FUND ALLOCATION AND SMALL AREA ESTIMATION IN THE WIC PROGRAM

Allen L. Schirm, Mathematica Policy Research, Inc. and Cindy Long, U.S. Department of Agriculture
Allen Schirm, Mathematica Policy Research, 600 Maryland Ave., S.W., Suite 550, Washington, DC 20024-2512

Key Words: Nutrition, Bayes, Shrinkage

1. INTRODUCTION

The Special Supplemental Nutrition Program for Women, Infants, and Children (WIC) is a federal grant program for states and Indian Tribal Organizations (ITOs) that is administered by the Food and Consumer Service (FCS) of the U.S. Department of Agriculture (USDA). The program provides nutrition and health assistance services for low-income childbearing women, infants, and children. Pregnant and postpartum women and children under the age of 5 who have family incomes at or below 185 percent of the applicable poverty guidelines ($27,380 for a family of four as of July 1, 1994) are "income-eligible" to participate. To be fully eligible, participants must be determined to be at nutritional risk based on a medical and/or nutritional risk factor identified by a competent health professional. Persons exhibiting a medical risk factor (such as anemia, underweight, or diabetes) receive higher priority than persons at risk of inadequate nutrition.

WIC participants receive a federally-prescribed package of foods designed to meet their specific nutritional needs. For example, a WIC package for a newborn infant would consist entirely of infant formula, while a child's package would contain such items as milk, cheese, peanut butter, cereal, and juice. WIC participants also receive nutrition education and counselling and access to health and social services.

Research has demonstrated that WIC has been successful in reducing infant deaths, low birthweight, premature births, and other health problems. The program has also improved the nutrition and health care use of participants.

WIC is not an entitlement program--the number of participants in each year is limited by the funds appropriated. Funds are allocated among states based on two formulas established by regulation, one for food and one for administration and nutrition services. The food funding formula has been modified several times since the program's inception in the early 1970s to reflect changes in the size of the program and changing program priorities.

The size of the WIC-eligible population in each state has always been a factor in determining state WIC grants, and the availability of timely and accurate data on the number of women, infants, and children income-eligible for WIC has long been an issue of concern. However, the relative importance of these data and other components in the funding formula has varied over time.

President Clinton's first budget highlighted the expansion of WIC as a major priority, and set as a goal "fully funding" the program, that is, providing enough funds to allow all eligible persons who want to participate to do so. The WIC program has already expanded significantly in recent years. From 1989 to 1994, annual WIC appropriation increases ranged from 9 to 12 percent per year. This program growth, coupled with the expected move toward full funding, prompted FCS to revise the funding formula in fiscal year 1994. This revision included a substantially increased emphasis on the eligibles data as a basis for funds allocation, and brought the need for timely and accurate data to the forefront.

2. EVOLUTION OF THE WIC FOOD FUNDING FORMULA

The WIC program was established in 1972, and for the first several years of operation, state grants were determined at the discretion of the USDA. In 1979, a food funding formula was formally established by regulation. This formula set forth two essential components that were the primary basis for food funds distribution until 1987. First, states were provided with their prior year funding level, plus some adjustment for inflation, assuming adequate funds were available. Any remaining funds were allocated based on a "growth" calculation, which attempted to direct funds to states on the basis of need for the program. Specifically, the formula considered each state's share of the estimated national population of income-eligible women, infants, and children, and its relative health status, as measured by the state's infant mortality rate and/or low birthweight rate relative to the national average.

The original 1979 funding formula placed equal weight on the income-eligibles data and the health data. However, in 1984, the formula was revised to place much greater emphasis on the income-eligibles data, increasing their weight in the growth calculation to a minimum of 80 percent (and in some cases much closer to 100 percent). This change reflected a belief that the eligibles data best indicated relative need among states.

In 1987, a major change occurred when a "targeting" component was added to the food funding formula. This component reflected concern that the limited WIC funds be allocated based not only on the estimated need for the program, but also on states' demonstrated ability to serve those at highest risk. In proposing the new targeting component, USDA stated its concern that the existing funding formula did not discriminate between state agencies that had targeted this population effectively and those that had not. Thus, the 1987 regulation required that after stability grants were made, half of any additional funds be distributed based on states' shares of the national population of "high priority" participants (defined as women, infants, and children with a demonstrated medical risk). The remaining half was distributed based on the growth calculation (that is, using the states' shares of the eligible population and the state health indicators). States with well-targeted programs according to the new measure also received a larger inflation adjustment.

