17 Latin America

The Standardized Census Sample Operation (OMUECE) of Latin America 1959-1982 [1992]: A Project of the Latin American Demographic Center (CELADE)

Robert McCaa
Dirk J. Jaspers-Faijer

Introduction

In 1959, barely five years after its founding, the Centro Latinoamericano de Demografía (CELADE) began one of the most ambitious international sample census microdata projects of the twentieth century, the Operación de Muestras de Censos (OMUECE). Thanks to the OMUECE project, the archived census microdata of Latin America are the most comprehensive of any region in the world. Of the United Nations' regional demographic centers, only CELADE undertook the mission to preserve, much less standardize, census microdata. Including many of the islands of the Caribbean and all the nations of Middle and South America, the CELADE project amassed 61 sets of high-caliber census microdata for almost two dozen countries. Of these, a total of thirty-two samples (all those

Robert McCaa is Professor of History at the University of Minnesota and a principal investigator on the IPUMS-International project at the Minnesota Population Center. Dirk J. Jaspers-Faijer is Chief of Division for Training on Population at CELADE in Santiago, Chile, where he is responsible for the development of the REDATAM initiative.

For more information, see http://www.eclac.cl/Celade-Eng.

collected prior to 1977), were incorporated into the OMUECE project (see Table 17–1).

The brain child of Carmen Miró, effectively the founding director of CELADE, the OMUECE project developed the largest, best documented (although not widely-used) census microdata library in the world. The project grew out of a 1959 seminar of Latin American census officials. Seven recommendations made at that time provided the orientation of the OMUECE project for nearly a quarter of a century. The Seminar on the Evalution and Utilization of Results of Latin

Table 17–1. Census Microdata Samples Collected by the Latin American Demographic Center (CELADE): Place and Year

	OMUECE Project		Other	
Place	**1960s**	**1970s**	**1980s**	**1990s**
Argentina	1960	1970	1980	1991
Bolivia	.	1976	.	1992
Brazil	1960	1970	1980	1991
Chile	1960	1970	1982	1992
Colombia	1964	1973	1985	1993
Costa Rica	1963	1973	1984	.
Dominican Republic	1960	1970	1981	1993
Ecuador	1962	1974	1982	1992
El Salvador	1961	1971	.	1992
Guatemala	1964	1973	1981	1994
Haiti	.	1971	1982	.
Honduras	1961	1974	.	.
Mexico	1960	1970	.	1990
Nicaragua	1963	1971	.	1995
Panama	1960	1970	1980	1990
Paraguay	1962	1972	1982	1992
Puerto Rico	1960	1970	1980	1990
Trinidad and Tobago	1963	1970	1980	.
Uruguay	1963	1975	1985	1995
Venezuela	1961	1971	1981	1990

American Population Censuses established the following objectives (CELADE, *Banco de Datos*, *Boletín Informativo* No. 4, noviembre 1970 [my translation]):

1. Increase the scope of national tabulation programs;
2. Prepare advanced tabulations, to facilitate the publication of certain results, before basic tabulations were concluded;
3. Prepare tabulations that, because of their particular nature, do not require the use of information for the entire population [that is, applicable to a subset of the population, such as individuals aged twelve years and over for the economically-active population];
4. Prepare tabulations that could be used by certain institutions for the study of specific topics of particular interest, even though they may not be included in national tabulation designs;
5. Assure the future availability of basic and detailed information on topics of demographic interest;
6. Permit institutions interested in scientific research to conduct studies based in census data;
7. Facilitate, by means of concentration in the Centro Latinoamericano de Demografía of samples corresponding to various countries, the realization of projects of demographic research, based in census data and that involve comparisons between countries of the Latin American area.

