# Description of The Acayucan Standards of Living Survey*

Marco González-Navarro
Princeton University

Climent Quintana-Domeque
Princeton University

This draft: January 2007

**Abstract**

This article documents the characteristics of the Acayucan Standard s of Living Survey. This household survey was collected in February-March 2006 in Acayucan Mexico. It contains information on labor supply, family composition, education, income, expenses, housing quality and expectations for 1,200 families.

**Keywords:** Mexico, data collection, household surveys, living standards.
**JEL Codes:** C8, I31

*We thank Anne Case, Angus Deaton, Bo Honoré, Alan Krueger, Chris Paxson, Cecilia Rouse, and all the seminar participants at the Princeton Labor and Development Lunches for their valuable comments and helpful suggestions. Thanks also to Jose Luis Palma and Hector Duran and the rest of people in INSAD for working so hard and putting all the best in this project. González-Navarro gratefully acknowledges financial support from the Fellowship of the Woodrow Wilson Scholars and Quintana-Domeque from the Bank of Spain. This project is funded by the Industrial Relations Section, the Research Program in Development Studies, and the Center for Health and Wellbeing. All errors are ours.

# 1 Introduction

The Acayucan Standards of Living Survey (ASLS) is a multipurpose survey that was applied in the city of of Acayucan, México during the months of February and March in 2006. Acayucan is a small city in southern México with a population slightly shy of 50,000 people. The survey consisted of a household questionnaire, which covers information on demographics, labor force behavior, income, health, education, housing, and expectations. In this paper, following a similar description to Peracchi (2002), we document some of the key characterisitcs of the survey.

The paper is laid out as follows. Section 2 describes the organization of the survey (coverage, sample design, questionnaire, fieldwork, and data editing). Section 3 is about confidentiality and data access issues. Section 4 discusses weighting and nonresponse. Section 5 concludes.

# 2 Organization of the Survey

The ASLS was fielded by survey company INSAD (Investigación en Salud y Demografía, *Research in Health and Demography*) under the supervision of González-Navarro and Quintana-Domeque. The data was collected in 2006 during the second half of February and the beginning of March. González-Navarro and Quintana-Domeque are associated with the Economics Department, and the Research Program in Development Studies and the Industrial Relations Sections, respectively. INSAD is a survey-research unit working on health and population issues in México. They are responsible for sample selection, questionnaire design, fieldwork, basic data processing and editing of the household questionnaire.

## 2.1 Coverage and Sample Design

The target population of the ASLS consists of all private households located on 56 clusters consisting of households without paved streets. Each cluster contains all households on a set of connected streets. These streets are mainly in the outer parts of the city and have the

lowest socioeoconomic indicators of the town, according to Census data. It should be noted that the target population is not representative of the country, nor the city. It merely consists of clusters of households in the more deprived areas of the city.

Within a dwelling, a household is defined as a group of people who "share both the same dwelling and food expenditures". INSAD hired a team of specialists in cartography who drew maps indicating and describing the location of each construction within a cluster. The constructions were marked as being inhabited dwelling, uninhabited dwelling and commercial purpose. This was done for two main reasons. First, as recognized by Deaton (1997), the use of out-dated or otherwise inaccurate sampling frames is an important source of error in survey estimates. To the best of our knowledge, no reliable sampling frame for dwellings was available in Acayucan. Second, we were interested in having a detailed census of all business and stores in the area for specific research purposes (e.g., to measure the level of economic activity within each street).

Once we had the cartography information of the target population, we randomly selected the number of dwellings to be interviewed. It is important to note that all the households within each dwelling were planned to be interviewed. The dwelling sampling procedure was done as clustered sampling, and weights, adjusted by non-response and representing the inverse of the probability that an observation is included to the sample, must be taken into account.

A total of 1,363 dwellings were randomly selected from the cartographic maps to participate in the ASLS. After eliminating mistakes in cartography (for example cases where cartographers marked a dwelling the field workers found an empty lot) and non-response, we obtained completed interviews in 1,201 dwellings. The number of households within these 1,201 dwellings is 1,239, i.e., 1.03 households per dwelling. It is worth noting that neither quota sampling nor substitution of non-responding households or indviduals (whether "refusals" or "non-contacts") were permitted at any stage. The households sampled constituted between 50 to 70% of households in the street clusters.

