



RACHSING and PRED* Variables: What they are, why they are needed, and how to use them.

March 23, 2022

QUESTIONS & ANSWERS

The following are questions received during the webinar and their answers. For additional questions or clarifications, contact IPUMS User Support at ipums@umn.edu.

1. Would PRED* differ for someone who reports "Chinese and White" versus someone who reports "White and Chinese," if the people were equal in other respects?

No. The data do not distinguish which answer they gave first.

2. Which years of data were used to set up individual-level predictions/regressions?

The last year of the four-year pooled NHIS data matches the year of the ACS or census. For example, the prediction model coefficients from the 2008-2011 NHIS were used for

-the 2011 single-year ACS and

-the 2011 3-year ACS (which is 2009-2011 pooled ACS) and

-the 2011 5-year ACS (which is 2007-2011 pooled ACS).

The NHIS stopped asking the key question in 2018, so the 2015-2018 prediction model coefficients are used for 2019 and later.

3. Will the working paper include the model coefficients so that we could, in theory, apply the model to other data sets?

Yes. That is a major goal of releasing the model coefficients!

4. Should we start moving towards using the PRED* variables as the main controls for race, treating it as more of a continuum instead of discrete categories?

Using the PRED* variables is preferable for two reasons:

-People who reported multiple races did this on purpose and the PRED* variables treat this as important information. They have non-zero values for each of the federal race groups reported by a multiple-race person.

-The PRED* variables are the output of the prediction equations and are not rounded. RACHSING introduces rounding error.

5. How would these variables impact analyses that would otherwise only look at single race American Indian/Alaska Native or single race Asian/Pacific Islander?

RACHSING and the PRED* variables are particularly valuable for analyses of these groups because a high proportion of American Indian, Alaska Native, and Pacific Islander people report multiple races. When “single race” categories are used, a non-representative half (or more) of these populations are left out of the group’s analysis (they are moved to the “two or more races” category, along with other

multiple-race populations). Using the PRED* variables or RACHSING allows all people with these responses to be included in analytically meaningful categories.

6. Can RACHSING be used to identify the different experiences given immigrant status, country of origin, or persons who have been in the United States for multiple generations?

RACHSING is not designed for this. You may be interested in the [BPL](#) variable and the [family interrelationship variables](#) designed by IPUMS to explore this important topic.

7. Does RACHSING help with the identification of heterogeneity within federally-defined single-race groups?

Only in a small way. The prediction equations are based on the person's age group, sex, and characteristics of where they live. If different subgroups in a larger group (e.g., immigrants from Africa as compared to Afro-Caribbean immigrants) are living in areas with different characteristics, the equations are able to give different results.

8. Where can I get more information about how specific responses to the race question might be classified?

Here are the [current federal guidelines](#). More detail is also available in a post from the [Census Bureau's Random Samplings blog](#) and this [Census Bureau topic page about race](#).