Separate and Unequal:
Observed and Self-Reported Race
in the General Social Survey[1]

by Aliya Saperstein
University of California-Berkeley
asaper@berkeley.edu

Department of Sociology
410 Barrows Hall
Berkeley, CA 94720-1980

Department of Demography
2232 Piedmont Avenue
Berkeley, CA 94720-2120

Abstract

Using unique national data from a recent experimental sub-sample of the General Social Survey (GSS) that includes multiple measures of race, I examine observed and self-reported racial classifications for patterns of inconsistency. Much as Telles concludes with similar Brazilian data (Telles 2002; Telles and Lim 1998), this study indicates that the two types of classification can yield different results when used in research because they describe demographically different groups of people. In particular, I find differences in immigrant status, educational attainment and age help explain inconsistencies between the two types of racial classification in the GSS sample. These findings have implications for improved methodology in social science research that relies on racial classification.

[1] This paper was prepared for presentation at the 2004 Population Association of America Annual Meeting. The research was supported by a training grant from the National Institute of Child Health and Human Development (T32HD07275-20). Many thanks go to Mike Hout, Claude Fischer, Leo Goodman, Sam Lucas, Rachel Sullivan and Ben Moodie for their helpful comments and suggestions.

Racial data, in the United States, is gathered in different and sometimes conflicting ways. An individual's physical appearance does not always match his or her identity, which in turn does not necessarily match his or her ancestry. Yet, scientific research has treated observed, self-reported and ancestrally inferred measures of race as interchangeable. Racial classifications, in everything from surveys to birth and death certificates used to compute vital rates, are recorded by different sources at different times and places (e.g., the interviewer, the respondent, the doctor, the parents, the spouse). This lack of uniformity can result in a hodge-podge of socially defined classifications based to an unknown extent on appearance, ancestry, social status or demeanor.

In the past, the assumption that these various measures of race described the same thing was understandable. Official definitions based on ancestry or blood, such as the "one-drop rule," which classifies anyone with any black ancestry as black, supposedly took away any guesswork. The boundaries between racial categories were thought to be firm; whatever race was, biologically determined or culturally inherited, you knew what race someone was if you saw them, and if you asked them what race they were, they would say the same thing. However, as a number of studies have shown, immigration, intermarriage and the increasing number of people claiming multiracial ancestry may be changing all of that.[2] A few also suggest that this new fluidity of identities and complexity of classifications is not new at all.[3]

Unfortunately these studies can only tell us that, in specific contexts, observed and self-reported measures of race are not always equivalent for certain groups in the United States. Their findings do not give us a sense of the accuracy of these measures on a larger scale. What are the implications of these potential inconsistencies for broad-based demographic research? In this study, I use unique national cross-sectional data from the General Social Survey (GSS) that includes observed and self-reported measures of race to extend previous group-specific findings. Is there inconsistency between the results of different measures of race in a representative national survey, and if so, is there

[2] E.g., Harris and Sim 2002; Montalvo and Codina 2001; Goldstein and Morning 2000; Xie and Goyette 1997.

[3] E.g., Nobles 2000; Haney Lopez 1996; Williamson 1995.

systematic bias that could result from using observed race over self-reported race (or vice versa)?

My results show that, while the overall inconsistency between the two measures of race is slight, it is not simply the product of random error. Characteristics such as parents' country of origin, educational attainment and age help predict for whom classification inconsistencies occur. And even when limiting the analysis of inconsistency to net effects, the distribution of parental nativity in the sample differs significantly by the type of classification used. These built-in associations can affect research conclusions, as I illustrate by using each classification to analyze income differences by race (Telles and Lim 1998).

The Mismeasure of Race

Theory suggests and empirical reality shows that there is nothing fixed or objective about race as a descriptive tool for either classification or identification. Historically, fluidity and complexity can be found in the ever-changing social status of Americans of mixed black-white ancestry (Williamson 1995), in the U.S. case law that defined, piecemeal, the limits of citizenship rights by race (Haney Lopez 1996), and in the fact that the U.S. Census has not employed the same racial category names or definitions from one decade to the next (Nobles 2000). Over time, and across countries, race has been defined by any combination of appearance, nationality, geography, ethnicity or religion, depending on time, place, and political context.[4] From such evidence, it is clear that inconsistency between measures of race should be expected, if for no other reason than because it is difficult to keep the latest socially constructed definition straight.

Two recent studies have attempted to identify potential analysis errors when dealing with such a fickle variable, as I do here. In their study of multiple measures of race from a Brazilian national survey, Telles and Lim (1998) find that self-identified pardos ("browns") have higher incomes that interviewer-classified pardos, such that there is a 10 percent difference between the predictions of the income gap between "whites"

[4] See Omi and Winant 1994

and "browns" in Brazil when using the interviewer's classification or the respondent's self-identification. Goldstein and Morning (2000) hint that similar over- or understatements are possible when working with racial data in the United States. They find using Current Population Survey and census data that people who report black, Asian or Native American as their primary race have higher household incomes if they go on to describe themselves by multiple races. While they note this could be capturing an empirical reality—particularly when the respondents' other race was white—Goldstein and Morning suggest another explanation could involve selective (multi) racial reporting in surveys.

Both of these studies show that there is enough inconsistency between measures of race in their respective surveys that the groups of people defined by these measures can differ significantly from one another on important characteristics. I use the GSS data and its measures of observed and self-reported race to extend the above findings to the general U.S. population, and to explore what some of those characteristics might be.

Indications of characteristics that might cause classification inconsistencies come from several recent studies of how people report their racial identities. For example, Montalvo and Codina (2001) show that self-reports of race can vary by immigrant status. Their study of the National Chicano Survey indicates that women are more likely to classify themselves as white if the survey is conducted in Spanish, while women given the survey in English are more likely to choose the color category "brown." The women with greater English proficiency also were more likely to have been in the United States longer, and Montalvo and Codina suggest that these women were responding based on a different understanding of race and the racial hierarchy in America than their Spanish-speaking counterparts. Harris and Sim (2002) take the context argument even further, showing that—at least among American multiracial teens—racial identity can vary by age, neighborhood racial composition, even by whether someone else is in the room when the question is asked.