In 1994, USDA determined that several components of the 1987 rule had become outdated. In close consultation with the states, USDA concluded that the program had expanded so significantly that it was no longer necessary to provide incentives for states to serve those at highest risk. In addition, the targeting component was viewed as an obstacle to achieving funding equity among states. Thus, the 1994 funding rule eliminated both the participation-based targeting component, as well as the portion of the growth component that measured relative health status. The new rule stated that after stability grants had been provided, all remaining funds were to be allocated based on states' estimated shares of the national eligible population.

The new funding formula was structured to allocate any growth funds to states whose current resources were less adequate for serving their estimated eligible populations relative to other states. This is accomplished using a "fair share" concept. A state's fair share of available funds is its share of the estimated national population of persons eligible for the program. Thus, a state with one percent of the eligible persons has a fair share of one percent of total available food funds. States whose stability grants are less than their fair shares receive growth funds. The amount of growth funds received by an "under fair-share" state is directly proportional to the difference between the stability grant and the fair share. States with stability grants in excess of their fair shares do not receive growth funds (unless all the "under fair-share" states decline to accept the full amount of growth funds available).

3. WIC ELIGIBLES DATA FOR STATES

The establishment of the fair share component of the funding formula, combined with the rapid growth of the program, heightened USDA's concern about the quality of state eligibles estimates. Under previous rules, census data were specifically identified as the source for calculating states' shares of the eligible population. Data from the 1980 census were used from the early 1980s until 1994, when 1990 census data were used. WIC-eligible infants and children were estimated directly from census counts of infants and children at or below 185 percent of the poverty line. Income-eligible pregnant and postpartum women were estimated indirectly based on the counts of income-eligible infants.

The lack of timeliness of the census data was a longstanding concern, and the 1987 funding rule noted that the census data were flawed in this respect. However, no alternative data source was put forth. In the 1994 rule, no specific source of WIC eligibles data was identified in order "to allow for the use of the most timely and reliable data as it becomes available." In discussions with the states, USDA committed to place high priority on developing an alternative source to the census for the eligibles data. State WIC agencies desired improvements in the eligibles data not only for allocating funds, but also for using the data as a benchmark to assess program performance.

FCS, through contract research, undertook to develop a new methodology for estimating the population income-eligible for WIC in each state. The research focused on the Current Population Survey (CPS) as the likely alternative to the census for state-level income data. FCS identified several criteria for the state-level eligibles estimates, including: consistent methodology and data sources across states; an understandable, technically sound methodology; an ability to update the estimates annually and to capture year-to-year changes in states' relative positions; and the use of data that are as current as possible.

4. EVOLUTION OF THE ESTIMATION METHODOLOGY

The methodology for estimating WIC eligibles has evolved more slowly than the formula for allocating WIC funds. The first generation estimator, which used census data, was essentially unchanged from 1979 to 1994. It was just last year that a second generation estimator was implemented, although a third generation estimator has been

developed this year. The next three sections of this paper describe the first and second generation WIC eligibles estimators and briefly introduce the third generation estimator.

4.1 The First Generation Estimator

According to the first generation WIC eligibles estimator, the current number of eligibles is the same as the number measured from the most recent census data available. Although census estimates are derived from large samples and, therefore, are precise, there was widespread dissatisfaction with the first generation estimator among policy analysts, WIC administrators, and advocates. They believed that census estimates were very inaccurate, failing to account for often rapidly changing economic conditions. According to census and CPS data, the recession of 1990-1991 was accompanied by a 20 percent increase nationwide in the number of eligible infants and children between 1989 (the year to which census income data pertain) and 1992. That increase, amounting to nearly 1.5 million infants and children, cannot be explained by population growth. The percentage of infants and children eligible for WIC rose by nearly 6 percentage points--from about 38 percent to 44 percent.

Such rapid growth in eligibles certainly creates difficulties for program planning and performance monitoring, key uses of WIC eligibles estimates. However, census data could still provide accurate estimates of fair shares for distributing program funds if the numbers of eligibles grew by the same proportion in every state. But, the recent growth in eligibles seems to have been spread very unevenly across the states. According to the 1990 census and the March 1993 CPS, the percentage of infants and children eligible rose between 1989 and 1992 by over 8 percentage points in Florida, New York, California, and New Jersey, but by under 3 percentage points in Massachusetts, Ohio, Michigan, and Pennsylvania.