For the 1960 round of censuses, fifteen nations and the island of Puerto Rico participated, including Argentina, Brazil, Chile, Colombia, Costa Rica, the Dominican Republic, Ecuador, El Salvador, Guatemala, Honduras, Mexico, Panama, Paraguay, Uruguay and Venezuela. In the 1970s, in addition to Portuguese-speaking Brazil, all Spanish nations of the region except Cuba and Peru joined the project, and most continued to provide microdata to CELADE to the present time. While the 1960 collection was made up of relatively low-density samples of individuals, amounting to 3% or less, in most cases, the 1970 datasets were of entire households with sample densities of 5, 10,

or even 25 percent (Table 17-2). Complete copies of entire census microdata tapes were provided by five nations in the 1970 round: Bolivia, Chile, Haiti (although it is unclear whether or not these data were included in OMUECE), Nicaragua and Panama.

Although in the early 1980s, budget restrictions forced the abandonment of the standardization project,[1] CELADE continued, indeed continues, to collect census microdata for Latin American nations. With the rapid expansion of computing facilities and expertise, a centralized, standardized operation like OMUECE was no longer as essential as in earlier decades. Meanwhile, thanks to the recommendations of the 1959 Seminar, CELADE amassed one of the largest collections of international census microdata in the world, along with the corresponding documentation.

Source Material

Documentation was the key to the success of the OMUECE program, as was recognized from the very beginning. CELADE librarians systematically collected four types of basic information about census operations:

1. original enumeration schedules, including all variations of forms used;
2. enumerator instructions, including training manuals and instructions to block captains and officials through-out the chain of command of the entire census operation;
3. data processing instructions, including attribution rules for missing data as well as details for critiqueing, correcting and standardizing codes;
4. codebooks of digitized information.
5. In addition, ephemeral materials, such as broadsides, posters, and other promotional literature were obtained, whenever possible.

[1] See CELADE. (1982). *Banco de Datos, Boletín Informativo* No. 9, abril.

Table 17–2. Summary of CELADE's OMUECE Samples: Place, Decade, Size, and Density

	1960s		1970s		1980s*	
	Size	Density	Size	Density	Size	Density
Argentina	499,264	2.50%	467,801	2.00%	558,943	2.00%
Bolivia	.	.	193,597	4.20%	.	.
Brazil	913,598	1.30%	1,104,697	1.00%	3,526,473	3.00%
Chile	87,933	1.20%	442,657	5.00%	554,791	5.00%
Colombia	349,563	2.00%	926,051	4.00%	2,644,275	10.00%
Costa Rica	80,723	6.00%	200,305	10.00%	240,826	10.00%
Dominican Republic	201,556	6.60%	390,151	9.80%	593,776	10.00%
Ecuador	135,999	3.00%	923,854	17.00%	834,790	10.00%
El Salvador	25,814	1.00%	165,729	5.00%	.	.
Guatemala	209,556	5.00%	289,699	5.00%	302,198	5.00%
Haiti	.	.	.	.	559,902	10.00%
Honduras	18,818	1.00%	278,472	10.00%	.	.
Mexico	502,702	1.50%	483,361	1.00%	.	.
Nicaragua	.	.	189,469	10.00%	.	.
Panama	53,398	5.30%	286,186	20.00%	362,797	20.00%
Paraguay	89,360	4.90%	233,672	10.00%	303,536	10.00%
Puerto Rico	240,000	10.20%	.	.	.	.
Trinidad and Tobago	.	.	69,936	7.30%	129,347	10.00%
Uruguay	127,935	4.90%	165,951	6.00%	584,665	20.00%
Venezuela	132,224	1.80%	439,185	4.00%	554,053	4.00%

*Samples were drawn for the 1980s but not placed in the OMUECE standard format.

All materials were catalogued according to the highest bibliographic standards and are still available for consultation at the CELADE library in Santiago, Chile. Although not entered into its electronic web-based catalogue, they can be found in the CELADE card catalogue under "censos—documentación" and the corresponding country and year. For example, in the case of Mexico for 1960, there are thirty-two items listed. Leading up to the census, we find the decrees which provided the basis for the census and two pieces of promotional material encouraging the population to cooperate with enumerators.

For census day there are nine items: the enumeration schedule, instructions to the enumerator, instructions to school teachers describing school activities for the week of the census, instructions on how to fill out the form, flyers describing the activities of block-captains, section chiefs, and district coordinators, log-books for supervising the activities of enumerators, and special instructions on how to fill out the housing form for vacant dwellings.

