THE 1993 NICARAGUAN LIVING STANDARDS

MEASUREMENT SURVEY:

DOCUMENTATION

Revised
August 2002

**DEVELOPMENT RESEARCH GROUP - POVERTY
WORLD BANK**

TABLE OF CONTENTS

APPENDICES

LIST OF TABLES

LIST OF FIGURES

DOCUMENTATION
LSMS DATA
NICARAGUA, 1993

I.   BACKGROUND

1   In the 1980s the World Bank designed a survey mechanism to measure the living standards of the populations in developing countries. The data collected using a Living Standards Measurement Study (LSMS) survey can be used to assess development policy, measure and identify poverty, and evaluate the effectiveness of various government interventions in the area of poverty alleviation, social services, labor and community infrastructure. The first of these LSMS surveys were carried out in the Ivory Coast and Peru in 1985 and, as of 2002, LSMS surveys have been carried out in over 30 countries.[1]

2   An LSMS survey was carried out in Nicaragua from February 21 through June 12, 1993. The Nicaraguan LSMS (LSMS_NICA_93) was carried out at a national level and consists of a household survey that was administered in urban and rural areas in all regions of Nicaragua as well as community and price questionnaires administered in rural areas. LSMS_NICA_93 represents a significant increase in the data available in Nicaragua: during the 1980's there was little survey research of any type undertaken in the country. No Population and Housing census was carried out during the decade[2] and the most recent national level survey dates from 1985. Several smaller surveys were carried out in 1989 but covered only the department of Managua.

3   The LSMS_NICA_93 survey was implemented by the National Institute of Statistics and Census of Nicaragua (Instituto Nacional de Estadísticas y Censos - INEC) and was supported by IDA with co-financing from the United Nations' Development Program (UNDP) and Children's Education Fund (UNICEF), the Swedish International Development Authority (SIDA), the United States' Agency for International Development (USAID), and the Government of Nicaragua (GON). Techical assistance was provided by IDA.

4   The purpose of the present document is to provide detailed documentation of the LSMS_NICA_93. The report is organized in seven sections. Section II contains a detailed outline of the types of data collected in the LSMS_NICA_93. Section III provides information on the sample design used. A description of the field work is found in Section IV, and in Section V there is a brief assessment of the data quality. Several of the aggregate consumption variables which were constructed for the poverty assessment of Nicaragua[3] are discussed in Section VI and the final section contains information on an anthropometric and nutrition data set that was created from the original data.

---

[1] See the LSMS website for a complete listing: http://www.worldbank.org/lsms/lsmshome.html

[2] The last Population and Housing Census was done in 1972. Additionally, the last Agricultural Census dates from 1962. (An Agricultural Census was carried out in 1972 but the results were lost in the 1972 earthquake.)

[3] See "Republic of Nicaragua: Poverty Assessment," Report No. 14038-NI, Country Department II, Country Operations Division, Latin America and the Carribean Regional Office, World Bank, Washington D.C., June 1, 1995.

II.   DATA COLLECTED

5   Three types of data were collected for the LSMS in Nicaragua: a) the main source of information was gathered through the administration of a questionnaire to the households in the sample; b) the second source of data is anthropometric information on all children under five in the households interviewed; and c) the third source is a community level questionnaire designed to obtain economic and infrastructure data as well as price data.

6   **Household questionnaire:** The household questionnaire covered a broad range of topics. An outline of the topics covered can be found in Table 1 and the remainder of the section provides more detailed information on the data collected. (A complete data dictionary is also available, see Appendix A: Related Documents).

7   Note that the household questionnaire used in LSMS_NICA_93 did not include an agricultural module. This was omitted due to the unstable land ownership situation and the extremely sensitive nature of such questions. It was feared that attempting to obtain this information would dramatically increase the refusal rate in rural areas (and, perhaps, cause problems for the interviewers themselves). The employment module, however, captures all income for agricultural employees and should approximate income for the self-employed in the agricultural sector.

[1] Household Roster:

This module collects basic demographic information on people in the household and a determination of whether or not an individual is a member of the household. The module was administered to the head of household or the head of household's spouse. Data for this section were collected for all people in the housing unit. Based on this information individuals are classified as household members or not.[4] Data in the remainder of the questionnaire are only collected for those who are household members.

Table 1: Household Questionnaire
1] Household Roster
2] Housing
3] Education
a] General
b] Vocational
c] Pre-school
4] Health
a] Children
b] General
c] Maternal
d] Disabilities
5. Employment
a] Activities
b] Principal Job, 7 days
c] Secondary Job, 7 days
d] Underemployment
e] Principal Job, 12 months
f] Secondary Job, 12 months
6. Migration
7. Independent Employment (Non-agricultural)
a] Identification
b] Expenditures/Income
c] Inventory and Capital
8. Consumption
a] 7 day reference period
b] 30 day reference period
c] 6 month reference period
d] 12 month reference period
9. Other Income

---

[4] A person is defined as a household member if s/he lived in the household for a minimum of three months during the previous 12 months. An exception is the head of household (as defined by the household itself) who is always classified as a household member regardless of length of absence. Domestic staff and paying boarders or all persons (except the household head) who were absent from the household for more than nine of the twelve months prior to the interview are not considered to be household members. Children under three months of age, of household members, are included even if they have not lived in the household.

[2] Housing:

In this module data were collected on the type, construction, size and legal situation of the housing unit. Expenditures associated with housing: rent (or estimate of rent if owner), electricity and other energy sources, water, phone were also gathered. There are also questions concerning the household's source of water, use of energy, type of sanitation and ownership of phones, televisions, radios and/or refrigerators. Three agricultural questions, designed to determine amount of agricultural land under cultivation by the household and the crops cultivated, were included. This module was administered to the head of household or his or her spouse.

[3] Education:

Part A: General education level: The three main areas of this section are: (i) basic literacy and educational attainment; (ii) current attendance in school and type of school; and (iii) costs associated with schooling. The section was administered to all household members age six and above.

Part B: Vocational/Training courses: Data were collected on attendance, type, expenditures, and relevance of such courses to present job. The section was administered to all household members age 15 and above.

Part C: Pre-School: Data were collected on attendance, type and expenditures. The section was administered for all children below the age of six (household members).

[4] Health:

Part A: Children's Health: This section includes data on vaccination records, the incidence of diarrhea in the 30 days prior to the interview, use of medical facilities and personnel and expenditures for the treatment of diarrhea. The section was administered for all children (members of household) below the age of six.

Part B: General Health: The section records data on illness and/or accident in the 30 days prior to the interview, use of medical facilities and personnel, hospitalization, medication, expenditures. The section was administered to all household members age six and above. It was also administered to all children (members of household) below the age of six who had a non-diarrheal disease in the 30 days prior to the interview.

