household questionnaire (*NEEMISIS2-HH*) this dummy variable reports if household members have worked or not during the year before the survey. Labour force participation questions depend on this variable.

<i>workedpastyear</i>			
<i>Yes</i> First job <i>workpastsevendays</i>		<i>No</i> Job aspiration not working past 7 days <i>everwork</i>	
<i>Yes</i> Job aspiration working past 7 days	<i>No</i> Search job last month Job aspiration not working past 7 days	<i>Yes</i> First job	<i>No</i> Search job since 15 years <i>searchjobsince15</i>
		<i>Yes</i> <i>nbermonthsearchjob</i>	<i>No</i> <i>reasondontsearchjobsince15</i>

everwork. If ego respondent hasn't worked over the past year, this variable captures if she/he has ever worked, including unpaid jobs and excluding housework.

- **First job.** Addressed to Egos who declared working the previous year or who have ever worked. Includes unpaid first jobs except housework. Respondents whose first job was unpaid (*kindofworkfirstjob* ≥ 5) are also asked the age when they started their first paid job (*agestartworkingpaidjob*) in order to account for the time between the entrance in the labour force and the access to paid work. However, part of respondents may never have had access to paid work such as those who remained unpaid workers or/and have stopped working. So this variable should be interpreted with caution and crossed with other labour force participation variables. 66, 99 values correspond to *not applicable* and *no response*.
- **Search job last month.** Respondents have worked the previous year but not the past seven days. They are questioned about job search/business initiative over the last month and the reason why they didn't work.
- **Search job since 15.** Respondents have never worked, including as unpaid workers but they can be doing housework. They are asked about job search/attempt to set up a business since 15 years old and the reason why they couldn't work.
- **Job aspirations not working past 7 days.** Respondents have already worked before but not the last 7 days or have never worked. Questions about availability to work and under what conditions and their job preferences.
- **Job aspirations working past 7 days.** Accounts for the gap between current job conditions and outcomes (salary, number of hours worked) and workers' aspirations.

First job

Details about the first job are directly addressed to Ego 1, 2, and 3.

It includes the kind of work (*kindofworkfirstjob*), age start working unpaid and paid jobs (*agestartworking* and *agestartworkingpaidjob*), and the social network to find the job (*methodfindfirstjob* and *snfindfirstjob*) for instance.

Main occupation variables

Details about occupations are now asked directly to Ego 1, 2, and 3 whereas only one household member (usually the head) addressed occupation questions about each household member in the household questionnaire (**NEEMSIS2-occupations**).

Following modules are consequently more precise and detailed. They focus on the respondent's "main occupation" when he/she is reported as working over the previous year in the household questionnaire (*workedpastyear=1*). For this purpose and using household questionnaire information, NEEMSIS-2 defined "main occupation" as the respondent's most time-consuming job over the past year, i.e., the occupation with the maximum of hours a year (*maxhoursayear*) since many interviewees have more than one occupation.

selected_occupname. Calculated variable. Occupation ID of Ego's most time consuming job over the past year.

Since that definition of "main occupation" may not be satisfactory, NEEMSIS-2 asks the respondent if this calculated main occupation is really his/her main occupation (*dummymainoccupation2*) and in case it is not, the respondent informs about which occupation he/she feels is the main one (*othermainoccupation2*).

dummyseasonalmainoccup. Is, the occupation, a regular one over the past year (fixed number of days a month and hours a day)? Regular in the sense a more or less fixed number of days a month over the year and hours a day. In the household questionnaire (**NEEMSIS2-occupations**), *daysamonth* and *hoursaday* variables give the average number of days a month and hours a day spent in an occupation but do not provide any information about the variability of days worked depending on months and of hours worked depending on days.

inddaysayear2. Effective working days a year if *dummyseasonalmainoccup=1*.

indhoursaday2. Effective working hours a day if *dummyseasonalmainoccup=1*.

indhoursayear2. Calculated variable. Effective working hours a year:

(*inddaysayear2*indhoursaday2*).

Main occupation schedule over the past year. If the respondent feels it is not regular (if *dummyseasonalmainoccup=0*), he/she can make a calendar of this "main occupation" over the previous year.

The initial objective of this module was to have more detailed variables than *monthsayear* (**NEEMSIS2-occupations**) that could inform about the months worked during the year in order to capture the jobs seasonality. And to have the number of days a month (*inddaysamonth*) and the number of hours a day (*indhoursaday*) worked over the previous year depending on each month.

We add suffix according to the month: *_chithirai*, *_vaikasi*, *_aani*, *_aadi*, *_aavani*, *_purataasi*, *_iypasi*, *_karthigai*, *_maargazhi*, *_thai*, *_maasi*, and *_panguni*.

