Guatemala Enterprise Informal Survey 2010 Data Set

1. Introduction

This document provides information on the Informal Enterprise Survey (IFS) collected in Guatemala from September 29th to November 20th, 2010. The surveys were conducted through the joint collaboration of the World Bank and the Inter-American Development Bank.

1.1. The Informal Enterprise Surveys (IFS) collect data on non-registered business activities in every region of the world. The IFS are implemented in parallel to the World Bank's Enterprise Surveys (ES), which interview formal, private, non-agricultural firms in countries around the world (www.enterprisesurveys.org).

The IFS use a standardized survey instrument designed to assess the business environment for non-registered businesses within a well-defined universe of activities, which have been identified using information from previous iterations of the studies. The IFS cover business environment topics including: general business characteristics, infrastructure, crime, sales & supplies, finance, labor, registration, business environment, and assets.

The IFS are conducted using a uniform sampling methodology in order to minimize measurement error and yield data that are comparable across the world's economies.

1.2. The Informal Surveys aim to accomplish the following objectives:

- To provide information about the state of the private sector for informal businesses in client countries;
- To generate information about the reasons of said informality;
- To collect useful data for the research agenda on informality; and
- To provide information on the level of activity in the informal sector of selected urban centers in each country

This report describes the survey population, the sampling methodology and the sample structure of the Guatemala informal survey; additionally, this report provides information that may be useful when using the data, including information on non-response rates and country-specific factors that may have affected survey implementation.

2. Survey Target Population

2.1. The primary sampling units of the IFS are un-registered business entities.[1] For consistency, "registration" is defined according to the established convention for the Enterprise Surveys in each country. In these surveys, the requirements for registration are defined on a country-by-country basis consulting information collected by Doing Business and information from the in-country contractors.

2.2. For Guatemala, informal firms were defined as those not registered with the Superintendencia de Administración Tributaria (SAT) or the Registro Mercantil.

2.3. At the beginning of each survey, a screening procedure is conducted in order to identify eligible interviewees. At this point, a full description of all the activities of the business owner or manager is taken; based on its principal activity, a business is then classified in the manufacturing or services stratum using a list of activities developed from previous iterations of the survey. Certain activities are excluded such as strictly illegal activities (e.g., prostitution or drug trafficking) as well as individual activities that are forms of selling labor like domestic servants or windshield washers.

3. Sampling Methodology

3.1. In each country, the IFS are conducted in two selected urban centers, which are intended to coincide with the locations for the implementation of the main Enterprise Surveys. The overall number of interviews is pre-determined, and these interviews are distributed between the two urban centers, according to criteria such as the level of business activity and each urban center's population, etc.

3.2. In Guatemala, the urban centers identified were Guatemala City and Quetzaltenango. The target sample for Guatemala City was 184 interviews and 120 interviews in Quetzaltenango.

3.3. Sampling in the IFS is conducted within clearly delineated **sampling areas**, which are geographically determined divisions within each urban center. **Sampling areas** are defined at the beginning of fieldwork, and are delineated according to the concentration and geographical dispersion of informal business activity. The distribution of informal business activity is estimated through discussions with local implementing contractors at the outset of fieldwork.

---

[1] Within the economics literature the concept of the informal firm is defined in many different ways. At times, combinations of these definitions are used in academic and policy work. In order to avoid confusion and also facilitate the standardization of methodology and data comparability, the IFS equate informality with non-registration.

3.4.   The number of **sampling areas**, and the geographical area they contain, is determined with the goal that each **sampling area** will yield four effective interviews.

3.5.   In Guatemala, each **sampling area** was designed to contain a physical area, on average, of no less than the equivalent of eight city blocks. These **sampling areas** may or may not correspond to the administrative districts of the urban center.

3.6.   Moreover, in order to ensure a degree of geographical dispersion, each urban center was divided into distinct **zones**. In both Guatemala City and Quetzaltenango, a **central zone** was identified as well as four quadrants, **north, south, east,** and **west**, as illustrated in the diagram below:

North

West   Central   East

South
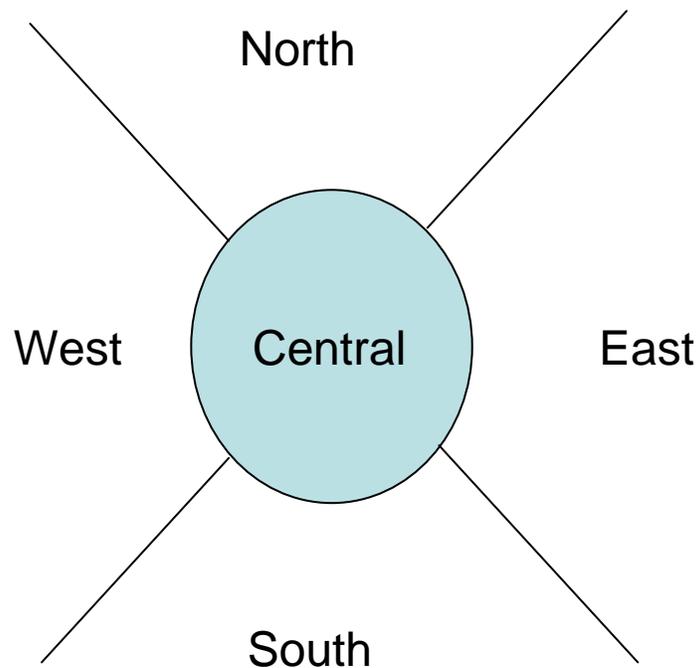
3.7.   Finally, due to issues of non-response in previous iterations of the IFS, a substitution procedure was determined before the launch of fieldwork, allowing for the possibility that each sampling area will not yield the target four interviews. Prior to fieldwork, a number of replacement sampling areas were assigned *by zone*. That is, in addition to **primary sampling areas**, each zone contained an allotted number of **replacement sampling areas** which were authorized to be used if the contractor encountered difficulty in obtaining the target number of interviews within each zone. This substitution procedure ensured that a degree of geographical representation was maintained. Once the total number of sampling areas in each zone was determined (i.e., primary plus

