Working paper 3
7 November 2018

UNITED NATIONS
ECONOMIC COMMISSION FOR EUROPE
CONFERENCE OF EUROPEAN STATISTICIANS

Expert meeting on measuring poverty and inequality
29-30 November 2018, Vienna, Austria
Session B: Improving response rate and sampling precision in surveys

Improving response rate and accuracy of the National Household Income and Expenditure Survey in Mexico

Prepared by Edgar Vielma Orozco, INEGI, Mexico

Abstract

In Mexico, the improvement of the response and the accuracy of the National Household Income and Expenditure Survey is approached from a conceptual, statistical and operational perspective. From the conceptual point of view, the way in which the measurement of income and expenditures is addressed gives greater elements to have more precise estimates. In the case of statistical design, the survey contemplates different elements that improve the accuracy of the estimates, such as the sampling frame, its stratification, the sample size, as well as the consideration of non-response rates differentiated by study domain. Operationally, there is an automated process of verification of households, where again the reported income must be reviewed and justified, which has allowed a more accurate capture of the phenomenon.

Introduction

Income surveys, both nationally and internationally, have historically detected that there is a sub-report of income on the part of informants. For the National Statistical Offices (NSO), it is a growing challenge to collect information on household income and expenditures. Mexico is not the exception, either due to unsafe environments as a result of the crime, for fear that the income information will be linked to a treasury authority or to the loss of stimuli from social programs. The above causes that the informants are more and more cautious in the information they provide.

The NSO from its creation mandates are entrusted with providing reliable and truthful information, so they work continuously to ensure that the measurements reflect better the reality of household income.

In this context, the experiences that are collected in the preceding exercises are fundamental, since, based on this information, innovation continues to be obtained in a better capture of the different variables, including income, without this implying a modification to the design conceptual and statistical that supports the project.

In the case of Mexico, once each statistical project of the Income and Expenditure National Household Survey (ENIGH) is completed, the National Institute of Statistics and Geography (INEGI) is given the task of reviewing operational, methodological and of information processing, with the purpose of taking pertinent measures that achieve that the next survey has better quality, as part of a process of continuous improvement. The aforementioned affect the improvement of the training process of the field personnel, thus improving the response rates and the quality with which the information is captured.

The fact of seeking to improve the effectiveness of the interviewers and reducing the non-responses by refusals is one of the aspects that INEGI considers imperative to solve for the lifting operation of each ENIGH.

Other critical opportunity areas that have been detected with the passage of time are:

- Persistence of a sample number of households with zero income or income not consistent with the socio-demographic and economic characteristics of the household members and their households, both in low and high strata and that these households did not have adequate justification. The previous situation reached an average of 2 448 households (3.5%) in the four surveys registered between 2008 and 2014. This type of response has a greater damage to the income measurement than the TNR (non-response rate) due to rejection, because it generates a measurement that moves away of the reality of households, causing an underestimation of income.

- Detection of inconsistencies in the information of some variables, such as age, kinship with the head of the household, the condition of economic activity and sources of income captured in the survey questionnaires.

In order to take advantage of the opportunity areas mentioned above, the need has been considered to strengthen the technical capacity of the operational staff to carry out the interview, as well as their conceptual training to better capture the information provided in the homes; the above without incurring changes in the methodology applied in the survey.

The above, since, if in an exercise the number of inconsistent zero were doubled or tripled due to field problems, the income estimate would collapse, without this having been a consequence of any methodological change. In this logic, the effort to bring down the zero represents an effort to maintain the quality of an instrument.

Therefore, in order to guarantee statistical quality in the collection of information in each of the ENIGH exercises, INEGI has established measures to strengthen the capacities of its operational staff:

- Development of a section in the *Interviewer's Manual* that provides instructions that allow interviewers to be more analytical, to encourage more consistent information gathering among the different questionnaires, including household income.
- Based on the foregoing, during the training, emphasize operational staff (especially interviewers and supervisors) on the inconsistencies detected in the information on variables such as age and kinship; descriptions of items such as "Others" in job seekers and people who are not economically active; description of the item "Other income not previously considered" and description of "Other social benefits at work", and from this, make a series of recommendations to avoid inconsistencies. Likewise, to emphasize to the State Supervising Instructor (ISE) that their functions also consist of detecting said incongruities and seeking their resolution.
- Incorporation of a validation criterion to review the cases of households that do not report income during the reference period or when the reported income is not consistent with the socio-demographic and economic characteristics of the household members and their dwellings, when the questionnaires do not there is an appropriate justification with the rest of the information.
- Correct identification and control of errors in capturing the information obtained from households, through the introduction in the Project Monitoring and Control System (IKTAN) of a report for the operative figure of the State Supervising Instructor (ISE), through the Identify interviewers who require closer supervision and retraining. Additionally, the implementation of a process of review of captured folios, by means of which the operational

figure Responsible for Capture and Validation (RCV) can identify the folios with capture errors that must be sent to recapture, as well as the capturists who commit a greater number of errors and that require more precise supervision.

