

Colombian Experience in the Quality Assurance of new sources

Juan Daniel Oviedo
Director - DANE

February 9th, 2021
52nd Statistical Commission



El futuro
es de todos

Gobierno
de Colombia



Agenda

1

Income poverty measurement: due to social distancing measurements implemented in Colombia during 2020, it was necessary to adapt the main face-to-face household survey to a phone survey between March and July, in which the questionnaire was reduced. For this reason, a complementary phone survey was applied in November and December to recover income data for the period March-July, 2020.

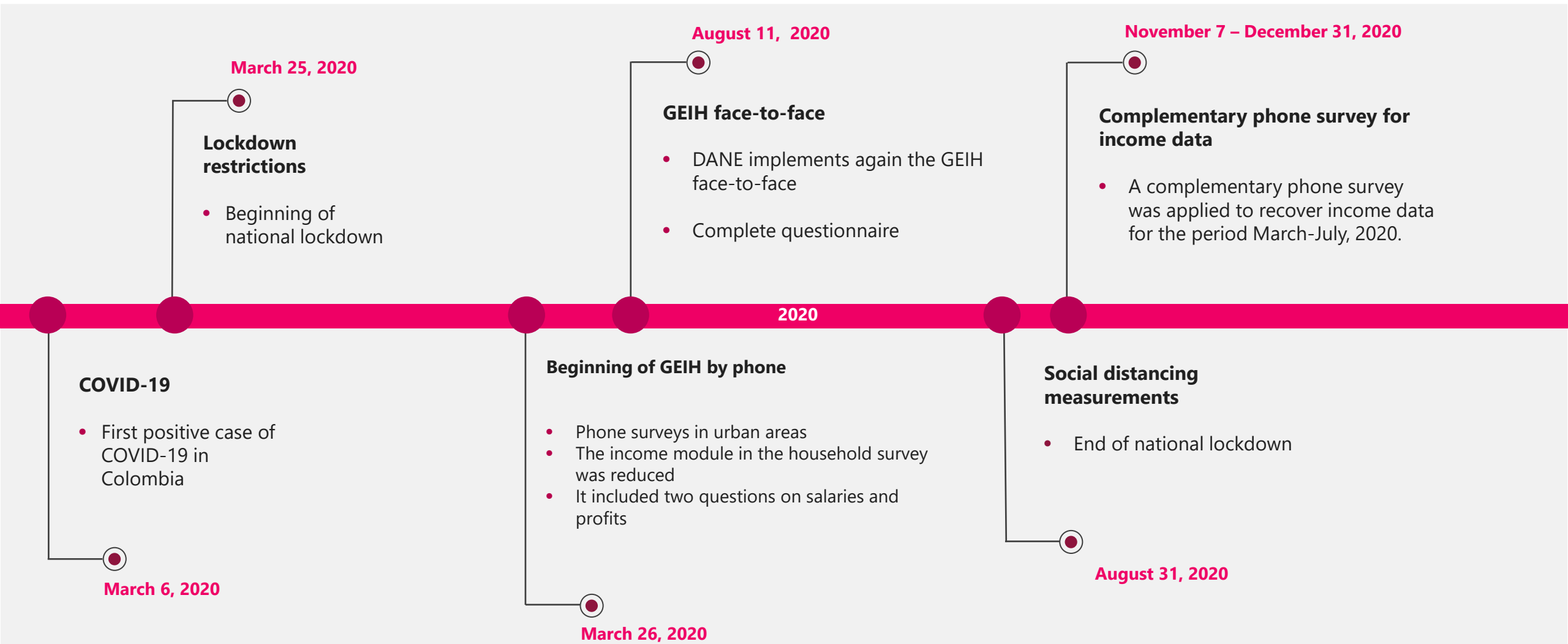
2

Consistency checks of income data to measure poverty in 2020: integrating the Household Survey (GEIH) with Administrative Records, in order to validate consistency of the data collected in November and December.



Timeline

Great Integrated Household Survey (GEIH) 2020





Income data for poverty measurement

Context:

Due to the social distancing measurements implemented associated with the pandemic, it was necessary to adapt the main face-to-face household survey (GEIH) to a phone survey between March and July. While the original questionnaire had to be reduced, we included questions on salaries and profits of self-employees, but information on other income sources were not collected during the period mentioned above.

For this reason, in November and December, DANE implemented a complementary phone survey to collect the household income data that was not previously collected between March and July. This data may have a report bias since the information was collected 5 months later. In order to address this, we intend to use Administrative Records such as the government allowances (public transfer programs) and social security register (PILA).