Xie and Goyette (1997) point to educational attainment as another possible source of variation in racial identity. They predict two possible outcomes for those with high educational attainment: one identity based on assimilation, the other on "awareness." On the one hand, members of minority groups—particularly the foreign born—might identify less with their culture of origin and more with that of the dominant group as their education (and therefore their assimilation) increases. On the other hand, increased education for minorities might lead to a heightened awareness of group differences and a strengthened racial or ethnic identity. Xie and Goyette find evidence of both processes in their study of the census classifications of biracial children with one Asian parent. They also find the likelihood of one or the other outcome depends on how long the Asian parent has been in the United States, and how large and concentrated the Asian population is where the family lives.

Lee (1994) shows the processes of assimilation and awareness can also work, in a sense, in reverse. In her study of Asian-American high school students, their assimilative or "oppositional culture" identities helped shape their perceptions of educational achievement. Thus, individuals might pursue higher education, or avoid it, depending on already formed identities. This reversal (or second generation) of Xie and Goyette's processes suggests a selection effect: that people with high educational attainment will often be those who prefer assimilative identities or lifestyles.

Other studies indicate that racial classification inconsistencies could be explained by characteristics or choices of the interviewers. For example, Hill (2002) shows interviewers perceive more variation in skin tone among respondents of their own race than among respondents of a different race. Telles (2002) also finds that Brazilian interviewers are more likely to "whiten" highly educated, darker-skinned respondents—that is, classify them as white instead of brown, or brown instead of black—than the respondents are to whiten themselves. Another, rather obvious, explanation for classification inconsistency from the side of interviewers is the inability to determine multiracial backgrounds from appearances. As one early 20th century study shows, visually distinguishing between someone with one black grandparent and someone homogenously white is nearly impossible (see Williamson 1995).

From these studies, it seems clear that age, immigrant status, and educational attainment are all plausible determinants of racial classification inconsistency. However, one caveat in this literature is the scope of the studies. Each of the analyses described above is based either in another country or on a small portion of the U.S. population. As such, their results cannot be generalized to the entire population of American adults, or used to inform research designs based on national-level data. Nor do any of the U.S. studies quantify the error risked by analyses that favor one type of classification over another. Thus, the findings I present below are a much-needed extension of our knowledge about variation in racial classification in the United States and its implications for social science research.

Data and methods

The standard procedure for measuring race in the General Social Survey[5] is to direct interviewers to designate a respondent's race under three categories: "White," "Black" and "Other (SPECIFY)." This coding by observation is stipulated "only if there is no doubt in your mind," otherwise the interviewer is to ask, "What race do you consider yourself?" In 1993 and 1994, as many as 14 percent of the responses were based on the respondent's self-report of race.

In part because the number of self-reports of race was on the rise, the GSS conducted a "race experiment" in 1996 and 2000 with the stated goal of developing "valid and reliable measures of race that also identify socially meaningful groups and are consistent or calibrated with current procedures for measuring race" (Smith 1997). For the experiment, racial classification was split into two parts, first the respondent's observed race—determined by the interviewer—and second, the respondent's self-reported race. Along with their observed racial classification, interviewers were prompted to code how confident they were in their assessment, on a four-point scale ranging from "no doubt" to "completely unsure."

Overall, the GSS race experiment yielded 2,869 cases, 1,419 from 2000 and 1,450 from 1996. In each year, the experiment included 50 percent of the total GSS sample for that year. Ninety-five percent of the experiment sample was consistently classified. This is a much higher percentage than the 79 percent found by Telles in Brazil, which likely reflects the comparative rigidity of the white-black racial divide in the United States.

However, GSS respondents in the ambiguous "Other" category were consistently classified in just 50 percent of the cases, and comparing the two types of classification results in very different estimates of the racial makeup in the sample (see Table 1). Using

[5] The General Social Survey has been conducted in the United States since 1972. It is an in-person survey of adults—18 years old and older—living in households, that includes wide-ranging background information and attitudinal measures. Since 1994, the GSS has conducted 3,000 interviews in even-numbered years, using full-probability sampling, and has maintained a 77 percent response rate.

observed race, the sample would be described as 80 percent "White," 15 percent "Black" and 5 percent "Other." Based on self-reported race, the percentage of "Others" nearly doubles to 9 percent with all but one-half of one percent of the change affecting the estimate of "Whites."[6]

<Table 1 about here>

To determine whether these measurement differences are patterned, I use logit regression techniques: analyzing classification consistency for the white and nonwhite groups, and alternating whether each is defined by the interviewer or respondent classification. The dependent variables are coded 1 when the classifications of the interviewer and respondent match and 0 when they do not. For cases in the self-reported "White" category, a 0 means the interviewer "darkened" the respondent to either "Black" or "Other." A 0 for self-reported "Nonwhites" corresponds to the interviewer "whitening" the respondent. The latter was the most common occurrence of inconsistency, accounting for 83 percent of the cases of disagreement. This, of course, flies in the face of many of the assumptions of racial classification literature. In the past, people have been prevented from claiming whiteness by both custom and law (Davis 2001; Haney Lopez 1996), yet these GSS respondents prefer to say they are not "white," even though the interviewer reports that is what they appear to be.

If observed and self-reported race are equivalent, then any classification inconsistencies that result from comparing the two measures should be randomly distributed across cases. Statistically significant coefficients on any of the independent variables, which include age, gender, region, educational attainment[7] and parents'

---

[6] The high degree of inconsistency for respondents in the "Other" category cannot be explained solely by people who are reporting Hispanic ethnicity as a race. Based on the respondents' answers to the GSS ancestry question, about half of those reporting "Other" as their race also reported Hispanic ancestry. Most of the others reported a western European or American ancestry. A similar comparison of responses indicates that interviewer-classified "Others" are predominantly of Asian ancestry.

[7] I combined the GSS "highest degree earned" variable into three groups: "Low," "Medium" and "High." "Low" for respondents with less than a high school degree, "Medium" for respondents with either a high school degree or a junior college degree, and "High" for those with Bachelors and graduate degrees. This coding is similar to the one used by Telles (2002)—though his cutoffs for Brazil are less than primary school, less than secondary school and secondary or more—and was necessary for meaningful analysis given the small number of inconsistent cases. In the GSS sub-sample, 16 percent of respondents fell into the "Low" degree category, 60 percent into "Medium" and 24 percent in "High." However, for the

country of origin,[8] indicate that the inconsistencies are patterned, and the two measures are not equivalent. Significant negative coefficients also indicate that respondents with those characteristics are less likely to be consistently classified.