From 2008 to 2014, the income and expenditure survey consisted of the application of the ENIGH instrument to a part of the sample, based on six questionnaires, while, for the rest, the instrument Socioeconomic Conditions Module (MCS) was applied, through four of the six questionnaires: Households and dwellings, Persons 12 years old and over, People under 12 years old and family businesses of the home. The MCS instrument is a subset of the ENIGH since the MCS measures only income and the ENIGH, income and, expenses.

The implementation of the ENIGH and the MCS separately, have a differentiated impact on the measurement of income, mainly in the non-monetary part. As mentioned, the ENIGH is a survey that investigates income and expenditures in households, unlike the MCS that only focuses on income. The hypothesis has arisen that spending measurement allows households to make a more accurate statement of income since simultaneously measuring both aspects allows them to make a better balance of income and expenses in the household, that is, the income and expenditure are measurements that by their nature complement each other.

The difference at the national level between the projects is not conclusive, that is, the combination of instruments allows estimating the same income levels at the national level, although in the case of the MCS subset at a lower cost than an ENIGH type survey. However, the difference is conclusive for the first five deciles in favour of the ENIGH.

Average quarterly current income by deciles for 2014

Decile	ENIGH*	Standard error	MCS-ENIGH	Standard error	MCS*	Standard error	Difference ENIGH vs MCS-ENIGH	Difference MCS-ENIGH vs MCS	Difference ENIGH vs MCS
National	**42 078**	**1 093**	**42 262**	**545**	**42 342**	**662**	**No**	**No**	**No**
I	7 286	115	6 266	73	5 797	86	Yes	Yes	Yes
II	12 728	148	11 723	98	11 142	122	Yes	Yes	Yes
III	16 991	166	16 073	111	15 480	139	Yes	Yes	Yes
IV	21 259	200	20 366	129	19 771	157	Yes	Yes	Yes
V	25 860	225	25 298	155	24 872	197	Yes	No	Yes
VI	31 251	283	31 089	194	30 919	232	No	No	No
VII	38 191	364	38 572	265	38 672	312	No	No	No
VIII	48 240	539	49 579	373	50 128	437	Yes	No	No
IX	66 471	967	68 524	589	69 364	679	No	No	No
X	152 504	9 219	155 135	4 692	157 273	4 910	No	No	No

Source: INEGI. Income and Expenditure National Household Survey (ENIGH). Mexico.

These evidence show that the MCS, when not investigating the expenditure in the households, captures less income in the lower part of the distribution, especially in the non-monetary income items. In an analysis by income categories, it is observed that between 60 and 70% remains in trend with previous years.

Therefore, it was shown that there are indications that the collection of income without the collection of expenditure generates that the amount of income is the result only from the perception of the informant.

Although some items of non-monetary income are differentiated by the instrument change, both national income and poverty indicators mostly remain without significant changes due to the difference in instruments.

On the other hand, from the statistical point of view, the INEGI adopts a series of measures that allow to give greater precision to the estimations of income and expenses, among which is the sampling frame, the stratification of the sample, the size of shows, the differentiation of non-response rates by federal entity and urban and rural domain. As for the sample, it has a design of more than 80 thousand houses with urban and rural state representation, which allows having a better statistical and conceptual measurement.

Next, the improvement of the Income and Expenditure National Household Survey is addressed through three sections, which differentiate perspectives from the point of view: conceptual, statistical and operational.

I. Improvement in conceptual terms

The joint measurement of income and expenditure aims to understand the essential economic problems that households present to meet their basic needs. The information presented allows its users to generate indicators and fundamental comparisons of the approximation of household disposable income, maximize the usefulness of the results, a high degree of detail, as well as a reliable measurement of income, consumption and global expenses.

INEGI in the household issue carries out the Income and Expenditure National Household Survey (ENIGH), which aims to provide a statistical overview of the behavior of household income and expenditure in terms of amount, origin and distribution, It also captures information on the occupational and socio-demographic characteristics of household members, as well as the characteristics of housing infrastructure and household equipment.

The methodology of this survey is based on international recommendations such as:

- The resolutions and reports of the 17 International Conferences on labour statistics, of the International Labour Organization (ILO).

