Investigating the Role of Hispanic Origin in Estimating the Number of Uninsured

Sheri Sharareh Craig*, Joanna Turner and Brett O'Hara
U.S. Census Bureau** Washington, DC 20233

Key Words: Small Area Estimates; Health Insurance Coverage; Hispanic Origin

1. Background

The United States Census Bureau's Small Area Health Insurance Estimates (SAHIE) program developed model-based health insurance coverage estimates for states and counties that will provide consistent and comparable estimates for all areas. The current model has shown that a variable indicating people who are of Hispanic origin is a strong predictor for people without health insurance coverage. More information is available at the SAHIE program's web site, http://www.census.gov/hhes/www/sahie.

The number of people in the United States without health insurance has grown from about 31 million (13 percent) in 1987 to 45 million (16 percent) in 2003 (DeNavas-Walt *et al.,* 2004). Hispanics were more likely to be uninsured than non-Hispanic Whites, 33 percent and 11 percent, respectively (DeNavas-Walt *et al.,* 2004).

The Hispanic population of the United States is growing fast and changing fast (PEW Hispanic Center, 2005). Hispanic uninsured rates (about 33 percent) have remained consistently high over past decades (Commonwealth Fund, 2003). There has been a consistent upward trend in the number of uninsured Hispanics from 6 million in 1987 to 13 million in 2003, and the proportion uninsured remained high at 31 percent in 1987 and 33 percent in 2003.

Hispanic groups are described in Census 2000 as Mexican, Mexican American, Puerto Rican, Cuban, Central or South American, or other Hispanic.

* Please contact Sheri Sharareh Craig with any questions by phone (301-763-5793) or by email (sheri.s.sharareh.craig@census.gov).
** This report is released to inform interested parties of ongoing research and to encourage discussion of work in progress. The views expressed on technical issues are those of the authors and not necessarily those of the U.S. Census Bureau.

Among the Hispanic-origin population, persons of Mexican origin were the most likely to lack health insurance coverage (39 percent) (NCHS, 2003a). The purpose of this research is to gain better understanding of the strong relationship between health insurance coverage and Hispanic origin and to investigate the development of other variables that could be used for modeling health insurance coverage. The correlation among variables such as, "years lived in the United States," "citizenship," "language of interview," "educational level," "region," "marital status," and "family size" explains variations in health insurance status among Hispanics. The results will be used to refine the health insurance coverage estimates in the SAHIE program's health insurance models.

Section 2 presents the data and sample design; Section 3 describes the methodology; Section 4 displays the results; and Section 5 discusses future research and gives the conclusion.

2. Data and Sample Design

The data used in this study are from the 2000-2002 National Health Interview Survey (NHIS) (NCHS, 2003b), which is sponsored by the Centers for Disease Control and Prevention's National Center for Health Statistics. NHIS data are collected annually for approximately 100,000 persons in 40,000 households. The overall household response rate for the 2000-2002 NHIS was 89 percent.

The NHIS is a nationally representative household survey that produces estimates of health conditions for the civilian, noninstitutionalized population of the United States. NHIS has a multistage area probability-based complex sample design, including clustering, stratification, and the assignment of unequal probabilities of selection. Three hundred fifty eight primary sampling units (PSUs) are drawn from approximately 1,900 geographically defined PSUs covering the 50 States and the District of Columbia. Hispanic/Latino and Black populations are over-sampled to increase the precision of estimates for those subgroups. Because of the complicated survey design, all of the results use the final person weight and the variances are corrected for stratification.

Health insurance information has been collected on the NHIS respondents periodically since 1960 and continuously since 1989. Being uninsured means that the person was not covered by any type of health insurance as well as those who had only Indian Health Service coverage or had only a private plan that paid for one type of service. The strengths of using the NHIS for a health insurance study are in its low non-response rate, cognitive testing, and the fact that the health plan names are obtained from health plan cards.

