# SURVEYS OF INFORMAL ECONOMIC UNITS[1]

Adriana Mata Greenwood

Department of Statistics, International Labour Office

## Introduction

The methodology presented in this paper is intended for countries with a recognizable informal economy. Conventional establishment surveys cannot cover enterprises which are informal and therefore in these countries, they will not be representative of the economy as a whole, but only of part of it. Even in countries with a small informal economy, conventional establishment surveys may not be confident to cover small enterprises, which constitute the majority of enterprises in all countries of the world.

This is because conventional establishment surveys, whether they obtain information about production and costs, or about employment, hours and earnings or labour costs, select their samples using a directory of establishments (or enterprises). This directory is constructed on the basis of registers kept by government agencies for different administrative purposes, or on the basis of the Economic Census.

When the registers are exhaustive and maintained up to date on a regular basis, the directory will be complete and samples selected from it can be considered representative of all establishments (or enterprises) in the country. Government registers will not cover enterprises in the informal economy, however, so in countries where they are substantial, registers will be incomplete. Even when informal enterprises are not common in a country, registers may not be exhaustive nor maintained up to date regularly. Often, they leave out small enterprise which tend to have higher birth and death rates, resulting in directories that can only confidently cover enterprises above a certain size.

In countries with a significant informal economy, administrative registers are only used to complement the Economic Census. Unfortunately, Economic Censuses are not always exhaustive either, as they tend to cover only enterprises above a certain size or only those with fixed visible premises, leaving out those that operate within the homes of their owners or which are itinerant. In many countries, these activities can be very numerous. Tailoring and food processing, for example, can be very common activities in many countries and are generally carried out within dwellings; other common activities are construction,

---

[1] Based on *Measuring Informality: a statistical manual on the informal sector and informal employment* (ILO 2013). The author welcomes comments and suggestions for improvements and correction at mata@ilo.org.

transport and ambulant trade, all carried out without a fixed location. In addition, the information provided by Economic Censuses becomes outdated very quickly, especially for small and informal enterprises, which have high turnover rates, seasonal activity and mobility.

Summing up, in many countries, surveys based on directories tend to cover only the more stable, larger enterprises. Most of the enterprises left out, if not all, are small, and those among them that are informal can constitute a significant share. Producing statistics about the contribution of small enterprises, and in particular those in the informal economy, is essential to prepare complete national accounts and to inform economic and social policies. Yet only limited statistics exist about the number of enterprises in the informal economy and their characteristics. Available statistics about informal workers are more common as they are based on household surveys. But these surveys are not able to produce statistics about informal enterprises: only surveys of establishments can. The need to have these types of statistics has become more pressing since the adoption of the ILO Recommendation concerning the transition from the informal to the formal economy[2], which requested the production of statistics on informal enterprises to monitor progress (paragraph 36 of ILO 2015).

This paper argues that it is possible to carry out establishment surveys that cover these small and informal enterprises. To do so, they have to abandon the use of directories to select their samples and resort to area sampling instead. This paper covers two types of surveys that can use area sampling to cover small and informal economic units. It describes how they can be combined in a coherent way with conventional surveys of establishments. One samples economic units while the other samples informal entrepreneurs[3]. Both methodologies are described in detail in the ILO Manual *Measuring Informality* (ILO 2013).

# Defining informal enterprises

The international definition of enterprises in the informal economy was adopted in 1993 by the 15th International Conference of Labour Statistician. Broadly speaking, informal enterprises are unincorporated enterprises that do not keep accounts (i.e., they are not quasi-corporations), that sell at least partly for the market, are not registered under national legislation, and engage few employees or no employees at all (c.f., paragraphs 8 and 9 of ILO, 1993 for the full international definition). Although informal enterprises are not registered in national bodies, they are not necessary *illegal* because the goods and services they produce are not forbidden by law, nor are they *concealed* from public authorities because many countries do not require them to register or to declare taxes.

---

[2] This Recommendation was adopted by the International Labour Conference in 2015 to facilitate the transition of informal workers and economic units into the formal economy.

[3] Entrepreneurs are defined here as employed persons identified as employers or own account workers. If well measured, employers and own account workers refer to owners of unincorporated enterprises who also manage the economic unit they own. They differ only in that employers will hire employees on a regular basis, while the own account workers will not. Both may work with unpaid contributing family workers.