- The final report and recommendations of the Canberra Group, a group of experts on "Household Income Statistics."
- Handbook of Household Surveys. Department of International Economic and Social Affairs, Office of Statistics. United Nations, New York, 1987.

Likewise, it is articulated with the System of National Accounts and with the Household Surveys that raise the INEGI and the information requirements of the different users, always taking care of the historical comparability.

The survey has a biennial periodicity (every two years).

The unit of selection is private housing, these homes are chosen through a meticulous statistical process that ensures that the results obtained from one (sample) can be generalized to the total.

The unit of observation is, therefore, the home and in turn, the unit of analysis will consequently be the home, the dwelling and the members of the household.

Current Income

There are several factors that affect the conformation of the level of welfare of households. The total current income in cash or in goods and services received by the members of the household is one of the most relevant, since the flow and amount of these incomes depend to a large extent on regular access by the household, to the goods and services that constitute their consumption basket in a given period.

The current income of households is formed by monetary and non-monetary entries that satisfy these three criteria:

Inclusion criteria

1. Regularity: they must be regular, that is, subject to an uninterrupted occurrence in a given period.
2. Availability: they must contribute to the current economic well-being, that is, they must be available for the acquisition of household consumption goods and services.

Exclusion criterion

3. Patrimony: these entries should not include the flows that modify the net worth or the net value of the assets or liabilities of the household.

The current income then establishes the maximum resources that the household can access on a regular basis and that are available for consumption of goods and services in a given period, maintaining the home's patrimony intact.

For the ENIGH, each of the items in which current income is presented contains both monetary and non-monetary elements, which in previous events were presented separately.

The current income in the ENIGH is disincorporated into five categories:

- Income from work;
- Rent of the property;
- Transfers;
- Estimation of the rent of the house; and
- Other current income.

Monetary current expenditure

Household spending is the expenditure that a household has to make to meet their needs and meet their commitments. The current expenditure accounts for the expenses regularly incurred in the home for the acquisition of its consumption basket, in addition to the regular expenditure not destined for consumption. Under this concept, the expense that modifies the patrimony of the household or one of any of the members of this one.

Consumption expenditure is measured taking into account the purchase value of goods and services; thus, the reported expense corresponds to the value of the goods acquired, regardless of whether they were paid or not in the reference period.

Currently (ENIGH 2016) in the expenditure section only monetary current expenditure is presented, since non-monetary current income is no longer presented separately, but is included in each of the different income items. The monetary current expenditure is, then, the sum of the regular expenses that households directly make in goods and services for their consumption.

Due to its composition, the monetary current expenditure of households is grouped into nine categories:

- Food, beverages, and tobacco;
- Clothing and footwear;
- Housing and conservation services; electric power and fuels;
- Goods and services for cleaning and care of the house; glassware, white goods, and household utensils; household goods and furniture;
- Healthcare;
- Transportation; acquisition, maintenance, accessories and services for vehicles; and communications;
- Education and entertainment items and services; tour packages and for parties, lodging and accommodation;

- Personal care; accessories and personal effects; and other miscellaneous expenses;
- Spending transfers.

Joint measurement of income and expenditure

Measuring household income and expenditure in the way that ENIGH does it allows:

- Evaluate the levels and trends of economic well-being, representing the total economic capacity of the household, that is, current and capital transfers, and the net balance of assets and liabilities.
- It allows generating income distributions and consumption expenditures among households, which allows the measurement of poverty.
- Study the population groups that are at the lowest level of that distribution (poverty) or measure their dispersion.
- Comparison of the measurement of poverty calculated by income and expenses.
- The impact of the prices of the basic basket on household consumption and nutrition patterns and their expenses.
- To have information on consumption expenditures that are the source for the preparation of the CPI (Consumer Price Index), which allows the determination of goods and services for the establishment of the basic basket, and thereby measure the impact of this in the income of households.
- Analysis of the economic conditions of the population with some disability.
- Analysis of households with access to certain social programs, such as for people with disabilities; agricultural day laborers; senior citizens, temporary employment, community canteens, nurseries, marketing and market development for businesses, among others and their consumption patterns.
- Analysis from a gender perspective in the aforementioned indicators.
- Analysis of redistributive public policies based on transfers of social programs received by households.
- All the previous analyses can be done characterizing both the housing and the residents of it.

In addition to the collection of income and expenditure together there are a number of topics of interest in the project, such as health and use of time, which allow analysis under the gender approach, among others.