Characteristics of main occupation in the past 12 months *beforemainoccup*. Kind of job of the previous “main occupation”. What is defined as a main occupation is left here to the appreciation of the respondent. Depending on the kind of job of the current main occupation (*mainoccupytype*), more details are asked about the previous and the current main occupation:

- *mainoccupytype=2* *Self-employed*.

Self-employment as main occupation, including:

- Questions about the previous job if it was a wage job (*dummypreviouswagejob*).
- Questions about the investment.
- Questions about unpaid (including household members) and employed workers in the business when it was set up and on the date of the interview.
- Details about paid business workers (number of paid workers, frequency of payments, total labour cost).
- Details about business income, expenses (gross receipt, fixed costs, production costs, social security).
- Details payment in kind.
- Details in recruitment of workers.

These questions exclude agriculture on own/lease land (*mainoccupytype=1*) since there is a module in household questionnaire dealing specifically with agricultural occupations (*NEEMSIS2-HH*).

One question is related to COVID-19 lockdown: *covdifficulties*.

- *mainoccupytype>=3* *Salaried job* or *Unpaid job*.

Wage job as main occupation, including:

- Details terms of wage job (size employing structure, contract, job type, social security).
- Details salary (net payment).
- Details payment in kind.

Most questions target only Egos who are salaried as a main occupation (*mainoccupytype=3* or *mainoccupytype=4*). A few variables also include unpaid workers (*mainoccupytype>=5*) such as *wagejobnbworkers* (only unpaid worker outside own farm or business), *wagejobpaymentinkind*, *wagejobpaymentinkindlist* and **Details payment in kind** module.

COVID-19 lockdown variables

A set of questions is dedicated to the effect of COVID-19 lockdown. This module consists of questions relating to the labour force participation, if there has been no COVID-19 lockdown.

covoccupationname. Name of occupation.

covkindofwork. What kind of work would be *covoccupationname*?

Then, a set of questions is asked about the hours worked, the income, and the effect of COVID-19 lockdown. We have renamed *annualincome* in the questionnaire to *covannualincome* in the data to avoid confusion with the variable of the **NEEMSIS2-occupations** database.

Decision to work

A new feature of NEEMSIS-2, is the adding of a set of questions related to the decision to work. These questions are asked to try to measure the bargaining power of individuals. These questions include, for instance:

- *decisionwork*. Who has the most say in decisions about your work?
- *opinionworkingwoman*. What do you think about a woman who takes the decision to work?

Aspirations

As with NEEMSIS-1, NEEMSIS-2 include questions related to the job aspirations of Ego 1, 2, and 3. Questions go from *readystartjob*: “If you were offered work, are you available to start within the next two weeks?” To *aspirationminimumwage2*: “If you were to take another job in this village or area, what is the lowest daily net (take-home) wage that you are willing to accept (in rupees)?”

Working conditions and discrimination

Another new feature of NEEMSIS-2, is the adding of a set of questions related to the satisfaction at work, working conditions and discrimination.

This new module includes questions related to the:

- Satisfaction at work (*respect*, *satisfactionsalary*, etc.).
- Working conditions (*executionwork1-9*, *problemwork1-10*, etc.).
- Discrimination (*discrimination1-9*, *resdiscrimination1-5*, etc.).

4.5.2 Social networks variables

Formal social capital variables

NEEMIS-2 measures formal social capital, i.e., individuals' involvement into public institutions, such as associations, political parties, trade unions, Self-Help Groups, youth unions, farmer unions, village councils (panchayat), religious groups, etc, with *associationlist*.

Variables in this module are numbered from 1 to 2 corresponding to the first and second association/institution Egos belong to (not to be confused with values in *associationlist*). In other words, one ego has a maximum of 2 associations.

associationtype1-2. The value in data corresponds to the kind of association/institution selected in *associationlist*.

associationname1-2. Name of association/institution. For instance, the name of the political party.

Further details are asked about each association, such as the degree of participation to that association, its size, the role of that association to find a job or to support Ego's business.

dummyassorecommendation1-2. This variable opens the name generator module if the answer is *Yes*.

Several questions are related to the COVID-19 lockdown: *covassociationhelp*, *covassociationhelplist*, *covassociationhelplistothe*, *covassociationhelptype*, *covassociationhelptypeothe*

Informal social capital variables

In NEEMIS-2, informal social capital data collection follows the name generator sociological method. Name generators consist of several questions designed to invite respondents (i.e., egos) to recall and elicit people (Alters) with whom they maintain certain types of direct relationships, to delineate the core members of their network.¹

In order to identify Ego's relationships, NEEMIS-2's individual questionnaire has chosen the criterion of interdependency or regular interaction among individuals involved in social relations, thus conveying resources needed in different dimensions of labour activities.^{2,3} The criterion chosen has the advantage of being clear and unequivocal, as it is less likely to be interpreted differently across respondents and we neither *a priori* define nor exclude specific social categories. The network composition, effective and potential, is the consequence of the nature of the network support that we observe.