Part C: Maternal Health: This section includes data on the number of children born to each woman and, if the women gave birth in the five years prior to the interview, data were collected on pre-natal care (timing, number of visits, medical facility and personnel used, expenditures), birth (location, medical personnel in attendance, expenditures) and breast-feeding and age of child when other liquids and foods introduced. This section was administered to all female household members age 15 to 49 inclusive.

Part D: Disabilities: Data collected on types of disabilities, severity and causes of them. This section was administered to all household members.

3

[5] Employment:

Part A: Activities of Household Members: The section identifies the activities of household members in the 7 days prior to the interview to determine if people were employed, unemployed (seeking work), or inactive. This section was administered to all household members age six and above.

Part B: Principal Job in Previous 7 Days: The section collects information on economic activity, sector and occupation, months, days and hours worked, seasonality, income (including all additional benefits in cash or kind), the presence of unions in the workplace, insurance and the size of the establishment. This section was administered to all household members age six and above who were classified as employed in Part A.

Part C: Secondary Job in Previous 7 Days: The section collects information on economic activity, sector and occupation, months, days and hours worked, seasonality, income (including all additional benefits in cash or kind), the presence of unions in the workplace, insurance and the size of the establishment, and income information on any additional jobs. The section was administered to all household members age six and above who indicated that they held a secondary job in Part B.

Part D: Underemployment: The data here concern any search for supplementary or substitute jobs. This section was administered to all household members age six and above who were classified as employed in Part A.

Part E: Principal Job Last 12 Months: The section collects information on economic activity, sector and occupation, hours worked per week, time in activity, seasonality, income, and additional benefits. The section was administered to all household members age six and older.

Part F: Secondary Job Last 12 Months: The data collected in this section are related to economic activity, sector and occupation, hours worked per week, time in activity, seasonality, income, and additional benefits. This section was administered to all household members age six and older.

[6] Migration:

This module includes data on the number of times a person moved in his or her lifetime, and details of the first and last move. The module was administered to all household members age 15 and older.

[7] Independent Employment (Non-agriculture) and/or businesses

Part A: Identification of Business: The section was designed to collect data on the economic activity of the business(es) owned by the household, and the number and identification of household members who work in the business. It was administered to the owner of the business (independently employed).

Part B: Expenditures and Income: This section collects data on the types of products/services produced, the length of time the business has existed, income from various sources,

expenditures for various items, number of non-household member employees, and insurance coverage. The section was administered to the owner of the business (independently employed).

Part C: Inventory and Capital: This section collects data on the ownership and value of various business equipment and similar items, and whether item is shared with household. The section was administered to owner of the business (independently employed).

[8] Consumption:

Part A: Expenditures in previous 7 days: The section is aimed at gathering data primarily on food expenditures for the seven days prior to interview. Data collected include: (i) whether item was purchased or obtained; (ii) how much of the item was purchased or obtained; (iii) how the item was obtained (purchase, home production, gift); and (iv) the value of the item obtained. Thirty-four items were included here. This section was administered to the head of the household or his or her spouse.

Part B: Expenditures in previous 30 days: The section is aimed primarily at gathering data on household and personal hygiene products for the 30 days prior to the interview. Data collected include: (i) whether item was purchased or obtained; (ii) how much of the item was purchased or obtained; (iii) how the item was obtained (purchase, home production, gift); and (iv) the value of the item obtained. Sixteen items were included here. This section was administered to the head of the household or his or her spouse.

Part C: Expenditures in previous 6 months: The section is aimed at gathering data on clothing, cookware, appliances, linens, furniture, housing repair and hotel expenditures for the six months prior to the interview. Data collected include: (i) whether item was purchased or obtained; (ii) how much of the item was purchased or obtained; (iii) how the item was obtained (purchase, home production, gift); and (iv) the value of the item obtained. Twenty-five items were included here. This section was administered to the head of the household or his or her spouse.

Part D: Transfer payments: Data collected refers to expenditures on transfer payments in previous 12 months: expenditures on pensions, alimony/childcare, financial support of individuals not living in the household, social security, taxes, association fees, donations, lotteries, weddings and funerals, domestic services, non-local transportation and others. Data were collected on whether the household spent money in any of the twelve categories, the amount of the last expenditure and how often expenditures were made annually. This section was administered to the head of the household or his or her spouse.

[9] Other Income:

The module was designed to collect data on income from non-labor sources such as: donations, lotteries, loans, interest on savings, remittances from within country and from abroad, inheritances, insurance and others. Data were collected on whether the household received income from any of the sources, the amount received and the frequency with which the amount was received. This module was administered to head of household or spouse.

8   **Anthropometric Questionnaire**: The anthropometric questionnaire was administered to all children under five years of age in the households. The following data were collected (The data dictionary for the household level questionnaire contains details of the anthropometric data, see Appendix A, Related Documents):

> Date measured
> Weight in kilograms
> Height (or length) in centimeters.

9   Note that information on the child's data of birth was collected in the household roster module. This date was verified at the time of the measurement. Where possible, the date was taken from the child's birth certificate.

10   **Community and Price Questionnaire**: The community questionnaire was administered to community leaders in all rural areas where household surveys were administered. It covered the following topics (a data dictionary for the community and price questionnaire is available, see Appendix A, Related Documents):

> Demographics (basic),
> Economic base of community,
> Transportation,
> Education facilities and use, and
> Health facilities

11   Additionally, in each rural community where this interview was carried out, prices were also collected for a basic consumer basket (58 items). Prices were collected in three stores or markets (when these existed). Where relevant, the brand names of the products were noted. **Note**: These data are of limited use due to errors in processing and data entry which make it impossible to merge much of the community data with household data. The price data are completely unuseable.

III.   SAMPLE

12   The following is a brief summary of the sample and sample frame used for the LSMS_NICA_93 and draws heavily on the report of Edmundo Berumen who designed the sample. Given the lack of a recent census, the design of the sample is somewhat more complex than that for other, similar surveys. (A complete copy of the document describing the sample design is available. See Appendix A, Related Documents.)

13   The size of the sample was preset at 3,600 households. The two criteria used for determining the overall sample size were: (i) to obtain a sufficient number of households for the sample to be representative at the urban and rural levels within the country's seven regions (this is one reason that the sample is not self-weighting, see paragraph 25); and (ii) to ensure a sufficient number of children under age five to carry out a detailed anthropometric analysis.

14   The political-territorial division of the country which existed in the early 1990s consisted of nine regions. For purposes of the LSMS_NICA_93, the first six regions were used and, because of the extremely low population density in the remaining three regions (RAAS, RAAN and Río San Juan) these were combined to form the seventh region used in the study. Each region used in the LSMS_NICA_93 is comprised of one to four departments. A list of the regions used in the survey can be found in Table 2.