These figures strongly suggest that by assuming nothing is changing over time, the first generation estimator may be badly biased. As an alternative to the census, the CPS provides the most timely data for developing annual state estimates of WIC eligibles. However, despite their timeliness, direct CPS estimates are imprecise because state samples of infants and children are small in all but a few states. This fundamental problem of small area estimation--the lack of data, that is, the small number of sample observations--led to the original adoption and continued use of the first generation estimator, an "indirect" estimator that "borrows strength" from the past. However, it seems that the gain in precision from using census data comes at the potential cost of substantial bias. Therefore, we sought an alternative estimator to minimize the tradeoff between bias from lack of timeliness and imprecision from lack of data. Based partly on the findings of Schirm (1994), who assessed the relative accuracy of several different estimators of state poverty rates, we began development of a second generation estimator that uses "shrinkage" methods.

4.2 The Second Generation Estimator

The second generation WIC eligibles estimator is a Bayesian shrinkage estimator that optimally averages CPS direct sample estimates and predictions from a regression model. As we will see, the shrinkage estimates obtained are more timely than census estimates, and substantially more precise than CPS estimates. This section describes our eight-step procedure for estimating the numbers of infants and children who were income eligible for WIC in each state in 1992. Additional technical details can be found in Schirm (1995).

Step 1: From the most recent census (1990), derive state estimates of the percentage of infants and children who were income eligible. Because the family income data collected in the census pertain to the preceding calendar year, the eligibility estimates are for 1989. We estimated the percentages, rather than the numbers, of eligible infants and children to standardize for state population size.

Because census samples for states are very large, the estimates are precise. However, they quickly become "old" if economic conditions have changed substantially since the census.

Step 2: From the most recent CPS (March 1993), derive state sample estimates of the percentage of infants and children who were income eligible. The most recent CPS that has income data for families provides more timely information than the census. That CPS was the March 1993 CPS when we were developing eligibles estimates for allocating fiscal year 1995 funds. Like the census, the CPS collects family income data for the prior year. Thus, the sample estimates pertain to 1992.

Although timely compared with the census estimates, the CPS sample estimates are relatively imprecise. The standard errors for the CPS estimates tend to be large, so our uncertainty is great. For example, according to widely used statistical standards, we can be confident only that the percentage of income-eligible infants and children in Delaware was between 22.5 percent and 41.6 percent. This range is so wide and our uncertainty so great because the CPS samples of

infants and children in each state are small. Indeed, that is why we derived an eligibility estimate for infants and children combined, rather than separate estimates, one for infants and one for children. In the March 1993 CPS, there are data for fewer than 30 infants for most states.

Step 3: Construct sample estimates of the change in the percentage eligible between 1989 and 1992. A sample estimate of the change in the percentage eligible between 1989 and 1992 was calculated by subtracting the census estimate for 1989 from the CPS estimate for 1992. We calculated sample estimates of change for use in the regression and shrinkage estimation described in the next few steps. Focusing on the change in the percentage eligible between 1989 and 1992, rather than just the percentage eligible in 1992, is a simple way to reflect a strong systematic relationship: states with a high percentage eligible in 1989 tend to have a high percentage eligible in 1992, and states with a low percentage eligible in 1989 tend to have a low percentage eligible in 1992. In principle, our shrinkage method obtains better estimates by using information on not only where a state "is," but also where it "began."

Step 4: Using a regression model, predict the change in the percentage eligible for each state based on observed changes in (i) Food Stamp Program (FSP) participation, (ii) Unemployment Insurance (UI) program participation, and (iii) per capita income. The main limitation of the sample estimates derived in the previous step is imprecision. Regression can reduce that imprecision. Regression estimates are predictions based on nonsample or highly precise sample data, such as census and administrative records data. The latter include government program case files and vital statistics.

Regression estimates are points on a regression line, a line obtained by regressing the sample estimates from the previous step on predictor variables. The three predictor variables we used measure the changes between 1989 and 1992 in (1) FSP participation, (2) UI program participation, and (3) per capita income. These three were selected as the best predictors from a longer list. As expected, the estimated regression shows that states with relatively large increases in FSP and UI program participation and large decreases in per capita income tend to have relatively large increases in the percentage of infants and children eligible for WIC. Standard errors for regression estimates are much smaller than standard errors for sample estimates.

