Developing a Low Response Score Model for Puerto Rico

Kelyvette Ortiz Fontanez, Ana I. Sánchez Rivera, Yazmin García Trejo US Census Bureau



This presentation is released to inform interested parties of research and to encourage discussion. The views expressed are those of the author(s) and not those of the U.S. Census Bureau.

What is the Low Response Score (LRS)?

Low Response Score (LRS)

- Measure developed by the Census Bureau to predict the percentage of people in a given census track or block group likely to not self-respond to the Decennial Census.
- The most recent version (2021) is based on an Ordinary Least Squares (OLS) regression of 25 predictors.
- Its values range from 0-100. Lower values mean a smaller proportion of the tract or block group is likely to not self respond.

Census Planning Database (PDB)

- The PDB contains the "Greatest hits" of ACS 5-year estimates and Decennial Census counts.
- Includes counts, percentages, and margins of error of several population, household, housing unit and operational measures



Why an LRS for Puerto Rico?

State Specific Models

• Contributes to the Census Bureau's broader initiative to develop state-specific LRS models.

ACS Puerto Rico Data

• ACS gathers annual data from Puerto Rico in preparation for the 2030 Census

Close a Gap

• Current LRS model excludes Puerto Rico estimates.

Advance Research

• Conducting advanced research on Puerto Rico's LRS applications and study usage



Research questions



RQ1: How accurately does the current LRS Model predict low response census tracts in Puerto Rico?



RQ2: What adjustments are needed to improve the LRS Model's prediction accuracy in Puerto Rico?



Method



Assess Model Performance

- Coefficient analysis
- Descriptive analysis and correlation tests
- Residual analysis



Explore Model Modifications

- Identifying predictors to remove
- Identify new predictor candidates



Evaluate Proposed LRS Puerto Rico Model



Step 1: Coefficient Analysis for Baseline Model

Predictor (n=25)	Estimate	P Value	
Intercept		0.00*	
Socio-Demographic Variables = 13			
Percent of males Percent Hispanic population Percent Non-Hispanic White alone Percent Non-Hispanic Black alone Percent of population aged 5-17 Percent of population aged 18-24 Percent of population aged 25-44 Percent of population aged 45-64 Percent of population aged 65+ Percent not high school graduates Percent college educated	0.00 0.04 -0.01 0.33 -0.03 -0.11 -0.19 -0.05 -0.26 0.02 -0.11	0.22 0.48 0.91 0.00* 0.52 0.00* 0.25 0.00* 0.03* 0.03*	
Median household income	-0.05	0.08*	
Percent persons below poverty level	0.01	0.20	
Household Composition Variables = 5			
Percent single person households	-0.10	0.00*	
Percent of females householder with no spouse present	-0.08	0.00*	
Percent married couple households Percent of related children under age 6 Average total persons in household	-0.01 0.01 -1.08	0.00* 0.11 0.00*	
Tenure and Housing Unit = 5			
Percent renter occupied housing unit Percent of householders moved within the past 5 years Percent of different housing unit 1 year ago Percent of single unit Percent of vacant unit	-0.01 0.00 0.00 0.00 0.21	0.39 0.51 0.75 0.32 0.00*	
Select Geographic Characteristics = 2			
Median house value Population density	0.07 -0.01	0.00* 0.09*	