To describe informal enterprises, the term *economic unit* is often preferred, instead of the well established term *enterprise*. This is because informal enterprises do not conform to the notion of what an enterprise is: many are operated within dwellings, without a distinctive sign, or without a fixed location. Using *economic unit* also avoids specifying whether it is an *establishment* which is part of a larger *enterprise*: it is very difficult to distinguish establishments from enterprises in the informal economy, given the lack of written records to make this distinction. Finally, using *economic unit* can avoid confusing statisticians, given that the 15th ICLS resolution defines an *enterprise* differently than the SNA. In the 15th ICLS resolution, the production units owned by different household members are considered as separate enterprises (ILO, 1993a, paragraph 12), but the SNA considers all production units owned by household members as part of the same enterprise (UN et al., 2009, paragraphs 4.5 and 24.30).

## Combining list samples with area samples

Conventional surveys can be combined with area samples of small and informal economic units, by selecting mutually exclusive samples. To do so, the statistical universe of economic units needs to be divided into two groups. The first group includes all economic units covered by administrative registers that feed the directory of establishments (or enterprises). These cover all the public institutions and enterprises, all private corporations, all non-profit organizations and all unincorporated economic units that keep formal accounts (i.e., all quasi-corporations). This group often includes all large and medium economic units, but it can also cover small economic units that are covered by the directory. The second group includes all the other economic units, which are often small-scale units that are not registered and do not keep formal accounts. Depending on the exhaustiveness of the directory or list frame, this group may also include economic units that are registered and keep some form of accounts. Informal economic units are either all or a significant subset of this group of economic units.

Economic units in the first group are covered by conventional establishment surveys. The sample is selected from the list frame and can follow a typical design, for example, it can include all large economic units with 100% certainty, and a random stratified sample of all other units.

For the second group of economic units, *area frames* need to be used, much in the same way as they are used to select a sample of households in household surveys. Area samples only require that the geographical concentration of in-scope units of different types remains relatively stable between the time of the creation of the frame and the survey. Area frames are more robust than list frames for samples that cover small and informal economic units or entrepreneurs of these units, and can guarantee their full coverage.

However, selecting area samples is more time consuming and costly than selecting a sample based on a list frame. Area samples require costly listing of enumeration areas and need to be larger to achieve similar precision in the estimates. As a consequence, area surveys of economic units or entrepreneurs are still not common, and only a few countries have implemented them, such as Brazil, Colombia, Afghanistan and India.

Two possible approaches can be used to select area samples of economic units in the second group. The first uses area samples that contain the economic units themselves and the second uses area samples that contain the households where their entrepreneurs reside. Both follow a similar design, with two stages of selection. The first stage consists of

selecting areas proportional to the number of units (whether it be economic units or households of entrepreneurs) and the second stage consists of selecting units within the selected areas. When the sampling unit is the economic unit, the survey is known as an *area-based survey of economic units* and when the sampling unit is the household of the entrepreneur, it is known as a *mixed survey of households and enterprises*. Both surveys obtain the same information from entrepreneurs about their economic units.

The combination of surveys is represented in Figure 1 below. This Figure shows how economic units are divided into two groups, one that is covered by the list frame or directory and the other by the area frame. A sample is selected from each group: the sample of the conventional surveys, made up of a census of large economic units and a sample of medium and small economic units selected from a list sample frame; and the sample of the area-based survey, made up of units selected from an area sample frame.