Table 2: REGIONS	
REGION	DEPARTMENTS
Segovias	Estelí, Madríz, and Nueva Segovia
Western	León and Chinandega
Managua	Managua
Southern	Granada, Masaya, Carazo, and Rivas
Central	Boaco and Chontales
Northern	Jinotega and Matagalpa
Atlantic Coast	Región Autonoma Atlántica Norte (RAAN) Región Autonoma Atlántica Sur (RAAS) Río San Juan

15   INEC's experiences with surveys in the 1980s had indicated that a refusal rate of 15 percent could be expected. As it was decided not to allow substitution of housing units in the field it was necessary to select a sample of 4,200 housing units to ensure that a sample of 3,600 households were included in the survey (4,200 * .85 = 3,600).

16   **Sample Frame**: When the sample was designed for the LSMS_NICA_93 there were two alternative sources for use as a sample frame. The first was the up-dated maps prepared for the Population and Housing Census which had been planned, but not carried out, in 1982. The second source was the voter registration lists for all persons aged 16 and older which had been prepared for the elections of 1990.

17   Both alternatives had advantages and disadvantages. The emphasis in the 1982 updating of census maps was on the quality of mapping. Nevertheless, the population data were out-of-date and INEC ran out of resources before completing the division of the country into census sectors. An additional drawback was that the data were not in electronic form.

18   In contrast to the census preparations, the compilation of voter registration lists emphasized complete coverage of the voting population. Precision in mapping and complete mapping of the national territory were of secondary importance. The data on voters were disaggregated at the level of Electoral Committees (Junta Receptoras de Votos - JRV) and was in electronic form.

19   As neither source was, by itself, adequate for use as a sample frame, it was decided to use a combination of the two sources. The population projections of INEC for 1991 (based on the previous census and updated in 1982) were used to determine the size of the sample within each region based on rural and urban areas. The JRV were used for the selection of enumeration areas within each region. Using this strategy, the enumeration areas were selected (JRV or groups of them) and then updated in the field, both in terms of geographic limits and enumeration of housing and other buildings within each area.[5]

---

[5] The fact that a complete enumeration of housing units and households was carried out within each JRV means that even if specific types of households or individuals were less likely to register to vote than others, these non-registered individuals or households had the same probability of being selected for the survey as other households.

20	**Sample Distribution**: Given the absence of estimations of variance at the regional level for the principal variables of the survey, it was decided to distribute the sample equally among all regions, independent of the population size of the regions. But, as there was particular interest in the Department of Managua (where the capital is located) it was decided to first allocate to Managua a proportion of interviews close to its proportion of national population (according to 1991 estimates) and then distribute the rest in equal fashion among the remaining six regions. Within each region, the sample was distributed proportionally in the urban and rural zones as is shown in Table 3.

Table 3: DISTRIBUTION OF THE SAMPLE (Number of Households)			
REGION	URBAN	RURAL	TOTAL
Segovias	240	280	520
Western	320	200	520
Managua	960	120	1080
Southern	320	200	520
Central	200	320	520
Northern	200	320	520
Atlantic Coast	240	280	520
-------	-----	----	-----
TOTAL	2480	1720	4200
	59%	41%	100%

21	**Sample Procedure**: Within each region, groups of contiguous municipalities were joined until an area encompassing a minimum of 60,000 and, on average, 80,000 inhabitants was created (according to INEC's population projections for 1991), except for the more populated municipalities which, by themselves, exceed this number. In the city of Managua contiguous JRVs were grouped until the desired population was reached. In the rest of the department of Managua the procedure was the same as in all other municipalities. In this fashion a total of 45 strata were formed. The JRV within each of these strata were categorized as urban or rural (except in the city of Managua where only one rural strata which groups the rural zones of the ten strata created was formed). Thus, there are a total of 81 substrata: 45 urban and 36 rural.

22	Once the JRV within each strata were organized, a rule of association was chosen which insured that all JRV had at least 220 registered voters and approximately 60 or more housing units (based on an estimate of 3.7 people 16 years or older per housing unit). Any JRV which did not fulfill this requirement, was joined to the previous JRV in the list (which was contiguous and pertained to the same strata). This association rule guaranteed that the JRV (or group of them) selected in the sample had enough housing units to support at least six surveys of the same size as the LSMS.[6]

---

[6] The sample frame was designed to support up to six household surveys.

23   In total, 420 JRV were selected with probability proportional to size (measured by the number of registered voters): 247 urban and 173 rural. The selection of the JRV was carried out independently within each region by ordering the JRV according to the strata and urban and rural substrata. The distribution of the JRV by regions can be seen in Table 4.

24   The selected JRV were updated in the field by INEC both in terms of mapping as well as the enumeration of the housing units and other buildings which they

Table 4: DISTRIBUTION OF JRV[1]			
REGION	URBAN	RURAL	TOTAL
Segovias	25	27	52
Western	32	20	52
Managua	96	12	108
Southern	31	21	52
Central	21	31	52
Northern	17	35	52
Atlantic Coast	25	27	52
-------	-----	----	-----
TOTAL	247	173	420
	59%	41%	100%

1/ Ten housing units were selected in each JRV.

contained. Within each enumeration area (JRV or group of them) in the sample, 10 housing units were selected randomly for the survey in two compact clusters of 5 housing units each.

25   Estimation Process: The sample is not self-weighting. There are different weights (expansion factors) for each JRV in the sample. This is due to the absence of good (up-to-date) population estimates for selecting the JRVs and the need to control the sample size. The expansion factors are included in the data set for each JRV in the sample.[7]

26   In general the probability of selecting JRV i for region h is given by:

$$p(\text{selecting JRV}_i \text{ in region h}) = \frac{K_h \, M_{hi}}{M_h} * \frac{2}{S_{hi}}$$

Where K_h is the number of JRVs (or group of JRVs) selected in region h, M_{hi} is the number of registered voters in the ith JRV of the region h, M_h is the sum of <u>all</u> the M_{hi} of the region h (not only the selected ones)[8] and S_{hi} is the number of compact segments formed in the ith JRV of region h. The expansion factors are determined by the inverse of these probabilities of selection.

27   The sample is also divided into four replicates. Therefore, the estimation of sample errors (variances) can be obtained simply by obtaining the four estimates of the parameter of interest based on the four replicates and contrasting each one of these with the average of the four. The formulas of estimation of means,

---

[7] Note that the sample weights and expansion factors were not available until late 1995 when much of the original analysis of the data was finished. The Poverty Assessment, as well as some other earlier studies, used rough weights based on total regional population estimates.

[8] In other words, M_{hi} is the population of registered voters in region h.

percentages, rates, proportions and ratios are the usual ones with the variables expanded by the corresponding expansion factors. The replicate information is also contained in the data set.

28   **Final Sample Size**: As can be seen in Table 5, the actual refusal rate for LSMS_NICA_93 was much lower than that predicted which led to an increase in the total sample size. Additionally, the existence of multiple households within housing units also served to increase the sample size.