There are some limitations to the GSS data. The broad categories of "White," "Black," and "Other," lump together the diverse self-images and classification histories among the "Other" groups (e.g., Hispanics, Asians, Native Americans). At the expense of further obscuring diversity, I grouped the respondents into whites and nonwhites to increase the statistical power of my analysis. However, this aggregation allows me to examine the inconsistency-causing characteristics that "Blacks" and "Others" may have in common. It also highlights characteristics that predict dominant group classification, which historically has proved difficult to claim (Davis 2001; Nobles 2000; Haney Lopez 1996).

Other design limitations that make this data less than ideal include the GSS method of matching interviewers and respondents by neighborhood, and the placement of the race measures after questions on ethnicity and a number of socioeconomic background characteristics.[9] However, the effect of the above limitations is generally in the direction of underestimating inconsistent racial classification.

"Nonwhite" group, the "Low" and "High" frequencies were reversed (23 percent and 17 percent, respectively).

[8] To capture immigrant status, I chose to use parents' country of origin, coded by whether both parents were born in the United States, one was, or neither were. Though the respondent's own immigrant status is a significant predictor classification inconsistency (analyses not presented here), the nativity of the respondent's parents accounts for more of the variation, perhaps because it combines information about both the respondent's self-report and the interviewer's classification. High rates of immigration from non-European countries, dating to the 1960s, make it more likely that those born outside the United States will be phenotypically different from "white" Americans. Thus, particularly in the case of respondents who had one foreign-born parent, parental nativity could be a proxy for mixed ancestry and/or a racially ambiguous appearance, which would make consistent classification less likely.

[9] The multi-part race experiment appeared in the same place on the survey as the racial classification in standard questionnaires. As Question 51, it followed queries on the respondent's country of origin, marital status, occupation and work status, education and degrees, self-reported class status, spouse's occupation and work status, parents' occupations, education and degrees and a number of attitudinal measures on topics such as women's roles and the death penalty. In the Brazilian survey analyzed by Telles (2002), the observed racial classification was filled out before anything else on the questionnaire, so the results would not be influenced by other reported information, including the respondent's self-classification. In the GSS experiment, the interviewer does code race before the respondent, but only after hearing a fair amount of background information. While this may result in an underestimation of inconsistency because the interviewers have more information about the respondents and how they think of themselves, it also may serve to highlight the characteristics that the interviewer is relying on to make their determination.

Findings

Since 1960, the U.S. census has allowed racial self-classification, and many scholars consider self-reports to be the most appropriate measures of race—for who would know the person's ancestry better than him or herself? The first model for self-reported nonwhites is based on this same assumption; the coefficients are the log odds predicting that the interviewer's classification will match the respondent's identification as nonwhite (see Table 2, column 1). Recall that this scenario accounts for the vast majority of the inconsistent cases in the GSS sample.

<Table 2 about here>

As expected, parental nativity, educational attainment and age[10] all help explain the inconsistent classification of self-reported nonwhites, and each of the significant coefficients from these variables have negative coefficients, which means self-reported nonwhite respondents with any of these characteristics have a lower likelihood of consistent classification. The largest negative effects come from having at least one foreign-born parent. These respondents are more likely to be classified as white by interviewers, compared to respondents who self-identify as nonwhite but have two U.S.-born parents. As illustrated in Figure 1, this translates into a predicted probability of consistent classification of 79 percent for an otherwise average respondent with two foreign-born parents who identifies as nonwhite, and 64 percent for an otherwise average respondent with one foreign-born parent.

<Figure 1 about here>

[10] I restricted the study sample to respondents aged 24 years and above to remove younger respondents who may not have completed their education yet.

In addition, as the interaction effects between educational attainment and age indicate, the older and more highly educated the self-identified nonwhite respondents are, the worse the interviewers do at consistently classifying them.[11] These findings could support the idea that educational attainment can increase “awareness” or ethnic consciousness (Xie and Goyette 1997). It also may indicate that the interviewers are more likely to assume well educated (i.e., well spoken) people are white—or at least, as goes the assumption in Brazil, that highly educated people would prefer to be classified as white. The finding regarding parental nativity suggests that immigrant status may influence consistency because the respondents have a different understanding of race than the interviewer; though they could claim to be white, for whatever reason, they choose not to.

Now, compare these findings to those for observed nonwhites (see Table 2, column 2). If observed race and self-reported race were equivalent measures, the results from this model and the previous one would be nearly identical; their cases would overlap and, again, any inconsistency would be random distributed. But the two models tell very different stories (compare Figures 1 and 2). As Figure 2 illustrates, among observed nonwhites there is very little difference in consistency between the parental nativity groups. There is significant inconsistency for the most highly educated respondents, such that an otherwise average observed nonwhite person with at least a bachelor’s degree would be consistently classified just 87 percent of time. However, unlike the self-reported nonwhites, this effect does not increase with age.

<Figure 2 about here>

Turning to the self-reported whites, yet another story emerges. Here, interviewers are less successful at matching classifications with respondents at either extreme of educational attainment, though the negative effect is strongest for self-reported whites with less than a high school degree. Consistent classification increases for older

[11] For example, the model predicts an otherwise average 25-year-old with less than a high school degree who identifies as nonwhite will be consistently classified 93 percent of the time, but a 65-year-old with at least a bachelor’s degree will be consistently classified just 78 percent of the time.

respondents who self-identify as white, but again parental nativity does not register as a significant effect.

For observed whites, the largest group in the sample,[12] respondents with less than a high school degree are again less likely to be consistently classified, though that deficit decreases for older respondents with the same level of education.[13] Having foreign-born parents also decreases the probability of consistent classification. This suggests that interviewers are more likely to identify native-born respondents and respondents with at least a high school degree as white. As was the case with the two models for nonwhites described above, a different set of characteristics[14] accounts for consistent classification among observed whites than self-reported whites. That the two models are different is, again, evidence that the two measures of race—observed and self-reported—are not equivalent.

