Public Use Samples of 1910 and 1920 Puerto Rico Censuses

Kristen Velyvis
Theresa Thompson-Colón
Halliman Winsborough
University of Wisconsin-Madison

Overview

1. The 1910 and 1920 Puerto Rican Census Project

2. How the sample files were produced and their quality

3. Interesting issues found in the process of making of the samples

1910 & 1920 Puerto Rican Census Project

- Large-scale data entry enterprise undertaken at the University of Wisconsin-Madison

- In collaboration with the Integrated Public Use Microdata Services (IPUMS) at the University of Minnesota

Final Products

- Two sets of public use sample files of individual and household records drawn from the Puerto Rican censuses of 1910 and 1920

Final Products

- Created by sampling and transcribing information from the microfilms of the original schedules of the first and second Censuses of Puerto Rico taken by the US Bureau of the Census

Microfilmed Census Schedule

Microfilmed Census Schedule

Files Are Available In Two Forms

- ICPSR versions
  - Include all variables unedited and unallocated for missing values
  - Include original input text variables exactly as transcribed
  - Available at http://www.icpsr.umich.edu

Files Are Available In Two Forms

- IPUMS versions
  - Conform to the IPUMS standard for census samples
    - Carried out data edits and allocation procedures
  - Strictly comparable to the US public use samples
  - Will be available at http://www.ipums.umn.edu/usa/

How the Files Were Produced

- The project took four and half years to complete starting in January of 2002
- Main project stages:
  - Pre-data entry
  - Data entry
  - Post-data entry

How the Files Were Produced

- Four data entry operators worked full time for 15 months to enter all the data

- Post-entry stage included checking and verification, data coding, format standarization to the get the data ready for IPUMS

- IPUMS conducted the data allocation

The Value of the Data

- Uniqueness
- Compelling time period they cover
- Comparability with data from other places
- Completeness

The Data in More Detail

- The population covered by the samples includes the entire population of the island of Puerto Rico in 1910 and 1920

- The data files consist of 1-in-10 samples for most of the island

  - Except the coffee-producing regions

  - And an African-descent enclave, the *municipio,* or county, of Loíza

Table 1: Distribution of the 1910 and 1920 Puerto Rican Samples

Census Year	Main Sample	Coffee region *	Loíza *	Total Sample
1910	95,124	40,092	2,675	137,891
1920	112,683	44,706	3, 097	160,484
Total	207,807	84,798	5,772	298,375

* Over-sample

Table 2: Race—1910

Race	Census Pop.	% of Pop.	Weighted Sample	% of Sample
White	732,555	65.5%	72,167	64.9%
Black	50,245	4.5%	5, 025	4.5%
Mulatto	335,192	30.0%	34,028	30.6%
Other	20	0.0%	3	0.0%
Total	1,118,012	100.0%	111,223	100.0%

Census & Preliminary ICPSR sample data (4/14/05)

Table 5b: Age X Sex –1920

Women				
Age group	Census Pop. Women	% of Pop.	Weighted Sample	% of Wght Sample
Under 5	99,357	15.2%	9,774	15.3%
5 to 9	95,981	14.7%	9,432	14.7%
10 to 14	82,352	12.6%	7,963	12.4%
. . .				
Total	651,984	100.0%	64,061	100.0%

Table 6a: Population per *Municipio*—1910

	Number of Families per *Municipio*				Number of Persons per *Municipio*			
Municipio	Census Popltn	% of Popltn	Wt'ed Sample	% of Sample	Census Popltn	% of Popltn	Wt'ed Sample	% of Sample
Adjuntas	3,250	1.48%	326.0	1.45%	16,954	1.52%	1,694.0	1.52%
Aguada	2,326	1.06%	246.0	1.10%	11,587	1.04%	1,183.0	1.06%
Aguadilla	4,218	1.92%	431.0	1.92%	21,419	1.92%	2,111.0	1.90%
Aguas Buenas	1,463	0.67%	142.0	0.63%	8,292	0.74%	836.0	0.75%
Aibonito	2,068	0.94%	194.0	0.86%	10,815	0.97%	995.0	0.89%
Añasco	2,934	1.33%	287.0	1.28%	14,407	1.29%	1,443.5	1.30%
Arecibo	8,453	3.85%	865.0	3.85%	42,429	3.80%	4,122.0	3.71%
Arroyo	1,576	0.72%	151.0	0.67%	6,940	0.62%	664.0	0.60%
Barceloneta	2,240	1.02%	213.0	0.95%	11,644	1.04%	1,069.0	0.96%
Barranquitas	1,818	0.83%	195.0	0.87%	10,503	0.94%	1,071.0	0.96%
Barros	2,578	1.17%	275.0	1.22%	15,028	1.34%	1,659.0	1.49%
Bayamón	6,062	2.76%	562.0	2.50%	29,986	2.68%	2,920.0	2.63%
Cabo Rojo	3,713	1.69%	380.0	1.69%	19,562	1.75%	1,971.0	1.77%