The remaining 19 items document data-processing activities: a 28-page typescript detailing the rules for checking enumeration schedules, a catalogue of codes and coding rules, specific codebooks for complicated variables, such as occupation, branch of economic activity, income, social structure, educational structure, and so on. Especially interesting are the enigmatic "Special instructions for criticizing column 41" and the "Reclassification of information contained in the cards of the 1.5% samples to be used in the work of the Economic and Geographic Research Seminar of the Colegio de México."

All the metadata on Latin American census operations currently held by CELADE are presently being digitized by a National Science Foundation subcontract administered by the Minnesota Population Center through the IPUMS-International project.

In the 1960s round of censuses, microdata were captured by means of individual initiatives conceived by national census authorities. CELADE played merely an advisory role. Before computerized census tabulation methods became the norm, sampling was a special operation, beyond the scope of the regular data-processing programs of most national census agencies. Many of the resulting samples were relatively small-scale, frugal undertakings. Invariably the sampling unit was the individual rather than the family, household or dwelling. Often a simple sample of every n^{th} individual was the rule, such as for the 1964 census of Colombia for which the fraction was 1:50. This sample design maximized efficiency and robustness. Even for minor administrative districts, sample statistics were often remarkably accurate. Thanks to this simple design, sample statistics could be compared with official figures to illuminate undocumented procedures for interpreting ambiguous data or even to correct mistakes in coding or in code design.

With the adoption of fully computerized census operations in the 1970s, samples were often drawn from complete data-tapes provided by national census authorities. This made feasible the use of the household as the sampling unit. Nevertheless, for these early samples, designs and procedures were not always thoroughly documented so a certain amount of detective work is required to assess the quality and comparability of at least some of these microdata.

Procedural History

The OMUECE materials were processed in three stages—cleaning, standardization, and tabulation—before they were made available for further research. After cataloguing, the original census microdata underwent a cleaning operation in which simple frequency tabulations were compared with official data, whether published or unpublished, to discover, correct and document inconsistencies. Twenty standardized variables were then constructed, consisting of 31 bytes of information ("columns" on a punch-card in the early days). From the standardized data, 34 basic tables were constructed. Many of these were 4-, 5- and even 6-dimensional tables with the most pertinent cross-classifications by age, sex, marital status, labor-force participation and so on. Together they summed some 5,857 cells for each country. The data were then archived and made available for further research. In some cases, the original microdata were preserved as well. In others, whether for reasons of confidentiality or economy, the original data are no longer available.

Electronic Formats

The census microdata in the CELADE collection are stored as ASCII text files and archived on 9-track tapes. Machine-readable copies of all OMUECE products are probably still usable. When the CELADE-Minnesota IPUMS-International preservation project is completed in 2002, a comprehensive catalogue of available census microdata will be available. In some cases, the CELADE microdata sets for specific countries and years are the only ones still extant.

Variable Availability

The twenty variables that have been identified for OMUECE standardization are described in Table 17-3. Fertility was represented by a single variable, but each of five core sets—location (with the addition of country), person, education, work, and migration—were represented by four variables each. The individual's location in administrative geography is determined by country (implicit since the datatapes were classified by country, but apparently for reasons of economy the data themselves were not encoded with a country variable), major administrative division, minor division, and urban/rural residence with a unique code for residence in the capital city.

Socio-demographic variables were standardized—sex, age, relation to household head, and marital status—as was information on education: literacy, school attendance, level of instruction, and the last year of instruction attained. Economic variables, based on international standards developed for COTA and CIUO, were of major concern. While only four variables were collected, these were alloted almost one-third of the space for standardization, including two columns for labor force participation, one for occupational category, four for occupation, and two for branch of work. Migration was also of considerable concern, including type and duration of current residence, place of prior residence, and place of birth. One fertility variable, number of children ever-born, completed the dataset.

OMUECE's standard coding scheme greatly facilitated comparative international research, as demonstrated by data availability by country for the 1960-1980 rounds of censuses (Table 17–4).

