

WIC Data mining survey: Survey and analysis methodologies

Technical report

WIC data mining survey methodologyⁱ

The WIC data mining survey, commissioned by Public Health Foundation Enterprises WIC Program (PHFE-WIC) and funded by First 5 Los Angeles, was conducted on a sample of about 5,000 adult WIC participants in Los Angeles County (LAC) in both 2005 and 2008. PHFE-WIC contracted with Field Research Corporation (FRC), an independent public opinion research organization, to administer the survey.

The survey's population of interest was WIC participants living in LAC, and thus phase 1 of the sampling consisted of taking a simple random sample of all WIC participants throughout the county. In phase 2 of the sampling, a simple random sample of residents of Antelope Valley, a sparsely populated region of LAC, was added to the survey population in order to increase the predictive validity of survey findings among residents in this area.

Before calling each of the WIC recipients, PHFE-WIC mailed out a post card to each randomly selected household alerting them that a Field interviewer would be calling on behalf of PHFE-WIC within the coming weeks. At the start of each telephone interview Field interviewers first verified that the WIC recipient was indeed a county resident, and for the Antelope Valley Sample that the recipient resided in a city or town within that portion of the county. In addition, each parent was asked to confirm that they or their child was still receiving WIC benefits. Those listings not meeting either the residency requirement or not receiving WIC benefits were dropped from the survey sample and replaced with other randomly selected WIC recipients.

The survey questionnaires were administered to WIC recipients in English and Spanish. After a penultimate version of the English language questionnaires was developed, it was translated into Spanish by professional translators working under Field's employ. These translations were then reviewed by a second professional translator, as well as the PHFE-WIC staff.

A cooperation rate describes the ratio of completed interviews as a percentage of all eligible units ever contacted. Household level cooperation rates are based on contacts made with a household member and are the most widely cited cooperation rates in RDD surveys. The broadest standard measure of household-level cooperation is the American Association for Public Opinion Research's (AAPOR's) cooperation rate 1 (COOP1), or minimum cooperation rate. Applying this to the PHFE-WIC survey yields an overall household cooperation rate of 92.6 percent.

A response rate is an even broader measure of sample implementation and is defined as the ratio of completed interviews to the total number of eligible reporting units. Using AAPOR's Response Rate 2 (RR2) for calculating the overall response rate for the PHFE-WIC survey yields a 66.5% response rate.

In order to combine the survey data collected from both the Cross-Section Sample and the Antelope Valley Sample Augment, it was necessary to develop appropriate sampling weights aligning the overall data file to PHFE-WIC's estimates of the total distribution of WIC recipients across the county's eight Service Planning Areas. These weights return the allocation of interviews conducted in the Antelope Valley region to their proper proportion of all WIC recipients countywide.

For more information on the WIC data mining survey, including data access questions, please contact Dr. Shannon Whaley and Lu Jiang of the PHFE WIC Data Mining team at datamining@phfewic.org, or visit their website at <http://www.phfewic.org/DataMiningNew>.

Analysis methodology for WIC brief "Latino Children Showing Higher Rates of Preschool Enrollment"

In order to estimate the true change in preschool enrollment among Latinos, rather than reflecting demographic changes in Los Angeles such as changes in employment rates or education levels, it was important to conduct a multivariate analysis where such demographic information was controlled for.

Each of the outcome variables, including preschool enrollment, difficulty enrolling in preschool, perceived cost barriers, and perceived access barriers, were predicted using the same model. This model incorporated the following independent predictor variables: Household composition, indicating whether both parents lived in the household, whether the mother is employed part-time or full-time, whether the mother was born outside of the United States, whether the mother completed a High School education, and mother's race. Mother's race was interacted with the year of the analysis in order to determine changes in different racial/ethnic groups from the 2005 to 2008 survey.

After the logistic regression analysis was conducted, predicted probabilities were calculated for each subgroup of interest. In this brief, we conducted predicted probabilities for Latinos in 2005 and 2008. These predicted probabilities are reported in the brief, "Latino Children Showing Higher Rates of Preschool Enrollment," which is posted on the First 5 Los Angeles website at <http://www.first5la.org/research/phfe-wic>.

For more information, including detailed logistic regression tables, please contact Teryn Mattox, at tmattox@first5la.org.

ⁱ In this section we have taken content extensively from Field Research Corporation's report to PHFE-WIC, *A Summary of the Survey Methods Used to Conduct the 2005 Survey of WIC Recipients in Los Angeles County*.