Discussion

What comes across clearly from these results is that the two measures of race are the most consistent for average respondents: middle-aged Americans with high school degrees and U.S.-born parents. Respondents who are atypical on any of those characteristics also are more likely to have atypical (inconsistent) racial classifications. One example is a 45-year-old woman of Mexican descent who did not complete high school and identifies as white, whom the interviewer is sure is not. Or a 31-year-old man with a graduate degree, who the interviewer has no doubt is white, but who reported his race as "Human."[15] Combining atypical characteristics makes racial classification

[12] This is because interviewers classified more respondents as white than respondents classified themselves.

[13] For example, this translates to predicted probabilities of consistent classification of 92 percent for an otherwise average 25-year-old with less than a high school degree whom the interviewer classified as white and 97 percent for a 65-year-old with less than a high school degree.

[14] Or, in the case, of respondents who have at least a bachelor's degree the effect of the same characteristic actually changes direction between models of the two groups.

[15] The public-use GSS data includes 134 verbatim responses from interviewers in 2000 who classified their respondents as "Other" and then specified a group such as "Asian," or "Hispanic." I obtained similar

inconsistency even more likely. For self-reported nonwhites, the 20 percent gap in the probability of consistent classification between respondents with two U.S.-born parents and respondents with one foreign-born parent in Figure 1 increases to 30 percent among the most highly educated in each group, as illustrated by Figure 3.

<Figure 3 about here>

However, much of this complexity is hidden when one examines only the net shifts between racial classification groups on one of the characteristics (see Table 3). Some of the effects identified above operate in opposite directions, and while they may be substantively interesting, in the aggregate they seem largely to cancel each other out. From this perspective, at most, the differences between classifications appear to be on the order of 3.5 percent—for under- or over-estimating the number of nonwhite respondents with native-born parents. Nevertheless, even with this more narrow (or pragmatic) definition of inconsistency bias, the disparity between the distributions of parental nativity by the type of racial classification is statistically significant.

<Table 3 about here>

To isolate the characteristic of parental nativity and illustrate this claim, I reverse the previous analysis and use parental nativity as the dependent variable in a multinomial logit regression. The model includes, as independent variables, the respondent's self-reported race and a categorical variable taking into account the three possibilities of classification consistency (i.e., it is coded "0" for a consistent classification as white, "1" for an inconsistent classification and "2" for a consistent classification as nonwhite). In setting up the analysis this way, I do not mean to imply that an individual's racial classification "predicts" his or her parents' country of origin; I am using the reversed formula to identify whether knowing both racial classifications fits the data on parental nativity better than knowing just one. The statistically significant coefficient for self-

---

verbatim responses from respondents from Tom W. Smith at the GSS, and the "Human" example comes from this previously unreleased data.

reported race in Table 4 indicates that there is, in fact, a difference between self-reported and observed race as a source of information about the respondent's immigrant status.

<Table 4 about here>

As noted above, second-generation Americans are significantly more likely to identify themselves as nonwhite than the interviewers are to so classify them. Thus, using self-reported race may "overestimate" the number of nonwhite Americans with foreign-born parents, or—conversely—using observed race may "overestimate" the number of nonwhites with native-born parents.[16]

Systematic biases also may exist for other variables not examined here. The small sample size for nonwhites made it necessary to limit my study to a few key variables, which were already identified in previous research to be linked to inconsistent classification. But it is likely that the distribution of other characteristics, such as political party identification or religion, might also be affected by the type of racial classification used, and help explain inconsistencies between the two measures of race. As I found above, I would expect respondents with atypical responses on these characteristics (e.g., identifying as an Independent, belonging to a non-Protestant denomination) would also be more likely than a "typical" American to have inconsistent racial classifications.[17]

Finally, to quantify the potential bias introduced by these various systematic differences, I present results from an analysis of family income differences by race (see Appendix, Table 5) similar to the one conducted by Telles and Lim (1998) using Brazilian survey data. As illustrated in Figure 4, using observed race,[18] GSS respondents in the "Other" category are predicted to have family incomes 3 percent greater than those

[16] A similar logit analysis (not presented here) on whether the respondent was foreign born, yielded results in the same direction (though of smaller magnitude) than those for the nativity of the respondent's parents.

[17] Cursory analyses indicate that the distributions of political views, marital status and the number of children the respondent thinks is ideal also differ significantly by type of racial classification.

[18] I use the original "Black" and "Other" categories for this analysis because combining them obscures obvious differences between the two groups with respect to family income.

of observed “Whites.”[19] Using self-reported race, respondents in the “Other” category are predicted to have family incomes 19 percent below those of self-reported “Whites.” The predictions of family income differences for “Blacks” vary little by type of classification, as might be expected by their low rates of classification inconsistency in this sample. But the biases that exist between the two measures of race—above and beyond common human capital and labor market controls—result not just in a 22 percent difference between estimates of the racial gap in earnings for respondents in the “Other” category, but in a substantive change in the conclusions drawn from the analysis.

<Figure 4 about here>

Conclusion

The various findings in this study suggest many things about racial classification in the United States: older people have different conceptions of race than younger ones; immigrants and their children define race differently than native-born Americans; racial identities are affected by educational attainment; and observed racial classifications may be responsive to status cues, as in Brazil and much of Latin America.[20] Though specifying the causal mechanisms behind inconsistencies in racial classification is beyond the scope of this study, the results above clearly illustrate the more general point that observed and self-reported racial classifications are not always equivalent—and that the biases introduced by one or the other measure are significant enough to register even in a nationally representative sample.

That conclusion alone has implications for improved methodology across the social sciences. Different research questions may require different types of racial data: observed race might be best for studies of discrimination, or access to goods and services,

[19] In fact, the coefficient for “Others” in the interviewer classification model is nonsignificant ($p<.66$), which indicates that the null hypothesis that their family incomes are not different from observed “Whites” cannot be rejected.

[20] Many of these conclusions have been established in previous studies with non-generalizable samples, including those mentioned earlier in the article.

while self-reported race might be more useful for studies of attitudes and motivations.[21] Teasing out the sources of race effects in this way would make research findings both more accurate and more useful: It would improve our ability to interpret differential outcomes by race, and help policy makers target resources more specifically to address those differentials.

The fact that various measures of race are potentially contradictory also has implications beyond methodology. For example, people's self-esteem may vary, in part, to the extent that their self-image is not accepted by society at large. This could affect behavior and achievement outcomes. Many scholars also have suggested that socioeconomic stratification exists by skin tone within racial groups not just between them (Keith and Herring 1991). Using multiple measures of race, such as those in the GSS experiment sample, one might be able to isolate these avenues to success: How does the income or educational attainment of those who think they are white but are seen as nonwhite compare to those who think they are nonwhite but are seen as white, and how do both groups' outcomes compare to people who are white according to both measures? That is to say, which matters more: socialization or discrimination, attitude or appearance?