					NON	
Table 5: INTERVIEWS CARRIED OUT BY REGION					
REGION	PLANNED SAMPLE	EXTRA INTER- VIEWS	NON-[1,2] RESPONSE	TOTAL INTER- VIEWS	NON RESPONSE/ ORIGINAL SAMPLE
Segovias	520	55	18	557	3.46%
Western	520	89	22	587	4.23%
Managua	1080	126	85	1121	7.87%
Southern	520	37	48	509	9.23%
Central	520	84	16	588	3.08%
Northern	520	97	63	554	12.12%
Atlantic Coast	520	51	33	538	6.35%
TOTAL	4200	539	285	4454	6.62%

[1] Non-response includes: refusal of interview by household, households in areas inaccessible to team due to violence, uninhabited housing units, household absent and/or mapping errors.
[2] The non-response rates for the Cental, Northern and Atlantic Coast Regions are calculations and may not be precise.

29   The total number of informants, both household members and non-members as well as the breakdown by gender and region can be found in Table 6.

Table 6: Number of Informants, by Region and Gender					
LOCATION	HSHOLDS	MEMBER	MALES	FEMALE	TOTAL
SEGOVIAS	558	75	1617	1636	3328
Urban	247	44	705	750	1499
Rural	311	31	912	886	1829
WESTERN	587	101	1539	1626	3226
Urban	356	60	899	938	1897
Rural	231	41	640	688	1369
MANAGUA	1121	203	2880	2964	6047
Urban	964	185	2406	2538	5129
Rural	157	18	474	426	918
SOUTHERN	509	72	1510	1565	3147
Urban	316	54	881	939	1874
Rural	193	18	629	626	1273
CENTRAL	587	71	1604	1721	3396
Urban	235	39	578	679	1296
Rural	352	32	1026	1042	2100
NORTHERN	554	59	1503	1488	3050
Urban	183	34	442	482	958
Rural	371	25	1061	1006	2092
ATLANTIC	538	44	1452	1432	2928
Urban	273	18	653	716	1387
Rural	265	26	799	716	1541
TOTAL	4454	635	12105	12422	25162

IV.   FIELD WORK

30   Interviewing for the LSMS_NICA_93 was done by region, with all of the interviewing completed in one region before a new region was started. Work began in the Southern Region of the country and ended in the Atlantic Coast Region. A regional field office was set up in a central location in the region where work was being done and all questionnaires were processed through these field offices. Coding of the few open-ended questions was done by INEC staff, data were entered on computers with the data entry program running automatic checks on range and consistency errors. Based on reports from the data entry, both the brigade-level supervisors and INEC general supervisory staff determined if questionnaires were complete or if follow-up visits were needed to the households.

31   Nine interviewing teams (brigades) did the interviewing at the household level.[9] Each interview team consisted of four interviewers, one supervisor and one driver (with vehicle). Interviewers and supervisors participated in a two-week training course and carried out the pilot study. Immediately prior to the actual field work, interviewers were given an additional week of training which focused on the problem areas identified in the pilot study and all changes which had been made to the questionnaire. Supervisors were given an additional week of training in using maps, locating clusters and housing units and specific supervision duties.

32   Working closely with the brigades were 9 data entry people. These people received one week of training and also entered the data from the pilot study. Originally the data entry personnel were assigned to work only with one brigade and to enter the data from the questionnaires completed by this brigade. After the first region was completed, however, it was determined that the work progressed more quickly if incoming questionnaires were assigned to the next available data entry person and not to any specific person.

33   In addition to the brigades, two other types of teams were fielded. Two anthropometric teams of four anthropometrists and one supervisor were formed. These people received one week of training by an expert in anthropometric measurement techniques (Irving Schorr) and visited hospitals and pre-schools to practice measuring children prior to the field work. The anthropometric personnel did not, however, participate in the pilot study. During the survey the anthropometric teams visited households after the household interview had been carried out. Only households with children under the age of five were visited.

34   The last 'team' consisted of two people who were responsible for the community and price questionnaires. The community questionnaire was administered to community leaders in rural areas in the sample and the price questionnaires were used to gather price information on a basic basket of products in up to three locations in the communities. The community/price team received one week of training and, although they did not participate in the pilot study, they were sent to various communities, prior to the field work, to get some field experience.

35   The LSMS_NICA_93 was carried out by the staff of the Survey Department of INEC. The staff involved in the actual implementation of the LSMS_NICA_93 consisted of: (i) three persons responsible for verifying the questionnaires and coding the few questions that were not pre-coded; (ii) one person responsible for the overall organization of the field work; (iii) one person responsible for supervising the data entry process; and (iv) one person, the Director of the Department, responsible for overall supervision and coordination. All other persons involved in the field work[10] were short-term employees hired for this specific survey.

36   The household interviews were designed to be completed in one visit to the household.[11] Once an interview was completed, supervisors were required to visit a random selection of households to check that the