## 2.2 Questionnaire

The household questionnaire collects detailed information on demographic characteristics, labor force behavior, health, education, housing, expectations, income, expenditures at both the household and the personal level. Both household and individual questions were answered by a "reference person" (in general, the household head or the spouse/partner of the head). This is a common technique used in several surveys and is justified on two main grounds: (i) it is both an easy and a cheap way to obtain information; (ii) the "reference person" usually has knowledge about the issues at hand, for example children's health. Due to budgetary reasons, we did not attempt to separately survey other adults in the household, except through the reference person proxy. Table –1 shows the distribution of the main respondents in the survey:

**[Insert Table –1 about here]**

It is well-known that the complexity of a questionnaire affects negatively both response rates and data quality. How long it takes to fill the questionnaire, the wording of the questions, the reference period for the questions, and the branching and skip patterns are among the several dimensions of questionnaire complexity.

To deal with the potential shortcomings associated with the complexity of the questionnaire, namely low response rates and low data quality, its design was based on existing questionnaires, mainly the Encel questionnaire which was used to determine PROGRESA (Oportunidades) eligibility in 1997, and suggestions from various participants in both the Labor (October 2005) and Development (January 2006) Lunches at Princeton University, the specific knowledge of INSAD on health and population issues in México, and a pre-test of the questionnaire on the field through a pilot study. This last source of information provided us with extremely valuable information to improve or to delete some questions from the final questionnaire.

To assess the degree of complexity involved in this survey and understand the design of the questionnaire, some key facts should be acknowledged. First, the ASLS questionnaire is long, though its advantage with respect to other long questionnaires is that the large variety of topics/issues of its questions makes it less boring/monotone for the interviewed person. The

median duration of household interviews was 28 minutes. Second, subjective and sensitive wording questions, such as *How satisfied do you feel with respect to the performance of the Federal Government?* (municipal/state), are not frequent. Third, the reference period varies considerably depending on the question being asked (last week, last 30 days, last month, last 3 months, last 6 months, last 12 months, last year, next 5 years, etc.). Finally, branching and skip patterns were clearly explained in training sessions, so none of these are likely to have created problems to interviewers.

Table – 2 shows the structure the questionnaire dealing with individual characteristics.

**[Insert Table –2 about here]**

Table – 3 shows the structure the questionnaire dealing with household characteristics.

**[Insert Table –3 about here]**

As shown in Table –3, interviewers were also asked to make an overall assessment regarding the presence of litter on the street and the cleanliness of the dwelling.

### 2.3 Fieldwork and Interviewers

The fieldwork period lasted 1 month, between the end of February 2006 and the middle of March 2006. The interviewing method was face-to-face personal interviewing and interviews were carried out using the conventional "paper and pencil" method.

The main criteria for pre-selecting the interviewers was their past experience (Census, Federal Electoral Institute, etc.). Once at INSAD, the pre-selected interviewers were given training on the working of the questionnaires, on how to get valid responses on sensitive items and the instructions in case of refusal to respond to certain questions. Sections of the questionnaire that required special attention were pointed out and explained carefully to interviewers. This training was done under the supervision of González-Navarro and Quintana-Domeque. After that, field interviewers were selected through a rigorous selection process.

There were four field teams, each was composed of one field supervisor and four interviewers. Interviewer's assignment size (workload) were 3-4 interviews per day. Quality control

back-checks were performed by three additional supervisors. Moreover, Quintana-Domeque was performing monitoring tasks in real time.

### 2.4 Data Editing

Most items in the questionnaire are pre-coded. A few socio-demographic items were recorded verbatim and subsequently coded by INSAD. INSAD performed a set of pre-specified range and logic checks following a well-established procedure to minimize codification errors: there were two independent processes of data entry. Once the preliminary version of the dataset was available, INSAD sent it to the researchers who checked for possible inconsistencies (June 2006). INSAD was informed about the results of these checking procedures (September 2006). INSAD was requested to review the data and edit items where possible through cross-checking with paper questionnaires. Reexamination of the revised version by the researchers was satisfactory (November 2006).

## 3 Unit Nonresponse and Weighting

Any study involving survey data needs to devote special attention to the issue of nonresponse. Nonresponse is one of the causes for the loss of precision of sample estimates. Furthermore, when nonresponse is not random, it is not clear whether the population parameters we are interested in can be estimated without bias using their sample estimates counterparts. The seriousness of the problem is directly proportional to the amount of nonresponse (Horowitz and Manski, 1998).