In **NEEMIS2-ego**, the following questions open the name generator module:

¹Marsden, P. V. (2005). Recent Developments in Network Measurement. In P. J. Carrington, J. Scott, & S. Wasserman (Eds.), *Models and Methods in Social Network Analysis* (pp. 8–30). Cambridge University Press. <https://doi.org/10.1017/CBO9780511811395.002>

²Campbell, K. E., & Lee, B. A. (1991). Name generators in surveys of personal networks. *Social Networks*, 13(3), 203–221. [https://doi.org/10.1016/0378-8733\(91\)90006-f](https://doi.org/10.1016/0378-8733(91)90006-f)

³Crossley, N., Bellotti, E., Edwards, G., Everett, M. G., Koskinen, J., & Tranmer, M. (2015). *Social Network Analysis for Ego-Nets*. SAGE Publications Ltd. <https://doi.org/10.4135/9781473911871>

1. What is the name of the lender for this business loan? *snbusinesslender*
2. If you need to recruit a worker, would you ask someone you are directly in contact with to help you to find him/her (excluding your household members)? *snrecruitworker*
3. Has someone recommended you to join this association? *dummyassorecommendation1-2*
4. Did someone help you to find your current main occupation? *snfindcurrentjobname*
5. If you needed to find a new job, is there someone you are directly in contact with you would feel comfortable to ask for help? *snfindjob*
6. In the last 12 months, did you recommend someone you are directly in contact with for a job opening that you knew of (excluding household members)? *snrecommendforjob*
7. Was anybody you are directly in contact with hired based on your recommendation (excluding household members)? *snrecojobsuccess*
8. Who are the 4 non-relatives you are talking the most? *sntalkthemost*
9. If you had a medical emergency and were alone at home whom you ask help for getting to a hospital? *snhelpemergency*
10. Name the 4 close relatives, aside those in this household, who also live in the village. *sncloserelouthh*

Also, two questions are related to the COVID-19 lockdown, and open the name generator:

11. *covsnhelpreceived*. Did you ask help to someone during the COVID-19 lockdown?
12. *covsnhelpgiven*. Did you give help to someone during the COVID-19 lockdown?

Last, in the **NEEMIS2-loans_mainloans** database, one question open the name generator module.

13. *loanlender*. From whom did you get loan?

Unlike NEEMIS-1, in NEEMIS-2 the data collected for interpersonal networks has been centralised at household level, meaning that as soon as a person in a household mentions an Alter, that person is recorded in the family name generator in order to estimate the extent to which the same person is linked to several respondents in the same household. In addition, NEEMIS-2 fills a gap in NEEMIS-1 by adding the variable “Did someone help you to find your current occupation?” Additionally, all the questions opening the name generator have been harmonised so that all questions linked to a network systematically have the same set of questions.

Some of the variables on interpersonal networks⁴ are not directly linked to the credit market or the employment market to verify the extent to which the Alters on the credit market or the employment market are the same types of Alters as those who share intimacy, provide regular help and share the respondents' daily lives.

Effective and potential networks

- **1, 3, 4, 6, 7, 8, 10, 11, 12, 13** invite egos to mention their effective network, i.e., people from whom they have effectively asked help or helped.
- **2, 5, 9** concerns potential network.

When the answer to any of these questions is *Yes* (except for *snbusinesslender*, *sntalkthemost*, *snclouserelouthh*, and *loanlender*) egos are invited to elicit Alters (household members are supposed to be excluded).

For the 4 exceptions (*snbusinesslender*, *sntalkthemost*, *snclouserelouthh*, and *loanlender*), egos should list alters directly.

- If *-namelist* value is 888, it means that ego mentions one or several “New person” (or new Alters) for the corresponding network.
- If *-namelist* displays another value, it means that Egos selected an alter he/she already mentioned in a previous network name generator question.

Since NEEMIS-2 was using tablets for data collection, questionnaire was coded so that alters are stored in the tablet. Then when egos need to list an alters, enumerators can list alters already cited by the egos, to avoid collecting the same information several times and to avoid lengthening the survey.

A specific database deals with alters data (*NEEMIS2-alters.dta*).

Other network variables

NEEMIS-2 also includes following network information:

- Network size: *nbcontactphone*, *nberpersonfamilyevent*.
- Resources of the network: contact with people who occupy specific positions or are leaders of large organisations and whom ego can contact easily (*contactlist*, *nbcontact*, *dummycontactleaders*, *contactleaders*).
- Network perception:

⁴For example, people to contact in the event of a medical emergency, the nearest people in the village outside the household, people who provided or received help during COVID-19, etc.

- Trust in the community: *networktrustemployees*, *networktrustneighborhood*, *networkpeoplehelping*, *networkhelpkinmember*.
- Questions related to the COVID-19 lockdown: *covnetworktrustneighborhood*, *covnetworkpeoplehelping*, *covnetworkhelpkinmember*, *covinstil-10*, and *covcontactinstitution*.