Comparing how the sample and regression estimators use data reveals how the regression estimator "borrows strength" to improve precision. When we derived sample estimates in Step 3, we used only data from Delaware to estimate the change in the percentage of eligible infants and children in Delaware, even though Delaware, like nearly all states, has a small CPS sample. Deriving regression estimates in this step, we estimated a regression line from sample and administrative records data for all the states and used the estimated line (with administrative records data for Delaware) to predict the change in WIC eligibles for Delaware. In other words, the regression estimator not only uses the sample estimates from every state to develop a regression estimate for a single state but also incorporates data from outside the sample, namely, administrative records data. The regression estimator improves precision by using more data to identify states with sample estimates that seem too high or too low because of sampling error, that is, error from drawing a sample that has a higher or lower percentage eligible than the entire state population has. For example, suppose a state had experienced stable FSP and UI program participation and rising per capita income. Our regression estimator would predict a stable or declining percentage of eligible infants and children, implying that a sample estimate showing a large increase in WIC eligibles is too high. The regression estimate will be lower than the sample estimate for such a state. On the other hand, if the sample data for a state show a much smaller increase in eligible infants and children than expected in light of the observed changes in FSP and UI program participation and per capita income, the regression estimate for that state will be higher than the sample estimate.

Step 5: Using "shrinkage" methods, average the sample estimates of change and the predictions of change. As noted, the limitation of the sample estimator is imprecision. The limitation of the regression estimator is bias. Some states really have larger or smaller increases in WIC eligibles than we expect (and predict with the regression estimator) based on changes in FSP and UI program participation and per capita income. Such errors in regression estimates reflect bias.

These limitations arise for the following reasons. The sample estimator uses only sample data for one state to obtain an estimate for that state. It does not use sample data for other states or administrative records data. Although the regression estimator borrows strength, using data from all the states and administrative records data, it makes no further use of the sample data after estimating the regression line. It treats the entire difference

between the sample and regression estimates as sampling error, that is, error in the sample estimate. No allowance is made for prediction error, that is, error in the regression estimate. Although not all, if any, true state values lie on the regression line, the regression estimator acts as though they do.

Using all of the information at hand, a shrinkage estimator addresses the limitations of the sample and regression estimators by combining the sample and regression estimates, striking a compromise. A shrinkage estimator takes a weighted average of the sample and regression estimates. We calculated weights using Bayesian methods, as described in Schirm (1995). Generally, the more precise the sample estimate for a state, the closer the shrinkage estimate will be to it. The larger samples drawn in large states support more precise sample estimates, so shrinkage estimates tend to be closer to the sample estimates for large states. Given the precision of the sample estimate for a state, the weight given to the regression estimate depends on how well the regression line fits, that is, whether we could find good predictors reflecting why some states have larger increases in WIC eligibles than other states. The shrinkage estimate will be closer to the regression estimate and farther from the sample estimate when we find good predictors and the regression line fits well. Striking a compromise between the sample and regression estimators, the shrinkage estimator strikes a compromise between imprecision and bias. The sample and regression estimates are optimally weighted to improve accuracy by minimizing a measure of error that reflects both imprecision and bias. By accepting a little bias, the shrinkage estimator may be substantially more precise than the sample estimator. By sacrificing a little precision, the shrinkage estimator may be substantially less biased than the regression estimator.

Step 6: Add the shrinkage estimate of the change between 1989 and 1992 to the census estimate of the percentage eligible in 1989 to get a shrinkage estimate of the percentage eligible in 1992.

Step 7: Multiply the shrinkage estimate of the percentage eligible by the state population of infants and the state population of children to get preliminary shrinkage estimates of the numbers of eligible infants and children. To obtain separate estimates for infants and children, we have assumed that the percentage of infants who were income eligible in a state is the same as the percentage of children who were income eligible, an assumption well-supported by census data. Our estimate of that percentage was obtained in Step 6.

To obtain estimated numbers from estimated percentages, we require state population estimates for both infants and children. The population estimates we used pertain to the resident population on July 1, 1992 and were developed by the U.S. Bureau of the Census from census and administrative records (mainly vital statistics) data.

Step 8: Control the preliminary state shrinkage estimates of the numbers of eligible infants and children to sum to the national totals for eligible infants and children obtained from the CPS. The most recent national CPS sample estimates are typically used to develop the budget for the WIC program. To obtain final shrinkage estimates for states that sum (aside from rounding error) to the national totals from the most recent CPS (March 1993), we ratio adjust the preliminary state shrinkage estimates. This ensures that the estimates used to allocate funds are consistent with the estimates generally used to determine total program funding. The adjustments were small, with ratios of about 0.99 and 1.03 for infants and children, respectively.