The NHIS groups respondents with Hispanic ancestry into many categories. We combined the NHIS categories into five groups: Puerto Rican, Mexican, Mexican-American, Central/South/Latin Americans, and other Hispanic groups[1]. We also created three categories for educational level: no high school diploma, high school diploma, or some college and above.

The age group selected for the majority of the study is 18-64 because persons ages 65 and over have high rates of insurance due to Medicaid and Medicare regardless of race and ethnicity. The combined data (NHIS 2000-2002) generated a larger sample size for each Hispanic/Latino subgroup. During 2000-2002, interviews were conducted for a total of 294,765 persons in 113,726 households, including 38,545 Hispanic/Latino persons ages 18-64. Using the sample weights, there were about 170 million adults age 18-64 per year and about 20 million were Hispanic.

3. Methodology

We investigated the relationship between "Hispanic origin" and "no health insurance coverage". Our question is "What is unique about Hispanic origin and the uninsured rate?" The following dependent variables were used in this model: citizenship, foreign-born, years lived in the United States, language of interview, educational level, poverty status, family size, marital status, region, sex, and employment status.

We compared health insurance rates of the Hispanic population with the non-Hispanic Black and non-Hispanic White population as well as selected Hispanic and non-Hispanic ancestries. The foreign-

---

[1] The other Hispanic groups category consists of multiple Hispanics, Cubans, Dominican Republic, other Latin American –type not specified, other Spanish and Hispanic/Latino/Spanish, non-specified type, type refused and not ascertained. This is because of the low sample size in these categories.

born Hispanic population is examined using the variable "years lived in the United States". "Persons born in the United States" includes persons born in the 50 states and the District of Columbia. Persons born in Puerto Rico, Guam, or outlying territories of the United States are not included in the native born category because of the NHIS classification scheme. The "others" group includes: do not know, not ascertained, unknown, and refused responses. This group was deleted from the graphs and tables, but not from the underlying analysis.

Because of the complex survey design, we used appropriate SAS procedures such as proc SurveyFreq and proc SurveyLogistic. By using SurveyFreq, figures and tables were created to visually describe relationships.

Because "being uninsured" is a yes or no response, a logistic regression was chosen. The odds ratio, an output of SurveyLogistic, is used for interpretation of the model. For model-based odds ratios, the odds are "adjusted" and are interpreted using the *ceteris paribus* assumption (e.g. all else held constant). The odds ratio has been recalculated to reflect the change in the odds.

All of the independent variables are categorical, so interpretation of the regression is straightforward. For example, if the increase in the odds for sex was 0.35, the odds of being uninsured are 35 percent higher for a male than a female, all else held constant. Negative changes in the odds are interpreted in the same way. The odds of being uninsured are 99% less for a Puerto Rican when compared with Mexican, all else held constant. The 90% confidence interval is calculated for the effects. Statistical significance is assumed if the interval does not include zero.

The overall "importance" of a given variable is measured by the additive increase in the fit of the regression using an R-squared (R^2) measure. The increase in R^2 is a combination of the mean value of an independent variable and the beta coefficient. To examine the influence of each independent variable on the explained variation in the health insurance model, we performed a stepwise regression. The stepwise regression chooses the most influential variables in order of importance.

4. Results

Tabular results (figures and tables)

All of the results are based on NHIS data from 2000-2002. Figure 1 shows 5 percent of Hispanics age 65 and over are uninsured, whereas non-Hispanic Whites and non-Hispanic Blacks age 65 and over had lower rates of being uninsured (1 percent and 2 percent, respectively).