Figure 1.  Combining different sample surveys to cover all economic units in the economy
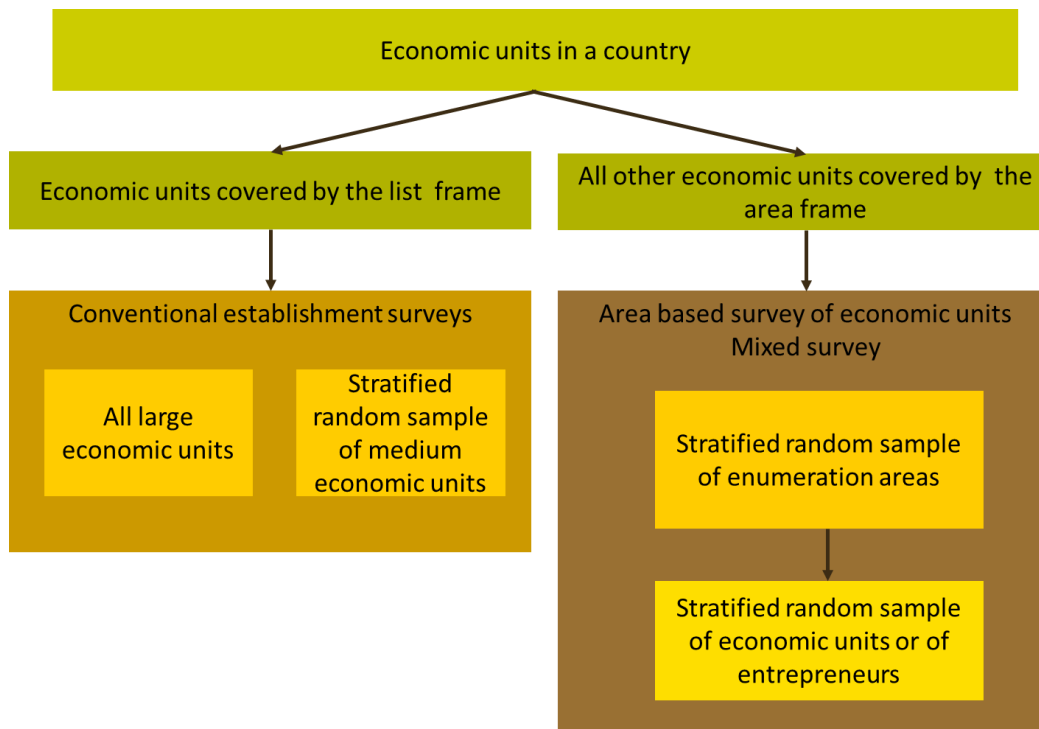


In order to avoid double counting of economic units, all units which belong to the list frame need to be excluded from the area frame. In order to have a consistent set of surveys, the survey based on an area frame needs to use the same concepts, definitions and classification systems as the conventional surveys of establishments. Such harmonization requires joint planning of regular establishment surveys and establishment surveys based on an area frame.

Combining conventional surveys of establishments with area based surveys maximizes the efficiency of the list-based conventional survey for units that it can cover with confidence, and allows for smaller units to be included as well. The result is full coverage of economic units, including of informal economic units, while at the same time avoiding double counting of economic units. As both surveys are consistent in terms of definitions and classifications, it is possible to compare and analyse the differences between sectors and to provide policy-makers with the information they require.

# Area-based surveys of economic units and mixed surveys

Both area-based surveys of economic units and mixed surveys aim at covering those economic units which are excluded from list-based establishment surveys. They differ in that the area-based establishment survey selects a sample of small and informal economic units while the mixed survey selects a sample of entrepreneurs of these economic units. Both types of surveys share the same need to simplify the data collection operation. The next sections briefly describe the selection of area samples of small and informal economic units, the selection of area samples of small and informal entrepreneurs, and those data collection issues that are most different from conventional establishment surveys.

Area based surveys of economic units can produce good statistics of the number and characteristics of small and informal economic units, as well as statistics of the number and characteristics of the workers in these units, thus producing coherent statistics of workers' conditions of work and of economic units' operations. They cannot, however, identify informal workers in formal units or in households producing exclusively for own final use (for example, as domestic employees). Also, because they sample economic units where they are located, they tend to miss the small activities operating within dwellings. In contrast, mixed surveys have the great advantage of being able to produce statistics about the number and characteristics of all informal workers, including those in formal units, as well as about the characteristics of small and informal units. They are particularly suitable to identify small activities which are mobile or that operate within dwellings, because they identify the activities of all persons residing in private dwellings, but are less suitable to identify enterprises operating outside of the dwelling, in fixed visible premises, given that they are fewer and require a larger sample than is usually selected in mixed surveys.

## Selecting an area sample for surveys of small and informal economic units

Surveys of small and informal economic units are based on area samples selected from an *area frame of small and informal economic units*. The best source for constructing an area frame of small economic units is the Economic Census[4] because it contains, for each enumeration area, information about the number of economic units that are excluded from

---

[4] Other sources can also be used to construct an area frame, such as the Population Census or even a list of villages and city blocks, as was done in Afghanistan in 2009. They will not contain the information that is required for a good sample design.

the list frame, their distribution by kind of activity and their size in terms of employment or production. The area frame should also include maps identifying the location of each in-scope economic unit at the time of the Census. As the coverage of economic units in the list frame varies from country to country, given that some will exclude only those economic units that are not registered while others will exclude all economic units below a certain size regardless of whether they are registered or not, this paper assumes that all small units are part of the area frame.