4.5.3 Personality traits and cognitive skills variables

Personality traits

The personality traits assessment relies on the “Big-5 taxonomy” questionnaire (35 questions), and further includes the socio-emotional skill “Grit” (the tendency to sustain interest in long-term goals and persistence). The Big Five taxonomy has since been replicated across cultures and developmental stages of the life course.^{5,6}

Also, a new feature of NEEMIS-2, is the adding of a set of questions that measure the “Locus of control” (*locuscontrol1-6*).⁷ It represents the degree to which people believe that they, as opposed to external forces, have control over the outcome of events in their lives. Locus of control is measured following the KOVALEVA scale, which consists of a set of six affirmative questions.⁸ As far as possible, these questions were addressed to the respondent alone with less people around. When it was not possible to interview ego alone (relatives in the house), enumerators warned them not to influence or comment ego answers.

Cognitive skills

The cognitive skills assessment includes:

- RAVEN Coloured Progressive Matrices, consisting of 36 matrices (*a1-12*, *ab1-12*, *b1-12*). RAVEN’s Coloured Progressive Matrices are a nonverbal test used to measure abstract reasoning. It consists of identifying the missing element from a set of answer possibilities that logically completes a pattern.
- A literacy test, consisting of four questions (*canread1a*, *canread1b*, *canread1c*, *canread2*).
- A numeracy test, consisting of six questions (*numeracy1-6*).

⁵John, O. P., & Srivastava, S. (1999). The Big Five Trait taxonomy: History, measurement, and theoretical perspectives. In L. A. Pervin & O. P. John (Eds.), *Handbook of personality: Theory and research* (pp. 102–138). Guilford Press.

⁶Soto, C. J., John, O. P., Gosling, S. D., & Potter, J. (2008). The developmental psychometrics of big five self-reports: Acquiescence, factor structure, coherence, and differentiation from ages 10 to 20. *Journal of Personality and Social Psychology*, *94*(4), 718–737. <https://doi.org/10.1037/0022-3514.94.4.718>

⁷Rotter, J. B. (1966). Generalized expectancies for internal versus external control of reinforcement. *Psychological Monographs: General and Applied*, *80*(1), 1–28. <https://doi.org/10.1037/h0092976>

⁸Kovaleva, A. (2012). *The IE-4: Construction and Validation of a Short Scale for the Assessment of Locus of Control* [Doctoral dissertation, GESIS - Leibniz-Institut für Sozialwissenschaften]. Köln, Germany. <https://doi.org/10.21241/ssar.37119>

4.6 Alter

Head questionnaire:	Individual questionnaire
File:	NEEMSYS2-alters.dta
ID:	HHID2020, INDID2020, alterid
Unit:	one observation is one alter

During the survey, the individuals surveyed were asked to list individuals, called “Alter”, who had a link with them. This is mainly the result of the network modules (i.e., “informal social capital variables” and “formal social capital variables”) in the individual questionnaire, but also informal lenders from the household questionnaire. This database consists of 12 055 observations, i.e. 12 055 alters. 95.61% of them were declared by egos, the rest by non-egos. In all the modules opening the name generator, individuals were able to mention several alters. Once an alter has been cited by a household member in the household or individual questionnaire, it is saved in the tablet so that it appears in the possible alters when a name generator appears. Thus, for each individuals (*HHID2020* and *INDID2020*), the variable *alterid* identifies different unique alters. The combination of *HHID2020*, *INDID2020*, and *alterid* is therefore the unique identifier for this database.

networkpurpose1 captures the first type of relation where the alter was mentioned for the first time. The options are:

1. *Loan*
2. *Business loan*
3. *Help recruit worker*
4. *Help join association*
5. *Find current job*
6. *Find future job*
7. *Recommend job opening*
8. *Hired on recommendation*
9. *Talk the most*
10. *Medical emergency*
11. *Close relatives*
12. *Receive help during COVID*

13. Give help during COVID

In other words, the variable captures the opening possibilities of the name generator. Individuals who are not egos had access to the generator only through the intermediary of the lenders, i.e., *Loan*. The other ways of accessing the name generator are for egos only.

Alters can be mentioned several times across social network modules, hence a set of new variables were created (*networkpurpose*) with suffixes from 2 to 11.

Name generators (*alterid* and *altername*) are followed by questions called “name interpreters” that gather information on alters’ attributes. On the one hand, we have information specific to alter such as *namealter*, *sex*, *age*, *castes*, *educ*, *occup*, *occupother*, *living*, and *phonenb*. On the other hand, we have information that is specific to the relationship between the individual who responded to the survey and his or her alter such as the relationship (*dummyfam*, *friend*, *wkp*, and *labourrelation*), *compared*, *duration*, *meet*, *meetother*, *meetfrequency*, *invite*, and *reciprocity1*, *intimacy*.