Collection of information through questionnaires

The different collection instruments used in the collection of household information are designed with the purpose of covering the different objectives of the ENIGH, by considering issues such as:

- Characteristics of the house;
- Residents and identification of households in the dwelling;
- Socio-demographic characteristics of the residents of the dwelling;
- Household equipment, services;
- The condition of activity and occupational characteristics of household members aged 12 and over;
- Current income (monetary and non-monetary) of households;
- Monetary current expenditure of households;
- Financial and capital perceptions of households and their members;
- Financial and capital expenditures of households.

All this subject is covered by the following questionnaires:

- Household and dwelling questionnaire;
- Questionnaire for individuals 12 years and older;
- Questionnaire for family businesses;
- Questionnaire for individuals younger than 12 years;
- f Daily expenditures booklet;
- Questionnaire on household expenditures.

Due to its specific design and its theme, the questionnaires provide advantages such as:

- They allow the collection of detailed information following the international recommendations for the collection of income and expenses.
- The informant is less affected when viewing small questionnaires instead of a questionnaire with a greater number of pages with the whole subject.
- They facilitate the identification of the appropriate informants by applying only questions corresponding to the different population groups according to their profile.
- They allow to apply all the corresponding questions to the informant in a single interview, thus avoiding to bother him constantly during the survey period.
- They facilitate the capture of information for interviewers by being able to follow sequences that should be reflected in the "passes" for the corresponding questions.
- There are greater control, supervision, and verification of the interview, being able to review complete questionnaires daily.
- It facilitates the recall of the informants having specific periods according to the frequency of acquisition of the products or services.

Due to the characteristic of the concepts to be captured, it may be necessary to use different time references or reference periods. In the questionnaires, they vary according to each section, section or question of the questionnaire.

For the collection of income for example are the salaries and regular entries (monthly, weekly or daily) and annual for some concepts (bonus and profit sharing), for the expenses the reference periods are monthly, quarterly, half-yearly and annual in the latter are those of agricultural type both income and expenditure due to their specific characteristics. For example, the monthly reference period in household expenses is used to record payments for services such as housing rental, conservation services, water, electricity, gas, education, culture, and recreation; the quarterly period is applied to expenditures for the acquisition of clothing, footwear, and healthcare; Half-yearly expenses are the least frequent and usually high, such as the acquisition of household goods, furniture, and maintenance of the home. And finally there are the daily expenses which are captured in a booklet for the best remembrance of the informant, and the expenses related to food, beverages, and tobacco are recorded; and transportation taking the last 7 days prior to the interview as a reference.

With the data obtained from ENIGH, it is possible to obtain breakdowns for specific groups of the population at the national and subnational levels:

- population under 12 years old,
- population from 12 to 29 years old,
- population from 30 to 60 years old,
- population of 60 years and over,
- population with physical or mental difficulties,
- indigenous population,
- population by size of locality,
- indigenous language speaking population,
- illiterate population,
- older adults without access to jubilations and/or pensions,
- population that subsists with transfers,
- employed population without access to social security.

II. Improvement in statistical terms

The sampling frame for the National Household Income and Expenditure Survey (ENIGH) is based on the information collected by the 2010 Population and Housing Census. With this information, a sampling frame updated at time t_0 was created, forming groups of dwellings with certain criteria of closeness and compactness, trying to include a homogeneous number of dwellings within each of these groups or clusters, with the restriction to include whole blocks. Clusters are named PSU (Primary Sampling Unit), as they are selected in a first sampling stage

All the clusters were classified into one of four groups at a national level, using multivariate stratification techniques with socio-demographic indicators at cluster level that allowed measuring the degree of marginalization of the population, with respect to household equipment, basic services housing, education, health, among others. Additionally, to the group or socio-demographic stratum that is assigned to each PSU, a geographic stratum is also identified for each one of them, considering the state, and the number of inhabitants of the city (town).

These four groups or socio-demographic strata sere then combined with the geographic strata to form the survey design strata. This classification of the units of the population is used when designing our probabilistic samples making sure they are probabilistic and representative surveys. At the same time, this stratification makes it possible to generate more precise estimates since, within these strata, the income and expenditure phenomena tend to be more homogeneous compared to a simple random sample, with the consequence of estimating smaller sampling errors.

Based on information from the 2010 Population and Housing Census, 245,279 PSU were formed throughout the national territory, from which 22,477 were selected with probability proportional to the number of dwellings in the PSU, to be part of the Master Sample Frame (MSF) which allows the design of each household survey. The clusters selected receive a continuous treatment which consists in updating the statistical indicators of those PSU.

The updating process of the MSF is a field operation that continuously measures statistical data like the number of dwellings inside the PSU (new constructed dwellings, demolition, use other than housing, etc.), the type of dwelling (inhabited, uninhabited or temporary use), the number of members of the dwelling, as well as the geographical changes of the PSU. With this updating process, it is guaranteed that the inclusion probabilities show the dynamics of the national population since the 2010 Population and Housing Census. From the set of PSU that are continuously updated, a subset of them is selected for the ENIGH survey. Inside those selected PSU, the dwellings will be chosen in a new sampling stage.