Ideally, the Economic Census should cover all economic units, including those that operate inside the dwelling of the business owner or without a fixed location, in addition to those above a certain size or which have fixed visible premises. This is best done when the Economic Census is carried out concurrently with the Population Census, as has been done in India since 1980, in Egypt since 1976, in Colombia since 2005 and in Indonesia since 1990.

A random stratified sample of enumeration areas is then selected with probability proportional to the number of small economic units. Strata can be by geographical location, concentration of in-scope units, kind of economic activities, etc. A list of all small economic units is then constructed in each selected enumeration area, by canvassing each building, whether commercial or living quarters, and collecting information about the economic units that operate there, their name and physical location, their kind of activity (name and code) and the number of workers engaged. It is important to cover all dwellings in order to include home-based activities and itinerant activities, which are very important among small and informal units. Itinerant activities are recorded as being carried out from the worker's home.

A very important task of canvassers at this stage consists of excluding those economic units in the enumeration area which are part of the list frame, in order to avoid double counting. It is therefore important that canvassers are equipped also with a list of economic units which belong to the list-base frame in each enumeration area selected.

The list of all small economic units in each selected enumeration area constitutes the sample frame to select a random sample of economic units, which may be stratified according to kind of economic activity, type of premises (fixed, within dwellings, itinerant) and the sex of the entrepreneur.

## Selecting an area sample for mixed surveys

Mixed surveys have a household component and an enterprise component. The household component identifies all households where employers and own account workers of small and informal units reside. The enterprise component obtains information about the economic units from a random sample of entrepreneurs identified in the household component. The enterprise component needs to be conducted simultaneously with or shortly after the household component, to facilitate the management and co-ordination of the two operations and to reduce attrition of the sample of entrepreneurs.

The household component in a mixed survey can be a full-fledged labour force or other household survey with an employment module (*mixed modular surveys*) or a listing of households in selected enumeration areas (*mixed independent surveys*). The enterprise component of *mixed modular* and *mixed independent surveys* is the same, namely a survey of entrepreneurs of small and informal units.

The first stage of a *mixed modular survey* is the household survey. The sample is selected following the standard household survey sample design, namely, by randomly selecting enumeration areas that contain households, then listing all households in these areas, in order to select a random sample of households. The household survey needs to include those data items that will allow the identification of small or informal entrepreneurs, namely, the status in employment (to identify own account workers and employers), the size of the economic unit (to identify small units) and bookkeeping and registration (to identify informal units). However, as this information is obtained through proxy response, its quality may not be sufficient to identify all relevant entrepreneurs with confidence. It is better, therefore, to identify as large a group as possible, say, all small entrepreneurs or all entrepreneurs, and to select a stratified random sample of entrepreneurs based on this large list of entrepreneurs. When detailed information about their economic units is obtained from the entrepreneurs themselves during the enterprise component, it will be possible to exclude with confidence all those entrepreneurs who are out of scope.

The first stage of a *mixed independent survey* is the selection of enumeration areas that contain households. As with mixed modular surveys, the Population Census is the ideal source for constructing the area frame, because it contains information about the number of persons residing in each area, by whether they are employed and by their status in employment. In contrast with the mixed modular surveys, the enumeration areas can be selected taking into account the density of small and informal entrepreneurs in each enumeration area. Within each selected area, a list of all small entrepreneurs is constructed by canvassing all living quarters, and obtaining information about the economic activities of each household member, their status in employment and other characteristics of their work. Based on the list of small entrepreneurs, a stratified random sample of entrepreneurs is then selected.

In both types of mixed survey, it may be that the sample of entrepreneurs will include some whose economic unit is covered by the conventional survey of establishments. In order to exclude from the final sample all of these cases, it is important to obtain information about the name, address and kind of economic activity of each economic unit and to compare it with the name, address and kind of economic activity of economic units in the list frame.

*Mixed modular surveys* are less complex and less expensive than *mixed independent surveys* and as a consequence are the most common. *Mixed independent surveys* on the other side, provide more precise estimates of the informal economy because their sample is specifically designed to produce estimates for each kind of economic activity about critical variables such as production, sales and costs.

## Data collection issues of surveys based on area samples

There are a number of data collection issues that are common to surveys of small and informal economic units and to mixed surveys, and which differ from those for conventional surveys of establishments. The main reason is that respondents to these surveys, namely, the small and informal entrepreneurs, may keep some accounts for personal reasons but not the detailed records and accounts that they can use to report the information. In addition, they may not be sufficiently literate to complete a questionnaire on their own. Small operators may have difficulty separating their operating expenses from their household expenses. Finally, addresses to identify the physical location of their economic units may be incomplete or unreliable.