---

  [9] Originally ten teams were to be formed. But, due to the long lag time between the pilot study and the actual field work, several interviewers dropped out. It was decided that it was more efficient to start the field work with the available interviewers rather than wait for new ones to be trained. However, concurrent to the start of the field work a small group of new interviewers were trained and began working in an 'apprenticeship' role. Once it was determined that they were capable of carrying out interviews on their own, they became part of the interviewing teams.
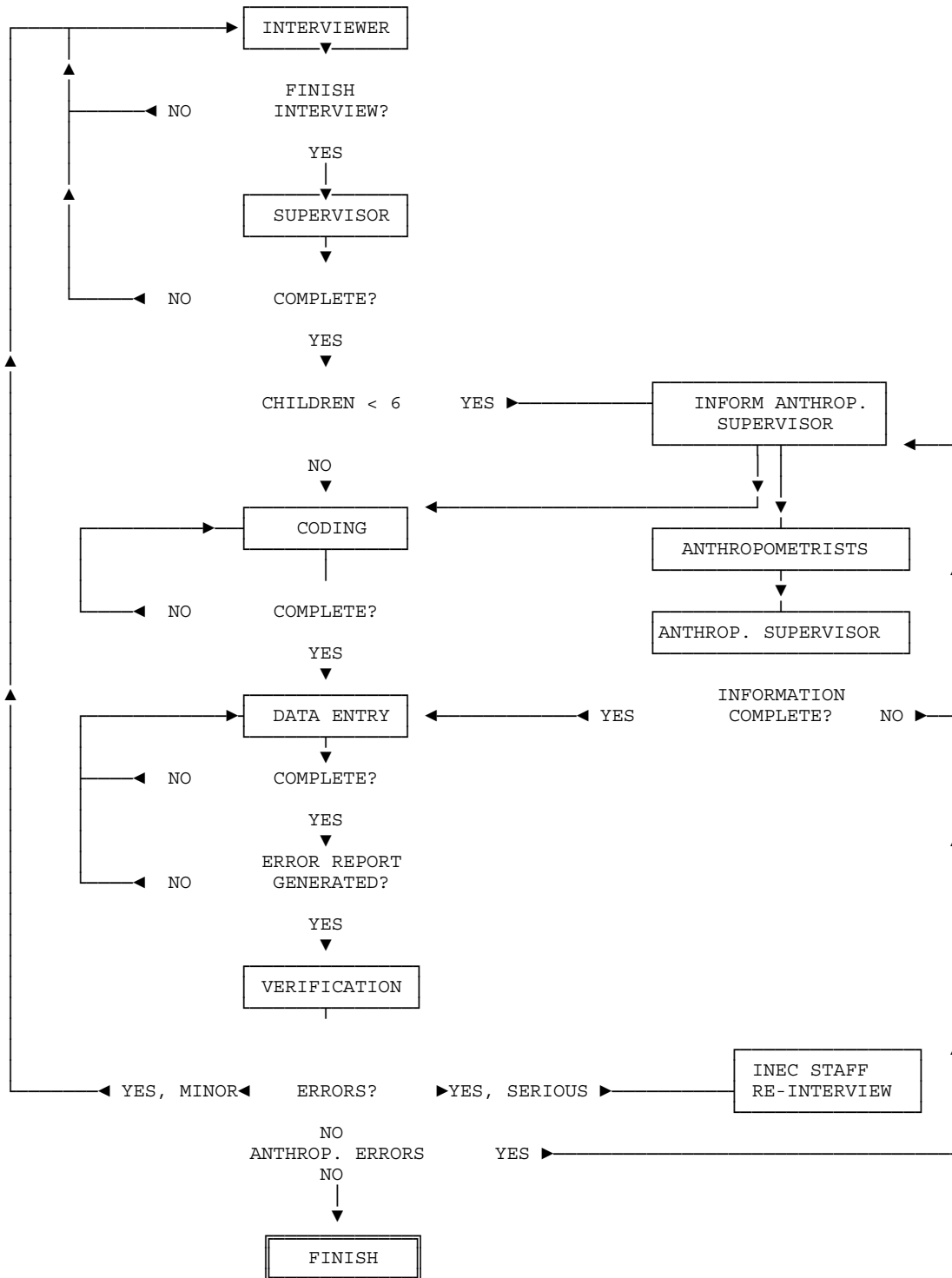
  [10] I.e., household interviewers, supervisors, data entry personnel, anthropometrists, anthropometric supervisors and community interviewers.

  [11] Only one visit was made to each household instead of two visits two weeks apart as has been done in other LSMS

interview had been done properly and to complete a short questionnaire designed to check the veracity of the data recorded by the interviewer. Records on the extent to which such supervisory activities were carried out are not, however, available. The overall flow of field work is outlined in Figure 1. Copies of the manuals used by the various teams are available (see Appendix A, Related Documents).

surveys. This decision was based on two considerations. First, the dearth of policy relevant, social-sector data in Nicaragua was creating problems both for the GON and the donor community in their poverty alleviation measures. Thus the speed with which the data could be collected was extremely important. Second, given INEC's lack of resources and other problems, the LSMS_NICA_93 was one of the more costly surveys. Some attempt was needed to hold costs down as much as possible. The standard two-interview process is both more time consuming and more costly than doing one interview. For these reasons, it was decided to do one interview, and obtain all of the data at that time. Note, however, that with either the one-visit or two-visit design, in reality more visits are usually made to the household to ensure that all members are interviewed.

Figure 1: Information and Supervision Flow Chart

INTERVIEWER

FINISH INTERVIEW?   ◄ NO

YES

SUPERVISOR

COMPLETE?   ◄ NO

YES

CHILDREN < 6   YES ►   INFORM ANTHROP. SUPERVISOR

NO

CODING   ◄   ANTHROPOMETRISTS

COMPLETE?   ◄ NO   ANTHROP. SUPERVISOR

YES

DATA ENTRY   ◄ YES   INFORMATION COMPLETE?   NO ►

COMPLETE?   ◄ NO

YES

ERROR REPORT GENERATED?   ◄ NO

YES

VERIFICATION

◄ YES, MINOR ◄   ERRORS?   ►YES, SERIOUS ►   INEC STAFF RE-INTERVIEW

NO
ANTHROP. ERRORS   YES ►
NO

FINISH

14

V.   DATA QUALITY

37   The overall data quality is considered to be good. However, several specific points need to be kept in mind when using the data set.

38   Data Set: First, when the analysis of poverty in Nicaragua was done (see Appendix D), approximately 9 percent of the households had to be dropped from the analysis due to incomplete expenditure data in one or more of the four modules of the household questionnaire where expenditure data was requested.[12] This result was based on the original data set from Nicaragua which had not been re-entered nor checked against the original questionnaire in the case of missing data. The documentation produced here refers to this 'uncleaned' data set. While it would be advantageous to use the later 'cleaned' data set as it should contain fewer inconsistencies and missing data, this data set is not distributed as: (i) there is no documentation of the cleaning process and it is unknown if any additional process was carried out beyond simply fixing data entry errors (such as replacement of outliers, imputations and the like); and (ii) the 'cleaned' data sets available are incomplete: missing both modules (anthropometrics and businesses) and specific variables (estimated rent payments and in-kind rent payments for example).

39   Southern Region: The second point that should be kept in mind is that the first region where interviewing occurred was the Southern Region followed by the Managua Region. In the Southern Region, the data entry process was not completely operational which means that many interviewing errors were not caught until it was too late to return to the household to for further interviewing. By the end of the time in the Managua Region, the data entry process was functioning. It is expected that the data in the following regions (Western, Segovias, Northern, Central and Atlantic Coast) are of higher quality than those of the first two regions.

40   Health: The third point that must be kept in mind refers to the Health Section, Part C: Maternal Health. Question #15 which reads: "At what age did you begin to give [Child's Name] liquids such as water, unpasteurized milk, whole milk, powdered milk, processed or pasteurized milk, soy milk etc?" was changed after the first region (Southern) was completed. In all other regions this question was changed to read: " At what age did you begin to give [Child's Name] liquids such as unpasteurized milk, whole milk, powdered milk, processed or pasteurized milk, soy milk etc?"   In other words, the latter version of the question does not ask the age at which water or other water-based liquids were introduced, only milk in its various forms.

41   Community and Price: For the community questionnaire, errors exist in the identification variables. This makes it impossible to match the household and community questionnaires in many communities. In other words, analysis can be done on the community questionnaires alone, but use of these variables with the household data set will be limited. Also the price data are not available due to a glitch in the date entry operation. It is not known if plans to re-enter and disseminate these data exist.

---

[12] To calculate total and per capita expenditures, expenditure data were drawn from four of the nine modules of the questionnaire. Due to the importance of health and education expenditure data for the analysis in the poverty profile, any missing data in these two categories, for any household member, caused the questionnaire to be dropped from the analysis. For food, personal and household goods, households were dropped if more than four items in total had missing data or if data were incomplete for major expenditure items.   In other words, a questionnaire that lacked data on expenditures on matches or ice was included while one that lacked data on expenditures for rice or beans (major food groups) were dropped.