The sample size of the most recent ENIGH has increased by area (urban, rural) at the state level, considering a minimum sample size required to estimate the quarterly current income per household average, quarterly income from work per household average, and the quarterly expenditure per household average based on information from the previous edition of the ENIGH. The national sample size of the 2016 ENIGH was 80,328 dwelling, distributed by 53,079 dwelling in urban areas and 27,252 dwellings in rural areas. For the 2018 ENIGH, the number are 87,828 dwellings distributed as 56,052 and 31,776 dwellings, in urban and rural areas respectively.

A differentiated non-response rate has been applied by state and area (urban, rural), starting from the non-response rate of the previous edition where we had a unique non-response rate of 15%, and considering differences in the expected non response rate for the 2018 version:

- 20% non-response rate:

Urban area: Campeche, Coahuila de Zaragoza, Colima, Chihuahua, Durango, Guerrero, Hidalgo, Mexico, Puebla, Quintana Roo, Tabasco, Tamaulipas, Tlaxcala and Veracruz de Ignacio de la Llave.

Rural area: Chihuahua

- 25% non-response rate:

Urban area: Jalisco, Nuevo Leon, Queretaro, San Luis Potosí and Yucatán

Rural area: Mexico City.

- The rest of the entities and areas remain at 15 percent.

The formula for the sample size adjusted by differentiated non response by domain and state is:

$$n'_{ed} = \left(\frac{1}{1 - TNR_{ed}^{ENIGH\text{-}2016}} \right) n_{ed}$$

where:

n'_{ed}	=	sample size adjusted for non-response in d-domain of the e-state for the ENIGH 2018.
n_{ed}	=	sample size calculated for the d-domain of the e-state
$TNR_{ed}^{ENIGH\text{-}2016}$	=	non response rate reported by the 2016 ENIGH for the d-domain of the e-state.

The probability weights of the sample elements are finally adjusted for the missing units, making an adjustment at the design stratum level.

III. Improvement in terms of field operation

The implementation of a validation criterion that reviews those incomes below the biannual threshold, arises due to the persistence of a sample number of households with zero income or income not consistent with the socio-demographic and economic characteristics of the household members and their homes, both in low and high strata, in addition to the fact that these households did not have adequate justification (as has always been requested), this situation causes a measurement that is far from the reality of households and therefore an underestimation of income.

It refers to a validation criterion, previously mentioned, for the review of all those incomes below the established semi-annual threshold of $ 6,120.00 biannual [1] pesos, that is, that considers households in a precarious situation.

To obtain households below the threshold, both monetary and non-monetary income of households captured in the ENIGH are taken into account, so the following items are considered:

- Gifts received.
- Remuneration in kind.
- Transfers.
- Household income.
- Income from non-agricultural businesses.
- Income from agricultural businesses.

The way to proceed for the aforementioned criterion is through the computer platform of the project, identifying the homes listed, then it must be verified that in the observations part of any of the questionnaires that make up the ENIGH there is some justification that supports the situation of the home regarding your income. In the absence of such observation, it is requested to return the questionnaire to the field in order to obtain a justification or, if applicable, to retrieve missing information.

When the capture of the observations is made, a review of each justification is carried out where three probable situations arise:

- The first of them, the characteristics of the dwelling, estimated amount of the dwelling, payment of rent or payment for housing, the composition of the household, the number of members that made up the home, attendance to the school, type of school are reviewed, level at which he attends, food conditions, household equipment, household expenses, and both monetary and non-monetary income received. If it is observed that it is a home with precarious conditions and the captured observation of that home agrees with each other, the questionnaire is accepted.
- The second, that in the justifications it is mentioned that new revenues were received. For these cases, we proceed to reject the questionnaire so that the entity performs the process of capture and validation of the information.

[1] Number established based on an internal analysis of the workloads of the various operational figures, of the aforementioned projects, which can review the information gathered in the field.

- Finally, for cases in which the analysis mentioned in the first situation is inconsistent with the observation captured, or cases such as observations that are not sufficiently substantiated, housing and household conditions are not precarious, negative on the part of the informant to give more information, the inability to contact the informant to verify the situation, are situations where a questionnaire can be rejected resulting in incomplete interview.

The application of this criterion led to a higher quality and accuracy of the information by considerably reducing the inconsistencies generated in the information, another important impact caused fuel change in the mentality of the operational figures that led to avoid an excessive burden on field returns.