Certainly, the small size of the GSS experiment sample and the broad racial categories available temper these remarks. Nevertheless, there is reason to believe that Americans' racial identities are becoming increasingly complex (Harris and Sim 2002; Omi 2001), and the small but patterned effects identified in this study may be a glimpse of the future of racial classification in the United States.

[21] Telles (2002) comes to much the same conclusion.

References

Davis, F. James. 2001. *Who is Black?: One Nation's Definition.* University Park, Penn.: PSU Press.

Goldstein, Joshua R. and Ann J. Morning. 2000. "The multiple-race population of the United States: Issues and estimates." *Proceedings of the National Academy of Sciences* 97 (11): 6230-6235.

Harris, David R. and Jeremiah Joseph Sim. 2002. "Who is Multiracial? Assessing the Complexity of Lived Race." *American Sociological Review* 67: 614-627.

Hill, Mark E. 2002. "Race of the Interviewer and Perception of Skin Color: Evidence from the Multi-city Study of Urban Inequality." *American Sociological Review* 67(Feb.): 99-108.

Keith, Verna M. and Cedric Herring. 1991. "Skin Tone and Stratification in the Black Community." *American Journal of Sociology* 97:760-78.

Lee, Stacey J. 1994. "Behind the Model Minority Stereotype: Voices of High- and Low-Achieving Asian American Students." *Anthropology and Education Quarterly* 25: 413-429.

Lopez, Ian F. Haney. 1996. *White by Law: The Legal Construction of Race*. New York: NYU Press.

Montalvo, Frank F. and G. Edward Codina. 2001. "Skin color and Latinos in the United States." *Ethnicities*1:321-341.

Nobles, Melissa. 2000. *Shades of Citizenship: Race and the Census in Modern Politics*. Stanford, Calif.: Stanford University Press.

Omi, Michael. 2001. "The Changing Meaning of Race," pp. 243-263 in *America Becoming: Racial Trends and Their Consequences*, edited by Neil J. Smelser, William J. Wilson, and Faith Mitchell. Washington, DC: National Academy Press.

______ and Howard Winant. 1994. *Racial Formation in the United States: from the 1960s to the 1990s.* New York: Routledge.

Smith, Tom W. 1997. "Measuring Race by Observation and Self-Identification" GSS Methodological Report No. 89. National Opinion Research Center, University of Chicago.

Telles, Edward E. 2002. "Racial Ambiguity among the Brazilian Population." *Ethnic and Racial Studies* 25: 415-441.

______ and Nelson Lim. 1998. "Does it Matter Who Answers the Race Question? Racial Classification and income Inequality in Brazil." *Demography* 34(4): 465-74.

Waters, Mary C. 1990. *Ethnic Options: Choosing Identities in America*. Berkeley, CA: University of California Press.

Williamson, Joel. 1995. *New People: Miscegenation and Mulattoes in the United States*. Baton Rouge, La.: Louisiana State University Press.

Xie, Yu and Kimberly Goyette. 1997. "The Racial identification of Biracial Children with One Asian Parent: Evidence from the 1990 Census" *Social Forces* 76(2): 547-70.

TABLE 1. Racial classification comparison in GSS experiment sample

Observed race	***Self-reported race*** White	Black	Other	Total
White	2182	11	109	2302
	95%	1%	5%	100%
	99%	3%	43%	**80%**
Black	11	399	17	427
	3%	93%	4%	100%
	1%	**97%**	7%	**15%**
Other	13	2	125	140
	9%	1%	89%	100%
	1%	0%	**50%**	**5%**
Total	2206	412	251	2869
	77%	**14%**	**9%**	100%
	100%	100%	100%	100%

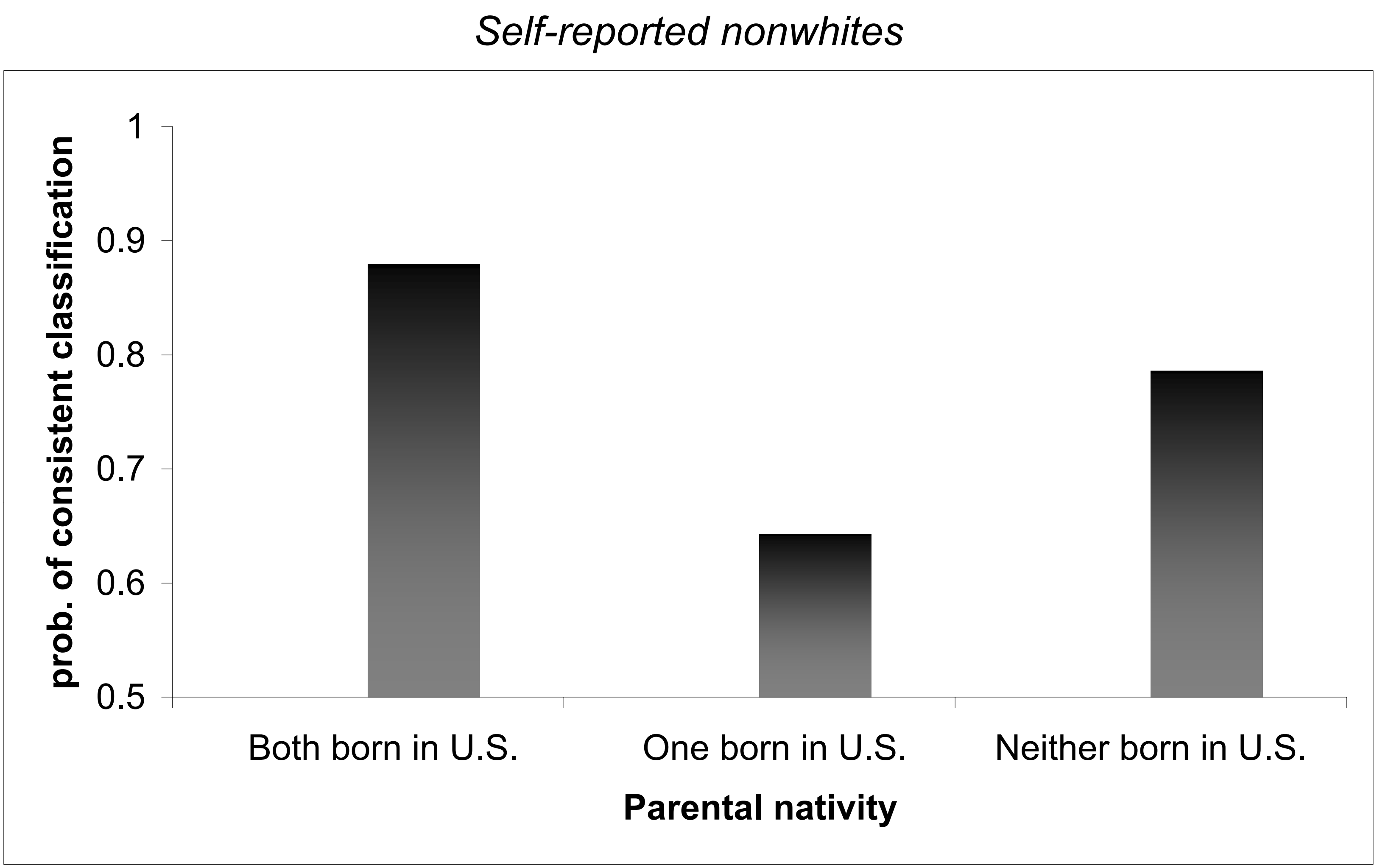
Note: Percent consistently classified on diagonal. Sample distribution by ID type highlighted in black and gray. Percents do not sum to 100 due to rounding.