VI.   CONSTRUCTED FILE: CONSUMPTION AGGREGATES

42   To carry out the analysis for the profile of poverty in Nicaragua (see Appendix D), a series of consumption aggregates were created for each household. These variables are available to the public on the condition that the World Bank and its staff are under no obligation to provide further services to users. These data files reflect methodological choices of individual researchers. Some users may prefer to calculate their own versions of these constructed files based on the original data. Note that these variables are only available for the households which were included in the poverty profile (approximately 91 percent of all households (see paragraph 38)). Additionally, the consumption aggregates as well as the nutrition aggregates were constructed using the data set before final cleaning was done. Due to this, if the consumption aggregates had been recalculated, some small differences could be expected to have occured. Also, it might have been possible to construct a consumption aggregate for slightly more households.

43   The variables found in the constructed data file available for the LSMS_NICA_93 data are found in Table 7. A brief summary of how each variable was constructed is presented in the following paragraphs. Further information can be found in the poverty assessment of Nicaragua. (A data dictionary for this file is available, see Appendix A, Related Documents.)

44   LOC: This is simply a location variable that was created by combing the regional variable (REGION) and the urban/rural variable (AREA). It was constructed by multiplying the regional variable by 10 and adding the urban/rural variable.

Table 7: Consumption Aggregates and Poverty Indicators	
QUESTNUM	Questionnaire Number
REGION	Region of country
AREA	Urban/Rural
LOC	(REGION*10)+AREA
JRV	Junta Receptora de Votos
POVERTY	Poverty Group
DECILE	Expenditure Decile
PCFOOD[1]	Per capita expenditures on Food
PC30	Total per capita expenditures
PCHOUSE	Per capita housing expenditures
PCHEALTH	Per capita health expenditures
PCEDUC	Per capita education expenditures
PC8EXP	Per capita expenditures from Section 8 of questionnaire
NOMEMB	Number of household members

Note: All expenditures are for 30 days.

45   Calculation of Expenditures: All expenditure data for food and goods includes both those items actually purchased by the household, as well as those items obtained through gifts, donations and home-production. For all goods and items which were not actually purchased, the respondent was asked to give a monetary value for the good, i.e., what it would have cost if the person had to purchase it.

46   As noted previously, the reference period for which data were collected varied based on the nature of the good in question. For food expenditures, the most frequent, a seven day reference period was used. For other goods which are consumed or purchased less often, the reference periods ranged from 30 days, to six months to 12 months.

47   All of the expenditure data were recalculated to conform to one reference period: 30 days. Expenditures collected for a seven day reference period were multiplied by 4.3 to obtain the thirty day figure, while those collected for a thirty day reference period were left as collected. The expenditures collected for the six month and 12 month period required slightly more adjustment to take into account the 20 percent devaluation of the cordoba at the end of January of 1993.

48   To obtain these expenditures in current cordobas it is necessary to inflate all those expenditures which occurred prior to February 1. Unfortunately, the exact timing of the expenditures is unknown. In order to take account of the devaluation, it was assumed that expenditures were continuous; in other words, the same amount was spent every month.[13] Using this assumption, expenditures for 6 (12) months were divided by 6 (12) and the expenditures in months prior to February 1 were inflated. These expenditures were then added to the expenditures for the months after February 1. The sum of the expenditures was then divided by 6 (12) to provide a monthly expenditure for the products or services.

49   This adjustment for devaluation was not carried out for those categories of education expenditures which were collected for a 12 month reference period. Due to the timing of the data collection process, it was not possible to determine if figures for tuition were for the previous school year or for the present one.[14] Thus some formal education expenditures may be underestimated. The same is true for vocational training and for pre-school expenditures.

50   For the calculation of total housing expenditures an additional step was required. There are five different expenditures related to housing for which data were collected: water, lights, cooking fuel, phone and rent. For the first four, expenditures are collected for a day, week or month reference period. To obtain 30 day expenditures these expenditures are simply multiplied by the appropriate number.

51   While this technique could be used to obtain 30 day rental expenditures for all those who rent, this leaves the vast majority of households without any rental expenditure; only 3.5 percent of the households are renters. To obtain the use value of the housing of the household an imputed rent figure had to be estimated for non-rental households.

52   To calculate an imputed rent figure for all households, the standard practice is to estimate a hedonic rent function correcting for selectivity, based on renters. Given the extremely small number of renters in the sample, however, this technique could not be used for the present analysis. Instead, imputed rents were calculated for urban and rural areas based on what homeowners estimated the rental value of their property to be. Imputed rents were computed separately for urban and rural households.

53   All per capita figures were calculated by summing the expenditures of the household for a given category and dividing by the number of household members. Thus, the per capita education expenditures variable (**PCEDUC**) does not indicate what a given student actually spent on education during the reference period. This must be calculated separately with the data. The same is true for health expenditures (**PCHEALTH**), housing (**PCHOUSE**), food (**PCFOOD**) and for all other expenditures (**PC8EXP**). Note also

---

[13]Clearly this is a logical assumption for some expenditures such as social security payments. For others, however, such as weddings and funerals the assumption is less valid.

[14]The data collection process began in February and finished in June. The academic year began in March.

that food expenditures refers to the actual expenditures during the reference period. There is no way to determine actual consumption of food for the household let alone for individuals within the household.

54   **POVERTY**: This variable is based on the poverty line which was constructed for Nicaragua. The variable takes on a value of one if the per capita 30 day expenditures (**PC30**) are below the extreme poverty line of C$101.32.[15] If the per capita expenditures are above the extreme poverty line, but below the overall poverty line of C$214.47, the variable takes on a value of 2. For individuals with per capita expenditures above the overall poverty line, the variable is coded as '3': non-poor. (A detailed description of how the poverty line was constructed can be found in Appendix B which summarizes the explanation found in the poverty assessment.)

55   **DECILES**: This variable is based on total per capita expenditures (**PC30**). It indicates the expenditure decile of the individual, and, by definition, each member of that person's household.

56   **NOMEMB**: This variable is the number of household members.

VII.   CONSTRUCTED FILE: CHILD NUTRITION

57   Using the data from the anthropometric measurements and the section on breastfeeding of the last born child, a special anthropometric and nutrition file was created. This includes z-scores for the standard nutrition indicators: weight-for-age, height-for-age and weight-for-height calculated using the anthropometric software ANTHRO. Additionally, for each child in the data set, variables on breastfeeding, age of introduction of other liquids and solid foods, and age of weaning are included. The specific variables in the file are outlined in Table 8. (A data dictionary for this file is available, see Appendix A, Related Documents.)

---

[15] The exchange rate at the time of the survey was US$ 1.00 = C. 6.00.