Table 1. Baseline Model: Predictor Coefficient And Signif	ficance
---	---------

Table 2. Baseline Model Fit Statistics

R2	Adjusted R2	MSE
0.72	0.71	0.0164



Source: U.S. Census Bureau, 2022 Planning Database

Step 1: Coefficient Analysis for Baseline Model

Predictor (n=25)	Estimate	P_Value
Intercept		0.00*
Socio-Demographic Variables = 13		
Percent of males	0.00	0.22
Percent Hispanic population	0.04	0.48
Percent Non-Hispanic White alone	-0.01	0.91
Percent Non-Hispanic Black alone	0.33	0.00*
Percent of population aged 5-17	-0.03	0.52
Percent of population aged 18-24	-0.11	0.00*
Percent of population aged 25-44	-0.19	0.00*
Percent of population aged 45-64	-0.05	0.25
Percent of population aged 65+	-0.26	0.00*
Percent not high school graduates	0.02	0.03*
Percent college educated	-0.11	0.00*
Median household income	-0.05	0.08*
Percent persons below poverty level	0.01	0.20
Household Composition Variables = 5		
Percent single person households	-0.10	0.00*
Percent of females householder with no spouse present	-0.08	0.00*
Percent married couple households	-0.01	0.00*
Percent of related children under age 6	0.01	0.11
Average total persons in household	-1.08	0.00*
Tenure and Housing Unit = 5		
Percent renter occupied housing unit	-0.01	0.39
Percent of householders moved within the past 5 years	0.00	0.51
Percent of different housing unit 1 year ago	0.00	0.75
Percent of single unit	0.00	0.32
Percent of vacant unit	0.21	0.00*
Select Geographic Characteristics = 2		
Median house value	0.07	0.00*
Population density	-0.01	0.09*

Table 1. Baseline Model: Predictor Coefficient And Significance

Table 2. Baseline Model Fit Statistics

R2	Adjusted R2	MSE
0.72	0.71	0.0164



Source: U.S. Census Bureau, 2022 Planning Database

Step 2: Removing Predictors





Distribution of tracts by Percent of Person Below Poverty Level: Puerto Rico vs United States







United States

Puerto Rico

Findings Step 2: Coefficient Analysis For Reduced Model

Table 3. Reduced Model: Predictor Coefficient And Significance

Predictor (n=16)	Estimate	P_Value
Intercept		0.00*
Socio-Demographic Variables = 9		
Percent Non-Hispanic Black alone	0.29	0.00*
Percent of population aged 5-17	-0.02	0.60
Percent of population aged 18-24	-0.10	0.00*
Percent of population aged 25-44	-0.16	0.00*
Percent of population aged 45-64	-0.03	0.50
Percent of population aged 65+	-0.24	0.00*
Percent not high school graduates	0.02	0.02*
Percent college educated	-0.12	0.00*
Median household income	-0.07	0.00*
Household Composition Variables = 4		
Percent single person households	-0.10	0.00*
Percent of females householder with no spouse present	-0.08	0.00*
Percent married couple households	-0.01	0.00*
Average total persons in household	-1.00	0.00*
Housing Unit = 1		
Percent of vacant unit	0.21	0.00*
Select Geographic Characteristics = 2		
Median house value	0.04	0.04*
Population density	-0.01	0.04*

Source: U.S. Census Bureau, 2022 Planning Database

Table 4. Reduced Model Model Fit Statistics

R ²	Adjusted R ²	MSE
0.72	0.71	0.0167

Table 5. Removed Predictors (N=8)

Percent of male	Percent renter occupied housing unit
Percent Hispanic population	Percent of householders moved within the past 5 years
Percent Non-Hispanic White alone	Percent of different housing unit 1 year ago
Percent persons below poverty level	Percent of single unit



Findings Step 3: Coefficient Analysis For Revised Model

Table 6. Proposed Model: Predictor Coefficient and Significance

Predictor (n=19)	Estimate	P_Value
Intercept		0.00*
Socio-Demographic Variables = 10		
Percent Non-Hispanic Black alone Percent of population aged 5-17 Percent of population aged 18-24 Percent of population aged 25-44 Percent of population aged 45-64 Percent of population aged 65+ Percent college educated Median household income Percent of population aged 5 years and over who speaks English less than "very well" and speaks Spanish or Spanish creole at home	0.25 -0.06 -0.14 -0.23 -0.10 -0.29 -0.09 -0.06 0.04	0.00* 0.13 0.00* 0.00* 0.02 0.00* 0.00* 0.00*
Percent of population that receive public assistance income	-0.01	0.07*
Household Composition Variables = 4		
Percent single person households Percent of females householder with no spouse present Percent married couple households	-0.11 -0.12 -0.01	0.00* 0.00* 0.00*
Average total persons in household	-0.97	0.00*
Housing Unit = 2		
Percent of vacant unit Percent of crowded occupied housing units	0.22 0.02	0.00* 0.00*
Select Geographic Characteristics = 1		
Median house value	0.05	0.01*
Technology and Connectivity = 2		
Percent of households with only smart phones Percent of households with no computing device	0.06 0.04	0.00* 0.00*

