Bridging the Gap Between the Old and New Race Categories: Comparing Population Estimates Prepared Using Two Different Sets of Bridging Factors.

Amy Symens Smith (Census Bureau) and Deborah D. Ingram (NCHS)

The switch from a "mark one race" question to a "mark one or more races" question in Census 2000 left many data users unable to measure changes by race and Hispanic origin between Census 2000 data and previous censuses, current surveys, administrative records data, and population estimates. One such example is the challenge faced by the National Center for Health Statistics (NCHS) to accurately measure post-2000 changes in births and deaths, with a reporting system that still collects vital events categorized using the old race standards.

To meet this need, NCHS led research in 2003 to design a method to "bridge" between the old and new race categories. NCHS' National Health Interview Survey (NHIS) data provides a unique opportunity to do such research because since 1982 respondents have been given the option to indicate more than one race. Respondents who do so are then asked the follow-up question "which single race best represents your race."

To estimate bridging factors, NCHS used the survey data to fit individual logistic and multi-logit models to the large multiple race groups with a composite model for the small groups. The bridging models included demographic covariates such as age, sex, and Hispanic origin and county-level contextual variables such as region, urbanization level, percent in single-race categories and percent multiple-race population. The outcome was a set of probabilities for each multiple-race group which "bridge" to the single-race component groups by state, county, age, sex, and Hispanic origin.

Once the bridging probabilities were determined, they were delivered to the Census Bureau and applied to the population estimates, producing estimates of the post-2000 population consistent with the old race standards. The population counts by race for each state, county, age, sex, and Hispanic origin group were then returned to NCHS to be used as the denominators in the published vital statistics rates. These same data were made available on NCHS' website to public and private data users.

The Census Bureau and NCHS want to update this research because literature suggests that individual's race reporting changes across time and circumstances. With this in mind new bridging factors are to be constructed based on updated models. These models will be applied to additional years of NHIS data, as well as, to the Census Bureau's Census Quality Survey (CQS).

The CQS is a stratified random sample of approximately 55,000 households conducted in the spring of 2001. The CQS used a split-panel design with two panels each contacted at two different points in time. Panel A received the "mark one or more races" question first followed by the "mark one race" question. Panel B received the opposite instruction, first being asked to "mark one race" and then asked to "mark one or more races". Similar

to the NHIS, the CQS allows researchers to determine how individuals "bridge" between the old and new race categories.

In the fall of 2005, NCHS will finalize the new logistic and multi-logit models and apply them to the CQS data and the NHIS data. From this, new CQS bridging probabilities will emerge which will be applied by the Census Bureau to the recent population estimates.

This paper will summarize the two-way comparison made of the population counts produced using: (1) the original NHIS bridging factors; (2) the new NHIS bridging factors. Mapping tools will be used to highlight differences in the national, state, and county race distributions. Once differences have been identified we will return to the probability files to determine why these differences occurred.

The final outcome of this research will be a decision by NCHS concerning which set of bridging factors will be the preferred set. Once this is determined the population estimates will be bridged and used by NCHS as population denominators to measure recent changes in birth and death rates and to construct measures of infant mortality and life expectancy, to name a few. These same data will be available to public and private data users needing to measure changes in race and Hispanic origin that occurred before and after the new race standards were implemented.