Alter database also includes *associationid*, i.e., specifies the kind of association (*associationlist* in **NEEMSIS2-ego**) for which Ego has received a recommendation in order to enter that association/institution.

4.7 Household and Ego tracking

Head questionnaire:	Tracking questionnaire
File:	NEEMISIS2-tracking.dta
ID:	HHID2022, INDID_mig, INDID_nonmig
Unit:	one observation is one household member

A household member is uniquely identified using *HHID2022* (unique household id) and *INDID_mig* (if the individual is a migrant), *INDID_nonmig* (if the individual is not a migrant). To make the data easier to handle, we create the dummy variable *dummymigrant* which identify whether the individual considered is a migrant (*Yes*) or not (*No*).

Also, we remove the prefix *mig* of variables from the **Migrant** module, in order to append migrant individuals with “new” members. For instance, we rename the variable *migname* in *name*, *migage* in *age*, or *migsex* on *sex*, and append “new” members with migrants.

Thus, in the *NEEMISIS2-tracking* database, one row is one household member, and *dummymigrant* identify whether the individual under consideration is a migrant or not.

4.7.1 Household identification

householdvillageoriginal. Allows the enumerator to filter responses of *householdidoriginal* to select the household ID and *householdmigrantlist* to select all the migrants.

rankingmigrant. Key variable to identify the main migrant, which is also the questionnaire respondent.

othermemberhhmig. Do other members currently share your household? Key variable because it conditions the opening of the module **New members**.

4.7.2 New members characteristics

The objective of this module is to capture the key characteristics of new household members. “New” referred to non-migrant individuals (*INDID_nonmig*).

It consists of a part of questions relative to:

- Socio-demographic characteristics (*age*, *sex*, etc.).
- Education (*everattendedschool*, *everattendedschool*, for instance).
- Gold, loan (*dummygold*, *dummyloan*).
- Employment (*dummyworkedpastyear*, etc.).
- Migration (*dummymigration*, *migrationarea*, etc.).

All these questions rely on the *NEEMISIS2-HH* questions.

4.7.3 Household characteristics

Here, the objective is to capture the household environment, and evaluate the wealth of the household. This module also relies on the **NEEMSIS2-HH** questions.

This module is composed of a part of questions on:

- Land (*dummyownland*, etc.).
- Livestock (*livestocklist*, etc.).
- Expenses (*foodexpenses*, *healthexpenses*, for instance).
- Good details (*numbergoods*, etc.).
- Housing and facilities (*house*, etc.).
- Schemes (*schemeslist*, etc.).

4.7.4 Migration step

migstepmigration. From the beginning of the migration process until today, how many steps did the migration consist of?

For instance, if the migrant first migrated from his home village to Tirupur and then in a second step he migrated from Tirupur to Chennai, then that is 2 steps migration.

migstepmigrationdetails provides details on the migration.

For instance, using the previous example, *migstepmigrationdetails=Tirupur, Chennai*.

Details for each step are provided in **NEEMSIS2-tracking_migpath.dta** database.

4.7.5 Migrant characteristics

This module contains the same questions as module **New members characteristics**, but with the prefix *mig* (*INDID_mig*). However, and as stated before, we remove the prefix *mig* in order to append migrant individuals with “new” members. For instance, we rename the variable *migname* in *name*, *migage* in *age*, or *migsex* on *sex*, and append “new” members with migrants.

Also, the **migrant characteristics** module contains additional information on:

- Education (cost, scholarship, etc.).
- Employment (self-employment, salaried job, advance, etc.).
- Migration (skill, etc.).
- Remittances sent and received (*dummyremittancesent* and *dummyremittancesreceived*).
- Employment module of individual questionnaire (**NEEMSIS2-ego**).

- Social network module of individual questionnaire (NEEMSI2-ego).
- Personality traits and cognitive skills module of individual questionnaire (NEEMSI2-ego).

As for NEEMSI2-ego, the **Social Network** module consists of a name generator. In NEEMSI2-tracking, the following questions open the name generator module:

1. Has someone recommended you to join this association? *migdummyassorecommendation*
2. If you need to borrow tools or machinery for your business, would you ask somebody you are directly in contact with (excluding your household members)? *migsntechnicalhelp*
3. If you have to go away from your business a couple of days, would you entrust your business to someone you are directly in contact with (excluding your household members)? *migsnentrustbusiness*
4. If you need to recruit a worker, would you ask someone you are directly in contact with to help you to find him/her (excluding your household members)? *migsnrecruitworker*
5. If you need to find new suppliers or customers, would you ask someone you are directly in contact with to help you to find them (excluding your household members)? *migsnfindsuppliers*
6. Did you lend tools or machinery within the past 12 months from someone you are directly in contact with (excluding household member)? *migsnlendtools*
7. If you needed to find a new job, is there someone you are directly in contact with you would feel comfortable to ask for help (excluding household members)? *migsnfindjob*
8. In the last 12 months, did you recommend someone you are directly in contact with for a job opening that you knew of (excluding household members)? *migsnrecommendforjob*
9. Was anybody you are directly in contact with hired based on your recommendation (excluding household members)? *migsnrecojobsuccess*