Table 7. Revised Model Fit Statistics

R ²	Adjusted R ²	MSE
0.74	0.73	0.0154

Source: U.S. Census Bureau, 2022 Planning Database



Findings Step 3: Coefficient Analysis For Revised Model

Table 6. Proposed Model: Predictor Coefficient and Significance

Predictor (n=19)	Estimate	P_Value
Intercept		0.00*
Socio-Demographic Variables = 10		
Percent Non-Hispanic Black alone Percent of population aged 5-17 Percent of population aged 18-24 Percent of population aged 25-44 Percent of population aged 45-64 Percent of population aged 65+ Percent college educated Median household income Percent of population aged 5 years and over who speaks English less than "very well" and speaks Spanish or Spanish creole at home	0.25 -0.06 -0.14 -0.23 -0.10 -0.29 -0.09 -0.06 0.04	0.00* 0.13 0.00* 0.00* 0.02 0.00* 0.00* 0.00*
Percent of population that receive public assistance income	-0.01	0.07*
Household Composition Variables = 4		
Percent single person households Percent of females householder with no spouse present Percent married couple households	-0.11 -0.12 -0.01	0.00* 0.00* 0.00*
Average total persons in household	-0.97	0.00*
Housing Unit = 2		
Percent of vacant unit	0.22	0.00*
Percent of crowded occupied housing units	0.02	0.00*
Select Geographic Characteristics = 1		
Median house value	0.05	0.01*
Technology and Connectivity = 2		
Percent of households with only smart phones	0.06	0.00*
Percent of households with no computing device	0.04	0.00*

Source: U.S. Census Bureau, 2022 Planning Database



Table 7. Revised Model Fit Statistics

R ²	Adjusted R ²	MSE
0.74	0.73	0.0154

Findings Step 3: Coefficient Analysis For Revised Model

Table 6. Proposed Model: Predictor Coefficient and Significance

Predictor (n=19)	Estimate	P_Value		
Intercept		0.00*		
Socio-Demographic Variables = 10				
Percent Non-Hispanic Black alone Percent of population aged 5-17 Percent of population aged 18-24 Percent of population aged 25-44 Percent of population aged 45-64 Percent of population aged 65+ Percent college educated	0.25 -0.06 -0.14 -0.23 -0.10 -0.29 -0.09	0.00* 0.13 0.00* 0.00* 0.02 0.00* 0.00*		
Median household income	-0.06	0.00*		
Percent of population aged 5 years and over who speaks English less than "very well" and speaks Spanish or Spanish creole at home Percent of population that receive public assistance income	0.04	0.00* 0.07*		
Household Composition Variables = 4				
Percent single person households Percent of females householder with no spouse present Percent married couple households Average total persons in household	-0.11 -0.12 -0.01 -0.97	0.00* 0.00* 0.00* 0.00*		
Housing Unit = 2				
Percent of vacant unit Percent of crowded occupied housing units	0.22 0.02	0.00* 0.00*		
Select Geographic Characteristics = 1				
Median house value	0.05	0.01*		
Technology and Connectivity = 2				
Percent of households with only smart phones Percent of households with no computing device	0.06 0.04	0.00* 0.00*		

Source: U.S. Census Bureau, 2022 Planning Database



Table 7. Revised Model Fit Statistics

R ²	Adjusted R ²	MSE
0.74	0.73	0.0154

Residual Analysis

Residual range	Baseline	Reduced	Revised
0 to 5	74.50%	74.50%	76.30%
5 to 10	21.80%	21.70%	20.70%
10 to 15	3.40%	3.40%	2.60%
15 to 20	0.20%	0.30%	0.30%
20+	0.10%	0.10%	0.00%

Table 8. Proposed Acs Predictors Specific To Puerto Rico

Fig 4. Baseline Model: Residuals Absolute Value By Tract



Fig 5. Revised Model: Residual Absolute Value By Tract





Results

Future Research

- Include geographical variables. Specifically, metropolitan vs non-metropolitan areas.
- Include Non-Response Follow Up Operational Statistics.
- Identify predictors not part of the current PDB
- Dynamic predictor selection



¡Gracias!

- Kelyvette Ortiz Fontanez <u>kelyvette.ortiz.fontanez@census.gov</u>
- Ana I. Sánchez-Rivera <u>ana.i.sanchez.rivera@census.gov</u>
- Yazmin Garcia Trejo <u>yazmin.argen.garcia.trejo@census.gov</u>