Table 6a: Population per *Municipio*—1910

Municipio	Number of *Families* per *Municipio*			
	Census Popltn	% of Popltn	Wt'ed Sample	% of Sample
Aguas Buenas	1,463	0.67%	142.0	0.63%

Table 6a: Population per *Municipio*—1910

Municipio	Number of *Person* per *Municipio*			
	Census Popltn	% of Popltn	Wt'ed Sample	% of Sample
Aguas Buenas	8,292	0.74%	836.0	0.75%

Table 7a: Population per *Municipio*—1920

Municipio	Number of Families per Municipio				Number of Persons per Municipio			
	Census Popltn	% of Popltn	Wt'ed Sample	% of Sample	Census Popltn	% of Popltn	Wt'ed Sample	% of Sample
Adjuntas	3,310	1.32%	336.5	1.33%	17,988	1.38%	1,858.0	1.45%
Aguada	2,604	1.04%	264.0	1.05%	12,981	1.00%	1,351.0	1.05%
Aguadilla	4,605	1.83%	448.0	1.77%	24,287	1.87%	2,390.0	1.87%
Aguas Buenas	1,834	0.73%	187.0	0.74%	10,741	0.83%	1,063.0	0.83%
Aibonito	2,306	0.92%	230.0	0.91%	13,264	1.02%	1,242.0	0.97%
Añasco	2,833	1.13%	283.0	1.12%	13,834	1.06%	1,357.5	1.06%
Arecibo	9,560	3.81%	956.0	3.78%	46,578	3.58%	4,702.0	3.67%
Arroyo	1,630	0.65%	160.0	0.63%	7,074	0.54%	717.0	0.56%
Barceloneta	2,609	1.04%	260.0	1.03%	13,442	1.03%	1,385.0	1.08%
Barranquitas	1,942	0.77%	194.0	0.77%	11,600	0.89%	1,128.0	0.88%
Barros	2,637	1.05%	251.0	0.99%	15,758	1.21%	1,560.0	1.22%
Bayamon	6,212	2.47%	638.0	2.53%	30,739	2.36%	3,142.0	2.45%
Cabo Rojo	4,334	1.73%	471.0	1.86%	22,412	1.72%	2,284.0	1.78%

Some Comments on Taking a U.S. Census in Puerto Rico in 1910

A Census is Culture-laden

- US Census Bureau has a method of census taking to produce data that contemporary users will find meaningful and useful.

- The process was developed based on shared cultural practices and shared meanings in the US.

- In PR, the cultural assumptions became obvious.

Three Important Areas

- Race

- Family

- Geography

Microfilmed Census Schedule

Race

"For census purposes, the term 'black' (B) includes all persons who are evidently full-blooded negroes, while the term 'mulatto' (Mu) includes all other persons having some proportion or perceptible trace of negro blood" (Departamento de Comercio y Trabajo 1910).

Family

1. A household defines a family.

2. Is the oldest male always the best reference for family relationships?

Household
Oldest Male
Consensual Union Wife
His Brother-in-law
Wife of Brother-in-law
Brother-in-law & his wife's daughter

Household	Census
Oldest Male	Head
Consensual Union Wife	Spouse
His Brother-in-law	Brother-in-law
Wife of Brother-in-law	*Concuñada*
Brother-in-law & his wife's daughter	Lodger

Household	Census	Alternative
Oldest Male	Head	Spouse
Consensual Union Wife	Spouse	Head
His Brother-in-law	Brother-in-law	Brother
Wife of Brother-in-law	*Concuñada*	Sister-in-law
Brother-in-law & his wife's daughter	Lodger	Niece

Family

1. A household defines a family.
2. Is the oldest male always the best reference for family relationships?
3. Kinship varies across cultures / languages.

Geography

- The US Census creates map-based enumeration districts

Geography

"These barrios, with the exception of a few urban barrios in the larger cities, have no fixed geographical boundaries, the division lines being purely imaginary, thus making it very difficult clearly to describe their limits. This being the case, it can easily be understood how difficult it would be to divide the island accurately into enumeration districts and to describe the limits of each district" (USNA, RG 29, Entry 254).

Summary of Comments on the Taking a U.S. Census in Puerto Rico in 1910

- The application of a practiced census to a new area can be revealing about both the new and the old areas.

- A professional census bureau will learn a lot in its first trial in a new area.

  - Comparison of the first and second instances are likely to confound social change and this bureaucratic learning.

Conclusion

Two contradictory goals

1. To produce a modern statistical file that is as true to the time and place as possible.

2. To create a resource that permits cross-time and cross-place comparisons.