4.8 Tracking occupations

Head questionnaire:	Tracking questionnaire
File:	NEEMISIS2-tracking_occupations.dta
ID:	<i>HHID2022, INDID_mig, INDID_nonmig, occupationid</i>
Unit:	one observation is one occupation

This module gathers information about all occupations during the past year, i.e., the twelve months before the interview date. For each occupation, the interviewees declared the type of occupation (*kindofwork*), an estimated number of days a month (*daysamonth*) and hours a day worked (*hoursaday*) –the number of hours a year calculated (*hoursayear*), an estimated annual income (*annualincome*), and the date they took up the job *datestartoccup*. It includes unpaid workers. Individuals who haven't worked the past year answer questions about the reasons they haven't worked (*reasonnotworkpastyear*, *stoppedworking*).

For migrants (*dummymigrant=Yes*), the entire occupation module is also asked (see section on **NEEMISIS2-occupations** database).

4.9 Alter tracking

Head questionnaire:	Tracking questionnaire
File:	<code>NEEMISIS2-tracking_alters.dta</code>
ID:	<code>HHID2022</code> , <code>INDID_mig</code> , <code>migrationstepid</code> , <code>alterid</code>
Unit:	one observation is one alter

`INDID_mig` corresponds to the same variable in `NEEMISIS2-tracking`. Migrants can mention several “Alters” referred as `alterid`. This database consists of 56 observations, i.e 56 “Alters”. Whenever migrants chose to declare a “New person”, this person is filed with `alterid` and `altername` in `NEEMISIS2-tracking_alter`.

Alters can be mentioned several times across social network modules, hence a new variable we constructed here (`networkpurpose2`) that is more intuitive than `-namelist` variables. For instance, `snsourse` is the network name generator question that have seen Alter mentioned by Ego for the first time. `networkpurpose2` is the second reason why Alter is mentioned.

Name generators (`alterid` and `altername`) are followed by questions called “name interpreters” that gather information on Alters’ attributes such as caste, age, gender, occupation, education, place of residence, economic situation, and about the relationship between the Egos and Alters (when and how they met, how often they interact, their level of intimacy).

All variables in the word file individual questionnaire are name interpreters variables. They have the following suffixes that correspond to following variables in `NEEMISIS2-tracking_alters.dta`:

- name: altername
- relationship: relationship
- sex: sex
- age: age
- castes: caste
- educ: education
- occup: occupation
- occupother: occupother
- living: place
- compared: ecocondition
- duration: durationknown

- meet: howmet
- meetother: howmetother
- meetfrequency: meetfrequency
- invite: egoinvite
- reciprocity1: alterinvite
- intimacy: intimacy

4.10 Migration path

Head questionnaire:	Tracking questionnaire
File:	NEEMISIS2-tracking_migpath.dta
ID:	HHID2022, migrationstepid
Unit:	one observation is one migration step

Each line corresponds to one migration step, identified by the variable *migrationstepid*. *migstepmigration*. From the beginning of the migration process until today, how many steps did the migration consist of?

For instance, if the migrant first migrated from his home village to Tirupur and then in a second step he migrated from Tirupur to Chennai, then that is 2 steps migration.

migstepmigrationdetails provides details on the migration.

For instance, using the previous example, *migstepmigrationdetails=Tirupur, Chennai*.

This module includes questions relatives to the:

- Reason and satisfaction.
- Help.
- Decision.
- Cost.

5 | Fieldwork procedure

5.1 Tasks for enumerators

The enumerator is the one who directly communicates and collects information from households and individuals. The role of enumerators, therefore, is critically important for the data set's quality and integrity. The NEEMSI-2 fieldwork team is composed of 11 members:

- Venkatasubramanian, the fieldwork supervisor;
- Vivek Raja, enumerator and assistant to the fieldwork supervisor as team leader;
- Chithra and Radhika, a pair of enumerators;
- Mayan, an enumerator;
- Pazani, an enumerator;
- Raichal, an enumerator;
- Rajalakshmi, an enumerator;
- Suganya and Malarvizhi, a pair of enumerators;
- Antoni, an enumerator.

The supervisors, Venkatasubramanian, and five enumerators took part to the NEEMSI-1 wave and RUME survey data collection so they already have an excellent knowledge of the fieldwork. Before conducting an interview, enumerators need to verify the list of households interviewed in 2016-17, their address, caste and basic characteristics in order to make sure they interview the same households. In the identification module of the questionnaire, the tablet asks for the village name and the village area (“Colony”/“Ur”) where the interview takes place. According to that preliminary selection, the list of households surveyed in 2016-17, with their 2016-17 household identifier displays. Enumerators then can start meeting available households in the list. The **General information** module allows to check if the household characteristics match with 2016-17 information. If enumerators meet a new household, i.e outside the 2016-17 list, he can

choose the “new household” option in the household ID list, and then fill the new household information in the **General Information** module.