TABLE 2. Logit models predicting consistent classification by race and ID type

	Nonwhites		*Whites*	
	Self-Reported	Observed	Self-Reported	Observed
Parental Nativity				
One born in U.S.	**-1.221**	.399	-.012	**-1.368**
	(.460)	(1.085)	(1.050)	(.384)
Neither born in U.S.	**-.620**	**1.584**	-.336	**-1.919**
	(.304)	(.775)	(.773)	(.282)
Age	-.012		**.041**	**.025**
	(.011)		(.018)	(.010)
Education				
Less than H.S.	-.876	-.225	**-1.342**	**-1.895**
	(.502)	(.617)	(.651)	(.498)
B.A. or more	.811	**-1.574**	-.874	.547
	(.602)	(.491)	(.527)	(.499)
Region				
Northeast	.397			.217
	(.430)			(.546)
South	.069			-.005
	(.345)			(.475)
Mountain	**-1.576**			**-1.911**
	(.480)			(.511)
Pacific	-.199			.153
	(.407)			(.593)
Female	.231	.738	.226	.166
	(.249)	(.446)	(.469)	(.542)
Interviewer Doubt	**-.933**		**-2.399**	**-2.281**
	(.316)		(.595)	(.313)
Interaction effects				
Education*Age				
Low*Age	**.041**			**.048**
	(.021)			(.020)
High*Age	-.047			-.008
	(.027)			(.021)
Region*Female				
Northeast*Female				.183
				(.799)
South*Female				-.724
				(.675)
Mountain*Female				1.707
				(.911)
Pacific*Female				-1.322
				(.783)
Constant	**2.112**	**2.895**	**4.450**	**3.415**
	(.411)	(.413)	(.535)	(.422)
N	562	511	1968	2044

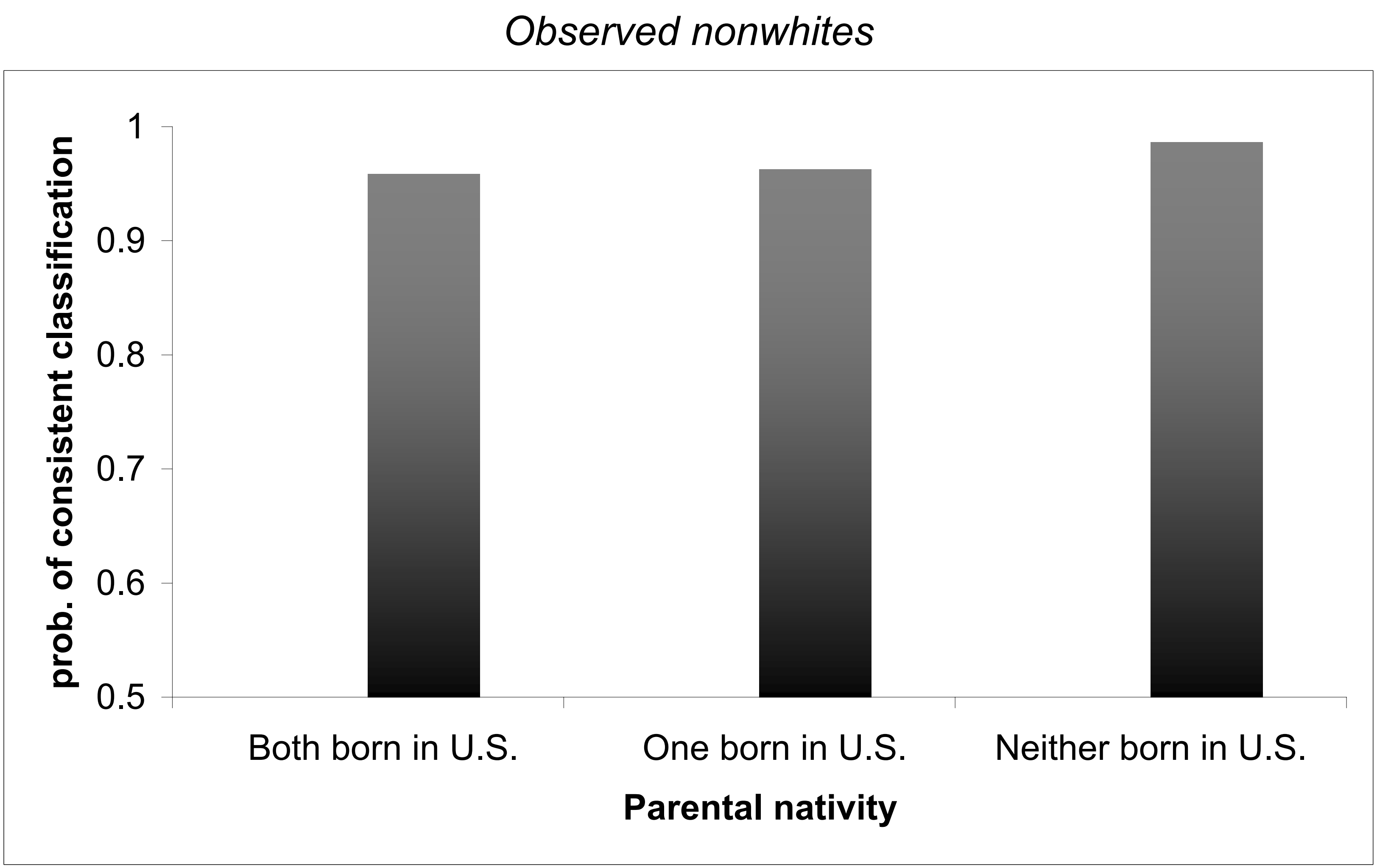
Note: Coefficients significant at p<.05 level in **boldface**. Numbers in parentheses are standard errors. Independent variables, except age, are coded 1 when respondents have the named charateristic and zero when they do not. Constant refers to the log odds of consistent classification for a male, age 24, from the Midwest who has native born parents and a H.S. degree.