replacement), each sampling area was then randomly designated as either primary or replacement.

3.8.   In Guatemala City, **46 primary sampling areas** (184/4 = 46) were required; **30 primary sampling areas** (120/4 = 30) were necessary in Quetzaltenango.

3.9.   The **sampling areas** were distributed among the **zones** according to estimates – determined in conversations with the local contractor – of the concentration of informal activity in each geographical zone. In Guatemala City, the distribution of the **sampling areas** was as follows: central – 12 sampling areas; north – 5 sampling areas; west – 5 sampling areas; south – 12 sampling areas; and east – 12 sampling areas. In Quetzaltenango, the distribution was: central – 7 sampling areas; north – 5 sampling areas; west – 7 sampling areas; south – 5 sampling areas; and east – 6 sampling areas.

3.10. Overall, the distribution of the sampling areas and the allotment of replacement sampling areas was as follows:

City	Zone	Sampling Areas	Replacements
Guatemala	1 - Central	12	3
Guatemala	2 – North	5	3
Guatemala	3 – West	5	2
Guatemala	4 – South	12	3
Guatemala	5 – East	12	3

City	Zone	Sampling Areas	Replacements
Quetzaltenango	1 - Central	7	3
Quetzaltenango	2 – North	5	1
Quetzaltenango	3 – West	7	3
Quetzaltenango	4 – South	5	2
Quetzaltenango	5 – East	6	1

3.11. In order to provide information on diverse aspects of the informal economy, the sample is designed to have equal proportions of services and manufacturing (50:50). These sectors are defined by responses provided by each informal business to a question on the business's main activity included in the screener portion of the questionnaire.

As a general rule, services must constitute an ongoing business enterprise and so exclude the sale of manual labor (e.g., windshield washers), as noted above. Manufacturing activity in the informal sector includes business activity requiring inputs and/or intermediate goods. Thus, for example, the processing of coffee, sugar, oil, dried fruit, or other processed foods is considered

manufacturing, while the simple selling of these goods falls under services. If an informal business conducts a mixture of these activities, the business is considered under the manufacturing stratum.

3.12.   Thus, each **sampling area** was designed with the goal of obtaining two interviews in services and two interviews in manufacturing.

3.13.   In order to ensure a degree of geographical dispersion *within* each **sampling area**, two **starting points** were identified.

3.14.   Each **starting point** was designed to correspond to four city blocks, which were numbered sequentially (see example below). The first **starting point** was identified as Starting Point A (blocks 1-4) and the second as Starting Point B (blocks 5-8).

3.15.   Proceeding from each **starting point**, interviewers were instructed to begin on the first block (i.e, 1 or 5), defining the starting block and corner. Each interviewer was instructed to attempt to achieve two interviews from each starting point, ideally one interview in manufacturing and one in services.

3.16.   Interviewers were instructed to proceed clockwise around block 1 from Starting Point A; if the target interviews were not achieved, interviewers proceeded to block 2, Starting Point A, and so forth until completing a circuit of block 4.

After achieving two interviews from **starting point A**, interviewers were instructed to cease work in the blocks assigned to that given starting point and repeat the same procedure from **starting point B**, beginning with block 5.

3.17.   Using the local knowledge of enumerators and the implementing contractor to help identify informal business activity, within each block all houses and shops were checked for unregistered businesses, following the pre-fixed route described above, until the allotted quota of interviews for the sampling area was reached. The implementing contractor reported that informal businesses were identified frequently as those that did not display commercial registration or sanitation permits, as required by Guatemalan law. Informal firms were also

frequently identified as those issuing receipts (if at all) without an NIT (Número de Identificación Tributaria), which Guatemalan law requires on sales records.

3.18. Due to the sampling procedures described above as well as the geographical dispersion of the targeted informal businesses, detailed maps using GPS coordinates and Google maps were produced.

3.18.1. Each **sampling area**, including its two **starting points**, were delineated using Google maps (or Google Earth), with the GPS coordinates of the starting points being systematically recorded.