Table 8: Constructed Nutrition Variables

ID	Unique child identification number:
GENDER	Child's gender
HAZ	Height-for-age z-score, calculated by ANTHRO using exact age.
WAZ	Weight-for-age z-score, calculated by ANTHRO using exact age.
WAZ	Weight-for-height z-score, calculated by ANTHRO using exact age.
AGECOMP	Age in months, compute by ANTHRO (measurement date - birthdate).
REGION	Region
STUNTED	If HAZ < -2.0
UNDERWT	Underweight: If WAZ < -2.0
WASTED	If WHZ < -2.0
EVEREXBF	Last born child ever exclusively breast-fed?
EVERMILK	Last born child ever consumed other liquids/artifical milk?
EVERFOOD	Last born child ever consumed other solid food?
STILLBF	Is last born child currently breastfed?
EVERBF	Was last born child ever breastfed?
REXBFMO	Duration of exclusive breastfeeding in months for lst born child.
RMILKMO	Age in months that other liquids/milks were introduced to last born child.
RFOODMO	Age in months that solid foods were introduced to last born child.
RWEANMO	Age in months that last born child was weaned.

APPENDIX A
RELATED DOCUMENTS

1. Household Questionnaire (in Spanish)

2. Anthropometric Questionnaire (in Spanish)

3. Community Questionnaire (in Spanish)

4. Data Dictionary: Household Questionnaire (includes Anthropometric Information also) (in English)

    Community Questionnaire (in English)

    Constructed Consumption Aggregates and Poverty Indicators (in English)

    Constructed Nutrition Variables (in English)

    Expansion Factors and Sample Information (in English)

5. Sample Design (in Spanish)

6. Interviewer's Manual (in Spanish)

7. Supervisor's Manual (in Spanish)

8. Data Entry Manual (in Spanish)

9. Abstract (in Spanish)

1   Poverty Line: The poverty line is defined as the level of total per capita monthly (30 day) expenditures at which an individual attains the minimum daily caloric requirement. Expenditures are used, as opposed to income, for three reasons: i) expenditure data are more accurate than income data; ii) income data are notoriously unreliable for workers in agriculture and the informal sectors of the economy; and iii) expenditures tend to vary less than income and are therefore more reflective of actual living standards.

2   The poverty line is estimated at the point where per capita expenditures are sufficient to obtain the minimum caloric requirements.[16] The estimation is based on Reutlinger and Selowsky [1976] and takes the form:

$$\ln C_j = \alpha + \beta \ln X_j + e_j \qquad \text{[Eq. 1]}$$

Where:
      $C =$ natural log of per capita (adult equivalent) caloric intake for the jth person
      $X =$ natural log of per capita total expenditures for the jth person

3   Note that the expenditure level at which the caloric requirements are met is <u>not</u> the total cost of calories nor the cost of a specific food basket. Instead it is the level of expenditures at which an individual, taking into account all types of expenditures that s/he must make, obtains the minimum required level of calories per day.[17]

4   Extreme Poverty Line: In contrast to the poverty line outlined above, the extreme poverty line is defined as the level of per capita <u>food</u> expenditures required to obtain the daily minimum caloric requirement. Essentially, this is the cost of the minimum daily adult requirement of calories. Below this level of expenditures individuals cannot, even if total income is spent on food, maintain the needed level of caloric consumption. This is estimated by:

$$\ln C_j = \alpha + \beta \ln F_j + e_j \qquad \text{[Eq. 2]}$$

Where:
      $C =$ natural log of per capita (adult equivalent) caloric intake for the jth person
      $F =$ natural log of per capita food expenditures for the jth person

---

[16]All data on caloric consumption and expenditures for food and all other goods and services were collected at the household level. To obtain an individual's level of calories and expenditures the household consumption and expenditures must be adjusted for household composition. Individual level expenditures are per capita (total household expenditures divided by the number of people in the household). Given the biologically different caloric requirements of different age groups, individual level caloric requirements are based on an adult equivalency scale (household adjusted equivalents) with the following weights: adults = 1, children 0-9 = .61 and young adults 10-17 = .91.

[17] No effort was made to assess other aspects of nutrition beyond total calories. While grams of protein is often included in calculations of the adequacy of a specific diet, research has shown that an adequate intake of calories is highly correlated with adequate protein intake and thus a separate analysis of protein in not required.

5   The extreme poverty line represents an absolute level of food deprivation; if per capita expenditures are not above this line, the individual cannot attain the minimum levels of calories required for full physical activity. Like the model for the poverty line, this estimation of the extreme poverty line does not make any assumptions about food consumption preferences and patterns.

6   Estimating the two poverty lines defined above required three steps: i) determination of the caloric intake requirement for Nicaragua; ii) calculation of per adult equivalent caloric intake; and iii) calculation of per capita expenditures. The procedures used to make these calculations are outlined briefly in the remainder of this Appendix.

7   Caloric Requirements: While there are various figures for the adult caloric requirements in Nicaragua, the Nutrition Program of Nicaragua (PAN) recommends that 1,850 calories should come from a list of eleven food products and 2,188 from a list of 16 products [Vio, 1990] for example, there does not appear to be one set figure for total calories per day in use in the country.

8   The minimum caloric requirement for adults per day was calculated here using the methodology of the Food and Agriculture Organization.[18] Caloric requirements were calculated for different age groups and for urban and rural areas. The age distinction is important due to the different physical requirements at different stages of development. Levels were calculated separately for urban and rural areas to take into account the significantly higher requirements of rural populations due to their higher levels of physical activity. The minimum daily caloric requirement for adults is a weighted average of the urban and rural requirements. Weighting the 'national' caloric requirement by urban and rural allows the model to take into account population shifts (from rural to urban areas for example) in future years and hence will be useful for comparisons over time.

9   Based on the calculations described above, the estimated minimum daily caloric requirements for an adult in Nicaragua is 2,226. Children under 10 require 61 percent of this requirement and those from age ten through 17 require 91 percent.

10   Calculations of Caloric Intake: Individual caloric intake was calculated in two steps. First, the caloric intake of the household was obtained by converting the data on the quantities and types of foods obtained[19] by the household in the 7 day reference period into caloric equivalents. Caloric equivalents were based on the calorie conversion tables included in "Food Composition Table for Use in Latin America."[20] Additional information was obtained from Dirección General De Atención Materno Infantil, Mexico.

---

[18]FAO, Human Energy Requirements: A Manual for Planners and Nutritionists, 1989. The software provided with this manual was used to calculate caloric requirements for Nicaragua.

[19] Note that caloric intake is based on the amount of calories obtained by the household, whether purchased, obtained through gifts or donations, or coming from household production. This latter category is particularly important in rural areas where a significant proportion of a household's calories may come from home or own-farm production.