Persons ages 18-64 have higher rates of being uninsured than persons 65 and older. Hispanics between the ages of 18-64 have a higher rate of being uninsured (40 percent) when compared with non-Hispanic Whites (14 percent) or non-Hispanic Blacks (23 percent).
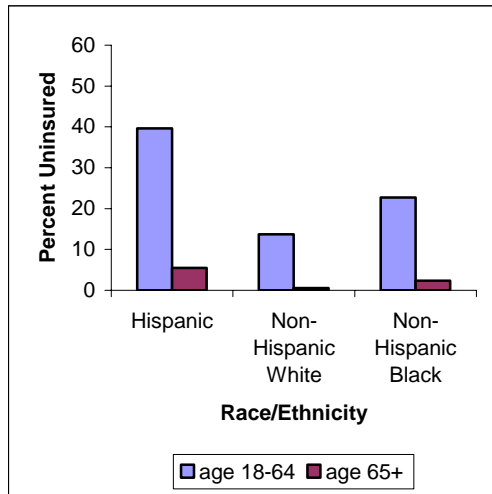

Figure 1: Percent uninsured by age and race and ethnicity: United States
Data Source: 2000-2002 National Health Interview Surveys

Within regional categories, Figure 2 shows that Hispanics living in the South and West had the highest uninsured rate as compared with the rest of the country. For all persons, the uninsured rate is highest in the South. Hispanics living in the South (46 percent), followed by the West, (39 percent) had the highest uninsured rate compared with other regions and other combinations of race and ethnicity.
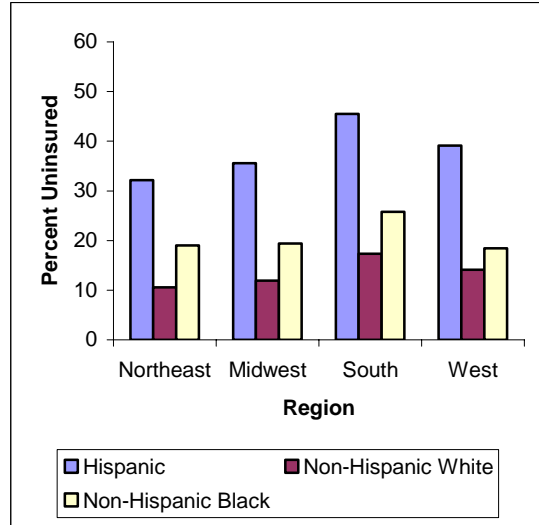

Figure 2: Percent uninsured by race and ethnicity and region, age 18-64
Data Source: 2000-2002 National Health Interview Surveys

As demonstrated in Figure 3, there is a strong relationship between education and uninsured status regardless of race and ethnicity. Hispanics had the highest level of uninsured within each category for education.
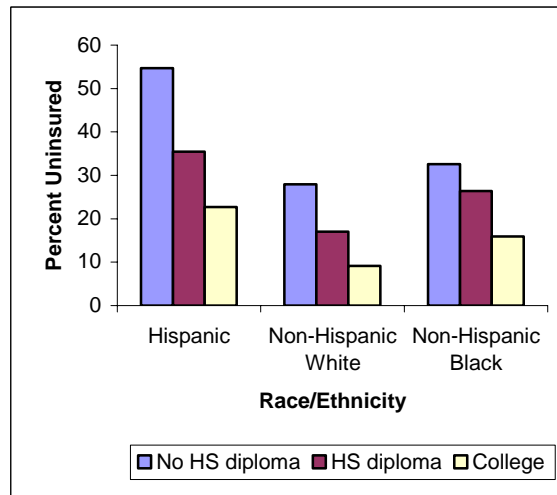

Figure 3: Percent uninsured by educational categories and by race and ethnicity, age 18-64
Data Source: 2000-2002 National Health Interview Surveys

We developed income to poverty ratios (IPR) by dividing the total family income by the poverty threshold. Table 1 shows the percent of uninsured people by race, ethnicity, and IPR. The percentage of uninsured for those in poverty (IPR<1.00) was 59 percent for Hispanics compared with 30 percent for non-Hispanic Whites. Hispanics with IPR of

3.00 or higher had a higher rate of being uninsured (31 percent).