Enumerators sometimes have to meet a household several times either to complete the household questionnaire, ego 1 individual questionnaire or to meet the second or third household member (ego 2, 3) depending on their availability. Enumerators, therefore, have to actively plan for appointment confirmation with each household and egos, who should be interviewed separately, and make use of time in communicating with them in order to meet the fieldwork schedule and fulfil the number of questionnaires assigned.

The eight enumerators working on NEEMSI-2 wave were trained on the questionnaire and the tablet before the beginning of the pilot survey and the research team regularly planned meeting with them in order to improve the questionnaire and clarify questions raised by the enumerators during fieldwork and thus to collect high quality data.

5.1.1 Working attitude

- Enumerators should maintain polite attitude with everyone (the respondents, their family and their friends, the team leader, the survey team members and other involved people). Interviewer’s behaviour greatly influences local people’s opinion as well as the survey results.
- Enumerators should be well dressed (neat, polite), showing professional style. Dressing style recommended is:
 - Male: shirt, trousers.
 - Avoid fancy dress, creating distance from the interviewees.
 - No alcohol/beer or smoking during work.
- Enumerators should create a friendly atmosphere to encourage communication and information sharing from the respondent. Enumerator, at the mean time, still has to comply with interview procedure and questionnaire contents.
- During the interview, the enumerator should not make any judgment, comment on respondent’s answers or behaviours that might affect respondent’s psychology and impose enumerators self opinions.

5.1.2 Working schedule

- In the initial stage of the survey, enumerators are divided into 3 teams. Each team consists of 2 members, including at least one of them who already has fieldwork experience (5 out of 6 enumerators were already enumerators for the 2016-17 wave). Teams of two members enable them to get used to the survey tool: with one tablet per team, one can focus more on

the technical part and the other one can keep track of the interview. After the initial stage, enumerators will start to go alone to interview households, i.e., they will have one tablet each.

- The team and then enumerators alone will complete one household plus the three individual questionnaires (when individuals are available) within a day or two days.
- Once all 2016-17 panel households in the 10 villages are completed, the enumerators will start tracking migrant households and to complete the around 100 new households in the 10 villages.
- Enumerators, after receiving the survey plan, should understand the enumeration areas and the list of households, and should meet and interview households or establish contacts prior to the interview to ensure to meet targeted respondents and obtain the best quality of interviews.
- Enumerators should strictly follow working procedures and schedule assigned. Interview and data input should follow the process taught in training courses. Enumerators should ensure no overlap or missed interviews of any respondent and no missing or wrong information recorded on the tablet.
- Enumerators are requested to finalise questionnaires every day on the tablet and to do fieldwork report to the coordinating team twice a week. The coordinating team can then send and save the finalised questionnaires on the Survey CTO server.

Important notice:

- Remember to “save” the file and data whenever “Save the form” note appears on the tablet and after completing the questionnaire to avoid data loss.
- Take note carefully during interviews if there are additional elements not included in the questionnaire.
- Check completed questionnaires, ensure the information’s integrity before finalising it.
- Maintain and preserve survey tools and Tablet PC, avoiding loss or misplacement.
- Report any incident and make a request to coordinating the team, especially in situations that the enumerator cannot solve himself/herself.
- Collect valid invoices and documents as required.

5.1.3 Interview principles

Enumerators should carefully follow instruction in this manual, details as follows:

1. **Ask the question exactly as stated in the questionnaire.** The questions were carefully designed to obtain the correct information for later analysis, pre-tested many times at the enumeration area and carefully translated in Tamil on the tablet so that there is no need for in live translation on the field. Enumerators should read the full text of the question. After the first clearly reading out loud, enumerator should wait for an answer. If the respondent does not provide an answer within a certain period of time, they might:
 - not hear the question clearly;
 - not understand the question;
 - do not know how to answer.

For all cases, an enumerator should repeat the question. If the respondent is still not able to answer the question, enumerator should ask whether respondent understands the question or not. If not, the enumerator should explain the question's content in another way. In principle, you will almost never read the list of the coded answers to the respondent even if he or she has difficulties to answer because very often he/she will agree with your proposal. Nevertheless, you must do it in some cases that this handbook will specify (for some "questions" ending with "...", the possible answers have to be enumerated as examples).