Second Generation Estimates. The strengths of the shrinkage estimates obtained from our second generation estimator are that they are more timely than census estimates and substantially more precise than direct CPS estimates. As documented in Schirm (1995), the shrinkage estimates have much smaller standard errors and much narrower confidence intervals than CPS estimates. A shrinkage confidence interval is, on average, 61 percent narrower than the corresponding direct sample confidence interval. According to rough calculations, that is about the same gain in precision that would be obtained from increasing the sample size of the CPS from under 60,000 households to nearly 400,000 households--a 6.5-fold increase.

While using a shrinkage estimator greatly narrows confidence intervals and reduces our uncertainty, using shrinkage estimates makes an important difference in how WIC funds are distributed. Table 1 shows that there are several large differences in fair shares when the fair shares are calculated using shrinkage rather than census estimates. Even a small difference in fair shares can affect a state's WIC grant, however, because the funding formula contains a threshold. Specifically, a state receives growth funds only if its stability funding is below its fair share. A small increase in a state's fair share may make it eligible for growth funds, while a small decrease might make it ineligible.

FCS used the shrinkage estimates of infants and children income-eligible for WIC in 1992 to

determine state WIC food grants for fiscal year 1995. Over $125 million in growth funds were distributed.

4.3 The Third Generation Estimator

Our second generation estimator borrows strength to improve precision. However, there is more strength to be borrowed. The second generation estimator uses census estimates for the "base" year (1989) and CPS estimates for the "current" year (1992 for the most recent set of estimates). Estimates for intervening years are not used, although CPS data for obtaining such estimates are available. With each intervening year, we are ignoring more information that could be relevant. An unusually large increase in WIC eligibles over three years, for example, would be more plausible and look less like sampling error if it appeared to consist of a series of modest increases rather than two small decreases followed by one enormous jump. Using data for only the base and current years, however, the second generation estimator cannot distinguish those two patterns of change. Also, with the second generation estimator, we can do little if our regression model seems to persistently under- or overpredict for a state, short of finding a predictor variable that explains why that state is different from all the other states.

An advantage of Bayesian shrinkage methods is that they allow additional data to be used in a systematic, rather than ad hoc, way. We have developed a third generation estimator that is both domain and time indirect, borrowing strength across not only states, but also time. The state WIC eligibles estimates for 1993 were derived from census data for 1989, as well as CPS data for 1990, 1991, 1992, and 1993. Administrative records data for all five years (1989-1993) were used. The third generation estimator takes account of correlations among sample estimates from different years and correlations among regression model prediction errors from different years.

5.0 CONCLUSION

Since its inception, the WIC program has undergone important changes. Over the last few years, funding for the program and the number of program participants have risen dramatically. The formula for allocating funds to the states was also revised significantly to place greater emphasis on the number of eligibles in each state as the basis for calculating state WIC grants. This new emphasis, coupled with rapid program growth, heightened the need for timely, accurate state estimates of WIC eligibles. Responding to that need, we have developed a Bayesian shrinkage estimator that optimally averages CPS sample estimates with predictions of WIC eligibles from a regression model. The predictions are based on observed changes in government program participation and other indicators of socioeconomic conditions. The shrinkage estimates obtained are more timely than census estimates, which had been used for fund allocation prior to the use of shrinkage estimates for fiscal year 1995, and substantially more precise than direct CPS estimates. The shrinkage estimator improves precision by borrowing strength, using data from all the states to derive each state's estimate. We have recently developed a new shrinkage estimator to take account of even more information and borrow strength across both space and time. The new estimator provides even better estimates for allocating WIC funds to the states.

REFERENCES

Schirm, Allen L. (1995), *State Estimates of Infants and Children Income Eligible for the WIC Program in 1992.* Washington, DC: Mathematica Policy Research.

Schirm, Allen L. (1994), "The Relative Accuracy of Direct and Indirect Estimates of State Poverty Rates," *1994 Proceedings of the Section on Survey Research Methods.* Alexandria, VA: American Statistical Association.

Table 1. Effect of Using Shrinkage Rather than Census Estimates

Percentage Point Change in Fair Share			States	
2.0	to	3.0	1	(CA)
1.0	to	2.0	0	
0.5	to	1.0	1	(FL)
0.2	to	0.5	3	(NY, NJ, MD)
0.1	to	0.2	1	(CT)
0.0	to	0.1	13	
−0.1	to	0.0	16	
−0.2	to	−0.1	11	
−0.5	to	−0.2	5	(OH, LA, PA, MI, WI)