Income to Poverty Ratio	Hispanic	Non-Hispanic White	Non-Hispanic Black
Under 1.00	59	30	38
1.00 - 1.49	55	34	35
1.50 - 1.99	46	26	31
2.00 - 2.99	34	18	21
3.00 or greater	31	10	17

Table 1: Percent uninsured by income to poverty ratio and race and ethnicity, age 18-64
Data Source: 2000-2002 National Health Interview Surveys

Figure 4 shows that Hispanics who spoke only Spanish in the interview had a higher uninsured rate (58 percent) compared with the English-speaking respondents (28 percent).
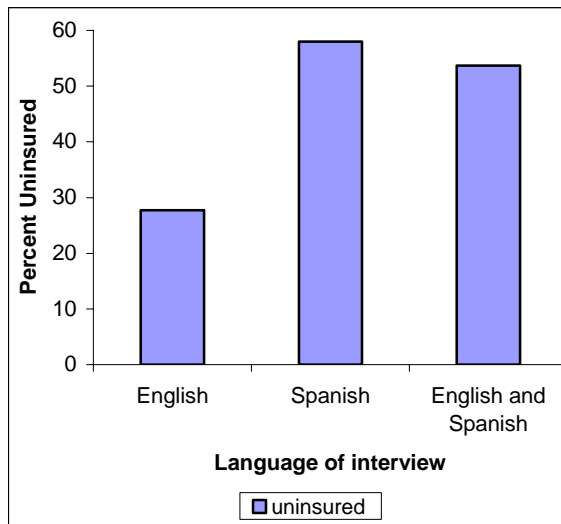

Figure 4: Percent uninsured Hispanic adults by language of interview, age 18-64
Data Source: 2000-2002 National Health Interview Surveys

Figure 5 shows that Hispanics' uninsured status varied substantially across ancestry. Hispanics with Mexican ancestry had the highest uninsured rate (54 percent) followed by Central/South/Latin Americans (43 percent).
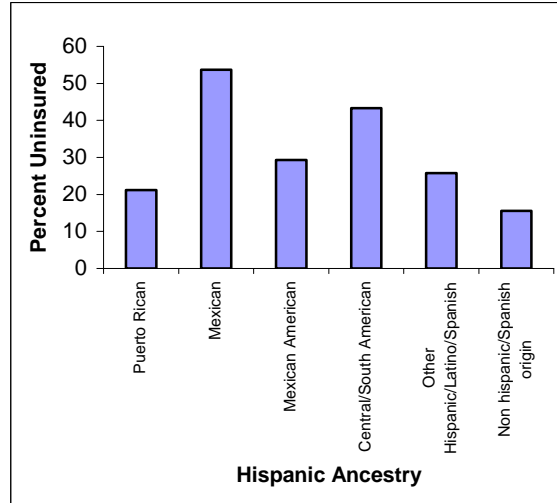

Figure 5: Percent uninsured Hispanic adults by ancestry, age 18-64
Data Source: 2000-2002 National Health Interview Surveys

Figure 6 shows that foreign-born Hispanics who are not citizens and have lived less than one year in the United States have the highest uninsured rate (82 percent) as compared with those who have lived in the United States longer. Foreign-born non-citizen Hispanics who lived in the United States for 15 or more years had an uninsured rate of 46 percent.
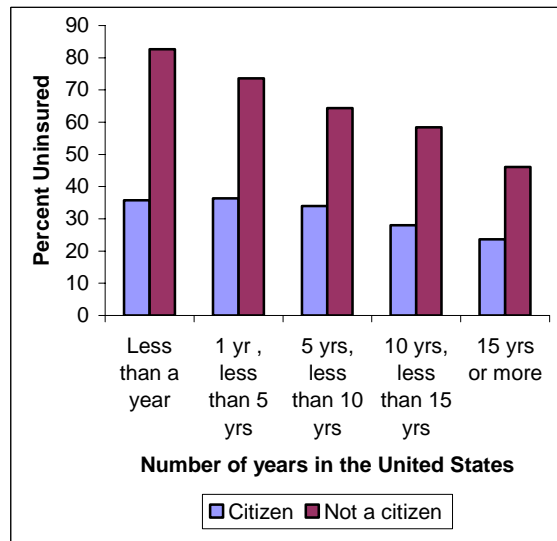

Figure 6: Percent uninsured foreign-born Hispanics by number of years in the United States and citizenship, age 18-64
Data Source: 2000-2002 National Health Interview Surveys

Adjusted odds ratios

Table 2, located at the end of the paper, shows that Mexicans are most likely to be uninsured compared

with other Hispanic ancestry groups. For example, the odds of being uninsured are 99 percent lower for Puerto Ricans and 25 percent lower for persons with Central/South/Latin American ancestry than for Mexicans.