In 1970 and subsequent census rounds, many Latin American countries adopted the core OMUECE standards as their own. It is important to note, however, that census concepts applied in various national contexts were not necessarily uniform. For example, with respect to economic activity, in some cases the question referred to at least one hour of paid work in the week preceding the enumeration, while in others the query referred to the previous year, or simply usual occupation. The OMUECE standardization design called for

Table 17–3. Twenty-One Standard OMUECE Variables and Their Values

Variable	Digits	Codes	Missing Values
Country	2		9
Major administrative division	2	Specific for each country	9
Minor administrative division	3	Specific for each major administrative division	9
Zone	1	1 = Capital 2 = Urban 3 = Rural	9
Sex	1	1= Male 2=Female	9
Age	2	0 = >l year 1 = 1 year old 2 = 2 years old … 98 = 98+ years old	99
Relation to household head	1	1 = Head 2 = Spouse 3 = Child 4 = Other relative 5 = Servants 6 = Non-relatives	9
Marital status	1	1 = Single 2 = Married 3 = In union 4 = Widowed 5 = Divorced 6 = Legally separated	9
Literacy	1	1 = Literate 2 = Not literate	9
School attendance	1	1 = Attends 2 = Does not attend	9

Table 17–3. Twenty-One Standard OMUECE Variables and Their Values (continued)

Variable	Digits	Codes	Missing Values
Level of instruction	1	0 = None 1 = Primary 2 = Secondary 3 = Post-secondary	9
Years of instruction	1	0 = None 1 = 1 year 2 = 2 years 3 = 3 years 4 = 4 years 5 = 5 years 6 = 6 years 7 = 7+ years	9
Labor-force participation	2	1 = In labor force 2 = Does not work for pay Second digit gives detail	99
Category of worker	1	0 = None 1 = Employer 2 = Self-employed 3 = Employee 4 = Family worker 5 = Coop worker	
Occupation	4	0 = None 1 = Professional 2 = Administrative 3 = Office 4 = Vendor 5 – Agriculture 6 = Mining 7 = Transport 8 = Artisan 9 = 99 = 9999 =	99
Branch of work	2		99

Table 17–3. Twenty-One Standard OMUECE Variables and Their Values (continued)

Variable	Digits	Codes	Missing Values
Duration of migration	1	1 = Resident 2 = Temporary 3 = Permanent	9
Permanent resident	1	0 = < 1 year 1 = 1 year 2 = 2 years 3 = 3 years 4 = 4 years 5 = 5-9 years 6 = 10-14 years 7 = 15-19 years 8 = 20+ years	9
Place of prior residence	1	1 = Same minor 2 = Different minor 3 = Not indicated 4 = Other major 5 = Other country 6 = This country; MAD not indicated	9
Birthplace	1	1 = Same minor 2 = Different minor 3 = Not indicated 4 = Other major 5 = Other country 6 = This country; MAD not indicated	9
Children ever born	2	1 = 1 child 2 = 2 children 3 = 3 . . .	99

applying the same coding scheme to these quite distinct concepts of economic activity. Modern principles for integrating census microdata require that these distinctions be retained.

Confidentiality Provisions

Access to OMUECE samples and other census microdata held by CELADE is strictly limited to CELADE personnel working at the CELADE offices in Santiago, Chile. CELADE personnel will compute statistical output and conduct demographic analysis for users on a fee basis. Any resulting statistics are checked by CELADE personnel to ensure that the highest standards of privacy and confidentiality are maintained.

Data Access

Researchers may analyze Latin American census microdata on CELADE computers, only after written permission is obtained from the relevant national statistical offices and a confidentiality agreement is signed. CELADE does not permit access to census microdata for any country without explicit written permission from the relevant census office. Moreover, CELADE does not assist in the approval process, even to the extent of providing a recommendation regarding access for a specific researcher.