Figure 1. Logit results predicting consistent classification



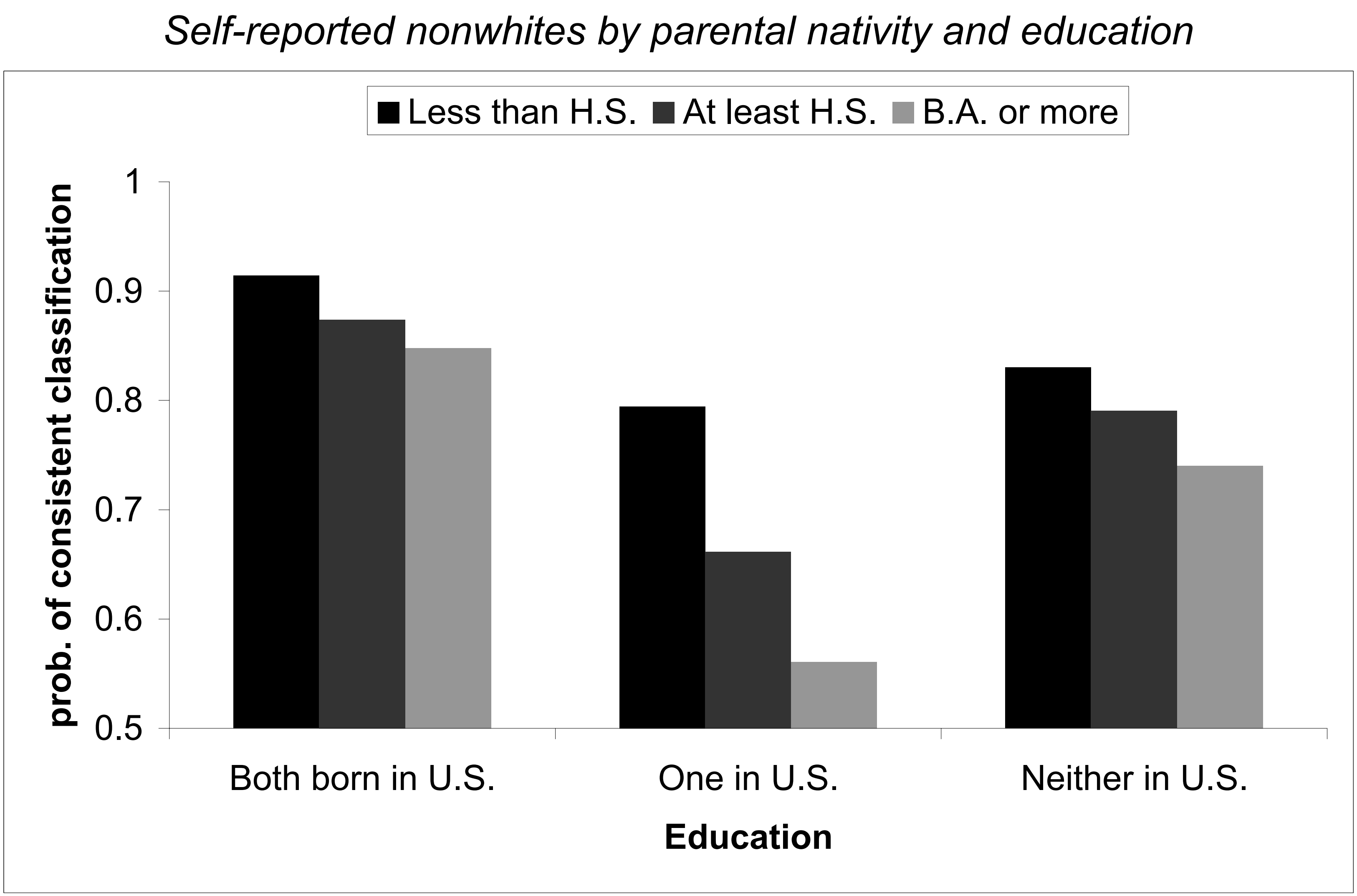
Note: All variables not shown were held to the sample mean.

Figure 2. Logit results predicting consistent classification



Note: All variables not shown were held to the sample mean.

Figure 3. Logit results predicting consistent classification



Note: All variables not shown were held to the sample mean.