Income data for poverty measurement

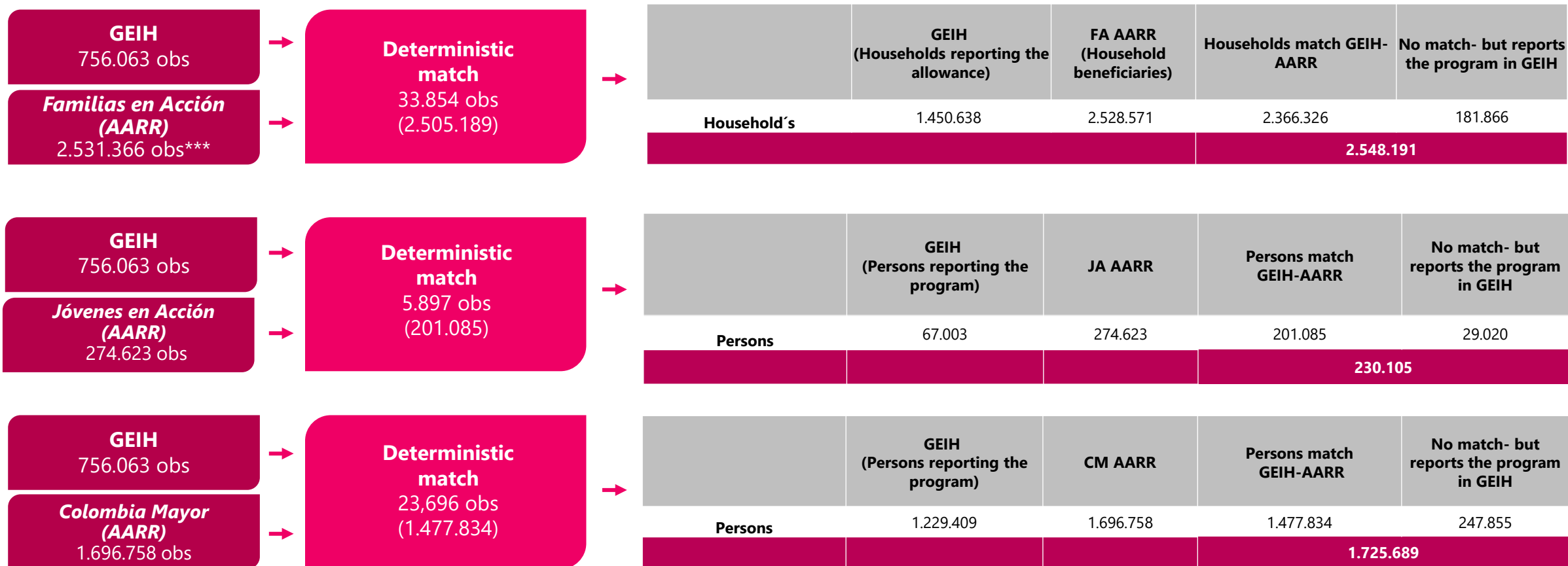
Challenges:

- **Integrability – unique keys:** The personal identification number (cédula) is not a mandatory question for respondents in the GEIH. Thus, it is troublesome to match the GEIH with Admin Records. In addition, statistical operations and administrative records do not have unique identifiers. For example, in Colombia the identification number changes as the person reaches the legal age. To increase the matching rate GEIH-AARR we use record linkage methods, using variables such as names, surnames and birth dates
- **Validity quality output**
 - **Accuracy:** : It is necessary to harmonize the information since the administrative records do not have statistical purposes, i.e standardizing the data sets and including the statistical unit
 - **Differences in the reference period:** The GEIH is a monthly operation while the Admin Records (i.e., government allowances) have a quarterly or annual reference period.
 - **Completeness:** In the GEIH we have a non-response issue. For instance, in the government allowances we need to carry out coverage assessments to define a correct imputation method.



Preliminary accuracy assessment - Government Allowances (Administrative Records-GEIH)

Deterministic match* (2019)





Government Allowances and Social Security Register (Administrative Records-GEIH)

Next steps:

- **Quality framework for administrative records:** DANE is designing a quality framework for administrative records to measure indicators such as coverage and integrability
- **Integrating administrative records of 2020 with the GEIH** by combining government allowances records and the social security register with the GEIH.
- **Imputation method:** checking the consistency of the information, the imputation method and the validity of outputs.

Colombian Experience in the Quality Assurance of new sources

Juan Daniel Oviedo
Director - DANE

February 9th, 2021
52nd Statistical Commission



El futuro
es de todos

Gobierno
de Colombia