In the ASLS, we can distinguish two types of nonresponse: unit nonresponse and item nonresponse. Unit nonresponse occurs when information is missing for a sample unit. In general, this is due to refusal or inability of the unit to participate. Item nonresponse occurs when a unit refuses to answer a particular question included in the survey.

Before computing unit nonresponse rates, we should exclude ineligible dwellings from our analysis:

- Unoccupied premises
- Non-residential address (e.g., solely business, school, office, factory, etc.)
- Address occupied, but not residential household (e.g., weekend homes)
- Address occupied by resident household, but not eligible respondent (no one aged 18+)

Table – 4 shows the distribution for the interview final results (after at least four interview attempts).

**[Insert Table –4 about here]**

Table – 5 shows a more schematic distribution.

**[Insert Table –5 about here]**

The response rate is 93.75%, and hence the nonresponse rate is 6.25%. In comparison to existing surveys, this nonresponse rate is very low, which makes us less concerned about the seriousness of nonresponse when analyzing the data from the ASLS.

In order to obtain reliable estimates of means at the street level, observations need to be weighed, taking into account unit nonresponse.

There are two kind of available weights in the ASLS: a dwelling weight and a household weight. These weigths (or expansions factors) are adjusted by nonresponse and represent the inverse of the probability that an observation (dwelling or household) is included to the sample. For each dwelling $j$, its corresponding weight $w_j$ is defined as

$$w_j^d = \frac{N_j^d}{n_j^d} \tag{1}$$

where $N_j^d$ is the number of total dwellings in street $j$ and $n_j^d$ is the number of dwellings in street $j$ with completed interviews. While for each household $i$, its corresponding weight $w_i$ is defined as

$$w_i^h = \frac{N_j^h}{n_j^h} \tag{2}$$

where $N_j^h$ is the number of total households in street $j$ and $n_j^h$ is the number of households in street $j$ with completed interviews.

## 4 Data Access

The ASLS is not currently available to the public and its use is restricted to the PI's, but will be made public in the near future. This research project, approved by the Princeton University Institutional Review Panel for Human Subjects (Protocol # 3104), was entirely funded by the Industrial Relations Section, the Research Program in Development Studies and the Center for Health and Wellbeing, all of them at Princeton University.

## 5 Discussion

This paper has offered a brief and preliminary description of the ASLS. This survey is not currently available to the public and its use is restricted to the PI's, but will be made public in the near future. Given the rich set of information included in this survey, there is no doubt it will become an important source of data for opening new avenues for further research in the context of developing countries.

## References


[1] Peracchi, F. (2002) "The European Community Household Panel: a review". *Empirical Economics* 27:63-90.

[2] Deaton, A. (1997) *The Analysis of Household Surveys:a microeconometric approach to development policy.* The Johns Hopkins University Press for the World Bank: Baltimore.

[3] Horowitz, J. and CF Manski (1998) "Censoring of outcomes and regressors due to survey nonresponse: identification and estimation using weights and imputations". *Journal of Econometrics* 84:37-58.

# TABLES

**Table – 1.**
**Main Respondent**

| | Freq. | Percent | Cum. |
|---|---|---|---|
| Head of the Household | 534 | 44.46 | 44.46 |
| Spouse or Partner | 608 | 50.62 | 95.09 |
| Son or Daughter | 37 | 3.08 | 98.17 |
| Father or Mother | 12 | 1 | 99.17 |
| Grandfather or Grandmother | 9 | 0.75 | 99.92 |
| Nephew or Niece | 1 | 0.08 | 100 |
| Total | 1,201 | 100 | |

**Table – 2**
**Questionnaire: I. Information on Individual Characteristics**

The 'reference person' reports information for

A. Everyone living in the dwelling about:
   - Age (I.3)
   - Sex (I.4)
   - Residence condition of each household member (I.5)
   - Relationship of each household member with the household head (I.6)
   - Health care coverage (I.7)
   - Self-reported health status (I.8)
   - Chronic Disease and Type (I.9) (I.10)
   - Physical or Mental Disability / Type (I.11) (I.12)
     - Demand for Health Services/ Where/ Which/ Transport/ Time (I.13)-(I.17)
   - Health Problems (last month): cough, fever, flu, bronchitis, vomit, diarrhea, stomach ache. (I.18)-(I.24)
   - Health Problems (last year): fungus, intestinal parasites, skin disease. (I.25)-(I.27)
   - Other Health Problems (last month) / Type (I.28) (I.29)