Foreign-born Hispanics who have lived in the United States for less than one year are more likely to be uninsured than other Hispanic immigrants are. The adjusted odds of being uninsured for these immigrants are 251 percent higher compared with Hispanics who were born in the United States. Hispanics who spoke Spanish in the interview are more likely to be uninsured; the odds are 55 percent higher than for Hispanics who spoke English only in the interview.

The variables that are important for determining if anyone has health insurance coverage (Hispanic or non-Hispanic) are also important for Hispanics alone. This research focuses only on Hispanics. The odds of being uninsured are 35 percent higher for a Hispanic male than a Hispanic female. Married Hispanics are less likely to be uninsured than non-married Hispanics. Hispanics with no diploma are more likely to be uninsured than Hispanic persons with some college. The odds of being uninsured are 108 percent lower for a person with some college than for those persons without a high school diploma. Hispanics with a family size of six or more are more likely to lack health insurance coverage than those with a family size of one or two. The odds of being uninsured are 103 percent lower for Hispanics in the highest income category (the income to poverty ratio of 300 percent or above) compared with those Hispanics living in poverty. The odds for Hispanics who live in the South being uninsured are 70 percent higher than those who live in the West are.

Ranking of the influence of variables by R^2

Table 3 shows the additive effect of adding variables in terms of R-squared (R^2). The model had a total R^2 of 30.1 percent. The contribution to R^2 from the citizenship variable is 16.5, or over half of the total explained variation. Although educational attainment and poverty status were important in the regression, these were only control variables that are shared between Hispanics and non-Hispanics. The years lived in the U.S. explained 7.6 percent (=2.3/30.1) of the total explained variation. The remainder of the independent variables that are specifically "Hispanic" explained little of the variation. The results of this test for importance suggest two

variables of interest, citizenship and years lived in the United States for foreign-born Hispanics.

Effect	Chi-Square	Total R^2	Contribution to R^2
Citizenship	4,887	16.5	16.5
Educational level	1,277	20.4	3.9
Poverty status	820	22.8	2.4
Years lived in the U.S.	748	25.1	2.3
Marital status	589	26.8	1.7
Region	399	27.9	1.1
Hispanic ancestry	321	28.8	0.9
Language of interview	245	29.5	0.7
Sex	89	29.7	0.2
Employment status	78	30.0	0.3
Family size	50	30.1	0.1

Table 3: Importance of independent variables measured by contribution to R^2
Data Source: 2000-2002 National Health Interview Surveys

5. Discussion

Citizenship, level of education, poverty status, years lived in the United States, marital status, region, Hispanic ancestry, language of interview, sex, employment status and family size are correlated with the uninsured rate among Hispanic subgroups living in the United States.

This research has indicated differences in the Hispanic population that may explain the "Hispanic" effect on uninsured rates. The results of logistic regression showed two variables are uniquely important for the Hispanic population and possibly for all immigrants: citizenship and foreign-born interacted with years lived in the United States. Because non-citizens may not be eligible for or able to obtain coverage through public assistance or employment based programs because of their immigration status, these findings should cause pause for health insurance researchers that rely solely on a Hispanic variable without further analyzing the citizenship status of the uninsured.

As the figures indicated, non-Hispanics are more likely to have health insurance coverage than Hispanics. These differences may continue to persist when all of the relevant Hispanic characteristics are accounted for, but this has not been shown in this research. For example, the return for higher levels of education may be lower

for Hispanics than non-Hispanics leading to a lower likelihood of being insured. Studying these differences through econometric methods is left for another study.