This means that these surveys need to rely on direct interviewing. It is not possible to send questionnaires by mail for self completion, as is done in conventional establishment surveys, much less can any type of registers be used to complement the information collected, as is done in many developed countries.

This also means that the questionnaire used needs to be very simple. It should request fewer data items and with less detail, and should avoid technical terms. Explanations that may be given to larger economic units in writing, should be transformed into verbal questions that entrepreneurs can understand. Careful and rigorous training of investigators and supervisors is crucial to collect good data from small informal entrepreneurs. As answers cannot be based on detailed records of production, accounting and payroll, it is important to train them to probe and guide the respondent on the meaning of each question.

As with conventional establishment surveys, separate questionnaires may be designed for each major industry group. The questionnaire (or questionnaires) should define the main data items in the same way as they are defined in conventional establishment surveys. The main data items should relate to the conditions of business operation, to workers' conditions of employment, and to production in order to complement conventional business surveys as well as conventional earnings surveys. In addition to data items included in conventional surveys, surveys for small operators need to include data items that can be used to identify informal units, such as whether any accounts are kept in the economic unit, whether they are registered with government authorities, etc. They may also want to include questions that relate to the business conditions of informal units, such as whether they have access to credit, whether they own their own assets, and about the type of constraints that they face, to name a few.

Finally, because small operators need to rely on recall to provide information, the questionnaires need to request information for short or flexible reference periods. The reference period should not be longer than one month (UN, 1994), except perhaps for agriculture, where the previous six months can be used (ADB, 2011). Ideally, it should be a flexible reference period that allows respondents to choose the period for which they prefer to provide information, for example, a day, a week, or a month. The information can be adjusted to a common reference period at the data processing stage. In order to produce estimates for the whole year, the questionnaire can, for example, request approximate information about how typical the reference period was relative to the whole year (for example, was business during the reference period higher, normal or lower than usual), and about the level of activity during each month in the previous year, using very broad answer categories (high, normal, low). It is also possible to distribute the sample over the reference period for which one wants to estimate statistics, but this complicates field operations.

Given sampling and data collection issues, area based surveys cannot be carried out frequently. Annual surveys or surveys every three to five years are the norm among countries that carry out these surveys. In contrast, conventional surveys can be carried out fairly frequently, for example once a month or quarter, especially if they cover a small range of data items. This frequent survey can be supplemented by a more in-depth annual survey covering a wider range of data items in greater detail.

# References

Asian Development Bank (ADB). 2001. *A handbook on using the mixed survey for measuring informal employment and the informal sector (*Philippines). Available at: http://www.adb.org/sites/default/files/publication/28294/measuring-informal-sector.pdf

International Labour Office (ILO). 1993a. "Resolution concerning statistics of employment in the informal sector, adopted by the Fifteenth International Conference of Labour Statisticians", in *Fifteenth International Conference of Labour Statisticians (Geneva, 19-28 January 1993),* Report of the Conference (Geneva, ILO). Available at http://www.ilo.org/wcmsp5/groups/public/---dgreports/---stat/documents/normativeinstrument/wcms_087484.pdf

-. 2013. *Measuring informality: a statistical manual on the informal sector and informal employment* (Geneva, ILO). Available at http://www.ilo.ch/wcmsp5/groups/public/---dgreports/---dcomm/---publ/documents/publication/wcms_222979.pdf

-. 2015. Recommendation concerning the transition from the informal to the formal economy (Geneva, ILO). Available at http://www.ilo.org/wcmsp5/groups/public/---ed_norm/---relconf/documents/meetingdocument/wcms_377774.pdf

United Nations. 1994. *Strategies for measuring industrial structure and growth,* Studies in Methods, Series F, No. 65 (New York, UN Statistical Office). Available at http://unstats.un.org/unsd/distributive_trade/SeriesF_65.pdf

-; European Commission, IMF, OECD, World Bank. 2009. *System of National Accounts 2008.* Available at http://unstats.un.org/unsd/nationalaccount/SNA2008.asp

United Nations Statistical Commission. 2007. *The results of the UNSD survey of country practices in economic census.* Available at http://unstats.un.org/unsd/economic_stat/Economic_Census/UNSD%20survey%20of%20country%20practices%20in%20economic%20census.pdf