[20] This study was carried out by the Institute of Nutrition of Central America and Panama (INCAP) and The Interdepartmental Committee on Nutrition for National Defense, National Institute of Health, U.S.A., June 1961.

11   In the second step, household total caloric intake was converted to individual level intake by dividing by the number of adult equivalency units in the household. This figure was then multiplied by 4.3 to give the per person caloric intake for a thirty day period.

12   <u>Calculations of Total Per Capita Expenditures</u>: Total expenditures include all expenditures made by the household, including: food, clothing, personal hygiene products, cleaning products, household goods and services, transfers such as taxes, association fees, alimony, transportation, education, health care and housing. For all goods and services obtained through some means other than monetary purchase such as home production, gifts and barter the respondent was asked to state the monetary value of the good or service in question. In this manner the value of all consumption by the household is included in the total expenditure figure.[21]

13   All expenditure data were adjusted to reflect a thirty day reference period. See paragraphs 47-49 in the main body of this document for details on how these calculations were made. Once this adjustment was made, a total household expenditure figure was obtained by summing all expenditures. This was then divided by the number of household members to provided a per capita expenditure figure.

---

[21] Expenditure data were not adjusted for any regional price differences which might exist.

1   The data set is available in ASCII, SAS Portable and STATA formats. SPSS versions may also be obtained. Table C.1 provides the names and list of contents of the data files. Note that all files can be merged using unique identifiers formed in the following manner:

Household Level data ==>   Join files by QUESTNUM (number of the questionnaire)

Individual Level data==>   Join files by a combination of QUESTNUM and IDCODE (household member identification code).   Create variable for merging: (QUESTNUM*100)+IDCODE.

Community Level data==   Join files by a combination of Municipality code and JRV

Expansion Factor data==>   Join files by a combination of Municipality code and JRV.

Table C1: DATA FILES			
NAME OF FILE[1]	CONTENTS	UNIT OF ANALYSIS	NUMBER OF VARIABLES[2]
DEMO	Household Roster, Section 1	Individual	28
HOUS	Housing, Section 2, Parts A,B	Household	69
EDUC	Education, Section 3, Parts A,B,C	Individual	47
HLTH	Health, Section 4, Parts A,B,C,D	Individual	97
EMPL	Employment, Section 5, Parts A,B,C,D,E,F	Individual	100
MIGR	Migration, Section 6	Individual	24
INDJ	Independent Employment, Section 7, Parts A,B,C	Business	70
CN7D	Consumption 7 days, Section 8, Part A	Household	171
CN30D	Consumption 30 days, Section 8, Part B	Household	73
CN6M	Consumption 6 months, Section 8, Part C	Household	116
CN12M	Consumption 12 months, Section 8, Part D	Household	59
OTHY	Other Income, Section 9	Household	60
ANTH	Anthropometric Data	Individual	22
COMM	Community Level Data	Community	137
CATEG_1	Constructed Variables	Household	13
NUTR	Constructed Nutrition Variables	Child	21
EXP_FACT	Sample information, expansion factors	JRV	18
NEWCOMP	Final expansion factors at the household level (use these expansion factors for any analyses)	Household	27

Note: All data file names will have extensions which indicate their format (i.e *.ASC for ASCII files, *.SSP for SAS export files, etc.)
1/ All files contain household (and member) identification codes to enable merging of files.
2/ The number of variables is based on the ASCII files. In SPSS/PC+, STATA and SAS files, additional identification variables used by the software are included.

2.   The Nicaragua data are the sole property of the Government of Nicaragua. The World Bank has been provided with unrestricted use of the LSMS_NICA_93 and is authorized by the Government of Nicaragua to transfer the LSMS_NICA_93 data to third parties. The data from the LSMS_NICA_93 will be made available for use by the general research community subject to the following restrictions:

a)   data users shall not transfer the data to third parties;
b)   in all uses of the data, due recognition of the National Institute of Statistics and Census of Nicaragua as the source of the data shall be made; and
c)   the researcher shall make copies of all publications stemming from the data available to the National Institute of Statistics and Census

Director General
National Institute of Statistics and Census
Fte. Hospital Lenin Fonseca
Managua, Nicaragua

3.   The data and documentation for the LSMS_NICA_93 can be downloaded free of charge from the LSMS web site, the address for which is:

http://www.worldbank.org/lsms/lsmshome.html

4.   Requests for the LSMS_NICA_93 data can also be made directly to the LSMS Office and should include a brief description of the intended research:

Living Standards Measurement Study
Development Research Group - Poverty
The World Bank
MSN MC3-306
1818 H Street, NW
Washington, DC 20433
USA
e-mail: lsms@worldbank.org

5.   There is a processing fee associated with the distribution of data sets through the LSMS Office. The World Bank provides them on CD in SAS portable (XPORT engine), STATA, SPSS and ASCII formats. The Development Research Group, Poverty Team of the World Bank requests copies of all reports and documents resulting from research that uses the data. The researcher should further note that once received, the data cannot be passed on to a third party for any reason. Other researchers must contact the World Bank directly for access to these data. Any infringement on this policy will result in the denial of future access to World Bank LSMS data.

APPENDIX D
LIST OF PAPERS AND REPORTS

1.   The following is a partial list of papers done using the Nicaragua 1993 LSMS data:

a]   "Republic of Nicaragua: Poverty Assessment". Report Number 14038-NI. Country Department II, Country Operations Division, Latin America and the Caribbean Regional Office, World Bank, Washington, D.C., June 1995.

b]   Scott, Kinnon. "Determinants of Educational Attainment in Nicaragua". Consultant Report for the World Bank, 1995.

c]   Piwoz, Ellen G. "Undernutrition in Nicaragua Preschool Aged Children: Prevalence, Determinants and Policy Implications", Economic Notes, Country Department II, Latin America and the Caribbean Region, World Bank, Washington, D.C., September 1995.

d]   Pessino, Carola. "The Informal Sector in Nicaragua". Economic Notes, Country Department II, Latin America and the Caribbean Region, World Bank, Washington, D.C., February 1996.

e]   Ministerio de Acción Social. Medición de la pobreza: Documento de Trabajo, Impresiones y Torqueles, S.A, Managua, Nicaragua, June 1996.

NICARAGUA

RAAN

Jinotega

Nueva Segovia

Madríz

SEGOVIAS   NORTHERN

Estelí   ATLANTIC

Matagalpa

Chinandega   RAAS

WESTERN

León

Boaco

MANAGUA   CENTRAL

Chontales

Managua   Masaya

SOUTHERN

Carazo   Granada

Rivas

Río San Juán

REGION	DEPARTMENTS
Segovias	Estelí, Madríz, and Nueva Segovia
Western Region	León and Chinandega
Managua	Managua
Southern	Granada, Masaya, Carazo, and Rivas
Central	Boaco and Chontales
Northern	Jinotega and Matagalpa
Atlantic Coast	RAAS, RAAN, and Río San Juan