Future research will be required in predicting the uninsured rates for counties. This research has been done at the person level, and there are county level implications that should be tested in the future. Further research will first investigate whether immigrants of Hispanic origin are different than other immigrants. We will then investigate incorporating the county-level variables for foreign-born non-citizens for inclusion in the SAHIE program's health insurance model. The number of persons that are foreign-born non-citizens can be easily incorporated into the SAHIE program's model using American Community Survey (ACS) data when they are available or population estimates.

References

Commonwealth Fund (2003). "Insurance, Access, and Quality of Care Among Hispanic Populations" For the National Alliance for Hispanic Health meeting, October 15-17, 2003.

DeNavas-Walt, C., Proctor, D. and Mills, R. (2004). *Income, Poverty, and Health Insurance Coverage in the United States: 2003.* U.S. Census Bureau, Current Population Reports, P60-226. Washington, DC: U.S. Government Printing Office.

National Center for Health Statistics (NCHS) (2003a). Health, United States, 2003. Hyattsville, MD: 2003. Available at: www.cdc.gov/nchs/data/hus/hus03.pdf.

National Center for Health Statistics (NCHS) (2003b). National Health Interview Survey (NHIS) Public Use Data Release, NHIS Survey Description. Hyattsville, MD. 2003. Available at: http://www.cdc.gov/nchs/about/major/nhis/quest_data_related_1997_forward.htm.

PEW Hispanic Center (2005). "Hispanics: A People in Motion" www.pewhispanic.org.

SAS software, Version 9.1 of the SAS System for PC. SAS Institute Inc., Cary, NC, USA.

Table2: Odds Ratio Estimates				
Characteristic	Variable	Increase in odds	90% Wald Confidence Limits	
Sex (comparison is Female)	Male	0.35	0.29	0.40
Hispanic Origin	Puerto Rican	-0.99	-1.18	-0.81
(comparison is Mexican)	Mexican American	-0.17	-0.25	-0.10
	Central/South/Latin Americans	-0.25	-0.33	-0.18
	Other Hispanics	-1.07	-1.24	-0.91
Educational Level	Some College	-1.08	-1.20	-0.97
(comparison is no diploma)	HS Diploma	-0.30	-0.37	-0.24
Marital Status	Divorced	0.55	0.43	0.69
(comparison is married)	Separated	0.58	0.43	0.75
	Single/never married	1.04	0.94	1.14
	Widowed	0.69	0.44	0.99
Family Size	3 ,4 and 5	0.04	-0.01	0.10
(comparison is 1 and 2)	6 or more	0.28	0.20	0.37
Employment Status (comparision is not working)	Working	-0.27	-0.34	-0.22
Income to poverty ratio	1.00 to 1.49	-0.04	-0.12	0.04
(comparison is "in poverty")	1.50 to 1.99	-0.45	-0.57	-0.34
	2.00 to 2.99	-0.83	-0.98	-0.70
	3.00 and over	-1.03	-1.16	-0.91
Citizenship (comparison is not a citizen)	Citizen	-1.08	-1.21	-0.95
Years Lived in U.S.	Less than 1 year	2.51	1.74	3.49
(comparison if born in the US)	1 yr., less than 5 yrs.	1.33	1.11	1.57
	5 yrs., less than 15 yrs	0.34	0.23	0.45
	15 years or more	-0.18	-0.26	-0.10
Language Spoken in Interview	Spanish	0.55	0.47	0.63
(comparison is English only)	English and Spanish	0.49	0.40	0.58
Region	Midwest	-0.05	-0.14	0.03
(comparison is West)	Northeast	0.14	0.06	0.23
	South	0.70	0.62	0.78

Note: The others group (don't know, not ascertained, unknown and refused responses) were deleted from the table, but not from the underlying analysis.

Data Source: 2000-2002 National Health Interview Surveys