TABLE 3. Descriptive statistics by racial group and ID type (study sample)

	Nonwhites		*Whites*	
	Self-reported	Observed	Self-reported	Observed
Parental Nativity				
Both born in U.S.	**73.5%**	**75.7%**	**85.9%**	**84.7%**
One in U.S.	**4.8%**	**3.9%**	**5.8%**	**6.0%**
Neither in U.S.	**21.7%**	**20.4%**	**8.3%**	**9.3%**
Age (mean)				
	43.2	43.1	48.9	48.7
Education				
Less than H.S.	23.0%	22.9%	13.5%	13.7%
At least H.S.	59.4%	58.9%	58.8%	59.0%
B.A. or more	17.6%	18.2%	27.7%	27.4%
N	562	511	1968	2044

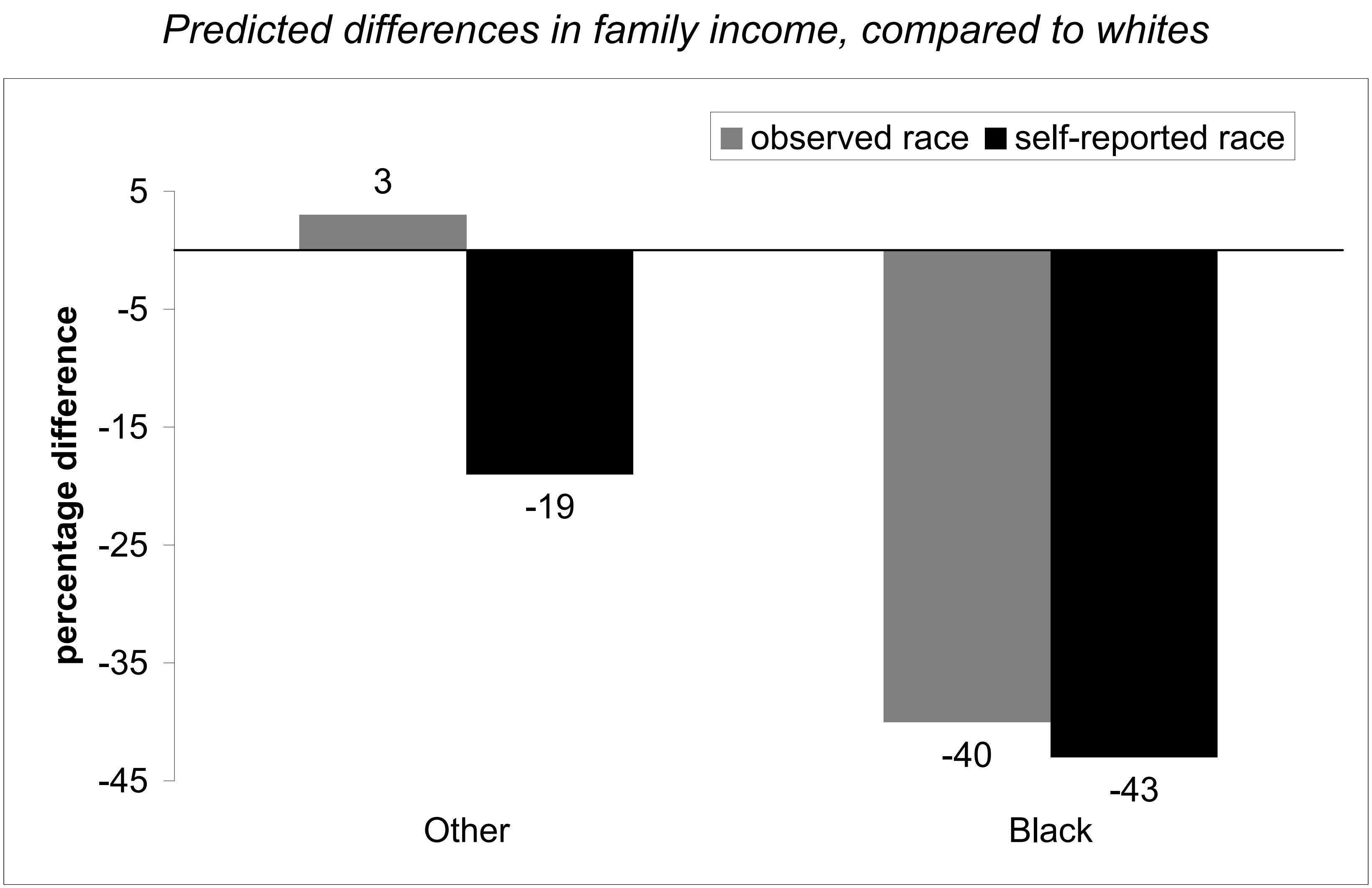
Note: Difference in nativity distributions by ID type is significant at $p<.001$ (see paper for details of test).

TABLE 4. Multinomial logit analysis of variation in parental nativity by racial ID type

	One parent born in the U.S.	*Neither born in the U.S.*
Consistency scale	**-1.306**	**-.568**
	(.364)	(.223)
Self-report as nonwhite	**2.415**	**2.248**
	(.628)	(.416)
Constant	**-2.769**	**-2.345**
	(.095)	(.078)

Note: All coefficients significant at p<.01 level. Numbers in parentheses are standard errors. Comparison group is respondents with two U.S. born parents. Constants refer to the log odds of a respondent consistently classified as white having one, or two foreign-born parents, respectively.

Figure 4. Income analysis by type of racial classification (study sample)



Note: All other variables not shown were held to the sample mean.

Appendix

Family income was calculated from the midpoints of the GSS income categories in each year and coded in thousands of dollars. Respondents in the open-ended top category for 1996 and 2000 were assigned values of $85,000 and $135,000, respectively. All values were then adjusted to constant 2000 dollars and logged.[22] The ordinary least squares regressions are limited to respondents 24 years old and above, as in the previous inconsistency analyses. Education and region are also coded in the same way.

I chose family income over respondent's income because it has fewer missing cases in the GSS data. However, similar analyses using respondent's income produced similar results.

TABLE 5. Ordinary least squares regression of logged family income by racial ID type

	Self-reported race	*Observed race*
Black	**-.414**	**-.396**
	(.049)	(.048)
Other	**-.228**	-.037
	(.062)	(.083)
Woman	**-.204**	**-.205**
	(.034)	(.034)
Age	**.035**	**.036**
	(.004)	(.004)
Age squared	**-.001**	**-.001**
	(.000)	(.000)
Education		
Less than H.S.	**-.609**	**-.616**
	(.050)	(.050)
B.A. or more	**.400**	**.404**
	(.039)	(.039)
Region		
Northeast	.066	.060
	(.050)	(.050)
South	**.129**	**.127**
	(.044)	(.044)
Mountain	.033	.004
	(.072)	(.072)
Pacific	**.209**	**.185**
	(.056)	(.056)
Constant	**3.321**	**3.299**
	(.056)	(.056)
N	2328	2328

Note: Coefficients significant at p<.01 level appear in **boldface**. Numbers in parentheses are standard errors.

[22] I am grateful to Mike Hout for suggesting this recoding of the GSS income variables.