B. Everyone aged 5+:

- Indigenous Language (I.30)
- Literate (I.31)
- Went to school in the past (I.32)
- Currently school enrolled (I.33)
    - Transport /Time/Which school (I.34)-(I.36)
    - Missing Days / Reason (I.37)-(I.38)
- Highest School Degree (I.39)

C. Everyone aged 8+:

- Worked last week (verification) (I.40)-(I.42)
    - Firm size (I.43)
    - Activity (I.44)
    - Job Position (I.45)
    - Labor Supply last week (days / hours per day) (I.46) (I.47)
    - Monthly labor income (I.48)
    - Transport / Time / Where (I.49)-(I.51)

D. Everyone aged 18+:

- Other income flows / Specific Sources / Amount (I.52)-(I.54)
- Loan / Who / Purpose / Amount (I.55)-(I.58)

**Table – 3.**
**Questionnaire: II. Information on Household Characteristics**

The 'reference person' (RP) reports information about

- The number of years she (if the RP is the household head) or her family (if the RP is not the household head) has been living in the dwelling (II.2) (II.3)
- Someone in the household receives some kind of government aid (II.4)
- Expectations about someone in the household leaving Acayucan for working/schooling reasons during the next 5 years (II.5) (II.6)
- Characteristics of the dwelling
    - Tenure status (II.7)
    - Property Title (II.8)
    - Material of the floor (II.9)
    - Material of the roof (II.10)
    - Material of the walls (II.11)
    - Number of rooms (II.12)
    - Water (II.13)
    - Sewage System (II.14)
    - Bathroom (II.15)
    - Electricity (II.16)
    - Flooded (II.17)
    - Squared meters of the plot of the dwelling (II.18)
    - Self-assessed price of the dwelling (II.19)
- Wealth Indicators
    - Bank account (II.20)
    - Credit card (II.21)
    - Assets (van, car, motorcycle, cell phone, etc.) (II.22)
- Materials bought in the last 6 months to make home improvements (II.23)
- Home improvements made in the last 6 months (II.24)
- Expectations about improvements in the next 2 years (II.25)
- Fuel to cook (II.26)
- Gas truck (II.27)
- Trash (II.28)
- Street sellers (II.29)
- Water source (II.30)
- Food is bought… (II.31)
- Livestock / Which kind / How many (II.32) (II.33)
- Public Safety
    - Walking during night (II.34)
    - Someone broke into your home (II.35)
    - Your vehicle was stolen (II.37)
- Public Services
    - Bus (II.38)
    - Transport to downtown (II.39)

  - Taxi (II.40)
  - Taxi Price (II.41)
- Satisfaction regarding living in Acayucan / Why (II.42) (II.43)
- Hours of sleep (II.44)
- Satisfaction with local government (II.45)
- Satisfaction with state government (II.46)
- Satisfaction with federal government (II.47)
- Public services (II.48)
- Responsible for paving streets (II.49)
- Monthly Household expenditure by item (food, rent, telephone, etc.) (II.50)
- Monthly Total Household expenditure (II.51)
- Someone started a business / Type / Where (II.52)-(II.54)
- Someone improved her business / How / Type / Where (II.55) (II.56)
- Interviewer #1 (II.61)
- Interviewer #2 (II.62)

**Table – 4.**
**Interview Final Results**

| | Freq. | Percent | Cum. |
|---|---|---|---|
| Completed | 1,201 | 88.11 | 88.11 |
| Postponed | 1 | 0.07 | 88.19 |
| No one in the dwelling | 14 | 1.03 | 89.21 |
| Refused | 19 | 1.39 | 90.61 |
| Not Eligible Respondent | 3 | 0.22 | 90.83 |
| Not Residential Household | 17 | 1.25 | 92.08 |
| Unoccupied | 31 | 2.27 | 94.35 |
| Weekend Homes | 31 | 2.27 | 96.63 |
| Other | 46 | 3.37 | 100 |
| Total | 1,363 | 100 | |

**Table – 5.**
**Response Rate**

| | Freq. | Percent |
|---|---|---|
| *Eligible Dwelling* | 1,281 | 100 |
| Response | 1,201 | 93.75 |
| Non-response | 80 | 6.25 |
| *Ineligible Dwelling* | 82 | 100 |