Publications Using these Data

Given these restrictions, it should not be surprising that the principal publications which resulted from the OMUECE project were basic tabulations of the data by CELADE experts. While too numerous to list here, the titles are readily available from the CELADE library web-page, using OMUECE as the key word http://www.eclac.cl/excite/Docpal. In addition several important comparative studies were published by CELADE and its mother institution, the Economic Commission of Latin America (CEPAL), on issues of reproduction (Aldunate and León B., *Comportamiento reproductivo*, 1975), educational opportunity (Filguiera, *Expansión educacional*, 1977; Terra, *Alfabetismo y escolarización*, 1979), and female labor force

Table 17–4. Data Availability by Country: OMUECE Standardized Variables, 1960 and 1970 Rounds of Censuses

Variables	Countries															
	Arg	Bra	Col	CR	Chi	Ecu	ES	Gua	Hon	Mex	Pan	Par	PR	RD	Uru	Ven
Major administrative unit	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Minor administrative unit	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Zone	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Sex	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Age	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Relation to household head	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Marital status	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Literacy	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
School attendance	X	X	.	X	X	X	X	X	X	X	X	X	X	X	X	X
Level of instruction	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Years of instruction	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Labor-force participation	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Category of worker	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Occupation	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Branch of work	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Duration of migration	X	X	X	X	X	X	.	X	.	X	X	X	X	.	X	.
Permanent resident	.	.	.	.	X	X	.	.	.	.	X	X	.	.	.	.
Place of prior residence	X	X	X	X	X	X	.	X	.	X	X	X	.	.	.	.
Birthplace	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Children ever born	X	X	.	.	X	.	.	X	.	X	.	X	X	.	X	.

participation (Barrera, *Participación femenina*, 1978). Recently there has been renewed interest in the OMUECE database for broadly comparative studies on the demography of the family and household (Arias and De Vos, "Using Housing Items," 1996; Arias and Palloni, "Prevalence and Patterns of Female Headed Households," 1999).

Research Possibilities

Most publications using these data focus on a single country and a single census (e.g., Pantelides, *Costa Rica*, 1972; Polo Najera, *Republica Dominicana*, 1977). The research potential of the OMUECE database is quite substantial, particularly for cross-national analysis over time. The comparative evolution of the labor force, educational opportunity, internal migration, fertility, marriage, and family are a few of the broad possibilities.

Expert Users

The CELADE research team constitutes the principal group of expert users. Dirk Jaspers Faijer is coordinating the IPUMS-International CELADE preservation project. In the United States, the University of Wisconsin demographers Alberto Palloni and Susan De Vos have considerable experience working with the OMUECE samples.

Data Expansion

The OMUECE project was terminated in 1982, and CELADE has no plans to revive it for the 1990 or subsequent censuses. The long term goal of the IPUMS-International project is to integrate much of the highest quality census microdata for Latin America. If successful, the proposal would be to return to the original census microdata, preserve as much of the original coding scheme and concepts as possible, and integrate both household and population variables, according to current international principles.

Bibliography

Aldunate, Rodolfo and Arturo León B. (1975). *Comportamiento reproductivo y hetereogeneidad estructural*. Santiago, Chile: CELADE.

Arias, Elizabeth and Susan De Vos. (1996). "Using Housing Items to Indicate Socioeconomic Status: Latin America," *Social Indicators Research*, 38: 53-80.

Arias, Elizabeth and Alberto Palloni. (1999). "Prevalence and Patterns of Female Headed Households in Latin America: 1970-1990," *Journal of Comparative Family Studies*, 30:2 (Spring) 257-79.

Barrera, Manuel. (1978). *Participación feminina en la actividad económica en América Latina*. Santiago, Chile: CELADE.

Centro Latinoamericano de Demografía. (1986). *Banco de Datos, Boletín Informativo*, números 1-11. Santiago, Chile: 1968-1986.

Filgueira, Carlos H. (1977). *Expansión educacional y estructura social en América Latina*. Buenos Aires: CEPAL.

Pantelides, Edith Alejandra. (1972). *Costa Rica: estudio de la migración interna a partir de una muestra del Censo de 1963 (OMUECE)*. Santiago: CELADE.

Polo Najera, Jose. (1977). *República Dominicana: analisis de la evolución de la población económicamente activa, 1960-1970*. San José: CELADE.

Terra, Juan Pablo. (1979). *Alfabetismo y escolarización básica de los jóvenes en América Latina*. Santiago, Chile: CEPAL.