3.18.2. Additionally, when obtaining a complete interview, the exact address of the informal business (or where the interview took place) was registered by the interviewer. Once in the office, this address was searched in Google maps, and its GPS coordinates were registered in a fieldwork report.

3.18.3. If no address was immediately available, using local knowledge, the GPS coordinates were determined using imaging via Google maps.

3.18.4. Documentation of this mapping procedure (if available) can be found in the accompanying appendices. In order to preserve confidentiality, the exact coordinates of businesses are not published.

4. Sampling Structure

4.1. In **Guatemala**, 303 interviews were realized. The breakdown of the final sample was:

City	Interviews
Guatemala City	184
Quetzaltenango	119
TOTAL	303

4.2. Moreover, the overall distribution by identified sectors is given in the following table (manufacturing sectors highlighted in grey):

Sample distribution by activity sector	Guatemala	Quetzal	Freq.
Manufacturing of clothes or shoes	4	4	8
Manufacturing of baked food (Baker)	22	4	26
Manufacturing of coffee, sugar, oil, dry fruits and other processed foods(exclude restaurants, street food-sellers,)	9	2	11
Manufacturing of handcrafts	2	9	11
Manufacturing of furniture	6	2	8
Manufacturing of metal products	1	1	2
Manufacturing of household items	1	0	1
Manufacturing of tools and instruments	0	0	0
Other manufacturing (not included above)	47	35	82
Selling food or Groceries (street food sellers, restaurants)	36	27	63
Selling of clothes or household items	13	8	21
Selling of business/computer/phone services	2	3	5
Selling of other goods	8	7	15
Transport Services	2	1	3
Cleaning and washing services	1	0	1
Hairdressers and barber shops	0	0	0
Professional Services (including internet services)	0	1	1
Repairing services	5	2	7
Other services (not included above)	22	10	32
Total	**184**	**119**	303

5. Non-response

5.1. The overall survey response rate among contacted, eligible businesses for the Guatemala IFS was estimated at 20%. The implementing contractor noted a number of factors that affected response rates:

5.1.1. Interviewers were occasionally confused with government officials; the implementing contractor suspected that this confusion lowered response rates.

5.1.2. Interviewers dressed formally to assure respondents of the formal nature of the study. Enumerators carried clear identification associating them with the implementing contractor.

5.1.3. The implementing contractor reported that informal businesses were more willing to respond in Quetzaltenango. The contractor additionally noted that a higher proportion of business activity was informal in Quetzaltenango.

6. Database Structure

6.1. One version of the questionnaire was used for all interviews; variable *a0* indicates the module of the questionnaire. There are two levels of distinction in the informals data set: the data are distinguished by both the urban center/region and as a manufacturing or services business or activity, as detailed above.

6.2. All variables are named using, first, the letter of each section and, second, the number of the variable within the section, i.e. *a1* denotes section *A*, question *1*. All variables are numeric with the exception of those variables with an "x" at the end of their names. The suffix "x" denotes that the variable is alpha-numeric.

6.3. There are 2 establishment/business identifiers, *idstd* and *id*. The first is a global-unique identifier. The second is a country-unique identifier.

6.4. As noted above, there are two levels of distinction: industry (manufacturing or services) and region. Variable *a3a* denotes the region[2] (here: urban center) where the interviewed business is located. The variable *sect* indicates whether a business or activity was treated as manufacturer or service for the purposes of the interview. It is important to note that this variable is based upon variable *Sc2b*, which codes in detail the activity of the business or activity for the purposes of the questionnaire. *Sc2a* codes in detail the **main** product of the business or activity (in terms of sales value). As noted above, if an informal business or service has a main activity that is a service, yet has a manufacturing product, then the business is treated as a manufacturer. This treatment allows certain questions to be asked that only pertain to manufacturers (i.e., regarding inputs).

6.5. The screener information section (section Sc.) contains questions about the business activity and basic physical location of informal businesses; section B provides general information on the business and its ownership; section C discusses location and infrastructure; section I contains questions on crime; section D information on sales and supplies; section K is on finance; section L poses questions on labor; section R contains questions on registration; section M the business environment; and section N includes questions on business productivity.

6.6. All data are made available using Stata software. Due to restrictions of this software, variable labels are often shortened in the full data set. ***Please consult***

---

[2] Term "region" used in order to maintain comparability with the formal Enterprise Surveys.

the full questionnaire for the complete text of questions and additional information.

7. Country-specific implementation comments

7.1. The implementing contractor noted the following relevant issues in Guatemala during fieldwork:

    7.1.1. Higher refusal rates resulted from target businesses suspecting that interviewers were from the government tax bureau. This confusion occurred both at the point of the initial interview as well as at the point of follow-up quality control contacts.

    7.1.2. The implementing contractor reported that interviewees were reluctant to give sales and salary information.

    7.1.3. As many of the respondents directly attend to customers, there were frequently breaks in conducting the interview.

    7.1.4. In Guatemala City, the contractor reported issues with the security situation, citing respondents' fears of frequent attacks and extortion.

    7.1.5. Over the course of fieldwork, the access road to Quetzaltenango was severely damaged by tropical storm Agatha, reducing many sales, particularly in businesses using agricultural products.