2. **Trying not to receive *I don't know* answers** by helping the respondent to estimate or answer or find the answer most closely resembling their point. In case figures are required (income, value of purchase, etc.) and the respondent seems to be reluctant to answer, ask for a range of value (about how much?), but without suggesting any amount. In case basic information is known, fill in the right information into the blank. If information is unclear or forecasted, enumerator should ask the respondent.
3. **If there is really no answer**, the enumerator can choose *No response* option or 99 if numerical variable (when don't want to give any answer), *Don't know* or 88 (when don't know a specific information about a household member for instance), or *Not applicable* or 66 (appears rarely because irrelevant questions are usually automatically skipped). The distinction between these three options is crucial for data analysis later.
4. **Maintain the pace of interview.** Enumerator has to administer the interview, yet listen carefully to respondent, and avoid displeasing the respondent. To do so, enumerators should refrain from lengthy discussion with the respondent; if respondents answer is inappropriate or too complex, enumerators should not interrupt suddenly but cleverly get respondent back towards the original question. Enumerators should avoid suggesting answer to respondent.

5. **Keep a completely neutral attitude regarding the topic of the interview.** Enumerators should not express their surprise, agreement or disagreement with the given answers. If the respondent seeks for an answer, the enumerator should not reveal his/her opinion, only provide an explanation of interview purpose is to gather the opinions of the respondents on this issue. Enumerators should not discuss his/her views with the respondent or hint an answer until the interview is completed.
6. **If you don't understand any question or process,** read this manual first, then ask for more detail if necessary. The enumerator should inform the coordinating team about these concerns who will gather this information in the report.
7. **Make sure that the respondent would not think that you are an inspector** (local tax authority, customs organisation. . .). A good way to make the interviewee more confident could be: “the information collected here is strictly confidential, will be used only for research purpose and won't be transmitted to government administration, according to the law of confidentiality of statistical surveys”. Indeed, we ask questions about forms of labour, indebtedness, which can be very sensitive for households, that's why he/she needs to be sure that it is not for police control or finance companies.
8. **Interview's Confidentiality.** All collected data is kept confidential. Any of information revealed to un-incharged person is considered as violation to interview's regulation. This principle is very important and fundamental to all statistical work. In principle, the presence of a third person can cause embarrassment and affect answers, confidentiality of information. For that reason, the interview should preferably take place inside the house. However, it is often difficult to prevent the presence of a third person during the interview. In such cases, the enumerator should suggest the respondent to persuade the third person to go elsewhere; or explain the importance of the privacy and the confidentiality of the interview in a proper manner. For individual questionnaires particularly, it is better to be alone (as much as possible) with the respondent because of personality and behaviour questions and cognitive tests (RAVEN's test). On the contrary, for the household questionnaire, the presence of other household members can be profitable since the enumerators can obtain more precise information about all members.
9. **Answer sincerely to the questions of the respondent.** Before accepting to be interviewed, the respondent can ask you some questions about the survey or the way he/she was selected to be interviewed. Be direct and friendly when you are answering. The respondent can also be worried by the duration of the survey. Tell him/her about the average duration of NEEMIS-2 interview, which is about half a day to one day for household plus the three individual questionnaires, and do not try to get it shorter. Always have on you your identity cards and your badge from the French Institute of Pondicherry (IFP) and do not hesitate to present on a simple request of your interlocutors.

5.2 Tasks for the team leader

Out of 8 enumerators, one is considered as a team leader, Vivek Raja, responsible for managing activities of teams. The team leader, therefore, besides acts as an enumerator, should have leading and communicating skills with other team members. This team leader works closely with the fieldwork supervisor, Venkatasubramanian.

5.2.1 Contact the enumeration area in advance

The team leader is responsible for contacting households, and in some cases head of villages or village administrative officers (VAO) to confirm the list of households to be interviewed and make working appointments for the team.

There are cases where the household address is not found, nobody at home, household relocated, new household residents in the address. The team leader should ask their contacts to households' neighbours or relatives and track these households if possible. Otherwise, fieldwork team leaders should contact with coordinating teams for final decision on the sample.

Notice for team leaders to communicate with the VAO:

- **Survey's purpose:** Collecting updated data in each of the 10 villages, so as to inform for policy-makers to formulate policy adjustments about labour and more generally rural livelihoods.
- **Team survey composition:** Number of members, tasks assigned.
- **Schedule:** Number of panel households to interviews, including migrants, plus around 10 new households by village (around 20 days by village over 6 months).
- **Survey procedure:** Interview duration is from 3 to 4 hours for the household questionnaire and around 2 hours for individual questionnaires depending on households.
- Announce survey schedule to the VAO and request for cooperation.

5.2.2 Overseeing enumeration area

- The team leader plays the role of enumeration area manager, closely communicates and works with enumerators to prevent possible errors in the early days, encourage the team to work at their best.
- Team leader should summarise, share the working experience among the team and report to coordinating team.
- Re-check in detail all filled questionnaire forms for data fulfilment and integrity.

- For any missing or errors found, team leader and coordinating team should discuss with enumerators to find solutions. The team has to come back to household if necessary ■